

# 分配不平等与再分配的最优税率结构<sup>\*</sup>

赵志君

**内容提要:**在给定均值和基尼系数的一般性约束条件下,本文得到了定义于有限区间上的社会福利函数、最优收入分布和最优税率结构,找到了收入低于平均值的概率与帕累托指数和基尼系数之间的关系,给出了相应的政策建议。根据常相对风险厌恶和常绝对风险厌恶偏好对应的最优分布——帕累托分布和指数分布,对社会上出现的“被平均”现象给出了理论解释,弄清了“二八现象”发生的条件。

**关键词:**再分配 最优分布 最优税率 基尼系数 社会福利

## 一、引言

在计划经济时代,公平或分配的平均主义是主流社会的价值取向,收入分配的基尼系数维持在0.2左右,造成了“干多干少一个样、干与不干一个样”的平均主义大锅饭局面。改革开放以后,主流社会价值逐步转向为“效率优先兼顾公平”,“让一部分人先富起来,逐步实现共同富裕”一度成为鼓励百姓发家致富的口号,随之而来的是伴随着经济高速增长的收入差距逐步扩大。根据国家统计局公布的统计数据,我国的收入分配基尼系数从20世纪80年代的0.21左右上升到2009年的0.49,2010年后基尼系数出现了逐年下降趋势:2010年为0.481,2011年为0.477,2012年为0.474,2013年为0.473,2014年为0.469,2015年为0.462<sup>①</sup>。西南财经大学甘犁等(2015)在《中国家庭金融调查系列报告》中指出,2010年中国的收入分配、农村家庭收入分配、城市家庭收入分配的基尼系数分别高达0.61、0.6和0.56。谢宇等(2014)在《中国民生发展报告2014》指出,我国财产分配的基尼系数已经从1995的0.45上升到0.73。目前,世界单个经济体的收入分配基尼系数和财富分配基尼系数的均值分别是0.44和0.7<sup>②</sup>(世界银行网站和美国中央情报局网站)。可见,与世界其他国家相比,我国的分配不均等程度比较严重。

在经济高速增长时期,由于就业充分,劳动力工资逐年上升,虽然收入和财富的不平等程度不断扩大,整个社会仍处于一种帕累托改进状态,社会矛盾不容易凸显出来。但是,在当下经济增长的减速期、结构调整的阵痛期和前期政策的消化期,一些人不仅跟周围的人比收入差距拉大了,而且跟过去的自己比收入水平也下降了,社会矛盾就凸显出来了,不满情绪就会发泄出来,如不适当应对,可能影响社会稳定。因此,在未来一段时期,社会稳定在政府决策考量中应该占有更大的比重,调节收入分配结构,强化社会保障、化解社会矛盾应成为政府责无旁贷的任务,也是供给侧改革顺利推进的重要保障。

## 二、理论回顾

从理论上讲,国内外对社会财富分配现象的研究由来已久,早在1897年,帕累托(Pareto,1897)就发现了社会财富分布的“幂法则”,又称“帕累托法则”。“二八现象”(即财富水平最低的80%的人只占有20%的财富,财富水平最高的20%的人占有社会总财富的80%的现象)是该法则的具体表现。萨缪尔森(Samuelson,1965)对“幂法则”给予了高度评价,把它视为社会科学的不变规律,他认为“无论什么时间和地点,收入分配都保持不变,制度变迁和税收都不能改变社会科学的常量”。那么,是什么力量造成了财富分布服从帕累托法则?它与收入分配

<sup>\*</sup> 赵志君,中国社会科学院经济研究所,邮政编码:100836,电子邮箱:zhaozj@cass.org.cn。本文为国家自然科学基金项目“中国—中亚货币合作与人民币区域化和国际化研究”(71263050)和中国社会科学院创新工程项目“收入分配与经济增长和社会福利”的阶段性成果。作者感谢匿名审稿人提出的宝贵意见,但文责自负。

和人们之间的能力差别有何种关系?

长期以来,为弄清收入分配的规律,经济学家从市场因素、随机因素、遗产因素、才能因素、政府因素等多方面进行了不懈的探索。Cantelli(1921,1929)从动态随机过程出发证明,如果人的能力服从指数分布,而收入随着能力的变化按指数递增的话,则收入分配是帕累托分布。然而,能力是否服从指数分布是不可观察的,而且财富分布也不是由收入单一因素决定的。Frechet(1939)的研究表明,如果才能服从拉普拉斯分布,那么收入分布密度函数是单峰的,在中位数之前是递增的,在中位数之后是递减的,这时收入分配并不严格地服从帕累托分布。Roy(1950)提出了一种才能服从对数正态分布的机制,在财富增长和收入之间的关系中,如果不考虑劳动收入对财富分布的影响,假设财富增长率等于财富收益率,且收益率是独立同分布的,则随着时间的推移财富分布的极限是对数正态分布。由于对数正态分布在尾端近似等于帕累托分布,Roy的工作对帕累托法则给出了部分说明。Champernowne(1953)通过一个带有向下漂移、反映阻力的财富积累过程产生了帕累托分布。Wold & Whittle(1957)通过一个反映生死过程的财富积累方程产生了帕累托分布。

前述研究由于忽视市场机制在财富分配中的作用而受到主流经济学家的批评。根据经济学鼻祖亚当·斯密的观点,在“看不见的手”的作用下,社会财富分配的结果大体上同时实现了个体利益和社会利益最优化。他在《国富论》中明确指出:“各个人都不断地努力为他自己所能支配的资本找到最有利的用途。固然,他所考虑的不是社会的利益,而是他自身的利益,但他对自身利益的研究自然会或者毋宁说必然会引导他选定最有利于社会的用途。——由于他管理产业的方式目的在于使其生产物的价值能达到最大程度,他所盘算的也只是他自己的利益。在这场合,像在其他场合一样,他受着一只看不见的手的指导,去尽力达到一个并非他本意想要到达的目的。他追求自己的利益,往往使他能比在真正出于本意的情况下更有效地促进社会的利益。”(亚当·斯密,1776)。李卉、陈成明(2010)将“看不见的手”解读为“上帝”,认为“看不见的手”会对人们的利己行为进行限制,从而协调个人利益和公共利益之间的关系,达到“为大家”的结果。因此,是否能从市场机制出发获得财富分配的理论是有待进行深入理论探讨的问题。

