

选择实验方法研究进展*

全世文

内容提要:选择实验近年来快速发展并被广泛应用于各个领域的价值评估。本文指出,在一系列陈述偏好方法中,选择实验具有两个核心优势:一是完备的经济理论基础和实证方法体系;二是与真实市场情景相吻合的高信息负荷。高信息负荷在为研究提供更多偏好信息的同时,也大幅提高了实验设计的难度,并在实验过程中引起了一系列偏差。本文从实验设计效率、现状偏差、实验设计维度、信息处理策略、排序效应和假想偏差六个方面介绍了目前学界对选择实验方法效度的研究进展。

关键词:选择实验 实验设计 现状偏差 信息处理策略 假想偏差

一、引言

由于显示偏好数据的局限性,陈述偏好方法被经济学家广泛用于分析消费者行为并在此基础上评价市场产品和非市场产品的价值^①,进而为制定政策和营销策略提供依据。陈述偏好方法可以分为两大类:条件价值评估(Contingent Valuation)^②与选择模型(Choice Modelling)。从20世纪80年代初,条件价值评估方法在备受争议的环境中仍然取得了长足的发展,争议的焦点集中在因评估方法的假想特征引起的一系列偏差。为了消除或者减小这些偏差,研究者就条件价值评估方法的操作流程达成了一系列规范性的共识,在此基础上,条件价值方法的评估结果成了很多公共政策制定的重要依据(Arrow et al, 1993)。

相比之下,选择模型从20世纪90年代末才逐渐被农业经济学家和资源环境经济学家采用;在一系列选择模型中,选择实验(Choice Experiment)是应用最为广泛的一种方法^③。在一个选择情景中,被试者需要在一组虚拟产品或选项之间做出选择,该过程被假设为满足效用最大化的原则。每个虚拟产品均由多个产品属性进行定义,属性取值的差异决定了产品差异。一个完整的选择实验通常要求被试者在多个选择情境中做出多次决策。依据被试者

的决策信息,研究者可以分析消费者对各个产品属性的偏好,并可以进一步评估产品属性的价值,据此为公共政策或营销策略的制定提供依据。这种决策模式也反映出了选择实验方法与条件价值评估方法的差异:前者同时对多个产品属性的价值进行评估,而后者主要就单一的产品目标进行价值评估。在实践中,由于营销决策或政策干预的目标通常都只涉及私人物品或公共物品的某个属性,而非物品整体的供给和需求,因此,选择实验在最近十年取得了快速的发展,在农业经济学、资源与环境经济学、交通经济学和市场营销学等各个领域中都得到了大量的应用。

Lancaster(1972)提出的微观效用理论为选择实验从属性层面上进行价值评估提供了理论基础。该理论假设产品可以被描述为一系列属性的组合,而消费者对产品的效用取决于构成产品的属性结构和属性水平,而非直接来源于产品本身。在此基础上,选择实验要求被试者通过对比属性进而选择产品的模式也被认为与消费者在真实市场情景中的决策过程更加吻合,而优于条件价值评估方法令消费者直接在产品和金钱价值之间进行权衡的模式(Lusk & Schroeder, 2004)。在实证层面上,基于McFadden(1974)随机效用理论(Random Utility Theory)的离散选择模型则完美地结合了选择实验

* 全世文,中国社会科学院农村发展研究所,邮政编码:100732,电子邮箱:quanshiwen@163.com。感谢匿名审稿人的修改建议,文责自负。

的理论模型与计量模型。近年来,离散选择模型的快速发展,尤其是 Logit 家族的不断丰富和完善,为构建选择实验的实证模型提供了有力的支撑。与此同时,计算机技术的发展大幅提高了运算速度,也为估算复杂的离散选择模型提供了保障。

从应用研究的现实意义来看,选择实验受到青睐的一个更重要的原因在于其更高的信息负荷,也就是说,研究者可以通过选择实验采集更多的消费者偏好信息。首先,多属性的实验设计使选择实验可以同时多个目标进行价值评估,研究者可以借此分析消费者对不同产品属性的偏好结构,了解产品属性间存在的相关关系。其次,选择实验通常会要求被试者连续在多个选择情境中进行决策,这也利于研究者更加准确地分析消费者偏好。正因为如此,研究者通过选择实验得到的数据通常是面板结构。信息潜力为选择实验提供了广阔的应用空间,使研究者可以在产品属性的层面上讨论更多的问题,因此大幅提高了选择实验方法对不同领域研究者的吸引力。

近年来,国内学者采用选择实验开展的应用研究也越来越多,尤其集中在资源与环境经济学和农业经济学领域。但国内学者在应用选择实验的过程中,对实验设计规范性的把握和实验结果效度的讨论都远远不够。然而,保证实验结果的效度恰恰是选择实验应用研究中最为关键的问题。Louviere et al(2000)指出,陈述偏好方法的核心问题不在于其经济学的理论基础和数据的可得性,而在于其分析结果的效度;因为结果的效度涉及研究者在多大层面上可以利用模拟的假想情景推测真实市场中的消费者行为与需求。

事实上,选择实验间接测度支付意愿(Willingness-to-Pay)的问题模式被早期的研究者认为可以回避或者减小条件价值评估方法中存在的一部分偏差。例如,选择实验可以克服二元选择问题中被试者过度回答“是”的偏差和条件价值评估在设定假想情景时因“框架效应”(Framing Effect)造成的偏差(Bateman et al,2002);又例如,由于被试者在选择实验中面临的主要任务是基于属性差异在虚拟产品中进行对比,而非直接赋予产品属性一个金钱价值,因此研究者认为选择实验的“范围效应”(Scope Effect)和“假想偏差”(Hypothetical Bias)都弱于条件价值评估。然而,随着行为经济学对认知心理学的引入以及经济学家对选择实验方法研究的逐渐深入,选择实验也被证明存在一系列新的偏差。例如,

消费者在选择情景中不仅会在虚拟产品之间进行权衡,还会将虚拟产品与其在真实市场中购买的产品进行比较。这意味着调查对象有可能弃选实验设计者提供的所有虚拟产品,而忽略了该情形则会引起偏差。又例如,调查对象由于认知能力有限,在面对高信息负荷的选择情景时,会根据“启发式”(Heuristics)采取一系列信息处理策略用来简化决策过程,而这些策略可能违背理性人假设,从而引起偏差。

伴随着应用研究的快速增加,近年来经济学家围绕选择实验方法的效度也展开了广泛而深入的研究。虽然现有的研究尚未达成普遍一致的共识,但是,已有的研究仍然为提高选择实验的效度做出了重要的贡献,更为关键的是,这些研究为后续研究提供了思考与改进的方向。而且,其中的一些研究结论也已经被应用研究中的实验设计者广泛采用。在上述背景下,本文试图从几个研究者关注的热点问题出发,系统地介绍改进选择实验方法效度的研究进展,重点说明这些问题产生的原因、影响和解决方法。

二、实验设计效率

选择实验方法的实施过程可以分解为五个步骤:(1)选择产品属性;(2)定义产品属性的水平;(3)构造虚拟产品,进而构造选择集;(4)开展调研与收集数据;(5)建模分析。其中,设计效率的问题主要集中在第三个环节,即根据已定义的产品属性及其水平构造选择情景。实验设计的目的首先在于识别模型,从而保证偏好参数可以被估计,其次在于获取统计学意义上更有价值的估计量(即效度更高的估计量)。实际上,所有的计量模型都涉及模型优化从而提高估计效率的问题,而这一问题之所以在选择实验的研究中更加突出,主要是由于研究者可以人为地设计虚拟产品的属性及其水平。这意味着研究者可以通过实验设计控制模型中的自变量,从而得到有效性更高的参数估计量。又因为选择实验本身的信息负荷较高,因此,实验设计被认为是选择实验应用中的核心环节(Scarpa & Rose,2008)。实验设计可以被认为是一个在有限约束条件下(包括样本容量、选择集数量、选择集中的产品数量等限制条件)获取最大信息量的过程(Carlsson & Martinsson,2003)。

在产品属性及其水平已定义条件下,全因子设计(Full/Complete Factorial Design)可以保证在产品构造层面上无信息损失(Louviere et al,2000)。但是,在多属性条件下,全因子设计会产生数量过多

的虚拟产品,因此,实践中更多地采用部分因子设计(Fractional Factorial Design)来定义虚拟产品。提高设计优度的第一个问题就在于如何通过部分因子设计选择一个合适的产品子集。Huber & Zwerina (1996)定义了部分因子设计的四个原则:正交性、属性水平的平衡、最小重叠和效用平衡。这些原则都有助于提高设计效率,但其局限性表现在:仅适用于评价线性模型,而且难以同时实现这四个原则。在后续的研究中,正交性和属性水平的平衡成了线性模型中进行实验设计的基本准则,同时满足这两个条件的部分因子设计被称为“正交设计表”(Orthogonal Array)(Kuhfeld et al,1994)。

