arXiv:2304.00678v1 [econ.EM] 3 Apr 2023

替代与互补模式的检验和识别*

Rui Wang†

2023年2月27日

摘要

本文研究了利用包含商品组合的面板多项选择模型，对两种商品之间的替代与互补模式进行半参数识别。模型允许两种商品互为替代品或互补品，并通过可观测特征体现异质性互补。首先，我提出了关于商品之间互补关系的可检验推论。随后，我刻画了模型参数的精确识别集，并给出了点识别的充分条件。识别分析通过在可观测特征与固定效应之间采用灵活的依赖结构，处理内生协变量，同时不对不可观测的偏好冲击施加任何分布假设。蒙特卡洛模拟表明，与参数化方法相比，我的方法表现得更为稳健。作为扩展，我允许互补性存在不可观测的异质性，探讨涉及两种以上商品的情形，并研究一类非可分效用函数。

关键词：替代模式与互补模式、半参数识别、检验、商品组合、面板多项选择模型

*我感谢 Sung Jae Jun、Joris Pinkse 和 Keisuke Hirano 提供的宝贵指导与支持。感谢 Andres Aradillas-Lopez、Jason Blevins、Jeffrey Campbell、Liyu Dou、Paul Grieco、Patrik Guggenberger、Marc Henry、Ming Li、Xiao Lin、Xun Lu、John Rehbeck、Xiaoxia Shi、Ewout Verriest、Jiahui Xu、Kaida Zhang，以及各类会议和研讨会的与会者。

†俄亥俄州立大学经济系。电子邮箱：wang.16498@osu.edu。

1 引言

商品之间的替代/互补关系已在多种应用场景中得到研究，例如在线新闻与纸质报纸、数字图书与传统图书，以及香烟与电子烟。该关系在消费者的选择中发挥着至关重要的作用；因此，理解替代模式对于预测某种商品的需求，以及分析例如两家公司合并或新商品推出所产生的福利效应十分重要（Petrin，2002；Goolsbee andPetrin，2004；Gentzkow，2007）。

标准的多项选择模型通常假设消费者一次只能购买一种商品，这一假设排除了互补性。然而，即使是一些传统上被视为替代品的商品，在不同情境下也被证明是互补品。例如，Zhao (2019)指出，香烟和电子烟可能是互补品；而Grzybowski and Pereira (2008)则证明了电话通话和短信之间存在互补性。

基于这些动机，我提出一种带有固定效应的半参数面板多项选择模型，以研究替代／互补模式。该模型允许消费者同时购买两种商品，从而考虑这两种商品可能互为替代品或互补品的情形。该模型还允许通过可观测特征体现异质性互补关系。

利用商品组合识别替代／互补模式面临若干挑战。首先，对一种商品的需求包括单独购买该商品的消费者，以及购买商品组合的消费者。因此，对某种商品的需求量较大，可能是因为消费者对该商品的效用较高，也可能是因为该商品与另一种商品具有互补性，或两者兼而有之。我们需要区分这两种来源，才能识别互补关系。其次，同时购买两种商品，可能是由于这两种商品互补，也可能是因为消费者对这两种商品的偏好存在未观测相关性。例如，消费者购买多种有机商品，可能是出于对有机商品的偏好，而非这些商品之间的互补性。区分商品间的互补关系与消费者对商品的偏好相关性可能颇具挑战，因为二者都不可观测，且会同时影响消费者的决策。

为应对这些挑战，我的论文利用了关于偏好冲击随时间变化的条件平稳性假设

，这使我们能够利用条件选择概率的跨期变异进行识别。我首先提出一种检验商品之间替代/互补关系的方法，然后推导出模型参数的精确识别集。

该检验方法利用两种商品需求与其协变量指标之间的关系，并考察商品之间存在替代或互补关系的情形。该方法无需估计任何模型参数，而是构造仅依赖观测变量的条件矩不等式。因此，可以直接检验这些矩不等式来检验互补性。

为推导精确识别集，本文利用条件选择概率的跨期比较，获得关于模型参数的识别约束。该结果不仅识别出互补性的符号，还对其取值加以界定。本文还证明了识别结果的尖锐性，这表明数据中关于模型参数的所有可用信息均已得到充分利用。此外，本文表明，在协变量满足大支持集条件且互补性采用线性设定时，可以实现点识别。

在估计和推断方面，我基于识别结果刻画条件矩不等式，并采用Shi、Shum 和 Song（2018年）的方法。在蒙特卡洛模拟中，我将本文方法的有限样本表现与参数化方法进行比较；后者假设误差项服从参数分布，并为固定效应设定线性模型。模拟结果表明，本文的半参数方法在不同的数据生成过程设计下表现稳健，而参数估计量则容易受到模型设定错误的影响。

我探讨了该模型的若干扩展。首先，我考虑互补关系中存在不可观测异质性的可能性，这使得不同个体之间的互补性可以存在差异，而这些差异无法由协变量捕捉。通过利用两种商品需求的变异，我刻画了关于两种商品互补人群比例的部分识别结果。在在线附录中，我还将模型扩展到包含两种以上商品，并在附加假设下建立了模型参数的部分识别结果。此外，我将该方法扩展到非可分效用函数以及横截面模型。

1.1 相关文献

本文为研究离散选择模型中商品组合的替代/互补模式的文献作出了贡献。Gentzkow (2007）构建了一个允许存在互补性的离散选择模型，用于研究在线新闻与纸质报纸之间的替代关系。他的论文对替代模式设定得较为灵活，同时也允许不同选择之间的误差项存在相关性。与 Gentzkow (2007）相比，我的论文不依赖参数假设，也允许协变量与固定效应之间存在灵活的依赖结构。 Dunker、Hoderlein 和 Kaido(2015）以及 Iaria 和 Wang(2020）通过扩展 Berry、Levinsohn 和 Pakes (1995）中的经典 BLP 方法，允许内生性，并对包含商品组合的模型提供识别结果。他们的方法使用需求反演和误差项的参数分布，并利用工具变量处理内生性。 Monardo(2021）研究了更为灵活的逆需求模型，并使用工具变量构造矩条件。我的论文主要利用面板数据中的跨期变异来推导精确识别结果，而该方法不需要需求反演或工具变量。

有几篇论文允许误差项分布未知，以研究替代模式。 Foxand Lazzati(2017) 在大支持假设和外生协变量条件下，研究了带有商品组合的离散选择模型的半参数识别。 Allen and Rehbeck(2022年) 允许互补性存在未观测异质性，并对认为这两种商品是互补品的人群比例进行部分识别。本文通过可观测协变量聚焦于异质性互补性，并在协变量具有有界支持集的条件下实现尖锐识别。本文的一项扩展在放宽外生协变量假设及 Allen and Rehbeck Allen and Rehbeck(2022年) 中的排除限制的同时，允许存在未观测的异质性互补性。

本文还与大量关于含固定效应面板多项选择模型的文献相关。Chamberlain (1980在扰动项服从逻辑斯蒂分布的假设下，为面板多项选择模型提出了一种条件固定效应 Logit 估计量。Manski (1987)以及Honor´e 和 Lewbel (2002)放宽了逻辑斯蒂分布假设，并研究了二元选择模型的半参数识别。Manski (1987)采用依赖于组平稳性假设的最大分数法，而Honor´e 和 Lewbel (2002)利用特殊回归变量的思路来识别面板二元选择模型。Honor´e 和 Weidner (2020),

Khan、Ponomareva 和 Tamer (2023年)以及Dobronyi、Gu 和 Kyoo(2021年)研究动态二元选择模型的识别。

Pakes and Porter(2022) 以及 Shi、Shum 和 Song (2018) 研究了一种面板多项选择模型。 Pakes and Porter (2022) 通过刻画条件矩不等式，推导出该模型的精确识别，而Shi、Shum 和 Song (2018) 使用循环单调性进行识别和估计。 Khan、Ouyang 和 Tamer(2021) 为多项选择模型提供了推断方法。 Gaoand Li(2020) 放宽了前述论文中的可分离效用函数假设，并研究了一类非可分效用函数。上述论文关注的是自身价格系数的识别，而非不同商品之间的替代模式。本文在这项文献的基础上，将商品组合纳入面板多项选择模型，并刻画替代模式的精确识别。

本文其余部分安排如下。第2 节介绍包含商品组合的面板多项选择模型。第3节给出关于两种商品互补性的可检验推论。第 4节刻画模型参数的精确识别集，并给出点识别的充分条件。第 5 节构建条件矩不等式，第6 节通过蒙特卡洛模拟考察有限样本表现。第 7 节研究一个扩展。第8 节作出总结。

2 面板多项选择模型

本节介绍一种允许商品组合的面板多项选择模型。考虑短面板结构：令 $i \in \mathcal{I}$ 表示消费者， $t \leq T$ 表示时期，其中面板长度 $T \geq 2$ 固定。由于本文关注两种商品之间的替代模式，因此考虑两种商品的情形： $\{A, B\}$。[1]不同于假设消费者只能购买商品 A 或只能购买商品 B，该模型允许消费者同时购买商品 A 和 B，并着重识别商品之间的互补性。

消费者的选择集为 $\mathcal{C} = \{A, B, AB, O\}$，其中 A （或 B）表示仅购买商品 A （或 B）， AB 表示同时购买 A 和 B，而O 表示外部选项。我假设消费者每种商品最多购买一件

[1]附录B.1 研究了商品数量多于两种的情形，并在附加假设下建立了部分识别结果。

goo，并从其选择集中选择效用最高的选项 .

令 u_{ijt} 表示消费者 i在时点 t选择 $j \in \mathcal{C}$ 所获得的效用。按照 Gentzkow (2007) 的做法，令 $\Gamma_{it} = (u_{iABt} - u_{iAt}) - (u_{iBt} - u_{iOt})$ 表示当商品 B 也被消费时，消费商品 A 所带来的增量效用。效用 u_{ijt} 具体设定为[2]

$$\begin{aligned} u_{iAt} &= X'_{iAt}\beta_0 + \alpha_{iA} + \epsilon_{iAt}, \\ u_{iBt} &= X'_{iBt}\beta_0 + \alpha_{iB} + \epsilon_{iBt}, \\ u_{iABt} &= u_{iAt} + u_{iBt} + \Gamma_{it}, \\ u_{iOt} &= 0, \end{aligned} \tag{1}$$

这里， $X_{ijt} \in \mathbb{R}^{d_x}$ 表示可观测特征向量，其中可能包括消费者 i的特征（例如收入）、产品 j的特征（例如价格）以及它们之间的交互项； $\alpha_{ij} \in \mathbb{R}$ 表示产品 j 未观测到的个体固定效应，该效应不随时间变化，例如消费者对某个品牌的忠诚度； $\epsilon_{ijt} \in \mathbb{R}$表示未观测且随时间变化的冲击，该冲击会影响消费者随时间变化的效用；而 $\beta_0 \in \mathbb{R}^{d_x}$ 是有限维未知参数向量。不失一般性，我们将外部选项的效用归一化为零，因此其余选择的效用均相对于外部选项的效用定义。

增量效用的符号 Γ_{it} 表示商品之间的互补关系。如果两种商品的联合效用大于其各自效用之和（$\Gamma_{it} > 0$），则称它们为互补品；反之，则称为替代品（$\Gamma_{it} < 0$）。正如第2.1节所述，这一定义等价于基于总需求的替代模式的另一种定义。在该模型中， Γ_{it} 可以为正、负或零，因此两种商品既可能是替代品，也可能是互补品。我给出了系数 β_0 和互补性的识别结果 Γ_{it}。

除协变量 X_{ijt}外，我假设消费者 i在时间 t 的选择是可观测的，记为 $Y_{it} \in \mathcal{C}$。消费者选择效用最高的选项，这意味着

$$Y_{it} = j \implies u_{ijt} \geq u_{ikt} \text{ for all } k \in \mathcal{C}.$$

为简化分析，我假设选择之间出现平局结果的概率为

[2]在附录B.4中，我讨论这一效用设定如何与传统的离散选择模型相联系。在附录B.3中，我还考察了一类非可分效用函数。

零，从而无需考虑此类情况。只要误差项中至少有一个服从连续分布，这一假设就能成立。即使以非零概率出现平局，只要平局规则始终固定，所有识别结果仍然成立。

本文的主要目标是根据消费者的选择 Y_{it} 和协变量 X_{ijt}，识别替代/互补关系及系数 β_0 。

接下来，我将介绍模型的一些假设。

假设1。 效用 u_{iABt} 中的增量项 Γ_{it} 规定为

$$\Gamma_{it} = \Gamma(Z_i),$$

其中， $Z_i \in \mathbb{R}^{d_z}$ 表示可观测特征向量， Γ 可以是 Z_i的未知函数。

假设1 要求互补性仅取决于可观测协变量 Z_i。函数 Γ 具有灵活性，研究人员也可以不知道其具体形式，因此能够容纳丰富的互补模式。协变量 Z_i 可以包括消费者的特征，例如收入和年龄，因此假设1 允许通过可观测特征形成异质性互补关系。我假设，对于任意 t，协变量 Z_i 在时间上保持不变。如果协变量随时间变化，则可以在协变量随时间取相同值的条件下进行相同的分析： $Z_{is} = Z_{it} = z$。

假设1 的一项限制是，它排除了互补性中的不可观测异质性，并假设具有相同特征 $Z_i = z$的个体具有相同的互补关系。这一结构不仅使我们能够识别互补性 $\Gamma(z)$的符号，还能在给定 $\Gamma(z)$的情况下，界定互补性 $Z_i = z$的大小。第7节中，我将考虑一种扩展，允许互补性 Γ_{it}存在不可观测异质性。在这种情况下，我们只能得到较弱的结果，并且只能部分识别 Γ_{it}符号的分布。

假设2(排除限制**)**。 在 $X_{it} = (X_{iAt}, X_{iBt})$ 中，至少存在一个特征X^*_{it} 不属于 Z_i，且其系数非零。

排除假设要求存在一个变量，该变量只影响商品 A 或 B 的效用，但不影响两种商品之间的互补性。此类变量的一个例子是商品 A 或 B的价格，它会影响

单个商品的效用，但可能不会影响两种商品之间的互补性。对于 X_{it}^{*} 的系数，研究人员仍可能无法确定其符号。此外，该假设不限制协变量 Z_i；任何影响互补性的变量都可以影响单个商品的效用。

最后一个假设是关于未观测冲击分布的平稳性条件。令 $X_{it} = (X_{iAt},\ X_{iBt})$ $\alpha_i = (\alpha_{iA},\ \alpha_{iB})$，以及 $\epsilon_{it} = (\epsilon_{iAt},\ \epsilon_{iBt})$ 分别汇集两个商品的协变量、固定效应和误差项。

假设3。（平稳性）给定（X_{is}、X_{it}、Z_i、α_i）时，ϵ_{it} 的分布随时间保持平稳；即，

$$\epsilon_{is} \mid X_{is}, X_{it}, Z_i, \alpha_i \overset{d}{\sim} \epsilon_{it} \mid X_{is}, X_{it}, Z_i, \alpha_i \quad for\ any\ s, t \leq T.$$

该假设是 Manski （1987年）中条件同质性假设的多项选择扩展。在面板多项选择模型文献中，该假设被广泛采用，包括 Pakes 和 Porter （2022年）以及 Shi、Shum 和 Song（2018年）的研究；他们在该假设下研究系数 β_0 的识别。假设3 要求 ϵ_{it} 的条件分布随时间保持平稳，但允许误差项 ϵ_{it} 在不同选择之间及随时间相关。此外，该假设不对未观测项 ϵ_{it}施加任何分布限制。因此，标准 Logit/Probit 模型以及误差项的独立同分布假设均可纳入假设 3之中。

假设3 的一个关键特征是，它允许固定效应α_{ij} 与协变量 X_{it}之间存在任意依赖结构，从而能够容纳内生协变量。内生性在需求估计中十分重要，因为产品的价格可能取决于该产品不可观测的异质性，例如产品的质量或消费者对该产品的偏好。Chesher 、 Rosen 和 Smolinski （ 2013 ）以及 Berry 和 Haile （ 2014 ）对在需求估计中允许内生性的重要性进行了更为详细的讨论。

当然，假设3确实施加了一些限制。例如，它排除了一些 ϵ_{it} 与协变量 X_{it}之间的依赖结构。考虑一下，如果 ϵ_{it}在任何时期 X_{it} 都只取决于 t，那么 ϵ_{is} 在ϵ_{it} 与 X_{is} 取不同值时，可能具有不同的分布 X_{it} 。不过，假设 ϵ_{it} 3仍允许 X_{it} 与 ϵ_{it} 之间的一些依赖结构：例如，如果 ϵ_{it} 取决于

协变量 s 以时间不变的形式出现，例如 $\frac{1}{T}\sum_{t=1}^{T} X_{it}'\beta_0$，假设3 仍然成立 .

2.1 替代模式

本节讨论两种不同的替代模式定义之间的关系。本文使用 $\Gamma(z)$的符号来表示两种商品之间的替代关系，该符号反映了消费商品组合相较于消费单一商品所带来的增量效用。引理1表明，这实际上等同于文献中此前使用的另一种替代定义。

替代模式的另一种定义着眼于商品 A （或 B）的需求如何受到商品 B （或 A）价格上涨的影响。如果商品 A 的需求增加，则两种商品为替代品；如果需求减少，则为互补品；如果需求不变，则为独立品。令 p_{jt} 表示其˜系数非零的商品 j 的价格，令 $X_{it} = X_{it} \setminus \{p_{Bt}\}$ 表示˜ ˜X_{it} 中除商品 B价格以外的剩余协变量。固定所有其他协变量 $X_{is} = X_{it} = \tilde{x}$ 不随时间变化，并比较商品 A 在商品 B价格 $p_{Bs} \neq p_{Bt}$ 不同时的条件需求。商品 A的需求来自两个来源：只购买商品 A 的个体，以及购买商品组合 AB的个体。令 $D_\ell = \{\ell, AB\}$ 包含所有含有商品 $\ell \in \{A, B\}$的选择。令 $\text{sign}(x) = \mathbb{1}\{x > 0\} - \mathbb{1}\{x < 0\}$ 表示符号函数。

给定协变量 $Z_i = z$ 时，替代模式 $s_{AB}(z)$定义为

$$s_{AB}(z) \equiv \text{sign}\left\{\frac{\Pr(Y_{is} \in D_A \mid p_{Bs}, p_{Bt}, \tilde{x}, z) - \Pr(Y_{it} \in D_A \mid p_{Bs}, p_{Bt}, \tilde{x}, z)}{p_{Bs} - p_{Bt}}\right\}.$$

$s_{AB}(z) \in \{-1, 0, 1\}$的值表示商品 A 和 B之间的互补关系。对于协变量为 $Z_i = z$的消费者，如果 $s_{AB}(z) = 1$，则这两种商品为替代品；如果 $s_{AB}(z) = 0$，则为独立品；如果 $s_{AB}(z) = -1$，则为互补品。在上述假设1-3下， $s_{AB}(z)$的值由任意两个时期 $s \neq t$定义时都相同，且除 z外独立于其他变量，因为互补项 $\Gamma(z)$仅取决于 z；因此， $s_{AB}(z)$可写成仅关于 z的函数。

通常很难直接根据$s_{AB}(z)$的定义研究替代模式。术语 $s_{AB}(z)$仅使用价格的变动，并要求固定所有其他协变量。这可能无法实现，因为其他协变量可能会

与价格同时变化，或者协变量可能包括时间虚拟变量等时变变量。此外，在某些产品价格随时间保持不变的情形下，可能无法获得价格变动。而且，由于定义 $s_{AB}(z)$ 涉及条件选择概率，直接估计 $s_{AB}(z)$的效果可能较差，尤其是在协变量维度较高时。

下一个引理确立了 $s_{AB}(z)$与增量效用 $\Gamma(z)$之间的关系。

引理1. 在假设*1-3*下，对于任意 z，以下结论成立：

$$\Gamma(z)s_{AB}(z) \leq 0.$$

引理1 表明， $s_{AB}(z)$的符号始终与增量效用项 $\Gamma(z)$的符号相反。该引理意味着，只要确定了增量效用 $\Gamma(z)$的符号，就能得知 $s_{AB}(z)$的符号。因此，确定互补项 $\Gamma(z$ 就足以研究由 $s_{AB}(z)$定义的替代模式。