Mincer(1958)、Becker & Tomes(1979)对收入

和财富分配研究中存在的忽视市场力量的现象也提出了批评。Mincer认为,用一个随机过程产生财富分配的机理缺乏财富分配的微观经济学基础。他写到:“从经济学家的观点看,随机过程模型与其他很多个收入分配模型一样,有一个令人不满的特征,它没有阐明分配过程的经济学,非经济因素固然对收入分配产生重要的影响,然而,除非否认理性最优化行为在经济活动中的合理性,否则个人收入分布就不能独立于经济活动,因而就不应该忽视个人选择在收入分配中的作用”。Becker & Tomes(1979)对从能力和运气出发研究财富分配、忽视财富的跨代转移的做法提出了批评:“不平等的跨代转移机制模型不包含父母对自己和孩子环境的最优化反应,极大地低估了资源禀赋的贡献,进而低估了家庭背景对不平等的影响”。

受Mincer(1958)、Becker & Tomes(1979)的观点的影响,20世纪90年代以来大量文献从异质经济人最优化行为出发研究财富积累动态过程的微观机理。Diaz et al(2003),Castaneda et al(2003)用Aiyagari(1994)的异质经济代理人和随机收益框架研究财富的稳定分布。Benhabib & Zhu(2008)在财富增长服从几何布朗运动的假设下,利用永葆青春的跨代生命周期模型证明财富服从双帕累托分布。Reed(2001)认为双帕累托分布比帕累托分布更符合实际。Blanchard(1985)用一个带有不变死亡率的永葆青春模型研究财富分配,可以看作Wold & Whittle模型的简化版。Carroll et al(2014)也利用具有微观基础的永葆青春模型研究了财富积累和分配。Piketty(2014)对如何缩小收入和财富差距提供了一些建设性的思路。

笔者认为,单从财富分配的微观基础出发研究分配也有明显的不足,特别是对像中国这样的转型国家来说更是如此。第一,现有的具有微观基础的经济模型,大多都采用了同质性偏好的假设,对异质偏好附加了很多限制条件。第二,新古典学派的经济模型所指的财富实际上是生产性物质资本而非金融财富,现实中的财富积累往往具有金融属性,金融财富的积累完全可能脱离生产领域而独立运行,积累的速度往往是物质资本积累速度的几倍或几十倍。例如,股票二级市场、期货、期权市场、债券交易市场、传销行为、汇率市场的交易行为基本是脱离实体经济而独立运行的,这些交易本身与经济增长没有直接的联系,其对财富分配的作用可能远远大于对经济增长的作用。很多富豪的财富积累是通过金

融市场的再分配过程实现的,金融市场的再分配造成的分配不公可能远远超过实体经济领域的分配不公,前述主流经济模型显然没有考虑到这一点。第三,由于主流经济学强调自由市场的作用,反对政府干预,所以财富分配的政治经济学很少在主流经济研究中得到体现。第四,一个显而易见的事实是,各国普遍采用的累进所得税政策肯定改变了收入和财富分配结构,财富分配依赖于财政政策,反过来财政政策也依赖于财富分配的不平等程度,但这方面的研究却非常匮乏(Benhbib & Bisin, 2016)。

从税收和社会的角度研究财富分配结构会遇到激烈的观念碰撞和价值判断的争议。一方面,税收和政府干预历来是市场派人士攻击的目标;另一方面,个体主义方法论在涉及群体行为和宏观问题时遇到了不可克服的困难。效用的可测性和人际间的效用可比性受到了新福利主义的质疑,福利的个人加总方法因此受到了挑战。阿罗不可能性定理(Arrow, 1951)认为,从个体理性出发得不到社会理性的结果,个体理性与集体理性的矛盾暴露了个体主义方法论的局限性。为了维护个体主义方法论的正统地位,主流经济学在处理宏观经济问题和社会福利问题时通常借助于一个虚拟的代表性微观个体代替宏观群体,把一个宏观问题和社会问题简化为微观问题,从而回避了个体行为的加总和阿罗不可能性定理的困境。近年来,一些对异质性问题的研究,包括对偏好异质性、收入异质性、约束异质性、信念异质性等问题的研究,也只是在微观个体进行分门别类后,对同类个体行为进行加总,而不同类别的个体行为如何加总,仍然是一个悬而未决的问题。

个体主义方法论也遭到了制度学派和演化经济学派的反对。霍布森(Hobson, 1929)认为:“一个有组织的单位或整体,并不能通过分析其各组成部分得到充分解释:它的整体性是一个全新的产品,尽管从某种意义上讲其属性是从其组成部分中衍生出来的,但却是不能从其组成部分中探知的”。霍布森进而指出:“制度可以像形容有机生物体那样进行描述,与生命体一样,制度也有自己的意志”(Dopfer, 2005)。布坎南(Buchanan, 1954)认为:“个人主义是建立在个体是唯一有价值和有目的的实体的哲学假设的基础上的。在那里,根本不存在社会和集体理性的问题,如此简单的社会价值尺度是不存在的。何况我们可以采用另一种有组织的哲学假设,在那里,集体是独立的实体并具有自己的价值排序。只

有根据这个价值排序来检验这个实体的理性或者非理性才是合理的。”