实验设计的目标是尽可能地降低参数估计量的标准误。例如,在线性模型 $y = X\beta + \epsilon$ 中,通过最小二乘法获得的参数估计量 $\hat{\beta}$,其协方差与信息矩阵 $X'X$ 的逆矩阵 $(X'X)^{-1}$ 成正比;因此,最大化 $X'X$ 或最小化 $(X'X)^{-1}$ 便是实验设计的目标。由于信息矩阵通常会因为样本量或者自变量个数的增加而快速增大,且单独考察信息矩阵不便于进行实验设计之间的比较,因此,一系列设计效率指标被研究者构造出来,用来量化地评价设计优度。经典的设计效率指标包括 A 效率, D 效率, G 效率等,最大化效率指标意味着最小化与之相对应的误差: A 误差, D 误差, G 误差等。其中,由于不受编码影响和便于计算的优点, D 效率的指标应用最为广泛,在线性模型中, D 效率被定义为(Kuhfeld et al,1994):

$$D\text{-efficiency} = \frac{100}{N|(X'X)^{-1}|^{\frac{1}{K}}} \quad (1)$$

其中, N 表示观测值数目, K 表示待估参数的数目。正交设计表意味着线性模型满足 100% 的 D 效率。许多统计分析软件(例如 SPSS, SAS 等)都开发了相应的模块用来提供在产品属性及其水平已定义条件下的正交设计表。

然而,在选择实验的概率选择模型中,参数的协方差不再单纯由 $X'X$ 的信息进行定义。对应的信息矩阵为 Fisher 信息矩阵,即对数似然函数的二阶导:

$$I(\beta, X) = \frac{\partial^2 \ln L}{\partial \beta \partial \beta'}$$

于是,渐进协方差矩阵的估计值为:

$$\Omega(\beta, X) = \{E[I(\beta, X)]\}^{-1} = \left(-\frac{\partial^2 \ln L}{\partial \beta \partial \beta'}\right)^{-1}$$

Scarpa & Rose (2008) 将其对应的 D 误差定义为:

$$D\text{-error} = |\Omega(\beta, X)|^{\frac{1}{K}} \quad (2)$$

最小化(2)式定义的 D 误差是选择实验应用中最常用的实验设计目标。式(2)同时也反映了选择实验设计效率异于线性模型的两个特征(Zwerina et al,1996):其一,参数的非线性导致效率指标不能单纯通过自变量矩阵进行优化,即效率指标依赖于参数 β ;其二,观测值的定义方式导致效率指标不仅依赖于构造的虚拟产品,还依赖于构造的选择集。因此,正交设计表在选择实验的应用研究中通常不能达到最优的设计效率,优化实验设计不仅需要依赖研究者对于参数 β 的预判信息,还依赖于研究者选择的实证模型。当无法预先获得关于参数 β 的信息时,一种简单的处理方式是假设 $\beta=0$ (Zwerina et al,1996; Carlsson & Martinsson,2003)。研究者也可以通过查阅相关资料或开展试调研获得关于参数 β 的信息,据此设计效率更高的选择实验(Ferrini & Scarpa,2007)。另一种常用的方法是假设参数 β 服从一个已知的分布,并据此定义一个贝叶斯 D 误差(Sándor & Wedel,2001):

$$D_B\text{-error} = E_{\beta}\{|\Omega(\beta, X)|^{\frac{1}{K}}\} = \int |\Omega(\beta, X)|^{\frac{1}{K}} \hat{f}(\beta) d\beta \quad (3)$$

其中, $\hat{f}(\beta)$ 表示对参数分布的预判断。这一思路被后续的研究广泛应用于 Logit 家族的其他模型设定,例如, Bliemer et al (2009) 利用 $\hat{f}(\beta)$ 展示了在嵌套 Logit 模型中进行最优设计的方法, Bliemer & Rose (2010) 进一步开发了在随机参数 Logit 模型中进行优化设计的方法,其研究结论都说明实验设计效率会显著受到预信息质量的影响, $\hat{f}(\beta)$ 的误设会显著降低设计效率。此外,由于选择实验的应用研究通常以获取支付意愿估计值作为直接目标,因此, Vermeulen et al (2011) 开发了基于 $\hat{f}(\beta)$ 优化以支付意愿估计方差为内涵的 $WTP_{B\text{-error}}$ 目标函数的方法,从而获得最优化支付意愿估计效率的实验设计。

由于研究者在实践中难以预先获取关于参数 β 的有效信息,而误设信息又会造成设计效率的下降,所以仍然有大量的应用研究仅根据正交设计表进行实验设计(Gao & Schroeder, 2009a, 2009b; De-Magistris et al, 2013)。虽然这种设计并不能在选择实验中实现最优化的设计效率,但是正如大量 Monte Carlo 分析所印证的,通过提高样本量可以有效地提高支付意愿的估计效率,这也意味着样本量可以在一定程度上替代设计效率的不足(Lusk & Norwood, 2005; Ferrini & Scarpa, 2007)。

事实上,选择实验的设计效率与实验设计过程中每个环节的信息定义都直接相关。例如,对于产品属性,研究者通常需要考虑的一个问题是:是否在定义虚拟产品时考察属性间的交互效应。从严格意义上讲,无信息损失的全因子设计需要同时包含所有产品属性之间的交互项;而在正交设计表中,添加全部交互项会导致由完美共线性造成的模型无法识别。从现实意义出发,研究者通常只会考虑一部分主要的交互作用。Louviere et al(2000)指出,遗漏交互项通常并不会造成有偏的模型估计,产品属性的主效应可以解释高达90%的行为选择,而二维交互项通常只能解释3%~6%,三维交互项更低,仅有1%~3%。因此,在没有特殊现实意义的条件下,研究者通常会仅根据属性的主效应进行实验设计。此外,在选择集中添加一个“都不选择”的选项被认为可以有效地提高设计效率(Louviere et al, 2000; Hensher et al, 2007)。但是,Vermeulen et al(2008)的研究发现,考虑“都不选择”选项下的实验设计与不考虑“都不选择”的普通实验设计在多元Logit模型和嵌套Logit模型中,参数的估计效度都没有显著的差异。

三、“维持现状”与“都不选择”

在实验设计者提供的虚拟产品集中,通常还会包含一个除虚拟产品以外的“都不选择”(No-Choice或Neither)选项或者“维持现状”(Status-Quo)的选项,由此捕获行为个体退出选择、拒绝回应、不知道如何选择和所列虚拟产品都不喜欢的情形。在环境项目或其他非市场产品的价值评估中,这个选项通常表示维持现状而不做任何改善的情形。而对于市场产品而言,这个选项可以表示消费者放弃选择虚拟产品或者在比较以后,消费者更倾向于选择其实际购买产品的情形。

研究认为,在选择集中提供一个维持现状或都不选择的选项更接近于真实的决策情景(Banzhaf et al, 2001),而且与消费者需求的经济理论更加一致,从而可以得到理论效度更高的福利评价(Bateman et al, 2002)。缺少这一选项意味着消费者被强制性地要求做出选择,那么消费者的真实偏好可能被扭曲,从而造成估计结果有偏(Dhar, 1997; Dhar & Simonson, 2003);在选择实验中,缺少这一选项会导致消费者选择所有虚拟产品的可能性均被高估,因此,希克斯消费者剩余也会被高估(Banzhaf et al, 2001)。

消费者选择“维持现状”或者“都不选择”选项的原因可以分为两类:其一,这种选择是符合理性人假设的最优决策,这意味着消费者在备选的虚拟产品中进行选择时,会有一个基于实际消费经验的保留效用,仅当实验设计者提供的假想产品带来的效用高于保留效用时,消费者才会选择假想产品。其二,这种行为是不符合理性假设条件的“现状偏差”(Status Quo Bias),即消费者会高估对现状的偏好。这种偏差在早期的消费行为研究中被认为与多种因素相关,这些因素包括:损失厌恶、禀赋效应、遗漏偏差、不作为偏好、避免后悔等(Kahneman et al, 1991; Tversky & Kahneman, 1991; Ritov & Baron, 1992)。

这些因素在选择实验的应用研究中得到了进一步的延伸。越来越多的研究将消费者选择“都不选择”或者“维持现状”的行为理解为消费者在面对复杂信息时的一种信息处理策略。例如,Dhar(1997)发现产品的相似度越高意味着消费者面临的不确定性越强,因而越难以做出选择,消费者就更容易退出选择;Dhar & Simonson(2003)进一步认为在产品集中没有明显的占优者时,消费者选择“维持现状”或“都不选择”可能仅仅为了节省时间和减少认知努力的投入;Boxall et al(2009)也发现随着选择实验信息负荷的上升(复杂程度的上升),被试者会更加倾向于选择“维持现状”或“都不选择”。

事实上,“维持现状”和“都不选择”是两个不同问题,两者表达的经济学意义亦存在差异。其中,前者更多地表达了理性选择和高估现状效用的偏差;而后者则更强调行为个体在面临不确定性和复杂信息时的策略行为。这两个不同的问题对模型估计结果的后续讨论可能带来显著的差异(Banzhaf et al, 2001; Kontoleon & Yabe, 2003)。但是,由于这两个问题的实验设计在进入模型的方法上高度一致,且经济意义上也存在一定程度的重合,因此,大多数研究会根据其研究目标的偏重,在实验设计中仅界定一种问题形式。目前,仅有Lancsar et al(2007)与Pedersen(2011)在实验设计中同时设定了“维持现状”和“都不选择”的选项,并在实证模型中对两者的效果进行了分离。

