

# 缴税还是捐赠?

## ——环境政策效益评估中的支付工具选择

全世文\*

**摘要:** 在环境政策的效益评估中,采用不同的支付工具得到的评估结果往往存在着显著差异,但是,经济学家关于支付工具效应产生的原因尚存在争议。本文首先为支付工具效应的成因提供了一种理论解释,然后,以北京市大气污染治理为例设计了多个随机的实验组,在调查样本的基础上对比分析了“缴税”和“捐赠”两种支付工具的效益评估结果。分析结果证实了支付工具效应的存在,缺乏结果性承诺的实验设计、民众对政策实施主体的信任度差异以及不认同环境治理责任都会引发支付工具效应。从总体上看,“捐赠”支付工具对实验设计因素更加敏感,而“缴税”支付工具对样本认知更加敏感。本文的研究结论意味着研究者和政策制定者在开展环境政策效益评估时需要谨慎地选择支付工具,不同的支付工具在实现其效度时对实验设计和样本的要求存在差异。

**关键词:** 环境政策;支付工具;选择实验;大气污染;价值评估

### 一、引言

环境污染对公共健康的负外部性是政府对生态环境进行政策干预的理论依据。在特定的发展阶段,政府需要在生态环境和经济发展之间进行权衡,环境治理的目标是将环境污染控制在一个经济合理的水平。例如,黎文靖和郑曼妮(2016)发现只有在经济发展压力较小的阶段,城市才会增加环境治理的投资用于改善空气质量,而当地方经济发展压力较大时,空气质量并不会显著影响城市环境治理的投资决策。在微观层面上,就某项特定的环境治理政策或治理项目而言,判断其可行性也需要依赖成本效益分析,效益(环境改善增加的福利)可以补偿成本(经济效率下降损失的福利)是环境政策得以实施的一个必要条件。测度环境公共物品对人类福利的影响一直是经济学研究中的重点问题<sup>①</sup>。由于缺乏可供观测的市场数据,因此,陈述偏好方法被广泛用于评估环境价值和环境政策的效益。

本文关注环境政策效益评估中的支付工具选择问题。“支付工具”(Payment Vehicle)是指采用陈述偏好方法进行价值评估或政策效益评估时为调查对象设计的支付方式(Morrison

\*全世文,中国社会科学院农村发展研究所,邮政编码:100732,电子信箱:quanshiwen@163.com。

本文得到中国社会科学院创新工程A类项目“农业农村绿色发展理论与政策研究”(项目编号:2018NFS01)资助。感谢匿名审稿人提出的建议,文责自负。

<sup>①</sup>在环境政策的效益评估中,环境质量的价值评估是一个重要的环节。相比之下,某项环境治理政策的成本比效益更容易测度,例如,为了节能减排关停落后产能企业带来的失业和产值下降都易于核算,但减排后环境质量改善的价值则难以直接核算。因此,环境质量的价值评估是研究者更为关注的问题。

et al., 2000)。支付工具的重要性表现在以下三个方面。首先,支付工具是评估环境政策效益的福利计价单位。研究者之所以要在效益评估中加入支付工具就因为支付工具本身充当着一般等价物,选择何种支付工具就意味着通过何种计价单位来测度环境价值或政策效益。在大多数情况下,研究者会选择货币类的支付工具,例如,收入、价格、税费,等等。

其次,支付工具的选择会对环境价值评估结果的效度产生影响。由于陈述偏好实验情景的假想特征,如何保证评估结果的效度一直是陈述偏好方法在应用研究中面临的核心问题(全世文,2016)。随着应用研究的不断积累,经济学家发现并证实了陈述偏好方法中存在多种偏差,进而提出了一系列经验性的实验设计规范用来避免或缓解这些偏差。其中,选择合适的支付工具就是实验设计中的一个重要环节。Mitchell 和 Carson(1989)指出真实性和中立性是选择支付工具的两个重要标准。Bateman 等(2002)认为有效的支付工具必须是调查对象熟悉、信任和可接受的,要符合评估对象的现实背景,而且应该是强制的。早期的大量研究证实了在条件价值评估方法中采用不同的支付工具会导致价值评估结果出现显著的差异<sup>①</sup>(Morrison et al., 2000; Akcira, 2013)。

此外,支付工具通常还具有直接的政策含义。在很多情况下,研究者开展环境价值评估的目的不仅是为了评价环境政策或环境项目理论上的可行性,更是为政策或项目的实际开展做准备,支付工具的设计可以直接反映政策或项目的实施主体与资金的筹集方式。选择与实际情况不同的支付工具得到的评估结果很可能与实际的支付效果存在差异。例如,就收费风景区的环境改造项目而言,研究者一般会选择最现实的门票价格作为支付工具,而不会选择收入、征税等其他支付工具。又例如,对政府的大型环境治理工程而言,征税是最常用的支付工具,但征收额外环境税和税收再分配同样是两种不同的筹资方案,也会使评估结果存在差异(Ivehammar, 2009)。

鉴于支付工具在环境价值评估中的重要意义,自从陈述偏好方法提出以来,国际上就有大量文献致力于支付工具的研究。这些研究验证了不同支付工具价值评估结果存在的差异(例如 Greenley et al., 1981; Morrison et al., 2000),并在大量经验研究结论的基础上总结了支付工具的设计规范(例如 Bateman et al., 2002)。但是,学界关于支付工具效应产生的原因仍然缺乏足够的阐释。现有研究讨论的两类典型的支付工具分别是强制型支付工具(以缴税为代表)和自愿型支付工具(以捐赠为代表)。Bateman 等(2002)认为采用自愿型支付工具时,调查对象会采取“搭便车”的策略性行为,导致环境价值会被低估。但是,这种解释并未被实证研究直接证实,而且,Carson(1997)也曾指出采用自愿型支付工具的相反效果:调查对象在自愿型支付工具下具有更强的动机夸大其实际的支付意愿。此外,应用研究关于两类支付工具所得价值评估结果的差异也并未得到一致的结论(Akcira, 2013)。相比之下,国内环境价值评估领域仅有少量应用研究考虑了不同支付工具的差异(例如高汉琦等, 2011; 刘海凤等, 2011; 全世文等, 2018),而大量文献在实验设计时甚至没有清晰地对支付工具进行界定。

本文拟以北京市的大气污染治理为例,对比分析“向政府缴税”和“向环境公益基金捐

---

<sup>①</sup>需要指出的是,“支付工具偏差”和“支付工具效应”是两个不同的概念,前者指支付工具的设计不合理导致调查对象认知错误或拒绝回答等情形,从而影响了评估结果的效度;后者指设计合理的不同支付工具同样会导致评估结果存在差异(Akcira, 2013)。本文主要关注“支付工具效应”。

赠”两种支付工具下的环境政策效益,并重点讨论其差异形成的原因。本文的主要贡献在于以下两点:第一,论证采用不同支付工具得到的价值评估结果存在差异的原因。本文提供了一种基于策略性行为发生的动机来阐释支付工具效应产生原因的理论框架。策略性行为是指调查对象在陈述偏好时有意扭曲自身真实偏好的行为。强制型和自愿型支付工具因其内涵的义务性差异和对支付的“结果性”差异,会导致调查对象采取不同性质的策略性行为,从而诱发支付工具效应。本文同时通过分组实验设计来验证上述解释是否成立。

第二,在选择实验方法中讨论支付工具的选择问题。现有研究对支付工具的讨论主要集中在条件价值评估方法中,而鲜有研究在选择实验方法中对比不同支付工具的评估结果。由于选择实验具有高信息负荷的优势,近年来该方法的应用研究快速上升(全世文,2016)。而且,许多研究已经证实,与条件价值评估方法要求调查对象在评估对象和金钱价值之间直接进行权衡的模式相比,选择实验要求调查对象在多个选项之间进行权衡进而间接进行价值评估的模式可以降低调查对象在实验过程中采取策略性行为的动机,因此评估结果具有更高的外部效度(例如 Carlsson and Martinsson, 2001; Lusk and Schroeder, 2004)。这一结论意味着选择实验方法中存在的支付工具效应可能有别于条件价值评估方法。

综上所述,下文首先在理论分析部分讨论了支付工具效应产生的原因,然后设计了多个随机的选择实验组,分别采用缴税和捐赠作为支付工具来评价北京市大气污染治理政策(项目)的效益。而且,本文在实验中重点收集了反映支付工具效应产生原因的相关信息,进而通过检验组间评估结果是否存在差异对这些原因进行验证。

## 二、理论分析

### (一) 基于选择实验评估环境政策效益的逻辑

考虑政策制定者或研究者需要对某项特定的大气污染治理政策或治理项目进行效益评估。记空气污染水平为  $\mathbf{A} = (A_1, A_2, \dots, A_K)$ , 表示由  $K$  种污染物反映的空气质量。不实施该项政策意味着空气污染水平将维持现状, 记为  $\mathbf{A}^0$ ; 实施了该项政策可以使空气污染水平下降到  $\mathbf{A}^1$ 。为了测度空气污染从  $\mathbf{A}^0$  下降到  $\mathbf{A}^1$  的价值, 政策制定者或研究者设计了一个选择实验, 并招募了  $N$  个被试者参与实验。简单起见, 假设选择实验中仅包含一个选择情景, 而且, 该选择情景仅包含两个选项<sup>①</sup>: “实施某项大气污染治理政策”(下文简称“实行政策”)和“维持现状”, 前者意味着空气质量为  $\mathbf{A}^1$ , 后者意味着空气质量为  $\mathbf{A}^0$ 。

假设空气质量仅作为一种需求品影响人类福利, 即对于任意被试者  $n \in N$ , 其效用函数为  $U_n = U_n(\mathbf{X}, \mathbf{A})$ , 其中,  $\mathbf{X}$  为私人物品向量。令  $\mathbf{C}$  表示私人物品对应的价格向量,  $M_n$  表示被试者  $n$  的收入, 预算约束为  $\mathbf{X}\mathbf{C}' = M_n$ , 那么, 间接效用函数为  $U_n = V_n(M_n, \mathbf{C}, \mathbf{A})$ 。于是, 该项治理政策在理论上对被试者  $n$  的价值可以通过希克斯补偿变差(Compensating Variation)进行反映, 即  $V_n(M_n, \mathbf{C}, \mathbf{A}^0) = V_n(M_n - CV_n, \mathbf{C}, \mathbf{A}^1)$ 。另外, 通过全微分可以求得污染物  $k$  对被试者  $n$  的边际价值为  $w_{nk} = dM_n/dA_k = -(\partial V_n/\partial A_k)/(\partial V_n/\partial M_n)$ 。

但是, 研究者通常无法直接获取  $CV_n$ <sup>②</sup>, 也不会将收入作为支付工具直接测度  $w_{nk}$ 。在选

①当模型扩展为多个选择情景和多个选项以后, 通过增加适当的假设条件, 分析结论并不会发生变化。

②早期研究者获取  $CV_n$  的一个直接方法是通过开放式问题直接询问被试者支付意愿或受偿意愿。但这种方法被公认为存在严重的偏差, 因此, 目前很少有研究直接使用这种方法开展价值评估。

