arXiv:2303.07488v2 [econ.TH] 7 Oct 2023

一致性的故事：弥合边沁与罗尔斯之间的鸿沟基础

Stéphane Gonzalez [1] 和 Nikolaos Pnevmatikos[2]

[1]里昂大学、圣艾蒂安让·莫内大学、GATE UMR 5824，F-42023圣艾蒂安，法国，stephane.gonzalez@univ-st-etienne.fr，[2]LEMMA，巴黎第二大学先贤祠-阿萨斯大学，Blaise Desgoffe 街 4 号，巴黎，法国，nikolaos.pnevmatikos@u-paris2.fr

2023年10月10日版本

摘要

本文在基数偏好的更广阔领域内探讨了边沁解与罗尔斯解的公理基础。研究揭示，这两种解概念都满足以下四条公理：非空性、匿名性、一致同意和连续性。为了充分刻画边沁解与罗尔斯解，本文提出了一致性准则的三种变体，并评估它们与其他公理的相容性。每一种一致性的表述都可以解释为不确定环境下决策的一种性质。

关键词：功利主义、平等主义、公理化、一致性

JEL分类：D71

1 引言

社会选择理论广泛研究了如何设计方法，以比较和评估特定社会可选的各种备选方案。社会福利准则常用于衡量每种备选方案对整个社会带来的效用或益处。道德哲学提出的两种准则影响尤为深远：边沁的(1789)功利主义准则（以下称边沁解），该准则考量所有个体的效用，并选择效用总和最大的备选方案；以及罗尔斯的(1971)平等主义准则（以下称罗尔斯解），该准则主张选择能为社会中处境最差的成员带来最大益处的备选方案。

在公平分配问题上，这两个概念的态度截然不同。功利主义对效用不平等持中立态度，而平等主义原则只关注处境最差的个体，因此极度厌恶不平等。这两种路径催生了社会选择领域的大量研究，其中包括试图依据理想的规范原则，为其价值提供依据的公理化刻画。Harsanyi (1955）关于功利主义的聚合结果在这方面具有里程碑意义。由此产生了一个重要的研究分支，专门探讨边沁解的公理化，尤其关注不确定性条件下的集体决策以及

传统经济分配问题。沿着这一思路，值得关注的例子包括 d'Aspremont 和 Gevers（1977）、Maskin（1978）、Blackorby 和 Donaldson（1982）、Mongin（1994）、Weymark（1990）以及 Weymark（2005）。相比之下，文献中对罗尔斯解进行公理化刻画的研究相对较少；例如，Hammond（1979）、d'Aspremont 和 Gevers（1977）、Gevers（1979）、Ou-Yang（2018）以及 Mongin 和 Pivato（2021）的研究。关于边沁解与罗尔斯解的公理化，读者可参阅 d'Aspremont 和 Gevers（2002）或 Kamaga（2018）的详尽综述。

首次尝试比较或调和边沁与罗尔斯的观点，旨在识别出能同时刻画这两种方案的公理，可见于d'Aspremont 和 Gevers（1977）的研究。作者指出，在社会福利泛函的背景下，功利主义与罗尔斯准则的词典序版本具有无关选项独立性、匿名性和强帕累托原则这几项共同性质，二者仅有一条公理不同：功利主义需要效用的个体起点，而平等主义则要求极端公平公理。在福利主义背景下也有类似结果：两种方案的刻画均可基于强帕累托原则、可分离性、匿名性以及适当的信息不变性公理推导得出；d'Aspremont 等人（1994年）指出，可以借助 Hammond（1976年）和 Strasnick（1976年）给出的词典序极小化的其他刻画证明一个定理（参见 d'Aspremont（1985年）或 d'Aspremont 与 Gevers（1977年）的证明）。这些刻画表明，归根结底，词典序极小化与经典功利主义之间唯一的区别，在于二者所依赖的信息类型和人际比较类型：前者以共序数性取代序数不可比性，而功利主义则要求基数性和单位可比性。

此后，许多文章沿着这一思路为这两种解提供公理化，例如 Sen (1979年)、Deschamps 和 Gevers (1978年)、Myerson (1981年)，或者旨在确定确保这两种解相交的条件，例如 Kamaga (2018年)。我们还提及 Bossert 和 Kamaga (2020)，他们研究这两种原则的凸组合，旨在从公理角度证明其合理性；此外，Gajdos 和 Kandil(2008年) 刻画了功利主义与平等主义准则的一种凸组合，其中个体的效用是基数可测且完全可比的[1]。更早些时候，Roemer (1996) 和 Roemer (1998) 将一种用于构建机会均等政策的机制形式化，这种机制可视为功利主义与平等主义准则之间的折衷。近来，Silvestre (2014) 分析了偏好为准线性或 Gorman 型偏好的消费者，并表明在某些技术条件下，功利主义最大值意味着效用相等。

对主体偏好的限制可能会成为争论的话题，不同观点会探讨各种方法的优缺点。其中一个争论焦点是，究竟应使用序数偏好还是效用函数来表示主体的偏好。

另一方面，序数偏好常被视为一种精简的结构，因为它只要求主体能够对备选项进行排序。这种简洁性有助于更好地理解和诠释在此框架下的模型所得出的结果。此外，序数偏好在文献中得到了广泛应用，已有大量结果和方法可用于这一框架。另一方面，效用函数的优势在于能够表达每个个体偏好的强度。这些额外信息能让我们更为详尽、细致地理解主体如何做出决策。与后者方向相关的重要论文包括 Maskin (1978年), Kamaga(2018年), Blackorby 和 Donaldson (1982). 基数偏好使我们能够得出与序数偏好框架下相似的结果，而这些结果在数学上更容易确立。利用效用函数表达偏好，有助于深入探讨

[1]作者得到的效用函数的归一化方式，与以下文献中出现的方式非常相似：Dhillon 和 Mertens (1999年)、Segal (2000)以及Moreno-Ternero 和 Roemer (2008年)。

决策问题，最终更深入地洞察主体如何做出决策。

本文旨在效用剖面的框架下，为边沁最优解和罗尔斯最大最小解提供一些新的公理化刻画，从而采用一种不同于以往研究的方法——以往研究通常依赖序数偏好剖面。我们的研究设定较为简单，只需具备关于效用函数剖面的基础知识和初等数学工具。