从进入估计模型的方式来看,研究者一般会通过添加一个“选项特定常数”(Alternative Specific Constant, ASC)来描述“都不选择”这种行为。而“维持现状”的问题则不仅可以通过设定ASC来测度,还可以通过显示偏好数据进行测度(例如, Bon-

nichsen & Ladenburg, 2010)。通常,当行为个体选择“维持现状”或“都不选择”的选项时,设定 $ASC=1$, 否则, $ASC=0$ 。从经济意义上看, ASC 反映了“维持现状”的基准效用, 或者所有未包含进模型的其他产品属性带来的平均效用 (Meyerhoff & Liebe, 2009; Bonnichsen & Ladenburg, 2010)。 ASC 既可以设定为一个恒定的参数 (意味着“维持现状”表现为一种通用属性), 也可以设定为随机参数 (意味着个体现状存在差异)。此外, 还可以在模型中设置 ASC 与个体特征的交互项, 从而反映不同特征的行为个体选择“维持现状”或“都不选择”的行为差异 (Meyerhoff & Liebe, 2009; Lanz & Provins, 2012)。

由于“维持现状”或“都不选择”的选项会同时影响到其他所有备选产品的选择概率, 因此, 多元 Logit 模型中关于“无关选项相互独立”(IIA) 的假设就无法成立 (Dhar & Simonson, 2003; Kontoleon & Yabe, 2003)。于是, 当选择实验中引入这一选项时, 通常需要采用嵌套 Logit 模型或随机参数 Logit 模型进行估计。Balcombe & Fraser (2011) 开发了一个根据嵌套 Logit 模型处理这一问题的通用框架, 并解释了选项之间的相似性会增加被试者放弃选择的概率。由于随机参数 Logit 模型更加宽松的假设条件和在捕获消费者偏好异质性上的优势, 近年来, 在选择实验的应用研究中, 更多的研究者采用了随机参数 Logit 模型进行估计, 自然地消除了 IIA 假设对于“维持现状”或“都不选择”这一选项的限制。

四、实验设计维度

与条件价值评估方法相比, 选择实验的一个核心优势表现在实验设计的高信息负荷。信息负荷的高低并没有一个统一的测量标准, 但选择实验的每个实验设计维度都会对信息负荷产生影响, 这些维度包括: 属性的数量、属性水平的数量、属性水平的定义区间、产品选项的数量、选择情景的数量和产品选项的相似程度等 (DeShazo & Fermo, 2002; Burton & Rigby, 2012)。尽管部分因子设计在很大程度上降低了全因子设计中的超高信息负荷, 但是, 即使是采用经过设计效率优化的部分因子设计, 选择实验的信息负荷仍然远高于条件价值评估。而且, 以上各个设计维度的提高都会大幅增加实验的总信息负荷。研究者需要关注的问题是每一个设计维度应该如何确定 (例如, 选择多少属性用于实验设计,

每个选择集中提供几个产品选项等)。根据经济学理论, 选择实验的评估结果不应该受到实验设计维度的影响。然而, 近年来的大量研究都得到了与此相悖的结论。

现有研究从不同的角度讨论了实验设计维度对选择实验的参数和支付意愿估计结果造成的影响。根据数据来源的差异, 这些研究可以分为两类: 一类采用 Monte Carlo 技术通过模拟数据来分析实验设计维度对实验结果产生的影响 (例如, Lusk & Norwood, 2005; Ferrini & Scarpa, 2007); 另一类则结合具体的应用研究, 通过调查数据来分析这一问题 (例如, Hensher, 2006; Islam et al, 2007; Louviere et al, 2008)。绝大多数研究证实了实验设计维度会显著地影响实验结果, 但研究者并未就影响效果达成一致性的结论 (参见综述文献 Burton & Rigby, 2012)。例如, 就属性的数量而言, 许多研究发现随着属性数量的上升, 一些属性的支付意愿会上升, 而另一些属性的支付意愿则会下降 (例如, Hensher, 2006; Islam et al, 2007); 而 Gao & Schroeder (2009b) 则发现支付意愿的估计值与属性数量之间具有一个二次型的关系。DeShazo & Fermo (2002) 发现, 随机误差会随着信息负荷的上升而上升, 而 Islam et al (2007) 在一些情况下得到了相反的结论, Swait & Adamowicz (2001) 则发现随机误差与信息负荷之间呈二次型关系。

实验设计维度对选择实验结果产生影响的原因可以分为两种: 其一是在统计学意义上由纯粹的数据结构变化产生的影响。选择哪些产品属性, 如何定义属性水平, 在每个选择集中提供多少虚拟产品, 等等, 这些维度的设计会直接影响用于估计实证模型的数据结构, 进而影响估计结果。上述通过 Monte Carlo 模拟数据进行的分析由于不涉及调查对象的真实偏好, 其分析结论可以被视为完全由统计学原因产生的影响。这些分析的结论同样证实了信息负荷会显著影响实验结果。这意味着间接效用函数中的随机效用部分不仅包括了不可观测的产品属性、消费者特征等因素, 还包括了实验设计的成分 (Louviere et al, 2008)。其二是调查对象的偏好在不同的实验设计维度下存在差异。标准微观效用理论假设消费者偏好在不同的决策情景和不同的决策时点具有一致性和稳定性, 而大量研究则证实选择实验中的消费者行为会违背这一假设。

造成消费者偏好依赖于实验设计维度的原因同样可以分为两种情况: 一种情况是实验设计本身存

在不合理因素。例如,在选择属性的阶段,实验设计者如果遗漏了对消费者决策具有重要影响的属性时,消费者可能会根据其他产品属性的水平推测该遗漏属性的水平(Tonsor,2011)。这意味着消费者对多个属性的偏好具有结构性特征,而不合理的实验设计会影响这种偏好结构。另一种情况是调查对象在实验情景中的决策违背了理性行为人的假设。选择实验的高信息负荷意味着研究者有机会采集更多的消费者偏好信息,但同时也意味着参与实验的被试者面临更大的认知负担(Swait & Adamowicz, 2001; Louviere et al, 2008; Burton & Rigby, 2012)。一方面,被试者个体的认知能力有限;另一方面,在虚拟的决策情境中,被试者缺乏足够的激励投入足够的认知资源(时间和精力);于是,认知缺失或认知不足导致被试者没有按照完全理性的假设条件在属性和产品间进行权衡。现有研究发现,选择实验中的被试者会根据启发式采取一系列信息处理策略来简化决策过程,下文将对这一问题进行综述。总之,在信息负荷不同的实验情景中,非理性的消费者可能会陈述完全不同的偏好。

实验设计维度会对实验结果造成影响,这一问题对经济学家采用选择实验方法开展应用研究提出了挑战。研究者为了得到有效的实验结果,就必须谨慎地考察各个实验设计维度。从统计学意义上讲,是否存在最优的实验设计维度,最优的设计维度应该如何确定,现有研究尚未回答这些问题。但通过专家访谈和焦点小组访谈等方法预先了解消费者关切的产品属性及这些属性水平的分布则可以有效避免一部分由设计维度定义不合理导致的偏差(Louviere, 2010)。相比之下,选择实验的研究者目前更重视的一个问题是调查对象的非理性决策问题。实验设计者面临着“完备性”与“简洁性”之间的权衡(Martinsson et al, 2001):设计一个高信息负荷的“完备性”实验可以获取更全面的偏好信息,但却会因为认知负担过重而导致被试者决策非理性程度更高;设计一个低信息负荷的“简洁性”实验虽然可以降低非理性行为的概率,却会拉大模拟情景与现实情景的差距,而且会增加被试者采取策略性行为的概率。

在实践中,并没有一个严格的理论标准用来指导研究者设计每一个维度,但现有研究仍然达成了一些经验性的共识来平衡实验设计的“完备性”与“简洁性”。例如,针对产品属性的数量,绝大多数研究都选择3~6个属性来设计实验,很少研究会选择

超过8个属性。当属性数量较多导致优化设计后的虚拟产品数量仍然很多时,研究者会采用“分块技术”(令多个调查对象分担完成一份实验设计问卷)来降低每个被试者所面对的认识负担。

五、信息处理策略

认知心理学将人脑理解为信息处理器,人体通过各种知觉过程将外部信息输送到信息处理器,然后储存起来或者进一步通过动作处理器控制行动。在面对大体量的信息时,由于认知资源(时间和精力)有限,人体在信息处理过程中通常不会采用缜密的算法,而是会采用“启发式”。用经济学的语言来表达,算法对应于理性经济人假设,而启发式则通常表现为对理性假设的违背^④。就选择实验而言,研究者广泛证实了在有限认知能力约束下的被试者会基于启发式采取一系列信息处理策略(Information Processing Strategies),用来简化决策过程(Hensher, 2007, 2010a)。研究者关注的问题是,被试者会采取哪些信息处理策略,如何识别这些策略,如何将其引入实证模型,引入以后是否会影响福利测度。