择实验中,实验设计者会在“实行政策”的选项中加入一个支付工具:税收  $T_n$  或者捐赠  $D_n$ ,由此测度  $CV_n$  和  $w_{nk}$ ,在“维持现状”的选项中,  $T_n = D_n = 0$ 。因此,被试者  $n$  需要对比两个选项对应的效用:  $V_n^0$  与  $V_n^1$ ,其中,  $V_n^0 = V_n(M_n, C, A^0)$ ,  $V_n^1 = V_n(M_n - T_n, C, A^1)$  或  $V_n(M_n - D_n, C, A^1)$ 。记被试者  $n$  的选择为  $B_n$ ,  $B_n = 1$  表示被试者  $n$  选择了“实行政策”,  $B_n = 0$  表示选择了“维持现状”。记被试者  $n$  以外的所有其他被试者的选择为  $\mathbf{B}_{-n}$ 。于是,研究者可以根据所有被试者的决策信息对政策效益进行估算,即  $\hat{CV} = F(B_n, \mathbf{B}_{-n}, e)$ ,  $\hat{w}_k = G(B_n, \mathbf{B}_{-n}, e)$ ,式中,  $e$  表示据以估算价值的其他外生信息,包括被试者的特征、空气污染水平、投标值,等等。

记被试者  $n$  基于真实偏好的“正确选择”应该为  $R_n$ ,  $R_n = 1$  表示  $V_n^1 \geq V_n^0$ ,  $R_n = 0$  表示  $V_n^1 < V_n^0$ 。函数  $F(\cdot)$  与  $G(\cdot)$  通常是建立在随机效用模型基础上的估计过程,根据 Train(2003) 可得,

命题 1: 当所有的被试者  $n \in N$  都有  $B_n = R_n$  时,通过  $F(\cdot)$  与  $G(\cdot)$  可以得到无偏且一致的估计量  $\hat{CV}$  与  $\hat{w}_k$ 。

命题 1 有两点含义:第一,研究者有能力通过开展选择实验得到有信息价值的估计量,在此基础上,研究者可以做进一步的效益转移分析,从而结合政策实施成本等其他因素决定该项环境政策或环境项目是否实施。第二,研究者得到有信息价值的估计量的一个必要条件是激励相容原则,即选择实验的实验设计需要保证被试者根据其真实的偏好进行决策。

## (二) 选择实验中的策略性行为

令  $\Delta \hat{CV} / \Delta B_n$  和  $\Delta \hat{w}_k / \Delta B_n$  表示被试者  $n$  选择“实行政策”选项对价值评估结果的边际影响。根据命题 1 不难证明:  $\Delta \hat{CV} / \Delta B_n \geq 0$ ,  $\Delta \hat{w}_k / \Delta B_n \geq 0$ ,即任一被试者从选择“维持现状”改变为选择“实行政策”都意味着价值评估的结果会有所上升,或至少保持不变。于是,当  $R_n = 0, B_n = 1$  时,环境政策的效益会被高估;而当  $R_n = 1, B_n = 0$  时,政策效益则会被低估。

设被试者  $n$  认知的政策实施概率函数为  $P_n = P_n(B_n, \mathbf{B}_{-n}, s_n, Z)$ ,式中,  $Z$  表示其他相关的外生信息,包括实验设计方法、环境政策的成本,等等;  $s_n$  表示被试者  $n$  的个人特征。相应的,被试者  $n$  认为政策不实施的概率为  $1 - P_n$ 。函数  $P_n$  反映了由被试者认知的实验过程对现实政策的影响力,  $\Delta P_n / \Delta B_n$  则表示被试者  $n$  个人的选择对政策实施概率的边际影响<sup>①</sup>。于是,被试者的目标可以被描述为最大化以下期望效用:

$$EV_n = P_n(B_n, \mathbf{B}_{-n}, s_n, Z) \cdot V_n^1 + (1 - P_n(B_n, \mathbf{B}_{-n}, s_n, Z)) \cdot V_n^0$$

Carson 和 Groves(2007) 提出了二元选择的实验设计满足激励相容的一系列条件<sup>②</sup>。条件 i: 被试者关心实验结果;条件 ii: 被试者认同环境质量的治理责任设定;条件 iii: 实验中的支付能够得到强制执行;条件 iv: 实验设计中的“实行政策”选项与其真实政策效果完全对应,且面临着一个二元离散选择的形式;条件 v:  $\Delta P_n / \Delta B_n > 0, \forall n \in N$ 。Vossler 等(2012) 证明了以下命题,

命题 2: 当条件 i - v 同时满足时,选择实验的设计是激励相容的,即被试者  $n \in N$  的占

①注意:函数  $P_n(\cdot)$  与函数  $F(\cdot)$ 、 $G(\cdot)$  的差异在于  $P_n(\cdot)$  是一个主观认知函数,而  $F(\cdot)$  和  $G(\cdot)$  是客观函数。一个正常的政策效益评估逻辑意味着政策实施的客观概率是  $\hat{CV}$  的增函数,因此,也是  $B_n$  和  $\mathbf{B}_{-n}$  的增函数,即:选择“实行政策”的被试者越多,评估的政策效益越高,政策实施的概率就越高。但被试者的主观概率函数  $P_n(\cdot)$  并不一定满足这一条件,这涉及到被试者对实验结果影响力的判断。

②Carson 与 Groves(2007) 的原文中没有提及条件 ii,但条件 ii 也是激励相容的一个前提条件。

优策略均为  $B_n = R_n$ 。

相反,违背了以上任意一个条件都会增加被试者采取策略性行为的动机,可以进一步证明(过程详见附录)。

命题3:当条件 i、ii、iv、v 满足而条件 iii 不满足时,被试者  $n \in N$  的占优策略是  $B_n = 1$ 。

命题4:当条件 i-iv 满足而条件 v 不满足,且  $\Delta P_n / \Delta B_n = 0$  或  $P_n = 0$  时,被试者  $n \in N$  的占优策略是  $B_n = 0$ 。

命题3意味着当被试者无需为其陈述的决策进行实际支付时,被试者会采取夸大空气质量价值的策略,从而提高环境政策的实施概率,因此,政策效益会被高估。命题4则意味着当被试者认为政策实施的现实决策与选择实验中被试者的决策无关时,被试者始终会选择“维持现状”的选项,即“搭便车”决策,因此,政策效益会被低估。

### (三) 支付工具设计对策略性行为的影响

下面,比较分别采用“向政府缴税”( $T_n$ )和“向环境公益基金捐赠”( $D_n$ )两种支付工具时被试者采取策略性行为动机的差异。

首先,考虑支付工具对条件 iii 的影响。条件 iii 意味着被试者需要为自己的选择承担真实的支付责任。违背条件 iii 的直接效果是“实施政策”选项中设计的支付工具  $T_n$  或  $D_n$  不再进入被试者  $n$  的效用函数,这就会导致命题3中的情形,即被试者有意地夸大环境价值或政策效益。严格来讲,条件 iii 应该被定义为被试者的主观认知:被试者认为当其选择了“实施政策”选项以后,支付会被强制执行。因此,附带强制支付的实验设计必须要提前告知被试者实验中的强制支付规则。相反,不要求强制支付的实验设计通常不会提前告知被试者无需强制支付的规则。也就是说,被试者在选择情景中面临着是否强制支付的不确定性,因此,被试者会根据实验设计方面的信息对强制支付的概率做出主观判断。显然,与捐赠概念的自愿性相比,税收概念的强制性会提高被试者对强制支付概率的判断,这就意味着采用支付工具  $D_n$  时,条件 iii 被违背的概率比支付工具  $T_n$  更高。相反,根据命题2,在强制支付的实验设计中,无论何种支付工具,被试者都更倾向于汇报真实偏好,因此,支付工具对价值评估结果的影响会减小。据此提出假说1。

假说1:当选择实验不要求被试者强制支付时,采用“缴税”支付工具较“捐赠”支付工具评估的环境政策效益更低;当选择实验要求被试者强制支付时,两种支付工具评估的环境政策效益差距会减小。

然后,考虑支付工具对条件 v 的影响。条件 v 意味着被试者需要认同自己在选择实验中的决策对环境政策的真正实施具有影响力。 $\Delta P_n / \Delta B_n = 0$  且  $P_n \neq 0$  的情况意味着政策实施概率外生于选择实验的评估过程。与强制支付的不确定性相似,当实验设计没有向被试者强化实验的现实意义时,被试者通常也面临着实验结果影响力的不确定性,而不同支付工具背后的政策实施主体很可能会对这一不确定性造成影响。普通民众与政府部门的接触显然多于与社会公益组织的接触,对支付工具及其政策实施主体的不熟悉会增加  $\Delta P_n / \Delta B_n = 0$  的概率,这就会导致命题4中的情形,即环境政策的效益被低估。但是,向被试者提前进行实验意义的信息干预则可以提高  $\Delta P_n / \Delta B_n > 0$  的概率。据此提出假说2。

假说2:当选择实验没有向被试者说明实验结果对环境政策的影响力时,采用“缴税”支付工具较“捐赠”支付工具评估的环境政策效益更高;而提前向被试者进行实验影响力的信息干预可以减小这一差异。

条件  $v$  被违背的另一种更强的假设是  $P_n = 0$ , 即被试者认为实验中设计的环境政策在现实中没有实施的可能性。排除实验设计不合理的因素, 出现这一情况的原因很可能在于被试者并不相信政策实施主体有能力实施实验中设计的环境政策。相比之下, 政府部门实施环境政策的权威性显然高于环境公益组织, 而且, 税收的强制性也保证了政府部门具有更强的筹资能力来实现既定的环境政策目标。这种信任度的差异意味着采用“捐赠”支付工具时,  $P_n = 0$  的概率更高。据此提出假说 3。

假说 3: 被试者对支付工具背后的政策实施主体的信任会影响环境政策效益的评估结果; 当更多的被试者对政府的信任度强于对环境公益组织的信任度时, 采用“缴税”支付工具较“捐赠”支付工具评估的环境政策效益更高。

最后, 考虑支付工具对条件 ii 的影响。条件 ii 意味着被试者拥有“维持现状”的治理责任, 并愿意以“支付”的形式获取空气质量改善带来的福利(支付意愿问题)。如果被试者不认同这一产权界定, 意味着被试者认为自己拥有享受清洁空气的产权, 因此, 应当以“受偿”的形式对空气污染的现状进行补偿(受偿意愿问题)。与捐赠相比, 缴税形式的支付工具与产权界定的相关性更高, 即政府是环境公共物品提供的一个权利主体, 而环境公益组织则无法承担这一功能。也就是说, “缴税”具有更强的义务性质, 反映了被试者在利己动机下的决策; 相比之下, “捐赠”则出于被试者纯粹或非纯粹的利他动机, 即“温光效应”(Andreoni, 1990)。因此, 当被试者  $n$  不认同实验中的治理责任界定时, 对任意  $T_n > 0$  都有  $V_n^1 < V_n^0$ , 被试者  $n$  倾向于做出  $B_n = 0$  的决策; 而对于一个特定范围内的捐赠  $0 < D_n < D^*$ , 利他动机仍有可能使  $V_n^1 > V_n^0$ , 由此使被试者做出  $B_n = 1$  的决策。相反, 当被试者认同治理责任界定时, 受最大支付意愿的限制,  $T_n$  和  $D_n$  会存在替代关系, 不同动机下支付额度的差异可能并不明显。据此提出假说 4。