我们的意图是聚焦于两个通常被视为相互冲突、但从公理化角度看却颇为相似的解概念。具体而言，我们表明，边沁解与罗尔斯解共有四条公理，它们在社会选择理论文献中广为人知：非空性、匿名性、一致同意和连续性。非空性确保至少有一个备选方案被选为解。匿名性保证所有主体受到平等对待。一致同意确保如果每个主体都偏好某个备选方案，那么该方案就会被选为解；这对应于帕累托最优概念的一个弱化版本。连续性要求解在效用剖面发生微小扰动时保持稳定。

我们提出一种一致性原则：如果两个不同的效用剖面都将同一个备选方案视为解，那么该备选方案也应当是聚合效用剖面中的解。该规则类似于混合保持性质，由Harsanyi (1955年) 提出，并广泛应用于刻画冯·诺依曼和摩根斯坦效用函数的文献中。该性质要求输入的凸组合所对应的聚合效用等于输出的凸组合，这是基数效用数学框架的一项基本性质。我们的原则与一致性性质密切相关，该性质最近由Brandl 和 Brandt(2019) 在双人零和博弈中进行了研究。在这一设定下，一致性通过如下要求来确立：如果两个参与者在给定的两个博弈中都选择同一策略对，那么在一个新博弈中也应采用该策略对，而新博弈的收益由原先两个博弈的收益随机化决定。

本文始终考虑这样一种情形：主体所面对的效用剖面取决于某个随机事件。我们根据主体对各种信念的类型，提出一致性的三种变体；主体据此认为每种效用剖面都有可能出现。

a) 第一种变体称为主观期望一致性。如果某个备选方案在两个不同的效用剖面中都被选中，那么无论主体对各个剖面出现的看法可能存在多大差异，该备选方案也应在期望效用剖面中被选中。当不同个体对某一特定事件发生可能性的个人信念或看法各不相同，且这些信念或看法受到个体的经历、知识和偏好影响时，这一原则似乎颇具说服力。具体而言，给定两个效用剖面，以及每个个体对它们赋予的概率 $p_i \quad i \in N$，期望效用剖面在 $\mathbf{p} = (p_1,\ldots,p_n)$ 下的解必须与这两个效用剖面解集的交集一致，除非该交集为空。主观期望一致性原则以主观概率这一概念为基础，也就是说，主观概率是根据个体对某一特定事件是否可能发生的个人判断或自身经验得出的一种概率。

b) 第二种变体称为客观期望一致性。它将前一种变体的思想限定于客观概率的情形。它与这样一种情况尤其相关：效用剖面由抛硬币等随机事件决定。它指出，如果某个备选方案在两个效用剖面中都被选中，那么在抛硬币之前也应选中该备选方案。形式上，给定两个效用剖面以及它们之间的概率 p，期望效用剖面的解集在 p 下必须

与这两个效用剖面解集的交集相同，除非该交集为空。客观期望一致性原则基于客观概率这一概念，简单来说，客观概率是与客观可观察事实相关、且独立于任何个体的信念或偏好的概率。

c) 一致性原则的第三种变体称为最小一致性，它用最坏情形剖面取代期望效用剖面。该原则规定：如果某个备选方案在两个效用剖面中都被选中，那么在这两个效用剖面的最坏情形下，该方案也应被选中。在最坏情形剖面中，对于任一备选方案，每个个体的效用都定义为两个效用剖面中较低的那个。尤其是，当个体对随机事件发生的概率没有任何判断时，这种变体很有用，因为它使人们能够在不依赖主观期望的不确定情境下进行决策。

示例说明

为了说明我们的结果，考虑这样一种情形：两位主体 Jeremy 和 John 合作开展一个研究项目。为此，他们在牛津会面并共同工作。他们可以选择在大学的实验室（ℓ）或附近的咖啡馆（c）工作。两位主体对工作地点的偏好取决于天气，可用他们对两个可能备选方案的效用表示。天气晴朗时，两位主体的效用如下：

$$\mathbf{u} = \begin{bmatrix} u_1(\ell) & u_1(c) \\ u_2(\ell) & u_2(c) \end{bmatrix}$$

下雨时，他们的效用为：

$$\mathbf{v} = \begin{bmatrix} v_1(\ell) & v_1(c) \\ v_2(\ell) & v_2(c) \end{bmatrix}$$

为了做出选择，Jeremy 和 John 需要设计一个决策过程 φ，将每个效用剖面$\mathbf{u}$ 与一个备选方案 $\varphi(\mathbf{u})$ 相关联，即实验室 (ℓ) 或咖啡馆 (c)。[2] 这两个主体认同此前讨论过的以下性质：非空性、匿名性、一致同意和连续性。最后一项理想标准——一致性——取决于三种不同的情境：

第一种情境。Jeremy 和 John 对他们逗留牛津期间的天气看法不同。Jeremy 认为天气晴朗的概率为 p_1，下雨的概率为 $1-p_1$；John 认为天气晴朗的概率为 p_2，下雨的概率为 $1-p_2$，其中 $p_1 \neq p_2$。他们的期望效用可以用以下矩阵表示：

$$\mathbf{p}_{\{1,2\}} \cdot \mathbf{u} + (I - \mathbf{p}_{\{1,2\}}) \cdot \mathbf{v} = \begin{bmatrix} p_1u_1(\ell) + (1-p_1)v_1(\ell) & p_1u_1(c) + (1-p_1)v_1(c) \\ p_2u_2(\ell) + (1-p_2)v_2(\ell) & p_2u_2(c) + (1-p_2)v_2(c) \end{bmatrix}$$

在这种情况下，主观期望一致性似乎是一种理想的属性。与我们的例子一致，如果 Jeremy 和 John 无论天气如何都同意去实验室，那么对于任何可能的天气信念，他们也应该选择相同的备选方案。也就是说，如果天气晴朗时选择备选方案 (ℓ)（分别为 (c)），即 ℓ（分别为 c）属于

[2] 我们允许 φ 在以下情况下返回两个备选项 tie.