为阐释引理1的直观含义，我将聚焦于增量效用为正的情形，即 $\Gamma(z) > 0$。在这种情况下，即使消费者从单独消费某一种商品中获得的效用较低，他们仍会购买商品组合，因为同时消费两种商品可以带来额外的正效用。当商品 B的价格上涨，使商品组合的效用下降时，一些消费者会从购买商品组合转而选择外部选项，因为他们从单独消费一种商品中获得的效用较低。因此，对商品 A 的需求会下降，这意味着 $s_{AB}(z) \leq 0$。

与引理1 类似的结果由 Gentzkow(2007) 在截面数据中得出。区别在于，Gentzkow (2007) 要求未观测误差项与可观测协变量满足独立性条件，因此其结果不适用于存在内生协变量的情况。本文利用平稳性假设，该假设允许存在内生协变量。因此，引理1 表明，即使存在内生协变量，通过利用协变量的跨期变异，两种替代模式定义之间的关系（$\Gamma(z)s_{AB}(z) \leq 0$）仍然成立。

3 互补性检验

在介绍识别结果之前，本节将提出一种检验互补关系的方法。该方法无需估计任何模型参数，而是直接构造仅取决于观测变量的可检验推论。

本文将重点检验两种商品之间的互补性；替代性的检验分析与此类似，详见附录 B.5。对于协变量 $Z_i = z$ 的消费者，检验商品之间的互补关系等价于检验 $\Gamma(z) \geq 0$。原假设 H_0 和备择假设 H_1 如下：

$$H_0 : \Gamma(z) \geq 0 \qquad H_1 : \Gamma(z) < 0.$$

互补性检验的核心思想是，改变商品 A 的协变量指数$X'_{iAt}\beta_0$，会因商品之间的替代/互补关系不同，对商品 B 的需求产生不同方向的影响。如果两种商品互为互补品（在 H_0条件下），提高商品 A 的协变量指数会促使消费者购买该商品组合，从而增加对商品 B的需求。然而，如果两种商品互为替代品（在 H_1条件下），消费者就会转而只选择商品A，使得对商品 B 的需求下降。因此，在这种情况下，如果观察到对商品 B 的需求下降，我们就可以拒绝 H_0。

然而，协变量指数 $X'_{ijt}\beta_0$ 的符号未知，因为其中包含未知参数 β_0。因此，检验互补性的第一步是利用观测变量（X_{it}、Z_i、Y_{it}）确定两种商品的协变量指数的符号。

令 $P_t(\{j\} \mid x_s、\ x_t、\ z)$ 表示在协变量（X_{is}、X_{it}）=（x_s、x_t）和 $Z_i = z$ 条件下，在时间 t 选择 $j \in \mathcal{C}$ 的概率，定义为

$$P_t(\{j\} \mid x_s, x_t, z) = \Pr(Y_{it} = j \mid x_s, x_t, z).$$

条件选择概率仅取决于可从数据中识别的观测变量。由于误差项的条件分布在不同时期相同（根据假设 3），选择概率的任何变化都只能源于协变量指标的变化。根据观测到的选择概率的变化，我们可以推断协变量指标的符号，并构造可检验推论。

令 $\xi^1_{s,t}(x_s, x_t \mid z)$ 表示所有选择的概率均增加的指示变量$j \in \{A, B, AB\}$，其条件为 $(X_{is}, X_{it}) = (x_s, x_t)$，且给定 $Z_i = z$:

$$\xi^1_{s,t}(x_s, x_t \mid z) = \mathbb{1}\Big\{P_s(\{j\} \mid x_s, x_t, z) - P_t(\{j\} \mid x_s, x_t, z) \geq 0,\ \forall j \in \{A, B, AB\}\Big\}.$$

当观察到所有选择的条件概率均上升时 $j \in \{A, B, AB\}$，这意味着两种商品的协变量指标均上升:

$$\xi^1_{s,t}(x_s, x_t \mid z) = 1 \Longrightarrow x'_{As}\beta_0 - x'_{At}\beta_0 \geq 0,\ x'_{Bs}\beta_0 - x'_{Bt}\beta_0 \geq 0.$$

上述关系利用多个选择的条件概率变化，来识别两种商品协变量指标的符号。如果我们观察到只有一种选择（例如商品A）的概率上升，这可能是由于商品A变得更好，也可能是由于商品B变差，因此我们无法区分这两种情况。相比之下，如果观察到所有选择的概率均上升 $\{A, B, AB\}$，则只有在两种商品都变好的情况下才会发生，因此我们能够确定两种商品协变量指标变化的符号。

现在，我们可以为原假设建立可检验条件。根据两种商品互为互补品这一原假设，两种商品的协变量指标上升会导致对这两种商品的需求增加。这种关系为原假设带来了可检验推论。如果观察到任一种商品的需求下降，就表明这两种商品是替代品，原假设因而被拒绝。

回忆一下，对于 $\ell \in \{A, B\}$，$D_\ell = \{\ell, AB\}$，商品 ℓ 的条件需求表示为

$$\Pr(Y_{it} \in D_\ell \mid x_s, x_t, z) = E[\mathbb{1}\{Y_{it} \in D_\ell\} \mid x_s, x_t, z].$$

命题 1. 在假设 *1-3*下，在原假设 H_0 给定 $Z_i = z$时，以下条件矩不等式成立。

$$E\Big[\xi^1_{s,t}(x_s, x_t \mid z)\big(\mathbb{1}\{Y_{is} \in D_\ell\} - \mathbb{1}\{Y_{it} \in D_\ell\}\big) \,\Big|\, x_s, x_t, z\Big] \geq 0,$$

对于任意 (x_s, x_t) $\ell \in \{A, B\}$, $s \neq t \leq T$。

命题 1通过刻画仅取决于观测变量的条件矩限制，为原假设 H_0 提供了可检验推论。因此，可以直接检验上述条件矩不等式，以检验原假设。该命题中的条件适用于任意两个时期和任意一对协变量，使我们能够利用协变量随时间的变化来检验互补性。

4 识别

除了检验 $\Gamma(z)$ 的符号外，我们也可能关注互补性 $\Gamma(z)$ 的取值以及效用系数 β_0。本节给出参数 $\theta_0 = (\beta_0, \Gamma)$ 的精确识别结果。观测变量包括协变量 (X_{it}, Z_i) 和消费者在每个时期的选择 $Y_{it} \in \mathcal{C}$。由于对消费者的决策而言，选择之间的相对效用才是关键，因此该参数只能识别到一个常数。为进行以下分析，我将参数 θ 的第一个元素 θ^1 归一化为 1：$\Theta = \{\theta : |\theta^1| = 1\}$。

对于任意子集 $K \subset \mathcal{C}$，令 $P_t(K \mid x_s, x_t, z)$ 表示在协变量 $(X_{is}, X_{it}) = (x_s, x_t)$ 和 $Z_i = z$ 给定时，在时点 t 出现 $Y_{it} \in K$ 的条件概率；也就是说，集合 K 中存在某一选择使其效用高于所有其他选择的概率：

$$\begin{aligned} P_t(K \mid x_s, x_t, z) &\equiv \Pr(Y_{it} \in K \mid x_s, x_t, z) \\ &= \Pr\left(\exists j \in K \text{ s.t. } \forall k \in \mathcal{C}\ u_{ijt} \geq u_{ikt} \mid x_s, x_t, z\right). \end{aligned}$$

当 $K = \{j\}$ 是单元素集合时，这就简化为 j的条件选择概率（CCP）。识别分析的主要思路是利用两个不同时期的条件选择概率的跨期变异，推导真实参数 θ_0 的识别约束。满足这些识别约束的所有参数构成真实参数的识别集。

令 $\delta_{\ell t} = x'_{\ell t}\beta_0$ 表示给定 $X_{i\ell t} = x_{\ell t}$时商品 $\ell \in \{A, B\}$ 的协变量指数。令 $\delta_{ABt} = \delta_{At} + \delta_{Bt}$ 和 $\delta_{Ot} = 0$ 分别表示商品组合 AB 和外部选项的协变量指数。令 $\Delta_{s,t}\delta_j = \delta_{js} - \delta_{jt}$ 表示选择 $j \in \mathcal{C}$ 的协变量指数在时期 s 与 t之间的变化。

在假设消费者一次只能购买一种商品的模型中，两种商品

只能是替代品。由于互补关系已知，影响条件选择概率的唯一未知因素是所有选择的协变量指标的变化。我的论文允许两种商品可能是替代品（$\Gamma(z) < 0$），也可能是互补品（$\Gamma(z) > 0$），且互补关系未知。因此，在本文中，有两个未知来源会影响条件选择概率：一个是协变量指标的变化，另一个是两种商品之间的互补关系。区分这两个来源给识别分析带来了挑战。

以下命题刻画了参数 θ_0 在假设1-3下的识别约束。令 $C_1 \vee C_2$ 表示条件 C_1 或 C_2 成立，或两者都成立；令 $C_1 \wedge C_2$ 表示 C_1 和 C_2 都成立。

命题2。在假设*1-3*下，对于任意$(x_s,\ x_t,\ z)$以及任意 $s \neq t$:

（*1*）选择 $j \in \mathcal{C}$的条件选择概率比较:

$$P_s(\{j\} \mid x_s,\ x_t,\ z) > P_t(\{j\} \mid x_s,\ x_t,\ z) \Longrightarrow \exists k \neq j \text{满足条件} \Delta_{s,\ t}\delta_j > \Delta_{s,\ t}\delta_k; \tag{ID1}$$

（*2*）商品 $\ell \in \{A, B\}$的需求比较:

$$\begin{aligned} &P_s\big(\{\ell, AB\} \mid x_s, x_t, z\big) > P_t\big(\{\ell, AB\} \mid x_s, x_t, z\big) \Longrightarrow \\ &\quad \{\Delta_{s,t}\delta_\ell > 0\} \vee \Big\{\Delta_{s,t}(\delta_\ell + \text{sign}(\Gamma(z))\delta_{\ell_{-1}}) > 0, |\Gamma(z)| > -\Delta_{s,t}\delta_\ell\Big\}, \end{aligned} \tag{ID2}$$

其中 $\ell_{-1} \in \{A, B\}$ 和 $\ell_{-1} \neq \ell$;

（*3*）两种选择的条件选择概率之和的比较:

$$\begin{aligned} &P_s(\{AB\} \mid x_s, x_t, z) + P_t(\{O\} \mid x_s, x_t, z) > 1 \Longrightarrow \\ &\qquad \Big\{\Gamma(z) > -\min\{\Delta_{s,t}\delta_A, \Delta_{s,t}\delta_B\}\Big\} \wedge \{\Delta_{s,t}(\delta_A + \delta_B) > 0\}, \\ &P_s(\{A\} \mid x_s, x_t, z) + P_t(\{B\} \mid x_s, x_t, z) > 1 \Longrightarrow \\ &\qquad \Big\{\Gamma(z) < \min\big\{\Delta_{s,t}\delta_A, -\Delta_{s,t}\delta_B\big\}\Big\} \wedge \{\Delta_{s,t}(\delta_A - \delta_B) > 0\}. \end{aligned} \tag{ID3}$$

命题 2 刻画了参数 θ_0 的识别约束，这些约束可以从数据中识别出的两个时期之间的条件选择概率比较得到。命题 θ_0 2中的识别约束不依赖于

不可观测项，例如固定效应 α_i 和误差项 ϵ_{it}。由于这些结果适用于面板数据的任意固定长度 T，我们可以利用任意两个时期的条件选择概率变化来识别 θ_0，并对识别集取交集。稍后我将基于命题 2中的识别约束构造条件矩不等式，用于对参数 θ_0进行估计和推断。

条件（ID1）给出了命题2 中针对系数 β_0的识别约束。该结果的直观解释如下：如果选择 j 的条件概率上升，那么选择 j 相对于所有其他选择（就协变量指数而言）变得更差是不可能的。因此，可以推断，选择 j 的协变量指数相对于至少一个其他选择应当上升。

命题2 中的另外两个条件为互补性参数 $\Gamma(z)$ 提供了新的识别结果。条件（ID2）通过比较两个商品随时间变化的条件需求，识别互补性 $\Gamma(z)$ 的符号，并为其绝对值设定界限。条件（ID3）则利用两个时期内两个不同选择的概率之和，为互补性 $\Gamma(z)$ 确定上下界。接下来，我将解释这两个条件背后的直观含义。

条件（ID2）主要利用这样一个思路：在两种商品之间存在不同互补关系时，一种商品的协变量指数与另一种商品需求之间的关系有所不同。当两种商品互为 A 互补品时，一种商品协变量指数的上升会促使消费者购买商品组合 AB，从而增加对商品 B的需求。因此，如果观察到任一种商品的需求下降，这两种商品就必定是替代品（$\Gamma(z) < 0$）。类似地，当商品 A 的协变量指数下降、商品 B 的协变量指数上升，同时观察到商品 A的需求增加时，可以识别出这两种商品互为互补品（$\Gamma(z) > 0$。

命题ID3中的条件2可以通过考察随时间变化的两种选择的条件概率之和，来界定互补性的值$\Gamma(z)$。例如，当同时购买两种商品和两种商品都不买的条件概率之和较大时，就可以得出互补性 $\Gamma(z$的一个下界。其直观解释是，当商品 A 变得不那么有吸引力时，如果两种商品互为互补品，消费者就会从同时购买两种商品转为两种商品都不买；但如果两种商品是替代品，他们就会转而只购买商品 B 。

因此，如果观察到同时购买两种商品和两种商品都不买的概率较大，我们就可以推断这两种商品具有互补性，并为 $\Gamma(z)$给出一个下界。同样，如果购买单一商品的条件概率之和较大，就可以得出 $\Gamma(z)$的值的上界。

命题 2 中的识别限制（ID1）-（ID3）刻画了一个关于 Θ_I 的 θ_0识别集，其定义如下：

$$\Theta_I = \{\theta \in \Theta : \text{conditions (ID1)} - \text{(ID3) hold with } \theta \text{ in place of } \theta_0\}.$$

定理 1. 在假设 *1-3*下，识别集 Θ_I 是尖锐的。

定理 1 表明，条件（ID1）-（ID3）已穷尽观测数据中关于参数 θ_0的所有可用信息。尖锐性的证明通过直接构造完成。对于识别集 Θ_I中的任意参数，我都构造一个满足假设 1-3 并与观测到的条件选择概率相匹配的数据生成过程，从而证明识别集 Θ_I的尖锐性。构造的主要难点在于，未知的数据生成过程涉及整个未观测误差项空间上的条件分布，而该空间是无限维的。

我通过构造“选择集”来解决这一难题；选择集是一些未观测项的集合，在给定协变量的条件下，其中只有一个选择会被选中。只需构造选择集上的分布，因为这些分布决定了观测到的选择概率。由于选择的数量有限，选择集的数量也有限；因此，我只需为有限个集合分配概率，从而简化构造。随后，本文表明，对于识别集 Θ_I 中的任意参数，都存在一个定义在选择集上的条件分布，它满足假设并生成观测到的选择概率。选择集上概率的构造取决于互补性 $\Gamma(z)$ 的符号以及协变量指数 $\Delta_{s,\ t}\delta_j$，详见附录 A.3。

下面对定理 1 作一些讨论。首先，与 Pakes 和 Porter(2022年类似，命题2中的识别分析以及 Θ_I的尖锐性结果可以推广到更一般的模型：$u_{i\ell t} = f(X_{i\ell t}, \beta_0) + g(\alpha_{i\ell}, \epsilon_{i\ell t})$，其中 f是一个已知函数（至多依赖于有限维参数） β_0 ，而 g 是一个对计量经济学家而言可以未知的函数。该效用函数允许存在无限维

固定效应和特质冲击，也允许二者之间存在任意交互。附录 B.3 还讨论了一类效用函数，这类函数在协变量与不可观测固定效应/误差项之间可以是非可分离的。识别集 Θ_I 在每个时期仅使用边际选择概率，但结果证明该识别集是尖锐的。因此，不同时期的联合选择概率不会为参数 θ_0提供任何额外信息。此外，尖锐性结果利用了不同选择之间以及不同时期误差项的相关性。因此，如果愿意对依赖结构施加附加假设（例如独立同分布），识别集 Θ_I 还可以进一步收紧。

4.1 点识别

本节研究模型参数按比例点识别的条件。分析取决于附加效用项 $\Gamma(Z_i)$ 的设定。我重点讨论互补性的线性设定下的点识别： $\Gamma(Z_i) = Z_i'\gamma_0$。

为简化记号，我考虑一个两期模型（$T = 2$）来说明这一思路。[3] 令 $\Delta X_{i\ell} = X_{i\ell 2} - X_{i\ell 1}$ 表示消费者i和商品 $\ell \in \{A, B\}$ 在两个时期之间可观测协变量的变化，并令 $\Delta X_i = (\Delta X_{iA}, \Delta X_{iB})$ 汇集两种商品的协变量变化。我用上标 k 表示向量的第 k 个元素，例如， ΔX_{iA}^k 表示向量 ΔX_{iA} 的第 k 个元素。

我首先介绍系数 β_0点识别的充分条件。系数 β_0 只能在尺度意义上实现点识别，因为将消费者对所有选择的效用乘以一个正的常数，不会改变观测到的选择。

假设4。 给定$(X_{i1}, X_{i2}, Z_i, \alpha_i)$ 时， ϵ_{it} 的密度在 $\mathbb{R}^2$ 上处处为正，其中 $t \in \{1, 2\}$。

假设5。 对于任意 $\ell \in \{A, B\}$，存在满足 $\beta_0^{k_\ell} \neq 0$的 k_ℓ 。令$\Delta\tilde{X}_i = \Delta X_i \setminus (\Delta X_{iA}^{k_A}, \Delta X_{iB}^{k_B})$ 表示 ΔX_i中的其余元素。给定$(\Delta\tilde{X}_i, Z_i)$ 时，$(\Delta X_{iA}^{k_A}, \Delta X_{iB}^{k_B})$ 的密度在 $\mathbb{R}^2$上处处为正。此外， $\Delta X_{i\ell}$ 的支撑集不包含于 $\mathbb{R}^{d_x}$的任何真线性子空间中。

假设4排除了这样一种无信息情形：尽管协变量指标随时间发生变化，条件选择概率却保持不变。

[3]当时期多于两个时，只要存在一对满足假设的时期，就可以直接实现点识别4-6。

假设 5 是关于协变量 ΔX_i 的支持条件。它要求每种商品至少有一个协变量具有大支持集，而其余协变量的支撑集则不受限制。大支持集条件保证协变量随时间有足够的变化，从而能够将真实参数与任何其他候选参数区分开来。

在这些假设下，β_0 可以利用命题2中的第一项识别限制（条件（ID1））实现点识别。对于任意满足 $b \neq k\beta_0$、$\forall k > 0$的参数 b，假设5中的大支持集条件意味着存在协变量的一个取值 Δx_ℓ，使得协变量指数 $\Delta x_\ell'\beta$ 在真实参数 β_0 和候选参数 b 下符号不同。于是，条件选择概率在 β_0 和 b 下会朝不同方向变化，从而识别参数 β_0。例如，假设对于任意 $\ell \in \{A, B\}$，协变量指数满足 $\Delta x_\ell'\beta_0 > 0$和 $\Delta x_\ell' b < 0$。那么在假设4下，购买商品组合 AB 的条件选择概率在真实参数 β_0下将严格增加，但在参数 b下将严格减少。因此，β_0 得到识别。

由于 β_0 是点识别的，协变量指数 $\Delta X_{ij}'\beta_0$ 的符号也得到识别。接下来，我将介绍互补性参数 γ_0点识别（按比例）的条件。

