基尼系数组群分解新方法研究:从城乡二亚组到多亚组 *

程永宏

内容提要:本文尝试在程永宏(2006)的基尼系数城乡分解方法基础上,论证一种适合多亚组的基尼系数组群分解新方法。主要内容是:第一,分析了现有分解方法不完善的原因;第二,推导出不含交叠项的多亚组情形下的分解方法,摆脱了交叠项的困扰;第三,提出一个组间不平等新指标并论证其理论依据;第四,明确解释了分解式和组间不平等新指标的社会福利含义。

关键词:基尼系数组群分解　组间不平等　社会福利损失

一、引　言

不平等[①]指标的组群分解是指:把一个收入总体(以下简称"总体")按某一属性划分为 m 个亚组,然后把总体的不平等表示成组内不平等和组间不平等两部分之和,其中组内不平等是各亚组不平等的加权和(Bourguignon,1979;Shorrocks,1980)。不平等指标组群分解的目的在于:度量组内和组间不平等对总体不平等的贡献,从而确认各种因素是如何影响总体不平等的(Shorrocks,1980;Fields,1979)。不平等指标组群分解的定义和目的,显然也适用于基尼系数。

自 20 世纪 60 年代以来,基尼系数组群分解方法持续受到国际学术界的关注。从 Soltow(1960)初步涉及这一问题,到 Das 和 Parikh(1982)对这类文献做出一个详细综述,其间 Bhattacharya 和 Mahalanobis(1967)做出重大发展。此后,其他作者分别基于不同的思路对 Bhattacharya 和 Mahalanobis 分解方法进行了局部改进或重新解释,或者将上述分解方法运用于实证研究,如 Silber(1989)、Sastry 和 Kelkar(1994)、Milanovic 和 Yizhaki(2002)、Lambert 和 Decoster(2005)等。

但是,基尼系数组群分解式的基本结构一直没有取得突破:[②]几乎所有分解式都是由组内不平等、组间不平等和交叠项三部分组成。其中交叠项的意义一直存在较大争议,如 Lambert 和 Decoster(2005)等;组间不平等定义也受到一些批评,如 Bourguignon(1979)、Dagum(1997)等;分解式的经济意义很难解释,其价值和说服力也因此受到怀疑。本文正是针对现有分解方法的缺陷,在程永宏(2006)的基础上进行拓展,建立一个适用于多亚组的分解方法。

二、基尼系数组群分解相关文献的回顾与评论

(一)关于符号的定义

本文涉及大量数学公式,为统一起见,首先对相关符号的含义做如下规定:设总体的个人收入随机变量为 I,相应的收入分布函数、基尼系数、总人口、总收入、平均收入分别为 $F(t)$、G、N、Y、

* 程永宏,中国人民大学公共管理学院,邮政编码:100872,电子信箱:chengyonghong@gmail.com。本文获得国家社科基金重大招标项目(05&ZD049)部分资助,以及中国人民大学科研基金(06XND001)部分资助。作者感谢联合国大学万广华教授的建议;同时也感谢匿名审稿人的宝贵意见和建议,但限于篇幅,部分建议没有采纳。

① 本文提到的"不平等",如果没有特别指明,均指相对不平等。

② Yitzhaki(1998)在 1998 年前后认为,基尼系数的组群分解问题还"处在发展的早期阶段"。

μ;第 i ($i=1,2,3,\cdots,m$) 亚组的个人收入随机变量为 I_i,相应的收入分布函数、基尼系数、总人口、总收入、平均收入、人口份额、收入份额分别为 $F_i(t)$、G_i、N_i、Y_i、μ_i、α_i、θ_i,则有:

$$N=\sum_{i=1}^{m}N_i,\ Y=\sum_{i=1}^{m}Y_i,\ \mu=Y/N,\ \mu_i=Y_i/N_i,\ \theta_i=Y_i/Y,\ \alpha_i=N_i/N \tag{1}$$

(二) 基尼系数组群分解相关文献回顾

基尼系数组群分解的相关文献基本上可以分为两类:第一类是直接提出、论证各种形式的分解方法;第二类是一般化地从理论上探讨不平等指标的可分解性问题。

1. 各种形式的基尼系数组群分解方法

Soltow(1960) 在分析教育、年龄等因素与美国个人收入差距的关系时,第一次未加证明地给出一个基尼系数组群分解公式:直接从总体基尼平均差中分离出组内不平等,即各亚组基尼平均差的加权平均,其中的权系数设定为亚组人口份额与亚组收入份额的乘积,其余部分作为组间不平等。

Bhattacharya 和 Mahalanobis(1967) 对 Soltow(1960) 的分解方法做出重大改进:他们首先定义了组间的基尼平均差,然后将总体的基尼平均差减去组间基尼平均差后的余项作为组内不平等,并把组内不平等又分解成两部分:第一部分是各亚组内部基尼平均差的加权平均,权系数是各亚组人口份额的平方;第二部分后来被称为"交叠项"(overlap term),它与亚组收入分布的交叠程度有关。

Rao(1969) 也给出另一种分解式,其组内不平等的权系数是人口份额,但仍然存在交叠项。Mangahas (1975) 未加证明地给出一个分解式:把总体基尼系数分解成组内不平等和组间不平等,其中组内不平等是各亚组基尼系数的加权平均,权系数是各亚组收入份额,且不含交叠项,这与本文即将论述的分解式完全相同,但二者共同点也仅限于此;在分解基础、组间不平等的定义及其经济意义、适用的数据形式等方面,二者都完全不同。为便于比较,这里先列出 Mangahas 分解式:

$$G=\sum_{i=1}^{m}\theta_i G_i+\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j(G_{ij}/\mu) \tag{2}$$

式中,G_{ij} 是 Mangahas 定义的第 i 组与第 j 组之间的"基尼差"(Gini difference),其表达式为:$G_{ij}=(f_i-f_j)'M(f_i-f_j)$。其中,$f_i$: $n\times1$ 维向量,n 为对各组人口按收入水平划分的阶层总数,f_i 中的第 k ($k=1,2,3,\cdots,n$) 个元素等于第 i 组中第 k 阶层的人口份额;$M=HP$: $n\times n$ 阶正定阵;P: $n\times n$ 阶对角矩阵,其主对角线上的第 k ($k=1,2,3,\cdots,n$) 个元素等于第 k 阶层的平均收入;H: $n\times n$ 阶三角矩阵,其主对角线上元素为 1,主对角线下方元素为 2,主对角线上方元素为 0。可见,Mangahas 以矩阵形式给出的组间不平等指标,结构比较复杂,经济意义不明确。分解式对原始数据要求也非常严格:总体划分成 m 个亚组后,还要对各亚组按相同收入标准划分成 n 个阶层。

Pyatt(1976) 利用博弈论方法,对基尼平均差做出新的解释,并给出一个新的分解公式,其特点是:组内不平等的权系数是各亚组收入份额与人口份额的乘积,交叠项被从组内不平等中分离出来。可以证明:Pyatt 与 Bhattacharya and Mahalanobis 的分解式本质上是相同的。Mookherjee 和 Shorrocks(1982) 以一种更简洁的方式重新表述了这一分解式 (式中 R 为交叠项):

$$G=\sum_{i=1}^{m}\frac{\mu_i}{\mu}\alpha_i^2G_i+\frac{1}{2N^2\mu}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j\,|\,\mu_i-\mu_j\,|+R \tag{3}$$