Deshazo & Fermo(2004)与Cameron & Deshazo(2010)提出了两个模型从理论层面上解释了被试者在选择实验中采取的信息处理策略:其一是“被动有限理性模型”(Passive Bounded Rationality Model),该模型假设实验参与者会考虑设计者提供的所有信息,但随着信息负荷的上升,实验参与者做错误决策的概率会随之上升。其二是“理性自适应行为模型”(Rationally Adaptive Behavioral Model),该模型假定被试者在决策过程中会考虑信息处理的认知成本和认知收益,被试者面临的问题是在有限的认知资源约束下,优化对不同信息的认知资源配置;认知成本取决于被试者的认知能力、时间的机会成本以及实验的信息负荷和设计形式等,而认知收益表现为被试者可以避免从非最优决策中造成的效用损失(Dellaert et al, 1999)。

在选择实验的实验设计维度中,“属性”和“选项”是两个最基本的设计单元,也是被试者采取信息处理策略的两个基本单元。也就是说,信息处理策略可以被划分为基于属性的策略和基于选项的策略(Mintz et al, 2013)。一种典型的基于选项的处理策略就是被试者过度地选择“维持现状”或者“都不选择”的选项,这意味着被试者不愿意在其他高信息负荷的选项之间投入认知努力进行决策。相比之下,研究者更重视基于属性的处理策略。Gilovich

et al(2002)在心理学中介绍了一系列由启发式引起的决策偏差,在此基础上,Hensher(2007,2010a)在选择模型中定义了三种属性处理策略,其中的“属性非参与”(Attribute Non-Attendance)策略随后在农业经济学、资源与环境经济学和交通经济学的应用研究中被广泛证实,而且,属性非参与策略与理性自适应行为模型的解释也相互一致;因此,属性非参与成为了目前选择实验的研究者关注最多的一种信息处理策略。

调查对象在实验过程中采取属性非参与策略主要有两个原因:一是调查对象对忽略的属性具有真实的零偏好(该属性的边际效用为零);二是调查对象用来简化实验的复杂性(Carlsson et al, 2010; Alemu et al, 2013)^⑤。识别非参与属性的传统方法是在选择情景的问题结束后直接询问实验参与者在刚刚的决策过程中是否忽略以及忽略了哪些产品属性,这种方式被称为“陈述性非参与”(Stated Non-Attendance)。处理陈述性非参与信息的标准方法是设定一个约束模型,其中,非参与属性的参数(边际效用)被限制为零值(例如, Hensher et al, 2005; Campbell et al, 2008)。然而,后续的研究却发现,调查对象陈述非参与的属性事实上并未被忽略,因为在非约束模型中获取的参数估计结果仍然显著(例如, Campbell & Lorimer, 2009; Carlsson et al, 2010)。对陈述信息信度的质疑促进了“推测性非参与”(Inferred Non-Attendance)方法的发展^⑥。这种方法使用概率技术在个体选择集的基础上推测每个被试者忽略的属性(例如, Hensher et al, 2012; Scarpa et al, 2013)。一些对比分析进一步发现,陈述性和推测性的非参与属性之间存在明显的差异(例如, Kragt, 2013; Scarpa et al, 2013)。

陈述性非参与信息的低信度可能由调查对象在认知或回忆过程中的不准确性引起(Scarpa et al, 2013),也可能源于调查对象故意的谎报信息(Kragt, 2013),这也是令一些研究者认为推测性信息优于陈述性信息的主要原因。然而,由于研究者永远无法获知调查对象真实的属性非参与信息,因此,研究者事实上很难判断陈述性与推测性信息孰优孰劣。正如 Hensher(2010b)所述,“目前,并不清楚将各种信息处理策略识别为某一概率的分析技术是否是对直接询问被试者属性处理策略的补充性问题的一个进步。”现有研究也证实了陈述性非参与并非一种噪音,而是具有较高的信息价值。尽管调查对象可能并没有完全忽略其声称忽略的属性,但一

种最可能的情形是调查对象在决策过程中赋予了这些属性更低的重要性权重(例如, Carlsson et al, 2010; Hess & Hensher, 2010)。

被试者忽略产品属性的策略意味着偏好的连续性公理被违背,因此,不考虑这一策略会造成估计结果有偏。在实践中,分析属性非参与行为是否以及如何影响模型估计结果和相关的福利测度是这一类文献的主要研究目标。一个典型的分析范式是对比在模型中解释属性非参与前后的估计结果,比较的对象包括模型的拟合效果和支付意愿的分布。将陈述性非参与与信息引入模型的方法除了上文提及的将非参与属性的边际效用约束为零值的传统方法以外,还包括其他一些改进的方法,例如 Hensher et al (2007)提出的决定性和随机性排除属性的方法以及 Hess & Hensher(2010)提出的“分离参数方法”,即对每个属性都估计两个系数,其中一个属于在决策过程中考虑该属性的样本,另一个属于忽略该属性的样本。

一个典型的推测属性非参与的方法是设定一个“约束潜类别模型”(Constraint Latent Class Model),该模型允许调查对象属于某一个潜类别,该潜类别的某一个或某几个属性的系数为零值;而其他属性的系数在不同的潜类别之间被约束为等值。该模型由 Scarpa et al(2009)提出,并在后续研究中(例如, Hensher et al, 2012; Scarpa et al, 2013)被广泛使用。此外,Scarpa et al(2009)还提出了一个基于贝叶斯方法的随机属性选择技术,该技术相对于上述约束潜类别模型而言,允许偏好异质性表现为连续分布。Hole(2011)则开发了一种“内生属性参与模型”(Endogenous Attribute Attendance Model),该模型假定被试者首先选择要使用哪些属性进行决策,进而根据这些选择的属性在虚拟产品之间进行权衡。另外, Hess & Hensher(2010)提出了一种更为直接的推测非参与属性的方法,即在个体水平的参数估计值基础上人为地设定一个“小概率”的阈值,将低于阈值的样本视为忽略某一属性。

无论是陈述性非参与还是推测性非参与,现有研究的结论普遍显示,在实证模型中解释了属性非参与行为以后,模型的拟合效度会有所改善(例如, Scarpa et al, 2009; Hensher et al, 2005; Hess & Hensher, 2010)。这说明属性非参与信息的引入对提高选择实验的内部效度具有促进作用。但是现有研究针对支付意愿的比较则没有得到一致的结论。一些研究发现,当模型中解释了属性非参与行为以

后,支付意愿会有所下降(例如,Hensher et al, 2005;Compbell et al,2008;Scarpa et al,2009),另一些研究却得到了上升的支付意愿(例如,Rose et al,2005;Hensher et al,2007),还有一些研究则没有发现显著的差异(例如,Carlsson et al,2010;Kragt,2013)。虽然属性非参与行为对支付意愿的影响是否存在规律还有待进一步探索,但这种显著的影响效果本身则说明,如果研究者忽略了调查对象的属性非参与策略,很可能会得到有偏的估计结果。

六、排序效应

选择实验高信息负荷的一个表现是:被试者需要在多个选择情景中连续地做出决策。标准经济理论假定被试者的偏好在这些不同的选择情景中具有稳定性,但近年来大量的研究结论都推翻了这一假设,即被试者在选择情景中陈述的偏好会受到这些选择情景排列次序的影响(例如,Bateman et al, 2008;Scheufele & Bennett,2012;Carlsson et al, 2012)。这种“排序效应”(Ordering/Sequence Effect)的存在使研究者必须慎重地思考在连续决策的选择实验中如何估算消费者的“真实”偏好。

事实上,排序效应在实验经济学中是一个普遍存在的问题,也是违背标准偏好理论的一个典型代表(Braga & Starmer,2005)。在条件价值评估中,早期的研究就发现了采用双边界二元选择方法时,调查对象在第二次选择中的偏好参数与第一次选择相比存在显著差异^⑦,据此估算的支付意愿也有所不同(例如,Hanemann et al,1991;Cameron & Quiggin,1994)。和条件价值评估相比,选择实验要求被试者做出更多次数的决策,这也就使排序效应在选择实验中成了一个更严重的问题。

现有研究对陈述偏好方法中存在排序效应的原因做出了充分的解释。Day et al(2012)将排序效应分为两类:“位置相依型”和“前例相依型”,前者指由某个选择情景出现的次序引起的排序效应,后者指由先出现的选择情景对后出现的选择情景的影响引起的排序效应。Day et al(2012)总结了现有研究对这两种排序效应的原因解释:导致“位置相依型”排序效应的主要原因是“机制学习”、“偏好学习”和“疲劳效应”。其中,“机制学习”指被试者对决策情景的运行机制不熟悉,需要在不断重复的情景中学习决策规则,从而表达自己的真实偏好。机制学习意味着排序靠后的陈述偏好具有更高的效度,而第一次