假设6。存在 k，使得 $\gamma_0^k \neq 0$。令 $\tilde{Z}_i = Z_i \setminus Z_i^k$表示 Z_i中的其余元素。给定$(X_{i1}, X_{i2}, \tilde{Z}_i)$ 时，Z_i^k 的密度在 $\mathbb{R}$上处处为正。此外，Z_i 的支撑集不包含于 $\mathbb{R}^{d_z}$的任何真线性子空间中。

与假设5类似，该假设要求协变量 Z_i满足大支持集限制。根据命题ID2中的条件（2），两个商品的条件需求随时间的变异可用于识别互补性$Z_i'\gamma_0$ 的符号。由于 $Z_i'\gamma_0$ 的符号已得到识别，因此在大支持假设下，参数 γ_0 也按比例点识别。对于γ_0 的分析与系数 β_0类似。对于任意满足 $\tilde{\gamma}$ 的候选参数 $\tilde{\gamma} \neq k\gamma_0$，$\forall k > 0$，假设6 确保存在某个协变量 Z_i 取值，使得在真实参数 γ_0 和候选参数 $\tilde{\gamma}$下，互补性$Z_i'\gamma$ 的符号不同。因此，参数 γ_0 可以得到点识别。

定理 2。在假设 *1-6* 以及 $\Gamma(Z_i) = Z_i'\gamma_0$下，参数 β_0 和 γ_0按比例点识别。

5 条件矩不等式

由条件（ID1）-（ID3）刻画的识别集 Θ_I 见命题 2，其定义较为抽象，而且要检验每个候选参数是否满足所有识别条件是一项颇具挑战性的任务。本节通过构造参数的条件矩不等式，对识别集 Θ_I 给出另一种刻画。基于这一刻画，文献中已发展出许多用于条件矩不等式推断的方法（例如，Andrews 和 Shi (2013),Chernozhukov、Lee 和 Rosen(2013），以及Armstrong(2015））。

命题 2 中的识别条件（ID1）—（ID3）具有相似的结构，涉及通过条件选择概率的跨期比较，为参数 θ_0 推导限制。我将重点讨论命题 2 中的第一个条件（ID1），以说明构造条件矩不等式的思路。令 $W_{ist} =$（X_{is}，X_{it}，Z_i）汇集两个时期（s，t）的所有协变量，并令$w_{st} =$（x_s，x_t，z）表示协变量 W_{ist} 的一个实现。

条件（ID1）利用单个选择 $j \in \mathcal{C}$ 的条件概率比较来推导参数限制。令 $\lambda^j_{s,t}$（w_{st}，θ）表示条件（ID1）中识别限制的指示变量指数，其定义为

$$\lambda^j_{s,t}(w_{st},\theta) = \mathbb{1}\{\exists\ k \neq j \text{ s.t. } \Delta_{s,t}x_j'\beta > \Delta_{s,t}x_k'\beta\}.$$

条件（ID1）通过随时间变化的选择选择 j 的条件概率正向变化，推导出识别限制 $\lambda^j_{s,t}$：

$$P_s(\{j\} \mid w_{st}) - P_t(\{j\} \mid w_{st}) > 0 \Longrightarrow \lambda^j_{s,t}(w_{st},\theta_0) = 1.$$

上述条件的逆否命题如下：如果识别限制 $\lambda^j_{s,t}$ 不成立，那么选择选项 j 的条件概率变化为非正。

$$\lambda^j_{s,t}(w_{st},\theta_0) = 0 \Longrightarrow P_s(\{j\} \mid w_{st}) - P_t(\{j\} \mid w_{st}) \leq 0.$$

代入条件选择概率 $P_t(\{j\} \mid w_{st}) = E[\mathbb{1}\{Y_{it} = j\} \mid W_{ist} = w_{st}]$ 的定义后，上述条件导出以下条件

适用于任意 w_{st} 的矩不等式，

$$g^j_{s,t}(w_{st},\theta_0) = E\big[(1-\lambda^j_{s,t}(w_{st},\theta_0))(\mathbb{1}\{Y_{is}=j\}-\mathbb{1}\{Y_{it}=j\}) \mid W_{ist}=w_{st}\big] \le 0.$$

上述条件矩不等式成立，因为二元指标成立$\lambda^j_{s,t}(w_{st},\theta_0)=1$ 因此矩函数 $g^j_{s,t}$ 为零；或者二元指标不成立 $\lambda_{s,t}(w_{st},\theta_0)=0$ 这意味着函数 $g^j_{s,t}$ 非正。我将使用条件矩不等式给出条件（ID1）的等价刻画。条件（ID2）—（ID3）的刻画也可以类似地构造。

条件（ID2）通过比较商品 $\ell \in \{A,B\}$ 的需求，推导出关于参数的限制。条件（ID2）中的识别限制指示变量 $\lambda^{D_\ell}_{s,t}(w_{st},\theta)$ 定义如下，令 $\ell_{-1}\in\{A,B\}$ 且 $\ell_{-1}\neq\ell$,

$$\lambda^{D_\ell}_{s,t}(w_{st},\theta) = \mathbb{1}\Big\{\{\Delta_{s,t}x'_\ell\beta>0\} \vee \Big\{\Delta_{s,t}(x_\ell+\text{sign}(\Gamma(z))x_{\ell_{-1}})'\beta>0, |\Gamma(z)|>-\Delta_{s,t}x'_\ell\beta\Big\}\Big\}.$$

通过比较商品 $\ell\in\{A,B\}$ 的需求，可以构造如下条件矩不等式：

$$g^{D_\ell}_{s,t}(w_{st},\theta_0) = E\big[(1-\lambda^{D_\ell}_{s,t}(w_{st},\theta_0))(\mathbb{1}\{Y_{is}\in D_\ell\}-\mathbb{1}\{Y_{it}\in D_\ell\}) \mid W_{ist}=w_{st}\big] \le 0.$$

条件（ID3）根据两个不同时期两种选择的条件概率之和，推导出互补性 $\Gamma(z)$ 的上下界。条件（ID3）中识别限制的二元指标定义为

$$\lambda^L_{s,t}(w_{st},\theta) = \mathbb{1}\Big\{\{\Gamma(z)>-\min\{\Delta_{s,t}x'_A\beta,\Delta_{s,t}x'_B\beta\}\}\wedge\{\Delta_{s,t}(x_A+x_B)'\beta>0\}\Big\},$$
$$\lambda^U_{s,t}(w_{st},\theta) = \mathbb{1}\Big\{\{\Gamma(z)<\min\{\Delta_{s,t}x'_A\beta,-\Delta_{s,t}x'_B\beta\}\}\wedge\{\Delta_{s,t}(x_A-x_B)'\beta>0\}\Big\}.$$

同样，基于命题2中的条件（ID3），构造条件矩不等式如下：

$$g^L_{s,t}(w_{st},\theta_0) = E\big[(1-\lambda^L_{s,t}(w_{st},\theta_0))(\mathbb{1}\{Y_{is}=AB\}+\mathbb{1}\{Y_{it}=O\}-1) \mid W_{ist}=w_{st}\big] \le 0,$$
$$g^U_{s,t}(w_{st},\theta_0) = E\big[(1-\lambda^U_{s,t}(w_{st},\theta_0))(\mathbb{1}\{Y_{is}=A\}+\mathbb{1}\{Y_{it}=B\}-1) \mid W_{ist}=w_{st}\big] \le 0.$$

我构造了条件矩不等式，它们等价于命题2中的识别条件（ID1)-(ID3)。令$g_{s、t} = (\{g^j_{s,t}\}_{j\in\mathcal{C}}、g^{D_A}_{s\,t}、、g^{D_B}_{s\,t}、、g^L_{s、t}、g^U_{s、t})'$表示由所有条件矩函数组成的向量。识别集 Θ_I 由满足条件矩不等式的参数集合刻画。

命题 3. 在假设 *1-3*下，以下结论成立：

$$\Theta_I = \{\theta \in \Theta : g_{s,t}(w_{st}, \theta) \leq 0 \quad \forall w_{st},\ \forall s, t \leq T\}.$$

命题 3利用条件矩不等式刻画了识别集。为了进行估计和推断，我们需要为函数 Γ 指定一种参数形式（例如线性函数），使未知参数的维度有限。随后，可以使用文献中针对条件矩不等式开发的方法，对该参数进行估计和推断。

6 模拟研究

本节将通过蒙特卡洛模拟，将本文的方法与一种参数化方法进行比较。该参数化方法对 ϵ_{ijt} 施加分布假设，并对 α_{ij}施加线性结构，具体内容将在后文详细说明。模拟结果表明，无论是参数化分布设定错误，还是协变量与固定效应之间的依赖结构设定错误，都会导致互补性参数的估计量产生误导性结果。

我研究互补性的线性设定 $\Gamma(Z_i) = Z_i\gamma_0$，并考察两期模型 $T = 2$。第 4.1 节已在协变量具有大支持集的条件下确立了点识别结果，因此本节聚焦于参数点识别的情形。附录 B.6 还展示了更多模拟结果，涉及较长面板下点估计量的表现（$T \geq 2$）以及协变量具有有界支持集时集合估计量的表现。

我采用Shi、Shum 和 Song(2018年) 中用于估计的准则函数方法。根据命题 3中的条件矩不等式，准则函数可构造如下：

$$\Omega(\theta) = \sum_{s\neq t\leq T} E\Big[\|\max\{(g_{s,t}(W_{ist}, \theta), 0\}\|_1\Big] \geq \Omega(\theta_0) = 0.$$

其中 $||x||_1 = \sum_j |x^j|$ 对于向量 $x = (x^1, \ldots\ldots, x^j)'$。

与Shi、Shum 和 Song (2018年)类似，基于上述准则函数构建了一个两步估计量。第一步使用非参数估计量 $P_t(\{j\} \mid w_{st})$估计条件选择概率 $\hat{} P_t(\{j\} \mid w_{st})$。我采用单层人工神经网络估计量；该估计量的渐近性质已由Chen 和 White (1999)确立。神经网络估计量易于计算和实现，而且已有现成的软件包可供使用（Bischl 等 (2016)）。令 $\hat{} g_{s,t}$ 表示估计的矩函数，以估计量 $P_t(\{j\} \mid w_{st})$替代条件选择概率 $\hat{} P_t(\{j\} \mid w_{st})$，据此构造样本目标函数 $\Omega_N(\theta)$ 如下：

$$\hat{\Omega}_N(\theta) = \frac{1}{N} \sum_i^N \sum_{s \neq t \leq T} \| \max\{\hat{g}_{s,t}(W_{ist}, \theta), 0\} \|_1.$$

通过最小化样本目标函数 Ω_N得到参数的第二步估计量。由于参数 β_0 和 γ_0 只能在给定尺度下实现点识别，因此将这两个参数的第一个元素归一化为 1。为公平比较这些方法，其他方法也采用了这一归一化处理。

为更好地评估本文两步估计量（两步估计）的表现，我采用模拟矩方法实施参数估计量（参数估计），以供比较。对于该参数估计量，假设误差项ϵ_{ijt} 服从标准 Gumbel 分布，在不同选择和时期之间相互独立，并且也独立于所有协变量。我允许固定效应 α_{ij} 到¯ ¯通过线性设定依赖于协变量：$\alpha_{ij} = \eta_0 + X'_{ij}\eta_1 + v_{ij}$，其中 $X_{ij} = \frac{1}{T}\sum_t X_{ijt}$ 表示平均协变量，而 $v_{ij} \sim \mathcal{N}(0, 1)$服从标准正态分布，并独立于所有协变量。当所有假设均成立时，该参数估计量 $\sqrt{N}$ 是相合的；但如果参数分布或固定效应的线性模型设定有误，则可能不相合。

对于系数 β_0，我还评估了另外两种不允许捆绑购买 $\Gamma_{it} = -\infty$的估计量，以便进行比较。一种是 Chamberlain 的条件固定效应 Logit 估计量（固定效应 Logit 估计）。该估计量假设 ϵ_{ijt} 服从标准 Gumbel 分布，而对固定效应 α_{ij} 的分布不作限制。另一种是在平稳性假设下开发的半参数估计量（半参数估计），但它假设不存在商品组合。因此，该估计量仅使用以下对象的条件选择概率

$\{A, B, O\}$ 来识别系数 β_0。

现在我来介绍模拟设定。令 $d_x = 2$ 和 $d_z = 2$ 分别表示协变量 X_{it} 和 Z_i 的维度。在每次模拟中，$X_{i\ell t}$ 从正态分布 $\mathcal{N}(\mathbf{0}, d_x)$ 中抽取，并且在各个选择 $\ell \in \{A, B\}$ 和时间$t \leq T$之间相互独立。令 Z_i的第一个元素从 $\mathcal{N}(2, 2)$ 中抽取，第二个元素从 $\mathcal{N}(0, 1)$ 中抽取。真实参数设定为：$\beta_0 = \gamma_0 = (1, 1)$。

我研究误差项 ϵ_{ijt} 和固定效应 α_{ij}的四种不同设计。第一种设计考虑参数估计量的正确设定：ϵ_{ij} 服从 Gumbel 分布，固定效应设定为 $\alpha_{ij} = X'_{ij}\beta_0/2 + v_{ij}$。在第二种设计中，误差项 ϵ_{it} 服从相关性为 $\rho = -0.7$的二元正态分布。因此，在这种设计中，ϵ_{it} 的参数分布设定错误。在第三种设计中，我允许固定效应 α_{ij} 以非加性形式依赖于¯ ¯另一种商品的协变量：$\alpha_{ij} = (X_{ij}/2 - X_{ik})'\beta_0 * (1 + v_{ij})$，其中 $j \in \{A, B\}$和 $k \neq j \in \{A, B\}$均如此。在这种设计中，参数估计量对固定效应 α_{ij} 采用了错误的模型。最后一种设计结合了第二种和第三种设计；对于参数估计量而言，它同时考虑了 ϵ_{it} 的分布设定错误和 α_{ij} 的模型设定错误。以下总结了这四种设计：

- 设计 1：正确设定

$$\begin{aligned}
&\epsilon_{ijt} \sim \text{Gumbel}(0, 1), \\
&\alpha_{ij} = \bar{X}'_{ij}\beta_0/2 + v_{ij}, \quad \text{where } v_{ij} \sim \mathcal{N}(0, 1).
\end{aligned}$$

- 设计 2：设定错误的分布

$$\begin{aligned}
&\epsilon_{it} \sim \mathcal{N}_2\big([2; -2], [1 \ \ -0.7; -0.7 \ \ 1]\big), \\
&\alpha_{ij} = \bar{X}'_{ij}\beta_0/2 + v_{ij}, \quad \text{where } v_{ij} \sim \mathcal{N}(0, 1).
\end{aligned}$$

- 设计 3：固定效应设定错误

$$\begin{aligned}
&\epsilon_{ijt} \sim \text{Gumbel}(0, 1), \\
&\alpha_{ij} = (\bar{X}_{ij}/2 - \bar{X}_{ik})'\beta_0 * (1 + v_{ij}), \quad \text{where } v_{ij} \sim \mathcal{N}(0, 1).
\end{aligned}$$

- 设计 4：分布设定错误且固定效应设定错误

$$\epsilon_{it} \sim \mathcal{N}_2\big([2;-2],[1 \ \ -0.7;-0.7 \ \ 1]\big),$$
$$\alpha_{ij} = (\bar{X}_{ij}/2 - \bar{X}_{ik})'\beta_0 * (1 + v_{ij}), \quad \text{where } v_{ij} \sim \mathcal{N}(0,1).$$

对于上述四种设计，我通过报告各方法的均方根误差（rMSE）和平均绝对偏差（MAD）来比较它们。对于互补性参数 γ_0，我还报告了两种估计量错误判断替代模式符号的概率（Err），其定义为

$$\text{Err} = E|\operatorname{sign}(Z_i\gamma_0) - \operatorname{sign}(Z_i\hat{\gamma})|.$$

我研究了三种不同的样本量 $N = \{1000, 2000, 4000\}$，并将模拟重复次数设为 $B = 500$。表 γ_0 和 β_0 分别展示了估计量的表现，见表 1和表 2。

表 1 比较了两步估计量与互补性参数 γ_0的参数估计量的表现。只有在其假设设定正确时（设计 1），参数估计量的表现才更好；而在模型设定错误时（设计 2–4），其表现则更差，尤其是在参数分布和固定效应模型均设定错误时。两步估计量在全部四种设计中的表现都一致，显示出其优势：面对不同的参数分布和固定效应模型设计时，均能保持稳健表现。此外，随着样本量增加，两步估计量的偏差和偏差均显著缩小。然而，在设计 2–4 中，参数估计量的偏差并未随样本量增加而减小，这表明该估计量在模型设定错误时不具有一致性。

表2 比较了两步估计量与前文介绍的另外三种估计量在四种设计下对系数 β_0 的估计表现。同样，在设计 2–4 中，两步估计量的表现始终优于另外三种估计量，而且随着样本量增加，差异愈发显著。总而言之，表1 和表2 表明，两步估计量在面对不同参数分布，或协变量与固定效应之间不同的依赖结构设定时，均具有稳健优势。

表 1：$\hat{\gamma}$ 的性能比较

N	设计	两步估计				参数估计			
		Err	SD	均方根误差	平均绝对偏差	Err	SD	均方根误差	平均绝对偏差
1000	设计 1	0.025	0.263	0.315	0.267	0.007	0.101	0.101	0.079
	设计 2	0.026	0.309	0.341	0.278	0.058	0.344	0.741	0.663
	设计 3	0.023	0.287	0.306	0.251	0.083	0.288	1.019	0.977
	设计 4	0.025	0.329	0.332	0.264	0.126	0.681	1.764	1.628
2000	设计 1	0.022	0.258	0.281	0.229	0.005	0.067	0.067	0.054
	设计 2	0.022	0.286	0.294	0.235	0.060	0.266	0.719	0.670
	设计 3	0.022	0.295	0.298	0.234	0.085	0.213	1.021	0.998
	设计 4	0.021	0.282	0.286	0.226	0.137	0.580	1.875	1.783
4000	设计 1	0.018	0.219	0.232	0.188	0.003	0.044	0.044	0.035
	设计 2	0.021	0.270	0.270	0.223	0.060	0.195	0.700	0.672
	设计 3	0.017	0.226	0.227	0.179	0.084	0.149	0.983	0.972
	设计 4	0.020	0.250	0.271	0.218	0.141	0.437	1.873	1.822

表 2：$\hat{\beta}$ 的性能比较

N	设计	带商品组合的估计量				假设不存在商品组合的估计量			
		两步估计		参数估计		固定效应 Logit 估计		半参数估计	
		均方根误差	平均绝对偏差	均方根误差	平均绝对偏差	均方根误差	平均绝对偏差	均方根误差	MAD
1000	设计 1	0.126	0.101	0.072	0.056	0.165	0.130	0.155	0.131
	设计 2	0.125	0.101	0.432	0.389	0.243	0.187	0.161	0.134
	设计 3	0.128	0.100	0.304	0.279	0.168	0.131	0.149	0.124
	设计 4	0.121	0.095	0.647	0.593	0.256	0.200	0.171	0.143
2000	设计 1	0.093	0.073	0.051	0.040	0.117	0.092	0.114	0.092
	设计 2	0.095	0.065	0.114	0.095	0.166	0.122	0.105	0.072
	设计 3	0.094	0.076	0.300	0.286	0.128	0.101	0.115	0.092
	设计 4	0.099	0.079	0.612	0.581	0.205	0.167	0.132	0.109
4000	设计 1	0.071	0.057	0.034	0.027	0.080	0.062	0.079	0.061
	设计 2	0.075	0.060	0.385	0.368	0.157	0.135	0.091	0.072
	设计 3	0.070	0.048	0.203	0.198	0.076	0.050	0.075	0.045
	设计 4	0.077	0.063	0.606	0.588	0.168	0.144	0.104	0.085

7 扩展：潜在互补性

前几节重点讨论了互补性Γ_{it} 中的异质性仅来自可观测协变量的情形： $\Gamma_{it} = \Gamma(Z_i)$。本节允许 Γ_{it}因个体而异，不论其协变量如何，从而考虑个体之间互补性的不可观测异质性。 Γ_{it}的符号刻画了每个个体两种商品之间的异质性互补关系。因此，我重点识别 Γ_{it}符号的分布，该分布表示两种商品为互补品或替代品的人群比例。

Next，我对互补性和误差项作出一些假设 .