20 世纪 80 年代以来,对基尼系数组群分解问题的研究一直没有中断,但研究方向主要集中于对上述分解式中交叠项 R 的解释和争论,如 Silber (1989)、Lambert 和 Aronson (1993)、Lambert 和 Decoster(2005) 等;或是对分解式各项的权系数做出略有不同的定义,以及对计算过程进行改进,并将其应用于实证研究,如 Silber(1989)、Sastry 和 Kelkar(1994)、Yitzhaki (1994)、Dagum (1997)、Yao (1999)、Suoniemi (2000)、Milanovic 和 Yizhaki(2002) 等。在分解式的结构方面,一直没有摆脱交叠项的困扰。

2. 不平等指标可分解性的理论研究

Bourguignon(1979)更一般化地研究了不平等指标的可分解性理论。Bourguignon 首先给出了总合性不平等指标(aggregative inequality measures)的定义:给定某个不平等指标,如果总体的不平等可以表示成各亚组内部不平等和亚组总量指标(总人口、收入均值)的函数,则称之为总合性不平等指标。在此基础上,Bourguignon 给出加和可分解性不平等指标(additive decomposability measures)的定义:给定某个不平等指标,如果总体不平等可表示成组内不平等与组间不平等之和,其中,组内不平等是各亚组内部不平等加权平均,组间不平等是各亚组总人口和收入均值的函数,并且与总体不平等指标定义相同(这是隐含的),则称之为"加和可分解性"不平等指标。Cowell (1980)、Shorrocks (1980)与 Bourguignon 几乎同时且独立地给出了类似的加和可分解性条件,并对组内不平等的权系数进行了深入研究,结论也是:基尼系数不满足加和可分解性条件。

Shorrocks(1984)则意识到加和可分解性条件作为普遍的标准,似乎过分严格,因此建议采用一个更弱的条件作为可分解性的标准,以便扩大可分解不平等指标的范围。Zheng(2007)进一步讨论了如何放松加和可分解性条件,以拓宽可分解指标的集合。

(三)对现有分解方法和分解理论的评论

上述分解方法基本上都是从离散数据出发,利用矩阵来表示分解结果,这使得分解结果存在几个共同的问题:第一,要么是组间不平等缺乏明确的经济意义,如 Soltow(1960)、Mangahas (1975)、Rao(1969),要么是组内不平等缺乏明确的经济意义,如 Bhattacharya 和 Mahalanobis (1967)、Pyatt (1976);第二,上述分解式中组内不平等部分的权系数一般都是人口份额或人口份额与收入份额的乘积,这使得分解式的经济意义难以解释;第三,几乎所有的分解式都存在或隐含着经济意义不清晰的"交叠项";第四,所有分解方法都没有对分解式的社会福利含义做出解释。

上述分解理论中的加和可分解性条件,与不平等指标分解的定义和目的并不完全一致,特别是其中的组间不平等定义存在严重缺陷和逻辑困境,具体情况分析如下。

第一,如果仅仅要消除组内不平等,则没有任何理由阻止我们"赋予"亚组内每个人任意相等的收入,亚组均值在这个意义上没有任何特殊地位。亚组内部每个人都获得亚组收入均值,是消除组内不平等的充分条件但不是必要条件,因为虚构的亚组个人收入无需受实际总收入约束。这一假设背后的逻辑似乎是:消除组内不平等的同时,如果各亚组收入均值不变,则可以保持组间不平等不变。那么,这就等于说,组间不平等是亚组收入均值的函数,但这恰恰是需要证明的。

第二,加和可分解性条件中隐含着"组间不平等与总体不平等定义相同"的要求,但其必要性没有被证明。从不平等指标组群分解的定义和目的看,这是不必要的:只要能够把总体不平等表示成组内不平等和组间不平等之和,并且二者都满足不平等指标的基本性质,即可达到分解的目的;二者定义相同,对实现分解目的并不具有任何特定价值。实际上,Bhattacharya and Mahalanobis (1967)基尼系数分解式中的交叠项就不是按基尼系数来定义的;Fei et al (1978)对基尼系数按收入来源进行的分解中,各个来源的不平等指标也不是基尼系数。

第三,以亚组收入均值的函数作为组间不平等的定义,还存在实际应用方面的困境。首先,组间不平等应该反映各亚组的整个收入分布之间的关系,这符合一个良好的不平等指标所必须具备的强洛伦兹一致性(万广华,2004)。其次,亚组均值显然不能代表亚组的全部收入分布信息,亚组均值的函数也就不能反映亚组之间的整体关系。例如,当两个亚组均值相同而收入分布不同时,亚组均值的函数测量到的亚组之间不平等为 0,这显然是不合理的。Blackorby 和 Donaldson (1980)也认为,以亚组均值的函数度量组间不平等的方法,忽视了组内不平等,会导致方法上的伦理任意性(ethically arbitrary procedure),主张组间不平等应该体现组内不平等。Blackorby et al (1981)提出以各

亚组“平均分配的等价收入”(equally distributed equivalent income, EDEI)[①] 的函数度量组间不平等的方法,其实质就是把亚组内部不平等转化为 EDEI 而纳入组间不平等指标。Dagum (1997) 也认为,单纯以亚组收入均值度量组间不平等是过于简单的,因为各亚组间的异方差性及亚组收入分布的不对称性被忽视了。Blackorby et al (1981)、程永宏(2007)分别构造一个例证充分揭示了这一悖论。

第四,Shorrocks (1980)、Foster 和 Shneyerov (2000) 还注意到,上述分解思想意味着,组间不平等对总体不平等的贡献存在着模棱两可的解释:一种解释是指,如果消除组内不平等而组间不平等不变,则总体不平等会是多少;另一种解释是指,如果消除组间不平等而组内不平等不变,则总体不平等会下降多少。这种不一致性正是用亚组均值定义组间不平等所导致的。

由此可见,传统的组间不平等定义以及建立在这一定义基础上的加和可分解性条件并非不容置疑。传统的组间不平等定义恰恰是基尼系数不满足加和可分解性条件的关键因素,也是出现交叠项的根本原因。因此,要实现基尼系数组群分解的突破,必须而且可以放弃对组间不平等指标的某些限定。下文将在重新定义组间不平等指标的基础上,实现不含交叠项的基尼系数组群分解。

三、基尼系数组群分解新方法的理论推导

现有分解方法都是基于基尼平均差定义导出的,本文则使用基于洛伦兹曲线的基尼系数定义,通过基尼系数与收入分布函数的关系导出分解式。

(一) 分解基础和基本假设

利用收入分组数据可以拟合出收入分布函数。再根据 Dorfman (1979)、程永宏 (2006) 可知,任一最高收入为 T、最低收入为 0 的收入分布函数 $F(t)$,相应的基尼系数 G 和平均收入 μ 可表示为:

$$G = \frac{1}{\mu}\int_0^T F(t)\,dt - \frac{1}{\mu}\int_0^T F^2(t)\,dt \tag{4}$$

$$\mu = T - \int_0^T F(t)\,dt \tag{5}$$

$$\text{记} \int_0^T F(t)\,dt = A,\ \int_0^T F^2(t)\,dt = B \tag{6}$$

代入(4)和式(5)式则有: $G = (A - B)/\mu,\ \mu = T - A,\ A - B = \mu G$ (7)