决策的误差最大(Braga & Starmer,2005)。“偏好学习”指被试者在进入决策情景之前并没有一个清晰定义的偏好,而是在重复决策的情境中逐渐“发现”自身的偏好。偏好学习是对标准偏好理论的违背,它意味着被试者在不同情景中的偏好可能是变化的,而这种变化过程也正是偏好被“发现”和“定义”的过程(Plott,1996;Bateman et al,2008)。“疲劳效应”指具有有限认知能力的被试者会在重复决策的过程中逐渐降低陈述偏好认真性(Bradley & Daly,1994;Day et al,2012)。与机制学习和偏好学习的效果相反,疲劳效应意味着被试者在排序靠后的选择情景中的决策具有更强的随机性。

造成“前例相依型”排序效应的主要原因包括“策略性行为”和“锚定效应”。其中,“策略性行为”指被试者有意识地错误表达自身的偏好,从而通过干预实验结果达到自身利益的最大化,当研究对象是公共物品时,这种策略性的错误陈述通常是占优策略(Carson & Groves,2007)。Day et al(2012)指出,被试者在实验过程中会记忆已经完成的决策,从而在后面的决策中寻找可以发生策略性行为的机会,被试者记忆的“最选项”会显著地影响消费者随后的决策。“锚定效应”则是违背标准偏好理论的又一种情形,它同样意味着被试者的真实偏好具有不确定性,被试者会将第一个决策情景中的属性水平(包括标的值)视为“合理值”,后期决策过程都会“锚定”到这个合理值。在条件价值评估中,这一效应被称为“起点偏差”。大量实证研究证实了在多选择情景的陈述偏好方法中存在这种锚定效应(例如,Herriges & Shogren,1996;Whitehead,2002)。

如何解决或者回避排序效应?由于引起排序效应的原因多种多样且影响机制存在明显差异,因此,并没有一个统一的解决方案可以在多情景的选择实验中完全消除排序效应的影响。例如,机制学习和偏好学习意味着研究者应当重视排序靠后的决策情景,而疲劳效应和策略性行为则意味着排序靠前的决策情景会提供效度更高的信息。目前,研究者主要采用了以下几种经验性的策略用来缓解排序效应对实验结果的影响。第一种策略是在被试者进入决策情景之前“提前披露”(Advanced Disclosure)实验情景中的信息(包括重复决策的次数、属性水平的值域等)。该设计方案被证实可以有效降低“位置相依型”的排序效应,但无法有效降低“前例相依型”的排序效应(Bateman et al,2009;Day et al,2012)。伴随着第一种策略,第二种策略是进一步提供一个决

策情景的示例供被试者“暖身”。这种策略也是目前选择实验应用研究中被广泛采用的标准设计方案。第三种策略是实验经济学提供的“抵消平衡”(Counterbalance)设计方案,即在问卷层面上平均分配重复情景的每种排序方案。由于选择实验中的选择情景数量较多,造成排序方案的种类非常多,因此,实践中,研究者通常采用随机排序方案作为抵消平衡方案的替代(Street & Burgess,2007)。

但是,正如上文所述,这些应对策略并不能完全消除排序效应的影响。例如,Meyerhoff & Glenk(2015)就发现,提前提供一个选择情景示例的设计方案会引起显著的起点偏差,因此他们认为这种设计弊大于利。由此可见,应用研究者在采用这些策略时,并不能想当然地断定这些策略可以提高实验结果的效度,而是应当有针对性地对模型中存在排序效应的原因进行检验,进而考虑如何降低排序效应的影响。例如,当实验中存在起点偏差而在实证模型中未加控制时必然导致估计量不一致,而控制了起点偏差以后,虽然后续决策的信息价值有所下降,但研究者至少可以得到一致的估计量(Herriges & Shogren,1996)。

七、假想偏差

通过陈述偏好方法估计的消费者偏好与其真实偏好之间的差异被称为“假想偏差”(Hypothetical Bias)。早期的大量实证研究在各个领域都证实了条件价值评估中的调查对象会显著地高估支付意愿(可参考 List & Gallet,2001;Murphy et al,2005)。假想偏差的存在降低了研究者使用陈述偏好推断行为个体真实市场行为的能力,因此也是陈述偏好方法受到一些研究者质疑的一个主要原因。在后续的研究中,许多缓解技术,例如,廉价对话(Cheap Talk)、确定性评分校准(Certainty Scale Calibration)等,都相继被研究者开发出来用来降低假想偏差(参考综述文献 Loomis,2014)。

假想偏差是研究者讨论陈述偏好方法效度时一个传统的研究问题。虽然经济学家在此领域积累了大量的实证研究(主要集中在条件价值评估方法),但迄今为止,学界依然没有形成一个被广泛接受的理论模型来解释假想偏差的成因(Murphy & Stevens,2004;Hensher,2010b;Loomis,2014)。这也导致目前研究者开发的用来降低假想偏差的技术基本都是建立在经验性结论的基础上,而缺乏一个严格的理论证明。研究者普遍认识到,假想偏差与实

验过程中的多项因素都有关系:是否采用真实支付,实验对象是私人物品还是公共物品,提问问题的技术,等等。“非结果性”(Inconsequentiality)是学界在解释假想偏差成因时使用的首要关键词,也就是说,调查对象缺乏对其陈述结果或选择结果的有效承诺,因此,调查对象没有激励在实验中陈述真实的偏好,这也是机制设计理论认为陈述偏好方法激励不相容的主要原因,可以参考 Carson & Groves(2007)对“非结果性”的详细阐释。在“非结果性”的实验中,理性的调查对象会采取“策略性行为”,用来有意识地影响实验结果,从而最大化自身的利益(Carson & Groves,2007;Day et al,2012)。通常情况下,在产品的提供与投标值正相关而不需要真实支付的情况下,占优策略是高估支付意愿;而在需要实际支付但产品的提供与投标值无关时,占优策略是低估支付意愿。

实证研究的结论还显示,假想偏差的程度在私人物品和公共物品之间存在显著差异(List & Gallet,2001),调查对象对实验标的产品的熟悉度和消费经验被用来解释这一差异(Schläpfer & Fischhoff,2012)。此外,假想偏差在不同的问题形式下也存在显著差异,在条件价值评估的方法系列中,开放式问题的假想偏差最高,而二元选择问题则被多数研究者认为是唯一激励相容的问题形式(例如,Arrow et al,1993;Carson & Groves,2007)。正因为如此,大多数选择实验的设计者在一个选择情景中仅提供两个虚拟产品(不含“退出选择”或“维持现状”的选项)供调查对象选择。

与条件价值评估方法相比,目前仅有少量的研究关注选择实验的外部效度问题。许多研究者认为选择实验令调查对象在产品之间进行选择的情景更逼真地模仿了真实市场情景,优于条件价值评估令调查对象在产品 and 金钱价值之间进行权衡的模式,因而不会存在严重的假想偏差(例如,Lusk & Schroeder,2004;Broadbent et al,2010)。正如 List et al(2006)所述,“选择实验可以通过一种有意义且令消费者熟知的方式构造价值形成的过程,因而可以实现其外部效度”。但是,目前有限的实证研究并没有就假想偏差是否应该成为选择实验应用研究的一个核心问题达成一致的结论。Lusk & Schroeder(2004)发现在假想的选择实验中,总支付意愿会被高估,但对单个食品属性的边际支付意愿则未被显著高估。Hudson et al(2005)和 Carlsson & Martinsson(2001)没有识别出选择实验的假想实验组

会高估支付意愿。但是, Loomis et al(2009)则证实了假想实验组非常显著地高估了消费者对降低瓶装水风险的支付意愿。此外, 还有一些研究致力于讨论如何降低假想偏差, 这类研究通常会将标准的选择实验组作为比较的基准, 而在这些实验组中也都证实了被显著高估的支付意愿(例如, De-Magistris et al, 2013; Moser et al, 2013)。但是, 在 Broadbent(2014)的一个同类研究中, 调查对象在选择实验中并未高估边际支付意愿, 反而采用了廉价对话和确定性评分校准技术以后的实验组却显著地低估了支付意愿。

事实上, 学界目前对选择实验外部效度的讨论完全沿用了条件价值评估方法外部效度的研究范式, 也就是说, 上述文献仿照条件价值评估方法的相关研究验证选择实验中是否存在假想偏差, 以及降低假想偏差的经验性技术是否在选择实验方法中同样有效。虽然这些研究提供了一些经验性的证据, 但对于从理论上理解假想偏差的成因则并无建树。由于陈述偏好方法共同的假想特征, 并没有一个有力的理论支撑可以确保选择实验的问题机制满足激励相容的标准而不存在假想偏差。虽然多数实证研究的结论显示, 选择实验方法即使存在假想偏差, 支付意愿被高估的程度也明显小于条件价值评估方法, 但是假想偏差的存在即意味着福利评价有偏, 当研究者进一步开展效益转移分析时, 这种偏差可能同样巨大。