假说 4: 当被试者不认同实验中的治理责任界定时, 采用“缴税”支付工具较“捐赠”支付工具评估的环境政策效益更低; 当被试者认同治理责任界定时, 采用两种支付工具评估的环境政策效益差距会减小。

### 三、实验设计

本文的实验设计包括两个环节: 第一, 以大气污染的治理效果为价值评估对象, 设计选择实验; 第二, 设计随机分组方案, 在每个实验组中采用不同的支付工具, 或添加不同的信息干预内容, 从而通过组间比较对上述假说进行验证。

首先介绍选择实验的设计方案。城市的空气污染物包括多种类别, 即存在多个备选属性用来构造向量  $\mathbf{A}$ 。由于属性过多会大幅增加实验的信息负荷, 从而导致被试者采取非理性的信息处理策略(全世文, 2016), 因此, 本文仅选择了可吸入颗粒物(PM10)和细颗粒物(PM2.5)作为价值评估对象, 前者是导致沙尘天气的主要污染物, 后者是导致雾霾天气的主要污染物。<sup>①</sup> 定义污染程度的度量单位为污染天气的年发生天数, 本文将 2014 年北京市 PM2.5 超标天数(即雾霾天气天数)和 PM10 超标天数(即沙尘天气天数)作为环境政策效果的现状基准, 前者为 180 天, 后者为 100 天。设定大气污染治理效果有三个水平: 分别将污

<sup>①</sup>根据实验设计的经验规范, 研究者应该首先选择被试者关切度更高的属性。根据课题组的预调查结果, 在六类空气污染的基本项目中, PM2.5 和 PM10 被北京市居民认为是最主要的两个空气污染物。

染天数下降 25%、50%和 75%。与此同时,在实验中加入一个支付工具(“缴税”或“捐赠”),参考同类研究对国内空气质量价值的评估结果,设计基准支付水平为 400 元/年,并通过增加 50%和降低 50%分别获得两个额外的支付水平:600 元/年和 200 元/年。于是,选择实验对应一个 3 属性 3 水平的设计方案。本文采用正交设计方案生成了 9 个反映大气污染治理效果的选项,然后根据属性水平平衡、效用平衡和最小重叠三个原则<sup>①</sup>配对了 4 组选择情景,每个情景中均包含了两个政策效果选项:效果 A 和效果 B。此外,每个选择情景中都加入了一个“维持现状”的选项,图 1 描述了一个选择情景的示例<sup>②</sup>。每个被试者都需要在 4 个不同的选择情景中进行多次选择,为了避免排序效应的影响,每份问卷中的 4 个选择情景都进行了随机排序。

第2个选择情景 (共4个)

|                                                                                                                    | 环境政策效果A                   | 环境政策效果B                  | 2014年北京市<br>空气质量状况<br>(维持现状) |
|--------------------------------------------------------------------------------------------------------------------|---------------------------|--------------------------|------------------------------|
| <br>每年雾霾 (PM <sub>2.5</sub> ) 天数  | 135天<br>(较2014年下降<br>25%) | 45天<br>(较2014年下降<br>75%) | 180天                         |
| <br>每年沙尘 (PM <sub>10</sub> ) 天数 | 50天<br>(较2014年下降<br>50%)  | 75天<br>(较2014年下降<br>25%) | 100天                         |
| <br>每年向政府额外上缴的专项税额              | 600元                      | 400元                     | 0元                           |
| 在上述三种状况中, 您更愿意接受哪种政策效果?<br>(1) 效果A    (2) 效果B    (3) 维持2014年的现状                                                    |                           |                          |                              |

图1 被试者面临的一个选择情景示例图

然后,表 1 介绍了本文的随机分组方案。依据实验前干预信息的差异和支付工具的差异,本文共设计了六个实验组。其中,前三个实验组以缴税作为支付工具,后三个实验组以捐赠作为支付工具。“信息 1”为选择实验对真实环境政策的影响力信息,内容设计如下:

“课题组会将研究成果递交给北京市环保局(反馈给某环境公益基金会),您在选择情景中的决策会对北京市政府制定的大气污染治理政策(环境公益基金会开展的大气污染治

①关于这三个原则的解释可以参考 Huber 和 Zwerina(1996)、全世文(2016)。

②图 1 是“缴税”支付工具的选择情景示例图,在采用“捐赠”支付工具的选择情景中,选项名称为“环境项目效果 A”、“环境项目效果 B”,支付工具为“每年向环境公益基金捐赠的专项金额”。此外,在所有实验组的被试者进入选择情景以前,问卷都向被试者提前介绍了一个选择情景的示例及填答说明,并向被试者展示了空气质量良好和空气污染的对比图。图 1 中展示的图片分别为本研究设定的标准“中度雾霾天气”和“中度沙尘天气”。

理项目)产生重要影响。<sup>①</sup>”

“信息 2”为强制支付信息,内容设计如下:

“您在正式选择情景中的决策会被实验后台记录,当实验结束以后,实验后台会随机选择一个选择情景,您在该选择情景中的支付(捐赠)会按照特定比例被强制执行,即实验结束后的补偿会根据您的选择扣除一定的金额,扣除金额的部分会转交于北京市环保局(某环境公益基金会)。”

表 1 分组实验设计方案

| 设计方案       | 第 1 组 | 第 2 组 | 第 3 组 | 第 4 组 | 第 5 组 | 第 6 组 |
|------------|-------|-------|-------|-------|-------|-------|
| 实验前信息干预    |       |       |       |       |       |       |
| 信息 1:政策影响力 | √     |       | √     | √     |       | √     |
| 信息 2:强制支付  |       | √     | √     |       | √     | √     |
| 实验设计选择的属性  |       |       |       |       |       |       |
| 雾霾天气       | √     | √     | √     | √     | √     | √     |
| 沙尘天气       | √     | √     | √     | √     | √     | √     |
| 支付工具       | 缴税    | 缴税    | 缴税    | 捐赠    | 捐赠    | 捐赠    |
| 实验后追问信息    |       |       |       |       |       |       |
| 信息 3:信任度   |       |       | √     |       |       | √     |
| 信息 4:治理责任  |       |       | √     |       |       | √     |

如表 1 所示,对第 1 组和第 4 组仅干预信息 1,对第 2 组和第 5 组仅干预信息 2,对第 3 组和第 6 组同时干预两种信息。此外,第 3 组和第 6 组的被试者会在选择实验结束后被追问信息 3 和信息 4。其中,“信息 3”为被试者对环境治理主体的信任,问题设计如下<sup>②</sup>:

“在开展大气污染的治理政策或治理项目时,您更信任以下哪类治理主体?

(1) 政府部门(国家环保部、各级政府环保局及其他环保相关的政府机构)

(2) 环境公益组织(非政府组织,例如中华环保基金会、自然之友,等等)”

“信息 4”为环境治理责任认知,问题设计如下,

“以下两种说法,您更赞同哪一种说法?

(1) 空气污染是全社会共同的责任,每个普通民众都有社会义务为治理空气污染做出贡献。

(2) 空气污染主要是排污企业或政府的责任,普通民众有权利享受清洁的空气,不需要为治理空气污染承担社会义务。”

在信息 4 中,选择(1)表示调查对象认同治理责任,选择(2)表示不认同治理责任。在上述实验设计基础上的组间比较可以对前文提出的假说进行验证。第 1 组和第 4 组为非强制支付的实验设计,用于验证假说 1;第 2 组和第 5 组是没有强调实验结果影响力的实验设计,用于验证假说 2;第 3 组和第 6 组为充分信息干预的标准实验设计,可以被视为控制组。此外,在第 3 组和第 6 组内可以通过信息 3 和信息 4 分别验证假说 3 与假说 4。表 2 列示了验证各个假说所需的子样本。

<sup>①</sup>在 *T* 实验组,括号内的文字被删除;在 *D* 实验组,信息干预内容使用括号内的文字,下文信息 2 同。

<sup>②</sup>原则上讲,“信息 3”中对不同主体的信任并不互斥,即可以允许多项选择或都不选择,但是,为了简化分析,本文要求被试者强制选择 1 个相对更信任的主体。这种设计就真实的“信任”而言可能有偏,但是,如果假设潜在的“都信任”或“都不信任”的群体在被要求做出强制选择时并不存在结构性的偏向性,那么,这种设计并不会对本文的分析结论造成较大影响。

表 2 各个假说对应的子样本

| 假说         | 子样本  | 子样本包含的实验组   |
|------------|------|-------------|
| 假说 1       | 样本 1 | 第 1、3、4、6 组 |
| 假说 2       | 样本 2 | 第 2、3、5、6 组 |
| 假说 3, 假说 4 | 样本 3 | 第 3、6 组     |

## 四、数据与实证方法

### (一) 调查实施与样本描述

本文分析所用的数据来源于课题组于 2016 年 10 月对北京市居民开展的问卷调查。调查问卷主要包括三个部分的信息:调查对象的个人特征、调查对象对大气污染的认知以及上一节设计的选择情景。课题组采用网络调查方法,委托某网络调查公司在其北京市样本库中投放调查问卷,调查持续 10 天,共收集有效样本 1 322 份。课题组限制调查对象均为在北京市常住时间超过 3 年的居民,从而保证调查对象对北京市空气污染状况具有良好的认知。表 3 汇报了各个实验组样本的统计特征。

表 3 各实验组样本的描述统计

| 变量            | 含义                           | 第 1 组                         | 第 2 组             | 第 3 组                         | 第 4 组                         | 第 5 组                         | 第 6 组                         |
|---------------|------------------------------|-------------------------------|-------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
|               |                              | 222                           | 215               | 217                           | 219                           | 223                           | 226                           |
| <i>gender</i> | 性别(男=0,女=1)                  | 0.514<br>(0.501)              | 0.460<br>(0.500)  | 0.488<br>(0.501)              | 0.553<br>(0.498)              | 0.570<br>(0.496)              | 0.527<br>(0.500)              |
| <i>age</i>    | 年龄                           | 34.351<br>(5.860)             | 34.549<br>(5.986) | 33.788<br>(5.546)             | 33.466<br>(5.254)             | 33.534<br>(5.764)             | 34.354<br>(5.527)             |
| <i>edu</i>    | 受教育年限                        | 15.347<br>(3.331)             | 15.219<br>(2.901) | 15.452<br>(3.188)             | 15.352<br>(4.053)             | 14.803<br>(2.966)             | 14.774<br>(3.109)             |
| <i>income</i> | 月收入水平(千元)                    | 7.081 <sup>T</sup><br>(5.028) | 6.805<br>(4.646)  | 7.065 <sup>T</sup><br>(4.382) | 7.959 <sup>T</sup><br>(6.429) | 5.969 <sup>T</sup><br>(3.945) | 5.858 <sup>T</sup><br>(3.640) |
| <i>child</i>  | 家庭是否有小孩<br>(是=1,否=0)         | 0.545<br>(0.499)              | 0.512<br>(0.501)  | 0.590<br>(0.493)              | 0.539<br>(0.500)              | 0.552<br>(0.498)              | 0.553<br>(0.498)              |
| <i>ill</i>    | 是否患过呼吸系统疾病<br>(是=1,否=0)      | 0.572<br>(0.496)              | 0.558<br>(0.498)  | 0.562<br>(0.497)              | 0.562<br>(0.497)              | 0.520<br>(0.501)              | 0.531<br>(0.500)              |
| <i>prof</i>   | 是否具有环保专业知识<br>背景(是=1,否=0)    | 0.095<br>(0.293)              | 0.116<br>(0.321)  | 0.115<br>(0.320)              | 0.132<br>(0.340)              | 0.139<br>(0.347)              | 0.111<br>(0.314)              |
| <i>trust</i>  | 信息 3(信任政府=1,信任<br>环境公益组织=0)  | —<br>—                        | —<br>—            | 0.751<br>(0.433)              | —<br>—                        | —<br>—                        | 0.735<br>(0.443)              |
| <i>ipr</i>    | 信息 4(认同治理责任=0,<br>不认同治理责任=1) | —<br>—                        | —<br>—            | 0.318<br>(0.467)              | —<br>—                        | —<br>—                        | 0.305<br>(0.462)              |

