

# 环境政策效益评估中的嵌入效应

——以北京市雾霾和沙尘治理政策为例

全世文, 黄波

**[摘要]** 嵌入效应会显著地影响经济学家对环境政策的成本效益分析,但目前国内鲜有研究在评估环境政策效益时对嵌入效应进行识别和讨论。本文首先在一个扩展的健康生产函数基础上分析了环境政策之间存在嵌入效应的原因,进而以北京市雾霾和沙尘治理政策为例,设计了三个选择实验组来验证嵌入效应。随后,本文采用随机参数 Logit 模型对各组样本进行了估计,污染物边际价值在组间的显著差异证实了嵌入效应的存在。如果忽略嵌入效应,降低雾霾和沙尘的平均边际价值分别会被高估 36.87%和 67.62%。在模型中控制调查对象的偏好异质性以后,高估程度有所下降,但嵌入效应仍不能完全消除。本文的研究结论意味着政策制定者和研究者在评估环境政策的效益之前,需要谨慎地处理嵌入效应,虽然不同的环境政策可能在技术目标上指向相互独立的不同污染物,但污染物之间的价值相关性仍然会导致嵌入效应,进而影响评估结果。

**[关键词]** 环境政策; 嵌入效应; 空气质量; 选择实验; 价值评估

**[中图分类号]**F124.5 **[文献标识码]**A **[文章编号]**1006-480X(2016)08-0023-17

## 一、引言

经济发展与环境质量之间的关系一直是学界和普通民众关心的一个重点话题。中国多年来的粗放型经济增长模式带来了严重的环境问题,与此同时,不断恶化的环境也逐渐成为了制约中国经济发展的一个紧约束条件。频频“爆表”的空气质量指数使大气污染问题成为了中国普通民众最为关切的环境问题<sup>[1]</sup>。近年来,北京市政府实施了一系列相应的政策用来改善城市空气质量,例如,限制工业排放、机动车尾号限行、鼓励新能源汽车、风沙源治理工程,等等。

经济学家评价环境政策是否可行的基本方法是成本效益分析。由于环境质量的公共物品特征及其价值构成的复杂性,评估环境政策的效益(也即评估环境质量的经济价值)成为了经济学家关注的主要问题。从数据来源看,价值评估方法可以分为两类:一是通过可观测的市场数据(即显示偏好数据)来评价空气质量价值,包括疾病成本法<sup>[2,3]</sup>、享乐价格法<sup>[4,5]</sup>,等等。显示偏好方法的优势在于外部效率较高,但是却无法评估空气质量价值中的非市场价值部分,因此,其评估结果往往只能作

**[收稿日期]** 2016-05-20

**[基金项目]** 中央高校基本科研业务费专项资金资助项目“基于选择实验的北京市雾霾治理公共政策研究”(批准号 16XNB023)。

**[作者简介]** 全世文(1987—),男,河南南阳人,中国社会科学院农村发展研究所助理研究员;黄波(1981—),男,吉林蛟河人,中国人民大学农业与农村发展学院讲师。通讯作者:全世文,电子邮箱:quanshiwen@163.com。

为空气质量某个价值构成的下限<sup>[6]</sup>。二是通过调查对象自我陈述的数据(即陈述偏好数据)来评价空气质量价值,包括条件价值评估方法<sup>[7,8]</sup>、选择实验,等等。陈述偏好方法可以对空气质量的综合价值或其中的某个特定的价值构成进行评估。但是,正如 Louviere et al.<sup>[9]</sup>和全世文<sup>[10]</sup>指出,陈述偏好方法在应用过程中的核心问题是如何保证评估结果的效度。事实上,由于数据的假想特征,学界对陈述偏好方法的结果在多大程度上可以反映被调查对象的真实偏好一直持谨慎的怀疑态度。随着应用研究的逐渐增加,经济学家发现并证实了陈述偏好方法从问题设计、实验实施到模型估计的过程中存在一系列的偏差。本文集中关注其中的一个主要偏差,即“嵌入效应”(Embedding Effect)。

嵌入效应由 Kahneman and Knetsch<sup>[11]</sup>在条件价值评估方法中发现,他们指出当某个公共物品被作为单一评估对象和作为复合评估对象中的一个构成要素时,价值评估的结果会出现非常显著的差异;也就是说,研究者将公共物品嵌入实验情景的方式会影响到评估结果。Kahneman and Knetsch<sup>[11]</sup>认为嵌入效应来源于“道德满足”,调查对象更像是在陈述一种“伦理偏好”而非“经济偏好”;即调查对象会根据自身对评估对象的关切度和兴趣做出主观陈述,而不会完全理性地关注评估对象发生边际变化的绝对值。后续研究在不同的应用领域广泛证实了条件价值评估方法中存在嵌入效应<sup>[12,13]</sup>。嵌入效应的存在显然会弱化价值评估结果的应用价值,被 Hausman<sup>[14]</sup>认为是条件价值评估方法的三个主要偏差<sup>①</sup>中最难处理的一个。

嵌入效应在环境政策的效益评估中具有一定的特殊性。特定环境政策的技术目标通常仅指向一种或有限的几种大气污染物。例如,尾号限行和推广新能源汽车都是为了减少尾气排放量,从而降低 PM<sub>2.5</sub>、一氧化碳和氮氧化物等,而对沙尘的缓解作用非常有限;相反,城市绿地建设和植树造林可以有效地缓解沙尘天气,而对降低 PM<sub>2.5</sub> 和臭氧污染则没有效果。因此,对某项环境政策的效益评价需要对降低特定污染物的价值进行评估。但是,空气质量通常是作为一个整体给个人带来效用。也就是说,特定污染物在空气质量的价值评估中是一种典型的嵌入结构,而嵌入方式很有可能会对评估结果造成影响。

事实上,早期的一些研究从评估对象的价值之间存在替代性的角度出发,认为所谓的嵌入效应并不违背标准的经济理论,因此,并不是一个偏差<sup>[15,16]</sup>。Hanemann<sup>[17]</sup>将嵌入效应归因于“次可加性”,即边际替代率递减规律会导致多个公共产品的边际变化之和高于复合公共产品的一个等量边际变化。在这种解释下,嵌入效应并不影响条件价值评估方法本身的效度,研究者只需要通过谨慎的实验设计避免情景设定存在的偏差,同时结合实际意义选择一个将评估对象嵌入实验情境的合理方式即可。但是,后续的实证研究并没有完全支持这种解释<sup>[18,19]</sup>。近年来的研究尝试从认知心理学的角度解释嵌入效应。除了“道德满足”的解释以外,Mørkbak et al.<sup>[20]</sup>认为有限认知能力和心智账户都可能导致嵌入效应。

综上所述,本文以雾霾和沙尘的治理政策为例,采用分组的选择实验方法来验证并分析环境政策效益评估中的嵌入效应。从政策意义上讲,识别嵌入效应可以更加准确地评估某项环境政策的效益,而且有助于把握多项环境政策之间的关系,从而为多元环境政策的设计提供科学依据。从学术意义上讲,本文存在以下两个创新之处:①现有研究对嵌入效应的经济学解释是被评估的项目或政策本身存在替代关系,也就是说,多个项目或政策的目标存在重合。这与不同大气环境治理政策的技术目标相对单一的情况并不吻合。本文试图为这种情况下的嵌入效应增加新的解释。②现有研究

① 这三个主要的偏差:一是假想偏差,即调查对象在虚拟的决策情景中倾向于高估其支付意愿;二是支付意愿和受偿意愿的分离;三是本文关心的嵌入效应。

对嵌入效应的讨论集中在条件价值评估方法上,而鲜有研究在选择实验方法中验证嵌入效应。目前主要有两篇文献,Mørkbak et al.<sup>[20]</sup>和 Jacobsen et al.<sup>[21]</sup>在选择实验中讨论嵌入效应,且两者得到了完全相反的结论。本文可以为选择实验中是否存在嵌入效应增添新的经验证据。

## 二、理论分析

健康生产函数由 Grossman<sup>[22]</sup>提出,用来解释人们的健康需求和健康投资行为。Harrington and Portney<sup>[23]</sup>扩展了健康生产函数,通过引入环境污染变量来分析人们对降低污染的支付意愿和采取防护措施(包括医疗活动)之间的关系。本文在此基础上做出进一步的扩展,即通过引入多个污染变量来分析环境治理政策在效益评估过程中存在的嵌入效应。