笔者认为,研究社会福利、构造社会福利目标函数时,霍布森和布坎南的观点应该重新受到重视。任何人(包括政府)都不可能掌握全社会上每一个人的偏好、收入和财富信息,对政府决策而言,一些样本特征信息,诸如平均值、最大值、最小值、方差、基尼系数等可能已经足够,由于政府目标的短期性和可变性,站在政府的角度研究社会福利和分配问题,静态和短期方法具有合理性。Mirrlees(1971)就是通过静态的同质性偏好模型获得了最优税率,其方法具有建设性,但他的收入分配和税收政策建立在能力差别的基础上,没有考虑分布的不确定性问题。赵志君(2011)在不考虑社会财富差别的具体原因的情况下,从财富分配的不确定性和政府的社会福利目标出发,把基尼系数视为社会风险的度量,讨论了最优分布结构和社会福利最大化问题,弄清了收入分配服从帕累托分布的条件,证明常相对风险厌恶偏好对应的最优分布恰好是帕累托分布。但是用帕累托分布作为现实收入分配的理论近似也存在明显的问题。一方面,由于帕累托分布的均值和基尼系数只有在帕累托指数大于1时才存在,帕累托分布的方差只有在帕累托指数大于2时才存在,这些条件显然限制了帕累托分布的应用。如果考虑到现实社会的任何群体都存在最低和最高收入者,则可以不必过分考虑帕累托指数的约束,最优分布在更一般的偏好下存在。根据赵志君(2011)的模型,如果效用函数是常替代弹性的(这是经济理论研究中最常用的假设),且收入分布定义在无限区间上,则当常替代弹性效用函数的参数 $\theta \leq 1$ 时,均值、方差和基尼系数都不存在,如果把收入分布考虑在有限区间上,则不论参数 $\theta$ 的取值区间范围是什么,均值、方差和基尼系数都可以计算。所以,这种扩充是有理论意义的。另一方面,把最优财富分配问题限制在有限区间上更符合实际情况。在特定时间点社会上一定有最高收入者,经过再分配后,最高收入者的收入也一定是有限的。对最低收入水平的限制,主要从同情心、生存的基本需要出发设计政策界限,但对最高收入的考虑,不是为了限制收入的无限提高,只是考虑客观现实,至于再分配后的最高收入则是通过社会福利目标函数的优化来决定的。本文的目的之一是在更一般的条件下获得再分配调节函数,赵志君(2011)没有得到一个令分配状态过渡到最优分配状态的再分配调节函数。

本文主要是从理论上探讨有限区间上的最优财富结构,提出调节收入差别、实现社会福利最大化的手段和途径,获得最优税率结构,对社会成员感到“被平均”的现象给出理论解释,弄清“二八现象”出现的条件。

### 三、偏好、基尼系数与最优财富分布

#### (一) 社会福利目标函数与约束条件

个体目标和集体目标是不同的,个体利益最优化目标不会因为政府和市场调控发生变化,但个体行为却受到政府行为和市场行为的影响。政府的目标是由社会整体利益所确定的,如果政府认为稳定压倒一切,则它要考虑缩小收入差距、降低社会风险,把收入差距控制在可接受的范围之内。采用的方法可能是发展产业、扩大就业、提高最低工资、扶贫,通过累进所得税进行税收调节。

作为由个体构成的组织,政府的偏好决定了政策取向,而政府的偏好又随着领导人的更替而发生变化。从这个角度讲,具有政策含义的社会福利函数是随着组织及其领导人的更替而不断发生变化的。如果政府代表最贫穷的社会阶层,对应的社会福利函数就是罗尔斯社会福利函数。如果政府代表最富有的社会阶层,社会福利函数就是精英者社会福利函数。如果政府注重社会和谐和社会稳定,就要在精英者和贫困阶层的利益之间寻找平衡点,社会福利函数就是罗尔斯社会福利函数和精英者社会福利函数的加权平均,或者是个体效用的加总和期望值。

本文定义的社会福利函数是在一个特定政府偏好下对社会福利的评价,这样的社会福利函数就有了政策含义和政策上的可操作性。这样的社会福利函数可能与个体对社会福利的评价不一致,当然,当两者发生同方向变化时,个体的评价也具有了政策含义。假设政府偏好函数  $U=U(y)$  是二阶连续可微的,即  $U=U(y)$  满足  $\frac{\partial U}{\partial y} > 0, \frac{\partial^2 U}{\partial y^2} \leq 0$ 。

当今社会一般都规定了最低生活保障水平。假设家庭初始收入(或财富)用随机变量  $X$  表示,其概率分布为  $F_0(x) = P(X < x)$ ,  $X$  的均值  $EX = \mu$ ,基尼系数  $G(X) = G_0$ 。  $F_0(x)$  在  $[x_m, x_M]$  上连续可微,且满足  $F_0(x_m) = 0, F_0(x_M) = 1$ ,  $x_m$  表示开始社会上最穷的人拥有的财富,  $x_M$  表示社会上最富的人拥有的财富。如果初始收入分配差距过大,需要通过收入再分配(税收或补贴)进行调整,调整后的最低收入水平为  $y_m$ ,最高收入水平为  $y_M$ ,基尼系数  $G(Y) = G$ 。调整后的收入  $Y$  与调整前的收入  $X$  之间的

关系为  $Y = f(X)$ 。

在社会成员总数一定的条件下,如果求得了社会平均福利水平,则社会成员总数与平均福利水平的乘积就是社会总福利,因此,可以用平均福利水平等价地表示社会的福利水平。为方便起见,假设总体足够大,经过再分配后收入分布近似地用连续函数  $F(y)$  表示,则社会福利可表示为:

$$SW = \int_{y_m}^{y_M} U(y) F'(y) dy \quad (1)$$

假如政府不愿意看到财富分配两极分化,试图通过某种再分配政策调整收入和财富分配,收入分配政策面临两个约束条件,第一个约束条件是社会财富的总水平在分配调节前后不发生改变,即:

$$\begin{aligned} EX = EY &= \int_{y_m}^{y_M} y F'(y) dy \\ &= \int_{x_m}^{x_M} x F_0'(x) dx = \mu \end{aligned} \quad (2)$$

其中,  $\mu$  是一个常数。

另一个约束条件是社会对收入和财富差别的容忍度。对收入差别的容忍度是指收入再分配受基尼系数大小的制约。基尼系数作为收入差距和社会风险的一种度量,既不能太大,也不能太小,太大会导致社会不稳定,太小会导致激励不足、影响效率。

连续分布的基尼系数可表示为洛伦兹曲线的积分。洛伦兹曲线的定义为:

$$L(F(y)) = \frac{\int_{y_m}^y t F'(t) dt}{\int_{y_m}^{y_M} t F'(t) dt} = \frac{1}{\mu} \int_{y_m}^y t F'(t) dt \quad (3)$$

