

利他主义、社会偏好与经济分析

文/何国卿 龙登高 刘齐平

“自利理性人”（self-interested rational man）假设一般被认为是以新古典经济学为核心的现代经济学理论的基石，包括了经济人的自利性假设以及理性选择假设。尽管该假设并不现实，但它被认为在私有产权和市场经济制度范畴中仍然是一种合理的科学抽象，而与之相反的研究假设就受到了正统范式的排斥。

事实上，从古典经济学创立开始，利他主义经济学思想就蕴含在诸如斯密、帕累托等重要经济学家的著作中，但由于他们的理论分析未能数学化，其原本丰富的人性和社会性思想就逐渐式微了。即便如此，对主流经济学人性假设的批判从未停止过。赫伯特·西蒙等人的有限理性理论改写了经济学理性人的内涵，而利他主义和社会偏好正是具有演化适应性的有限理性之一种。

利他主义及其对经济学理论的意义

在本文中，利他主义动机指的是经济社会行动者在决策时不仅考虑自身当前的物质利益，也同时考虑他人、社会集体或社会价值观的物质或精神利益。利己和利他两种利益动机会同时进入行动者的效用函数，并具有不同的权重系数。我们只将自身物质利益界定为利己主义动机，而产生于他人和社会价值观的自身精神利益（如荣誉感）甚或具有间接的长远的物质利益（如沽名钓誉）则属于利他主义动机的范畴。之所以这样规定，是因为我们希望研究各种各样的亲社会行为，而不是将那些“不纯粹”的亲社会行为（如互利互惠）跟明显与之不同的纯粹自利行为混为一谈。

尽管利他主义假设长期未被主流经济学接受，但近年来经济学理论和方法在多个方向上的新进展使其重新回到了经济学研究的前沿领域。

首先，利他主义被认为属于社会、政治或宗教范畴，不在经济学的研究范围之内。然而，在20世纪后期，社会科学中出现了一股将理性选择方法向非市场、非经济研究领域扩展的“经济学帝国主义”浪潮，暴露了经济学基本假设的局限性。加里·贝克尔因此承认，无论是自利还是非自利的动机，都只不过是偏好的不同表现形式，都可以纳入理性选择模型进行分析。

其次，从本体论上来说，利己主义被认为是资本主义市场经济主体的基本特征，符合经济学所研究的主要事实。然而，实验经济学的最后通牒博弈、信任博弈、礼物交换博弈、公共品博弈等实验研究有力地反驳了自利理性人假设，使得构建新的经济学理论去解释这些违背传统假设的公平、信任与合作动机变得重要起来。

再次，从方法论上来说，利己主义假设被认为与个体主义方法论是完全吻合的，而利他主义假设则并没有切断个体与群体之间千丝万缕的联系，带上了“整体论”的不良印记。但经济学方法论研究的进展已经突破了方法论个体主义与整体主义的二元对立，引入了在个体中蕴含整体的涌现属性、但整体仍然通过个体动力学来解释的系统论方法。

最后，从规范经济学的角度来说，利他主义曾认为是扭曲理性人的最优决策的，因此是不值得提倡的。与新古典经济学预设的、在完备的产权和法治制度下相互脱嵌的孤立个体不同，现实世界中的经济活动往往是在一定的社会关系的辅助下完成的。利己主义假设不仅不符合现实，而且在有利益冲突的情况下，也并不总是最优的，囚徒困境博弈、关系型合约理论和社会资本理论就很好地说明了这一点。人性在历史上不断地经受严酷的演化选择至今，而利他主义没有消亡，反而在多种经济活动中起着不可替代的作用，这也说明了利他主义充分适应了人类的经济社会环境。

利他主义假设对新古典理性人假设的拓展是为了增强经济学理论在运用于社会、政治、宗教等领域时的经验适当性。为此，既要吸收借鉴前人的思想和成果，也要努力将之数学化，以构建一个可用的社会经济人决策模型。下文将借鉴斯密、帕累托的利他主义经济学思想和科斯的指令经济思想，建立一个包含慈善、互惠、公平、身份认同、社会规范等社会偏好现象的适应性偏好效用模型。

利他主义、社会偏好与适应性效用函数

（一）利他主义与公平偏好

亚当·斯密在《道德情操论》中将利他主义看作是

人类的本性之一，提出了同情心理论和“公正的旁观者”理论。“公正的旁观者”近似于弗洛伊德心理学中的“超我”；利他主义体现了行动者的“自我”，这种“自我”属性可以是偏向自利的，也可以是偏向利他的，还能以对亲朋利他、对外人利己的“偏私”的形式将利己和利他这两种相反的行为倾向集于一身；最后，行动者的原始需求和欲望则可归于“本我”的范畴。行动者最终的决策将是利己主义的本我、利他主义的自我和公正旁观的超我的偏好叠加的结果。