to $\varphi(\mathbf{u})$，且天气下雨时选择 ℓ（分别为 c），即 ℓ（分别为 c）属于$\varphi(\mathbf{v})$，那么对于期望效用剖面 $\mathbf{p}^{\{1,2\}}\cdot\mathbf{u}+(I-\mathbf{p}^{\{1,2\}})\cdot\mathbf{v}$，$\ell$（分别为 c）也会被 φ 选中。形式化地说，

$$\varphi(\mathbf{u})\cap\varphi(\mathbf{v})\subseteq\varphi(\mathbf{p}_{\{1,2\}}\cdot\mathbf{u}+(I-\mathbf{p}_{\{1,2\}})\cdot\mathbf{v}).$$

第二种情境。 Jeremy 和 John 根据天气预报行事；预报显示，晴天的概率为 p，下雨的概率为 $1-p$。他们的期望效用可用以下矩阵表示：

$$p\mathbf{u}+(1-p)\mathbf{v}=\begin{bmatrix}pu_1(\ell)+(1-p)v_1(\ell) & pu_1(c)+(1-p)v_1(c)\\ pu_2(\ell)+(1-p)v_2(\ell) & pu_2(c)+(1-p)v_2(c)\end{bmatrix}$$

如果是这样，客观期望一致性似乎是一项理想的性质。也就是说，如果无论天气如何，Jeremy 和 John 都同意去实验室，那么对于天气的任何客观概率，他们也应该选择相同的备选方案。请注意，这项性质对应于一种较弱形式的主观期望效用，因为只有在 $p_1=p_2$ 时才要求决策保持一致。沿着同样的论证思路，我们得到：

$$\varphi(\mathbf{u})\cap\varphi(\mathbf{v})\subseteq\varphi(p\mathbf{u}+(1-p)\mathbf{v}).$$

第三种情境。 Jeremy 和 John 对天气一无所知，因此没有为两种可能事件指定任何概率。面对这种情况，Jeremy 和 John 采取了更为悲观的方法，比较每种备选方案下各自的效用，并考虑最坏情形剖面。这个效用剖面可用以下矩阵表示：

$$\mathbf{u}\wedge\mathbf{v}=\begin{bmatrix}u_1(\ell)\wedge v_1(\ell) & u_1(c)\wedge v_1(c)\\ u_2(\ell)\wedge v_2(\ell) & u_2(c)\wedge v_2(c)\end{bmatrix}$$

其中，对于任意实数 a 和 b，记号 $a\wedge b$ 表示 a 和 b 中的较小值。在这种情况下，最小一致性似乎是一项理想的性质。也就是说，如果不论天气如何，Jeremy 和 John 都同意去实验室，那么在最坏情形剖面中，他们也应该选择相同的备选方案。以同样的方式，我们得到：

$$\varphi(\mathbf{u})\cap\varphi(\mathbf{v})\subseteq\varphi(\mathbf{u}\wedge\mathbf{v}).$$

Harsanyi (1955年）的聚合定理旨在建立一个社会效用函数，用以汇总主体的序数偏好。该定理保证，只要主体遵循冯·诺依曼和摩根斯坦期望效用模型（下文简称冯·诺伊曼—摩根斯坦期望效用模型），那么社会效用函数也符合该模型，并且只要它关于个体的效用函数是仿射的，就满足帕累托条件。该定理一直备受讨论，衍生出许多版本和扩展，主要涉及证明中的技术细节，或探讨结果仍然成立的适用范围[3]。学界曾多次尝试简化这一结果的证明，例如Hammond (1992年）和Mandler (2005年所做的工作。尽管付出了这些努力，Harsanyi (1955年）的定理仍然需要相当繁复的证明，理解起来颇为困难。定理1立足于基数偏好，与这一结果有相似之处，而且更容易证明。我们不构造社会效用函数，而是

[3]关于这些争论的文献综述，请参阅Mongin (1994年)。

提出设计一个选择函数；对于该选择函数，我们所提出的非空性、客观期望一致性和连续性公理，发挥着类似于社会效用函数所需的冯·诺伊曼—摩根斯坦期望效用模型假设的作用，而我们提出的一致同意假设则取代了帕累托条件。最后，我们的匿名性公理与Harsanyi (1955年）的结果一样，导出一条平等对待所有主体的规则。

相反，如果个体遵循的不是 vNM 期望效用模型，而是最大最小规则，那么，沿用类似的论证，并以设计选择函数为目的，我们将证明，这种聚合方式应当符合最小一致性原则。后者也体现在定理2中，该定理同样是在基数偏好的背景下确立的。换言之，定理2 表明，遵循最大最小规则的个体与这样一种选择函数相一致：该函数将最大最小规则应用于最坏情形下的效用剖面。据我们所知，文献中从未在平等主义原则的框架下采用过这一做法。

本文给出三个主要结果，分别对应这三种不同的参照框架。命题1表明，以下性质无法同时兼容：主观期望一致性、非空性、匿名性、一致同意以及连续性。定理1 和定理2 克服了这一限制。一方面，定理1通过以客观期望一致性取代主观期望一致性，刻画了边沁解。另一方面，定理2通过采用最小一致性而非客观期望一致性，刻画了罗尔斯解。定理1 的证明与定理2 的证明类似。

文章结构：第2 节介绍形式化定义和符号。在第3节，我们陈述并证明主要结果。第4节给出结论与评述。

2 预备知识

本节首先介绍基本符号，并给出边沁解概念和罗尔斯解概念的定义。随后，我们提出并讨论这两种解概念共同满足的公理。

2.1 定义和符号

对于任意 $n \geq 2$，令 $N = \{1, \ldots, n\}$ 为主体的有限集合，并令 S 为社会备选方案（例如状态、候选人、行动剖面）的有限集合。给定有限集合 J 和实数 a_j，对于任意 $j \in J$，记号 $\bigwedge_{j \in J} a_j$ 表示 $\{a_j : j \in J\}$ 中的最小值。每个个体 $i \in N$ 都配有一个效用函数 $u_i : S \to \mathrm{D}$，其中 D 是 $\mathbb{R}$ 的非空凸子集 [4]。我们进一步用 $\mathbf{u} = (u_i)_{i \in N}$ 表示由此导出的效用剖面，并用 $\mathrm{U} = (\mathrm{D}^S)^N$ 表示所有可能效用剖面的集合。