做变换:  $F(t) = q, F(y) = p, t = F^{-1}(q), y = F^{-1}(p)$ , 则洛伦兹曲线可以写成:

$$L(p) = \frac{\int_0^p F^{-1}(q) dq}{\int_0^1 F^{-1}(q) dq} = \frac{1}{\mu} \int_0^p F^{-1}(q) dq \quad (4)$$

基尼系数可表示为:

$$\begin{aligned} G &= 1 - 2 \int_0^1 L(p) dp \\ &= 1 - \frac{2}{\mu} \int_0^1 \left( \int_0^p F^{-1}(q) dq \right) dp \end{aligned} \quad (5)$$

基尼系数给定条件下的约束条件还可以表示为:

$$\int_{y_m}^{y_M} y(1 - F(y)) F'(y) dy = \frac{1}{2} \mu(1 - G) \quad (6)$$

关于均值和基尼系数的约束条件还有如下关系:

$$\int_{y_m}^{y_M} (1 - F(y)) dy = \mu - y_m \quad (7)$$

$$\int_{y_m}^{y_M} (1 - F(y))^2 dy = \mu(1 - G) - y_m \quad (8)$$

约束条件(6)的含义是:对应给定均值 $\mu$ 和基尼系数 $G$ ,存在无穷多个分布于区间 $(y_m, y_M)$ 上的随机变量和无穷多条洛伦兹曲线。那么在这无穷多条洛伦兹曲线当中,哪一个是最优的分布呢?即哪一个是在约束条件(2)和(6)下使社会福利目标函数(1)最大化的分布呢?

## (二)最优分布与社会福利函数

$SW = \int_{y_m}^{y_M} U(y) F'(y) dy$  是一个在积分约束条件(2)和(6)下的关于分布函数 $F(y)$ 的积分泛函,其中的 $F(y)$ 是不确定的。这种概率不确定条件下的最优化结果是一种非线性数学期望(Peng, 1997; Chen & Epstein, 2002)。

命题1:在约束条件(2)和(6)下,使社会福利目标函数(1)最优的分布为:

$$F(y) = 1 - \frac{U'(y) - U'(y_M)}{U'(y_m) - U'(y_M)}$$

社会福利函数由最低收入、平均收入、最高收入和基尼系数共同决定:

$$SW = U(y_m) + U'(y_m)(\mu - y_m) - [U'(y_m) - U'(y_M)]\mu G$$

证明:首先根据目标函数(1)、约束条件(2)和(6),构造拉格朗日被积函数:

$$L = U(y) F'(y) + \lambda_1 [y F'(y)] + \lambda_2 [y(1 - F(y)) F'(y)] \quad (9)$$

一般而言,拉格朗日乘子是依赖于变量 $y$ 的,但在等周约束条件下,拉格朗日乘子是常数。蒋中一(1999)在《动态最优化方法》一书第170~171页中专门讨论了等周约束问题,证明在积分等于常数的约束条件下拉格朗日乘子必定是常数。据此,在解下面的变分问题时本文直接假设拉格朗日乘数 $\lambda_1$ 和 $\lambda_2$ 为常数。

对(9)式求导可得该最优化问题的一阶条件:

$$\frac{\partial L}{\partial F} = -\lambda_2 y F'(y)$$

$$\frac{\partial L}{\partial F'(y)} = U(y) + \lambda_1 y + \lambda_2 y(1 - F(y))$$

$$\frac{d}{dy} \left[ \frac{\partial L}{\partial F'(y)} \right] = U'(y) + \lambda_1 + \lambda_2(1 - F(y)) - \lambda_2 y F''(y)$$

以上三式代入欧拉公式 $\frac{d}{dy} \left[ \frac{\partial L}{\partial F'(y)} \right] - \frac{\partial L}{\partial F} = 0$ 得:

$$U'(y) + \lambda_1 + \lambda_2(1 - F(y)) = 0 \quad (10)$$

将边界条件 $F(y_m) = 0, F(y_M) = 1$ 代入(10)式得:

$$U'(y_m) + \lambda_1 + \lambda_2 = 0 \quad (11)$$

$$U'(y_M) + \lambda_1 = 0 \quad (12)$$

解得:

$$\lambda_1 = -U'(y_M) < 0 \quad (13)$$

$$\lambda_2 = -[U'(y_m) - U'(y_M)] < 0 \quad (14)$$

$$F(y) = 1 - \frac{U'(y) - U'(y_M)}{U'(y_m) - U'(y_M)} \quad (15)$$

显然, $F(y) = 1 - \frac{U'(y) - U'(y_M)}{U'(y_m) - U'(y_M)}$ 满足分布函数的性质。

由于 $\frac{\partial^2 L}{\partial F^2} = 0, \frac{\partial^2 L}{\partial F'^2} = 0, \frac{\partial^2 L}{\partial F \partial F'} = \frac{\partial^2 L}{\partial F' \partial F} = -\lambda_2 y, \frac{d}{dy} \left( \frac{\partial^2 L}{\partial F \partial F'} \right) = -\lambda_2 = U'(y_m) - U'(y_M) > 0$ ,所以,对于任意满足 $p(y_m) = 0$ 和 $p(y_M) = 0$ 的扰动曲线 $p(y)$ 和任意小的 $\epsilon$ ,定义 $F(y) = F^*(y) + \epsilon p(y)$ ,则:

$$\begin{aligned} \frac{D^2 SW}{d\epsilon^2} &= \int_{y_m}^{y_M} \left[ \left( \frac{\partial^2 L}{\partial F^2} - \frac{d}{dy} \frac{\partial^2 L}{\partial F \partial F'} \right) p^2(y) \right. \\ &\quad \left. + \frac{\partial^2 L}{\partial F'^2} p'^2(y) \right] dy \\ &= \int_{y_m}^{y_M} [\lambda_2 p^2(y)] dy < 0 \quad (16) \end{aligned}$$