### 1. 污染降低的边际价值

考虑一个如式(1)所示的效用函数形式:

$$u=u(X,f,s,c_1,c_2) \quad (1)$$

在式(1)中,个体通过对货币化物品  $X$ (将其价格标准化为 1)、休闲  $f$  的消费获取效用; $\partial u/\partial X > 0$ ,  $\partial u/\partial f > 0$ 。与此同时,生病时间  $s$  和大气污染水平  $c$  会对效用造成负影响; $\partial u/\partial s < 0$ ,  $\partial u/\partial c < 0$ 。这里,大气污染(简单起见,假设仅有两种大气污染物  $c_1$  和  $c_2$ )通过视觉和知觉舒适性的下降对效用产生直接影响。除了直接影响以外,大气污染还会通过两个间接渠道对效用产生影响:

$$f=f(c_1,c_2) \quad (2)$$

$$s=s(c_1,c_2,a_1,a_2) \quad (3)$$

式(2)表示大气污染会影响休闲时间; $\partial f/\partial c < 0$ 。式(3)表示大气污染水平会对个人健康造成影响; $\partial s/\partial c > 0$ ,  $a_1$  和  $a_2$  分别表示个人为应对两种污染物所采取的防护措施; $\partial s/\partial a < 0$ ;这些防护措施包括医疗活动、购买空气净化器、购买口罩、安装过滤窗和搬迁,等等。个人在预算约束(4)下通过选择  $X$ ,  $f$ ,  $a_1$  和  $a_2$  实现式(1)的效用最大化。

$$I+p_w(T-f-s(c_1,c_2,a_1,a_2))=X+p_1 \cdot a_1+p_2 \cdot a_2+p_f \cdot f \quad (4)$$

在式(4)中, $I$ 表示非劳动收入, $p_w$ 为工资率, $T$ 为所有可利用的时间, $p_1$ 和  $p_2$ 分别是防护措施  $a_1$  和  $a_2$  的价格, $p_f$ 为休闲活动的价格。于是,参照 Harrington and Portney<sup>[23]</sup>可以求得:

$$w_1=p_w \cdot \frac{ds}{dc_1} - \frac{\partial u/\partial s}{\lambda} \cdot \frac{ds}{dc_1} + p_1 \cdot \frac{\partial a_1^*}{\partial c_1} + p_2 \cdot \frac{\partial a_2^*}{\partial c_1} - p_f \cdot \frac{df}{dc_1} - \frac{\partial u/\partial c_1}{\lambda} \quad (5)$$

在式(5)中, $a_1^*$ 和  $a_2^*$ 表示两种防御措施的均衡需求。式(5)说明降低大气污染  $c_1$  的边际价值由四个部分构成:①可观测的生病天数减少所获得的工资收入和由生病天数减少直接提高效用的货币等价;②可观测的防护措施(包括医疗活动)支出的减少;③可观测的由休闲活动增加所产生的价值;④通过视觉和知觉舒适性的改善引起效用上升的货币等价。同理,降低大气污染  $c_2$  的边际价值为:

$$w_2=p_w \cdot \frac{ds}{dc_2} - \frac{\partial u/\partial s}{\lambda} \cdot \frac{ds}{dc_2} + p_1 \cdot \frac{\partial a_1^*}{\partial c_2} + p_2 \cdot \frac{\partial a_2^*}{\partial c_2} - p_f \cdot \frac{df}{dc_2} - \frac{\partial u/\partial c_2}{\lambda} \quad (6)$$

### 2. 政策效益评估中的嵌入机制

考虑两项环境政策  $q_1$  和  $q_2$ ,目标分别是降低两种污染物水平  $c_1$  和  $c_2$ ,即  $c_1=c_1(q_1)$ ,  $c_2=c_2(q_2)$ ,

且 $\partial c/\partial q < 0$ <sup>①</sup>。那么,间接效用函数为 $v = v(I, p_w, p_1, p_2, p_f, c_1(q_1), c_2(q_2))$ 。环境政策 $q_1$ 的边际效益为:

$$w_{q_1} = \frac{dI}{dq_1} = -\frac{\partial v/\partial q_1}{\partial v/\partial I} = -\frac{\partial v/\partial c_1}{\lambda} \cdot \frac{\partial c_1}{\partial q_1} = w_1 \cdot \frac{\partial c_1}{\partial q_1} \quad (7)$$

同理,环境政策 $q_2$ 的边际效益为:

$$w_{q_2} = w_2 \cdot \frac{\partial c_2}{\partial q_2} \quad (8)$$

给定函数 $c(q)$ ,在式(7)和式(8)中讨论环境政策间的嵌入效应等价于在式(5)和式(6)中讨论环境边际价值之间的嵌入效应<sup>②</sup>。用 $c$ 表示两种污染物的综合水平,与式(5)和式(6)同理,降低污染综合水平的边际价值为:

$$w_{1+2} = p_w \cdot \frac{ds}{dc} - \frac{\partial u/\partial s}{\lambda} \cdot \frac{ds}{dc} + p_1 \cdot \frac{\partial a_1^*}{\partial c} + p_2 \cdot \frac{\partial a_2^*}{\partial c} - p_f \cdot \frac{df}{dc} - \frac{\partial u/\partial c}{\lambda} \quad (9)$$

根据可加性原则,嵌入效应可以被定义为 $w_{1+2} \neq w_1 + w_2$ <sup>[21]</sup>。下面,本文从边际价值的各项构成中说明嵌入效应产生的原因。①考虑大气污染通过影响生病时间产生的价值: $p_w \cdot (ds/dc) - (ds/dc) \cdot (\partial u/\partial s)/\lambda$ 。对函数 $s(c_1, c_2, a_1, a_2)$ 中 $c_1$ 和 $c_2$ 关系的判断需要依赖医学方面的专业知识。一种直观的理解是两种污染物在致病时间上具有重合效果,也就是说,污染物数量对致病时间具有边际递减的影响效果,从而有 $ds/dc < ds/dc_1 + ds/dc_2$ 。②考虑大气污染通过防护措施支出产生的价值: $p_1 \cdot (\partial a_1^*/\partial c_2) + p_2 \cdot (\partial a_2^*/\partial c_1)$ 。假设针对每种污染物的防护措施都不存在交叉影响: $\partial a_1^*/\partial c_2 = \partial a_2^*/\partial c_1 = 0$ ,即个体可以明确地区分每种防护措施的价值,那么,防护支出就不会引起嵌入效应。但这一假设显然与现实情况并不一致。例如,个体为减少吸入PM<sub>2.5</sub>而购买的口罩也可以同时起到屏蔽沙尘的作用,而个体为了躲避大气污染举家搬迁的措施则可以同时减少暴露于各类污染物的概率。由于个体无法清晰地将防护措施针对每种污染物的价值进行分离,即个体在实验情景中通常会将某项防护措施的价值对不同污染物进行“重复记账”,这就会导致嵌入效应的发生,即 $\partial a_i^*/\partial c < \partial a_i^*/\partial c_1 + \partial a_i^*/\partial c_2, \forall i=1,2$ 。③考虑大气污染通过影响休闲活动和个体舒适性产生的价值: $-p_f \cdot df/dc - \partial u/\partial c$ 。在这两项价值构成中,大气污染物 $c_1$ 和 $c_2$ 具有互补效果。也就是说,仅当两种污染物的水平都降低时,个体才会增加休闲时间,个体视觉和知觉的舒适性才可以得到提高。这就意味着 $-df/dc > -df/dc_1 - df/dc_2$ 且 $-\partial u/\partial c > -\partial u/\partial c_1 - \partial u/\partial c_2$ 。

通过上述分析可知,即使多项环境政策分别以不同的污染物为治理目标,环境政策的效益仍然会存在嵌入效应。嵌入效应并不是直接来自于环境政策的技术效果之间的替代关系,而是来自于污染物价值之间的相关性。这种相关性表现在:不同污染物在致病时间上具有重合效果、个体无法明确分离针对每种污染物的防护支出、不同污染物对休闲和个体舒适性具有互补效果。其中,前两种关系说明污染物价值之间具有替代关系,因此,对空气质量整体的评估价值会低于对单个污染物的评估价值之和。相反,互补关系则会导致相反的嵌入效应,即综合价值会高于单项价值之和。因此,

① 函数被称为“干预函数”,反映了干预政策或措施在技术层面上对污染物水平的影响。当然,在环境经济学的研究中,干预函数本身就是一个研究目标,它是评价环境政策效益的基础。例如,曹静等<sup>[24]</sup>讨论了北京市尾号限行政策对改善空气质量的影响,席鹏辉和梁若冰<sup>[25]</sup>讨论了油价变动对空气污染的影响。由于干预函数并非本文的研究目标,本文假设这一函数所反映的技术效果是提前给定的。