定义在 U 上的解概念 是一个对应 $\varphi : \mathrm{U} \rightrightarrows S$，即 φ 为每个效用剖面 $\mathbf{u} \in \mathrm{U}$ 指定一个社会备选项集合 S 的子集 $\varphi(\mathbf{u})$。给定 $\mathbf{u} \in \mathrm{U}$，边沁解 $B : \mathrm{U} \rightrightarrows S$ 定义为

$$B(\mathbf{u}) = \operatorname*{argmax}_{s \in S} \sum_{i \in N} u_i(s)$$

[4] 凸性假设是获得第3.1 节和3.2 节证明所必需的，但在第3.3节的证明中可以放宽。

此外，给定$\mathbf{u} \in \mathrm{U}$，罗尔斯解 $R: \mathrm{U} \rightrightarrows S$ 定义为

$$R(\mathbf{u}) = \operatorname*{argmax}_{s \in S} \bigwedge_{i \in N} u_i(s).$$

2.2 常见公理性质

本节介绍公理化这两种解概念所涉及的公理。第一条公理通常构成社会选择函数定义的一部分。

非空性：对于任意$\mathbf{u} \in \mathrm{U}$，均有 $\varphi(\mathbf{u}) \neq \emptyset$。

我们用记号 $\mathfrak{S}(N)$表示 n 参与者的所有排列组成的集合。对于任意 $\sigma \in \mathfrak{S}(N)$，令 $u_i^\sigma(s) = u_{\sigma(i)}(s)$。诱导出的效用剖面记为$\mathbf{u}_\sigma = (u_i^\sigma)$。第二条公理确保所有个体都受到平等对待。

匿名性：对于任意$\mathbf{u} \in \mathrm{U}$，均有 $\varphi(\mathbf{u}) = \varphi(\mathbf{u}_\sigma)$。

以下公理弱于标准的帕累托最优性公理。它指出，只要存在备选项受到所有人一致偏好，解就必须选择其中一个。给定$\mathbf{u} \in \mathrm{U}$，令

$$M(\mathbf{u}) = \{s^* \in S : \forall i \in N, \forall s \in S, u_i(s^*) \geq u_i(s)\}$$

一致同意：如果 $M(\mathbf{u}) \neq \emptyset$，那么 $\varphi(\mathbf{u}) \subseteq M(\mathbf{u})$。

下一个公理确保，给定一个效用剖面，我们总能找到另一个与其足够接近、且与原问题共享解的效用剖面。

连续性：对于每个 $a \in S$，集合 $\mathrm{U}_a = \{\mathbf{u} : \ a \in \varphi(\mathbf{u})\}$ 是闭集。

3 主要结果

本节将展示我们的主要结果。首先，我们给出一个不可能性结果，指出不存在任何解概念能够同时满足非空性、匿名性、一致同意和主观期望一致性。随后，我们通过在两种解中加入连续性公理，并分别以客观期望一致性公理和最小一致性公理取代最后一项一致性性质，从公理角度刻画边沁解与罗尔斯解。

3.1 不可能性结果

为正式引入一致性公理的第一种变体，我们用 P_N 表示以下形式的对角矩阵集合：

$$\mathbf{p}_N = \begin{bmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_n \end{bmatrix},$$

其中 $p_i \in (0,\ 1)$ 对每个 $i \in N$。我们将 P_N 中的矩阵记为 I，其中对于任意 $i \in N$，$p_i = 1$.

如引言所述，主观期望一致性公理要求在主体利用由 P_N 中的矩阵表示的主观概率时，集体行为保持一致。我们考虑两个不同的效用剖面 $\mathbf{u}$ 和$\mathbf{v}$ ，使得推荐给主体的备选方案集合存在交集，并假设每个个体 $i \in N$ 都认为$\mathbf{u}$ 发生的概率为 p_i，而$\mathbf{v}$ 发生的概率为 $1 - p_i$。那么，主观期望一致性指出，由于在两个效用剖面下都存在一个为主体所选的备选方案，因此无论个体概率信念如何，该备选方案都应当被选中。该原则可以形式化如下：

主观期望一致性：对于所有$\mathbf{u}$，$\mathbf{v} \in \mathrm{U}$，若 $\varphi(\mathbf{u}) \cap \varphi(\mathbf{v}) \neq \emptyset$，则对于所有$\mathbf{p}_N \in \mathrm{P}_N$：

$$\varphi(\mathbf{u}) \cap \varphi(\mathbf{v}) \subseteq \varphi\left(\mathbf{p}_N \cdot \mathbf{u} + (I - \mathbf{p}_N) \cdot \mathbf{v}\right).^{5}$$

现在我们已经具备了提出命题 1 的条件。

命题 1. 不存在同时满足非空性、匿名性、一致性和主观期望一致性的解概念。

证明. 为导出矛盾，假设存在一个解 φ，满足该定理的公理。为了证明我们的不可能性结论，只需给出一个情形示例，说明这种 φ 的存在会导致矛盾。为此，我们回到引言中列出的 Jeremy 和 John 的例子，并假设$u_1(\ell) = v_2(\ell) = u_2(c) = v_1(c) = \alpha$ 以及 $u_1(c) = v_2(c) = u_2(\ell) = v_1(\ell) = \beta$。因此，我们有：

$$\mathbf{u} = \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix} \text{ and } \mathbf{v} = \begin{bmatrix} \beta & \alpha \\ \alpha & \beta \end{bmatrix}$$

请注意，$\mathbf{v}$ 对应于主体效用的置换。于是，利用匿名性，我们得到 $\varphi(\mathbf{u}) = \varphi(\mathbf{v})$，而由非空性显然可知：

$$\varphi(\mathbf{u}) \cap \varphi(\mathbf{v}) \neq \emptyset$$

因此，根据主观期望一致性，我们得到：

$$\varphi(\mathbf{u}) \cap \varphi(\mathbf{v}) \subseteq \varphi\left(\begin{bmatrix} p_1\alpha + (1-p_1)\beta & p_1\beta + (1-p_1)\alpha \\ p_2\beta + (1-p_2)\alpha & p_2\alpha + (1-p_2)\beta \end{bmatrix}\right)$$

