第 8 卷第 1 期　　经 济 学（季 刊）　　Vol. 8, No. 1
2008 年 10 月　　China Economic Quarterly　　October, 2008

一个新的基尼系数子群分解公式
——兼论中国总体基尼系数的城乡分解

洪兴建*

摘　要　基尼系数的子群分解公式有多种形式，不同分解公式的实证结论有较大差异。本文分析了群内基尼系数的权数以及群间不平等指标所应具备的基尼系数内涵，提出了一个新的群间不平等指标，从而给出了一个新的基尼系数子群分解公式。利用这个新的基尼系数子群分解公式，本文实证分析了近年来中国总体基尼系数的城乡分解。实证结果表明，农村群间不平等对总体基尼系数的贡献率最大，城镇群内不平等的贡献率不断上升，这是促使中国总体基尼系数不断上升的两个重要因素。

关键词　基尼系数，子群分解，群间不平等

一、引　　言

分解不平等指标主要有两个方面含义：一是对总体按子群（Subgroup）分解，即将总体按一定标准分为若干个子群，分别研究各个子群内部不平等及其相互之间的不平等，并计算它们对总体不平等的贡献，比如将某个国家的总人口按不同种族、不同年龄或者不同地区等分解的结果就是这种形式；二是按收入成分或来源（Component）进行分解，即将总收入分解为所有收入来源之和，分析每个收入来源的分布对总体不平等的贡献。不平等指标子群分解的一个重要作用是，可以计算群内不平等以及群间不平等对总体不平等的贡献，从而为有关政策的制定提供参考价值。在不平等指标中，方差和泰尔指数（Theil Index）等大部分指标的子群分解公式均是唯一的，而基尼系数的子群分解却有多种形式，也因而成为众多学者争论的一个热点问题，譬如 Rao（1969）、Pyatt（1976）、Mookherjee and Shorrocks（1982）、Shorrocks（1984）、Silber（1989）、Lambert and Aronson（1993）、Dagum

* 浙江工商大学统计学院。通信地址：杭州市下沙高教园区浙江工商大学统计学院，310018；电话：(0571)28008085；E-mail：hongxj18@163.com。作者感谢匿名审稿人及姚洋教授对本文提出的富有建设性的建议，当然文中所有错误和疏漏之处皆由作者负责。感谢国家社会科学基金(07CTJ008)、国家自然科学基金(70671093)、浙江省高校人文社科重点研究基地(浙江工商大学统计学)以及国家统计局重点科研项目(2006B14)的资助。

（1997）以及 Cowell（2000）等。对于 20 世纪 80 年代以前的基尼系数子群分解方法，Das and Parikh（1982）进行了很好评述；徐宽（2003）则对基尼系数的计算方法做了很好综述，并对基尼系数的分解进行了简要评论。由于中国是一个典型的二元经济社会，居民收入数据天然地分成了城镇居民收入和农村居民收入两套数据，这给研究全体居民收入差距带来了一定困难，也给学者们创造了一个争论的话题。譬如李实（2002）和陈宗胜（2002）就对我国总体基尼系数的估算进行了有益探讨。此外，程永宏（2006）推导出了一个关于基尼系数的城乡分解公式，董静和李子奈（2004）分析了对数正态分布条件下城乡加权基尼系数的计算方法，胡祖光（2004）根据中国城乡数据的特点利用一个简易公式估算了中国总体基尼系数。一个非常令人困惑的问题是，针对中国总体基尼系数的城乡分解结果，关于城乡之间差距的贡献率却有很多结论，各种版本的城乡差距贡献率大约分布在 20%到 80%之间（具体数值参见陈宗胜和周云波（2002）、Shorrocks and Wan（2005）以及程永宏（2007）的结论）。如果说不同的不平等指标实证结果不一致还是可以理解的，那么同样是基尼系数的分解结果却产生很大分歧，就很难理解和接受了。导致这一问题的部分原因是采用了不同数据或数据处理方法不同，但根本原因还是基尼系数子群分解公式的不同，因此有必要探究科学的分解方法。

不平等指标的子群分解思想主要来源于统计分析中的方差分析，即群内方差的加权平均反映了群内变异对总体变异的影响，而以每个子群算术平均数计算的群间方差反映了群间变异总体变异的影响。但是基尼系数子群分解的难点主要存在于两个方面：一是如何确定群内基尼系数的权数，二是如何构造群间不平等指标。本文致力于分析这两个难点问题，在此基础上提出一个新的基尼系数子群分解公式，并运用这个分解公式对中国总体基尼系数的城乡分解进行实证分析。