该最优化问题的充分条件也成立,即 $F(y) = 1 - \frac{U'(y) - U'(y_M)}{U'(y_m) - U'(y_M)}$ 是使 $SW$ 最优的概率分布。有了最优收入分布函数,下面来求对应的社会福利表达式。

$$\begin{aligned} SW &= \int_{y_m}^{y_M} U(y) F'(y) dy \\ &= \int_{y_m}^{y_M} U(y) \left( 1 - \frac{U'(y) - U'(y_M)}{U'(y_m) - U'(y_M)} \right)' dy \\ &= U(y_m) + U'(y_m)(\mu - y_m) - [U'(y_m) - U'(y_M)]\mu G \quad (17) \end{aligned}$$

式(15)表明,最优收入分配依赖最低收入、最高收入和偏好的具体形式。式(17)的推导过程用到了(7)、(8)、(15)式。从式(17)可知,社会福利由最低收入、最高收入、平均收入和基尼系数四个变量确定,它是最低收入和平均收入的增函数,是最高收入和基尼系数的减函数。

## (三)最优税率与税收起征点

关于最优税率的定义,本文的考虑为:假设给定

初始分布  $F_0$ , 经过最优化再分配过程,  $F_0$  变成了  $F$ , 从  $F_0$  到  $F$  的转化是通过税收函数  $T(x)$  来实现的, 由于  $F$  是给定约束条件下的最优分布, 所以称税收函数  $T(x)$  为最优税收函数, 对应的边际税率  $\tau(x)$  为最优税率。

因为  $P(Y < y) = P[f(X) < y] = P(X < f^{-1}(y)) = F_0[f^{-1}(y)] = F_0(x)$ , 所以  $F(y) = F_0(x)$ ,  $y = F^{-1}(F_0(x))$ 。

由  $F(y) = 1 - \frac{U'(y) - U'(y_M)}{U'(y_m) - U'(y_M)} = F_0(x)$  可得:

$$U'(y) = (U'(y_M) + [U'(y_m) - U'(y_M)][1 - F_0(x)])$$

税收前后的收入关系:

$$y(x) = U'^{-1}(U'(y_M) + (U'(y_m) - U'(y_M))(1 - F_0(x)))$$

设边际税率为  $\tau(x)$ , 则:

$$y(x) = x - T(x) = x - \int_{x_m}^x \tau(x) dx, \int_{x_m}^x \tau(x) dx = x - y(x), \tau(x) = 1 - y'(x)。$$

命题 2: 通过最优税收函数  $T(x)$  将初始分布  $F_0(x)$

变换到分布  $F(y) = 1 - \frac{U'(y) - U'(y_M)}{U'(y_m) - U'(y_M)}$  后, 有:

$$(1) \text{ 税后收入 } y(x) = U'^{-1}(U'(y_M) + [U'(y_m) - U'(y_M)][1 - F_0(x)]);$$

$$(2) \text{ 边际税率 } \tau(x) = 1 - y'(x);$$

(3) 税收起征点满足方程

$$\frac{dU'^{-1}\{U'(y_M) + [U'(y_m) - U'(y_M)][1 - F_0(x)]\}}{dx} = 1。$$

设效用函数是最常见的常相对风险厌恶的:

$$U(y) = \frac{y^{1-\theta}}{1-\theta}, \theta > 1 \quad (18)$$

当  $y_M \rightarrow \infty$  时, 我们有:

$$F(y) = 1 - \frac{U'(y) - U'(y_M)}{U'(y_m) - U'(y_M)} = \frac{1 - (\frac{y_m}{y})^\theta}{1 - (\frac{y_m}{y_M})^\theta} \rightarrow 1 - (\frac{y_m}{y})^\theta \quad (19)$$

假设税前收入服从帕累托分布, 即  $F_0(x) = 1 - (\frac{x_m}{x})^\theta$ ,

税后收入服从帕累托分布, 即  $F(y) = 1 - (\frac{y_m}{y})^\theta$ , 税

前和税后基尼系数分别为  $G_0$  和  $G (G < G_0)$ , 则  $\theta > \theta_0$

(因为  $G = \frac{1}{2\theta - 1}$ ),  $y(x) = y_m x_m^{-\frac{\theta_0}{\theta}} x^{\frac{\theta_0}{\theta}}$ ,  $y'(x) = \frac{\theta_0}{\theta} y_m$

$$x_m^{-\frac{\theta_0}{\theta}} x^{\frac{\theta_0}{\theta} - 1}, \tau(x) = 1 - y'(x) = 1 - \frac{\theta_0}{\theta} y_m x_m^{-\frac{\theta_0}{\theta}} x^{\frac{\theta_0}{\theta} - 1}。$$

令  $\tau(x) = 0$ , 解得  $\bar{x} = (\frac{\theta_0}{\theta})^{\frac{\theta}{\theta - \theta_0}} y_m^{\frac{\theta}{\theta - \theta_0}} x_m^{-\frac{\theta_0}{\theta - \theta_0}}$  为税收起征点。当  $x > \bar{x}$  时, 税率为正; 当  $x < \bar{x}$  时, 个体收入达不到起征点, 实行负税率计划, 即对低收入进行补贴。又由于  $\tau'(x) = -y''(x) = \frac{\theta_0}{\theta} \frac{\theta - \theta_0}{\theta} y_m x_m^{-\frac{\theta_0}{\theta}} x^{\frac{\theta_0}{\theta} - 2} > 0$ , 所以最优边际税率是累进的。

#### 四、“被平均”和“二八现象”的理论依据

众所周知, 经济学家在理论和实证分析中用得最多的两种偏好是常相对风险厌恶和常绝对风险厌恶的效用函数。如果偏好是常相对风险厌恶的,  $U(y) = \frac{y^{1-\theta}}{1-\theta}$ , 则分布函数的极限是帕累托分布

$\lim_{y_M \rightarrow \infty} F(y) = \lim_{y_M \rightarrow \infty} (1 - \frac{U'(y) - U'(y_M)}{U'(y_m) - U'(y_M)}) = 1 - (\frac{y_m}{y})^\theta$ ; 如果偏好是常绝对风险厌恶的,  $U(y) = -e^{-ay}$ , 则分布函数的极限是指数分布  $\lim_{y_M \rightarrow 0} \lim_{y_M \rightarrow \infty} F(y) = \lim_{y_M \rightarrow 0} \lim_{y_M \rightarrow \infty} (1 - \frac{U'(y) - U'(y_M)}{U'(y_m) - U'(y_M)}) = 1 - e^{-ay}$ 。这两个结果与赵志君(2011)的结论是一致的。