另外,本文沿用并改进程永宏(2006)关于收入分布函数的几个假定:对第 i 个亚组,设其最低收入为 0(只为方便起见,非必需),最高收入为 t_i,其他人的收入分布在 $(0, t_i)$ 内黎曼可积。则收入等于 0 的人数为 1,相应的概率为 $1/N_i$;收入小于 0 的人数为 0,相应的概率亦为 0;收入大于 0 的人数为 $N_i - 1$,相应的概率为 $1 - 1/N_i$,因此,收入分布函数 $F_i(t)$ 在 $t = 0$ 处有一个跳跃间断点。同理,收入分布函数 $F_i(t)$ 在 $t = t_i$ 处也有一个跳跃间断点,即

$$\lim_{t\to 0^-} F_i(t) = F_i(0^-) = 0,\ \lim_{t\to 0^-} F_i(t) = F_i(0^+) = 1/N_i \tag{8}$$

$$\lim_{t\to t_i^+} F_i(t) = F_i(t_i^+) = 1,\ \lim_{t\to t_i^-} F_i(t) = F_i(t_i^-) = 1 - 1/N_i \tag{9}$$

易知,整个总体的最低收入为 0,最高收入为 $T = \max\{t_i\}$。为方便起见,不妨设:

$$t_1 \geq t_2 \geq t_3 \geq \cdots \geq t_m \tag{10}$$

则有:

$$T = \max\{t_i\} = t_1 \tag{11}$$

由上述假定可知: 当 $t > t_i$ 时,$F_i(t) \equiv 1$;当 $t > T$ 时,$F(t) \equiv 1$ (12)

① 对于给定的社会福利函数 W,从实际收入分布 F 转变为每个人都获得收入 ξ 的完全平等的收入分布,如果社会福利保持不变,则称 ξ 是 F(在社会福利函数 W 下)的“平均分配的等价收入”。

(二) 分解方法的理论推导

利用上述假设,我们将推导出新的基尼系数组群分解式。由概率论可知,总体收入分布函数与各亚组收入分布函数的关系为(参见 Bhattacharya and Mahalanobis(1967)):

$$F(t) = \sum_{i=1}^{m} \alpha_i F_i(t) \tag{13}$$

把(13)式代入(4)式,得到整个总体的基尼系数为:

$$G = \frac{1}{\mu}\int_0^T F dt - \frac{1}{\mu}\int_0^T F^2 dt = \frac{1}{\mu}\int_0^{t_1}\left(\sum_{i=1}^{m}\alpha_i F_i\right) dt - \frac{1}{\mu}\int_0^{t_1}\left(\sum_{i=1}^{m}\alpha_i F_i\right)^2 dt \tag{14}$$

把(14)式分子内第二个积分中的平方项展开,并交换求和与积分的次序,得到:

$$G = \frac{1}{\mu}\sum_{i=1}^{m}\int_0^{t_1}\alpha_i F_i dt - \frac{1}{\mu}\left(\sum_{i=1}^{m}\int_0^{t_1}\alpha_i^2 F_i^2 dt + 2\sum_{i=1}^{m}\sum_{j>1}^{m}\int_0^{t_1}(\alpha_i\alpha_j F_i F_j)\, dt\right) \tag{15}$$

由定积分的基本性质及(8)式—(12)式关于 $F(t)$、$F_i(t)$ 和 t_i 性质的假定可知,当 $j > i$ 时,$t_i \geq t_j$;当 $t > t_i \geq t_j$ 时,$F_i = F_j \equiv 1$,于是有:

$$\int_0^{t_1}\alpha_i F_i dt = \int_0^{t_i}\alpha_i F_i dt + \int_0^{t_1}\alpha_i F_i dt = \int_0^{t_i}\alpha_i F_i dt + \alpha_i(t_1 - t_i) \tag{16}$$

$$\int_0^{t_1}\alpha_i^2 F_i^2 dt = \int_0^{t_i}\alpha_i^2 F_i^2 dt + \int_{t_i}^{t_1}\alpha_i^2 F_i^2 dt = \int_0^{t_i}\alpha_i^2 F_i^2 dt + \alpha_i^2(t_1 - t_i) \tag{17}$$

$$\int_0^{t_1}(\alpha_i\alpha_j F_i F_j)\, dt = \int_0^{t_i}(\alpha_i\alpha_j F_i F_j)\, dt + \alpha_i\alpha_j(t_1 - t_i) \tag{18}$$

把(16)式、(17)式、(18)式代入(15)式,并重新组合,得到:

$$G = \frac{\sum_{i=1}^{m}\left(\int_0^{t_i}(\alpha_i F_i - \alpha_i^2 F_i^2)\, dt\right) + \sum_{i=1}^{m}\left(\alpha_i(t_1 - t_i) - \alpha_i^2(t_1 - t_i)\right) - 2\sum_{i=1}^{m}\sum_{j>1}^{m}\left(\int_0^{t_i}(\alpha_i\alpha_j F_i F_j)\, dt + \alpha_i\alpha_j(t_1 - t_i)\right)}{\mu} \tag{19}$$

为方便起见,我们记: $\int_0^{t_i} F_i dt = A_i$,$\int_0^{t_i} F_i^2 dt = B_i$,$\int_0^{t_i} F_i F_j dt = C_{ij}$ (20)

则(19)式可转化为:

$$G = \frac{\sum_{i=1}^{m}\alpha_i(A_i - \alpha_i B_i) + \sum_{i=1}^{m}\alpha_i(t_1 - t_i)(1 - \alpha_i) - 2\sum_{i=1}^{m}\sum_{j>1}^{m}(\alpha_i\alpha_j C_{ij} + \alpha_i\alpha_j(t_1 - t_i))}{\mu} \tag{21}$$

下面我们对(21)式分母中的三个求和表达式分别进行化简。

考虑到 $\sum_{j=1}^{m}\alpha_j = 1$,$1 - \alpha_i = \sum_{j=1}^{m}\alpha_j - \alpha_i = \sum_{j\neq i}^{m}\alpha_j$,$i, j = 1, 2, \cdots, m$,记 $\sum_{j\neq i}^{m}\alpha_j = \overline{\alpha}_i$,则有:

$$\alpha_i = 1 - \overline{\alpha}_i,\ \overline{\alpha}_i = 1 - \alpha_i,\ \overline{\alpha}_i = \sum_{j\neq i}^{m}\alpha_j \tag{22}$$

又由(7)式可知:$A_i - B_i = \mu_i G_i$,再利用(22)式,把(21)式分母中第一个求和表达式转化为:

$$\begin{aligned}\sum_{i=1}^{m}\alpha_1(A_i - \alpha_i B_i) &= \sum_{i=1}^{m}\alpha_i(A_i - (1 - \overline{\alpha}_i) B_i) = \sum_{i=1}^{m}(\alpha_i(A_i - B_i) + \alpha_i\overline{\alpha}_i B_i)\\ &= \sum_{i=1}^{m}\alpha_i\mu_i G_i + \sum_{i=1}^{m}\alpha_i\overline{\alpha}_i B_i = \sum_{i=1}^{m}\alpha_i\mu_i G_i + \sum_{i=1}^{m}\sum_{j\neq i}^{m}\alpha_i\alpha_j B_i\end{aligned} \tag{23}$$