注:(1)报告的统计量为各变量的样本均值,括号内为对应的标准差;(2)上标 T 表示组间均值 T 检验可以在 1% 的统计水平上拒绝均值相等的原假设,这些检验组集中在 *income* 变量,具体包括:(1,5)、(1,6)、(3,5)、(3,6)、(4,5)、(4,6);除此以外,对所有变量,任意两组均值 T 检验的结果都无法在 1% 的统计水平上拒绝均值相等的原假设。

从表 3 来看,六个实验组的样本分配基本均匀。样本特征与课题组在 2015 年底组织的类似调查所获样本的特征基本一致,和北京市居民的情况相比,调查样本同样存在女性比例略高、平均年龄偏低、平均受教育年限偏高和平均收入偏高的问题,这也是网络调查样本的

共同特点。也就是说,调查样本在对北京市居民的代表性上存在一定的缺陷。但是,随着移动互联网的快速普及,网络调查样本的偏差正在不断缩小。而且,与全世文和黄波(2016)的研究相似,本研究对调查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。从表3中各组样本的统计检验来看,各实验组样本在多数统计指标上不存在显著的差异,仅第5组和第6组样本在收入指标上偏低。综合来看,调查样本基本满足组间随机分配的要求,可以较好地服务于本文的研究目标。

## (二) 实证方法

本文的实证过程包括两个环节:第一,在各实验组样本中估算空气质量(雾霾天气和沙尘天气)的价值;第二,检验空气质量价值在组间是否存在显著差异,从而对前文提出的假说进行验证。首先考虑估算空气质量价值的方法,令间接效用函数  $V(\cdot)$  对支付工具  $pv$  和空气质量属性  $\mathbf{A} = (haze, sand)$  表现为一个线性函数:

$$V = \beta_0 + \beta_1 \cdot haze + \beta_2 \cdot sand + \beta_3 \cdot pv + \varepsilon \quad (1)$$

(1)式中: $\beta_1$  和  $\beta_2$  分别表示雾霾天气发生天数  $haze$  和沙尘天气发生天数  $sand$  的边际负效用,即  $\beta_1 < 0, \beta_2 < 0$ ;  $pv$  表示支付额度(可以选择缴税  $T$  或者捐赠  $D$ ),  $\beta_3$  表示支出的边际负效用,即  $\beta_3 < 0$ 。那么,雾霾天气和沙尘天气减少的边际价值分别为  $w_1 = \beta_1 / \beta_3$  和  $w_2 = \beta_2 / \beta_3$ , 记总价值为  $w_3 = w_1 + w_2$ 。本文采用随机参数 Logit 模型对(1)式中的参数  $\beta$  进行估计,详细可以参考 Train(2003)。为了得到理论上呈正态分布的边际价值  $w$ , 在估计(1)式时,本文限制  $\beta_3$  为固定参数,并假设  $\beta_1$  和  $\beta_2$  为服从正态分布的随机参数。分别在每组样本中采用以上随机参数 Logit 模型对(1)式进行估计,即可得到基于每组样本估算的边际价值  $w^i, \forall i = 1, 2, \dots, 6$ 。

然后,本文采用两种方法检验边际价值  $w$  在组间是否存在差异。第一种方法:首先在随机参数 Logit 模型的估计结果基础上,根据贝叶斯法则估算所有特定选择序列的被试者的偏好参数(估计方法可以参考 Train(2003)),由此可以进一步计算得到每个被试者对应的边际价值(边际支付意愿);然后基于表2中所示的各个子样本将不同实验组样本作为合并数据(pool data)分别进行如下式所示的 OLS 回归<sup>①</sup>:

$$w = \alpha_0 + \alpha \cdot s + \lambda_0 \cdot D + \lambda_1 \cdot coer + \theta_1 \cdot coer \cdot D + \mu \quad (2)$$

$$w = \alpha_0 + \alpha \cdot s + \lambda_0 \cdot D + \lambda_2 \cdot cons + \theta_2 \cdot cons \cdot D + \mu \quad (3)$$

$$w = \alpha_0 + \alpha \cdot s + \lambda_0 \cdot D + \lambda_3 \cdot trustb + \theta_3 \cdot trustb \cdot D + \lambda_4 \cdot ipr + \theta_4 \cdot ipr \cdot D + \mu \quad (4)$$

(2)-(4)式中: $s$  表示被试者的个人特征,  $\alpha$  为其对应的待估参数;  $D$  表示支付工具虚拟变量,缴税作为支付工具时取值为0,捐赠作为支付工具时取值为1,  $\lambda_0$  为其对应的待估参数。(2)式用于验证假说1,所用样本为表2中的“样本1”,  $coer$  变量是含义为“是否没有要求强制支付”的虚拟变量,强制支付时取值为0,未强制支付时取值为1,  $\lambda_1$  为其对应的待估参数,  $\theta_1$  为  $D$  与  $coer$  交互项的待估参数<sup>②</sup>。于是,假说1可以定义为:  $\lambda_0 + \theta_1 > 0, \theta_1 > 0$ 。同理,(3)式

①第一种方法的优点在于相对简单易行,从被试者个体的边际支付意愿出发进行分析也更加直观。但是,第一种方法同时也存在缺陷。(2)-(4)式中的因变量是一个估计量,而非真实值,因此,从严格意义上讲,待估参数的显著性需要根据因变量的标准差进行矫正。但是,通过模拟运算得到每个被试者边际支付意愿的标准差需要进行大量的运算,因此,本文并未进行这一操作。

②在(2)式中,因为  $coer$  和  $D$  是严格外生的(由随机分组的实验设计所决定),所以,不引入  $s$  也可以得到无偏且一致的估计结果,但是,引入  $s$  可以提高估计结果的效度。(3)式与此相同。但是,(4)式中的  $trustb$  和  $ipr$  并非由实验设计随机生成,所以,必须要引入  $s$  以避免由遗漏变量导致的内生性问题。

用于验证假说2,所用样本为表2中的“样本2”,*cons*变量的含义是“是否没有进行实验政策影响力的信息干预”,有信息干预时取值为0,无信息干预时取值为1, $\lambda_2$ 和 $\theta_2$ 分别是*cons*变量和交互项的待估参数。于是,假说2可以定义为: $\lambda_0+\theta_2<0, \theta_2>0$ 。(4)式同时用于验证假说3和假说4,所用样本为表2中的“样本3”,*trustb*变量表示“被试者是否信任所在实验组的支付工具对应的环境治理主体”,不信任时取值为0,信任时取值为1<sup>①</sup>, $\lambda_3$ 和 $\theta_3$ 分别是*trustb*变量和交互项的待估参数。*ipr*变量的含义是“被试者是否不认同实验设计中的环境治理责任”,该变量根据被试者对“信息4”的回答直接生成,认同治理责任时取值为0,不认同治理责任时取值为1, $\lambda_4$ 和 $\theta_4$ 分别是*ipr*变量和交互项的待估参数。于是,假说3可以定义为: $\lambda_3>0, \theta_3=0$ ;假说4可以定义为: $\lambda_0+\theta_4>0, \theta_4>0$ 。

第二种方法:首先,在各组样本对(1)式的估计结果的基础上,采用Krinsky和Robb(1986)提出的限制参数Bootstrapping方法在各个实验组都模拟1000个偏好参数,进而计算出1000个模拟的边际价值*w*;然后,采用组间T检验方法和Poe等(2005)提出的“非参数完全组合方法”(后文称为“Poe检验”)对各组计算的边际价值进行组间均值检验。于是,假说1可以定义为: $w^1<w^4, w^4-w^1>w^6-w^3$ <sup>②</sup>;假说2可以定义为: $w^2>w^5, w^2-w^5>w^3-w^6$ 。另外,为了采用该方法验证假说3和假说4,本文将第3组和第6组样本依据被试者对信息3和信息4的回答进一步分解为两个子样本;记第3组和第6组样本中选择更信任政府部门的子样本组分别为“第31组”和“第61组”,选择更信任环境公益组织的子样本组分别为“第32组”和“第62组”;选择认同治理责任的子样本组分别为“第33组”和“第63组”,选择不认同治理责任的子样本组分别为“第34组”和“第64组”。于是,假说3可以被定义为: $w^{31}>w^{32}, w^{61}<w^{62}$ ;假说4可以被定义为: $w^{34}<w^{64}, w^{64}-w^{34}>w^{63}-w^{33}$ 。

## 五、结果与讨论

### (一) 空气质量对个体的边际价值及其组间比较

首先,基于各实验组的样本采用随机参数Logit模型对(1)式进行估计,结果如表4所示。根据表4可知,各实验组的估计模型都具有良好的拟合效果,基本上所有的偏好参数在5%的统计水平上都显著不为零。而且,所有均值参数的符号均为负,与理论预期相符合,说明空气污染和货币支出都会给被试者带来负效用。与此同时,高度显著的标准差系数也证实了被试者对空气污染的偏好异质性,说明了选用随机参数Logit模型进行估计较普通的多元Logit模型更恰当。此外,对比表4中各实验组的估计结果可知,参数估计值处于相同的数量级,即(1)式中的 $\varepsilon$ 在不同实验组中不存在明显差异,说明不同实验组的被试者在选择情景中面临着程度相近的不确定性,或者说,被试者具有程度相近的决策理性。这一结论与各实验组具有相同信息负荷的实验设计相一致,可以排除后文验证的组间差异来自组间的信息负荷差异,或被试者在组间的信息处理策略差异。

①注意,*trustb*变量不同于表3中的*trust*变量,*trustb*变量并非由被试者对“信息3”的回答直接生成,而需要同时考虑被试者所在的实验组。

②采用组间均值检验的方法验证假说时,不仅涉及对边际价值的直接检验,还涉及对边际价值差值的检验。在计算边际价值的差值时,本文将不同组计算的边际价值进行随机配对,然后相减得到差值的分布。