由此可见, 帕累托分布和指数分布是两个最典型的分布函数。下面的命题 3 针对帕累托分布情况给出了收入低于平均值的概率和洛伦兹曲线的表达式, 从而给出了“被平均”现象的理论根据和“二八现象”发生的条件。

命题 3: 设  $Y$  是服从帕累托分布  $F(y) = 1 - (\frac{y_m}{y})^\theta$  的随机变量,  $p$  是收入少于  $y$  的人口占总人口的比例,  $L$  是收入少于  $y$  的人口的收入之和占社会总收入的比例, 则有:

$$(1) \text{ 均值 } EY = \frac{\theta y_m}{\theta - 1}$$

$$(2) \text{ 几何均值 } g(Y) = y_m e^{\frac{1}{\theta}}$$

$$(3) \text{ 调和均值 } H(Y) = y_m (1 + \frac{1}{\theta})$$

$$(4) \lim_{\theta \rightarrow \infty} P(Y < EY) = 1 - e^{-1}$$

$$(5) P(Y < g(Y)) = 1 - e^{-1}$$

$$(6) \lim_{\theta \rightarrow \infty} P(Y < H(Y)) = 1 - e^{-1}$$

$$(7) \text{ 洛伦兹曲线 } L = L(p) = 1 - (1 - p)^{\frac{\theta - 1}{\theta}}$$

命题 3 的证明:

$$(1) EY = \int_{y_m}^{\infty} y dF(y)$$

$$= \int_{y_m}^{\infty} y \theta y_m^\theta y^{-\theta - 1} dy = \frac{\theta y_m}{\theta - 1}$$

$$\begin{aligned}
 (2) \ln[g(Y)] &= E[\ln Y] = \int_{y_m}^{\infty} \ln y d\left(-\left(\frac{y_m}{y}\right)^{\theta}\right) \\
 &= \left[\ln y \left(-\left(\frac{y_m}{y}\right)^{\theta}\right)\right]_{y_m}^{\infty} + \int_{y_m}^{\infty} \left(\frac{y_m}{y}\right)^{\theta} \frac{1}{y} dy \\
 &= \ln y_m + \frac{1}{\theta}, g(Y) = e^{(\ln y_m + \frac{1}{\theta})} = y_m e^{\frac{1}{\theta}}
 \end{aligned}$$

$$\begin{aligned}
 (3) H(Y) &= 1/(E\left(\frac{1}{y}\right)) = 1/\int_{y_m}^{\infty} \frac{1}{y} (-y_m^{\theta}) dy^{-\theta} \\
 &= 1/\int_{y_m}^{\infty} \theta y_m^{\theta} y^{-\theta-2} dy = y_m \left(1 + \frac{1}{\theta}\right)
 \end{aligned}$$

$$\begin{aligned}
 (4) \lim_{\theta \rightarrow \infty} P(Y < EY) &= 1 - \left(\frac{y_m}{EY}\right)^{\theta} \\
 &= \lim_{\theta \rightarrow \infty} \left[1 - \left(1 - \frac{1}{\theta}\right)^{\theta}\right] = 1 - e^{-1}
 \end{aligned}$$

$$\begin{aligned}
 (5) P(Y < g(Y)) &= 1 - \left(\frac{y_m}{g(Y)}\right)^{\theta} = 1 - (e^{-\frac{1}{\theta}})^{\theta} \\
 &= 1 - e^{-1}
 \end{aligned}$$

$$\begin{aligned}
 (6) \lim_{\theta \rightarrow \infty} P(Y < H(Y)) &= \lim_{\theta \rightarrow \infty} \left(1 - \left(\frac{y_m}{H(Y)}\right)^{\theta}\right) \\
 &= \lim_{\theta \rightarrow \infty} \left[1 - \left(1 + \frac{1}{\theta}\right)^{-\theta}\right] \\
 &= 1 - e^{-1}
 \end{aligned}$$

$$(7) \text{假设 } F(y) = 1 - \left(\frac{y_m}{y}\right)^{\theta} = q, \text{ 则 } F^{-1}(q) = y = y_m$$

$(1-q)^{-\frac{1}{\theta}}$ , 代入(4)解得:

$$\begin{aligned}
 L(p) &= \frac{\int_0^p F^{-1}(q) dq}{\int_0^1 F^{-1}(q) dq} = \frac{1}{\mu} \int_0^p F^{-1}(q) dq \\
 &= \frac{1}{\theta y_m} \int_0^p y_m (1-q)^{-\frac{1}{\theta}} dq \\
 &= 1 - (1-p)^{\frac{\theta-1}{\theta}}
 \end{aligned}$$

命题3告诉我们,帕累托分布族中个体收入低于平均收入的概率的下界是0.63212。在一般情况下,  $P(Y < EY) > 1 - e^{-1} \approx 0.63212$ 。表1列出了对1-3之间的帕累托指数进行的模拟分析的部分结果,从中可以看出,帕累托指数为3时,基尼系数是0.2,对应的收入分配是相当平均的,我国只有在计划经济时期和改革开放初期才达到这样的平等程度,然而,即便在这样的情况下,低于平均收入的人数仍然占到70%。在当前基尼系数达到0.48或更高的情况下,可能有80%以上的人的收入达不到平均数,这也难怪大多数人感觉“被平均”了。从财富分配情况看,目前世界财富分配的基尼系数是0.8,这意味着在世界范围内,80%的低收入者仅占有16%的财富,比“二八现象”对应的分配不平等还严重。就单个经济体而言,各经济体的平均基尼系数是0.7,对应于80%的低收入者仅占24%的财富。