同样地,利用(22)式,把(21)式分母中第二、第三两个求和表达式分别转化为:

$$\sum_{i=1}^{m}\alpha_i(t_1 - t_i)(1 - \alpha_i) = \sum_{i=1}^{m}\alpha_i(t_1 - t_i)\overline{\alpha}_i = \sum_{i=1}^{m}\sum_{j\neq i}^{m}\alpha_i\alpha_j(t_1 - t_i) \tag{24}$$

$$-2\sum_{i=1}^{m}\sum_{j>i}^{m}(\alpha_i\alpha_j C_{ij}+\alpha_i\alpha_j(t_1-t_i)) = -\sum_{i=1}^{m}\sum_{j>1}^{m}2\alpha_i\alpha_j C_{ij}-\sum_{i=1}^{m}\sum_{j>i}^{m}2\alpha_i\alpha_j(t_1-t_i) \quad (25)$$

把(23)式、(24)式、(25)式代入(21)式,得到:

$$G=\frac{\sum_{i=1}^{m}\alpha_i\mu_i G_i+\sum_{i=1}^{m}\sum_{j\neq i}^{m}\alpha_i\alpha_j B_i+\sum_{i=1}^{m}\sum_{j\neq i}^{m}\alpha_i\alpha_j(t_1-t_i)-\sum_{i=1}^{m}\sum_{j>1}^{m}2\alpha_i\alpha_j C_{ij}-\sum_{i=1}^{m}\sum_{j>1}^{m}2\alpha_i\alpha_j(t_1-t_i)}{\mu} \quad (26)$$

为了进一步简化(26)式,我们需要而且可以证明以下两个恒等式(证明可向作者索取):

$$\sum_{i=1}^{m}\sum_{j\neq i}^{m}\alpha_i\alpha_j B_i=\sum_{i=1}^{m}\sum_{j>i}^{m}\alpha_i\alpha_j(B_i+B_j) \quad (27)$$

$$\sum_{i=1}^{m}\sum_{j\neq i}^{m}\alpha_i\alpha_j(t_1-t_i)=\sum_{i=1}^{m}\sum_{j>1}^{m}\alpha_i\alpha_j((t_1-t_i)+(t_1-t_j))=\sum_{i=1}^{m}\sum_{j>1}^{m}\alpha_j\alpha_i(2t_1-t_i-t_j) \quad (28)$$

把(27)式和(28)式代入(26)式得到:

$$G=\frac{\sum_{i=1}^{m}\alpha_i\mu_i G_i+\sum_{i=1}^{m}\sum_{j>i}^{m}\alpha_i\alpha_j(B_i+B_j)+\sum_{i=1}^{m}\sum_{j>i}^{m}\alpha_i\alpha_j(2t_1-t_i-t_j)-\sum_{i=1}^{m}\sum_{j>i}^{m}2\alpha_i\alpha_j C_{ij}-\sum_{i=1}^{m}\sum_{j>i}^{m}2\alpha_i\alpha_j(t_1-t_i)}{\mu} \quad (29)$$

把(29)式分母中含 B_i+B_j 的第 2 项与含 C_{ij} 的第 4 项进行合并,并利用(20)式将 B_i、B_j 和 C_{ij} 还原成定积分的形式,然后利用本文(8)和(9)式关于 $F_i(t)$ 性质的假定进行整理,得到:

$$\sum_{i=1}^{m}\sum_{j>i}^{m}\alpha_i\alpha_j(B_i+B_j)-\sum_{i=1}^{m}\sum_{j>i}^{m}2\alpha_i\alpha_j C_{ij}=\sum_{i=1}^{m}\sum_{j>i}^{m}\alpha_i\alpha_j\int_0^{t}(F_i-F_j)^2\,dt-\sum_{i=1}^{m}\sum_{j>i}^{m}\alpha_i\alpha_j(t_i-t_j) \quad (30)$$

把(30)式代入(29)式得到:

$$G=\sum_{i=1}^{m}\frac{\alpha_i\mu_i}{\mu}G_i+\sum_{i=1}^{m}\sum_{j>i}^{m}\frac{\alpha_i\alpha_j}{\mu}\int_0^{t}(F_i-F_j)^2\,dt \quad (31)$$

由(1)式可知:$\alpha_i\mu_i/\mu=(N_i\mu_i)/N(\mu)=Y_i/Y=\theta_i$,于是(31)式可以简化为:

$$G=\sum_{i=1}^{m}\theta_i G_i+\sum_{i=1}^{m}\sum_{j>i}^{m}\frac{\alpha_i\alpha_j}{\mu}\int_0^{t}(F_i-F_j)^2\,dt \quad (32)$$

令 Y_{ij} 表示第 i、j 两亚组总收入之和,N_{ij} 表示这两个亚组总人口之和,μ_{ij} 表示这两个亚组合并后的平均收入,即:

$$Y_{ij}=Y_i+Y_j,\ N_{ij}=N_i+N_j,\ \mu_{ij}=Y_{ij}/N_{ij}; \quad (33)$$

再令:

$$\alpha_{ij}=N_{ij}/N,\ \theta_{ij}=Y_{ij}/Y,\ \beta_{ij}=N_i/N_{ij},\ \beta_{ij}=N_i/N_{ij} \quad (34)$$

于是(32)式可以转化为(推导过程从略):

$$G=\sum_{i=1}^{m}\theta_i G_i+\sum_{i=1}^{m}\sum_{j>i}^{m}\alpha_{ij}\beta_{ij}\beta_{ij}\frac{1}{\mu_{ij}}\int_0^{t}(F_i-F_j)^2\,dt \quad (35)$$

至此,我们初步获得了一个新的基尼系数组群分解公式:总体基尼系数可以分解成两部分,分别是各亚组基尼系数的加权平均和各亚组分布函数的泛函。第一部分可以解释成组内不平等,这是没有疑义的。第二部分如果能够被合理地解释成组间不平等,则可以说(35)式完成了不含交叠项的基尼系数组群分解。因此,分解式是否有价值,关键在于:把第二部分解释为组间不平等是否具有可靠的理论依据。另外,组间不平等指标和分解式本身是否具有明确的经济意义也是至关重要的(Das and Parikh, 1982)。以下两节将分别讨论这两个问题,并给出最终的分解形式。

四、组间不平等指标的重新定义及其理论依据

(一) 组间不平等指标的基本性质

一般不平等指标的基本性质在过去的文献中有详细论述(Kolm, 1976; Ebert, 1984; Shorrocks, 1984) ,但组间不平等指标的基本性质并没有被专门讨论。不过根据常识,我们可以合理判定组间(绝对和相对)不平等指标应满足以下三个公理:

(R1) 当第 i、j 两个亚组收入分布函数全同时,组间不平等取最小值 0。

(R2) 当第 i、j 两个亚组收入分布函数不全同时,组间不平等大于 0,即使第 i、j 亚组均值相同也应如此。这是因为,均值相同而收入分布不同的一对亚组,在所有"收入分布函数不完全相同的一对亚组"构成的集合中,不具有任何特殊地位,因此其组间不平等不应该取特殊值 0,否则必将出现 Blackorby et al (1981) 、程永宏(2007) 所演示的悖论。

(R3) i、j 亚组的组间不平等,与 j、i 亚组的组间不平等应该相等。这是一般不平等指标所必须具备的基本性质 ——对称性(Kolm, 1976; Ebert ,1984; Shorrocks ,1984 等) 。