表 4 各实验组随机参数 Logit 模型的估计结果

| 变量              | 第 1 组                | 第 2 组                | 第 3 组                | 第 4 组                | 第 5 组                | 第 6 组                |
|-----------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 均值              |                      |                      |                      |                      |                      |                      |
| <i>haze</i>     | -2.835 **<br>(0.354) | -2.380 **<br>(0.258) | -2.702 **<br>(0.340) | -2.277 **<br>(0.328) | -2.355 **<br>(0.383) | -4.201 **<br>(0.348) |
| <i>sand</i>     | -1.117 **<br>(0.375) | -0.782 **<br>(0.286) | -1.058 **<br>(0.386) | -0.786 *<br>(0.338)  | -0.673<br>(0.373)    | -1.296 **<br>(0.312) |
| <i>pv</i>       | -0.356 **<br>(0.042) | -0.364 **<br>(0.039) | -0.360 **<br>(0.043) | -0.194 **<br>(0.035) | -0.503 **<br>(0.051) | -0.654 **<br>(0.056) |
| 标准差             |                      |                      |                      |                      |                      |                      |
| <i>haze</i>     | 3.932 **<br>(0.432)  | 2.767 **<br>(0.293)  | 3.988 **<br>(0.457)  | 3.573 **<br>(0.386)  | 4.297 **<br>(0.469)  | 2.851 **<br>(0.308)  |
| <i>sand</i>     | 2.854 **<br>(0.451)  | 1.203 *<br>(0.470)   | 2.847 **<br>(0.487)  | 2.026 **<br>(0.480)  | 1.902 **<br>(0.500)  | 1.243 *<br>(0.543)   |
| 观测值             | 2 664                | 2 580                | 2 604                | 2 628                | 2 676                | 2 712                |
| Chi2            | 355.5                | 221.7                | 344.9                | 327.9                | 389                  | 206.8                |
| P 值             | 0.000                | 0.000                | 0.000                | 0.000                | 0.000                | 0.000                |
| McFadden' $R^2$ | 0.209                | 0.140                | 0.209                | 0.192                | 0.235                | 0.150                |

注:(1)观测值=样本量 $\times$ 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),各组的样本量详见表3;(2)括号内为标准误,\*\*和\*分别表示在1%和5%的统计水平上显著;(3)为了增加系数的量级,在估计时,*pv*的单位被调整为“百元”,*haze*和*sand*的单位被调整为“百天”。

在上述估计结果的基础上,进一步根据贝叶斯法则估算各实验组中每一个被试者的偏好参数,然后计算边际价值,结果统计如表5所示。表5报告的统计结果可以用来反映被试者对降低空气污染的边际支付意愿。以第三组样本为例,被试者对减少一个标准的中度雾霾天和中度沙尘天的平均支付意愿分别为7.904元/天和2.895元/天,这一结果与部分同类研究的估算结果相一致(例如,Tan and Zhao,2014;全世文、黄波,2016)。各组估算的雾霾边际价值都显著高于沙尘的边际价值,这一结论也符合近年来普通民众对空气污染的认知,说明雾霾已经成为民众最为关切的空气污染物,其带来的健康价值损失也最为严重。直观上看,各组估算的边际价值之间存在明显的差异,说明信息干预和支付工具的选择会对价值评估的结果造成影响。为了更科学地验证这一结论,分别基于表2中不同的子样本对(2)-(4)式进行OLS回归,结果如表6所示。

表 5 各实验组空气质量对被试者边际价值的统计结果

| 分组    | 雾霾边际价值( $w_1$ ) |        | 沙尘边际价值( $w_2$ ) |       | 总价值( $w_t$ ) |        |
|-------|-----------------|--------|-----------------|-------|--------------|--------|
|       | 均值              | 标准差    | 均值              | 标准差   | 均值           | 标准差    |
| 第 1 组 | 7.836 **        | 9.399  | 2.931 **        | 4.934 | 10.767 **    | 10.776 |
| 第 2 组 | 6.586 **        | 6.287  | 2.187 **        | 1.299 | 8.773 **     | 6.655  |
| 第 3 组 | 7.904 **        | 9.263  | 2.895 **        | 4.815 | 10.800 **    | 10.756 |
| 第 4 组 | 11.644 **       | 15.665 | 4.041 **        | 5.674 | 15.685 **    | 17.056 |
| 第 5 组 | 4.747 **        | 7.422  | 1.314 **        | 1.796 | 6.061 **     | 7.760  |
| 第 6 组 | 6.407 **        | 3.565  | 1.970 **        | 0.733 | 8.378 **     | 3.778  |

注:(1)\*\*表示在1%的统计水平上显著;(2)边际价值的单位为“元/天”。

表 6 空气质量对被试者边际价值的 OLS 回归结果

| 变量                     | 模型 1     |       | 模型 2      |       | 模型 3      |       |
|------------------------|----------|-------|-----------|-------|-----------|-------|
|                        | 系数       | 标准误   | 系数        | 标准误   | 系数        | 标准误   |
| <i>gender</i>          | 1.215    | 0.730 | 0.561     | 0.497 | 1.047     | 0.673 |
| <i>age</i>             | -0.083   | 0.066 | 0.060     | 0.043 | 0.136 *   | 0.061 |
| <i>edu</i>             | 0.506 ** | 0.137 | 0.344 **  | 0.103 | 0.239     | 0.135 |
| <i>income</i>          | 0.525 ** | 0.095 | 0.287 **  | 0.076 | 0.178     | 0.108 |
| <i>child</i>           | 0.555    | 0.736 | 1.161 *   | 0.498 | 0.911     | 0.677 |
| <i>ill</i>             | 1.720 *  | 0.735 | 0.603     | 0.499 | 1.354 *   | 0.674 |
| <i>prof</i>            | 1.124    | 1.153 | 1.317     | 0.762 | 0.789     | 1.065 |
| <i>D</i>               | -1.366   | 1.035 | -1.830 ** | 0.702 | -2.348    | 1.239 |
| <i>coer</i>            | 0.059    | 1.035 |           |       |           |       |
| <i>D · coer</i>        | 5.679 ** | 1.469 |           |       |           |       |
| <i>cons</i>            |          |       | -1.810 *  | 0.708 |           |       |
| <i>D · cons</i>        |          |       | -0.553    | 0.993 |           |       |
| <i>trustb</i>          |          |       |           |       | 2.790 *   | 1.125 |
| <i>D · trustb</i>      |          |       |           |       | -0.406    | 1.546 |
| <i>ipr</i>             |          |       |           |       | -8.663 ** | 1.042 |
| <i>D · ipr</i>         |          |       |           |       | 5.299 **  | 1.462 |
| 常数项                    | 0.076    | 3.025 | -0.012    | 2.138 | 0.003     | 2.974 |
| $\lambda_0 + \theta_1$ | 4.313 ** | 1.037 |           |       |           |       |
| $\lambda_0 + \theta_2$ |          |       | -2.383 ** | 0.707 |           |       |
| $\lambda_0 + \theta_4$ |          |       |           |       | 2.952 *   | 1.425 |
| 样本量                    | 884      |       | 881       |       | 443       |       |
| <i>P</i> 值             | 0.000    |       | 0.000     |       | 0.000     |       |
| 修正 $R^2$               | 0.162    |       | 0.116     |       | 0.248     |       |

注:(1) \*\* 和 \* 分别表示在 1% 和 5% 的统计水平上显著;(2) 模型 1、模型 2 与模型 3 分别对应于(2)式、(3)式和(4)式,所用子样本分别为表 2 中的样本 1、样本 2 和样本 3;(3)各模型所用因变量均为  $w_3$ 。

表 6 中,三个模型的整体拟合效果都表现良好。被试者个体特征对空气质量价值的影响效果符合经验研究的结论,例如,受教育年限和收入水平的提高都会提高个人对空气质量的支付意愿。以下重点考虑支付工具的影响效果。在模型 1 中, $\lambda_0 + \theta_1$  显著为正,说明第 4 组样本评估的空气质量价值显著高于第 1 组,也就是说,在没有要求强制支付的实验设计中,捐赠支付工具比缴税支付工具评估的价值更高。与此同时, $\theta_1$  也显著为正,说明第 3 组和第 6 组的价值评估差异小于第 1 组和第 4 组的差异,即强制支付的实验设计可以降低由支付工具差异引起的价值评估差异。据此,假说 1 得到验证。在模型 2 中, $\lambda_0 + \theta_2$  显著为负,说明第 5 组样本评估的空气质量价值显著低于第 2 组,也就是说,如果不强调实验结果的政策影响力,捐赠支付工具比缴税支付工具评估的价值更低。但是, $\theta_2$  并不显著,说明第 3 组和第 6 组的价值评估差异并没有显著地低于第 2 组和第 5 组差异,也就是说,缺失政策影响力信息导致价值被低估的程度在两种支付工具中的差异并不明显。据此,假说 2 部分得证。在模型 3 中,*trustb* 变量系数显著为正,说明被试者对支付工具背后的环境政策实施主体的信任度会对价值评估的结果产生正面的影响。而且,交叉项 *D · trustb* 并不显著,说明信任对于价值评估的影响效果在两种支付工具中也不存在显著的差异。据此可知,因为有更多的被试者更加信任政府部门,而非环境公益组织(见表 3),因此,在充分信息干预的条件下,缴税支付工具评估的空气质量价值高于捐赠支付工具(第 3 组高于第 6 组)。据此,假说 3 得到证实。另外, $\lambda_0 + \theta_4$  显著为正,说明当被试者不认同治理责任的设定时,捐赠支付工具比

缴税支付工具评估的价值更高。与此同时,  $\theta_4$  也显著为正, 说明在认同治理责任的子样本中, 两种支付工具评估的价值差异小于不认同治理责任的子样本。据此, 假说 4 得到证实。

根据表 2 的计算, 作为两种支付工具控制组的第 3 组和第 6 组评估的空气质量总价值相差了 2.422 元/天, 该差值在 1% 的统计水平上显著。如模型 3 所示, 当在模型中控制了被试者特征以后, 支付工具效应仍然存在 (2.348 元/天), 只是统计显著性有所下降 (10% 的统计水平上显著)。这一结果说明, 即使在实验设计过程中进行了充分的信息干预, 支付工具效应仍然可能存在, 信任因素和治理责任认知因素都是导致支付工具效应的潜在原因。但是两个因素的作用机制并不相同。由于信任对价值评估的促进作用在两种支付工具之间没有显著差异, 所以, 信任导致的支付工具效应并不体现在个体层面上, 而是体现在总体层面上。也就是说, 由于居民对不同的环境政策实施主体的整体信任度分布存在差异, 所以导致了政策效益评估结果存在差异。而治理责任认知对价值评估的影响作用在两种支付工具之间存在显著差异, 当个体认为普通居民具有享受清洁空气的治理责任时, 基于支付意愿开展价值评估本身就存在偏差。此时, 由于“温光效应”的存在, 采用捐赠支付工具较缴税支付工具评估的价值会更高。

## (二) 空气质量的模拟边际价值及其组间比较

进而考虑直接采用组间均值检验的方法对假说进行验证。如前文所述, 假说 1 和假说 2 可以分别通过样本 1 和样本 2 内部的组间比较进行验证, 但是, 对假说 3 和假说 4 的验证需要进一步根据被试者对信息 3 和信息 4 的回答将第 3 组和第 6 组样本划分为多个子样本。表 7 首先汇报了基于不同子样本的随机参数 Logit 模型估计结果。