表1 低于平均水平的概率与“二八现象”发生的条件

情形	帕累托指数	基尼系数	收入低于均值的概率	90%低收入者收入占比	80%低收入者收入占比	70%低收入者收入占比
1	1	1	1	0	0	0
2	1.01	0.980392	0.990546	0.02254	0.015809	0.01185
3	1.02	0.961538	0.981875	0.044145	0.031065	0.023331
4	1.03	0.943396	0.973805	0.064866	0.045795	0.034459
5	1.04	0.925926	0.966238	0.084753	0.060024	0.045251
6	1.05	0.909091	0.959105	0.103849	0.073777	0.05572
7	1.06	0.892857	0.952355	0.122199	0.087074	0.065879
8	1.07	0.877193	0.945948	0.13984	0.099937	0.075743
9	1.08	0.862069	0.939849	0.156809	0.112385	0.085322
10	1.09	0.847458	0.934033	0.173142	0.124438	0.094629
11	1.1	0.833333	0.928473	0.188869	0.136112	0.103675
12	1.11	0.819672	0.923151	0.204022	0.147425	0.11247
13	1.12	0.806452	0.918048	0.218629	0.158391	0.121023
14	1.13	0.793651	0.913148	0.232717	0.169026	0.129346
15	1.14	0.78125	0.908438	0.24631	0.179343	0.137445
16	1.15	0.769231	0.903905	0.259432	0.189356	0.14533
17	1.16	0.757576	0.899537	0.272105	0.199078	0.153009
18	1.17	0.746269	0.895324	0.28435	0.208519	0.160489

续表 1

情形	帕累托指数	基尼系数	收入低于均值的概率	90%低收入者收入占比	80%低收入者收入占比	70%低收入者收入占比
19	1.18	0.735294	0.891257	0.296186	0.217693	0.167779
20	1.19	0.724638	0.887328	0.307633	0.226608	0.174884
21	1.2	0.714286	0.883529	0.318708	0.235276	0.181811
22	1.21	0.704225	0.879852	0.329427	0.243705	0.188567
23	1.22	0.694444	0.876292	0.339805	0.251906	0.195158
24	1.23	0.684932	0.872843	0.349858	0.259887	0.20159
25	1.24	0.675676	0.869497	0.3596	0.267655	0.207868
26	1.25	0.666667	0.866252	0.369043	0.27522	0.213997
27	1.26	0.657895	0.863101	0.3782	0.282589	0.219982
28	1.27	0.649351	0.860041	0.387083	0.289768	0.225829
29	1.28	0.641026	0.857067	0.395704	0.296765	0.231542
30	1.29	0.632911	0.854175	0.404072	0.303586	0.237125
31	1.3	0.625	0.851362	0.412198	0.310238	0.242582
32	1.31	0.617284	0.848624	0.420092	0.316726	0.247918
33	1.32	0.609756	0.845958	0.427763	0.323056	0.253136
34	1.33	0.60241	0.843361	0.435219	0.329234	0.25824
35	1.34	0.595238	0.84083	0.442469	0.335264	0.263235
36	1.35	0.588235	0.838363	0.449521	0.341152	0.268122
37	1.36	0.581395	0.835957	0.456382	0.346902	0.272905
38	1.37	0.574713	0.833609	0.463058	0.352519	0.277589
39	1.38	0.568182	0.831318	0.469558	0.358007	0.282174
40	1.39	0.561798	0.829081	0.475888	0.363371	0.286666
41	1.4	0.555556	0.826897	0.482053	0.368615	0.291066
42	1.41	0.549451	0.824763	0.488059	0.373742	0.295376
43	1.42	0.543478	0.822677	0.493913	0.378756	0.299601
44	1.43	0.537634	0.820638	0.499619	0.38366	0.303741
45	1.44	0.531915	0.818645	0.505183	0.388459	0.3078
46	1.45	0.526316	0.816695	0.51061	0.393154	0.31178
47	1.46	0.520833	0.814787	0.515904	0.39775	0.315683
48	1.47	0.515464	0.81292	0.52107	0.40225	0.319511
49	1.48	0.510204	0.811092	0.526112	0.406655	0.323266
50	1.49	0.505051	0.809303	0.531034	0.41097	0.326951
51	1.5	0.5	0.80755	0.535841	0.415196	0.330567
52	1.51	0.49505	0.805833	0.540536	0.419337	0.334116
53	1.52	0.490196	0.80415	0.545122	0.423395	0.3376
54	1.53	0.485437	0.802501	0.549604	0.427371	0.34102
55	1.54	0.480769	0.800885	0.553984	0.431269	0.344379
56	1.55	0.47619	0.7993	0.558266	0.435091	0.347677
57	1.56	0.471698	0.797745	0.562452	0.438839	0.350917
58	1.57	0.46729	0.79622	0.566546	0.442514	0.3541
59	1.58	0.462963	0.794724	0.570551	0.44612	0.357228



续表 1

情形	帕累托指数	基尼系数	收入低于均值的概率	90%低收入者收入占比	80%低收入者收入占比	70%低收入者收入占比
60	1.59	0.458716	0.793256	0.574469	0.449657	0.360301
61	1.6	0.454545	0.791815	0.578303	0.453127	0.363321
62	1.7	0.416667	0.77874	0.612532	0.484547	0.390888
63	1.8	0.384615	0.767688	0.640619	0.510957	0.414389
64	1.9	0.357143	0.758215	0.664018	0.533438	0.434646
65	2	0.333333	0.75	0.683772	0.552786	0.452277
66	2.1	0.3125	0.742804	0.700642	0.569599	0.467756
67	2.5	0.25	0.721145	0.748811	0.619269	0.514407
68	3	0.2	0.703704	0.784557	0.658005	0.55186

## 五、结语

本文的主要贡献在于:第一,在一般性假设下得到了有限区间上最优收入分配,拓展了赵志君(2011)的结果。本文证明,如果效用函数是常相对风险厌恶的,有限区间上最优财富分布是截尾帕累托分布,帕累托分布是其极限;如果效用函数是常绝对风险厌恶的,则最优分布是截尾指数分布,指数分布是其极限。截尾帕累托分布与帕累托分布相比存在明显的优点,它不需要像帕累托分布那样为保证方差和基尼系数的存在对帕累托指数的取值范围进行严格的限制,对截尾帕累托分布而言,帕累托指数只要大于0,均值、方差、基尼系数就存在。