为了提高选择实验的外部效度, 现有研究的普遍做法是设计具有结果承诺的实验。具备“结果性”的设计方案包括以下两种: 第一, 使用可观测的真实市场数据, 即非实验性的显示偏好数据; 第二, 设计具有真实支付承诺的选择实验^⑥。在实践中, 第一种方案明显地受制于显示偏好数据的缺陷(详见 Louviere et al, 2000, pp. 21—23), 因此很少被研究者采用。根据笔者了解, 在农业经济学领域, 仅有 Hudson et al(2005)采用了显示偏好数据作为参照讨论选择实验是否存在假想偏差。事实上, 如果某项产品的显示偏好数据满足价值评估的需要, 也就不需要再使用陈述偏好方法进行研究。

相比之下, 第二种方案的应用更为广泛, 这种方案在实验设计层面上完全等同于陈述性的选择实验, 唯一的不同在于被试者需要为其决策发生真实的支付^⑦。现有研究在验证选择实验是否存在假想偏差时基本上都选择了这种附加真实支付承诺的实验结果作为参照基准(例如, Lusk & Schroeder,

2004; Moser et al, 2013)。而且, 许多经济学家在开展应用研究时直接默认这种设计方案可以保证实验结果的外部效度(例如, Michaud et al, 2013)。但是, 这种设计方案依然存在明显的局限性。首先, 为了发生真实支付行为, 实验设计者必须要保证选择情境中提供的并非虚拟产品, 而是真实市场产品(或者, 虚拟产品至少应该有对应的相似市场产品)。显然, 陈述偏好方法的绝大多数研究对象都无法满足这一条件。其次, 这种设计方案同样属于实验数据, 而人为设定的实验情景在模仿真实市场情景中必然存在信息损失, 而且, 令被试者重复选择的设计方案本身就因为排序效应而被 Carson & Groves(2007)认为是激励不相容的。因此, 正如 Carlsson & Martinsson(2001)所述, “我们对外部效度的检验并不能被视为真实的显示偏好, 而应视为对假想和真实实验之间的差异检验, ……获取真实的显示偏好应该需要其他的激励措施。”最后, 这种需要发生真实支付的实验也会明显提高调查成本。

八、总结与启示

近年来, 选择实验方法在农业经济学、资源与环境经济学和交通经济学等领域的应用研究都得到了快速增长。与传统的条件价值评估方法相比, 选择实验方法的一个核心优势是与真实市场情景相吻合的高信息负荷, 意味着研究者有希望通过选择实验获取个体更多的偏好信息。但是, 选择实验的高信息负荷也增加了实验设计的难度, 而且也实验过程带来了一系列偏差。围绕着这些偏差, 近年来有大量的研究致力于通过改进实验设计或实证模型来提高选择实验的效度。在此背景下, 本文对选择实验方法效度的研究进行了综述, 以期应用研究者提供实验设计层面的指导, 并为后续致力于提高实验结果效度的研究指明方向。

(一) 研究总结

根据本文的综述, 现有研究对选择实验方法效度的讨论可以做出以下归纳:

(1) 基于线性模型开发的正交设计方案通常不能实现选择实验中设计效率的优化目标。选择实验非线性的实证模型意味着实验设计的优化方案不仅与每个实验设计维度的信息定义有关, 还与偏好参数的实际值有关。如果研究者可以提前获知偏好参数的信息, 通过最小化贝叶斯 D 误差可以得到更有效的估计结果。相反, 如果研究者无法提前获取有效的偏好信息, 尚没有统一的结论说明哪种实验设

计最优。目前,仍然有很多研究者在应用中选择正交设计方案,这意味着研究者需要通过提高样本量来弥补设计效率的损失。

(2)研究者在设计选择情景时必须要考虑调查对象弃选所有虚拟选项的情形,否则可能引起由强制选择造成的偏差。排除实验设计维度定义不合理的因素,退出选择可能是由于调查对象确实具有更高的保留效用,也可能是调查对象高估了维持现状的偏好,也可能只是调查对象用来简化决策过程的一种信息处理策略。在实证模型中,研究者可以通过引入一个特定选项的常数项或调查对象的显示偏好来处理这一问题。由于该情形意味着 IIA 假设被违背,因此,选用传统的多元 Logit 模型并不恰当;研究者可以考虑使用嵌套 Logit、随机参数 Logit 或其他模型进行估计。

(3)每个实验设计维度的信息定义都会对选择实验的估计结果以及在此基础上计算的支付意愿造成影响。这种影响一方面来自统计学意义上自变量数据结构的差异,另一方面来自高信息负荷下调查对象采取的信息处理策略。目前的研究尚没有说明在统计学意义上是否存在对实验设计维度的最优设定方法。信息处理策略的存在意味着实验设计者必须在实验的“完备性”和“简洁性”之间做出权衡;保证了完备性意味着调查对象会根据启发式采取更多非理性的信息处理策略,而保证了简洁性则可能引起遗漏变量偏差和策略性偏差。研究者需要结合具体的应用领域谨慎地做出选择。

(4)调查对象在选择实验中会采用信息处理策略,现有研究发现的一个最广泛的策略是“属性非参与”,即调查对象会有意地忽略某一个或几个属性。属性非参与行为违背了偏好的连续性公理,因此忽略这种行为会导致估计结果有偏。研究者可以通过在选择情景之后设计一个问题追问调查对象的属性非参与信息进而引入实证模型,也可以直接在实证模型中识别属性非参与行为。目前的研究并没有一致的结论说明哪种途径更优,而且,两种途径识别的非参与属性往往并不一致。

(5)选择实验的多选择情景设计会引起排序效应。造成排序效应的原因众多,如机制学习、偏好学习、疲劳效应、策略性行为 and 锚定效应等。而且这些因素对调查对象决策的影响效果存在差异。目前,绝大多数实验设计者都会采用以下几种方式来试图降低或消除排序效应:向调查对象提前披露实验设计维度的定义,提供一个选择情景的示例供调查对

象“暖身”,在每一份问卷中将各个选择情景随机排序。但是,这些经验性的策略并不能完全保证排序效应会减弱或消除,应用研究者仍然需要对排序效应进行检验。

(6)选择实验方法比条件价值评估方法具有更高的外部效度,但研究者并不能据此排除选择实验方法仍然存在假想偏差的可能性。应用研究者仍然需要通过实验设计保障估计结果的外部效度。除了在实验过程中引入“廉价对话”或“确定性评分校准”等方法以外,目前学界公认的一种更为有效的方法是直接设计附带有结果性承诺的实验,即调查对象需要为其陈述的选择发生实际支付。但这种方法的应用具有局限性,而且会大幅增加调查成本。

(二)研究启示

自从陈述偏好方法被提出以来,学界对其效度的质疑和批评就一直存在。其中的核心问题是,调查对象在虚拟情景中陈述的偏好是否可以用来分析现实问题。然而,由于显示偏好数据的局限性,陈述偏好方法在这种质疑声中仍然取得了快速发展,并被广泛应用于各个领域的实证研究。其中的代表就是条件价值评估方法,研究者逐渐证实了条件价值评估方法中存在的一系列偏差,并在实验设计和实证模型构建上形成了一系列经验性的操作规范来消除或缓解这些偏差。在国际上,只有谨慎地遵守这些操作规范并对实验结果的效度进行充分论证以后,条件价值评估方法所得的结果才可能具有应用价值。正是在这种规范下,条件价值评估方法在国外才会被广泛地用于公共政策的制定。但国内基于条件价值评估的应用研究对实验设计规范性和方法效度论证的重视都远远不够,这样的分析结果在学术上难以被认可,在实践中也就更难以被政策制定者所采纳。

选择实验近年来逐渐兴起,并成了目前国际上最热门的一种陈述偏好方法,为研究者带来了广阔的应用空间。与此同时,选择实验的方法效度同样是应用研究中最为核心的问题。但是,选择实验方法的效度问题与条件价值评估方法不尽相同。两种方法在情景设定方式上的差异典型地反映在了信息负荷上。条件价值评估中的情景设定令调查对象在金钱价值和产品之间进行权衡,实验设计中的信息负荷通常比较低;而选择实验的情景设定则是令调查对象在产品属性的基础上对虚拟产品进行权衡,实验设计中的信息负荷远高于条件价值评估方法。选择实验方法中存在的一系列新的偏差都围绕着这

一高信息负荷的特点。

现有研究为减小这些偏差也从实验设计的角度提供了一些建议。但从本文的综述可以发现,目前的研究总体上“破”有余而“立”不足,即用来改进实验设计的统一结论尚未广泛形成,这就为大量致力于改进方法效度的学者提供了研究契机。而且,由于信息负荷的差异,同一种偏差在选择实验中的形成机制和解决办法也可能完全有异于条件价值评估。后续的研究应对这种差异予以重视,而不能想当然地认为条件价值评估研究中的实验设计范式对改进选择实验的方法效度同样有效。

从学科层面上看,选择实验的广泛应用与近年来行为经济学的兴起密不可分。选择实验的方法效率度面临的核心挑战是调查对象在高信息负荷情景中会提高非理性决策的概率,这与行为经济学对传统微观经济理论中理性人假设的突破是相一致的。研究者在选择实验中证实的许多偏差(如现状偏差、锚定效应等)都来自行为经济学。当然,研究者在选择实验方法效率度上开展的研究也已经而且将继续为行为经济学的理论体系构建做出贡献。