表 7 信任和治理责任子样本组随机参数 Logit 模型的估计结果

| 变量              | 第 31 组               | 第 32 组               | 第 61 组               | 第 62 组               | 第 33 组               | 第 34 组               | 第 63 组               | 第 64 组               |
|-----------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 均值              |                      |                      |                      |                      |                      |                      |                      |                      |
| <i>haze</i>     | -3.094 **<br>(0.582) | -1.975 **<br>(0.363) | -3.276 **<br>(0.332) | -6.806 **<br>(1.242) | -5.018 **<br>(0.633) | -0.22<br>(0.385)     | -4.709 **<br>(0.441) | -2.289 **<br>(0.491) |
| <i>sand</i>     | -2.645 **<br>(0.496) | 0.907<br>(0.543)     | -0.993 **<br>(0.347) | -2.077 *<br>(0.856)  | -1.467 *<br>(0.598)  | -0.704<br>(0.530)    | -1.346 **<br>(0.369) | -1.162 *<br>(0.531)  |
| <i>pv</i>       | -0.364 **<br>(0.052) | -0.343 **<br>(0.072) | -0.627 **<br>(0.060) | -0.701 **<br>(0.156) | -0.406 **<br>(0.062) | -0.309 **<br>(0.061) | -0.586 **<br>(0.065) | -0.748 **<br>(0.105) |
| 标准差             |                      |                      |                      |                      |                      |                      |                      |                      |
| <i>haze</i>     | 6.078 **<br>(1.010)  | 1.541 **<br>(0.362)  | 2.594 **<br>(0.319)  | 0.751<br>(1.374)     | 4.215 **<br>(0.577)  | 2.184 **<br>(0.421)  | 2.145 **<br>(0.330)  | 2.937 **<br>(0.534)  |
| <i>sand</i>     | 2.047 *<br>(0.877)   | 1.159<br>(0.779)     | 1.086<br>(0.716)     | 3.102<br>(1.649)     | 4.547 **<br>(0.859)  | 1.010<br>(1.113)     | 1.501<br>(0.880)     | 0.221<br>(1.366)     |
| 样本量             | 163                  | 54                   | 166                  | 60                   | 148                  | 69                   | 157                  | 69                   |
| 观测值             | 1 956                | 648                  | 1 992                | 720                  | 1 776                | 828                  | 1 884                | 828                  |
| Chi2            | 350.1                | 25.2                 | 141.8                | 2.0                  | 238.1                | 54.1                 | 53.6                 | 75.5                 |
| P 值             | 0.000                | 0.000                | 0.000                | 0.367                | 0.000                | 0.000                | 0.000                | 0.000                |
| McFadden' $R^2$ | 0.282                | 0.064                | 0.129                | 0.012                | 0.242                | 0.098                | 0.069                | 0.159                |

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) 括号内为标准误, \*\* 和 \* 分别表示在 1% 和 5% 的统计水平上显著; (3) 为了增加系数的量级, 在估计时, *pv* 的单位被调整为“百元”, *haze* 和 *sand* 的单位被调整为“百天”; (4) 各子样本组的定义详见第四节实证方法部分。

根据表 7 汇报的结果,除了第 62 组以外,各模型的整体拟合效果都表现良好<sup>①</sup>。结合表 4 的估计结果可知,第 3 组和第 6 组的估计结果可以被视为对任一信息维度划分的两个子样本的估计结果进行“加权平均”的结果。从总体上看,各模型  $pv$  变量估计系数的量级并没有因为子样本的划分发生明显的改变,但  $haze$  与  $sand$  变量的均值估计系数在两个子样本之间却表现出了明显的差异,而且,这种差异在标准差估计系数中的表现同样明显。这说明信任和治理责任认知确实会对被试者的空气质量偏好产生影响,这一结果间接地支持了表 6 中模型 3 所反映的结论。

分别基于表 4 和表 7 的估计结果,采用限制参数的 Bootstrapping 方法在各样本组中模拟 1 000 个空气质量的偏好参数,并据此分别计算出 1 000 个对应的雾霾边际价值、沙尘边际价值和总价值,统计情况如表 8 所示。表 8 报告的第 1 组-第 6 组的边际价值均值与表 5 报告的统计结果高度一致,以第三组样本为例,一个中度雾霾天气和一个中度沙尘天气的平均边际价值分别为 7.600 元/天和 2.886 元/天。但是,由于模拟偏好参数是控制了样本异质性以后的均值偏好参数,因此,模拟边际价值的标准差明显小于表 5 报告的统计结果。与表 7 反映的信息相一致,表 8 中第 3 组和第 6 组模拟的边际价值均值也可以被视为对子样本模拟边际价值进行“加权平均”的结果。而且,在任一信息维度下的两个子样本中,空气质量的边际价值都表现出了明显的差异,再次说明了信任和治理责任认知都会对被试者的空气质量偏好产生影响。

表 8 各实验组空气质量模拟边际价值的统计结果

| 分组     | 雾霾边际价值( $w_1$ ) |       | 沙尘边际价值( $w_2$ ) |       | 总价值( $w_t$ ) |       |
|--------|-----------------|-------|-----------------|-------|--------------|-------|
|        | 均值              | 标准差   | 均值              | 标准差   | 均值           | 标准差   |
| 第 1 组  | 8.063 **        | 1.059 | 3.087 **        | 0.944 | 11.149 **    | 1.355 |
| 第 2 组  | 6.608 **        | 0.755 | 2.101 **        | 0.694 | 8.709 **     | 0.894 |
| 第 3 组  | 7.600 **        | 1.028 | 2.886 **        | 0.974 | 10.486 **    | 1.339 |
| 第 4 组  | 12.103 **       | 2.544 | 3.977 **        | 1.611 | 16.079 **    | 3.032 |
| 第 5 组  | 4.717 **        | 0.685 | 1.288 **        | 0.686 | 6.005 **     | 0.811 |
| 第 6 组  | 6.454 **        | 0.437 | 1.956 **        | 0.406 | 8.410 **     | 0.545 |
| 第 31 组 | 8.655 **        | 1.726 | 7.266 **        | 1.256 | 15.921 **    | 2.306 |
| 第 32 组 | 5.957 **        | 1.278 | -2.997 **       | 2.055 | 2.960 **     | 1.480 |
| 第 61 组 | 5.248 **        | 0.437 | 1.547 **        | 0.489 | 6.796 **     | 0.543 |
| 第 62 组 | 9.939 **        | 1.484 | 2.957 **        | 1.140 | 12.896 **    | 2.149 |
| 第 33 组 | 12.590 **       | 1.835 | 3.559 **        | 1.362 | 16.149 **    | 2.373 |
| 第 34 组 | 0.699 **        | 1.217 | 2.128 **        | 1.624 | 2.827 **     | 1.523 |
| 第 63 组 | 8.086 **        | 0.647 | 2.263 **        | 0.545 | 10.349 **    | 0.819 |
| 第 64 组 | 3.090 **        | 0.559 | 1.492 **        | 0.602 | 4.582 **     | 0.659 |

注:(1) \*\* 表示在 1% 的统计水平上显著;(2) 边际价值的单位为“元/天”;(3) 第 1 组-第 6 组的模拟值来自表 4 的估计结果,第 31 组-第 64 组的模拟值来自表 7 的估计结果。

<sup>①</sup>根据笔者采取的似然比检验结果,第 62 组随机参数 Logit 模型的估计结果并不显著优于普通的多元 Logit 模型,也就是说,第 62 组样本并没有显著的偏好异质性。采用多元 Logit 模型估计得到的 Chi2 为 364.1,  $P$  值为 0.000, McFadden'  $R^2$  为 0.310, 三个均值估计系数依次为 -5.968、-1.939、-0.674, 均在 1% 的统计水平上显著。采用此估计结果模拟得到的边际价值分布与随机参数 Logit 模型没有显著差异。因此,为了避免因模型设定导致的组间差异,后文直接采用随机参数 Logit 模型的估计结果进行模拟运算。

根据表 8 模拟的边际价值,本文分别采用两种均值检验方法对组间差异进行验证,检验结果如表 9 所示。由于模拟边际价值的标准差较小(见表 8),因此,所有 T 检验都在 1% 的统计水平上拒绝了“等式左侧=等式右侧”的原假设。下文的讨论主要基于 Poe 检验的结果。从整体上看,雾霾价值相对于沙尘价值更容易拒绝原假设,在 10% 的统计水平上,雾霾价值有 7 组检验拒绝了原假设,而沙尘价值有 3 组检验拒绝了原假设。这一结论与现有研究的结论相符合,说明被试者更关切的属性价值更容易受到实验设计因素的影响(例如, Gao and Schroeder, 2009)。正因为如此, Bennett 和 Blamey (2001) 才建议在实验设计时应当首先考虑引入被试者更加需求的和政策相关度更高的属性。原假设  $w^3 = w^6$  在总价值的检验中被拒绝,这与表 6 中模型 3 的估计结果相一致,说明即使是充分的信息干预,支付工具效应也同样存在,结合表 8 第 3 组和第 6 组的结果计算可知,采用缴税支付工具评估的空气质量价值较捐赠支付工具高出了 24.68%。但是,在雾霾价值和沙尘价值的检验中并没有拒绝该原假设,可以理解为,强制型支付工具较自愿型支付工具会系统性地高估各项环境属性的价值,但高估程度在单个环境属性中并不稳健。

表 9 空气质量模拟边际价值的组间均值检验结果

| 原假设                                 | 雾霾边际价值( $w_1$ ) |               | 沙尘边际价值( $w_2$ ) |               | 总价值( $w_t$ ) |               |
|-------------------------------------|-----------------|---------------|-----------------|---------------|--------------|---------------|
|                                     | T 检验<br>P 值     | Poe 检验<br>P 值 | T 检验<br>P 值     | Poe 检验<br>P 值 | T 检验<br>P 值  | Poe 检验<br>P 值 |
| $w^3 = w^6$                         | 0.000           | 0.146         | 0.000           | 0.185         | 0.000        | 0.071         |
| $w^4 = w^1$                         | 0.000           | 0.049         | 0.000           | 0.312         | 0.000        | 0.052         |
| $w^4 - w^1 = w^6 - w^3$             | 0.000           | 0.000         | 0.000           | 0.019         | 0.000        | 0.000         |
| $w^2 = w^5$                         | 0.000           | 0.030         | 0.000           | 0.202         | 0.000        | 0.012         |
| $w^2 - w^5 = w^3 - w^6$             | 0.000           | 0.123         | 1.000           | 0.589         | 0.000        | 0.213         |
| $w^{31} = w^{32}$                   | 0.000           | 0.097         | 0.000           | 0.000         | 0.000        | 0.000         |
| $w^{62} = w^{61}$                   | 0.000           | 0.000         | 0.000           | 0.118         | 0.000        | 0.000         |
| $w^{31} - w^{32} = w^{62} - w^{61}$ | 1.000           | 0.977         | 0.000           | 0.000         | 0.000        | 0.009         |
| $w^{64} = w^{34}$                   | 0.000           | 0.031         | 1.000           | 0.667         | 0.000        | 0.139         |
| $w^{64} - w^{34} = w^{63} - w^{33}$ | 0.000           | 0.000         | 0.000           | 0.309         | 0.000        | 0.000         |

注:(1)所有检验的 P 值均为单尾 P 值,备择假设为:等式左侧>等式右侧;(2)所有对差值的均值检验中,差值均通过随机配对相减得到;(3)所有 T 检验均为非配对 T 检验,且标准差不相等。