[5]给定$\mathbf{w} \in \mathrm{U}$ 和$\mathbf{q}_N \in \mathrm{P}_N$，记号$\mathbf{q}_N \cdot \mathbf{w}$ 表示$\mathbf{q}_N$ 和$\mathbf{w}$的标准矩阵乘积。

反转 $\mathbf{u}$ 和$\mathbf{v}$ 的作用，并再次使用主观期望一致性，我们还可以得到 e:

$$\varphi(\mathbf{u}) \cap \varphi(\mathbf{v}) \subseteq \varphi\left(\begin{bmatrix} p_1\beta + (1-p_1)\alpha & p_1\alpha + (1-p_1)\beta \\ p_2\alpha + (1-p_2)\beta & p_2\beta + (1-p_2)\alpha \end{bmatrix}\right)$$

合并最后三个公式，我们得到：

$$\varphi\left(\begin{bmatrix} p_1\alpha + (1-p_1)\beta & p_1\beta + (1-p_1)\alpha \\ p_2\beta + (1-p_2)\alpha & p_2\alpha + (1-p_2)\beta \end{bmatrix}\right) \cap \varphi\left(\begin{bmatrix} p_1\beta + (1-p_1)\alpha & p_1\alpha + (1-p_1)\beta \\ p_2\alpha + (1-p_2)\beta & p_2\beta + (1-p_2)\alpha \end{bmatrix}\right) \neq \emptyset \quad (1)$$

现在假设 $p_1 \in (1/2, 1]$, $p_2 \in [0, 1/2)$ 且 $\alpha \geq \beta$。于是，我们得到：

$$\begin{cases} p_1\alpha + (1-p_1)\beta \geq p_1\beta + (1-p_1)\alpha \\ p_2\beta + (1-p_2)\alpha \geq p_2\alpha + (1-p_2)\beta \end{cases}$$

因此可以推出，在该期望效用剖面中：

$$\begin{bmatrix} p_1\alpha + (1-p_1)\beta & p_1\beta + (1-p_1)\alpha \\ p_2\beta + (1-p_2)\alpha & p_2\alpha + (1-p_2)\beta \end{bmatrix}$$

示例中与备选方案（ℓ）对应的左列受到一致同意的偏好。相反，在以下期望效用剖面中：

$$\begin{bmatrix} p_1\beta + (1-p_1)\alpha & p_1\alpha + (1-p_1)\beta \\ p_2\alpha + (1-p_2)\beta & p_2\beta + (1-p_2)\alpha \end{bmatrix}$$

示例中与备选方案（c）对应的右列受到一致同意的偏好。因此，根据一致同意和非空性，我们得到：

$$\varphi\left(\begin{bmatrix} p_1\alpha + (1-p_1)\beta & p_1\beta + (1-p_1)\alpha \\ p_2\beta + (1-p_2)\alpha & p_2\alpha + (1-p_2)\beta \end{bmatrix}\right) = \{\ell\}$$

此外，

$$\varphi\left(\begin{bmatrix} p_1\beta + (1-p_1)\alpha & p_1\alpha + (1-p_1)\beta \\ p_2\alpha + (1-p_2)\beta & p_2\beta + (1-p_2)\alpha \end{bmatrix}\right) = \{c\}$$

因此显然可以得出：

$$\varphi\left(\begin{bmatrix} p_1\alpha + (1-p_1)\beta & p_1\beta + (1-p_1)\alpha \\ p_2\beta + (1-p_2)\alpha & p_2\alpha + (1-p_2)\beta \end{bmatrix}\right) \cap \varphi\left(\begin{bmatrix} p_1\beta + (1-p_1)\alpha & p_1\alpha + (1-p_1)\beta \\ p_2\alpha + (1-p_2)\beta & p_2\beta + (1-p_2)\alpha \end{bmatrix}\right) = \emptyset$$

这与公式（1）相矛盾。 ■

命题1 的一个重要假设是主体信念具有异质性。考察相关文献可知，Hylland 和 Zeckhauser(1979年) 针对基数偏好，以及 Mongin（1995年) 针对序数偏好，均表明，当主体信念各不相同时，主观期望效用与帕累托最优性会导致不可能性结果。命题1 指出，与帕累托最优性紧密相关的性质——一致同意公理——与一种混合保持性质——主观期望一致性——之间存在不相容性。这一结果与在序数偏好框架下确立的其他不可能性结果保持一致，而那些结果的证明显然更具挑战性。

此后，我们将研究 $p_1 = p_2 = ... = p_n$ 的情形，这可以看作主体对可能发生的事件具有客观概率的情况。

3.2 边沁解的公理化刻画

本节中，我们将使用非空性、一致同意、匿名性、连续性公理，以及一致性公理的一种变体——客观期望一致性，对边沁解进行公理化刻画。后一性质基于一个解概念在相关效用剖面中的一致行为。

为此，我们研究两个推荐给主体的备选项集合存在交集的效用剖面。考虑这样一种情形：通过抛硬币决定采用两个效用剖面中的哪一个。那么，客观期望一致性规定，由于两个效用剖面都为主体选出了某个备选项，因此在抛硬币之前也应选出该备选项。也就是说，给定两个效用剖面及其上的概率 p，其期望效用剖面的解集在 p 包含这两个解集的交集，除非交集为空。形式化地说，

客观期望一致性：对于所有 $\mathbf{u}, \mathbf{v} \in \mathrm{U}$，若 $\varphi(\mathbf{u}) \cap \varphi(\mathbf{v}) \neq \emptyset$，则对于所有 $p \in (0,1)$：

$$\varphi(\mathbf{u}) \cap \varphi(\mathbf{v}) \subseteq \varphi(p\mathbf{u} + (1-p)\mathbf{v})$$

引理 1 将用于定理1 的证明。

引理 1. 假设 φ 满足客观期望一致性。令 $m \geq 2$，并考虑$\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m \in \mathrm{U}$。若存在 $s \in S$，使得对每个 $i \in \{1, \ldots, m\}$，都有 $s \in \varphi(\mathbf{u}_i)$，则

$$s \in \varphi\left(\frac{1}{m} \sum_{i \in \{1,\ldots,m\}} \mathbf{u}_i\right).$$