注:

- ①显示偏好数据局限于现有市场的技术边界,因而无法反映当技术边界发生移动时的情形(Louviere et al, 2000, p. 23)。也就是说,显示偏好数据无法测度新产品的市场价值,例如,生产者提供的一个经过改良的或具有特殊功能的市场产品。此外,在成熟的市场中,显示偏好数据还具有变量的变异度过低、变量间的共线性过高等缺陷,这同样会导致研究者无法有效地评价市场产品的价值(Louviere et al, 2000, pp. 21-22)。
- ②在部分文献中也被译为“或有价值评估”,本文统一称为“条件价值评估”。
- ③在此之前,选择模型的一个起源是营销学和交通经济学中常用的“联合分析”(Conjoint Analysis)方法,早期的一些研究甚至对选择实验和联合分析的概念不做明确的区分。Louviere et al(2010)对联合分析和选择实验进行了详尽的对比,并认为起源于“联合测度”(Conjoint Measurement)理论的联合分析是一个纯粹的数学系统,与行为经济理论并不一致,相反,选择实验则具有明确和系统的经济理论基础。
- ④事实上,关于启发式是否应当被视为非理性在学界仍然存在争论,但启发式在简化决策过程的同时可能会造成决策偏差则是一个不争的事实。
- ⑤Carlsson et al(2010)发现,根据简化信息和零偏好两种原因获得的WTP估计结果之间存在显著差异,据此建议对这两种不同原因的样本进行分离。Alemu et al(2013)使用一个补充性问题来获取实验参与者忽略属性的原因,结

果显示,两种最常被选中的原因分别是简化信息和属性无关;而且,同样发现了两个不同原因的样本之间得到的估计结果存在显著差异。

- ⑥Hess & Hensher(2013)认为,陈述性信息的另一个缺点是内生性问题,即行为个体报告的非参与信息可能与其他不可观测的因素存在相关性。
- ⑦在双边界二元选择方法中,大多数应用研究采用二元Probit模型进行估计。根据经济学理论,偏好的稳定性要求通过两次选择的估计式得到的系数不应存在显著差异,但实践中很少研究能够得到这一结论。
- ⑧另一种近年来被部分研究者使用的设计方案是合并陈述偏好数据与显示偏好数据,前者的作用表现在优化数据结构,后者的作用表现在保证外部效度。关于此问题的详细内容可以参考Louviere et al(2000)第8章内容。
- ⑨通常情况下,在选择实验的多个选择情境中,实验设计者会随机抽选一个情境中的决策要求被试者真实支付。当然,上述试验程序需要提前告知被试者。

参考文献:

- Alemu, M. H. et al(2013), “Attending to the reasons for attribute non-attendance in choice experiments”, *Environmental and Resource Economics* 54(3):333-359.
- Arrow, K. et al(1993), “Report of the NOAA panel on contingent valuation”, *Federal Register* 58(10):4602-4614.
- Balcombe, K. & I. Fraser(2011), “A general treatment of ‘don’t know’ responses from choice experiments”, *European Review of Agricultural Economics* 38(2):171-191.
- Banzhaf, M. R. et al(2001), “Opt-out alternatives and anglers’ stated preferences”, in: J. Bennett & R. Blamey (eds.), *The Choice Modelling Approach to Environmental Valuation*, Edward Elgar.
- Bateman, I. J. et al(2002), *Economic Valuation with Stated Preference Techniques: A Manual*, Edward Elgar.
- Bateman, I. J. et al(2008), “Learning design contingent valuation(LDCV): NOAA guidelines, preference learning and coherent arbitrariness”, *Journal of Environmental Economics and Management* 55(2):127-141.
- Bateman, I. J. et al(2009), “Procedural invariance testing of the one-and-one-half-bound dichotomous choice elicitation method”, *Review of Economics and Statistics* 91(4):806-820.
- Bliemer, M. C. J. et al(2009), “Efficient stated choice experiments for estimating nested logit models”, *Transportation Research Part B: Methodological* 43(1):19-35.
- Bliemer, M. C. J. & J. M. Rose(2010), “Construction of experimental designs for mixed logit models allowing for correlation across choice observations”, *Transportation Research Part B: Methodological* 44(6):720-734.
- Bonnichsen, O. & J. Ladenburg(2010), “Reducing status quo bias in choice experiments: An application of a protest

- reduction entreaty”, Working Paper, University of Copenhagen, Department of Food and Resource Economics.
- Boxall, P. et al(2009), “Complexity in choice experiments: Choice of the status quo alternative and implications for welfare measurement”, *Australian Journal of Agricultural and Resource Economics* 53(4):503—519.
- Bradley, M. & A. Daly(1994), “Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data”, *Transportation* 21(2):167—184.
- Braga, J. & C. Starmer(2005), “Preference anomalies, preference elicitation and the discovered preference hypothesis”, *Environmental and Resource Economics* 32(1):55—89.
- Broadbent, C. D. et al(2010), “Testing for hypothetical bias in a choice experiment using a local public good: Riparian forest restoration”, *International Journal of Ecological Economics and Statistics* 19(F10):1—19.
- Broadbent, C. D. (2014), “Evaluating mitigation and calibration techniques for hypothetical bias in choice experiments”, *Journal of Environmental Planning and Management* 57(12): 1831—1848.
- Burton, M. & D. Rigby(2012), “The self selection of complexity in choice experiments”, *American Journal of Agricultural Economics* 94(3):786—800.
- Cameron, T. A. & J. Quiggin(1994), “Estimation using contingent valuation data from a ‘dichotomous choice with follow-up’ questionnaire”, *Journal of Environmental Economics and Management* 27(3):218—234.
- Cameron, T. A. & J. R. DeShazo(2010), “Differential attention to attributes in utility-theoretic choice models”, *Journal of Choice Modelling* 3(3):73—115.
- Campbell, D. et al(2008), “Incorporating discontinuous preferences into the analysis of discrete choice experiments”, *Environmental and Resource Economics* 41(3):401—417.
- Campbell, D. & V. S. Lorimer(2009), “Accommodating attribute processing strategies in stated choice analysis”, European Association of Environmental and Resource Economists Annual Conference, Amsterdam.
- Carlsson, F. & P. Martinsson(2001), “Do hypothetical and actual marginal willingness to pay differ in choice experiments? Application to the valuation of the environment”, *Journal of Environmental Economics and Management* 41(2):179—192.
- Carlsson, F. & P. Martinsson(2003), “Design techniques for stated preference methods in health economics”, *Health economics* 12(4):281—294.
- Carlsson, F. et al(2010), “Dealing with ignored attributes in choice experiments on valuation of Sweden’s environmental quality objectives”, *Environmental and Resource Economics* 47(1):65—89.
- Carlsson, F. et al(2012), “The first time is the hardest: A test of ordering effects in choice experiments”, *Journal of Choice Modelling* 5(2):19—37.
- Carson, R. & T. Groves(2007), “Incentive and informational properties of preference questions”, *Environmental and Resource Economics* 37(1):181—210.
- Day, B. et al(2012), “Ordering effects and choice set awareness in repeat-response stated preference studies”, *Journal of Environmental Economics and Management* 63(1):73—91.
- Dellaert, B. G. et al(1999), “The effect of attribute variation on consumer choice consistency”, *Marketing Letters* 10(2):139—147.
- De-Magistris, T. et al(2013), “On the use of honesty priming tasks to mitigate hypothetical bias in choice experiments”, *American Journal of Agricultural Economics* 95(5):1136—1154.
- DeShazo, J. R. & G. Fermo(2002), “Designing choice sets for stated preference methods: The effects of complexity on choice consistency”, *Journal of Environmental Economics and Management* 44(1):123—143.
- DeShazo, J. R. & G. Fermo(2004), “Implications of rationally-adaptive pre-choice behavior for the design and estimation of choice models”, Working Paper, School of Public Policy and Social Research, University of California, Los Angeles.
- Dhar, R. (1997), “Consumer preference for a no-choice option”, *Journal of Consumer Research* 24(2):215—231.
- Dhar, R. & I. Simonson(2003), “The effect of forced choice on choice”, *Journal of Marketing Research* 40(2):146—160.
- Ferrini, S. & R. Scarpa(2007), “Designs with a priori information for nonmarket valuation with choice experiments: A Monte Carlo study”, *Journal of Environmental Economics and Management* 53(3):342—363.
- Gao, Z. & T. C. Schroeder(2009a), “Consumer responses to new food quality information: Are some consumers more sensitive than others?”, *Agricultural Economics* 40(3):339—346.
- Gao, Z. & T. C. Schroeder(2009b), “Effects of label information on consumer willingness-to-pay for food attributes”, *American Journal of Agricultural Economics* 91(3):795—809.
- Gilovich, T. , D. Griffin & D. Kahneman(2002), *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge University Press.
- Hanemann, M. et al(1991), “Statistical efficiency of double-bounded dichotomous choice contingent valuation”, *American*