以下考虑对各个假说的验证(主要基于雾霾价值和总价值的 Poe 检验结果)。原假设  $w^4 = w^1$  被拒绝,说明在非强制支付的实验设计中,捐赠支付工具比缴税支付工具评估的空气质量价值更高;原假设  $w^4 - w^1 = w^6 - w^3$  被拒绝,说明强制支付的实验设计可以降低两种支付工具的价值评估差异;据此,假说 1 得证。原假设  $w^2 = w^5$  被拒绝,说明如果不强调实验结果的政策影响力,捐赠支付工具较缴税支付工具评估的价值更低;但是,原假设  $w^2 - w^5 = w^3 - w^6$  同样不能被拒绝,说明第 2 组与第 5 组的差异并不能因为政策影响力信息的干预而降低;据此,假说 2 部分得证。原假设  $w^{31} = w^{32}$  和  $w^{62} = w^{61}$  均被拒绝,说明信任因素对价值评估具有显著的促进作用;原假设  $w^{31} - w^{32} = w^{62} - w^{61}$  也被拒绝,说明信任因素的促进作用在两种支付工具中具有显著差异,而且,该差异在雾霾价值和沙尘价值中方向相反;据此,假说 3 得证。原假设  $w^{64} = w^{34}$  被拒绝,说明当被试者不认同治理责任设定时,捐赠支付工具较缴税支付工具评估的价值更高;原假设  $w^{64} - w^{34} = w^{63} - w^{33}$  被拒绝,说明认同治理责任可以降低两种支付工

具的价值评估差异;据此,假说4得证。在上述检验结果中,除了假说3以外,其他结果均与表6中的3个模型的估计结果相一致。

### (三) 对检验结果的进一步讨论

综合上述分析,采用选择实验方法开展环境政策效益评估时广泛存在支付工具效应,即采用不同的支付工具评估的政策效益会存在显著差异。这就意味着研究者或政策制定者开展环境价值评估或政策效益评估都需要谨慎地考虑以何种支付工具作为价值评估标准,甚至以何种支付工具作为筹资方式,而筹资方式的背后可能还涉及对政策实施主体的选择。

支付工具效应产生的第一个原因在于“非结果性”的实验设计。因为陈述偏好方法存在假想性特征,因此,为了弥补其外部效度的不足,附带“结果性”承诺被作为一种标准的实验设计规范(全世文,2016)。实验设计的“结果性”包括“支付结果性”和“政策结果性”,前者即强制支付规则,后者即实验结果对政策实施具有实际影响力,同时满足这两种“结果性”是保证实验设计激励相容的必要条件(Herriges et al., 2010)。正如前文所述,违背任意一种“结果性”的实验设计都会诱发被试者采取策略性行为,导致价值评估的结果出现偏差,而这种偏差与支付工具的设计可能会存在交互效果。非强制支付的实验设计与捐赠支付工具相结合时将导致政策效益被严重高估,而且,这两种设计对于价值高估存在显著的“互补”效果。不具有政策影响力的实验设计与捐赠支付工具相结合时则会导致政策效益被严重低估,但是,这两种设计对于价值低估未被证实存在“互补”效果。该结论是否可以推广到其他应用领域尚有待验证,但是,正如大量经验研究所示,支付工具的选择会显著地影响到实验设计的激励相容机制(例如, Bennett and Blamey, 2001),而本文说明了这种影响效果首先来源于不同的支付工具本身具有不同的“结果性”内涵。强制型支付工具内涵的“支付结果性”强于自愿型支付工具,而被试者对政策实施主体的熟悉度或了解度存在差异又导致了缴税支付工具的“政策结果性”也强于捐赠支付工具。

附带结果性承诺的实验设计虽然可以有效降低不同支付工具评估结果的差异,但是,支付工具效应并不能完全消除。具体表现为,满足“结果性”条件的第3组与第6组之间的差异虽然有所下降,但该差异依然显著。本文给出了支付工具效应的另外两个原因,一是信任度差异,二是治理责任认知差异。被试者对支付工具背后的环境政策实施主体的信任度会提高被试者对于“政策结果性”的认知,从而提高评估价值。从总体来看,因为民众对政府治理环境的信任度高于对公益组织治理环境的信任度,所以,缴税支付工具比捐赠支付工具评估的价值更高。治理责任认知差异导致的支付工具效应并不在于样本认知上的差异,而在于样本认知与支付工具之间的交互效果。如果被试者认为自己对环境治理没有货币支出方面的义务,那么,出于支付工具本身在义务性内涵上的差异,捐赠支付工具评估的价值会高于缴税支付工具。上述两个原因都可以被归纳为认知因素的差异,因此,一个简单的推论是,任何与支付工具相关的认知因素都有可能引发支付工具效应。与实验设计因素有所不同,在应用研究中,样本认知因素很少被研究者主动地采取信息干预措施。一个可能的原因是效益转移分析面临的矛盾:样本认知的改变并不意味着总体认知的改变,因此,以认知发生改变的样本价值推测总体价值必然会引起偏差。但是,这并不意味着干预认知因素缺乏政策含义,恰恰相反,为了获得政策效益评估的最大值,政策制定者应该结合目标支付工具从总体层面上设计干预措施,例如,提高全社会对政策实施主体治理大气污染的信心、强化

民众对大气污染治理的责任意识,等等。

本文的结论从另外一个视角可以总结为:捐赠支付工具对实验设计因素更加敏感(外部效度更弱),而缴税支付工具对样本认知因素更加敏感(内部效度更弱)。具体表现为,在6个实验组中,缴税支付工具的评估价值变差小于捐赠支付工具(表5、表8);而在第3组和第6组的8个子样本组中,缴税支付工具的评估价值变差大于捐赠支付工具(表8)。由于选择实验特殊的问题设计方式,许多经验研究认为选择实验比其他陈述偏好方法具有更高的外部效度(例如,Lusk and Schroeder,2004;Broadbent et al.,2010)。即使经验研究证实了“非结果性”的选择实验同样存在偏差,偏差的程度也明显小于条件价值评估方法(全世文,2016)。从本文的分析结果来看,上述结论的一个前提条件是选择强制型支付工具;当选择自愿型支付工具时,“非结果性”的选择实验在实现其外部效度时同样会存在较大的风险,实验设计者仍需重视“结果性”承诺。然而,当实验设计者采用缴税支付工具时,需要更加重视价值评估结果对样本的敏感性。传统方法在开展价值评估时会谨慎地处理被试者拒绝回答或放弃回答的情形,标准的选择实验则将这种情形归于“维持现状”选项(图1)。不同处理方式对评估结果效度的影响尚有待后续研究进行探索。但是,一个符合经验的实验设计规范的建议是,应当在开展实验以前对具体环境政策或项目的概括(包括实施主体、实施效果等)进行科学的陈述,从而减小被试者的认知差异。

## 六、结论与启示

成本效益分析是判断环境政策是否具备可行性的基本方法,效益可以补偿成本是环境政策得以实施的一个必要前提。而支付工具是环境政策效益评估中一个关键的实验设计要素,选择恰当的支付工具对于获取有效的价值评估结果至关重要。本文首先以被试者采取策略性行为作为纽带,从理论上阐述了支付工具效应产生的原因。实验设计因素、被试者对政策实施主体的信任因素和对大气污染治理的治理责任认知因素会导致不同的支付工具诱发策略性动机的概率存在差异,从而使政策效益的评估结果存在差异。而且,不同因素引起支付工具效应的机制和方向也不尽相同。

本文以北京市的大气污染治理为例,重点考察了两种支付工具,一是由政府部门主导环境政策的“缴税”,二是由公益组织主导环境项目的“捐赠”,前者代表了强制型支付工具,后者代表了自愿型支付工具。然后,本文设计了6个随机的选择实验组,在北京市常住居民调查样本的基础上估算了各组的政策或项目效益,进而通过组间对比验证了支付工具效应形成的原因。支付工具效应产生的第一个原因是实验设计因素,在不要求强制支付的实验设计中,捐赠支付工具评估的效益比缴税支付工具高出了约46%;相反,在缺乏政策影响力的实验设计中,缴税支付工具评估的效益比捐赠支付工具高出了约45%。

虽然强制支付和政策影响力的信息干预可以有效地减小两种支付工具效益评估结果的差异,但是,支付工具效应并不能完全消除。支付工具效应产生的第二个原因是民众对环境政策实施主体的信任度存在差异,信任因素对效益评估结果具有显著的促进作用,与公益组织主导的环境项目相比,更多的居民信任政府部门主导的环境政策,因此,缴税支付工具评估的效益会高于捐赠支付工具。第三个原因是治理责任认知因素,当居民认为自己不应该为污染治理承担社会义务时,捐赠支付工具评估的效益比缴税支付工具高出了约62%。

本文的分析结论意味着研究者和环境政策的制定者在开展政策效益评估时需要谨慎地选择支付工具。因为“缴税”比“捐赠”具有更强的强制支付内涵和政策影响力,所以,本文建议在没有明确的政策实施主体指向时,环境政策的效益评估应首先考虑使用以“缴税”为代表的强制型支付工具,在这种情况下,即使采用非强制支付的实验设计也不至于引起严重的偏差。而当待评估项目的实施主体明确为非政府机构时,采用以“捐赠”为代表的自愿型支付工具同样可行,但是,在这种情况下,研究者必须采用强制支付的实验设计,并通过信息干预,强调实验结果对项目实施的重要性,否则,评估结果可能会严重有偏。

无论采用何种支付工具,研究者都需要谨慎地考察效益评估结果对样本的敏感性,尤其是,当民众对支付工具及其背后的政策实施主体存在较大的认知差异时,这种敏感性会较强,而且,采用强制型支付工具评估的效益对样本的敏感性更强。本文建议在开展环境政策效益评估之前,首先向民众披露政策实施主体、实施效果、政策意义等方面的信息,从而减小民众的认知差异。其次,因为环境政策的效益在理论上来自民众的“最大支付意愿”,所以,提高民众对环境政策实施主体的信任度以及强化民众的责任意识同样必要。

## 附录

### 1.附录 1:命题 2 证明

给定  $\mathbf{B}_{-n}$ , 被试者  $n$  做出选择  $B_n$  和  $R_n$  的期望效用差异为:

$$\begin{aligned}\Delta EV_n &= [P_n(B_n, \mathbf{B}_{-n}, s_n, Z) \cdot V_n^1 + (1 - P_n(B_n, \mathbf{B}_{-n}, s_n, Z)) \cdot V_n^0] - \\ &\quad [P_n(R_n, \mathbf{B}_{-n}, s_n, Z) \cdot V_n^1 + (1 - P_n(R_n, \mathbf{B}_{-n}, s_n, Z)) \cdot V_n^0] \\ &= [P_n(B_n, \mathbf{B}_{-n}, s_n, Z) - P_n(R_n, \mathbf{B}_{-n}, s_n, Z)] \cdot (V_n^1 - V_n^0)\end{aligned}\quad (\text{F1})$$