(二) 分布函数距离与组间不平等指标的关系

上述公理 R1、R2 正是函数距离数学性质中的非负性,R3 是函数距离的对称性(参见江泽坚等(1994)) 。这表明:组间不平等应该与分布函数之间的距离有关。事实上,1970 年代以来,已经有不少文献提出以分布函数的距离度量组间不平等的思想。根据 Vinod (1985) ,美国人口普查局较早提到收入分布之间的差距,并建议用分布函数的交叠面积度量这一差距。Gastwirth (1975) 对此提出批评,并给出一个新的指标: $TPROB = \int_0^{\infty}(1 - F_1(t))\,dF_2(t)$ 。Dagum (1980) 给出两个收入分布之间"经济距离"和"相对经济富裕度"的精确定义,用来度量一组人口相对于另一组人口的富裕程度。Shorrocks (1982) 对 Dagum 的方法提出批评,认为 Dagum 对经济距离的定义不满足自反性等距离的数学性质,因而缺乏逻辑上的一致性。Vinod (1985) 则认为 Shorrocks 的批评过分强调了距离定义的数学性质,他提出了度量两组人口之间经济不平等的替代指标"经济优势"。Butler 和 McDonald (1987) 提出"分布函数不完全矩"的概念,用两个亚组收入分布函数的 0 阶和 1 阶不完全矩之差来度量"组间不平等"。以上作者的工作表明:组间不平等可以由亚组收入分布函数之间的某种距离来表征,这种距离决不是亚组均值的某个函数,而是亚组分布函数的泛函。

Ebert (1984) 更系统地研究了两个收入分布函数之间的距离与组间不平等的关系。设 $F_1^{-1}(t)$、$F_2^{-1}(t)$ 分别表示两个亚组的收入分布函数的反函数,则这两个收入分布之间的距离被定义为:

$$\rho_r(F_1, F_2) = \left(\int_0^1 \left| F_1^{-1}(t) - F_2^{-1}(t) \right|^r dt\right)^{1/r}, r \geq 1 \qquad (36)$$

Ebert (1984) 证明,上述定义满足以下性质 P1 - P6: (P1) 零标准化(自反性) ; (P2) 对称性; (P3) 三角不等式; (P4) 人口无关性; (P5) 基础无关性(base independence) ; (P6) 线性齐次性[①]。

Ebert 根据这 6 条性质得出结论: (36) 式所定义的分布函数距离是一个良好的组间绝对不平等指标。不过,Chakravarty 和 Dutt (1987) 批评这一指标不能反映亚组间的福利差异。

(三) 分布函数距离和组间绝对不平等指标的新定义

根据本文分解结果的特点,我们定义另外一个分布函数距离。在 (35) 式中,记:

① 其中,P1、P2、P4 是绝对不平等指标的基本性质,其含义与下文相对不平等指标的相应性质 Q1、Q2、Q4 完全相同; P5 是指每个人增加相同数量的收入,不平等水平不变,P6 是指每个人的收入扩大相同倍数,不平等水平也提高相应倍数,二者也是绝对不平等指标一般应该具有的性质;另外,P1、P2、P3 也正好分别是数学上的函数距离所必须满足的非负性、对称性、三角不等式(其中 P1 是非负性的一部分。参见江泽坚等(1994)) 。

$$\left(\int_0^{t_i}(F_i-F_j)^2dt\right)^{1/2}=\rho(F_i,F_j) \tag{37}$$

由(8)式—(10)式关于亚组分布函数 $F_i(t)$ 性质的假定可知,函数 F_i-F_j 必定是可测集$[0,t_i]$上的勒贝格 p 方可积函数,根据泛函分析中的距离空间理论,可以定义 F_i 与 F_j 之间的距离为:

$$\left(\int_0^{t_i}|F_i-F_j|^p dt\right)^{1/p}=\rho_p(F_i,F_j),p\geq 1 \tag{38}$$

并且,可以证明(参见江泽坚、吴智泉(1994),第 218—223 页),这一定义满足距离的基本数学性质 P1—P3,因此,自然也满足上文的公理 R1—R3。当 $p=2$ 时,$\rho(F_i,F_j)$ 就成为本文(37)式定义的 $\rho(F_i,F_j)$。由此可见,$\rho(F_i,F_j)$ 也是对收入分布函数距离的合理定义。

根据 Ebert (1984)、Dagum (1997)等的结论:满足性质 P1—P3 的函数距离定义都可以作为组间绝对不平等指标,因此,$\rho(F_i,F_j)$ 必然也可以作为组间绝对不平等指标。

另外,容易证明:$\rho(F_i,F_j)$ 还满足上述人口无关性 P4 和基础无关性 P5。当然,$\rho(F_i,F_j)$ 不满足线性齐次性 P6。但由(37)式可知,当所有人的收入都倍乘 λ 时,$\rho(F_i,F_j)$ 倍乘 $\sqrt{\lambda}$。可见 $\rho(F_i,F_j)$ 与 λ 仍然是单调增的关系,只不过不是线性关系,这意味着 $\rho(F_i,F_j)$ 的"刻度"不是均匀的。但刻度均匀并不是绝对不平等指标的必要条件,现实中很多工程仪表的刻度也是不均匀的。由此可见:不满足线性齐次性,并不影响 $\rho(F_i,F_j)$ 作为组间绝对不平等指标的合理性。Shorrocks (1984)给出的绝对不平等指标性质也不包含线性齐次性 P6。

进一步地,我们注意到以下两个特征:1) $\rho^2(F_i,F_j)$ 是 $\rho(F_i,F_j)$ 的连续、严格增变换;2) 当且仅当 $\rho(F_i,F_j)=0$ 时 $\rho^2(F_i,F_j)=0$。根据 Shorrocks (1984)可知,在具备这两个特征的前提下,如果 $\rho(F_i,F_j)$ 是合理的不平等指标,则 $\rho^2(F_i,F_j)$ 同样也是合理的不平等指标。另外,$\rho^2(F_i,F_j)$ 不仅与 $\rho(F_u,F_j)$ 一样具备性质 P1、P2、P4、P5(这由 $\rho^2(F_i,F_j)$ 与 $\rho(F_i,F_j)$ 的关系可以直接得证),而且具备线性齐次性 P6(证明可向作者索取)。另外,$\rho^2(F_i,F_j)$ 显然也满足本文的公理 R1—R3。因此,我们最终把作为本文的组间绝对不平等指标定义,记之为:

$$d_{ij}=\int_0^{t_i}(F_1-F_j)^2dt=\rho^2(F_i,F_j) \tag{39}$$

(四)组间相对不平等指标的新定义及其性质

以上论证了 d_{ij} 作为组间绝对不平等指标的依据,下面进一步证明:d_{ij}/μ_{ij} 具备一般相对不平等指标的 5 个性质〔参见 Shorrocks (1984)等〕,因而可以作为组间相对不平等指标。

(Q1)零标准化(自反性):是指不平等指标最小值为 0,当且仅当自变量全同时取得。根据本文公理 R1 和 R2,对于组间不平等指标而言,这应该是指:当且仅当亚组分布函数 F_i 与 F_j 全同时组间不平等指标取最小值 0。由 d_{ij} 的自反性可知,d_{ij}/μ_{ij} 满足这一性质。