二、基尼系数子群分解方法简述

为方便说明，我们先定义一些数学符号。设总共 n 个单位的收入按升序排列为 $X=(x_1,x_2,\cdots,x_n)$，平均收入为 $\mu=\frac{\sum_{i=1}^{n}x_i}{n}$，基尼系数为 G。假如总体被分成了 s 个子群，其中第 k 个子群记为 $\Omega_k(k=1,2,\cdots,s)$，该子群的第 i 个单位为 x_{ki}，单位数为 n_k，平均收入为 μ_k，人口份额为 $\nu_k=\frac{n_k}{n}$，收入份额为 $\theta_k=\frac{n_k\mu_k}{n\mu}$，基尼系数为 G_k。反映子群内部收入差距与子群之间收入差距的项分别记为 G_W 和 G_B。

关于基尼系数的子群分解大致有三种思路。第一种思路为先求出群内不平等的贡献额，然后将剩下部分归结为群间不平等的贡献额。虽然学者们对于群内基尼系数的形式完全一致，但针对群内基尼系数的权数却有不同主张。Rao（1969）主张组内基尼系数的权数为各群人口份额，即

$$G = \sum_{k=1}^{s}\nu_k G_k + G_B. \tag{1}$$

Mangahas（1975）则认为组内基尼系数的权数为各群的收入份额，即

$$G = \sum_{k=1}^{s}\theta_k G_k + G_B. \tag{2}$$

当然式（1）与式（2）的组间不平等 G_B 并不相同，每个 G_B 有着不同的计算形式[1]，我们说它们是总体基尼系数减去群内不平等的剩余部分，主要是从数值方面以及 G_B 的结构方面讲的。

第二种思路为先求出群间不平等的贡献额，然后将剩下部分归结为群内不平等的贡献额。比如 Bhattacharya and Mahalanobis（1967）提出的分解方法是先求出群间基尼系数贡献额，即 $G_B = \frac{1}{2}\sum_{k=1}^{s}\sum_{l=1}^{s}\nu_k\nu_l\left|\frac{\mu_k}{\mu}-\frac{\mu_l}{\mu}\right|$，再用总体基尼系数减去 G_B 表示组内不平等 G_W 的贡献额，即总体基尼系数为

$$G = \frac{1}{2}\sum_{k=1}^{s}\sum_{l=1}^{s}\nu_k\nu_l\left|\frac{\mu_k}{\mu}-\frac{\mu_l}{\mu}\right| + G_W. \tag{3}$$

Das and Parikh（1982）认为上述三个式子或者是先计算群内不平等，然后将总体不平等减去群内不平等后的剩余项作为群间不平等；或者是先计算群间不平等，用总体不平等减去群间不平等后的剩余项作为群内不平等，这些都不是真正的统计分解。他们认为一个严格的统计分解应该是这样的：群间不平等 I_B 仅仅取决于各个子群的平均收入，群内不平等 I_W 仅仅取决于群内不平等，即 $I_B = f(\mu_1,\mu_2,\cdots,\mu_s)$，$I_W = g(I_1,I_2,\cdots,I_s) = \sum_{k=1}^{s} w_k I_k$，其中 w_k 表示一个合适的权数。Das and Parikh（1982）认为，各种分解方法的价值取决于其背后的经济解释和说服力。

第三种思路为分别求出群内和群间不平等的贡献额，再加上一个可能的剩余项 R，其中群内不平等的贡献额为群内基尼系数的加权，群间不平等就是以各群算术平均数计算出的群间基尼系数，剩余项 R 是由于各群之间可能的交叠（Overlapping）而产生的。比如 Mookherjee and Shorrocks（1982）给

[1] 由于这两个等式中的 G_B 比较复杂，涉及一些其他符号，因为这些不是本文重点，所以没有将它们完整地表示出来，欲详细了解分解形式的读者可以参看 Rao（1969）、Mangahas（1975）以及 Das and Parikh（1982）等。

出的一个公式为

$$G=\sum_{k=1}^{s}\nu_k\theta_k G_k+\frac{1}{2}\sum_{k=1}^{s}\sum_{l=1}^{s}\nu_k\nu_l\left|\frac{\mu_k}{\mu}-\frac{\mu_l}{\mu}\right|+R. \tag{4}$$