② 另一种情况是,多项环境政策的目标或效果存在重叠,即 $\partial c_2/\partial q_1 < 0, \partial c_1/\partial q_2 < 0$ 。在这种情况下,两项政策本身的替代关系就会直接导致嵌入效应。



对单项环境政策的效益进行评估时,如忽略嵌入效应,很有可能会高估或低估政策效益,导致评估结果的效度下降。对大气污染的治理政策而言,效益被高估和低估的可能性同时存在。

### 三、实验设计

根据国家环保部发布的《环境空气质量标准》(GB 3095—2012),空气污染物的基本项目包括六类:二氧化硫(SO<sub>2</sub>)、二氧化氮(NO<sub>2</sub>)、一氧化碳(CO)、臭氧(O<sub>3</sub>)、可吸入颗粒物(PM<sub>10</sub>)和细颗粒物(PM<sub>2.5</sub>)。出于调查对象对这些污染物的认知有限以及实验设计过于复杂的考虑,本文仅选择两种污染物:PM<sub>2.5</sub>和PM<sub>10</sub>,用来设计选择实验。其中,PM<sub>2.5</sub>是导致雾霾天气的主要成分,而PM<sub>10</sub>是导致沙尘天气的主要成分。本文采用一个分组实验设计方案:第一实验组仅评价雾霾治理政策,第二实验组仅评价沙尘治理政策,第三实验组同时评价雾霾和沙尘治理政策,见表1。通过组间比较可以用来分析多元环境政策中的嵌入效应。本文选择“征税”作为价值评估的“支付工具”(Payment Vehicle)。于是,前两组选择实验基于两个属性进行设计,第三组实验基于三个属性进行设计。

现有研究在评估空气质量价值时对污染程度的定义有两种方法:以浓度为单位<sup>[26-28]</sup>和以天数为单位<sup>[29,30]</sup>。在陈述偏好方法中,不同的调查对象对污染物浓度的认知可能会存在显著的差异,而认知的不一致显然会弱化评估结果的效度。因此,本文以污染天气的发生天数为单位进行实验设计。为了使调查对象对污染程度有一致的认知,这里设定了两个“污染标准”,分别用来反映中度雾霾天气和重度沙尘天气,并在选择情境中提供了对应的图片<sup>①</sup>。

本文统计了北京市环境保护检测中心在2014年发布的PM<sub>2.5</sub>和PM<sub>10</sub>浓度数据,在359个观测日中,PM<sub>2.5</sub>超标天数为174天,PM<sub>10</sub>超标天数为99天<sup>②</sup>。为了便于认知,这里将这两个数据分别约等为180天和100天,并以此作为设计环境政策效果的现状基准。设定政策效果在每种污染物上具有3个属性水平:分别将污染天数降低25%、50%和75%。同时,参考同类研究的价值估算结果<sup>[8,26]</sup>,设定基准税额为400元/年,并分别通过增加50%和降低50%获得两个税额标准:200元/年和600元/年。于是,前两个实验组是一个2属性3水平的设计方案,第三组实验是一个3属性3水平的设计方案,见表1。

表1 选择实验的分组实验设计方案

	第一实验组	第二实验组	第三实验组
雾霾天气(3个水平)	√		√
沙尘天气(3个水平)		√	√
征税额度(3个水平)	√	√	√
属性数量	2	2	3
全因子设计的选项数量	3×3=9	3×3=9	3×3×3=27
设计方案及选项数量	全因子设计,9	全因子设计,9	正交设计,9
选择情境数量	4	4	4

资料来源:作者整理。

① 在正式进行选择以前,本文还设计了一个关于北京市大气污染状况的描述说明,在此过程中,向调查对象展示了空气质量良好和空气污染的对比图。

② 根据国家环保部颁布的《环境空气质量指数(AQI)技术规定(试行)》(HJ 633—2012),AQI轻度污染对应的PM<sub>2.5</sub>浓度为75ug/m<sup>3</sup>,对应的PM<sub>10</sub>浓度为150ug/m<sup>3</sup>。

对前两个实验组,本文首先根据全因子设计方案,提取了9个由空气质量和上缴税额组合而成的选项。对第三实验组,这里根据正交设计方案同样提取了9个选项。因此,本文在9个选项中根据属性水平平衡、效用平衡和最小重叠三个原则在三个实验组中都配对了4个选择情景。根据Carson and Groves<sup>[31]</sup>,陈述偏好方法中仅有二分选择题项可以保证激励相容的设计要求。于是,所有选择情景都包括两个政策效果:效果A和效果B,此外,选择情景加入了一个“维持现状”(Status Quo)的选项用来避免调查对象对两种政策效果都不满意的情形。于是,调查对象需要在4个连续的选择情景中进行决策,为了避免“排序效应”(Ordering Effect),本文将每个实验组中的4个选择情景均进行了随机排序。在正式进入选择情景前,问卷还向调查对象提供了一段文字,用来说明导致雾霾天气和沙尘天气的不同原因以及对应的不同治理政策。其中,治理雾霾天气的政策主要包括调整能源结构、减少工业排放和降低机动车尾气排放;治理沙尘天气的政策则主要包括植树种草、退耕还林、绿地建设和风沙源防控。然后,问卷提供了一个选择情景示例,并向调查对象说明,效果A和效果B均为两类污染治理政策的综合效果。

#### 四、数据与计量模型

##### 1. 调查实施与样本描述

本文分析所用的数据来源于课题组在2015年11月对北京市居民开展的问卷调查。问卷共包括三个部分:调查对象的个人特征、调查对象对大气污染的认识以及上一节设计的4个选择情景。课题组采用网络调查方法,委托某大型网络调查公司向其北京市样本库投放调查问卷,在为期一周的时间内共收集了有效样本690份,三个实验组的样本分配基本均匀。为了保证调查对象对北京市大气污染状况有良好的认知,课题组限制调查对象均为在北京市常住时间超过3年的样本。表2分别报告了三个实验组中样本的统计特征。

检验结果显示,三个实验组中的样本在各指标上的统计特征均没有显著差异,样本的随机分布为后文的组间对比提供了保障。由表2可知,与北京市城镇居民的基本统计指标相比,调查样本存在女性比例偏高、年龄偏低、受教育水平和收入水平偏高的特点。从这个角度看,本文的调查样本对北京市城乡居民的代表性存在一定的欠缺。事实上,如何选择调查方法和样本框是研究者一直关注的话题。虽然网络调查方式面临着样本有偏的风险,但随着近年来移动互联网技术的快速发展,这种偏差正在大幅下降。与传统的面对面调查方法相比,网络调查可以排除因调查员主观引导造成的偏差,对样本的随机性也更容易控制<sup>[32,33]</sup>。近年来,在以城市居民作为调查对象时,网络调查方法正在被学界广泛采用。与评估环境价值的研究目标相比,本文更关注评估价值在不同实验组的差异。因此,本文对样本代表性的要求相对弱于对组间样本分配随机性的要求,这是本文宜于采用网络调查方法的另一个原因。在表2中,三个实验组中的样本在统计特征上的高度相似性说明调查样本可以较好地服务于研究目标。

##### 2. 计量模型

将间接效用函数 $v(\cdot)$ 对税收和污染 $c$ (包括雾霾发生天数 $haze$ 和沙尘发生天数 $sand$ )写为一个线性函数的形式,那么,三个实验组对应的间接效用函数分别为:

$$v^1 = \beta_0^1 + \beta_1^1 \cdot haze + \beta_3^1 \cdot tax + \varepsilon^1 \quad (10)$$

$$v^2 = \beta_0^2 + \beta_2^2 \cdot sand + \beta_3^2 \cdot tax + \varepsilon^2 \quad (11)$$

$$v^3 = \beta_0^3 + \beta_1^3 \cdot haze + \beta_2^3 \cdot sand + \beta_3^3 \cdot tax + \varepsilon^3 \quad (12)$$