(Q2)对称性:是指互换任意两个自变量在不平等指标定义中的位置,不平等指标值不变。根据本文公理 R3,对组间不平等指标而言,这应该是指:互换 F_i、F_j 的位置,组间不平等指标值不变。由 d_{ij} 和 μ_{ij} 的对称性可知,d_{ij}/μ_{ij} 满足这一性质。

(Q3)零次齐次性:是指所有自变量都倍乘时不平等指标值不变。对组间不平等指标而言,这应该是指:对各亚组内每一个体收入都倍乘 λ,组间不平等指标值不变。由 d_{ij} 和 μ_{ij} 的线性齐次性可知,d_{ij}/μ_{ij} 满足这一性质。

(Q4)人口无关性:这是指把每一个体复制相同次数不平等指标值不变。对组间不平等指标而言,这应该是指把所有亚组整体复制相同次数,组间不平等指标值不变。由 d_{ij} 和 μ_{ij} 的人口无关性可知,d_{ij}/μ_{ij} 满足这一性质。

(Q5) Pigou-Dalton 转移原理:是指个体间适当的收入转移可以降低不平等指标值。对组间不平等指标而言,这应该是指亚组间适当的收入转移可以降低组间不平等指标值。由(39)式易知,当两个亚组收入分布函数不相交(最低收入点和最高收入点除外),则其中一个亚组的平均收入必然高于另一个亚组,这时从高收入组向低收入组进行的收入转移,如果规模没有大到使得两个亚组收入分布函数位置发生互换的程度,且同一亚组内每个成员的收入变化都是同方向的,则由(39)式可知:d_{ij}/μ_{ij}必然下降;当两个亚组收入分布函数相交时,情况比较复杂,但从高收入组向低收入组进行适当的收入转移和组内再分配,仍然能够使 d_{ij}/μ_{ij}下降。

这里有必要对组间相对不平等的"人口无关性"Q4 做进一步分析。将人口无关性应用于组间相对不平等指标时,存在两种解释:第一种解释正如 Ebert(1984)所定义的那样,对两个亚组内的每一个体分别进行 k 次和 q 次复制,组间相对不平等保持不变;第二种解释是把每个亚组当作一个不可分的"个体",将所有亚组整体复制相同次数,组间相对不平等保持不变。如果我们把"组间相对不平等"理解为两个亚组整体之间的关系,则后一种解释更符合相对不平等指标人口无关性的原始定义。这两种解释关键区别在于对各亚组是分别复制不同次数还是相同次数。前一种解释意味着组间相对不平等与亚组间人口结构无关,后一种解释则相反。我们可以用城乡差距为例,比较二者的合理性:设想进行这样一种亚组内收入分布不变的"人口复制",使得农村人口份额趋向于 0,城镇人口份额趋向于 1;此时,按第一种解释,城乡差距不变,按第二种解释,城乡差距改变。显然,第二种解释更合理,因为即使城乡各自内部收入分布不变,全国人口向城镇集中也意味着城乡差距缩小。从这个角度看,对于组间相对不平等指标,人口无关性应该是指不平等水平与两个亚组人口之和无关,而不是与亚组间人口结构无关。因此,为了反映组间相对不平等与人口结构的关系,有必要把(35)式中亚组人口结构指标 $\beta_{ij}\beta_{ij}$ 纳入组间相对不平等指标。于是,我们最终把 $\beta_{ij}\beta_{ij} d_{ij}/\mu_{ij}$ 作为第 i、j 组的"组间相对不平等"指标,记之为:

$$\beta_{ij}\beta_{ij} d_{ij}/\mu_{ij} = D_{ij} \tag{40}$$

易知,D_{ij}只与第 i、j 组的特征有关,且与 d_{ij}/μ_{ij}一样,符合相对不平等指标的所有性质 Q1—Q5(Q4 按上述第二种解释定义),因此作为第 i、j 亚组的"组间相对不平等"指标是合理的。

(五)新分解式的最终形式

利用(40)式的组间不平等指标定义,(35)式简化为:

$$G = \sum_{i=1}^{m}\theta_i G_i + \sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_{ij}\theta_{ij}D_{ij} \tag{41}$$

至此,我们获得了最终的基尼系数分解式:总体基尼系数被完全地分解成组内不平等和组间不平等两部分;其中组内不平等是各亚组内部基尼系数的加权平均,权系数为各亚组的收入份额;组间不平等是所有亚组"两两之间"相对不平等的加权和,权系数是相应的两亚组合并后在总体中的人口份额与收入份额之积。下面讨论组间不平等指标和分解式的社会福利含义。

五、新分解式和组间不平等指标的社会福利含义和经济意义

(一)基尼系数的社会福利含义

Sen(1973,1997)、Blackorby 和 Donaldson(1978)等研究了基尼系数的社会福利含义,徐宽(2003)对这类研究进行了综述。设总收入 Y 按向量 $y=\{y_i\}$,$i=1,2,\cdots,N$ 的离散形式分配,其中,y_i 按非减顺序排列,N 为总人口,平均收入设为 μ,Sen(1973)给出了该离散分配的基尼系数表达式:

$$G = 1 + 1/N - 2/(N^2\mu)\sum_{i=1}^{N}(N+1-i)y_i \tag{42}$$

Sen(1973,pp. 32)指出,上述表达式隐含着一个线性齐次社会福利函数;Blackorby and Donaldson

(1978) 进一步研究了基尼系数等不平等指标的社会福利含义,给出"基尼社会福利函数"W 的定义:

$$W(y) = (1/N^2)\sum_{i=1}^{N}(2N - 2i + 1)y_i \tag{43}$$

则由(42) 式可知:

$$G = 1 - W(y)/\mu \tag{44}$$

(43) 式中,W 显然是 y 的线性齐次函数。在这一社会福利函数下,由 y 的平均分配等价收入 ξ (EDEI,参见前文第二节脚注) 的定义可知:

$$W(\xi \cdot e) = W(y) \tag{45}$$

其中是所有分量都是 1 的 N 维向量。由于 W 是线性齐次函数,则由(45) 式得到:

$$W(\xi \cdot e) = \xi W(e) = W(y) \tag{46}$$

可以证明:$W(e) = 1$(参见:徐宽,2003) ,则由(46) 式得到:

$$\xi = W(y) \tag{47}$$

把(47) 式代入(44) 式得到:

$$G = 1 - W(y)/\mu = 1 - \xi/\mu \tag{48}$$

又根据 EDEI 的定义可知:一定的总收入 $Y = N\mu$ 在基尼系数为 G 的不平等分配时所得到的社会福利,与另一总收入 $N\xi$ 在完全平等分配时获得的社会福利是相同的,这意味着不平等分配导致总收入 $Y = N\mu$ 产生的社会福利损失百分比 l 为:

$$l = (N\mu - N\xi)/N\mu = 1 - \xi/\mu \tag{49}$$

由(48) 和(49) 式可知:$G \equiv l$。这意味着:基尼系数 G 恰好反映了一定收入 Y 因不平等分配所造成的社会福利损失的比例;对应的社会福利损失绝对量(以下简称"福利损失") L 为:

$$L = Y - N\xi = Yl = YG \tag{50}$$

以上结论是基于离散分布得出的,但由于基尼系数社会福利含义是基尼系数固有的属性,因此在连续情形中这一含义不会改变,只不过此时基尼社会福利函数成为收入分布函数的泛函。