应该说，20 世纪 80 年代以来，式（4）的分解形式得到了多数学者的认同，相当部分文献围绕着式（4）进行探讨，一方面利用新的方法推导该式（Silber，1989；Yao，1999），另一方面从不同方面解释剩余项 R。对于剩余项的含义则经历了一个从不了解到相对了解的过程，Mookherjee and Shorrocks（1982）认为不可能得到剩余项的精确含义，这是一个仅仅起维持等式平衡作用的项；Silber（1989）认为剩余项是由于不同子群之间的交错造成的交互影响项（Interaction Effect），反映了子群收入分布的交错程度，而 R 的这个特征又被 Yitzhaki and Lerman（1991）用来度量收入分层（Stratification），R 越大反映收入分层的程度越低；Lambert and Aronson（1993）则通过图形解释了剩余项的含义。

此外，将所有子群按 μ_k 从小到大排序后，Dagum（1997）给出的一个分解公式为：

$$G=\sum_{k=1}^{s}\nu_k\theta_k G_k+\sum_{k=2}^{s}\sum_{l=1}^{k-1}(\nu_k\theta_l+\nu_l\theta_k)G'_{kl}, \tag{5}$$

其中 G'_{kl} 定义为第 k 群与第 l 群的广义基尼比率（Extended Gini Ratio），计算公式为 $G'_{kl}=\dfrac{\Delta_{kl}}{\mu_k+\mu_l}=\sum_{i=1}^{n_k}\sum_{r=1}^{n_l}\dfrac{|x_{ki}-x_{lr}|}{n_k n_l(\mu_k+\mu_l)}$。Dagum（1997）在定义经济富裕（Economic Affluence）概念[2]的基础上将式（5）化为：

$$G=\sum_{k=1}^{s}\nu_k\theta_k G_k+\sum_{k=2}^{s}\sum_{l=1}^{k-1}(\nu_k\theta_l+\nu_l\theta_k)G'_{kl}D_{kl}+\sum_{k=2}^{s}\sum_{l=1}^{k-1}(\nu_k\theta_l+\nu_l\theta_k)G'_{kl}(1-D_{kl}), \tag{6}$$

其中式（6）右边的第二项和第三项分别定义为群间不平等的净贡献率与转移变异贡献率，这两项之和表示群间不平等的贡献率。可以证明，式（6）的三

[2] 首先将平均绝对差距作如下变形：

$$\Delta_{kl}=\frac{\sum_{i=1}^{n_k}\sum_{j=1}^{n_l}|x_{ki}-x_{lj}|}{n_k n_l}=\frac{\sum_{i=1}^{n_k}\sum_{x_{ki}>x_{lj}}(x_{ki}-x_{lj})}{n_k n_l}+\frac{\sum_{i=1}^{n_k}\sum_{x_{ki}<x_{lj}}(x_{lj}-x_{ki})}{n_k n_l}$$

该式右边的第一项和第二项分别记为 d_{kl} 和 p_{kl}，其中 d_{kl} 定义为第 k 个子群与第 l 个子群之间总的经济富裕（Gross Economic Affluence），p_{kl} 定义为转移变异（Transvariation，即个体差值 $x_{ki}-x_{lr}$ 与子群平均数差值 $\mu_k-\mu_l$ 符号不一致）的一阶矩，这样 $d_{kl}-p_{kl}$ 就表示第 k 个子群相对于第 l 个子群的净经济富裕（Net Economic Affluence），$D_{kl}=(d_{kl}-p_{kl})/\Delta_{kl}$ 表示第 k 个子群相对于第 l 个子群的相对经济富裕（Relative Economic Affluence）。

项在数量上与式（4）的三项分别相等，从这个角度讲，式（6）的第三项也是对 R 的一种诠释。

不难发现，按照式（1）、（2）或（3）的分解方法，G_W 和 G_B 总有一个缺乏明确的经济意义，实证分析时虽然简便但解释能力较差。应该说，式（4）的三项分解比式（3）的两项分解能提供更多的信息，式（4）中的 G_W 和 G_B 也均有明确的经济意义，但美中不足的是 G_B 度量群间收入不平等不尽合理（详细说明见第四部分），由此产生的剩余项 R 即便有其他方面的含义（比如可以度量收入分层），也因为这些含义与不平等不是同一个概念，致使式（4）中 G_B 和 R 对总体收入不平等 G 的解释是不明晰的。式（6）利用经济富裕解释群间不平等很有新意，但是 G'_{kl} 与基尼系数的内涵存在一定差异（详细说明见第五部分），解释效果与式（4）其实也是等价的。因此，一个理论上可行、实际意义明确并且与基尼系数内涵相一致的分解公式将是非常有用的。