考虑被试者  $n$  谎报真实偏好的两种情形。第一种情形为  $R_n = 0$  而  $B_n = 1$ , 此时,  $V_n^1 - V_n^0 < 0$ , 因为  $\Delta P_n / \Delta B_n > 0$ , 所以  $P_n(B_n, \mathbf{B}_{-n}, s_n, Z) - P_n(R_n, \mathbf{B}_{-n}, s_n, Z) > 0$ , 由此可得  $\Delta EV_n < 0$ , 这意味着被试者  $n$  谎报偏好会导致期望效用的下降。第二种情形为  $R_n = 1$  而  $B_n = 0$ , 此时,  $V_n^1 - V_n^0 \geq 0$ ,  $P_n(B_n, \mathbf{B}_{-n}, s_n, Z) - P_n(R_n, \mathbf{B}_{-n}, s_n, Z) < 0$ , 同样可以得到  $\Delta EV_n \leq 0$ , 即被试者  $n$  同样无法通过谎报偏好来增加期望效用。于是, 汇报真实偏好, 即  $B_n = R_n$  是被试者  $n$  的占优策略。

### 2.附录 2:命题 3 证明

条件 iii 不成立意味着选择实验不再是“结果性”的实验设计, 被试者无需为其陈述的偏好付出真实的成本, 即“政策实施”选项中设计的支付工具不具有实际意义。当被试者  $n$  认识到这一情况以后, 给定  $\mathbf{B}_{-n}$ , 其做出选择  $B_n = 0$  和  $B_n = 1$  的期望效用差异为:

$$\begin{aligned}\Delta EV_n &= [P_n(0, \mathbf{B}_{-n}, s_n, Z) \cdot V_n(M_n, \mathbf{C}, \mathbf{A}^1) + (1 - P_n(0, \mathbf{B}_{-n}, s_n, Z)) \cdot V_n(M_n, \mathbf{C}, \mathbf{A}^0)] - \\ &\quad [P_n(1, \mathbf{B}_{-n}, s_n, Z) \cdot V_n(M_n, \mathbf{C}, \mathbf{A}^1) + (1 - P_n(1, \mathbf{B}_{-n}, s_n, Z)) \cdot V_n(M_n, \mathbf{C}, \mathbf{A}^0)] \\ &= [P_n(0, \mathbf{B}_{-n}, s_n, Z) - P_n(1, \mathbf{B}_{-n}, s_n, Z)] \cdot [V_n(M_n, \mathbf{C}, \mathbf{A}^1) - V_n(M_n, \mathbf{C}, \mathbf{A}^0)]\end{aligned}\quad (\text{F2})$$

在 (F2) 式中, 支付工具  $T_n$  或  $D_n$  没有发挥作用, 无论被试者作何选择, 都不必发生实际支付。只要被试者对空气质量具有正向的偏好, 就有  $V_n(M_n, \mathbf{C}, \mathbf{A}^1) - V_n(M_n, \mathbf{C}, \mathbf{A}^0) > 0$ 。而由于  $\Delta P_n / \Delta B_n > 0$ , 所以  $P_n(0, \mathbf{B}_{-n}, s_n, Z) - P_n(1, \mathbf{B}_{-n}, s_n, Z) < 0$ 。于是,  $\Delta EV_n < 0$  恒成立, 即  $B_n = 1$  是被试者  $n$  的占优策略。

### 3.附录 3:命题 4 证明

当条件 v 修改为  $\Delta P_n / \Delta B_n = 0$  且  $P_n \neq 0$  时, 意味着被试者认为选择实验的结果并不会对环境政策制定者的实际决策造成影响, 也就是说, 政策概率函数  $P_n(\cdot) = P_n$ , 独立于  $B_n$  与  $\mathbf{B}_{-n}$ 。此时, 给定支付工具  $T_n > 0$ , 被试者  $n$  做出选择  $B_n = 0$  和  $B_n = 1$  的期望效用差异为:

$$\begin{aligned}\Delta EV_n &= [P_n \cdot V_n(M_n, \mathbf{C}, \mathbf{A}^1) + (1 - P_n) \cdot V_n(M_n, \mathbf{C}, \mathbf{A}^0)] - \\ &\quad [P_n \cdot V_n(M_n - T_n, \mathbf{C}, \mathbf{A}^1) + (1 - P_n) \cdot V_n(M_n, \mathbf{C}, \mathbf{A}^0)]\end{aligned}$$

$$= P_n \cdot [V_n(M_n, C, A^1) - V_n(M_n - T_n, C, A^1)] \quad (F3)$$

在(F3)式中,  $P_n$  外生于选择实验过程, 因此,  $\Delta EV_n$  反映了搭便车策略的收益: 当  $B_n = 0$  时, 被试者  $n$  无需支付  $T_n$  即可以  $P_n$  的概率享受  $A^1$  状态下(政策实施)的福利; 而当  $B_n = 1$  时, 被试者  $n$  必须支付  $T_n$  才可以在同等概率下享受  $A^1$  的福利。因此, 只要收入的边际效用为正, 必然有  $\Delta EV_n > 0$ , 即  $B_n = 0$  是被试者  $n$  的占优策略。

当条件  $\mathbf{v}$  修改为  $P_n = 0$  时, 意味着被试者认为实验中设计的环境政策并不真实, 没有实施的可能性, 即无论被试者作何选择, 空气污染水平始终为  $A^0$ 。此时, 被试者的目标不再是最大化期望效用, 而是确定性效用。给定支付工具  $T_n > 0$ , 被试者  $n$  做出选择  $B_n = 0$  和  $B_n = 1$  的效用差异为:

$$\Delta V_n = V_n(M_n - T_n, C, A^0) - V_n(M_n, C, A^0) < 0 \quad (F4)$$

因此,  $B_n = 0$  是被试者  $n$  的占优策略。

### 参考文献:

1. 高汉琦、牛海鹏、方国友、梅泽勇, 2011:《基于 CVM 多情景下的耕地生态效益农户支付/受偿意愿分析——以河南省焦作市为例》,《资源科学》第 11 期。
2. 黎文靖、郑曼妮, 2016:《空气污染的治理机制及其作用效果——来自地级市的经验数据》,《中国工业经济》第 4 期。
3. 刘海凤、郭秀锐、毛显强、金建君, 2011:《应用 CVM 方法估算城市居民对低碳电力的支付意愿》,《中国人口·资源与环境》第 S2 期。
4. 全世文、秦光远、王昌海, 2018:《北京市城市湿地价值评估》,《中国人口·资源与环境》第 7 期。
5. 全世文, 2016:《选择实验方法研究进展》,《经济学动态》第 1 期。
6. 全世文、黄波, 2016:《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》,《中国工业经济》第 8 期。
7. Akcura, E. 2013. "Mandatory vs Voluntary Payment for Green Electricity." Energy Policy Research Group Working Paper, No.1316. University of Cambridge.
8. Andreoni, J. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *The Economic Journal* 100(401): 464-477.
9. Bateman, I.J., R.T. Carson, B. Day, W.M. Hanemann, N. Hanley, T. Hett, M.J. Lee, G. Loomes, S. Mourato, E. Özdemiroglu, D.W. Pearce, R. Sugden, and J. Swanson. 2002. *Economic Valuation with Stated Preference Techniques: A Manual*. Cheltenham, UK: Edward Elgar.
10. Bennett, J., and R. Blamey. 2001. *The Choice Modelling Approach to Environmental Valuation*. Cheltenham, UK & Northampton, MA, USA: Edward Elgar Publishing.
11. Broadbent, C.D., J.B. Grandy, and R.P. Berrens. 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.
12. Carlsson, F., and 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.
13. Carson, R.T. 1997. "Contingent Valuation Surveys and Tests of Insensitivity to Scope." In *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*. Edited by R. J. Kopp, W. W. Pommerehne and N. Schwarz, 127-163. Netherlands: Springer.
14. Carson, R. T., and T. Groves. 2007. "Incentive and Informational Properties of Preference Questions." *Environmental and Resource Economics* 37(1): 181-210.
15. Gao, Z., and T.C. Schroeder. 2009. "Effects of Label Information on Consumer Willingness-to-Pay for Food Attributes." *American Journal of Agricultural Economics* 91(3): 795-809.
16. Greenley, D.A., R.G. Walsh, and R.A. Young. 1981. "Option Value: Empirical Evidence from a Case Study of

- Recreation and Water Quality.” *The Quarterly Journal of Economics* 96(4): 657–673.
- 17.Herriges, J., C.Kling, C. Liu, and J.Tobias.2010.“What Are the Consequences of Consequentiality?” *Journal of Environmental Economics and Management* 59(1): 67–81.
- 18.Huber, J., and K.Zwerina.1996.“The Importance of Utility Balance in Efficient Choice Design.” *Journal of Marketing Research* 33(3): 307–317.
- 19.Ivehammar, P.2009.“The Payment Vehicle Used in CV Studies of Environmental Goods Does Matter.” *Journal of Agricultural and Resource Economics* 34(3): 450–463.
- 20.Krinsky, I., and A.L.Robb.1986.“On Approximating the Statistical Properties of Elasticities.” *The Review of Economics and Statistics* 68(4): 715–719.
- 21.Lusk, J.L., and 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.
- 22.Mitchell, R.C., and R.T.Carson.1989.*Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington D.C., USA: Resources for the Future Press.
- 23.Morrison, M.D., R.K.Blamey, and J.W.Bennett.2000.“Minimising Payment Vehicle Bias in Contingent Valuation Studies.” *Environmental and Resource Economics* 16(4): 407–422.
- 24.Poe, G.L., K.L.Giraud, and J.B.Loomis.2005.“Computational Methods for Measuring the Difference of Empirical Distributions.” *American Journal of Agricultural Economics* 87(2): 353–365.
- 25.Tan, J.,and J.Zhao.2014.“The Value of Clean Air in China: Evidence from Beijing and Shanghai.” *Frontiers of Economics in China* 9(1): 109–137.
- 26.Train, K.E.2003.*Discrete Choice Methods with Simulation*.New York, USA: Cambridge University Press.
- 27.Vossler, C.A., M.Doyon, and D.Rondeau.2012.“Truth in Consequentiality: Theory and Field Evidence on Discrete Choice Experiments.” *American Economic Journal: Microeconomics* 4(4): 145–171.

## Tax or Donation? Choice of Payment Vehicles in Environmental Policy Evaluation

Quan Shiwen

(Rural Development Institute, Chinese Academy of Social Sciences)

**Abstract:** Payment vehicle effect (PVE) in benefit analysis of environmental policy has been extensively found in empirical studies. However, the reason of PVE still remains controversial. This paper first provides a theoretical explanation of PVE from the perspective of strategic behaviors. Taking Beijing’s air pollution control policy as example, this paper designs several random experimental groups, makes a comparative analysis between evaluation results of tax and donation vehicles. The results confirm that inconsequential experimental design, different degree of trust between people and policy implementer, and disidentification of environmental governance responsibility could result in PVE. Generally, donation vehicle is more sensitive to experiment design, while tax vehicle is more sensitive to sample perception. The results imply that when conducting benefit analysis of environmental policies, researchers or policy makers should choose payment vehicle discreetly. Different vehicles require different standards on experimental design and sample to achieve its validity.

**Keywords:** Environmental Policy, Payment Vehicle, Choice Experiment, Air Pollution, Valuation

**JEL Classification:** Q51, Q58, C91

(责任编辑:陈永清)