(二) 新分解式和组间不平等指标的社会福利含义

利用(50) 式可以对本文组间不平等指标 D_{ij} 和最终分解式(41) 式的含义做出清晰的福利经济学解释。对(41) 式两边同乘以总收入 Y 得到:

$$YG = \sum_{i=1}^{m}(Y\theta_i)G_i + \sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_{ij}(Y\theta_{ij})D_{ij} = \sum_{i=1}^{m}Y_iG_i + \sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_{ij}Y_{ij}D_{ij} \tag{51}$$

由(50) 式易知:YG 等于总收入因总体不平等产生的社会福利损失 L,$\sum_{i=1}^{m}Y_iG_i$ 等于各亚组内部不平等产生的社会福利损失之和 $\sum_{i=1}^{m}L_i$,记之为 L_w。则 $\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_{ij}\alpha_{ij}Y_{ij}D_{ij} = L - L_w$ 必然是全部组间不平等产生的社会福利损失,记之为 L_b。显然,社会福利损失 L_b 是 $Y_{ij}D_{ij}$ 的加权和,其中权系数 α_{ij} 是人口份额,其量纲为 1,不具备任何社会福利含义,因此,社会福利损失只能来自 $Y_{ij}D_{ij}$。又由于 $Y_{ij}D_{ij}$ 是 Y_{ij} 和 D_{ij} 的线性增函数,且只与 i、j 亚组有关,因此,$Y_{ij}D_{ij}$ 度量的正是 i、j 亚组收入之和 Y_{ij} 由于组间不平等 D_{ij} 的存在而产生的社会福利损失。把 $Y_{ij}D_{ij}$ 记为 L_{ij},则有 $D_{ij} = L_{ij}/Y_{ij}$,由此可知,我们定义的组间不平等指标 D_{ij},恰好度量了第 i、j 亚组的组间不平等产生的社会福利损失 L_{ij} 占这两个亚组收入之和 Y_{ij} 的比例。这与基尼系数的社会福利含义完全相同,也正是一个理想的不平等指标所应该具备的性质。Atkinson (1970) 、Kolm (1969) 、Sen (1973) 等都认为,具有社会福利含义的不平等指标才是最理想的。

本文分解式(41) 式的社会福利含义也是非常清楚的 ——总体不平等产生的社会福利损失 L,等于全部组内不平等产生的社会福利损失 L_w 与全部组间不平等产生的社会福利损失 L_b 之和,即:

$$L = L_w + L_b \text{,其中,} L = YG, L_w = \sum_{i=1}^{m} Y_i G_i = \sum_{i=1}^{m} L_i, L_b = \sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_{ij} Y_{ij} D_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_{ij} L_{ij} \tag{52}$$

利用(52)式,可以计算出:在总体不平等引起的社会福利损失中,组内和组间不平等的贡献率分别是多少。这进一步证明:基尼系数组群分解式中,组间不平等与总体不平等的定义相同是没有必要的;只要二者福利含义相同,分解式便可以提供社会福利含义明确的结果,具有实际应用价值。

从上述论证过程可以看出:当且仅当分解式中组内不平等的权系数是亚组收入份额时,分解式才会有清晰的社会福利含义,这正是本文分解式的优点之一。

(三)组间不平等指标与基尼系数的关系

设总体有N个人,每个人的收入分别是 $x_1, x_2, \cdots, x_n$。若我们以每个人为一个亚组,可以证明(可向作者索取):

$$G = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\alpha_i \alpha_j}{\mu} \int_0^t (F_i - F_j)^2 dt = \frac{1}{\mu N^2} \sum_{i=1}^{N}\sum_{j=1}^{N} | x_j - x_i | \tag{53}$$

(53)式最右边正好是用基尼平均差定义的基尼系数,而最左边 G 则是用洛伦兹曲线定义的基尼系数。这意味着:对于此种收入分布情形,本文组间不平等部分将转化为用基尼平均差定义的基尼系数,同时,基于洛伦兹曲线和基于基尼平均差的两种基尼系数的定义是等价的〔Kendall (1948) 曾证明,转引自 Soltow(1960)〕。

六、总　结

本文分解式中,亚组一致性(subgroup consistency)不成立,但这不一定是缺点。满足亚组一致性的不平等指标分解式,是以组间不平等指标的缺陷为代价的,如泰尔指数分解式。

本文分解方法的主要优点是:不存在交叠项;分解式具有明确的社会福利含义;组间不平等指标具有可靠的理论基础和社会福利含义,克服了以亚组收入均值定义组间不平等的缺陷。本文基于分组数据的计算方法,相对于使用离散数据的算法,在实际应用方面具有一个优点:离散数据要求被调查者精确报告自己的实际收入,这在实际调查中经常会遇到一定阻力,导致较大误差;而本文算法所需要的分组数据只要求被调查者报告自己收入所在区间,因此更易获得被调查者的配合,容错性能更好。

本文定义的组间不平等与基尼系数的关系值得深入探讨。另外,本文分解方法对收入分布函数的拟合精度提出了更高的要求,这方面问题有待于进一步研究。

参考文献

程永宏,2006:《二元经济中城乡混合基尼系数的计算与分解》,《经济研究》第1期。

程永宏,2007:《改革以来全国总体基尼系数的演变及其分解》,《中国社会科学》第4期。

江泽坚、吴智泉,1994:《实变函数论》(第二版),高等教育出版社。

万广华,2004:《收入分配的度量与分解:一个对于研究方法的评介》,《世界经济文汇》第1期。

徐宽,2003:《基尼系数的研究文献在过去八十年是如何拓展的》,《经济学(季刊)》第2卷第4期。

Atkinson, A. B., 1970, "On the Measurement of Inequality", *Journal of Economic Theory*, Vol. 2, pp244—263.

Bhattacharya, N. and B. Mahalanobis, 1967, "Regional Disparities in Household Consumption in India", *Journal of the American Statistical Association*, Vol. 62, No. 317 (Mar.), pp. 143—161.

Blackorby, C., and D. Donaldson, 1978, "Measures of Relative Equality and Their Meaning in Terms of Social Welfare", *Journal of Economic Theory*, Volume 18, No. 1 (Jun.), pp.59—80.

Blackorby, C., and D. Donaldson, 1980, "A Theoretical Treatment of Indices of Absolute Inequality", *International Economic Review*, Vol. 21, No. 1. (Feb.), pp. 107—136.

Blackorby, C., D. Donaldson, and M. Auersperg, 1981, "A New Procedure for the Measurement of Inequality within and among Population Subgroups", *Canadian Journal of Economics*, Vol. 14, No. 4, pp. 665—685.

Bourguignon, F., 1979, "Decomposable Income Inequality Measures", *Econometrica*, Vol. 47, No. 4 (Jul.), pp. 901—920.

Butler, J. R. and J. B. McDonald, 1987, "Interdistributional Income Inequality", *Journal of Business & Economic Statistics*, Vol. 5, No. 1. (Jan.), pp. 13—18.

Cowell, F., 1980, "On the Structure of Additive Inequality Measures", *Review of Economic Studies*, Vol. 47, No. 3 (Apr.), pp. 521—531.

Chakravarty, S. R. and B. Dutt, 1987, "A Note on Measures of Distance between Income Distrubution", *Journal of Economic Theory*, Vol. 41: pp. 185—188.