假设7。 给定$(\epsilon_{it}, \Gamma_{it})$ 条件下，$(\alpha_i, X_{is}, X_{it})$ 的联合分布随时间保持平稳：

$$(\epsilon_{is}, \Gamma_{is}) \mid X_{is}, X_{it}, \alpha_i \overset{d}{\sim} (\epsilon_{it}, \Gamma_{it}) \mid X_{is}, X_{it}, \alpha_i \quad for\ any\ s, t \leq T.$$

假设7与假设3类似，但还假设互补性 Γ_{it}满足平稳性条件。在互补性Γ_{it} 仅取决于可观测协变量（假设1）的情况下，假设7 退化为假设3中的平稳性条件，因为 Γ_{it} 在给定协变量时为常数。

假设7仅要求 Γ_{it} 的分布随时间保持平稳，但仍允许每个个体的互补性 Γ_{it} 随时间变化。此外，该假设不限制互补性 Γ_{it} 与不可观测项之间的依赖关系，包括固定效应 α_i 和误差项ϵ_{it}。

令 $X_i = (X_{it})_{t=1}^T$ 收集所有时期的协变量。

假设8。互补性 Γ_{it} 在给定固定效应 α_i的条件下，与协变量 X_i 独立：$\Gamma_{it} \perp\!\!\!\perp X_i \mid \alpha_i$。

假设8假定互补性与 Γ_{it} 所有时期的协变量向量相互独立。在此条件下，所有协变量的变异都可用于识别互补性的分布 $\Pr(\Gamma_{it} \geq 0)$。这一假设可以放宽，以适应存在一部分

与 Γ_{it} 独立、而其他协变量可能与Γ_{it}相关的情况。在这种情况下，可以针对可能与互补性相关的协变量进行条件分析。

在上述假设下，我得以识别两种商品互为互补品的个体所占比例，记为 $\eta = \Pr(\Gamma_{it} \geq 0)$。根据假设 7， Γ_{it} 的分布随时间保持平稳，因此 η不依赖于 t。$\Pr(\Gamma_{it} < 0)$ 的识别结果可直接由公式 $\Pr(\Gamma_{it} < 0) = 1 - \Pr(\Gamma_{it} \geq 0)$ 推导得出，因此此处省略。

识别策略 η 的直观解释如下。一种商品的条件需求可以表示为两类人群的混合：两种商品互为互补品的人群（$\Gamma_{it} \geq 0$）和两种商品互为替代品的人群（$\Gamma_{it} < 0$）。当商品 A 的协变量指数上升时，会以不同方向影响这两类人群对商品 B 的需求。这种关系有助于识别这两类人群各自所占的比例。

与第3节类似，第一步是根据条件选择概率推导协变量指数（$\Delta_{s,t}\delta_A$,$\Delta_{s,t}\delta_B$）变化的符号。令 $\xi^1_{s,t}(x_s, x_t)$ 和$\xi^2_{s,t}(x_s, x_t)$ 定义为

$$\xi^1_{s,t}(x_s, x_t) = \mathbb{1}\big\{P_s(\{j\} \mid x_s, x_t) - P_t(\{j\} \mid x_s, x_t) \geq 0,\ \forall j \in \{A, B, AB\}\big\},$$
$$\xi^2_{s,t}(x_s, x_t) = \mathbb{1}\big\{P_s(\{j\} \mid x_s, x_t) - P_t(\{j\} \mid x_s, x_t) \geq 0,\ \forall j \in \{A, AB, O\}\big\}.$$

根据观测到的选择概率的变化，可以识别两种商品的协变量指数变化的符号。当观察到所有选择的概率 $\{A, B, AB\}$ 均增加时，可以推断两种商品的协变量指数都应增加。类似地，除选择 B 外，所有选择的概率都增加，这表明商品 A 的协变量指数增加，而商品 B的协变量指数降低。

令 $\mathcal{X}^1_{s,t} = \{(x_s, x_t) \mid \xi^1_{s,t}(x_s, x_t) = 1\}$ 以及 $\mathcal{X}^2_{s,t} = \{(x_s, x_t) \mid \xi^2_{s,t}(x_s, x_t) = 1\}$ 表示协变量集合，使得 $\xi^1_{s,t}(x_s, x_t) = 1$ 和 $\xi^2_{s,t}(x_s, x_t) = 1$ 分别成立，这意味着

$$(x_s, x_t) \in \mathcal{X}^1_{s,t} \implies \Delta_{s,t}\delta_A \geq 0,\ \Delta_{s,t}\delta_B \geq 0,$$
$$(x_s, x_t) \in \mathcal{X}^2_{s,t} \implies \Delta_{s,t}\delta_A \geq 0,\ \Delta_{s,t}\delta_B \leq 0.$$

根据协变量指标的符号，利用两种商品需求的变化，可以识别两种商品为互补品的人群比例 $\eta =$

$\Pr(\Gamma_{it} \geq 0)$。为说明这一思路，我先考虑商品 A和 B的协变量指标均增加的情况：$(x_s,\ x_t) \in \mathcal{X}^1_{s,\ t}$。在这种情况下，对于两种商品为互补品的人群，两种商品的需求都会增加；但对于两种商品为替代品的人群，需求可能下降。因此，数据中任一种商品的需求下降，只可能来自满足 $\Gamma_{it} < 0$的人群，这有助于确定满足 $\Gamma_{it} < 0$ 的人群比例的下界，从而确定满足 $\Gamma_{it} \geq 0$的人群比例的上界。类似地，当协变量满足 $(x_s,\ x_t) \in \mathcal{X}^2_{s,\ t}$时，可以给出$\eta = \Pr(\Gamma_{it} \geq 0)$的下界。

下一个命题刻画了 $\eta = \Pr(\Gamma_{it} \geq 0)$ 的部分识别结果。

命题 4. 在假设 7-8下，η 可界定为 $\eta \in [L_\eta, U_\eta]$，其中

$$L_\eta = \sup_{(x_s,x_t)\in\mathcal{X}^2_{st},\ell\in\{A,B\},s,t\leq T} \Big\{(-1)^{\mathbb{1}\{\ell=A\}}[P_s(D_\ell \mid x_s, x_t) - P_t(D_\ell \mid x_s, x_t)]\Big\},$$
$$U_\eta = \inf_{(x_s,x_t)\in\mathcal{X}^1_{st},\ell\in\{A,B\},s,t\leq T} \Big\{P_s(D_\ell \mid x_s, x_t) - P_t(D_\ell \mid x_s, x_t)\Big\} + 1.$$

命题 4 通过利用不同协变量指标组合下两种商品需求的变化，为 η 确定了上下界。根据上下界的定义，始终有 $L_\eta \geq 0$，$U_\eta \leq 1$ 且这些界利用了任意两个时期内协变量所有取值的变化。界的范围取决于两种商品条件需求的变化；变化越大，界越紧。命题4的结果还为假设 7-8 提供了可检验推论，因为它意味着上界不应小于下界：$U_\eta \geq L_\eta$。命题 4 的证明见附录 B.7。

此前的研究Allen andRehbeck (2022年）也研究了潜在互补性，并利用截面数据为人群中两种商品互为互补品的比例给出了界限。他们利用排除限制，以及协变量与不可观测项之间的独立性假设。我的方法通过考察面板数据情形并利用跨期变异，补充了他们的方法。我允许协变量与不可观测固定效应之间存在任意依赖。此外，我不施加排除限制，在两种商品的协变量同时变化时仍能对 η 进行部分识别。

8 结论

本文采用带商品组合的面板多项选择模型，研究商品之间的替代与互补关系。该模型不对特质误差项施加任何参数假设，并通过允许可观测协变量与不可观测固定效应之间存在灵活的依赖结构来处理内生性。我提出了有关替代与互补关系的可检验推论，并推导出模型参数的尖锐识别集。

主要的识别策略是利用条件选择概率的跨期变异，对未知参数推导识别约束，而这些条件选择概率可由数据识别。我构造条件矩不等式，以刻画用于估计和推断的识别集。通过蒙特卡洛模拟，本文表明，与参数化方法相比，该方法在固定效应和误差项分布的不同设定下表现得更加稳健。在扩展部分，本文允许互补性存在不可观测异质性，并建立了部分识别结果。在在线附录中，我还研究了商品数量超过两种的情形、一类非可分效用函数，以及横截面模型。

本文聚焦于一种静态面板多项选择模型，其中消费者对商品的效用仅取决于同一时期的特征。值得进一步研究的是，如何在动态模型中识别互补关系；在这种模型中，消费者的选择也可能取决于过去的选择。由于必须区分以往选择的影响与当前互补性的影响，分析会更加复杂。此外，探讨如何在选择集异质且未知的情况下识别互补性和效用系数，也可能很有价值。

参考文献

R. Allen 和 J. Rehbeck。组合模型中的潜在互补性。《计量经济学杂志》，228(2):322–341，2022年。D. W. Andrews 和 X. Shi。基于条件矩不等式的推断。《计量经济学》，81(2):609–666，2013年。

T. B. Armstrong. 条件矩不等式模型中的渐近精确推断. 《计量经济学杂志》, 186(1):51–65, 2015. S. Berry, J. Levinsohn, and A. Pakes. 市场均衡中的汽车价格.《计量经济学》, pages 841–890, 1995. S. T. Berry and P. A. Haile. 利用市场层面数据识别差异化产品市场. 《计量经济学》, 82(5):1749–1797, 2014. B. Bischl, M. Lang, L. Kotthoff, J. Schiffner, J. Richter, E. Studerus, G. Casalicchio, and Z. M. Jones. mlr：R中的机器学习. 机器学习研究杂志, 17(1):5938–5942, 2016. G. Chamberlain. 定性数据的协方差分析. 经济学研究评论, 47:225–238, 1980. X. Chen and H. White. 非参数神经网络估计量的改进收敛速率与渐近正态性. *IEEE*信息论汇刊, 45(2):682– 691, 1999. V. Chernozhukov, H. Hong, and E. Tamer. 计量经济模型中参数集合的估计与置信域. 《计量经济学》, 75(5):1243–1284, 2007. V. Chernozhukov, S. Lee, and A. M. Rosen. 交集界：估计与推断. 《计量经济学》, 81(2):667–737, 2013. A. Chesher, A. M. Rosen, and K. Smolinski. 多重离散选择的工具变量模型. 《数量经济学》, 4(2):157–196, 2013. C. Dobronyi, J. Gu, and i. K. Kyoo. 带固定效应的动态面板 Logit 模型的识别. *arXiv*预印本 *arXiv:2104.04590*, 2021. F. Dunker, S. Hoderlein, and H. Kaido. 内生异质性总需求模型的非参数识别：互补品、商品组合与市场层面. *cemmap*工作论文, 2015. J. T. Fox and N. Lazzati. 关于商品组合离散选择模型与二元博弈识别的说明. 《数量经济学》, 8(3):1021–1036, 2017.

W. Y. Gao 和 M. Li。面板多项选择模型中的稳健半参数估计。可在 *SSRN 3282293* 获取，2020年。M. Gentzkow。互补性模型中的新商品价值：在线新闻。《美国经济评论》，97(3):713–744，2007年。A. Goolsbee 和 A. Petrin。直播卫星带来的消费者收益及其与有线电视的竞争。《计量经济学》，72(2):351–381，2004年。L. Grzybowski 和 P. Pereira。移动电话中通话和短信的互补性。《信息经济学与政策》，20(3):279–287，2008年。B. E. Honor´e 和 A. Lewbel。不存在严格外生解释变量的半参数二元选择面板数据模型。《计量经济学》，70(5):2053–2063，2002年。B. E. Honor´e 和 M. Weidner。具有固定效应的动态面板 Logit 模型的矩条件。*arXiv* 预印本 *arXiv:2005.05942*，2020年。A. Iaria 和 A. Wang。捆绑产品需求的识别与估计。工作论文，2020年。S. Khan、F. Ouyang 和 E. Tamer。半参数多项响应模型的推断。《数量经济学》，12(3):743–777，2021年。S. Khan、M. Ponomareva 和 E. Tamer。动态二元响应模型的识别。技术报告，即将在《计量经济学杂志》发表，2023年。C. F. Manski。基于二元面板数据的随机效应线性模型半参数分析。《计量经济学》，第357–362页，1987年。J. Monardo。利用灵活需求模型测量替代模式。可在 *SSRN 3921601* 获取，2021年。A. Pakes 和 J. Porter。具有固定效应的多项选择的矩不等式。即将在《数量经济学》发表，2022年。A. Petrin。新产品收益的量化：以小型厢式车为例。《政治经济学杂志》，110(4):705–729，2002年。X. Shi、M. Shum 和 W. Song。利用循环单调性估计半参数面板多项选择模型。《计量经济学》，86(2):737–761，2018年。

J. Zhao。电子烟和香烟：互补品还是替代品？工作论文，2019。

A 附录

在以下证明中，我将省略协变量 Z_i ，并用 Γ_0 表示增量效用 $\Gamma(Z_i)$。

A.1 引理1

证明。 引理 1 包含两个待证结果： $\Gamma_0 \geq 0$ 蕴含 $s_{AB} \leq 0$，以及$\Gamma_0 \leq 0$ 蕴含 $s_{AB} \geq 0$。我将证明第一个结果，第二种情形也可用相同思路证明。

假设互补项为正 $\Gamma_0 \geq 0$，需要证明$s_{AB} \leq 0$。根据 s_{AB}的定义，证明 $s_{AB} \leq 0$ 等价于证明：若$p_{Bs} > p_{Bt}$，则 $\Pr(Y_{is} \in D_A \mid p_{Bs},\ p_{Bt},\ \tilde{x}) \leq \Pr(Y_{it} \in D_A \mid p_{Bs},\ p_{Bt},\ \tilde{x})$。在此证明中省略协变量$\tilde{x}$ ，因为它在时间上固定，不会影响条件选择概率的变化。

令 $\beta_{0,p} \leq 0$ 表示价格 $p_{\ell t}$的系数。令 $v_{i\ell t} = \alpha_{i\ell} + \epsilon_{i\ell t}$ 表示 $\ell \in \{A, B\}$的系数。商品 B 的效用可表示为 $u_{iBt} = p_{Bt}\beta_{0,p} + v_{iBt}$ ，而商品 A 的效用为$u_{iAt} = v_{iAt}$ ，因为省略了商品 A 的所有其他协变量。

令 $\mathcal{V}_{D_A}(p_{Bt})$ 表示满足以下条件的 $v = (v_A, v_B)$ 的集合：在价格为 p_{Bt}的条件下，存在一个选择被选中于 $D_A = \{A, AB\}$ 中。集合 $\mathcal{V}_{D_A}(p_{Bt})$ 包含两种情况：选择 A 或选择 AB 的效用高于所有不属于 D_A的其他选项。集合 $\mathcal{V}_{D_A}(p_{Bt})$可表示如下：

$$\begin{aligned}\mathcal{V}_{D_A}(p_{Bt}) = &\left\{v \mid v_A \geq p_{Bt}\beta_{0,p} + v_B,\ v_A \geq 0\right\} \equiv \mathcal{V}_1(p_{Bt}) \\ &\cup \left\{v \mid v_A + \Gamma_0 \geq 0,\ v_A + p_{Bt}\beta_{0,p} + v_B + \Gamma_0 \geq 0\right\} \equiv \mathcal{V}_2(p_{Bt}).\end{aligned}$$

给定固定效应和价格后，商品 A 的需求可表示为

$$\Pr(Y_{it} \in D_A \mid \alpha_i, p_{Bs}, p_{Bt}) = \Pr(v_{it} \in \mathcal{V}_{D_A}(p_{Bt}) \mid \alpha_i, p_{Bs}, p_{Bt}).$$

在假设 3 （平稳性）下， v_{it} 的条件分布随时间保持平稳，因为 ϵ_{it} 的条件分布在不同时间相同，并且

固定效应 α_i 保持不变。因此，更大的集合将意味着更高的条件概率，如下所示：

$$\mathcal{V}_{D_A}(p_{B\,s}) \subseteq \mathcal{V}_{D_A}(p_{Bt}) \Longrightarrow \Pr(Y_{is} \in D_A \mid \alpha_i, p_{Bs}, p_{Bt}) \leq \Pr(Y_{it} \in D_A \mid \alpha_i, p_{Bs}, p_{Bt}).$$

在给定协变量的条件下，对固定效应 α_i 取期望，上述条件可得

$$\mathcal{V}_{D_A}(p_{Bs}) \subseteq \mathcal{V}_{D_A}(p_{Bt}) \Longrightarrow \Pr(Y_{is} \in D_A \mid p_{Bs}, p_{Bt}) \leq \Pr(Y_{it} \in D_A \mid p_{Bs}, p_{Bt}).$$

证明引理 1 等价于证明，更高的价格 $p_{Bs} > p_{Bt}$ 意味着$\mathcal{V}_{D_A}(p_{Bs}) \subseteq \mathcal{V}_{D_A}(p_{Bt})$。我将通过证明 $\mathcal{V}_{D_A}(p_{Bs}) \subseteq \mathcal{V}_{D_A}(p_{Bt})$，来说明对于任意元素 $v \in \mathcal{V}_{D_A}(p_{Bs})$，它都满足 $v \in \mathcal{V}_{D_A}(p_{Bt}$。证明将通过讨论以下两种情形展开： $v \in \mathcal{V}_1(p_{Bs}$和 $v \in \mathcal{V}_2(p_{Bs})$。

情形1： $v \in \mathcal{V}_1(p_{Bs})$。如果 v 满足 $v_A \geq p_{Bt}\beta_{0,\ p} + v_B$ 那么 $v \in \mathcal{V}_1(p_{Bt})$。否则，$v$ 应满足

$$v_A < p_{Bt}\beta_{0,p} + v_B,\ v_A \geq 0.$$

由于 $\Gamma_0 \geq 0$，可得以下推论：

$$v_A + \Gamma_0 \geq 0,\ (v_A + \Gamma_0) + p_{Bt}\beta_{0,p} + v_B > (v_A + \Gamma_0) + v_A \geq 0.$$

因此我们知道 $v \in \mathcal{V}_2(p_{Bt}) \subseteq \mathcal{V}_{D_A}(p_{Bt})$。

情形2： $v \in \mathcal{V}_2(p_{Bs})$。根据集合 $\mathcal{V}_2(p_{Bs})$的定义，在给定$\beta_{0,\ p} \leq 0$的条件下，它随 p_{Bs} 递减。由于 $p_{Bs} > p_{Bt}$，这意味着 $v \in \mathcal{V}_2(p_{Bs}) \subseteq \mathcal{V}_2(p_{Bt})$。

我已经证明，对于任意元素 $v \in \mathcal{V}_{D_A}(p_{Bs})$，当$p_{Bs} > p_{Bt}$时，它满足 $v \in \mathcal{V}_{D_A}($ $p_{Bt})$。因此，我们可以得出 $\Gamma_0 \geq 0$ 蕴含 $s_{AB} \leq 0$。

□

A.2 命题的证明2

证明。令 $v_{i\ell t} = \alpha_{i\ell} + \epsilon_{i\ell t}$ 表示 $\ell \in \{A, B\}$ 的固定效应与误差项之和。对于任意集合 $K \subset \mathcal{C}$，令 $\mathcal{V}_K(x_t)$ 表示满足以下条件的 $v = (v_A,\ v_B)$ 的集合：给定 $X_{it} = x_t$ 时，存在一个被选择的 $K \subset \mathcal{C}$。令 $v_{AB} = v_A + v_B + \Gamma_0$和 $v_O = 0$ 分别表示商品组合 AB 和外部选项的误差项。

集合 $\mathcal{V}_K(x_t)$ 可以表示为

$$\mathcal{V}_K(x_t) = \big\{v \mid \exists j \in K \text{ s.t. } \delta_{jt} + v_j \geq \delta_{kt} + v_k \text{ for all } k \in K^c\big\},$$