表2 三个实验组中样本的描述性统计

变量	含义	第一实验组	第二实验组	第三实验组	北京市
gender	样本量	231	227	232	
	性别(男性=0,女性=1)	0.589 (0.493)	0.533 (0.500)	0.565 (0.497)	0.486 (0.500)
age	年龄	31.009 (8.986)	30.890 (9.104)	31.418 (9.433)	40.506 (13.465)
	受教育年限	16.325 (1.928)	15.965 (2.037)	15.897 (2.355)	12.100 (3.744)
inc	月收入水平(千元)	7.629 (5.583)	6.930 (4.874)	7.079 (5.011)	3.659
	家庭是否有儿童(是=1,否=0)	0.602 (0.491)	0.599 (0.491)	0.616 (0.487)	
illness	是否有过呼吸道疾病(是=1,否=0)	0.654 (0.477)	0.573 (0.496)	0.651 (0.478)	
	是否有环保专业知识背景(是=1,否=0)	0.225 (0.419)	0.207 (0.406)	0.203 (0.403)	
aqi	是否了解“空气质量指数(AQI)”(是=1,否=0)	0.641 (0.481)	0.652 (0.477)	0.720 (0.450)	
	对雾霾和沙尘的知识水平(取值范围:1—8)	3.710 (1.379)	3.749 (1.281)	3.448 (1.261)	

注:①报告的统计量为样本均值,括号内为对应的标准差;②所有变量任意两组均值T检验的结果都无法在5%的统计水平上拒绝均值相等的原假设;③北京市数据来源于北京市统计局发布的2014年度数据,其中,年龄数据根据20—69岁的分段频率数据计算,受教育年限根据受教育程度的分类数据折算,人均月收入根据人均年度可支配收入(43910元)计算。  
资料来源:作者整理。

根据第三节的实验设计,在(10)—(12)式中,  $tax$  表示个体需要额外上缴的用于实施专项环境治理政策的税额(即收入的减少);  $\beta_3$  表示税收的边际效用,且  $\beta_3 < 0$ ;  $\beta_1$  和  $\beta_2$  表示空气污染的边际效用,且  $\beta_1 < 0, \beta_2 < 0$ 。那么,在三个实验组中可以计算得到污染降低的四个边际价值:  $w_1^1 = -\beta_1^1 / \beta_3^1$ ,  $w_2^2 = -\beta_2^2 / \beta_3^2$ ,  $w_1^3 = -\beta_1^3 / \beta_3^3$  和  $w_2^3 = -\beta_2^3 / \beta_3^3$ 。

然后,本文采用随机参数 Logit 模型对各组样本中的参数进行估计。定义向量  $\beta$  服从密度函数为  $g(\beta|\theta)$  的随机分布,定义自变量向量  $x = (tax, haze, sand)$ 。参照 Train<sup>[34]</sup>,对任意一组的某个调查对象  $i$  而言,在4次选择情景的选择空间  $R$  中做出选择集  $R_i$  的概率为:

$$P(R_i | x_i, \theta) = \int \prod_{t=1}^4 \frac{\exp(\beta'_i x_{i,i(t),t})}{\sum_{j=1}^3 \exp(\beta'_i x_{i,j,t})} \cdot g(\beta|\theta) d\beta, \quad i(t,r) = 1, 2, 3; \quad t = 1, 2, 3, 4 \quad (13)$$

其中,  $i(t,r)$  表示  $i$  个体在第  $t$  个选择情景中做出的选择。通过式右侧的积分运算,式(13)中的  $P(\cdot)$  可以表达为参数  $\theta$  的函数。通过拟最大似然估计可以得到参数  $\theta$  的估计量  $\hat{\theta}$ ,进而可以对边际价值  $w$  做出估算。在估计过程中,本文参考同类研究将参数  $\beta_3$  约束为固定参数,将  $\beta_1$  和  $\beta_2$  设定为服从正态分布的随机参数,从而得到一个服从正态分布的边际价值  $w$ <sup>[35,36]</sup>。进而,本文通过以下三组检验对降低雾霾和沙尘边际价值之间的嵌入效应进行验证:  $H_1^0 : w_1^1 = w_1^3, H_1^1 : w_1^1 \neq w_1^3; H_2^0 : w_2^2 = w_2^3, H_2^1 : w_2^2 \neq w_2^3; H_3^0 : w_1^1 + w_2^2 = w_1^3 + w_2^3, H_3^1 : w_1^1 + w_2^2 \neq w_1^3 + w_2^3$ 。

为了进行上述检验<sup>①</sup>,本文首先根据 Krinsky and Robb<sup>[37]</sup>提出的限参 Bootstrap 方法在随机参数 Logit 模型估计结果的基础上模拟了 2000 个样本的偏好参数,并据此计算出 2000 个边际价值。然后,本文同时采用组间 T 检验法和 Poe et al.<sup>[38]</sup>提出的“非参数完全组方法”(后文称“Poe 检验”) 对边际价值进行组间均值检验。

## 五、降低空气污染的经济价值及嵌入效应检验

### 1. 污染物降低的边际价值及其组间对比

首先采用上文设定的随机参数 Logit 模型对三个实验组的样本分别进行估计,结果如表 3 所示。似然比检验的结果显示,在第三实验组中,随机参数  $\beta_1$  和  $\beta_2$  之间并没有显著的相关性,因此,第三实验组仅报告了独立参数模型的估计结果。从总体上看,三个实验组的拟合效果良好,每个自变量均在 1% 的统计水平上高度显著。所有参数的符号均与预期相吻合; $tax$  的系数为负,说明税收会降低个人效用, $haze$  和  $sand$  的均值系数为负,说明空气污染也会产生负效用。

高度显著的标准差系数验证了个体对空气质量的偏好具有异质性。根据随机参数的估计值可以估算样本中“厌恶”空气污染的样本比例,例如,在第一实验组和第二实验组的样本中,分别有 80.38% 和 69.19% 的样本对减少雾霾天气和减少沙尘天气具有正向的边际支付意愿。这一结果与曾贤刚等<sup>[29]</sup>对北京市居民的调查结果基本吻合。个体不愿意为降低空气污染进行支付的一个主要原因是对产权界定的质疑。部分调查对象认为自己拥有享受清洁空气的初始产权,这就意味着政府或排污企业需要为空气污染承担责任,而居民应当因空气被污染而接受补偿。

在表 3 报告的估计结果基础上,本文采用限参 Bootstrap 方法对每组样本都模拟了 2000 个边际价值,见表 4。表 4 的统计结果可以被描述为:当单独评价雾霾天气和沙尘天气时,北京市居民对减少一个标准的中度雾霾天和中度沙尘天的平均边际支付意愿分别为 8.680 元/天和 4.199 元/天;而当同时评价两种污染时,平均边际支付意愿分别降低为 6.342 元/天和 2.505 元/天。根据第三实验组的统计结果,如果在 2014 年的基础上将年度雾霾和沙尘发生天数分别降低 50%,个体的平均支付意愿分别为 570.78 元和 125.25 元。该分析结果与部分相关研究的结论基本吻合<sup>[7,8]</sup>。

对比可知,无论是单独对每种污染物进行评价还是同时评价两种污染物,减少雾霾天气的边际价值都显著地高于沙尘天气,这与近年来在直观感受和舆论宣传下公众关于雾霾天气对公共健康的损害较沙尘天气更严重的普遍意识相吻合。这一结论同时也说明本文在调查中不存在“范围不敏感”(Scope Insensitivity)<sup>②</sup>,因此,可以拒绝范围不敏感是导致组间差异的原因。根据表 4 报告的检验结果,边际价值在组间相等的三个原假设均被拒绝,这说明嵌入方式可以显著地影响对降低空气污染的价值评估。导致边际价值在组间存在显著差异的另一个可能的原因是不同组在实验设计中的信息负荷存在差异,即第一实验组和第二实验组的信息负荷低于第三实验组。现有研究的一个共识

① 在第三个检验  $H_3^0$  中,由于第一实验和第二实验组中的样本并未一一匹配,所以  $w_1^1$  和  $w_2^2$  无法直接相加。但后文在进行该检验时是基于模拟的边际价值数据,由于  $w_1^1$  和  $w_2^2$  均服从正态分布,所以,随机相加的  $w_1^1 + w_2^2$  也服从正态分布。而  $w_1^3$  和  $w_2^3$  却可能因偏好相关导致  $w_1^3 + w_2^3$  不服从正态分布,但后文的估计并未验证出个体对雾霾和沙尘的偏好具有显著的相关性。

② 范围不敏感是指调查对象不关注评估对象被定义的“范围”(或“度量单位”),仅关注评估对象的相对优劣,这会导致评估结果的效度大幅下降<sup>[14]</sup>。本文在实验设计中对雾霾和沙尘的区别仅在发生天数上存在差异(见表 3),那么,如果本文的结论支持范围不敏感的说法,应该有  $w_2/w_1 \approx 1.8$ ;而估计结果显示  $w_1$  显著高于  $w_2$ ,据此可以认为调查过程中不存在范围不敏感。