证明。 给定 $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m \in \mathrm{U}$，假设存在 $s \in S$，使得对于任意 $i \in \{1, \ldots, m\}$,$s \in \varphi(\mathbf{u}_i)$。我们对 m 使用归纳法。若 $m = 2$，则由客观期望一致性可直接得出结论。假设该结论对 $m > 2$ 成立。于是，

$$s \in \varphi\left(\frac{1}{m} \sum_{i \in \{1,\ldots,m\}} \mathbf{u}_i\right).$$

令 $\mathbf{w} = \frac{1}{m} \sum_{i \in \{1,\ldots,m\}} \mathbf{u}_i$，并考虑 $\mathbf{u}_{m+1} \in \mathrm{U}$，使得 $s \in \varphi(\mathbf{u}_{m+1})$。因此，根据客观期望一致性，我们得到

$$s \in \varphi\left(\frac{m}{m+1}\mathbf{w} + \frac{1}{m+1}\mathbf{u}_{m+1}\right) = \varphi\left(\frac{1}{m+1} \sum_{i \in \{1,\ldots,m+1\}} \mathbf{u}_i\right)$$

由此完成证明。 ■

在其开创性论文中，Harsanyi (1955年) 提出了一种功利主义的公理化基础，其中一项与客观期望一致性 密切相关的准则与帕累托最优性共同蕴含：社会效用是个体效用的组合。

随后，我们给出定理1，对边沁解作公理化刻画。它显然带有Harsanyi (1955年）关于序数偏好聚合结果的几分神韵；如前所述，该结果的证明相当困难。而在基数偏好的语境下，我们的证明则更为简洁。

定理1. 一个解概念满足非空性、匿名性、一致同意、连续性和客观期望一致性，当且仅当它是边沁解。

证明。 我们只证明这些公理的充分性。证明分两步进行。

步骤 1。 如果 φ 满足匿名性、一致同意和客观期望一致性，那么我们有$\varphi(\boldsymbol{u}) \subseteq B(\boldsymbol{u})$.

设 a属于 $\varphi(\mathbf{u})$，并设 $\sigma \in \mathfrak{S}(N)$ 是定义在 N上的循环置换： $\sigma(1) = 2, \sigma(2) = 3, \ldots, \sigma(n-1) = n$，以及 $\sigma(n) = 1$。令 σ^k表示复合 σ k次后得到的置换，使得若 $k + r \leq n$，则 $\sigma^k(r) = k + r$；若 $k + r > n$，则 $\sigma^k(r) = n - k + r$。根据匿名性，对每个 $k \in \{1, \ldots, n\}$, $a \in \varphi(\mathbf{u}_{\sigma^k})$。由于根据假设， φ满足客观期望一致性，结合 $m = n$ 应用引理 1，可得，

$$a \in \varphi\left(\frac{1}{n}\sum_{k=1}^{n} \mathbf{u}_{\sigma^k}\right).$$

令 $\mathbf{v}$ 为由 $\mathbf{v} = \sum_{k=1}^{n} \mathbf{u}_{\sigma^k}$定义的效用剖面。容易观察到，对于每个备选方案 $s \in S$ 和每个个体 $i \in N$,

$$v_i(s) = \sum_{j \in N} u_j(s).$$

因此 ，对于任意 $i, j \in N$ 和任意 $s \in S$，我们有 $v_i(s) = v_j(s)$。由一致同意可知 t

$$\varphi\left(\frac{1}{n}\sum_{k=1}^{n} \mathbf{u}_{\sigma^k}\right) \subseteq B(\mathbf{u}),$$

从而得到 $a \in B(\mathbf{u})$。

步骤 2. 若 φ 满足非空性、匿名性、一致同意、连续性和客观期望一致性，则有 $B(\boldsymbol{u}) \subseteq \varphi(\boldsymbol{u})$。

令 $a \in B(\mathbf{u})$ 和 $(\mathbf{w}^n)_{n \in \mathbb{N}^*}$ 为一列效用剖面，对每个 $i \in N$ 和$n \in \mathbb{N}^*$，分别定义为 $w_i(a) = u_i(a) + \frac{1}{n}$；对任意 $s \neq a$，定义为 $w_i(s) = u_i(s)$。显然，对于任意$n \in \mathbb{N}^*$， $B(\mathbf{w}^n) = \{a\}$。非空性公理和步骤 1 意味着，对于每个 $n \in \mathbb{N}^*$，$\varphi(\mathbf{w}^n) = \{a\}$。由于当 n 趋于无穷时，$(\mathbf{w}^n)_{n \in \mathbb{N}^*}$ 收敛于 $\mathbf{u}$，连续性公理意味着 $a \in \varphi(\mathbf{u})$，如所需。

■

备注 1。 值得注意的是，命题 1 可由以下观察推导得出：

a) 根据证明的第一步，任何满足匿名性、一致同意和客观期望一致性的解，都是边沁解的子解。b) 满足主观期望一致性的解也满足客观期望一致性。c) 然而，边沁解不满足主观期望一致性。

为支持最后一点，我们可以引用一个反例，它与用于确立命题1时所用的反例类似。

3.3 罗尔斯解的公理化刻画

本节中，我们使用非空性、一致同意、匿名性、连续性公理，以及名为最小一致性的一致性公理变体，对罗尔斯解进行公理化刻画。据此，我们找出一条能够区分罗尔斯解与边沁解的公理。

为此，我们提出了一种一致性公理的替代表述，它更适用于完全不确定的情形。在这种情况下，我们无法获知某个效用剖面（$\mathbf{u}$ 或$\mathbf{v}$）出现的概率，因此无法使用期望效用。文献中一种自然且广泛使用的方法是采取悲观方法，假定主体的效用会出现最坏的可能结果；具体而言，对于任意参与者$i \in N$和任意备选方案 $s \in S$，最坏情形剖面以$\mathbf{u} \wedge \mathbf{v}$表示，其定义如下：

$$[\mathbf{u} \wedge \mathbf{v}]_i(s) = u_i(s) \wedge v_i(s)$$

以下针对不确定环境中集体决策的准则指出，只要存在一个备选方案在两个效用剖面中都被社会选中，那么该备选方案也应在最坏情形剖面$\mathbf{u} \wedge \mathbf{v}$中被社会选中。该原则可形式化表述如下：

最小一致性：对于所有$\mathbf{u}, \mathbf{v} \in \mathrm{U}$，若 $\varphi(\mathbf{u}) \cap \varphi(\mathbf{v}) \neq \emptyset$，则：

$$\varphi(\mathbf{u}) \cap \varphi(\mathbf{v}) \subseteq \varphi(\mathbf{u} \wedge \mathbf{v})$$

定理 2. 一个解概念满足非空性、匿名性、一致同意、连续性和最小一致性，当且仅当它是罗尔斯解。

证明. 我们只证明这些公理的充分性。证明分两步进行。

步骤 1. 如果 φ 满足非空性、匿名性、一致同意和最小一致性，那么我们有 $\varphi(\boldsymbol{u}) \subseteq R(\boldsymbol{u})$.