Dagum, C., 1980, "Inequality Measures between Income Distributions with Applications", *Econometrica*, Vol. 48, No. 7. (Nov.), pp. 1791—1803.

Dagum, C., 1997, "A New Approach to the Decomposition of the Gini Income Inequality Ratio", *Empirical Economics*, Vol. 22, pp. 515—531.

Das, T and Parikh, A. 1982, "Decomposition of Inequality Measures and a Comparative Analysis", *Empirical Economics*. Vol. 7, No. 1—2, pp. 23—48.

Dorfman, R., 1979, "A Formula for the Gini Coefficient", *Review of Economics and Statistics*, Vol. 61, No. 1 (Feb.), pp. 146—149.

Ebert, U., 1984, "Measures of Distance between Income Distributions", *Journal of Economic Theory*. Vol. 32, No. 2 (Apr.), pp. 266—274.

Fei, J. C. H., G. Ranis, S. W. Y. Kuo, 1978, "Growth and the Family Distribution of Income by Factor Components", *Quarterly Journal of Economics*, Vol. 92, No. 1 (Feb.), pp. 17—53.

Fields, G. S., 1979, "Decomposing LDC Inequality", *Oxford economic papers*, Vol. 31 (3) Nov. pp437—459.

Foster, J. E. and A. A. Shneyerov, 2000, "Path Independent Inequality Measures", *Journal of Economic Theory*, Vol. 91, No. 2, (Apr.), pp. 199—222.

Gastwirth, J. L., 1975, "Statistical Measures of Earnings Differentials", *American Statistician*, Vol. 29, No. 1 (Feb.), pp. 32—35.

Kendall, M. G., 1948, The Advanced Theory of Statistics, Vol. 1, 4th ed., London, pp. 42—46.

Kolm, C., 1969, "The Optimal Production of Social Justice", in Public Economics, ed. by J. Margolis and H. Guitton, London and New York: Macmillan.

Kolm, C., 1976, "Unequal Inequalities", *Journal of Economic Theory*, Vol. 12, Issue 3, pp416—442.

Lambert, P. J. and Decoster, A., 2005, "The Gini Coefficient Reveals More", Working Paper in University of Oregon and University of Leuven, Feb.

Lambert, P. J., and J. Richard Aronson, 1993, "Inequality Decomposition Analysis and the Gini Coefficient Revisited", *Economic Journal*, Vol. 103, No. 420 (Sep.), pp. 1221—1227.

Mangahas, M., 1975, "Income Inequality in the Philippines: A Decomposition Analysis", World Employment Programme Working Paper, No. 12, Geneva, ILO.

Milanovic, B, and S. Yizhaki, 2002, "Decomposing World Income Distribution: Does the World Have a Middle Class?" *Review of Income and Wealth*, Vol. 48, No. 2., pp. 156—178.

Mookherjee, D. and A. F. Shorrocks, 1982, "A Decomposition Analysis of the Trend in UK Income Inequality", *Economic Journal*, Vol. 92, No. 368 (Dec.), pp. 886—902.

Pyat, G., 1976, "On the Interpretation and Disaggregation of Gini Coefficients", *Economic Journal*, Vol. 86, No. 342 (Jun.), pp. 243—255.

Rao, V. M., 1969, "Two Decompositions of Concentration Ratio", *Journal of the Royal Statistical Society*, Series A (General), Vol. 132, No. 3., pp. 418—425.

Sastry, D. V. S. and U. R. Kelkar, 1994, "Note on the Decomposition of Gini Inequality", *Review of Economics and Statistics*, Vol. 76, No. 3 (Aug.), pp. 584—586.

Sen, A., 1973, On Economic Inequality, Oxford: Clarendon Press.

Sen, A., 1997, On Economic Inequality, Enlarged Edition with a Substantial Annexe, Oxford: Clarendon Press.

Shorrocks, A. F., 1980, "The Class of Additively Decomposable Inequality Measures", *Econometrica*, Vol. 48, No. 3, pp. 613—626.

Shorrocks, A. F., 1982, "On the Distance between Income Distributions", *Econometrica*, Vol. 50, No. 5, pp. 1337—1339.

Shorrocks, A. F., 1984, "Inequality Decomposition by Population Subgroup", *Econometrica*, Vol. 52, No. 6, pp. 1369—1386.

Silber, J., 1989, "Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality", *Review of Economics and Statistics*, Vol. 71, No. 1 (Feb.), pp. 107—115.

(下转第 144 页)

The Specification of Population Regression Models in Econometric Application Studies

Li Zinai

(School of Economics and Management, Tsinghua University)

Abstract: Taking the target of econometric population models specification as the starting point, the paper analyses and evaluates the study target guidance, the economics theory guidance and the data guidance of the population models specification, and advances some important criterions about the population models specification. For example, the population model must be sole, general and realistic, and must pass the statistical testing by the data. Finally, the paper raises a new theory, i.e. the population models specification must be directed by the dynamic relationship in economy system.

Key Words: Econometric Models; Population Regression Models; Economics Theory Guidance; Data Guidance; Dynamic Relationship Guidance

JEL Classification: C100, C500, A200

(责任编辑:松　木) (校对:晓　鸥)

(上接第 135 页)

Soltow, L., 1960, "The Distribution of Income Related to Changes in the Distributions of Education, Age, and Occupation", *Review of Economics and Statistics*, Vol. 42, No. 4. (Nov.), pp. 450—453.

Suoniemi, I., 2000, "Decomposing the Gini and the Variation Coefficients by Income Sources and Income Recipients", Working Paper, in Labour Institute for Economic Research, Finland, Oct..

Vinod, H. D. 1985, "Measurement of Economic Distance between Blacks and Whites", *Journal of Business & Economic Statistics*, Vol. 3, No. 1 (Jan.), pp. 78—88.

Yao, Shujie, 1999, "On the Decomposition of Gini Coefficients by Population Class and Income Source, a Spreadsheet Aproach and Application", *Applied Economics*, Vol. 31, pp. 1249—1264.

Yitzhaki, S., 1994, "Economic Distance and Overlapping of Distributions", *Journal of Econometrics*, Vol. 61, pp. 147—159.

Yitzhaki, S., 1998, "More than a Dozen Alternative Ways of Spelling Gini", in Research on Economic Inequality, eds by D.J. Slottje, JAI Press, pp. 13—30.

Zheng, B., 2007, "Unit-Consistent Decomposable Inequality Measures", *Economica*, Vol. 74, issue 293, pp. 97—111.

A New Decomposition of Gini-Coefficient by Population Subgroups

Cheng Yonghong

(School of Public Administration, Renmin University of China)

Abstract: This paper demonstrates a new decomposition of the Gini-coefficient under multi-subgroups. The main contents are as follows. First, the fundamental reasons of imperfections on present decomposition methods are analyzed. Second, a new decomposition of Gini-coefficient under multi-subgroup without overlap term is proposed. Third, the reliable theoretical basis and the clear economic significance of the between-group inequality index are proposed and proved. Lastly, the meaning of welfare economics behind the decomposition and the between-group inequality index are explored.

Key Words: Gini-coefficient; Between-group Measure of Inequality; Welfare Losses

JEL Classification: D63, D31, C43, O15

(责任编辑:荆　岩) (校对:子　璇)