- Journal of Agricultural Economics* 73(4):1255–1263.
- Hensher, D. A. et al(2005), “The implications on willingness to pay of respondents ignoring specific attributes”, *Transportation* 32(3):203–222.
- Hensher, D. A. (2006), “Revealing differences in willingness to pay due to the dimensionality of stated choice designs: An initial assessment”, *Environmental and Resource Economics* 34(1):7–44.
- Hensher, D. A. (2007), “Attribute processing in choice experiments and implications on willingness to pay”, in: B. Kanninen(ed.), *Valuing Environmental Amenities Using Stated Choice Studies*, Springer.
- Hensher, D. A. et al(2007), “The implications on willingness to pay of a stochastic treatment of attribute processing in stated choice studies”, *Transportation Research Part E: Logistics and Transportation Review* 43(2):73–89.
- Hensher, D. A. (2010a), “Attribute processing, heuristics and preference construction in choice analysis”, *State-of Art and State-of Practice in Choice Modelling*, Emerald Press.
- Hensher, D. A. (2010b), “Hypothetical bias, choice experiments and willingness to pay”, *Transportation Research Part B: Methodological* 44(6):735–752.
- Hensher, D. A. et al(2012), “Inferring attribute non-attendance from stated choice data”, *Transportation* 39(2):235–245.
- Herriges, J. A. & J. F. Shogren(1996), “Starting point bias in dichotomous choice valuation with follow-up questioning”, *Journal of Environmental Economics and Management* 30(1):112–131.
- Hess, S. & D. A. Hensher(2010), “Using conditioning on observed choices to retrieve individual-specific attribute processing strategies”, *Transportation Research Part B: Methodological* 44(6):781–790.
- Hess, S. & D. A. Hensher(2013), “Making use of respondent reported processing information to understand attribute importance: A latent variable scaling approach”, *Transportation* 40(2):397–412.
- Huber, J. & K. Zwerina(1996), “The importance of utility balance in efficient choice design”, *Journal of Marketing research* 33(3):307–317.
- Hudson, D. et al(2005), “Hypothetical(non) bias in choice experiments: Evidence from freshwater prawns”, EconW-PA Working Paper.
- Islam, T. et al(2007), “Modeling the effects of including/excluding attributes in choice experiments on systematic and random components”, *International Journal of Research in Marketing* 24(4):289–300.
- Kahneman, D. et al(1991), “Anomalies: The endowment effect, loss aversion, and status quo bias”, *Journal of Economic Perspectives* 5(1):193–206.
- Kontoleon, A. & M. Yabe(2003), “Assessing the impacts of alternative ‘opt-out’ formats in choice experiment studies”, *Journal of Agricultural policy and Resources* 5(1):1–43.
- Kragt, M. E. (2013), “Stated and inferred attribute attendance models: A comparison with environmental choice experiments”, *Journal of Agricultural Economics* 64(3):719–736.
- Kuhfeld, W. F. et al(1994), “Efficient experimental design with marketing research applications”, *Journal of Marketing Research* 31(4):545–557.
- Lancaster, K. J. (1972), *Consumer Demand: A New Approach*, Columbia University Press.
- Lancsar, E. J. et al(2007), “Using discrete choice experiments to investigate subject preferences for preventive asthma medication”, *Respirology* 12(1):127–136.
- Lanz, B. & A. Provins(2012), “Do status quo choices reflect preferences? Evidence from a discrete choice experiment in the context of water utilities’ investment planning”, CEPE Center for Energy Policy and Economics, ETH Zurich.
- List, J. & C. Gallet(2001), “What experimental protocol influence disparities between actual and hypothetical stated values?”, *Environmental and Resource Economics* 20(3):241–254.
- Loomis, J. B. et al(2009), “A comparison of actual and hypothetical willingness to pay of parents and non-parents for protecting infants’ health: The case of nitrates in drinking water”, *Journal of Agricultural and Applied Economics* 41(3):697–712.
- Loomis, J. B. (2014), “Strategies for overcoming hypothetical bias in stated preference surveys”, *Journal of Agricultural and Resource Economics* 39(1):34–46.
- Louviere, J. J. et al(2000), *Stated Choice Methods: Analysis and Applications*, Cambridge University Press.
- Louviere, J. J. et al(2008), “Designing discrete choice experiments: Do optimal designs come at a price?”, *Journal of Consumer Research* 35(2):360–375.
- Louviere, J. J. et al(2010), “Discrete choice experiments are not conjoint analysis”, *Journal of Choice Modelling* 3(3):57–72.
- Lusk, J. L. & T. C. Schroeder(2004), “Are choice experiments incentive compatible? A test with quality differentiated beef steaks”, *American Journal of Agricultural Economics* 86(2):467–482.
- Lusk, J. L. & F. B. Norwood(2005), “Effect of experimental design on choice-based conjoint valuation estimates”, *American*

- Journal of Agricultural Economics* 87(3):771—785.
- Martinsson, P. et al(2001), “Using choice experiments for non-market valuation”, Göteborg University, Working Papers in Economics, No. 52.
- McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior, *Frontiers in Econometrics*, Academic Press.
- Meas, T. et al(2014), “Substitutes or complements? Consumer preference for local and organic food attributes”, *American Journal of Agricultural Economics*, doi: 10.1093/ajae/aau108.
- Meyerhoff, J. & U. Liebe(2009), “Status quo effect in choice experiments: Empirical evidence on attitudes and choice task complexity”, *Land Economics* 85(3):515—528.
- Meyerhoff, J. & K. Glenk(2015), “Learning how to choose: Effects of instructional choice sets in discrete choice experiments”, *Resource and Energy Economics* 41:122—142.
- Michaud, C. et al(2013), “Willingness to pay for environmental attributes of non-food agricultural products: A real choice experiment”, *European Review of Agricultural Economics* 40(2):313—329.
- Mintz, O. et al(2013), “Information processing pattern and propensity to buy: An investigation of online point-of-purchase behavior”, *Marketing Science* 32(5):716—732.
- Moser, R. et al(2013), “Testing hypothetical bias with a real choice experiment using respondents’ own money”, *European Review of Agricultural Economics* doi: 10.1093/erae/jbt016.
- Murphy, J. J. & T. H. Stevens(2004), “Contingent valuation, hypothetical bias, and experimental economics”, *Agricultural and Resource Economics Review* 33(2):182—192.
- Murphy, J. J. et al(2005), “A meta-analysis of hypothetical bias in stated preference valuation”, *Environmental and Resource Economics* 30(3):313—325.
- Pedersen, L. B. (2011), “Soccer attendees’ preferences for facilities at Fionia park stadium: An application of the discrete choice experiment”, *Journal of Sports Economics* 12(2):179—199.
- Plott, C. R. (1996), “Rational individual behavior in markets and social choice processes: The discovered preference hypothesis”, in: K. Arrow et al(eds), *The Rational Foundations of Economic Behavior*, MacMillan.
- Ritov, I. & J. Baron(1992), “Status-quo and omission biases”, *Journal of Risk and Uncertainty* 5(1):49—61.
- Rose, J. M. et al(2005), “Recovering costs through price and service differentiation: Accounting for exogenous information on attribute processing strategies in airline choice”, *Journal of Air Transport Management* 11(6):400—407.
- Sándor, Z. & M. Wedel(2001), “Designing conjoint choice experiments using managers’ prior beliefs”, *Journal of Marketing Research* 38(4):430—444.
- Scarpa, R. & J. M. Rose(2008), “Design efficiency for non-market valuation with choice modelling”, *Australian Journal of Agricultural and Resource Economics* 52(3):253—282.
- Scarpa, R. et al(2009), “Modelling attribute non-attendance in choice experiments for rural landscape valuation”, *European Review of Agricultural Economics* 36(2):151—174.
- Scarpa, R. et al(2013), “Inferred and stated attribute non-attendance in food choice experiments”, *American Journal of Agricultural Economics* 95(1):165—180.
- Scheufele, G. & J. Bennett(2012), “Response strategies and learning in discrete choice experiments”, *Environmental and Resource Economics* 52(3):435—453.
- Schlöpfer, F. & B. Fischhoff(2012), “Task familiarity and contextual cues predict hypothetical bias in a meta-analysis of stated preference studies”, *Ecological Economics* 81(0):44—47.
- Street, D. J. & L. Burgess(2007), *The Construction of Optimal Stated Choice Experiments: Theory and Methods*, John Wiley & Sons, Inc.
- Swait, J. & W. Adamowicz(2001), “Choice environment, market complexity, and consumer behavior”, *Organizational Behavior and Human Decision Processes* 86(2):141—167.
- Tversky, A. & D. Kahneman(1991), “Loss aversion in riskless choice: A reference-dependent model”, *Quarterly Journal of Economics* 106(4):1039—1061.
- Vermeulen, B. et al(2011), “Bayesian conjoint choice designs for measuring willingness to pay”, *Environmental and Resource Economics* 48(1):129—149.
- Zwerina, K. et al(1996), “A general method for constructing efficient choice designs”, Durham, NC: Fuqua School of Business, Duke University.

(责任编辑:刘新波)