其中 $\delta_{\ell t} = x'_{\ell t}\beta_0$ 对于 $\ell \in \{A, B\}$、$\delta_{ABt} = \delta_{At} + \delta_{Bt}$ 和 $\delta_{Ot} = 0$。

集合 K 中存在一个选择被选中的条件概率可表示如下：

$$\Pr(Y_{it} \in K \mid \alpha_i, x_s, x_t) = \Pr\big(v_{it} \in \mathcal{V}_K(x_t) \mid \alpha_i, x_s, x_t\big).$$

在假设 3 （平稳性）下， v_{it} 的条件分布随时间保持平稳，对任意 $s \neq t$ 均如此，因为固定效应 α_i 不随时间变化。因此，较大的集合意味着随时间变化的条件选择概率较大：

$$\mathcal{V}_K(x_s) \subseteq \mathcal{V}_K(x_t) \implies \Pr(Y_{is} \in K \mid \alpha_i, x_s, x_t) \leq \Pr(Y_{it} \in K \mid \alpha_i, x_s, x_t). \quad (2)$$

接下来，我们将给出参数 θ_0 满足集合关系 $\mathcal{V}_K(x_s) \subseteq \mathcal{V}_K(x_t$的充分条件；该关系意味着集合 K 的条件概率会随时间下降。于是，通过逆否命题，当选择概率随时间上升时，可以推导出 θ_0 的识别限制。命题 2 包含三部分识别限制，每一部分的证明将在下文依次给出。

第1部分：随时间比较选择 $j \in \mathcal{C}$ 的条件概率。根据集合 $\mathcal{V}_j(x_t)$ 的定义，它随 $\delta_{jt} - \delta_{kt}$ 的变化而增加（$k \neq j$）。因此，当选择 j 相对于所有其他选择的协变量指数随时间降低时，可推出以下集合关系：

$$\delta_{js} - \delta_{ks} \leq \delta_{jt} - \delta_{kt} \ \forall k \neq j \implies \mathcal{V}_j(x_s) \subseteq \mathcal{V}_j(x_t).$$

注意，对于选择 AB 在 $\delta_{ABt} = \delta_{At} + \delta_{Bt}$时，上述关系同样成立。代入记号 $\Delta_{s,\ t}\delta_j = \delta_{js} - \delta_{jt}$ 以及条件（2），可得以下推论：

$$\Delta_{s,t}\delta_j - \Delta_{s,t}\delta_k \leq 0 \ \forall k \neq j \implies \Pr(Y_{is} = \{j\} \mid \alpha_i, x_s, x_t) \leq \Pr(Y_{it} = \{j\} \mid \alpha_i, x_s, x_t).$$

由连续性重新定位并对 α_i 取关于 (x_s, x_t) 的条件期望，可得

命题中第一个识别限制 2:

$$P_s(\{j\} \mid x_s, x_t) > P_t(\{j\} \mid x_s, x_t) \Longrightarrow \exists\, k \neq j \text{ s.t. } \Delta_{s,t}\delta_j - \Delta_{s,t}\delta_k > 0.$$

第2部分: 随时间比较对商品 $\ell \in \{A, B\}$ 的需求。以商品 A 为例说明证明。集合 $\mathcal{V}_{D_A}(x_t)$ 可表示为以下两个集合的并集: 选择 A 的集合, 以及效用高于不属于 D_A 的其他选择的选择 AB 的集合,

$$\begin{aligned}\mathcal{V}_{D_A}(x_t) = &\left\{v \mid \delta_{At} + v_A \geq \delta_{Bt} + v_B,\ \delta_{At} + v_A \geq 0\right\} \equiv \mathcal{V}_1(x_t)\\ &\cup \left\{v \mid \delta_{At} + v_A + \Gamma_0 \geq 0, \delta_{At} + \delta_{Bt} + v_{AB} \geq 0\right\} \equiv \mathcal{V}_2(x_t).\end{aligned}$$

为证明条件（ID2）, 命题2中, 我考察（ID2）的逆否命题, 表述如下:

$$\left\{\Delta_{s,t}\delta_A \leq 0,\ \Delta_{s,t}(\delta_A + \text{sign}(\Gamma_0)\delta_B) \leq 0\right\} \vee \left\{|\Gamma_0| \leq -\Delta_{s,t}\delta_A\right\} \Longrightarrow \mathcal{V}_{D_A}(x_s) \subseteq \mathcal{V}_{D_A}(x_t). \tag{3}$$

若上述条件成立, 则同理, 可通过逆否命题并对 α_i取条件期望来证明条件（ID2）。式（3）中关于参数θ_0 的条件还取决于互补性 Γ_0的符号。我着重讨论 $\Gamma_0 > 0$ 的情形, 相关思路也适用于 $\Gamma_0 \leq 0$的情形。

当 $\Gamma_0 > 0$时, 式（3）中对参数 θ_0 的限制包括两部分:
$C_1 = \left\{\Delta_{s,t}\delta_A \leq 0, \Delta_{s,t}(\delta_A + \delta_B) \leq 0\right\}$ 和 $C_2 = \{\Gamma_0 \leq -\Delta_{s,t}\delta_A\}$。现在需要证明两个条件 $C_1 \vee C_2$ 之一蕴含 $\mathcal{V}_{D_A}(x_s) \subseteq \mathcal{V}_{D_A}(x_t)$。只需证明在条件 C_1 或 C_2下, 属于 $\mathcal{V}_{D_A}(x_s)$的任意元素 v 也属于 $\mathcal{V}_{D_A}(x_t)$即可。对于任意元素 $v \in \mathcal{V}_{D_A}(x_s)$, 我分两种情况讨论: $v \in \mathcal{V}_1(x_s)$和 $v \in \mathcal{V}_2(x_s)$。

情形*1*: $v \in \mathcal{V}_1(x_s)$。如果 v满足 $\delta_{At} + v_A \geq \delta_{Bt} + v_B$, 那么 $v \in \mathcal{V}_1(x_t)$, 因为条件 C_1 或 C_2 中的任一条件都蕴含 $\delta_{As} \leq \delta_{At}$。否则, v 应满足以下不等式:

$$\delta_{Bt} + v_B > \delta_{At} + v_A \geq \delta_{As} + v_A \geq 0.$$

由于 互补性为非负数 $\Gamma_0 \geq 0$, 以下条件成立 :

$$\delta_{At} + v_A + \Gamma_0 \geq 0,\ (\delta_{At} + v_A) + (\delta_{Bt} + v_B) + \Gamma_0 \geq 0.$$

因此， $v \in \mathcal{V}_2(x_t) \subseteq \mathcal{V}_{D_A}(x_t)$。

情形*2*： $v \in \mathcal{V}_2(x_s)$。首先考虑条件 C_1 成立的情况。根据集合 $\mathcal{V}_2(x_s)$ 的定义，当指标 δ_{As} 和 $\delta_{As} + \delta_{Bs}$ 均增加时，该集合会扩大。条件 C_1 意味着协变量指标 $\delta_{At} \geq \delta_{As}$ 和$\delta_{At} + \delta_{Bt} \geq \delta_{As} + \delta_{Bs}$增加，因此 $v \in \mathcal{V}_2(x_t)$。

现在考虑条件 C_2 成立的情况。对于任意元素 $v \in \mathcal{V}_2(x_s)$，条件 C_2意味着以下条件：

$$\delta_{At} + v_A \geq \delta_{As} + \Gamma_0 + v_A \geq 0.$$

如果 v还满足 $\mathcal{V}_2(x_t)$ 中的第二个条件，即 $\delta_{At}+\delta_{Bt}+v_A+v_B+\Gamma_0 \geq 0$，那么 v 属于集合 $\mathcal{V}_2(x_t)$： $v \in \mathcal{V}_2(x_t)$。否则， v 应满足

$$\delta_{Bt} + v_B < -(\delta_{At} + v_A + \Gamma_0) \leq \delta_{At} + v_A.$$

这意味着 $v \in \mathcal{V}_1(x_t)$。我已证明，无论 $v \in \mathcal{V}_{D_A}(x_s)$，在条件 C_1 或 C_2下，它都满足$v \in \mathcal{V}_{D_A}(x_t)$。

第 3 部分：随时间变化的两种选择的条件概率之和的比较。条件（ID3）包含两部分识别约束：一是购买单一商品的条件概率之和，二是购买组合商品和外部选项的条件概率。我重点讨论使用购买单一商品的条件概率之和的条件。

同样，我考察命题ID3中的条件2的逆否命题。令 $C_3 = \{$最小值 $\{\Delta_{s,t}\delta_A, -\Delta_{s,t}\delta_B\} \leq \Gamma_0\}$ 以及 $C_4 = \{\Delta_{s,t}(\delta_A - \delta_B) \leq 0\}$，条件（ID3）的逆否命题可表示为

$$C_3 \vee C_4 \implies \mathcal{V}_A(x_s) \subseteq \mathcal{V}_{\{A,AB,O\}}(x_t),$$

其中，集合 $\mathcal{V}_A(x_t)$和$\mathcal{V}_{A,\ AB,\ O}(x_t)$表示为

$$\mathcal{V}_A(x_t) = \{v \mid \delta_{At} + v_A \geq 0,\ \delta_{At} + v_A \geq \delta_{Bt} + v_B,\ \delta_{Bt} + v_B + \Gamma_0 \leq 0\},$$
$$\mathcal{V}_{A,AB,O}(x_t) = \{v \mid \delta_{At} + v_A \geq \delta_{Bt} + v_B \text{ or } \delta_{At} + v_A + \Gamma_0 \geq 0 \text{ or } 0 \geq \delta_{Bt} + v_B\}.$$

Fir先考虑条件 C_3 成立的情形，它具有以下含义：

$$\delta_{As} \leq \delta_{At} + \Gamma_0 \quad \text{or} \quad \delta_{Bs} + \Gamma_0 \geq \delta_{Bt}.$$

对于任意元素 $v \in \mathcal{V}_A(x_s)$，条件 C_3 意味着

$$\delta_{At} + v_A + \Gamma_0 \geq 0 \quad \text{or} \quad 0 \geq \delta_{Bt} + v_B.$$

因此得出结论：$v \in \mathcal{V}_{A,\ AB,\ O}(x_t)$。

当条件 C_4 成立时，可推出 $\delta_{As} - \delta_{Bs} \leq \delta_{At} - \delta_{Bt}$。那么，满足 $v \in \mathcal{V}_A(x_s)$ 的元素v 也满足 $\delta_{At} + v_A \geq \delta_{Bt} + v_B$，由此可得 $v \in \mathcal{V}_{A,\ AB,\ O}(x_t)$。购买该商品组合和选择外部选项的条件概率之和的分析类似，因此这里省略。

□

A.3 定理1的证明

证明。为证明尖锐性，我需要说明，对于识别集 Θ_I 中的任意参数 θ，都可以构造一个数据生成过程，使其与观测到的选择概率相符并满足假设。

令 $X_i = (X_{it})_{t=1}^T$ 和 $Y_i = (Y_{it})_{t=1}^T$ 分别汇集所有时期的协变量和选择变量。令 $F_{Y|X}(j_1,\ j_2、\ \ldots\ldots、\ j_T \mid x)$ 表示给定 $X_i = x$ 时所有时期 $t \leq T$ 均选择 $j_t \in \mathcal{C}$ 的联合选择概率，该概率可由数据识别。我将固定效应设为零 $\alpha_i = 0$，并着重构造误差项 $\epsilon_i \mid x$ 的条件分布。

尖锐性的第一项要求是，根据所有选择的效用函数，所构造的误差项分布能够匹配观测到的选择概率 $F_{Y|X}(j_1,\ j_2,\ \ldots\ldots j_T \mid x)$：

$$F_{Y|X}(j_1, j_2, ..., j_T \mid x) = \Pr(u_{ij_tt} \geq u_{ik_tt} \quad \forall k_t \neq j_t, \forall t \leq T \mid x). \tag{4}$$

左侧项表示数据中的观测到的选择概率，右侧项表示由模型生成的选择概率，该概率取决于所构造的误差项分布。

第二个要求是，所构造的误差项分布满足假设 3（平稳性）；给定（X_{is}，X_{it}），这等价于以下条件 $=$（x_s，x_t）：

$$\Pr(\epsilon_{is} \in K \mid x_s, x_t) = \Pr(\epsilon_{it} \in K \mid x_s, x_t) \quad \text{for any set } K. \tag{5}$$

为了构造满足上述两个要求的误差项条件分布，第一步是构造选择集。由于数据中只能观测到消费者的选择，我将定义能够确定观测到的选择的选择集，并构造 $\epsilon_i \mid x$ 在这些选择集上的条件分布。其他选择集上的分布对观测变量并不重要，只要满足假设即可任意构造。令 $\mathcal{E}_K(x_t)$ 表示满足以下条件的 $\epsilon = (\epsilon_A, \epsilon_B)$ 集合：给定 $X_{it} = x_t$ 时，该集合中存在一个选择 K 被选中，定义为

$$\mathcal{E}_K(x_t) = \{\epsilon \mid \exists j \in K \text{ s.t. } \delta_{jt} + \epsilon_j \geq \delta_{kt} + \epsilon_k \quad \forall k \in K^c \mid x_t\},$$

其中 $\epsilon_{AB} = \epsilon_A + \epsilon_B + \Gamma_0$ 和 $\epsilon_O = 0$。当 $K = \{j\}$ 是单元素集合时，$\mathcal{E}_j(x_t)$ 是满足给定 x_t 时选择 j 的误差项集合。

四个选择集 $\{\mathcal{E}_j(x_t)\}_{j\in\mathcal{C}}$ 构成了在给定x_t条件下 ϵ_{it} 空间的划分。选择选择 j 的条件概率可表示如下：

$$\Pr(Y_{it} = j \mid x_t) = \Pr(\epsilon_{it} \in \mathcal{E}_j(x_t) \mid x_t).$$

对于任意 $j_t \in \mathcal{C}$，构造 $\epsilon_i \mid x$ 在集合 $\mathcal{E}_j(x_t)$ 上的条件分布如下，即可满足第一个要求（条件 (4)）：

$$F_{Y|X}(j_1, j_2, ..., j_T \mid x) = \Pr(\epsilon_{i1} \in \mathcal{E}_{j_1}(x_1), ..., \epsilon_{iT} \in \mathcal{E}_{j_T}(x_T) \mid x). \tag{6}$$

通过匹配观测到的选择概率，便确定了 $\epsilon_i \mid x$ 在选择集 $\mathcal{E}_j(x_t$上的联合分布。现在只需验证条件 (5) 中的平稳性假设可以满足。为此，我将构造 $\epsilon_{it} \mid (x_s, x_t)$ 在较小集合上的边际分布，使其在任意两个时期 $s \neq t$ 之间保持平稳，并且与根据方程 (6) 推导出的选择集分布相合。

式（6）限制了每个 $t \leq T$下，$\epsilon_{it} \mid (x_s, x_t)$在选择集$\mathcal{E}_j(x_t)$上的边际分布。选择集 $\mathcal{E}_j(x_t)$取决于 x_t，因此当协变量 x_t 变化时，它也会随时间变化。由于两个分布在不同的时期定义于不同的集合上，因此很难对其进行比较并验证平稳性假设。为解决这一难题，我构造了 $\epsilon_{is} \mid (x_s, x_t)$和 $\epsilon_{it} \mid (x_s, x_t)$在两个选择集 $\mathcal{E}_j(x_s)$和 $\mathcal{E}_j(x_t)$的交集上的边际分布，从而使它们的分布定义于同一集合上。令$J_{j,k}(x_s, x_t)$表示这两个集合的交集

$\mathcal{E}_j(x_s)$ 和 $\mathcal{E}_j(x_t)$，定义如下：

$$J_{j,k}(x_s, x_t) = \mathcal{E}_j(x_s) \cap \mathcal{E}_k(x_t).$$

令 $P_t(j \mid x_s, x_t) = \Pr(Y_{it} = j \mid x_s, x_t)$表示选择 j 在时间 t 的边际概率，该概率可由数据识别。关于 $\epsilon_{it} \mid (x_s, x_t)$在集合$J_{j,k}(x_s, x_t)$上的条件分布的要求，等价于以下条件：对于任意 $j, k \in \mathcal{C}$ 和任意 $s \neq t$,

$$\begin{aligned} \Pr(\epsilon_{is} \in J_{j,k}(x_s, x_t) \mid x_s, x_t) &= \Pr(\epsilon_{it} \in J_{j,k}(x_s, x_t) \mid x_s, x_t), \\ \sum_k \Pr(\epsilon_{is} \in J_{j,k}(x_s, x_t) \mid x_s, x_t) &= P_s(j \mid x_s, x_t), \\ \sum_j \Pr(\epsilon_{it} \in J_{j,k}(x_s, x_t) \mid x_s, x_t) &= P_t(k \mid x_s, x_t). \end{aligned} \tag{7}$$

第一个方程保证了条件平稳性假设（假设3），另外两个方程则确保所构造的 $\epsilon_{it} \mid (x_s, x_t)$边际分布与观测到的选择概率相合。

现在我需要证明，在集合 $J_{j,k}(x_s, x_t)$ 上存在非负的 $\epsilon_{it} \mid (x_s, x_t)$ 概率，使得条件（7）成立。令 $r_{j,k}(x_s, x_t) \geq 0$ 表示集合 $J_{j,k}$ 上的条件概率：

$$r_{j,k}(x_s, x_t) = \Pr(\epsilon_{is} \in J_{j,k}(x_s, x_t) \mid x_s, x_t) = \Pr(\epsilon_{it} \in J_{j,k}(x_s, x_t) \mid x_s, x_t).$$

由于概率$r_{j,k}(x_s, x_t)$不随时间变化，因此满足平稳性假设。为简化记号，在以下分析中，我将省略条件概率$r_{j,k}(x_s, x_t)$ 和 $P_t(j \mid x_s, x_t)$中的协变量（x_s, x_t）。我只需构造非负概率$r_{j,k} \geq 0$，使得对于所有 $j, k \in \mathcal{C}$，条件（7）中的最后两个条件成立：

$$\begin{aligned} \sum_k r_{j,k} &= P_s(j), \\ \sum_j r_{j,k} &= P_t(k). \end{aligned} \tag{8}$$

此证明聚焦于情形 $\Gamma_0 \geq 0$，其思路同样适用于情形 $\Gamma_0 < 0$。$r_{j,k}$ 的构造取决于协变量指标与互补项 $\{\Delta_{s,t}\delta_A, \Delta_{s,t}\delta_B, \Delta_{s,t}\delta_{AB}, \Gamma_0\}$之间的关系，因为这种关系

决定了选择集之间的关系。我讨论以下情形，以展示 $r_{j,k}$的构造。

情形1：$\Delta_{s,\ t}\delta_A \geq \Delta_{s,\ t}\delta_{AB} \geq 0 \geq \Delta_{s,\ t}\delta_B$，以及 $\Gamma_0 \geq$ 最小值 $\{\Delta_{s,t}\delta_A, -\Delta_{s,t}\delta_B\}$。根据 A.2 中命题 2 的证明，协变量指标与互补项之间的关系意味着以下集合包含关系：

$$\begin{aligned}
&\mathcal{E}_J(x_t) \subseteq \mathcal{E}_J(x_s) \quad \text{for any } J \in \{\{A\}, \{A, AB\}, \{A, AB, O\}\}, \\
&\mathcal{E}_B(x_t) \subseteq \mathcal{E}_{\{B,AB,O\}}(x_s).
\end{aligned}$$

根据 $J_{j\text{的定义},\ k}$，上述集合包含关系意味着以下集合为空：

$$J_{k_1,A} = J_{k_2,AB} = J_{B,O} = J_{A,B} = \emptyset \quad \text{for } k_1 \neq A, k_2 = \{B, O\}.$$

鉴于协变量指标与互补项之间的关系，命题 2 中条件（ID1）-（ID3）的逆否命题等价于以下不等式：

$$\begin{aligned}
P_t(A) &\leq P_s(A), \\
P_t(B) &\geq P_s(B), \\
P_t(A) + P_t(AB) &\leq P_s(A) + P_s(AB), \\
P_t(B) + P_s(A) &\leq 1.
\end{aligned} \tag{9}$$