表 3 三个实验组的随机参数 Logit 模型估计结果

	第一实验组		第二实验组		第三实验组	
	系数	标准误	系数	标准误	系数	标准误
均值						
tax	-0.401**	0.068	-0.332**	0.050	-0.479**	0.047
haze	-3.399**	0.431			-3.009**	0.363
sand			-1.412**	0.418	-1.200**	0.388
标准差						
haze	3.974**	0.352			4.164**	0.439
sand			2.817**	0.266	2.857**	0.498
观测值	2772		2724		2784	
Chi2	512.3		128.8		385.1	
P 值	0.000		0.000		0.000	
McFadden' R <sup>2</sup>	0.268		0.067		0.222	

注:①观测值=样本量×12,其中,12表示4个选择情景中的选项总数(每个选择情景有3个选项);②\*\*和\*分别表示在1%和5%的统计水平上显著;③为了增加估计系数的量级,tax的单位被调整为“百元”,haze和sand的单位被调整为“百天”。资料来源:作者整理。

表 4 模拟边际价值的统计结果及其组间均值检验

2000 个模拟边际价值的均值及标准差					
$w_1^1$	$w_2^2$	$w_1^3$	$w_2^3$	$w_1^3+w_2^3$	$w_1^1+w_2^2$
8.680**	4.199**	6.342**	2.505**	8.847**	12.879**
(1.201)	(0.845)	(0.767)	(0.757)	(0.984)	(1.609)
组间均值检验					
原假设	T 检验 P 值		Poe 检验 P 值		
$w_1^1=w_1^3$	0.000		0.035		
$w_2^2=w_2^3$	0.000		0.070		
$w_1^1+w_2^2=w_1^3+w_2^3$	0.000		0.014		

注:①T检验为非配对T检验方法;②边际价值的单位为“元/天”;③\*\*表示在1%的统计水平上显著不为零;④T检验的备择假设是“等式左侧≠等式右侧”,P值为双尾P值;Poe检验的备择假设是“等式左侧>等式右侧”,P值为单尾P值。资料来源:作者整理。

是调查对象的认知能力有限,因而会在高信息负荷下根据启发式采取信息处理策略,从而影响分析结果<sup>[10,39]</sup>。但是,本文的估计结果基本上可以排除信息负荷的影响。原因在于:一是本文实验设计的信息负荷相对较低(见第三节),由此降低了调查对象采取信息处理策略的动机。二是从表3的估计结果可以看出,三个实验组的系数估计值处于同一个数量级,而且不存在显著的差异,这意味着三个实验组的调查对象在不同的选择情境中面临着程度相近的不确定性。

因此,降低空气污染的边际价值在组间存在的差异可以被归因于嵌入方式的差异,这也就验证了嵌入效应的存在。由于在进入实验情景前,问卷向调查对象强化了雾霾和沙尘治理政策相互独立的信息,因此,可以排除嵌入效应是因为调查对象认为两种治理政策之间具有重合效果。据此可以认为嵌入效应来源于降低两种污染物的边际价值之间存在关联。同时评价两种污染得到的边际价

值低于单独评价的结果,说明如果忽略了嵌入效应,环境政策的效益将会被高估,这与目前大多数验证嵌入效应的研究所得到的结论一致<sup>[18]</sup>。结合本文的理论分析,这一结果意味着两种污染物的价值重合大于价值互补。但是,嵌入效应对两种环境治理政策效益评价的影响程度则存在差异。以第三实验组的估计结果为基础,忽略嵌入效应将导致雾霾天气减少的边际价值被高估 36.87%,而沙尘天气减少的边际价值则被高估了 67.62%,后者接近前者的两倍。造成这一差异的主要原因在于降低雾霾较降低沙尘的边际价值更高。因此,价值重合部分对雾霾价值评估的相对影响更小。

## 2. 控制偏好异质性后的边际价值

为了进一步把握调查对象的偏好异质性,本文在(10)—(12)式中加入调查对象个体特征与空气污染的交叉项后重新估计模型,结果如表 5 所示。与表 3 对比可知,加入了交叉项以后,各组样本的  $R^2$  均得到了提高,说明调查对象的个体特征可以在一定程度上解释对空气污染的偏好异质性。

在各模型中,统计显著的交叉项系数均为负,说明空气污染给个体带来的负效用随着受教育水平的提高、收入的提高、对空气质量指数了解程度的提高和对雾霾与沙尘辨识度的提高而加强。这些结果与现有研究的基本结论相一致。例如,蔡春光和郑晓瑛<sup>[9]</sup>同样证实了受教育水平和收入更高的调查对象对降低空气污染的支付意愿更高。陈永伟和史宇鹏<sup>[27]</sup>得到了同样的结论,并据此认为向居民征收的环境治理专项税应设计为累进税制。

根据表 5 报告的估计结果,本文同样模拟了 2000 个偏好参数,并在一个“典型居民”的基础上计算出 2000 个边际价值,其统计情况如表 6 所示。表 6 报告的估计结果与表 4 基本一致,但第三实验组估算得到的边际价值有所提高:雾霾由 6.342 元/天提高到 7.567 元/天,沙尘由 2.505 元/天提高到 3.193 元/天。但提高后的边际价值仍然低于单独评价每种污染时的结果,说明两种污染物的价值重合具有稳健性。虽然如此,表 6 报告的边际价值的组间差异较表 4 仍然有明显下降。雾霾价值在第一实验组中被高估的程度下降到 8.26%,沙尘价值在第二实验组中被高估的程度下降到 31.94%。现有研究验证嵌入效应的基本方法是通过设计分组的实验来检验组间差异,而本文的这一结果说明由组间差异测度的嵌入效应在一定程度上可以通过控制偏好异质性而得到缓解。根据前文的理论分析,在价值重合和价值互补的共同作用下,调查对象在不同嵌入结构的选择情景中会有不同的偏好结构,而这种组间的偏好差异如果能够通过控制偏好异质性而缩小,也就意味着现有研究可能高估了由价值重合或价值互补引起的嵌入效应。

## 3. 基于子样本估算空气污染降低的边际价值

结合现有研究的经验性结论,表 5 中统计显著的交叉项对应的个体特征均可以用来反映调查对象对环境污染的认知,即受教育水平更高、收入更高、对 AQI 了解程度更高、对雾霾和沙尘的知识水平更高的群体可以代表对大气污染认知度更高的群体。因此,这些变量可以作为大气污染认知程度的“代理变量”。表 5 的结论可以被引申为:认知度越高,大气污染的边际负效用越高。在此基础上,本文试图进一步讨论个体认知对嵌入效应带来的影响。为此,依据上述四个特征将每组样本都分割为两个子样本(低认知组和高认知组),进而基于每个子样本进行估计,并通过组间对比来讨论嵌入效应在低认知组和高认知组是否存在差异。基于子样本共估计了 24 个模型,如表 7 所示。

对比可知,表 3 中报告的估计结果可以被视为对表 7 中两个子样本的估计结果进行“包络”的结果。在所有子样本中,税收和雾霾的估计系数均显著为负值,说明其负效用在不同的子样本中具有稳健性。在各个子样本估计结果的基础上,本文同样模拟了 2000 个边际价值,其统计结果如表 8 和表 9 所示。从表 8 和表 9 的检验结果可知,在 8 个低认知组样本中有 2 个验证了嵌入效应,此外, $know > 4$  的子样本对沙尘的边际价值验证出了反方向的嵌入效应;而在 8 个高认知组样本中有 5 个

表5 考虑个体特征与环境污染交叉效应后的随机参数Logit模型估计结果

主效应	第一实验组		第二实验组		第三实验组			
	系数	标准误	系数	标准误	系数		标准误	
<i>tax</i>	-0.403**	0.068	-0.333**	0.050	-0.455**		0.047	
<i>haze</i>	12.446**	2.776			8.278**		2.856	
<i>sand</i>			15.079**	2.172	4.121		3.008	
交叉效应	<i>haze</i>		<i>sand</i>		<i>haze</i>		<i>sand</i>	
	系数	标准误	系数	标准误	系数	标准误	系数	标准误
<i>gender</i> ×	-0.101	0.553	-0.259	0.396	-0.128	0.611	0.062	0.655
<i>age</i> ×	-0.059	0.032	-0.101**	0.023	-0.060	0.031	-0.026	0.037
<i>educ</i> ×	-0.462**	0.143	-0.453**	0.107	-0.180	0.145	0.045	0.139
<i>inc</i> ×	-0.328**	0.064	-0.150**	0.046	-0.434**	0.080	-0.279**	0.085
<i>child</i> ×	0.123	0.561	-0.089	0.430	-0.717	0.621	-0.073	0.713
<i>illness</i> ×	-0.567	0.577	-0.257	0.403	0.132	0.644	0.074	0.731
<i>prof</i> ×	0.432	0.629	-0.569	0.475	-1.655*	0.727	-0.917	0.852
<i>aqi</i> ×	-2.093**	0.567	-0.429	0.426	-1.245	0.676	-1.748*	0.817
<i>know</i> ×	-0.671**	0.201	-1.168**	0.162	-0.608*	0.238	-0.597*	0.284
观测值	2772		2724		2784			
Chi2	341.70		39.23		251.98			
P值	0.000		0.000		0.000			
McFadden' R <sup>2</sup>	0.311		0.126		0.274			