假设当前所考虑的个体行动者*i*对另一个行动者*j*具有*f_{ij}*的“同情度”或“友好度”（-1 ≤ *f_{ij}* ≤ 1），该系数表示*j*相对于*i*自身来说所具有的效用权重：例如，*f_{ij}*=1表示“爱人如己”，*f_{ij}*=0表示“漠不关心”，*f_{ij}*<0则表示“视同寇仇”。为了让社会系统的网络结构能够演化，我们考虑的是能够按照互动博弈的结果进行调整的互惠利他主义（reciprocal altruism）。我们需要一个能满足两个特点的效用模型：（1）给定各效用项的预期效用，友好度向量的各分量可以调整，以使行动者的预期效用最大化；（2）友好度是行动者对其他人进行资源配置的依据，服从某种不受*f_{ij}*正负号影响的约束条件。以下模型符合这些条件，其中友好度与相应用项的预期效用成正比（证明从略）：

$$U_i = \pi_i + \sum_{j \neq i} f_{ij} \cdot [\pi_j + \tilde{f}_{ji} \cdot \pi_i] + r_i^s \cdot \sum_j f_{ij}^s \cdot \pi_j \quad \text{并满足 } (r_i^s)^2 + \sum_{j \neq i} f_{ij}^2 = 1 \quad (1)$$

其中，上式的第一项是行动者自己的自利效用；第二项是他人自利和利他的总效用；最后一项是社会评价项，也就是说，以社会的通常眼光来看，一次收益分配结果（不考虑利他效用）所产生的总效用为*f_{ij}^s* · *π_j*的和。这一效用的权重系数是*r_i^s*，表示行动者*i*对社会评价模式*s*的认同度（可以有許多社会评价模式共存，上式只列出了一个）。社会评价模式的偏好系数*f_{ij}^s*不能按行动者的意愿随时调节，但行动者对该社会评价模式的认同度*r_i^s*却是行动者自身可以自由调节的，行动者就像衡量一件商品那样衡量某种社会价值观的优劣取舍，由此实现个体价值观的更新和演化。

（二）社会身份认同与内在规则

维弗雷多·帕累托（1848-1923）对新古典经济学领域以外的社会科学的贡献很少为经济学界所知。他从诸种社会偏好的一团乱麻中辨认出了：（1）身份认同，包括国家、民族、宗教等对大规模社会行为具有重要意义的社会对象；（2）社会规范，即一些康德式的道德律令，往往依附于某种身份认同。我们在效用函数最后加上一项“身份认同”项（可以有多个）：

$$U_i = \pi_i + \sum_{j \neq i} f_{ij} \cdot [\pi_j + \tilde{f}_{ji} \cdot \pi_i] + r_i^s \cdot \sum_j f_{ij}^s \cdot \pi_j + r_i^t \cdot \tilde{U}_i \quad (2)$$

其中，*r_i^t*仍然是行动者对社会身份*t*的认同度，而 \tilde{U}_i 是社会身份*t*的虚拟“效用”。虽然社会身份并非是一个可以感知和行动的个体，但我们仍然假想其为一个从支持者中获得“效用”或社会价值。此外，社会身份还能产生额外的社会资本，以及依附于身份认同的社会规范体系，即宗教。对这种社会规范的违背将引起对身份认同的损害，进而破坏违规者在身份群体内的社会资本，这种巨大的潜在损失有可能是宗教道德规范约束力强大的原因之一。社会规范的虚拟效用函数如下：

$$\tilde{U}_t = \sum_n r_n^t \cdot U_n + \sum_j r_j^t \cdot r_j^t \cdot \pi_j + \sum_m r_m^t \cdot \chi_m^t \quad (3)$$

其中，我们假设身份认同群体中的一个人*n*当期的总效用*U_n*乘以他对目标对象*t*的认同度*r_n^t*，就是他对目标对象*t*所具有的价值，加总所有的*n*之后就是目标对象*t*的社会价值。*r_i^t* · *r_j^t*是圈内好友*j*在圈中的价值，因而*r_i^t* · *r_j^t*就是*j*对*i*的价值。*r_m^t*为社会对第*m*个规则的平均认同度，*χ_m^t*是一个布尔数，表示是否遵守了规则*m*，遵守规则记1，违背规则记0。在宗教团体内部，社会规范的认同度是变化缓慢的集体变量而非可调节的个体变量，所以群体压力和约束力更大，个体的自由选择空间更小。

（三）指令

罗纳德·科斯（1910-2013）主张企业的本质是以中心化的计划管理指令替代市场价格机制来配置资源。政治经济组织中的行动指令与上文中的康德式道德律令具有相似性，因而规则和指令都可以抽象为一种“布尔商品”。如果这种布尔商品是内在于行动者内心中的，那么它是规则；如果它是由一个行动者移交给另一个行动者的，那么它就是指令。事实上，违背这种社会规范导致的亵渎神圣，就好像拒绝老板的命令使得老板恼怒一样。对于指令的接受者来说，指令发出者的效用函数和（3）式一样，只是用指令*m*对老板的价值*v_m^b*来取代*r_m^t*。

进一步的研究方向

上文参照前人的思想建立了适应性社会偏好的效用模型。尽管它比传统的新古典决策模型复杂很多，难以用纯数学的方法得出具有概括性的定理证明，但演化经济学，尤其是计算机仿真技术的发展为复杂决策模型对经济系统行为影响的研究带来了曙光。构建包含了慈善、公平、互惠、身份认同、社会规范等社会偏好的利他主义效用模型还只是拓展经济学分析框架的第一步。在此基础上进一步寻求对声誉、信任、欺诈、社会网络、权力关系等更复杂的社会偏好现象的解释才是更重要的。^⑤

（何国卿系清华大学社会科学学院经济所博士生，龙登高系清华大学社会科学学院经济所教授，刘齐平系华中师范大学信息管理学院博士生；摘自《经济动态》2016年第7期）