现在需要证明，当上述限制（9）成立时，在非空集合 $r_{j,\ k} \geq 0$ 上存在非负概率 $J_{j,\ k}$，使得（8）成立。

非空集合上的概率$J_{j,\ k}$ 构造如下：

$$\begin{aligned}
&r_{O,B} = \min\{P_t(B) - P_s(B), P_s(O)\}, \quad r_{AB,B} = P_t(B) - P_s(B) - r_{O,B}, \\
&r_{O,O} = P_s(O) - P_{O,B}, \quad r_{A,AB} = \min\{P_s(A) - P_t(A), P_t(AB)\}, \\
&r_{A,O} = P_s(A) - P_t(A) - r_{AB,A}, \quad r_{AB,AB} = P_t(AB) - r_{AB,A}.
\end{aligned}$$

最后一个概率$r_{AB,\ O}$ 可按如下方式构造：

$$r_{AB,O}=\begin{cases}1-P_t(B)-P_s(A) & \text{if } P_s(A)\geq P_t(\{A,AB\}), P_t(B)\geq P_s(\{B,O\}),\\ P_s(AB) & \text{if } P_s(A)\geq P_t(\{A,AB\}), P_t(B)\leq P_s(\{B,O\}),\\ P_t(O) & \text{if } P_s(A)\leq P_t(\{A,AB\}), P_t(B)\geq P_s(\{B,O\}),\\ P_s(\{A,AB\})-P_t(\{A,AB\}) & \text{if } P_s(A)\leq P_t(\{A,AB\}), P_t(B)\leq P_s(\{B,O\}).\end{cases}$$

根据定义，可以验证所有构造出的概率均非负；在条件（9）下，概率 $r_{AB,\ O}$ 也非负。对于以下情形，构造非负概率 $r_{j,\ k}$ 的思路类似。

情形2：$\Delta_{s,\ t}\delta_A \geq \Delta_{s,\ t}\delta_{AB} \geq 0 \geq \Delta_{s,\ t}\delta_B$ 且 $\Gamma_0 < \min\{\Delta_{s,t}\delta_A, -\Delta_{s,t}\delta_B\}$。这意味着以下集合包含关系：

$$J_{k_1,A}=J_{k_2,AB}=J_{k_3,O}=\emptyset \quad \text{for } k_1\neq A, k_2=\{B,O\}, k_3=\{B,AB\}.$$

根据协变量指标与互补项之间的关系，命题2 中条件的逆否命题推出以下不等式：

$$\begin{aligned}P_t(A)&\leq P_s(A),\\ P_t(B)&\geq P_s(B),\\ P_t(A)+P_t(AB)&\leq P_s(A)+P_s(AB),\\ P_t(B)+P_t(AB)&\geq P_s(B)+P_s(AB).\end{aligned} \tag{10}$$

概率 $r_{j,\ k}$ 可构造如下：

$$\begin{aligned}&r_{B,B}=P_s(B), && r_{A,A}=P_t(A),\\ &r_{O,O}=\min\{P_t(O),P_s(O)\}, && r_{O,B}=P_s(O)-r_{O,O}, && r_{A,O}=P_t(O)-r_{O,O},\\ &r_{AB,AB}=\min\{P_t(AB),P_s(AB)\}, && r_{AB,B}=P_s(AB)-r_{AB,AB}, && r_{A,AB}=P_t(AB)-r_{AB,AB}.\end{aligned}$$

根据构造方式，上述概率均非负。最后一个概率确定为 $r_{A,\ B}=P_s(A)+P_t(B)-1+r_{AB,\ AB}+r_{O,\ O}$，且可证明，当条件（10）成立时，该概率非负。

其余情形的分析也类似，只需交换 A 和 B 的顺序，或交换 AB 和 O。

□

A.4 定理2的证明

证明。 我先证明系数 β_0 的点识别；对于参数 γ_0，思路类似。首先需要证明，对于任意候选参数 $b \neq k\beta_0$，都存在某个协变量取值，使得在真实参数 β_0 和候选参数 b 下，商品 $\ell \in \{A, B\}$ 的协变量指数 $\Delta X'_{i\ell}\beta$ 符号不同。下面以商品 A 为例说明这一思路。

根据假设5，$\Delta X_{iA}^{k_A}$ 的条件密度处处为正。令$\Delta \tilde{X}_{iA} = \Delta X_{iA} \setminus \Delta X_{iA}^{k_A}$表示 ΔX_{iA} 中的剩余协变量， $\tilde{\beta}_0$ 表示其系数。考虑$\Delta X_{iA}^{k_A}$的系数为正 $\beta_0^{k_A} > 0$的情况，分析适用于 $\beta_0^{k_A} < 0$情形。对于任意候选 b，我讨论三种情形：$b^{k_A} < 0$、$b^{k_A} = 0$和 $b^{k_A} > 0$。

情形1： $b^{k_A} < 0$。当协变量 $\Delta X_{iA}^{k_A}$ 取较大的正值 $\Delta X_{iA}^{k_A} = \Delta x_A^{k_A} \to +\infty$ ，且所有其他协变量在其支撑集中取有界值时，这意味着 $\Delta x'_A\beta_0 > 0$ 和 $\Delta x'_A b < 0$ ，因为真实系数与候选系数符号不同 $\beta_0^{k_A} > 0$ 和 $b^{k_A} < 0$。

情形2： $b^{k_A} = 0$。对于任意值 $\Delta X_{iA} = \Delta x_A$， $\Delta x'_A b$ 的值要么为正，要么为非正。首先考虑 $\Delta x'_A b > 0$ 为正的情况。当 $\Delta x_A^{k_A}$取一个足够大的负数 $\Delta x_A^{k_A} \to -\infty$ ，使得 $\Delta x'_A\beta_0 < 0$，其符号与$\Delta x'_A b$不同。类似地，如果 $\Delta x'_A b \leq 0$，则存在 $\Delta x_A^{k_A} \to +\infty$ 使得 $\Delta x'_A\beta_0 > 0$。

情形3： $b^{k_A} > 0$。假设5 指出，协变量 ΔX_{iA} 的˜支撑集不包含于任何真线性子空间，因此存在 $\Delta X_{iA} = \Delta \tilde{x}_A$ 使得$\Delta \tilde{x}'_A\tilde{\beta}_0/\beta_0^{k_A} \neq \Delta \tilde{x}'_A\tilde{b}/b^{k_A}$。假设 $\Delta \tilde{x}'_A\tilde{\beta}_0/\beta_0^{k_A} - \Delta \tilde{x}'_A\tilde{b}/b^{k_A} = k > 0$，那么当协变量取值为 $\Delta X_{iA}^{k_A} = -\Delta \tilde{x}'_A\tilde{b}/b^{k_A} - \epsilon$ 且 $0 < \epsilon < k$时。协变量指数的符号满足： $\Delta x'_A\beta_0 = \beta_0^{k_A}(k - \epsilon) > 0$ 且 $\Delta x'_A b = -b^{k_A}\epsilon < 0$。当 $k < 0$时，构造方法类似。

我已经证明，存在某个协变量取值，使得指数的符号 $\Delta X'_{i\ell}\beta$ 在真实参数 β_0 和候选参数b下不同。根据假设4，至少存在一个选择，使得该选择的条件概率在 β_0 和 b 下变化方向严格相反，因此β_0 可识别。例如，如果 $\Delta x'_\ell\beta_0 > 0$ 且 $\Delta x'_\ell b \leq 0$ 对 $\ell \in \{A, B\}$而言，则选择 AB 的条件概率在 β_0 下严格增加，而在 b 下减少。其他情形的分析类似。

现在我们来看参数 γ_0的证明，其与上述分析类似。协变量指数$\Delta x'_\ell\beta_0$ 变化的符号对于 $\ell \in \{A, B\}$

是可点识别的，因为 β_0 是点识别的。对于任意 $\gamma \neq k\gamma_0$，我们知道，在 Z_i 的大支持假设（假设6）下，存在某个协变量取值 $Z_i = z$，使得互补性 $z'\gamma_0 \neq 0$ 的符号与 $z'\gamma$ 的符号不同。于是，在给定某些协变量取值时，商品 A 或商品 B 的条件需求在 γ_0 和 γ 下会朝着截然不同的方向变化。例如，假设商品 A 的协变量随时间保持不变，而商品 B 的协变量增加$\Delta x'_B\beta_0 > 0$。如果 $z'\gamma_0 > 0$ 和 $z'\gamma \leq 0$，那么根据命题2中的条件ID2，商品 A 的需求在 γ_0下会严格增加，而在 γ 下会减少。对于另一种情况 $z'\gamma_0 < 0$ 和 $z'\gamma \geq 0$，结论则相反。因此，参数 γ_0 是点识别的。

□

B 在线附录

附录 B.1 扩展：两种以上商品

虽然许多应用都关注两种商品之间的替代关系，但探讨如何扩展这一方法以适用于两种以上商品也很有意义。在本节中，我允许商品数量为任意有限数量 $\mathcal{L} = \{A, B, ..., J\}$。用字母 J表示最后一种商品，其他字母则可用于表示更多商品。令 Γ_{ibt}表示选择商品组合 b 所带来的增量效用，即该效用与单独选择商品 $\ell \in \mathcal{L}$ 在 b中的效用之和相比：

$$\Gamma_{ibt} = u_{ibt} - \sum_{\ell \in b} u_{i\ell t}.$$

我对 Γ_{ibt} 采用与假设 1相同的假设，即增量效用仅取决于观测到的协变量 Z_i:

$$\Gamma_{ibt} = \Gamma_b(Z_i).$$

函数 Γ_b 可以因商品组合而异，因此不同商品组合的增量效用可以具有异质性。在本附录中，我省略协变量 Z_i ，因此用 Γ_b 表示商品组合 b的增量效用。

将分析推广到两种以上商品时，问题会变得更加复杂，并带来若干挑战。若允许商品的所有可能组合，可行选择的数量将呈指数级增长。与单独购买一种商品相比，所有商品组合的增量效用可能各不相同，且其正负皆有可能。对某种商品的需求取决于所有商品的协变量指标以及所有可能商品组合的增量效用。在这种情况下，某种商品的需求与所有可能商品组合的增量效用以及所有协变量指标之间的关系将变得难以处理。

鉴于这些挑战，为了简化对两种以上商品的分析，我考虑了两个假设。第一个条件要求消费者在一个时期内最多只能购买两种商品。作出这一假设后，包含两种以上商品的商品组合便被排除，从而减少了消费者可行选择的数量。这一假设或许是合理的，因为消费者可能受到预算或库存限制，无法同时购买两种以上

商品。在这一条件下，消费者的选择集简化为
$\mathcal{C} = \{O, j, j_1 j_2 \ \forall j \in \mathcal{L}, j_1 \neq j_2 \in \mathcal{L}\}$。其次，我考虑的重点是商品 A 和 B之间的替代关系，并允许其增量效用Γ_{AB} 为正或负。假设所有其他商品组合的增量效用均为负： $\Gamma_{j_1 j_2} \leq 0$ 对于 $j_1 \neq j_2$ 和 $j_1 j_2 \neq AB$。这一条件仍允许消费者购买商品组合 $j_1 j_2$，仅限制其增量效用的符号。

在上述两个假设下，以下命题刻画了 (β_0， Γ_{AB}) 的识别约束。

命题5。在假设*1*-*3*下，以下条件成立：对于任意 $s \neq t \leq T$，

- 商品 $j \in \mathcal{L}$ 的条件选择概率比较：

$$P_s(\{j\} \mid x_s, x_t) > P_t(\{j\} \mid x_s, x_t) \Longrightarrow \{\Delta_{s,t}\delta_j > 0\} \vee \{\exists k \neq j \in \mathcal{L},\ \Delta_{s,t}\delta_k < 0\};$$

- 针对 $j_1 \neq j_2 \in \mathcal{L}$ 的商品组合 $j_1 j_2$ 的条件选择概率比较：

$$P_s(\{j_1 j_2\} \mid x_s, x_t) > P_t(\{j_1 j_2\} \mid x_s, x_t) \Longrightarrow \\ \{\exists j \in \{j_1, j_2\},\ \Delta_{s,t}\delta_j > 0\} \vee \{\exists k \notin \{j_1, j_2\},\ \Delta_{s,t}\delta_k < 0\};$$

- 针对 $\ell \in \{A, B\}$ 的 $D_\ell = \{\ell, AB\}$ 的条件选择概率比较：

$$P_s(D_\ell \mid x_s, x_t) > P_t(D_\ell \mid x_s, x_t) \Longrightarrow \{\Delta_{s,t}\delta_\ell > 0\} \vee \{\Delta_{s,t}(\delta_\ell + \text{sign}(\Gamma_{AB})\delta_{\ell_{-1}}) > 0\} \\ \vee \{\exists k \in \{C, .., J\},\ \Delta_{s,t}(\delta_{\ell_{-1}} - \delta_k) > 0\} \vee \{\exists k \in \{C, .., J\},\ \Delta_{s,t}\delta_k < 0\},$$

其中 $\ell_{-1} \in \{A, B\}$ 和 $\ell_{-1} \neq \ell$。

命题5 将两种商品的结果推广至任意有限数量的商品。它仍然利用条件选择概率的跨期变异，推导出关于效用系数 β_0 和增量效用 Γ_{AB} 的识别限制。由于消费者的可行选择数量更多，且需要考虑所有商品的协变量指标变异，识别条件也更加复杂。

第一部分利用购买单一商品的变化来识别效用系数 β_0。当我们观察到购买某种单一

商品的概率上升时，例如商品 A，可以推断出该商品的协变量指数 A上升，或者某些其他商品的协变量指数下降。同样，第二部分考察随时间变化的购买组合行为。如果选择某个组合（例如 AB）的概率上升，则表明该组合中某种商品（ A 或 B）的协变量指数上升，或者组合之外某种商品的协变量指数下降。

命题的最后一部分5 利用选择集合 $D_A = \{A, AB\}$ （或 D_B）的条件概率变化来推断增量效用 Γ_{AB}的符号。为说明这一思路，我先假设所有其他商品的协变量$\{C, ..., J\}$ 保持不变。如果商品 A 和 B 都有所改善，但购买概率$\{A, AB\}$ 下降，这表明一些个体转而只购买商品 B ，而不购买商品组合 AB ，且增量效用为负 $\Gamma_{AB} < 0$。如果其他商品的协变量 $\{C, ..., J\}$ 也可能变化，那么选择 $\{A, AB\}$ 的概率上升，可能是因为商品 $\{C, ..., J\}$的协变量指标下降，导致个体转而选择 $\{A, AB\}$。这解释了命题 $\Delta_{s5t}\delta_k < 0$ 最后一部分中，针对某些 $k \in \{C, ..., J\}$ 所提出的另一项限制 。

B.2 扩展：横截面模型

在本文中，我研究了面板数据下的多项选择模型。本节简要介绍如何将识别策略应用于横截面多项选择模型。

现在考虑截面数据下的模型。消费者 i选择商品 $j \in \{A, B\}$ 的效用 u_{ij} 取决于协变量 X_{ij} 和误差项 ϵ_{ij}，其形式如下：

$$u_{ij} = X'_{ij}\beta_0 + \epsilon_{ij}.$$

外部选项的效用归一化为零： $u_{iO} = 0$。选择商品组合 AB 所带来的增量效用 Γ_{iAB} 表示为

$$\Gamma_{iAB} = u_{iAB} - u_{iA} - u_{iB}.$$

我采用与假设 1相同的假设，即增量效用仅取决于协变量 Z_i：

$$\Gamma_{iAB} = \Gamma(Z_i).$$

协变量 Z_i 被省略，因此我们用 Γ_0 表示 $\Gamma(Z_i)$。对于横截面数据，假设 3 中的平稳性条件需要作相应调整。令$\epsilon_i = (\epsilon_{iA},\ \epsilon_{iB})$ 和 $X_i = (X_{iA},\ X_{iB})$。我假设误差项 ϵ_i 独立于X_i。这一假设对误差项 ϵ_i与协变量 X_i 之间施加了独立性条件，但仍允许误差项在不同选择之间自由相关。

在这一假设下，可以利用协变量的变异推导出与命题 2类似的识别结果。给定协变量$X_i = x$ 和 $X_i = \tilde{x}$的两个不同取值，令 $\delta_j = x_j'\beta_0$ 和 $\delta_j = \tilde{x}_j'\beta_0$。对于任意 $(x, \tilde{x})$，以下条件均成立：

1. 比较选择 $j \in \mathcal{C}$ 的条件选择概率：

$$P(\{j\} \mid x) > P_t(\{j\} \mid \tilde{x}) \Longrightarrow \exists k \neq j \text{ s.t. } \delta_j - \tilde{\delta}_j > \delta_k - \tilde{\delta}_k;$$

2. 比较商品 $\ell \in \{A, B\}$ 的需求：

$$P\big(\{\ell, AB\} \mid x\big) > P\big(\{\ell, AB\} \mid \tilde{x}\big) \Longrightarrow$$
$$\{\delta_\ell - \tilde{\delta}_\ell > 0\} \vee \Big\{(\delta_\ell - \tilde{\delta}_\ell + \text{sign}(\Gamma_0)(\delta_{\ell_{-1}} - \tilde{\delta}_{\ell_{-1}}) > 0, |\Gamma_0| > -(\delta_\ell - \tilde{\delta}_\ell)\Big\},$$

其中，$\ell_{-1} \in \{A, B\}$ 和 $\ell_{-1} \neq \ell$；

3. 比较两个选择的条件选择概率之和：

$$P(\{AB\} \mid x) + P_t(\{O\} \mid \tilde{x}) > 1 \Longrightarrow$$
$$\Big\{\Gamma_0 > -\min\{\delta_A - \tilde{\delta}_A, \delta_B - \tilde{\delta}_B\}\Big\} \wedge \{\delta_A - \tilde{\delta}_A + \delta_B - \tilde{\delta}_B > 0\};$$
$$P(\{A\} \mid x) + P(\{B\} \mid \tilde{x}) > 1 \Longrightarrow$$
$$\Big\{\Gamma_0 < \min\{\delta_A - \tilde{\delta}_A, -(\delta_B - \tilde{\delta}_B)\}\Big\} \wedge \{\delta_A - \tilde{\delta}_A - (\delta_B - \tilde{\delta}_B) > 0\}.$$

B.3 扩展：非可分效用函数

本文关注文献中离散选择模型常用的加性可分效用函数。本节研究一类更一般的效用函数，其中可观测协变量与不可观测异质性之间可以是不可分离的。这类效用函数允许可观测协变量、固定效应和误差项之间存在灵活的交互作用。

设有两种商品 $\{A, B\}$。效用 u_{ijt} 对于消费者 i 在时间 $j \in \{A, B\}$ 选择商品 t 时，仍取决于三个关键组成部分：协变量 X_{ijt}、未观测的个体层面固定效应 α_{ij}和未观测误差项ϵ_{ijt}。与模型（1）不同，本节研究如下更一般的效用函数：对于 $j \in \{A, B\}$,

$$u_{ijt} = u(X'_{ijt}\beta_0, \alpha_{ij}, \epsilon_{ijt}). \tag{11}$$

函数 u 对研究人员而言可能是未知的，并且在这三个组成部分之间可能是不可分离的。对函数 u的唯一限制是假设9中所述的单调性条件。这类效用函数可以体现消费者更丰富的偏好。

外部选项 O 的效用归一化为零：

$$u_{iOt} = 0.$$

我们仍然关注消费商品组合 AB相较于选择单一商品所带来的增量效用 Γ_{it} ：

$$\Gamma_{it} = u_{iABt} - u_{iAt} - u_{iBt}.$$

我对互补性作出与假设1相同的假设，即互补性只能取决于可观测协变量 Z_i 。协变量 Z_i 被省略，因此互补性 Γ_0 被写作常数。

(11)中的效用函数也由 Gao and Li (2020研究过，不过他们的论文侧重于系数 β_0的识别。我的论文允许购买商品组合 AB ，并给出对两种商品之间互补性$\Gamma(z$ 的识别结果。遵循 Gao andLi (2020)，我假设效用函数关于协变量指数 $X'_{ijt}\beta_0$满足单调性条件。