注:①T检验为非配对T检验方法;②边际价值的单位为“元/天”;③\*\*表示在1%的统计水平上显著不为零;④限定所有交叉项的系数为固定参数,仅空气污染主效应的参数为随机参数;⑤篇幅原因,表中没有报告随机参数标准差的估计结果,如有需要可以向作者索取。

资料来源:作者整理。

表6 控制个体特征后模拟边际价值的统计及组间均值检验

2000个模拟边际价值的均值及标准差					
$w_1^1$	$w_2^2$	$w_1^3$	$w_2^3$	$w_1^3+w_2^3$	$w_1^1+w_2^2$
8.192**	4.213**	7.567**	3.193**	10.761**	12.879**
(1.023)	(0.730)	(0.855)	(0.862)	(1.250)	(1.288)
组间均值检验					
原假设	T检验P值		Poe检验P值		
$w_1^1=w_1^3$	0.000		0.319		
$w_2^2=w_2^3$	0.000		0.178		
$w_1^1+w_2^2=w_1^3+w_2^3$	0.000		0.177		

注:①T检验为非配对T检验方法;②边际价值的单位为“元/天”;③\*\*表示在1%的统计水平上显著不为零;④T检验的备择假设是“等式左侧≠等式右侧”,P值为双尾P值;Poe检验的备择假设是“等式左侧>等式右侧”,P值为单尾P值;⑤各组模拟的边际价值都基于一个相同的“典型居民”,其各项特征为表4中三组样本总体的平均值。

资料来源:作者整理。

验证出了嵌入效应。据此可以说明高认知组的样本更容易受到嵌入效应的影响。在表8和表9中,最右侧一栏的检验可以进一步验证这一结论。在对雾霾的4个检验中,3组检验的结果显示低认知

表 7 基于子样本的随机参数 Logit 模型估计结果

	第一实验组		第二实验组		第三实验组	
	系数	标准误	系数	标准误	系数	标准误
<i>edu</i> ≤ 15	观测值=648		观测值=600		观测值=720	
<i>tax</i>	-0.376**	0.123	-0.359**	0.115	-0.500**	0.101
<i>haze</i>	-1.692*	0.706			-2.376**	0.702
<i>sand</i>			-0.450	1.021	-1.002	0.748
<i>edu</i> > 15	观测值=2124		观测值=2124		观测值=2064	
<i>tax</i>	-0.414**	0.082	-0.324**	0.055	-0.471**	0.054
<i>haze</i>	-4.016**	0.528			-3.131**	0.415
<i>sand</i>			-1.640**	0.457	-1.397**	0.452
<i>inc</i> ≤ 8	观测值=1728		观测值=1848		观测值=1812	
<i>tax</i>	-0.356**	0.077	-0.335**	0.060	-0.439**	0.055
<i>haze</i>	-1.968**	0.438			-1.606**	0.373
<i>sand</i>			-0.865	0.501	-0.462	0.445
<i>inc</i> > 8	观测值=1044		观测值=876		观测值=972	
<i>tax</i>	-0.566**	0.146	-0.328**	0.088	-0.517**	0.092
<i>haze</i>	-7.131**	1.126			-5.838**	0.931
<i>sand</i>			-2.636**	0.754	-3.295**	0.812
<i>aqi</i> =0	观测值=996		观测值=948		观测值=780	
<i>tax</i>	-0.382**	0.107	-0.372**	0.085	-0.508**	0.100
<i>haze</i>	-2.258**	0.663			-3.368**	0.787
<i>sand</i>			-1.762*	0.716	-0.025	0.950
<i>aqi</i> =1	观测值=1776		观测值=1776		观测值=2004	
<i>tax</i>	-0.416**	0.088	-0.311**	0.061	-0.472**	0.054
<i>haze</i>	-4.158**	0.569			-2.906**	0.441
<i>sand</i>			-1.215*	0.516	-1.680**	0.418
<i>know</i> ≤ 4	观测值=1248		观测值=1260		观测值=1452	
<i>tax</i>	-0.517**	0.102	-0.269**	0.074	-0.425**	0.062
<i>haze</i>	-2.967**	0.624			-2.131**	0.549
<i>sand</i>			0.481	0.628	-0.870	0.515
<i>know</i> > 4	观测值=1524		观测值=1464		观测值=1332	
<i>tax</i>	-0.301**	0.093	-0.381**	0.067	-0.518**	0.070
<i>haze</i>	-3.766**	0.602			-3.709**	0.541
<i>sand</i>			-2.984**	0.562	-1.911**	0.538

注:①观测值=样本量×12,其中,12表示4个选择情景中的选项总数(每个选择情景有3个选项);②\*\*和\*分别表示在1%和5%的统计水平上显著;③为了增加估计系数的量级,*tax*的单位被调整为“百元”,*haze*和*sand*的单位被调整为“百天”;④限于篇幅,未报告随机参数标准差的估计结果和各模型的拟合效果信息,如有需要,可向作者索取。

资料来源:作者整理。

组样本边际价值的组间差异显著地低于高认知组样本;而在对沙尘的4个检验中,仅有1组支持这一结论,而且*aqi*=0的子样本组间差异反而显著地高于*aqi*=1的子样本。



表 8 基于子样本估计结果模拟的雾霾边际价值统计结果及组间均值检验

	$w_1^{0,1}$	$w_1^{0,3}$	$w_1^{0,1}=w_1^{0,3}$ (P 值)	$w_1^{1,1}$	$w_1^{1,3}$	$w_1^{1,1}=w_1^{1,3}$ (P 值)	$\Delta w_1^0=\Delta w_1^1$ (P 值)
<i>edu</i>	4.287** (6.947)	4.866** (1.469)	0.579	9.911** (1.585)	6.619** (0.914)	0.018	0.087
<i>inc</i>	5.563** (0.933)	3.645** (0.841)	0.059	13.261** (3.586)	11.448** (2.013)	0.322	0.553 <sup>T</sup>
<i>aqi</i>	6.093** (1.772)	6.785** (1.650)	0.631	10.353** (1.946)	6.217** (0.955)	0.007	0.050
<i>know</i>	5.802** (0.943)	5.075** (1.359)	0.323	13.899** (9.415)	7.223** (1.042)	0.008	0.056

注:① $w$ 的第一个上标为 0 表示低水平的子样本,上标为 1 表示高水平的子样本, $w$ 的第二个上标表示组别;② $\Delta w$ 由两组样本对应的  $w$  模拟值随机配对相减获得;③\*\*表示在 1%的统计水平上显著;④T 检验的 P 值全部为 0(除上标 T 所在的检验以外),表中未再报告,表中报告的 P 值为 Poe 检验的 P 值;⑤Poe 检验 P 值为 0.553 所在的假设检验中,T 检验的 P 值为 0.264。资料来源:作者整理。

表 9 基于子样本估计结果模拟的沙尘边际价值统计结果及组间均值检验

	$w_2^{0,2}$	$w_2^{0,3}$	$w_2^{0,2}=w_2^{0,3}$ (P 值)	$w_2^{1,2}$	$w_2^{1,3}$	$w_2^{1,2}=w_2^{1,3}$ (P 值)	$\Delta w_2^0=\Delta w_2^1$ (P 值)
<i>edu</i>	0.341 (11.825)	1.958** (1.475)	0.628	5.011** (0.894)	2.950** (0.902)	0.053	0.158
<i>inc</i>	2.396** (1.271)	1.026** (1.010)	0.184	8.173** (1.658)	6.395** (1.487)	0.196	0.450
<i>aqi</i>	4.597** (1.370)	-0.045 (2.031)	0.027	3.774** (1.221)	3.565** (0.790)	0.413	0.945
<i>know</i>	-2.663** (4.118)	1.969** (1.172)	0.948	7.864** (0.810)	3.686** (0.954)	0.001	0.001