设 a属于 $\varphi(\mathbf{u})$，并设 $\sigma \in \mathfrak{S}(N)$ 是定义在 N上的循环置换，其定义与定理 1中步骤 1 的定义相同。根据 匿名性，对每个 $k \in \{1, \ldots, n\}$， $a \in \varphi(\mathbf{u}_{\sigma^k})$。由于假设 φ 满足 最小一致性，对 n 作简单归纳即可推出，

$$a \in \varphi\left(\bigwedge_{k=1}^{n} \mathbf{u}_{\sigma^k}\right).$$

设 $\mathbf{v}$ 为由 $\mathbf{v} = \bigwedge_{k=1}^{n} \mathbf{u}_{\sigma^k}$定义的效用剖面。容易看出，对每个备选方案$s \in S$ 和每个个体 $i \in N$，我们有

$$v_i(s) = \bigwedge_{j \in N} u_j(s).$$

因此，对于任意 i， $j \in N$ 以及任意 $s \in S$，我们有 $v_i(s) = v_j(s)$。然后，根据一致同意，

$$\varphi\left(\bigwedge_{k=1}^{n} \mathbf{u}_{\sigma^k}\right) \subseteq R(\mathbf{u}),$$

于是我们得到 $a \in R(\mathbf{u})$。

步骤 2. 如果 φ 满足非空性、匿名性、一致同意、最小一致性和连续性，那么我们有 $R(\boldsymbol{u}) \subseteq \varphi(\boldsymbol{u})$.

设 $a \in R(\mathbf{u})$ 和 $(\mathbf{w}^n)_{n\in\mathbb{N}^*}$ 是一个效用剖面序列，对每个 $i \in N$ 和$n \in \mathbb{N}^*$，由 $w_i(a) = u_i(a) + \frac{1}{n}$ 定义；对每个 $s \neq a$，由 $w_i(s) = u_i(s)$定义。显然，对任意$n \in \mathbb{N}^*$，$R(\mathbf{w}^n) = \{a\}$。非空性公理和步骤 1意味着，对每个 $n \in \mathbb{N}^*$，$\varphi(\mathbf{w}^n) =\{a\}$。根据连续性公理，由于当 n 趋于无穷时，$(\mathbf{w}^n)_{n\in\mathbb{N}^*}$ 收敛于$\mathbf{u}$，我们得出 $a \in \varphi(\mathbf{u})$，如所求。■

定理 1 和定理 2 的证明采用了相当相似的论证和步骤。很容易看出，这两个证明都可以采用弱化版的 一致同意 公理，而不影响其有效性。该弱化版本如下：如果所有个体共享同一个效用函数，以 $\mathbf{u} = (u_i)_{i\in N}$表示，即 $u_i = u_j$对所有 $i, j \in N$那么$\varphi(\mathbf{u}) \subseteq M(\mathbf{u})$。这一替代公理类似于 Young(1974) 为 Borda选择对应提出的忠实性概念，显然可由 一致同意 公理推出。

3.4 公理的独立性

本节将证明定理1 和定理2 中所涉及公理的独立性（形式证明留给读者）。因此，我们需要定义一些额外的解概念。首先回顾，一致解 M 对每个$\mathbf{u} \in \mathcal{U}$ 的定义如下：

$$M(\mathbf{u}) = \{s^* \in S : \forall i \in N, \forall s \in S, u_i(s^*) \geq u_i(s)\}.$$

令 $i \in N$ 为任意一个个体，并称其为独裁者。若对每个$\mathbf{u} \in \mathsf{U}$，$dict^i : \mathsf{U} \rightrightarrows S$ 是相对于 i 的独裁解，则称之为

$$dict^i(\mathbf{u}) = \operatorname*{argmax}_{s\in S} u_i(s).$$

若对每个$\mathbf{u} \in \mathsf{U}$，$AB : \mathsf{U} \rightrightarrows S$ 是反边沁解，则称之为

$$AB(\mathbf{u}) = B(-\mathbf{u}).$$

此外，我们定义 $R : \mathsf{U} \rightrightarrows S$ 为 $\prec$-次罗尔斯解，其中 $\prec$ 是S的一个线性序。用$\min_{\prec} X$ 表示 $X \subseteq S$ 关于 $\prec$的最小元素。R 对每个$u \in \mathsf{U}$定义为 $R(\mathbf{u}) = \min_{\prec} R(\mathbf{u})$。

此外，我们定义 $\widetilde{B}$ 为 $\widetilde{\prec}$-$subBentham$ 解，其中 $\prec$ 是一个线性序，定义在 S上。B 对每个 $u \in \mathcal{U}$ 定义为 $B(\mathbf{u}) = \min_{\prec} B(\mathbf{u})$。下表展示了定理 1 和定理 2中公理的独立性。事实上，对于每种公理化，我们总能找到另一种不同时满足全部四条公理的解概念。

	NE	A	U	C	O	EC	MC	
M	no	yes	yes	yes	yes	yes		
$dict^i$	yes	no	yes	yes	yes	yes		
AB	yes	yes	no	yes	yes	no		
$\widetilde{R}$	yes	yes	yes	no	no	yes		
$\widetilde{B}$	yes	yes	yes	no	yes	no		
R	yes	yes	yes	yes	no	yes		
B	yes	yes	yes	yes	yes	no		