假设9。（单调性）对于每一种实现值（δ, α, ϵ），效用 u 随指数δ 弱单调递增，即（α, ϵ）。

$$for\ any\ \ (\alpha, \epsilon), \quad u(\tilde{\delta}, \alpha, \epsilon) \geq u(\delta, \alpha, \epsilon) \quad if\ \tilde{\delta} \geq \delta.$$

假设9 仅要求效用关于协变量指数单调，但不对不可观测固定效应和误差项施加任何限制。加性可分离

方程（1）中的耕地效用函数属于假设9的特例。非可分效用函数不仅允许可观测特征与不可观测异质性之间存在灵活的交互作用，还允许协变量采用非线性函数X_{ijt}，例如指数函数或高阶多项式函数。

回顾一下，对于 $\ell \in \{A, B\}$，$D_\ell = \{\ell, AB\}$。下一个命题将刻画参数 $\theta_0 = (\beta_0, \Gamma_0)$ 的识别约束。

命题6。 在假设 *1-3* & *9*下，对于任意 (x_s, x_t)和 $s \neq t \leq T$，

(1) 对选择 $j \in \mathcal{C}$ 的条件选择概率进行比较，

$$P_s(\{j\} \mid x_s, x_t) > P_t(\{j\} \mid x_s, x_t) \Longrightarrow \\ \left\{(-1)^{\mathbb{1}[j \in D_A]} \Delta_{s,t} \delta_A < 0\right\} \vee \left\{(-1)^{\mathbb{1}[j \in D_B]} \Delta_{s,t} \delta_B < 0\right\};$$

(2) 对商品 $\ell \in \{A, B\}$ 的需求进行比较：

$$P_s(D_\ell \mid x_s, x_t) > P_t(D_\ell \mid x_s, x_t) \Longrightarrow \\ \{\Delta_{s,t} \delta_\ell > 0\} \vee \left\{ \operatorname{sign}(\Gamma_0) \Delta_{s,t} \delta_{\ell_{-1}} > 0\right\},$$

其中 $\ell_{-1} \in \{A, B\}$ 和 $\ell_{-1} \neq \ell$。

与命题2类似，命题6 利用任意两个时期条件选择概率的跨期变异，推导参数 θ_0 的识别条件。主要区别在于，式（1）中的加性可分效用函数对协变量指标如何影响个体效用 u_{ijt}的方向和程度都施加了限制，而式（11）中的非可分离函数仅假设满足单调性条件，但对协变量指标影响效用的程度不作限制。因此，命题 6 中的识别集比命题 2中的结果更宽泛，而命题 6 中的结果对效用函数的设定具有稳健性。命题 6 的证明与附录A.2类似，故此处省略。

在命题6中，第一个条件可以通过单个选择的条件概率识别系数 β_0。直觉如下：如果选择 j 的条件概率上升，那么必然是商品 j

变得更好（协变量指数上升），或者其他商品变得更差（协变量指数下降）。

命题6 中的第二个条件利用条件需求随时间的变化来确定 Γ_0 的符号。与命题（ID2）中的条件2类似，其思路是利用不同互补关系下协变量指数与需求之间的关系。当两种商品互为互补品时，商品 A 的协变量指数上升会推高商品 B 的需求，因为人们会转而选择商品组合 AB。因此，如果观察到商品 B 的需求下降，就可以推断这两种商品是替代品 $\Gamma_0 \leq 0$。由于在这个非可分效用函数中，协变量指数对效用的影响未知，我们无法为互补性的大小设定界限。

B.4 效用设定的讨论

本节讨论本文的效用设定与传统离散选择模型之间的联系。具体而言，考虑 Pakes 和 Porter （2022年）模型的三商品版本：

$$\begin{aligned} u_{iAt} &= X'_{iAt}\beta_0 + \alpha_{iA} + \epsilon_{iA}, \\ u_{iBt} &= X'_{iBt}\beta_0 + \alpha_{iB} + \epsilon_{iB}, \\ u_{iCt} &= X'_{iCt}\beta_0 + \alpha_{iC} + \epsilon_{iCt}, \\ u_{iOt} &= 0. \end{aligned} \tag{12}$$

可以将 $C = AB$ 重新命名，并将商品组合 AB 解释为一种额外的商品。

在我的模型中，我则通过将 AB 的效用与单个商品的效用联系起来，直接对其进行规定，具体如下：

$$\begin{aligned} u_{iAt} &= X'_{iAt}\beta_0 + \alpha_{iA} + \epsilon_{iA}, \\ u_{iBt} &= X'_{iBt}\beta_0 + \alpha_{iB} + \epsilon_{iB}, \\ u_{iABt} &= u_{iAt} + u_{iBt} + \Gamma(Z_i), \\ u_{iOt} &= 0. \end{aligned} \tag{13}$$

上述设定假定，与仅选择单个商品的效用之和相比，同时选择两种商品所带来的增量效用仅取决于可观测特征 Z_i。这一设定可视为施加了以下条件

于（12）：

$$X'_{iABt}\beta_0 + \alpha_{iAB} + \epsilon_{iABt} = (X_{iAt} + X_{iBt})'\beta_0 + \alpha_{iA} + \alpha_{iB} + \epsilon_{iAt} + \epsilon_{iBt} + \Gamma(Z_i). \quad (14)$$

请注意，假定同时消费两种商品所产生的互补性 $\Gamma(Z_i)$ 仅取决于可观测特征，但其来源可以是协变量指数、固定效应和未观测冲击的任意组合。[4] 例如，假设 ϵ_{ijt} 表示产品的未观测质量。商品组合的效用中由质量带来的部分 ϵ_{iABt}，可能不等于 $\epsilon_{iAt} + \epsilon_{iBt}$，即质量对 u_A 和 u_B 的效用贡献之和。这反映了由未观测质量产生的替代／互补性。然而，假定差值 $\epsilon_{iABt} - \epsilon_{iAt} - \epsilon_{iBt}$ 仅取决于可观测变量。这一结构使我们能够检验并识别商品之间的替代与互补关系。

条件（14）假设互补性仅取决于可观测特征。一个更一般的设定是在互补性中纳入不可观测异质性，具体如下：

$$X'_{iABt}\beta_0 + \alpha_{iAB} + \epsilon_{iABt} = (X_{iAt} + X_{iBt})'\beta_0 + \alpha_{iA} + \alpha_{iB} + \epsilon_{iAt} + \epsilon_{iBt} + \Gamma_{it},$$

其中 Γ_{it} 是一个不可观测的个体特定误差项。因此，该设定允许个体之间存在未被可观测特征捕捉的异质性互补关系。第 7 节探讨了这一设定，但只能得到较弱的识别结果。

B.5 替代性检验

在某些应用中，我们可能认为两种商品是替代品，并希望检验它们之间的替代关系。检验替代性的思路与第 3 节中的分析类似。检验替代性等价于检验 $\Gamma_0 \leq 0$（协变量 Z_i 被省略）。那么，原假设 H'_0 和备择假设 H'_1 如下：

$$H'_0 : \Gamma_0 \leq 0 \qquad H'_1 : \Gamma_0 > 0.$$

[4] 在此模型中，我们只能识别源自全部三个组成部分的总体互补性，因为任意将互补项划分为这三个部分都不会影响消费者的选择。

主要思路仍是在原假设 H_0' 下构造仅依赖于观测变量的矩不等式。为构造矩约束，我们将利用两种商品存在替代关系时，它们的需求与协变量指标之间的关系。直观解释如下：如果两种商品是替代品（在 H_0' 下），商品 A的改善和商品 B 的下降应导致商品 A 的需求增加、商品 B的需求减少。如果观察到商品 A 的需求反而下降，或商品 B 的需求反而增加，则可以推断这两种商品是互补品，且原假设 H_0' 不成立。

为刻画可检验推论，第一步是推断两个商品的协变量指数的符号，使得商品 A 变得更好，而商品 B 变得更差。令 $\xi^2_{s,t}(x_s, x_t)$ 为所有选择 $\{A, AB, O\}$ 的概率均增加的指示变量，其定义如下：

$$\xi^2_{s,t}(x_s, x_t) = \mathbb{1}\big\{P_s(\{j\} \mid x_s, x_t) - P_t(\{j\} \mid x_s, x_t) \geq 0,\ \forall j \in \{A, AB, O\}\big\}.$$

当观察到所有选择 $j \in \{A, AB, O\}$ 的条件概率均增加时，我们可以推断：

$$\xi^2_{s,t}(x_s, x_t) = 1 \Longrightarrow \Delta_{s,t}\delta_A \geq 0,\ \Delta_{s,t}\delta_B \leq 0.$$

由于除商品 B 外，所有选择的概率均有所增加，因此只能推断商品 A 的协变量指数上升，而商品 B 的协变量指数下降。根据上述两个商品协变量指数的符号，在原假设“两种商品互为替代品”成立的条件下，商品 A 的条件需求应当增加，而商品 B 的需求应当减少。当观察到商品 A 的需求增加或商品 B 的需求下降时，我们可以得出结论：这两种商品是互补品，并拒绝原假设。以下命题刻画了原假设 H_0' 的可检验推论。

命题7。 在假设 *1-3*下，原假设 H_0'下成立以下条件矩不等式：

$$\begin{cases} E\Big[\xi^2_{s,t}(x_s, x_t)\big(\mathbb{1}\{Y_{is} \in D_A\} - \mathbb{1}\{Y_{it} \in D_A\}\big) \,\Big|\, x_s, x_t\Big] \geq 0; \\ E\Big[\xi^2_{s,t}(x_s, x_t)\big(\mathbb{1}\{Y_{is} \in D_B\} - \mathbb{1}\{Y_{it} \in D_B\}\big) \,\Big|\, x_s, x_t\Big] \leq 0, \end{cases}$$

对于任意(x_s, x_t)以及任意 $s \neq t \leq T$。

B.6 更多模拟结果

B.6.1 较长面板 $T \geq 2$

本节研究点估计量的表现，但考虑较长的面板$T = \{2, 3, 4\}$。我将两步估计量与参数估计量以及论文中描述的另外两种 β_0 估计量进行比较。协变量采用相同的模拟设定，固定效应和误差项采用相同的四种数据生成过程设计。样本量为 $N = 1000$ ，模拟重复次数为 $B = 500$。

表 3和表 4比较了较长面板下两步估计量对互补性参数 γ_0 和效用系数 β_0 的估计表现。随着面板长度增加，两步估计量的表现有所改善，且在所有设计中都稳健。较长面板下，参数估计量表现更好，但只要存在任何模型设定错误，它仍可能产生较大偏差（设计2至4）。无论时期长度如何，两步估计量在设计2至4中的表现均优于其他估计量，并继续保持更稳健的优势。

表 3：$\hat{\gamma}$（较长面板）的性能比较

T	设计	两步估计				参数估计			
		Err	SD	均方根误差	平均绝对偏差	Err	SD	均方根误差	平均绝对偏差
$T=2$	设计 1	0.025	0.263	0.315	0.267	0.007	0.101	0.101	0.079
	设计 2	0.026	0.309	0.341	0.278	0.058	0.344	0.741	0.663
	设计 3	0.023	0.287	0.306	0.251	0.083	0.288	1.019	0.977
	设计 4	0.025	0.329	0.332	0.264	0.126	0.681	1.764	1.628
$T=3$	设计 1	0.018	0.204	0.233	0.187	0.006	0.079	0.079	0.063
	设计 2	0.022	0.291	0.293	0.238	0.063	0.311	0.782	0.721
	设计 3	0.018	0.233	0.233	0.187	0.056	0.232	0.674	0.634
	设计 4	0.021	0.284	0.286	0.229	0.115	0.527	1.533	1.440
$T=4$	设计 1	0.018	0.181	0.231	0.193	0.006	0.075	0.075	0.059
	设计 2	0.021	0.287	0.290	0.230	0.066	0.282	0.796	0.746
	设计 3	0.015	0.194	0.200	0.162	0.041	0.179	0.487	0.453
	设计 4	0.017	0.229	0.229	0.181	0.108	0.385	1.379	1.324

$\hat{}$ 表4：β 的性能比较（较长面板）

		带捆绑的估计量				假设不存在捆绑的估计量			
T	设计	两步估计		参数估计		固定效应 Logit 估计		半参数估计	
		均方根误差	平均绝对偏差	均方根误差	平均绝对偏差	均方根误差	平均绝对偏差	均方根误差	MAD
$T=2$	设计 1	0.126	0.101	0.072	0.056	0.165	0.130	0.155	0.131
	设计 2	0.125	0.101	0.432	0.389	0.243	0.187	0.161	0.134
	设计 3	0.128	0.100	0.304	0.279	0.168	0.131	0.149	0.124
	设计 4	0.121	0.095	0.647	0.593	0.256	0.200	0.171	0.143
$T=3$	设计 1	0.087	0.069	0.052	0.042	0.107	0.084	0.130	0.108
	设计 2	0.083	0.064	0.385	0.360	0.178	0.149	0.123	0.103
	设计 3	0.090	0.070	0.214	0.198	0.106	0.084	0.124	0.102
	设计 4	0.085	0.069	0.539	0.509	0.189	0.157	0.131	0.109
$T=4$	设计 1	0.081	0.066	0.045	0.036	0.079	0.062	0.117	0.095
	设计 2	0.069	0.050	0.081	0.058	0.145	0.125	0.095	0.060
	设计 3	0.075	0.060	0.165	0.153	0.080	0.064	0.116	0.096
	设计 4	0.074	0.059	0.488	0.465	0.163	0.138	0.119	0.097

B.6.2 部分识别

本节考察协变量具有有界支持集时集合估计量的表现。同时也实施参数方法作为比较；无论协变量的支撑集如何，该方法都能实现点识别。

模拟设定与之前类似，但部分协变量的支撑集有界。令 $d_x = 2$ 和 $d_z = 2$ 分别表示协变量 $X_{i\ell t}$ 和 Z_i 的维度，其中 $\ell \in \{A, B\}$。在每次模拟中，X_{iAt} 服从均匀分布$U[-3, 3]$，而 X_{iBt}服从正态分布 $\mathcal{N}(0,\ 2)$。不同选择 $\ell \in \{A, B\}$ 和时期 $t \leq T$之间的协变量相互独立。令Z_i的第一个元素服从均匀分布 $U[0, 4]$，第二个元素服从 $U[-2, 2]$。真实参数设为 $\beta_0 = \gamma_0 = (1,\ 1)$。我们研究与之前相同的四种固定效应和误差项设定。设样本量为 $N \in \{1000, 4000\}$，时期长度为 $T = 2$，重复次数为$B = 500$。

由于 X_{iAt} 和 Z_i 的支撑集有界，因此参数只能被部分识别。根据 Chernozhukov、Hong 和 Tamer (2007)，识别集 Θ_I 的集合估计量定义为

$$\hat{\Theta}_{\hat{c}_N} = \left\{\theta \in \Theta : \hat{\Omega}_N(\theta) \leq \inf_{\theta \in \Theta} \hat{\Omega}_N(\theta) + \hat{c}_N / a_N \right\},$$

ˆ 其中，a_N 表示样本目标函数 Ω_N的一致收敛速率；采用神经网络估计量作为第一步估计时，该速率为 $a_N \approx N^{1/4}$；ˆc_N取为 $10^{-4}\log(N)$，以满足 ˆ$c_N/a_N \to 0$。我将区间 $[-5, 5]$划分为$k = 100$ 个网格，并通过网格搜索求得识别集。

ˆ ˆ设 (β_l, β_u) 和 $(\hat{}\gamma_l, \hat{}\gamma_u)$ 分别表示 β_0 和 γ_0的下界与上界。ˆ设 $\beta_{par}, \hat{}\gamma_{par}$ 表示 β_0 和 γ_0的参数估计量。为比较这两种方法，我仍报告标准差（SD）、均方根误差（rMSE）和平均绝对偏差（MAD）。

表 5 和 6 比较了本文下界/上界与参数方法的表现。两种方法的比较呈现出与点识别部分相似的模式。尽管在大支撑集情形下，集合估计量的表现略逊于点估计量，但相较于参数估计量，其在固定效应或误差项分布假设的设定方面仍表现得更为稳健。参数方法

在正确设定（设计 1）下具有较小的偏差，但在模型设定错误（设计2至4）时可能产生严重偏差。本文提出的集合估计量在不同的固定效应和误差项设定下均表现稳健。

表5：$\hat{}\gamma$的性能比较

N	设计	两步估计						参数估计		
		$\hat{\gamma}_l$			$\hat{\gamma}_u$			$\hat{\gamma}_{par}$		
		SD	均方根误差	平均绝对偏差	SD	均方根误差	平均绝对偏差	SD	均方根误差	平均绝对偏差
1000	设计 1	0.309	0.353	0.285	0.318	0.483	0.410	0.083	0.083	0.067
	设计 2	0.365	0.367	0.292	0.397	0.492	0.394	0.389	0.946	0.864
	设计 3	0.376	0.510	0.418	0.406	0.739	0.635	0.241	1.061	1.032
	设计 4	0.408	0.538	0.419	0.456	0.766	0.634	0.690	2.670	2.580
4000	设计 1	0.205	0.208	0.170	0.211	0.334	0.283	0.043	0.043	0.034
	设计 2	0.221	0.239	0.189	0.231	0.459	0.403	0.197	0.745	0.718
	设计 3	0.212	0.242	0.195	0.233	0.538	0.488	0.149	1.047	1.036
	设计 4	0.204	0.290	0.238	0.232	0.588	0.543	0.409	2.767	2.734

$\hat{}$ 表 6：β 的性能比较

N	设计	两步估计						参数估计		
		$\hat{\beta}_l$			$\hat{\beta}_u$			$\hat{\beta}_{par}$		
		SD	均方根误差	平均绝对偏差	SD	均方根误差	平均绝对偏差	SD	均方根误差	平均绝对偏差
1000	设计 1	0.121	0.129	0.107	0.132	0.132	0.107	0.068	0.069	0.054
	设计 2	0.125	0.136	0.113	0.131	0.131	0.106	0.160	0.186	0.148
	设计 3	0.125	0.129	0.104	0.133	0.138	0.110	0.138	0.460	0.438
	设计 4	0.127	0.131	0.109	0.134	0.136	0.109	0.253	0.627	0.573
4000	设计 1	0.071	0.090	0.077	0.079	0.081	0.069	0.032	0.032	0.026
	设计 2	0.074	0.094	0.079	0.077	0.077	0.066	0.085	0.092	0.070
	设计 3	0.069	0.079	0.067	0.076	0.086	0.073	0.069	0.424	0.419
	设计 4	0.072	0.084	0.071	0.079	0.086	0.072	0.132	0.596	0.581

B.7 命题4的证明

证明。一种商品的条件需求可以表示为两类人群的混合：一类人群认为这两种商品是互补品 $\Gamma_{it} \geq 0$ ，另一类人群认为这两种商品是替代品 $\Gamma_{it} < 0$。因此，在给定协变量和固定效应的条件下，对商品 A （或 B）的需求如下所示：

$$\begin{aligned}\Pr(Y_{it} \in D_A \mid \alpha_i, x_s, x_t) = \Pr(Y_{it} \in D_A \mid \alpha_i, x_s, x_t, \Gamma_{it} \geq 0) \Pr(\Gamma_{it} \geq 0 \mid \alpha_i) + \\ \Pr(Y_{it} \in D_A \mid \alpha_i, x_s, x_t, \Gamma_{it} < 0)[1 - \Pr(\Gamma_{it} \geq 0 \mid \alpha_i)].\end{aligned}$$

上述等式成立，是因为假设8 （条件独立性）意味着 $\Pr(\Gamma_{it} \geq 0 \mid \alpha_i, x_s, x_t) = \Pr(\Gamma_{it} \geq 0 \mid \alpha_i)$。