注:① $w$ 的第一个上标为 0 表示低水平的子样本,上标为 1 表示高水平的子样本, $w$ 的第二个上标表示组别;②  $\Delta w$ 由两组样本对应的  $w$  模拟值随机配对相减获得;③\*\*表示在 1%的统计水平上显著;④T 检验的 P 值全部为 0,表中未再报告,表中报告的 P 值为 Poe 检验的 P 值。资料来源:作者整理。

现有研究在讨论陈述偏好方法的偏差时,普遍认为对价值评估对象认知度或熟悉度更高的样本偏差更小。例如,Dupont<sup>[40]</sup>以环境公共产品为评估对象,将调查对象分为主动使用者、潜在主动使用者和被动使用者,研究发现主动使用者在条件价值评估中的排序效应显著地低于潜在主动使用者和被动使用者。但本文得到了相反的结论,即调查对象对空气污染的认知度越高,评估的环境价值受到嵌入效应的影响越大。造成这一差异的主要原因在于偏差的起因存在差异。调查对象的认知能力有限和决策非理性被用来解释很多陈述偏好方法的偏差,也就是说,调查对象并没有根据一个清晰定义的偏好来进行陈述<sup>[31,41]</sup>。因此,对评估对象认知度或熟悉度更高的样本在陈述偏好时的不确定性更低,决策时的理性程度更高,因而受选择情景的表达方式和排序效应等偏差的影响就更小,调查对象采取策略性行为的概率也更低。相反,本文所验证的嵌入效应在很大程度上并非源自

于调查对象对决策情境的信息认知偏差或非理性决策,而是源自于雾霾天气和沙尘天气在致病损害和防护措施支出上的替代关系。从本文的分析结果看,对空气污染认知度更高的居民很可能投入了更多的防护措施,而在单独评价每种污染物时,这类调查对象基于应对空气污染的总防护支出进行回答,导致高估环境价值的程度更高,也即嵌入效应更强。

#### 4. 对检验结果的进一步讨论

综合上述分析可知,采用选择实验方法对雾霾和沙尘进行价值评估时会受到将评估对象嵌入实验情景的方式的影响,即单独基于每种污染物设计实验得到的评估结果显著地高于同时评估两种污染物得到的结果。这种嵌入效应具有稳健性,而且,基于空气污染认知度更高的居民验证的嵌入效应更强。这一结论首先为选择实验方法存在嵌入效应提供了新的证据。但嵌入效应的形成原因并不是组间实验设计方案存在的信息负荷差异、调查对象在不同实验设计中的信息认知差异或偏好不确定性,而是调查对象在不同嵌入结构下对污染物的偏好确实发生了变化。根据本文的理论分析,个体偏好受到污染物嵌入方式的影响机制表现在多个层面上,其中,多个污染物在致病时间和防护支出上的价值重合意味着单独评估污染物会高估环境价值,而多个污染物在休闲和个体舒适性上的互补效果则意味着单独评估污染物会低估环境价值。本文验证的嵌入效应说明雾霾和沙尘的价值重合效果大于互补效果。这意味着后续的应用研究者在评估某个环境产品的价值时必须要充分把握该产品与其他环境产品的价值关联,从而在实验设计时谨慎地处理可能存在的嵌入效应,提高估计结果的效度。

嵌入效应的存在对环境政策的效益评估更加重要。根据本文的分析,尽管雾霾和沙尘的治理政策各异且不存在目标重合,但由于雾霾和沙尘的价值本身存在相关性,经济学家在分析每种环境政策的效益时仍然会受到嵌入效应的影响。这对环境政策的制定提出了更为严格的要求。某项环境政策治理污染物的技术效果相对单一,因此政策的实施要求对降低单一污染物的效益进行评价。但缺乏专业知识的普通居民通常无法对各类污染物进行明确区分,单一污染物通常是通过总体的污染状况给居民造成福利损失,也就是说,多个污染物价值相关的现象普遍存在。因此,政策制定者在对某项环境政策的效益进行预先评估或事后评价时需要同时考虑与其相关的其他环境政策。尤其是,新制定环境政策时要充分考虑到现行政策的技术效果已经覆盖的价值,避免高估政策效益。

在成本效益分析中如果忽略了环境政策之间的这种间接相关性很可能导致由嵌入效应引起的偏差。但宽泛地引入过多环境政策或环境公共产品虽然可以充分把握评估对象之间的相关结构,却会严重增加陈述偏好方法的信息负荷,而过高的信息负荷会因为调查对象认知能力有限而引起一系列其他偏差<sup>[10,39]</sup>。因此,在实践中,研究者面临的问题是为这种相关结构确定一个合理的边界。在这个过程中,研究者首先要把握环境公共物品为个体带来价值的主要途径,通过这些价值途径确定哪些环境政策存在间接相关的关系以及相关程度有多高。而且,研究者还需要充分了解调查对象对环境公共物品的认知,本文的结论说明研究者要更加重视认知度更高的调查对象,这类调查对象由于认知的环境价值更高,因而受嵌入效应的影响也更大。

## 六、结论与启示

环境政策在技术目标性上通常指向相对明确和单一的污染物,但环境质量却经常是作为一个整体给普通民众带来价值。这个矛盾意味着对环境政策的效益评估会受到嵌入效应的影响。本文首先在一个扩展的健康生产模型的基础上分析了环境政策中嵌入效应的形成机制。尽管环境政策的技术目标单一,但降低不同污染物的价值存在关联。这种关联表现在多个污染物在致病损失和防护

支出上的替代效果,以及在休闲和个体舒适性上的互补效果。

以雾霾和沙尘的治理政策为例,本文采用选择实验方法设计了一个通过组间对比方法来识别嵌入效应的实验方案,前两个实验组分别单独评价降低雾霾和沙尘的经济价值,第三个实验组则对两者同时进行评价。结论显示,单独评价每种污染物时的经济价值显著地高于同时评价两种污染得到的结果。在单独评价时,降低雾霾和沙尘的平均边际价值(8.680元/天和4.199元/天)分别被高估了36.87%和67.62%。通过在估计模型中控制个体对环境偏好的异质性,高估的程度分别下降到了8.26%和31.94%。此外,对大气污染认知度更高的居民对降低污染的边际支付意愿也更高;在子样本回归结果基础上的假设检验进一步说明,认知度更高的居民在组间识别出的嵌入效应也更强。以上结果验证了雾霾治理政策和沙尘治理政策之间存在稳健的嵌入效应,而且,降低雾霾和降低沙尘的经济价值之间,替代效果强于互补效果。

成本效益分析是评价环境政策是否具有可行性的基本方法。本文验证的嵌入效应意味着政策制定者或研究者在对环境政策的效益进行评价之前,必须要谨慎地把握环境政策之间的嵌入效应。这种嵌入效应不仅来源于环境政策之间的直接相关性,还可能来源于间接相关性。其中,直接相关的政策是指污染治理的技术目标具有重合作用的环境政策,而间接相关的政策则是指虽然技术目标不存在重合,但目标污染物的经济价值存在相关性的环境政策。忽略了嵌入效应会导致环境政策效益评估的结果有偏。但有偏的方向需要根据政策效益之间是替代效果还是互补效果进行判断。就本文的雾霾和沙尘治理政策而言,单独评价其中的一种污染治理政策会导致政策效益被高估。

本文的研究结论支持后续的研究者采用选择实验方法处理多元环境政策间的嵌入效应。近年来选择实验在资源与环境经济学和农业经济学等领域的应用研究均呈快速增长趋势。研究者青睐于选择实验方法的一个重要原因在于选择实验可以同时多项相关的私人物品或公共物品进行价值评估,这也就意味着研究者可以通过设计恰当的选择情境来处理嵌入效应。但应用研究者通过引入更多相关产品处理嵌入效应时,还需要谨慎地处理选择实验复杂程度上升带来的风险,因为信息负荷过高的选择实验容易引起其他一系列的偏差。

#### [参考文献]

- [1]石庆玲,郭峰,陈诗一. 雾霾治理中的“政治性蓝天”——来自中国地方“两会”的证据[J]. 中国工业经济, 2016(5):40-56.
- [2]Greenstone, M., and R.Hanna. Environmental Regulations, Air and Water Pollution, and Infant Mortality in India[J]. American Economic Review, 2014,104(10):3038-3072.
- [3]陈硕, 陈婷. 空气质量与公共健康:以火电厂二氧化硫排放为例[J]. 经济研究, 2014,(8):158-169.
- [4]Won Kim, C., T. T. Phipps, and L.Anselin. Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach[J]. Journal of Environmental Economics and Management, 2003,45(1):24-39.
- [5]陈永伟, 陈立中. 为清洁空气定价:来自中国青岛的经验证据[J]. 世界经济, 2012,(4):140-160.
- [6]Freeman, A. M., J. A. Herriges, and C. L.Kling. The Measurement of Environmental and Resource Values: Theory and Methods[M]. Washington: Third Edition. Routledge, 2014.
- [7]Tan, J., and J. Zhao. The Value of Clean Air in China: Evidence from Beijing and Shanghai [J]. Frontiers of Economics in China, 2014,9(1):109-137.
- [8]蔡春光, 郑晓瑛. 北京市空气污染健康损失的支付意愿研究[J]. 经济科学, 2007,(1):107-115.
- [9]Louviere, J. J., D. A. Hensher, and J. D. Swait. Stated Choice Methods: Analysis and Applications [M]. New York: Cambridge University Press, 2000.
- [10]全世文. 选择实验方法研究进展[J]. 经济学动态, 2016,(1):127-141.