4 结论

本文介绍了广为人知的一致性性质的三种变体。运用第一种变体，即 主观一致性，我们得到了一个不可能性结果；运用第二种变体，即 客观一致性，我们建立了经典功利主义的公理化；运用最后一种变体，即 最小一致性，我们刻画了平等主义原则。研究还表明，尽管这两种解在哲学上相互冲突，其公理化却密切相关；也就是说，以下四种性质： 非空性、一致同意、匿名性和连续性 为这两种解概念所共有，而且其公理基础的证明遵循相当相似的推理。因此，两种刻画的分歧仅在于各自概念所对应的恰当一致性变体的选择。一方面，功利主义要求满足 客观一致性 性质，因此，只要个体遵循冯·诺伊曼—摩根斯坦期望效用模型，就有可能与 Harsanyi(1955年) 的奠基性研究建立联系。另一方面，平等主义涉及 最小一致性 公理，因此，只要个体遵循最大最小准则，便可以解释选择函数应如何汇总基数偏好。

致谢

本研究由法国国家研究署（ANR）通过“ANR-22-CE26-0019-01” Citizens项目资助。谨此衷心感谢 MODMAD 提供的资金支持。作者还要感谢 Marco Li Calzi、Eric Rémila 和 Vassili Vergopoulos 提供的宝贵意见。

参考文献

Jeremy Bentham 的《道德与立法原理导论》（1789），由 j.*H Burns* 和 *HLA Hart* 编，伦敦，1789。

Charles Blackorby 和 David Donaldson。无领域限制下比率尺度与平移尺度的完全人际可比性：可容许的社会评价函数。《国际经济评论》，第249–268页，1982年。Walter Bossert 和 Kohei Kamaga。混合功利主义—最大最小社会福利排序的公理化。《经济理论》，69(2)：451–473，2020年。Florian Brandl 和 Felix Brandt。通过一致性证明最优博弈的合理性。《理论经济学，14(4)：1185–1201，2019年。Claude d’Aspremont 和 Louis Gevers。公平与集体选择的信息基础。《经济研究评论》，44(2)：199–209，1977年。Claude d’Aspremont 和 Louis Gevers。社会福利泛函与人际可比性。《社会选择与福利手册》，1：459–541，2002年。

Robert Deschamps 和 Louis Gevers。词典序极小化与功利主义规则：联合刻画。《经济理论杂志》，17(2)：143–163，1978年。Amrita Dhillon 和 Jean-François Mertens。相对功利主义。《计量经济学》，67(3)：471–498，1999年。Claude d’Aspremont。社会福利排序的公理。《社会目标与社会组织：纪念 *Elisha Pazner* 文集》，第19–76页，1985年。Claude d’Aspremont 等。福利主义与人际比较。《经济研究》，18(1)：3–17，1994年。Thibault Gajdos 和 Feriel Kandil。无知观察者。《社会选择与福利》，31(2)：193–232，2008年。Louis Gevers。论人际可比性与社会福利排序。《计量经济学：计量经济学会杂志》，第75–89页，1979年。Peter J Hammond。公平、阿罗条件与罗尔斯差别原则。《计量经济学：计量经济学会杂志》，第793–804页，1976年。Peter J Hammond。大经济体中的直接个体激励相容性。《经济研究评论》，46(2)：263–282，1979年。Peter J. Hammond。《赫萨尼功利主义定理：一个更简明的证明及其若干伦理意涵》，第305–319页。Springer Berlin Heidelberg，柏林，海德堡，1992年。John C Harsanyi。基数福利、个人主义伦理学与效用的人际比较。《政治经济学杂志》，63(4)：309–321，1955年。Aanund Hylland 和 Richard Zeckhauser。分别聚合信念与价值的贝叶斯群体决策不可能性。《计量经济学：计量经济学会杂志》，第1321–1336页，1979年。Kohei Kamaga。功利主义与平等主义何时对评价达成一致？一种交集方法。《数学社会科学》，94：41–48，2018年。Michael Mandler。通过线性规划实现赫萨尼功利主义。《经济学快报》，88(1)：85–90，2005年。Eric Maskin。功利主义定理。《经济研究评论》，45(1)：93–96，1978年。Philippe Mongin。赫萨尼聚合定理：多剖面版本与未决问题。《社会选择与福利》，11(4)：331–354，1994年。Philippe Mongin。一致的贝叶斯聚合。《经济理论杂志》，66(2)：313–351，1995年。Philippe Mongin 和 Marcus Pivato。罗尔斯差别原则与最大最小分配规则：一项新分析。《经济理论》，71(4)：1499–1525，2021年。Juan D Moreno-Ternero 和 John E Roemer。无知之幕违背优先性。《经济学与哲学》，24(2)：233–257，2008年。Roger B Myerson。社会选择问题中的功利主义、平等主义与时机效应。《计量经济学：计量经济学会杂志》，第883–897页，1981年。

Kui Ou-Yang. 广义罗尔斯主义。《社会选择与福利》，50(2):265–279，2018年。约翰·罗尔斯。《正义论》。剑桥（马萨诸塞州），1971年。John E Roemer。分配正义理论。哈佛大学出版社，1996年。John E Roemer。机会平等。1998年。Uzi Segal。让我们一致认为，所有独裁制度都同样糟糕。《政治经济学杂志》，108 (3):569–589，2000年。阿马蒂亚·森。功利主义与福利主义。《哲学杂志》，76(9):463–489，1979年。Joaquim Silvestre。准线性或 Gorman 型偏好下的功利主义与平等。《社会选择与福利》，42:781–791，2014年。Steven Strasnick。社会选择与罗尔斯差别原则的推导。《哲学杂志》，73(4):85–99，1976年。约翰·冯·诺伊曼和奥斯卡·摩根斯特恩。《博弈论与经济行为》，普林斯顿大学出版社，普林斯顿，1944年。编程辅助工具。John A Weymark。对哈萨尼—森之争中功利主义问题的重新考察。不列颠哥伦比亚大学，1990年。John A Weymark。测量理论与功利主义基础。《社会选择与福利》，25:527–555，2005年。H Peyton Young。博尔达规则的公理化。《经济理论杂志》，9(1):43–52，1974年。