第二,本文得到了使社会福利达到最大化的最优税率函数,对制定收入和财富分配政策具有参考价值。最优税率函数对政府制定遗产税、个人所得税的政策含义在于:如果社会初次分配造成的收入差距过大,则政府可以通过累进所得税和遗产税或其他政策缩小基尼系数。由于社会福利函数是由最低收入、平均收入、最高收入和基尼系数共同决定的,政府在制定政策时只需考虑偏好的具体形式调节一些样本总量指标。

第三,根据帕累托分布计算的收入低于均值的概率对社会上普遍存在的“被平均”现象给出的理论解释是本文一个意外收获,本文得到了帕累托分布族中收入低于均值的概率的下限。帕累托分布的洛伦兹曲线告诉我们“二八现象”近似反映了当今社会财富分配状况,但“二八现象”不是一种必然现象。

注:

①数据来源:国家统计局网站([http://www.stats.gov.cn/tjsj/sjjd/201603/t20160308\\_1328214.html](http://www.stats.gov.cn/tjsj/sjjd/201603/t20160308_1328214.html))。

②数据来源:世界银行网站(<http://data.worldbank.org/indicator>),美国中央情报局网站(<https://www.cia.gov/library/publications/the-world-factbook/rankorder/2172rank.html>)。

③蒋中一在其《动态最优化基础》(商务印书馆1999年版)一书中介绍了变分问题的四种约束条件,一般而言,拉格朗日乘子依赖于变量 $x$ ,但在等周约束条件下,拉格朗日乘子被证明是常数。

### 参考文献:

- 甘犁 尹志超 谭继军,2015:《2014 中国家庭金融调查系列报告》,西南财经大学出版社。
- 蒋中一,1999:《动态最优化基础》,商务印书馆。
- 李卉 陈成明,2010:《论“两只看不见的手”及其相互关系》,《郑州航空工业管理学院学报》第5期。
- 谢宇等,2014:《中国民生发展报告》,北京大学出版社。
- 亚当·斯密,1776:《国民财富的性质和原因的研究》(中译本,1972),商务印书馆。
- 赵志君,2011:《收入分配与社会福利函数》,《数量经济技术经济研究》第9期。
- Aiyagari, S. R. (1994), “Uninsured idiosyncratic risk and aggregate savings”, *Quarterly Journal of Economics* 109 (3):659—684.
- Arrow, K. J. (1951), *Social Choice and Individual Values*, Yale University Press.
- Becker, G. S. & N. Tomes (1979), “An equilibrium theory of the distribution of income and intergenerational mobility”, *Journal of Political Economy* 87(6):1153—1189.
- Benhabib, J. & A. Bisin (2016), “Skewed wealth distributions: Theory and empirics”, <https://www.researchgate.net/publication/301890892>.
- Benhabib, J. & S. Zhu (2008), “Age, luck and inheritance”, NBER Working Paper No. 14128.
- Blanchard, O. J. (1985), “Debt, deficits, and finite horizons”, *Journal of Political Economy* 93(2):223—247.
- Buchanan, J. M. (1954), “Social choice, democracy, and free markets”, *Journal of Political Economy* 62(2):114—123.

- Cantelli, F. P. (1921), “Sulla deduzione delle leggi di frequenza da considerazioni di probabilita”, *Metron* 1(3):83—91.
- Cantelli, F. P. (1929), “Sulla legge di distribuzione dei redditi”, *Giornale degli Economisti e Rivista di Statistica* 69 (11): 850—852.
- Carroll, C. D. et al(2014), “Wealth inequality and the marginal propensity to consume”, European Central Bank Working Paper No. 1655.
- Castaneda, A. et al(2003), “Accounting for the U. S. earnings and wealth inequality”, *Journal of Political Economy* 111(4):818—857.
- Champernowne, D. G. (1953), “A model of income distribution”, *Economic Journal* 63(250):318—351.
- Chen, Z. & L. Epstein(2002), “Ambiguity, risk and asset returns in continuous time”, *Econometrica* 70(4): 1403—1443.
- De Nardi, M. (2004), “Wealth inequality and intergenerational links”, *Review of Economic Studies* 71(3): 743—768.
- Diaz, A. J. et al (2003), “Precautionary savings and wealth distribution under habit formation preferences”, *Journal of Monetary Economics* 50(6):1257—1291.
- Dopfer, K. (2005), “Evolutionary economics: A theoretical framework”, in: K. Dopfer(ed), *The Evolutionary Foundations of Economics*, Cambridge University Press.
- Frechet, M. (1939), “Sur les formules de repartition des revenus”, *Revue del’Institut International de Statistique* 7(1):32—38.
- Hobson, J. A. (1929), *Wealth and Life: A Study in Values*, London: Macmillan.
- Levy, M. & S. Solomon(1996), “Power laws are logarithmic Boltzmann laws”, *International Journal of Modern Physics C* 7(1):65—72.
- Levy, M. (2005), “Market efficiency, the Pareto wealth distribution, and the Levy distribution of stock returns”, in: S. Durlauf & L. Blume (eds), *The Economy as an Evolving Complex*, Oxford University Press.
- Mirrlees, J. A. (1971), “An exploration in the theory of optimum income taxation”, *Review of Economic Studies* 38 (2):175—208.
- Pareto, V. (1897), *Cours d’Economie Politique, II*, F. Rouge, Lausanne.
- Peng, S. (1997), “BSDE and related g—expectation”, Pitman Research Notes in Mathematics Series 364:141—159.
- Piketty, T. (2014), *Capital in the Twenty—First Century*, Harvard University Press.
- Reed, W. J. (2001), “The Pareto, Zipf and other power laws”, *Economics Letters* 74(1):15—19.
- Roy, A. D. (1950), “The distribution of earnings and of individual output”, *Economic Journal* 60(239):489—505.
- Samuelson, P. A. (1965), “A fallacy in the interpretation of the Pareto’s law of alleged constancy of income distribution”, *Rivista Internazionale di Scienze Economiche e Commerciali* 12:246—250.
- Wold, H. O. A. & P. Whittle(1957), “A model explaining the Pareto distribution of wealth”, *Econometrica* 25(4): 591—595.

(责任编辑:陈建青)