主要策略是利用商品 A条件需求的跨期变异，来界定概率 $\Pr(\Gamma_{it} \geq 0 \mid \alpha_i)$。可以证明，当 $(x_s, x_t) \in \mathcal{X}^1_{s,t}$ 时，条件需求 $\Pr(Y_{it} \in D_A \mid \alpha_i, x_s, x_t, \Gamma_{it} \geq 0)$ 在时间 s 高于时间 t，这一点将在后文证明。给定 $\Gamma_{it} < 0$ 时条件需求的变动可以界定为 $[-1, 1]$ ，因为它是两个概率之差。因此，商品 A总需求的变动可以界定为：对于任意 $(x_s, x_t) \in \mathcal{X}^1_{s,t}$，

$$\Pr(Y_{is} \in D_A \mid \alpha_i, x_s, x_t) - \Pr(Y_{it} \in D_A \mid \alpha_i, x_s, x_t) \geq 0 + (-1) * [1 - \Pr(\Gamma_{it} \geq 0 \mid \alpha_i)].$$

在给定协变量的条件下，对固定效应 α_i 取期望，概率 $\eta = \Pr(\Gamma_{it} \geq 0)$可按如下方式给出上界：对于任意 $(x_s, x_t) \in \mathcal{X}^1_{s,t}$，

$$\eta \leq P_s(D_A \mid x_s, x_t) - P_t(D_A \mid x_s, x_t) + 1.$$

由于概率 η 不取决于协变量，并且根据假设 7，其随时间保持平稳，因此可以通过对协变量的所有取值以及任意两个时期取下确界来对其进行界定。此外，还可以利用商品B 需求的变化，以类似方式对概率 η 进行界定。因此， η 的上界可建立如下：

$$\eta \leq \inf_{(x_s,x_t) \in \mathcal{X}^1_{s,t}, \ell \in \{A,B\}, (s,t) \leq T} \Big\{ P_s(D_\ell \mid x_s, x_t) - P_t(D_\ell \mid x_s, x_t) \Big\} + 1 = U_\eta.$$

现在需要证明，当协变量满足条件（x_s， x_t）$\in \mathcal{X}^1_{s,\ t}$时，条件概率 $\Pr(Y_{it} \in D_A \mid \alpha_i,\ x_s,\ x_t,\ \Gamma_{it} \geq 0)$ 在时期 s 比时期 t 增大。令 $v_{it} = \epsilon_{it} + \alpha_i$，并令 $\mathcal{V}^{\Gamma}_{D_A}(x_t)$ 表示满足以下条件的（v， $\Gamma \geq 0$）的集合：选择 A 或 AB：

$$\begin{aligned}\mathcal{V}^{\Gamma}_{D_A}(x_t) =&\{(v, \Gamma \geq 0) \mid \delta_{At} + v_A \geq \delta_{Bt} + v_B,\ \delta_{At} + v_A \geq 0\} \equiv \mathcal{V}^{\Gamma}_1(x_t),\\ &\cup\{(v, \Gamma \geq 0) \mid \delta_{At} + v_A + \Gamma \geq 0,\ \delta_{At} + v_A + \delta_{Bt} + v_B + \Gamma \geq 0\} \equiv \mathcal{V}^{\Gamma}_2(x_t).\end{aligned}$$

条件需求 $\Pr(Y_{it} \in D_A \mid \alpha_i,\ x_s,\ x_t,\ \Gamma_{it} \geq 0)$ 可以表示为集合 $\mathcal{V}^{\Gamma}_{D_A}(x_t)$ 的条件概率：

$$\Pr(Y_{it} \in D_A \mid \alpha_i, x_s, x_t, \Gamma_{it} \geq 0) = \Pr((v_{it}, \Gamma_{it}) \in \mathcal{V}^{\Gamma}_{D_A}(x_t) \mid \alpha_i, x_s, x_t, \Gamma_{it} \geq 0).$$

假设7 意味着，给定 $(\alpha_i,\ X_{is},\ X_{it},\ \Gamma_{it})$ 时，分布 $(v_{it},\ \Gamma_{it})$ 随时间保持平稳。接下来只需证明，当协变量满足 $(x_s,\ x_t) \in \mathcal{X}^1_{s,\ t}$ 时，$\mathcal{V}^{\Gamma}_{D_A}(x_t) \subseteq \mathcal{V}^{\Gamma}_{D_A}(x_s)$，其推论如下：

$$\begin{aligned}&\mathcal{V}^{\Gamma}_{D_A}(x_t) \subseteq \mathcal{V}^{\Gamma}_{D_A}(x_s) \Longrightarrow\\ &\qquad \Pr(Y_{it} \in D_A \mid \alpha_i, x_s, x_t, \Gamma_{it} \geq 0) \leq \Pr(Y_{is} \in D_A \mid \alpha_i, x_s, x_t, \Gamma_{it} \geq 0).\end{aligned}$$

为证明 $\mathcal{V}^{\Gamma}_{D_A}(x_t) \subseteq \mathcal{V}^{\Gamma}_{D_A}(x_s)$，我将证明，对于任意元素 $(v, \Gamma) \in \mathcal{V}^{\Gamma}_{D_A}(x_t)$，当 $(x_s, x_t) \in \mathcal{X}^1_{s,t}$ 时，它也满足 $(v, \Gamma) \in \mathcal{V}^{\Gamma}_{D_A}(x_s)$。如前所述，任意 $(x_s, x_t) \in \mathcal{X}^1_{s,t}$ 都满足 $\delta_{At} \geq 0, \delta_{Bt} \geq 0$。为证明这一结论，我将讨论两种情况：$(v, \Gamma) \in \mathcal{V}^{\Gamma}_1(x_t)$ 和 $(v, \Gamma) \in \mathcal{V}^{\Gamma}_2(x_t)$。

情形1：$(v,\ \Gamma) \in \mathcal{V}^{\Gamma}_2(x_t)$。根据集合 $\mathcal{V}^{\Gamma}_2(x_t)$ 的定义，它随着商品 A 和商品 B 的协变量指数增加。因此，我们知道 $(v,\ \Gamma) \in \mathcal{V}^{\Gamma}_2(x_s)$，因为当 $(x_s,\ x_t) \in \mathcal{X}^1_{s,\ t}$ 时，商品 A 和 B 的协变量指数都会增加。

情形2：$(v,\ \Gamma) \in \mathcal{V}^{\Gamma}_1(x_t)$。如果 v 满足 $\delta_{As} + v_A \geq \delta_{Bs} + v_B$，则意味着 $(v,\ \Gamma) \in \mathcal{V}^{\Gamma}_1(x_s)$，因为商品 A 的协变量指数在时间 s 相对于时间 t 有所增加。否则，v 应满足 $\delta_{As} + v_A < \delta_{Bs} + v_B$。此外，互补性非负 $\Gamma \geq 0$，这意味着以下条件：

$$\delta_{As} + v_A + \Gamma \geq \delta_{At} + v_A \geq 0,\ \delta_{As} + v_A + \Gamma + \delta_{Bs} + v_B \geq 2(\delta_{As} + v_A) \geq 0.$$

上述条件意味着$(v,\ \Gamma) \in \mathcal{V}_2^{\Gamma}(x_s) \subseteq \mathcal{V}_{D_A}^{\Gamma}(x_s)$。

总之，我已证明，对于任意$(x_s,\ x_t) \in \mathcal{X}_{s,\ t}^1$，$\mathcal{V}_{D_A}^{\Gamma}(x_t) \subseteq \mathcal{V}_{D_A}^{\Gamma}(x_s)$，这意味着条件概率 $\Pr(Y_{it} \in D_A \mid \alpha_i,\ x_s,\ x_t,\ \Gamma_{it} \geq 0)$ 在时间 s 相较于时间 t有所上升。η 的下界证明可以类似地得出，因此此处省略。

□

B.7.1 命题5的证明

证明。该证明与命题2的证明类似。我们首先推导条件选择概率下降的充分条件。然后，当观察到条件选择概率上升时，可以通过逆否命题推导参数的识别限制。

令 $v_{ijt} = \alpha_{ij} + \epsilon_{ijt}$ 表示 $j \in \mathcal{L}$的固定效应与误差项之和。令 $v_{j_1 j_2 t} = v_{j_1 t} + v_{j2t} + \Gamma_{j_1 j_2}$ 和 $v_O = 0$ 分别表示商品组合 $j_1 j_2$ 和外部选项对于 $j_1 \neq j_2 \in \mathcal{L}$的固定效应与误差项之和。给定协变量$X_{ijt} = x_{jt}$，令 $\delta_{jt} = x'_{jt}\beta_0$、$\delta_{j_1 j_2 t} = \delta_{j_1 t} + \delta_{j_2 t}$和 $\delta_{Ot} = 0$。

对于任意集合 $K \subset \mathcal{C}$，令 $\mathcal{V}_K(x_t)$ 表示满足以下条件的 $v = (v_A,\ \ldots\ldots,\ v_J)$ 的集合：给定 $X_{it} = x_t$ 时，$K \subset \mathcal{C}$ 中至少有一个选择被选中。集合 $\mathcal{V}_K(x_t)$ 可表示为

$$\mathcal{V}_K(x_t) = \left\{ v \mid \exists j \in K \text{ s.t. } \delta_{jt} + v_j \geq \delta_{kt} + v_k \text{ for all } k \in K^c \right\},$$

集合 K 中至少有一个选择被选中的条件概率可表示如下：

$$\Pr(Y_{it} \in K \mid \alpha_i, x_s, x_t) = \Pr\left(v_{it} \in \mathcal{V}_K(x_t) \mid \alpha_i, x_s, x_t\right).$$

在假设 3 （平稳性）下，v_{it} 的条件分布随时间保持平稳，对任意 $s \neq t$ 均如此，因为固定效应 α_i 不随时间变化。因此，集合越大，条件选择概率随时间越大：

$$\mathcal{V}_K(x_s) \subseteq \mathcal{V}_K(x_t) \implies \Pr(Y_{is} \in K \mid \alpha_i, x_s, x_t) \leq \Pr(Y_{it} \in K \mid \alpha_i, x_s, x_t).$$

我将给出上述集合包含关系成立的充分条件，该关系意味着条件选择概率下降。然后，在观察到条件选择概率上升时，可以推导出识别约束

。命题 5 包含三部分识别约束，各部分的证明如下。

第1部分：比较随时间变化的选择商品 $j \in \mathcal{L}$ 的条件概率。我以商品 A 为例。集合 $\mathcal{V}_A(x_t)$ 定义为

$$\mathcal{V}_A(x_t) = \big\{v \mid \delta_{At} + v_A \geq 0, \delta_{At} + v_A - \delta_{j_1 t} - v_{j_1} \geq 0, \delta_{j_1 t} + v_{j_1} + \Gamma_{Aj_1} \leq 0, \\ \delta_{At} + v_A - \delta_{j_1 j_2 t} - v_{j_1 j_2} \geq 0 \text{ for all } j_1 \neq j_2 \in \{B, ..., J\}\big\},$$

注意，根据定义，$\delta_{j_1 j_2 t} = \delta_{j_2 j_1 t}$ 和 $v_{j_1 j_2} = v_{j_2 j_1}$，因此交换商品组合的两个指数顺序不会造成任何差异。对于 $k \in \{B, .., J\}$，集合 $\mathcal{V}_A(x_t)$ 随 δ_{At} 增大，随 δ_{kt} 减小。因此，以下集合关系成立：

$$\Delta_{s,t}\delta_A \leq 0,\ \Delta_{s,t}\delta_k \geq 0\ \forall k \in \{B, .., J\} \implies \\ \Pr(Y_{is} = \{A\} \mid \alpha_i, x_s, x_t) \leq \Pr(Y_{it} = \{A\} \mid \alpha_i, x_s, x_t).$$

根据逆否命题，并对 α_i 在给定 (x_s、x_t) 条件下取期望，可得命题 5 中的第一个识别限制：

$$P_s(\{A\} \mid x_s, x_t) > P_t(\{A\} \mid x_s, x_t) \implies \{\Delta_{s,t}\delta_A > 0\} \vee \{\exists k \in \{B, .., J\},\ \Delta_{s,t}\delta_k < 0\}.$$

第2部分：比较 $j_1 \neq j_2 \in \mathcal{L}$ 时选择商品组合 $j_1 j_2$ 的条件概率随时间的变化。我以商品组合 AB 为例。集合 $\mathcal{V}_{AB}(x_t)$ 定义为

$$\mathcal{V}_{AB}(x_t) = \big\{v \mid \delta_{ABt} + v_{AB} \geq 0,\ \delta_{\ell t} + v_\ell + \Gamma_{AB} \geq 0, \delta_{ABt} + v_{AB} - \delta_{j_1 t} - v_{j_1} \geq 0, \\ \delta_{ABt} + v_{AB} - \delta_{j_1 j_2 t} - v_{j_1 j_2} \geq 0,\ \text{for all } \ell \in \{A, B\}, j_1 \in \{C, ..., J\}, j_2 \neq j_1 \in \mathcal{L}\big\}.$$

根据 $\mathcal{V}_{AB}(x_t)$ 的定义，当协变量指数 $\delta_{\ell t}$ 增加且 δ_k 减少时，对于 $\ell \in \{A, B\}$ 和 $k \in \{C, ..., J\}$，集合 $\mathcal{V}_{AB}(x_t)$ 会扩大。因此，以下限制成立：

$$\Delta_{s,t}\delta_\ell \leq 0, \Delta_{s,t}\delta_k \geq 0\ \ \forall \ell \in \{A, B\}, k \in \{C, ..., J\} \implies \\ \Pr(Y_{is} = \{AB\} \mid \alpha_i, x_s, x_t) \leq \Pr(Y_{it} = \{AB\} \mid \alpha_i, x_s, x_t).$$

通过逆否命题，并对 α_i 在给定 (x_s, x_t的条件下取期望，可得命题5中的第二个识别限制：

$$P_s(\{AB\} \mid x_s, x_t) > P_t(\{AB\} \mid x_s, x_t) \Longrightarrow \\ \{\exists \ell \in \{A, B\},\ \Delta_{s,t}\delta_\ell > 0\} \vee \{\exists k \in \{C, ..., J\},\ \Delta_{s,t}\delta_k < 0\}.$$

第 3 部分：随时间变化的选择 $D_\ell = \{\ell, AB\}$ 而非$\ell \in \{A, B\}$ 的条件概率比较。以集合 $D_A = \{A, AB\}$ 为例说明证明。我们可以将集合 $\mathcal{V}_{D_A}(x_t)$表示为两个集合的并集：选择 A 或选择 AB 的效用高于不属于 D_A的其他选择，

$$\mathcal{V}_{D_A}(x_t) = \mathcal{V}_1(x_t) \cup \mathcal{V}_2(x_t),$$

其中， $\mathcal{V}_1(x_t)$和 $\mathcal{V}_2(x_t)$定义如下

$$\mathcal{V}_1(x_t) = \big\{v \mid \delta_{At} + v_A \geq 0,\ \delta_{At} + v_A \geq \delta_{kt} + v_k,\ \delta_{j_1 t} + v_{j_1} + \Gamma_{Aj_1} \leq 0, \\ \delta_{At} + v_A \geq \delta_{j_1 t} + \delta_{j_2 t} + v_{j_1 j_2},\ \text{for all } k, j_2 \in \{B, ..., J\}, j_1 \neq j_2 \in \{C, ..., J\}\},$$

and

$$\mathcal{V}_2(x_t) = \big\{v \mid \delta_{ABt} + v_{AB} \geq 0,\ \delta_{ABt} + v_{AB} \geq \delta_{kt} + v_k,\ \delta_{Bt} + v_B + \Gamma_{AB} \geq \delta_{j_1 t} + v_{j_1} + \Gamma_{Aj_1}, \\ \delta_{ABt} + v_{AB} \geq \delta_{j_1 j_2 t} + v_{j_1 j_2} \text{ for all } k, j_2 \in \{B, ..., J\}, j_1 \neq j_2 \in \{C, ..., J\}\big\}.$$

我们考察命题5中最后一个条件的逆否命题：

$$\{\Delta_{s,t}\delta_A \leq 0\} \wedge \{\Delta_{s,t}(\delta_A + \text{sign}(\Gamma_{AB})\delta_B) \leq 0\} \wedge \{\Delta_{s,t}(\delta_B - \delta_k) \leq 0\} \\ \wedge \{\Delta_{s,t}\delta_k \geq 0\}\ \forall k \in \{C, .., J\} \Longrightarrow \mathcal{V}_{D_A}(x_s) \subseteq \mathcal{V}_{D_A}(x_t). \tag{15}$$

现在，只需证明上述条件即可证明命题5中的最后一个限制。参数条件取决于互补性 Γ_{AB}的符号。我们说明 $\Gamma_{AB} \geq 0$这一情形的证明，思路同样适用于 $\Gamma_{AB} < 0$。

当 $\Gamma_{AB} \geq 0$时，式（15）中的参数条件变为

$$\{\Delta_{s,t}\delta_A \leq 0\} \wedge \{\Delta_{s,t}(\delta_A + \delta_B) \leq 0\} \wedge \{\Delta_{s,t}(\delta_B - \delta_k) \leq 0\} \wedge \{\Delta_{s,t}\delta_k \geq 0\}\ \forall k \in \{C, .., J\}. \tag{16}$$

根据上述条件，我们希望证明任意属于

$\mathcal{V}_{D_A}(x_s)$ 的元素也属于 $\mathcal{V}_{D_A}(x_t)$，由此可证明 $\mathcal{V}_{D_A}(x_s) \subset \mathcal{V}_{D_A}(x_t)$。对于任意元素 $v \in \mathcal{V}_{D_A}(x_s)$，我们讨论两种情况： $v \in \mathcal{V}_2(x_s)$ 和 $v \in \mathcal{V}_1(x_s)$。

情形*1*： $v \in \mathcal{V}_2(x_s)$。根据集合 $\mathcal{V}_2(x_s)$ 的定义，当所有 $k \in \{C, ..., J\}$ 的索引 $(\delta_{As}, \delta_{As} + \delta_{Bs}, \delta_{Bs} - \delta_{ks}, -\delta_{ks})$ 增大时，该集合也会增大。因此，我们知道在条件（16）下， $\mathcal{V}_2(x_s) \subset \mathcal{V}_2(x_t)$。

情形*2*： $v \in \mathcal{V}_1(x_s)$。如果 v满足 $\delta_{At} + v_A \geq \delta_{Bt} + v_B$且 $\delta_{At} + v_A \geq \delta_{Bt} + \delta_{kt} + v_{Bk}$对所有 $k \in \{C, .., J\}$均成立，那么在条件（16）下， $v \in \mathcal{V}_1(x_t)$。否则， v应满足以下条件之一：

$$\{\delta_{Bt} + v_B > \delta_{At} + v_A\} \text{ or } \{\exists k \in \{C, .., J\},\ \delta_{Bt} + \delta_{kt} + v_{Bk} > \delta_{At} + v_A\}.$$

由于我们假设对所有 $k \in \{C, .., J\}$， $\Gamma_{Bk} \leq 0$，上述条件中的任意一个都将意味着

$$\delta_{Bt} + v_B > \delta_{At} + v_A \geq 0,\ \delta_{Bt} + v_B + \Gamma_{AB} \geq 0.$$

然后我们可以验证， v 在条件（16）下满足集合 $\mathcal{V}_2(x_t)$ 中的所有条件，因此我们知道 $v \in \mathcal{V}_2(x_t) \subseteq \mathcal{V}_{D_A}(x_t)$。

现在我们已经证明了条件（16）。通过逆否命题并对 α_i 取期望，可得命题5中的第三个限制：

$$P_s(D_A \mid x_s, x_t) > P_t(D_A \mid x_s, x_t) \Longrightarrow \{\Delta_{s,t}\delta_A > 0\} \vee \{\Delta_{s,t}(\delta_A + \delta_B) > 0\}$$
$$\vee \{\exists k \in \{C, .., J\},\ \Delta_{s,t}(\delta_B - \delta_k) > 0\} \vee \{\exists k \in \{C, .., J\},\ \Delta_{s,t}\delta_k < 0\}.$$

□