- [11]Kahneman, D., and J. L. Knetsch. Valuing Public Goods: The Purchase of Moral Satisfaction [J]. *Journal of Environmental Economics and Management*, 1992,22(1):57-70.
- [12]Shiell, A., and L. Gold. Contingent Valuation in Health Care and the Persistence of Embedding Effects without the Warm Glow[J]. *Journal of Economic Psychology*, 2002,23(2):251-262.
- [13]Loomis, J., M. Lockwood, and T. DeLacy. Some Empirical Evidence on Embedding Effects in Contingent Valuation of Forest Protection[J]. *Journal of Environmental Economics and Management*, 1993,25(1):45-55.
- [14]Hausman, J. Contingent Valuation: From Dubious to Hopeless [J]. *Journal of Economic Perspectives*, 2012,26(4):43-56.
- [15]Randall, A., and J. P. Hoehn. Embedding in Market Demand Systems [J]. *Journal of Environmental Economics and Management*, 1996,30(3):369-380.
- [16]Carson, R. T., and R. C. Mitchell. Sequencing and Nesting in Contingent Valuation Surveys [J]. *Journal of Environmental Economics and Management*, 1995,28(2):155-173.
- [17]Hanemann, W. M. Valuing the Environment Through Contingent Valuation [J]. *The Journal of Economic Perspectives*, 1994,8(4):19-43.
- [18]Sælensminde, K. Embedding Effects in Valuation of Non-market Goods[J]. *Transport Policy*, 2003,10(1):59-72.
- [19]Loomis, J., A. Gonzalez-Caban, and R. Gregory. Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates?[J]. *Land Economics*, 1994,70(4):499-506.
- [20]Mørkbak, M. R., T. Christensen, D. Gyrd-Hansen, and S. B. Olsen. Is Embedding Entailed in Consumer Valuation of Food Safety Characteristics[J]. *European Review of Agricultural Economics*, 2011,38(4):587-607.
- [21]Jacobsen, J. B., T. H. Lundhede, L. Martinsen, L. B. Hasler, and B. J. Thorsen. Embedding Effects in Choice Experiment Valuations of Environmental Preservation Projects [J]. *Ecological Economics*, 2011,70(6):1170-1177.
- [22]Grossman, M. On the Concept of Health Capital and the Demand for Health [J]. *Journal of Political Economy*, 1972,80(2):223-255.
- [23]Harrington, W., and P. R. Portney. Valuing the Benefits of Health and Safety Regulation [J]. *Journal of Urban Economics*, 1987,22(1):101-112.
- [24]曹静,王鑫,钟笑寒. 限行政策是否改善了北京市的空气质量? [J]. *经济学(季刊)*, 2014,13(3):1091-1126.
- [25]席鹏辉,梁若冰. 油价变动对空气污染的影响:以机动车使用为传导途径[J]. *中国工业经济*, 2015,(10):100-114.
- [26]曾贤刚,谢芳,宗佳. 降低 PM<sub>2.5</sub> 健康风险的行为选择及支付意愿——以北京市居民为例[J]. *中国人口·资源与环境*, 2015,25(1):127-133.
- [27]陈永伟,史宇鹏. 幸福经济学视角下的空气质量定价——基于 CFPS2010 年数据的研究[J]. *经济科学*, 2013,(06):77-88.
- [28]Hammit, J., and Y. Zhou. The Economic Value of Air-Pollution-Related Health Risks in China: A Contingent Valuation Study[J]. *Environmental and Resource Economics*, 2006,33(3):399-423.
- [29]曹怀术,廖华. 公众改善环境之意愿的层次贝叶斯分析[J]. *管理学报*, 2013,(2):274-279.
- [30]Alberini, A., M. Cropper, T. Fu, A. Krupnick, J. Liu, D. Shaw, and W. Harrington. Valuing Health Effects of Air Pollution in Developing Countries: The Case of Taiwan [J]. *Journal of Environmental Economics and Management*, 1997,34(2):107-126.
- [31]Carson, R., and T. Groves. Incentive and Informational Properties of Preference Questions [J]. *Environmental and Resource Economics*, 2007,37(1):181-210.
- [32]Windle, J. Comparing Responses from Internet and Paper-Based Collection Methods in more Complex Stated



- Preference Environmental Valuation Surveys[J]. *Economic Analysis and Policy*, 2011,41(1):83-97.
- [33]Olsen, S. Choosing Between Internet and Mail Survey Modes for Choice Experiment Surveys Considering Non-Market Goods[J]. *Environmental and Resource Economics*, 2009,44(4):591-610.
- [34]Train, K. E. *Discrete Choice Methods with Simulation*[M]. Cambridge: Cambridge University Press, 2003.
- [35]Revelt, D., and K. E. Train. Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level[J]. *The Review of Economics and Statistics*, 1998,80(4):647-657.
- [36]De -Magistris, T., A. Gracia, and R. M. Nayga. On the Use of Honesty Priming Tasks to Mitigate Hypothetical Bias in Choice Experiments [J]. *American Journal of Agricultural Economics*, 2013,95 (5):1136-1154.
- [37]Krinsky, I., and A. L. Robb. On Approximating the Statistical Properties of Elasticities [J]. *The Review of Economics and Statistics*, 1986,68(4):715-719.
- [38]Poe, G. L., K. L. Giraud, and J. B. Loomis. Computational Methods for Measuring the Difference of Empirical Distributions[J]. *American Journal of Agricultural Economics*, 2005,87(2):353-365.
- [39]Burton, M., and D. Rigby. The Self Selection of Complexity in Choice Experiments [J]. *American Journal of Agricultural Economics*, 2012,94(3):786-800.
- [40]Dupont, D. CVM Embedding Effects When There Are Active, Potentially Active and Passive Users of Environmental Goods[J]. *Environmental and Resource Economics*, 2003,25(3):319-341.
- [41]Day, B., I. J. Bateman, R. T. Carson, D. Dupont, J. J. Louviere, S. Morimoto, R. Scarpa, and P. Wang. Ordering Effects and Choice Set Awareness in Repeat-Response Stated Preference Studies [J]. *Journal of Environmental Economics and Management*, 2012,63(1):73-91.

## Embedding Effects in Evaluation of Multiple Environmental Policies ——Evidences from Beijing's Haze and Sand Control Policies

QUAN Shi-wen<sup>1</sup>, HUANG Bo<sup>2</sup>

- (1. Institute of Rural Development CASS, Beijing 100732, China;  
2. School of Agricultural Economics and Rural Development, Renmin University of China, Beijing 100872, China)

**Abstract:** This paper first, in theory, explores the causes of embedding effects among multiple environmental policies on the basis of an extended health production function. And then taking haze and sand control policies in Beijing as example, we design a split sample choice experiment in order to verify embedding effects. Random parameter logit model is adopted to conduct the estimation, and we simulate 2000 marginal values according to the estimates in each split sample. Results of mean tests confirm the significant and steady embedding effects, and if the effect is ignored, marginal values of reducing haze and sand would be overestimated by respectively 36.87% and 67.62%. The results indicate that before conducting cost and benefit analysis of an environmental policy, the policy makers or researchers should carefully examine the possible embedding effects in multiple policies, especially those targeting different pollutants but economic values of the pollutants are correlated or integrated.

**Key Words:** environmental policy; embedding effect; air quality; choice experiment; valuation

**JEL Classification:** Q51 Q58 C91

〔责任编辑:马丽梅〕