三、群内基尼系数的权数

从前面列举的分解公式可以看出，群内基尼系数的权数有 ν_k、θ_k 和 $\nu_k\theta_k$ 等三种形式，因而有必要探究较为合理的权数。为了便于论述，我们以两个子群为例进行说明，设低收入群和高收入群分别为 $X'=(x'_1,x'_2,\cdots,x'_m)$ 和 $X''=(x''_1,x''_2,\cdots,x''_{n-m})$，群内收入均按升序排列，$n>m$，且低收入群和高收入群的洛伦茨（Lorenz）曲线分别如图 1 和图 2 中的 D 与 U 所示。

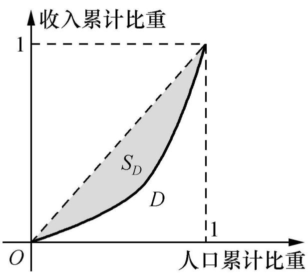

图 1　低收入群洛伦茨曲线

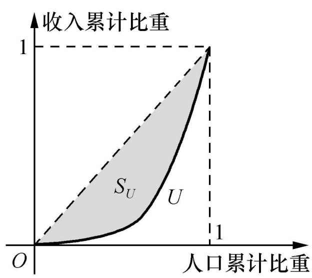

图 2　高收入群洛伦茨曲线

首先，如果两个子群之间没有交错现象，即 $X'=(x_1,x_2,\cdots,x_m)$，$X''=(x_{m+1},x_{m+2},\cdots,x_n)$。若任意一个收入 x_i 属于低收入群，则它在 D 上的坐标为 $\left(\dfrac{i}{m},\dfrac{\sum_{j=1}^{i}x_j}{\sum_{j=1}^{m}x_j}\right)$，其中 $1\leqslant i\leqslant m$；但若从总体角度看，它在图 3 中总体洛伦茨曲线 T 上的坐标应该为：

$$\left(\frac{i}{n},\frac{\sum_{j=1}^{i}x_j}{\sum_{j=1}^{n}x_j}\right)=\left(\frac{m}{n}\cdot\frac{i}{m},\frac{\sum_{j=1}^{m}x_j}{\sum_{j=1}^{n}x_j}\cdot\frac{\sum_{j=1}^{i}x_j}{\sum_{j=1}^{m}x_j}\right)=\left(\nu_1\cdot\frac{i}{m},\theta_1\cdot\frac{\sum_{j=1}^{i}x_j}{\sum_{j=1}^{m}x_j}\right). \tag{7}$$

若任意一个收入 x_i 属于高收入群，则它在 U 上的坐标为 $\left(\frac{i-m}{n-m},\frac{\sum_{j=m+1}^{i}x_j}{\sum_{j=m+1}^{n}x_j}\right)$，其中 $m+1\leqslant i\leqslant n-m$，但是该收入在总体洛伦茨曲线 T 上的坐标为：

$$\left(\frac{i}{n},\frac{\sum_{j=1}^{i}x_j}{\sum_{j=1}^{n}x_j}\right)=\left(\frac{m}{n}+\frac{n-m}{n}\frac{i-m}{n-m},\frac{\sum_{j=1}^{m}x_j}{\sum_{j=1}^{n}x_j}+\frac{\sum_{j=m+1}^{n}x_j}{\sum_{j=1}^{n}x_j}\cdot\frac{\sum_{j=m+1}^{i}x_j}{\sum_{j=m+1}^{n}x_j}\right)$$

$$=\left(\nu_1+\nu_2\cdot\frac{i-m}{n-m},\theta_1+\theta_2\cdot\frac{\sum_{j=m+1}^{i}x_j}{\sum_{j=m+1}^{n}x_j}\right). \tag{8}$$

由式（7）不难看出，当图 3 中的横坐标小于等于 ν_1 时，T 上任意点的横坐标和纵坐标分别相当于图 1 中 D 上某个对应点的 ν_1 和 θ_1 倍；而由式（8）可以发现，当图 3 中的横坐标大于 ν_1 时，T 上任意点的横坐标和纵坐标分别减去 ν_1 和 θ_1 后，相当于图 2 中 U 上某个对应点的 ν_2 和 θ_2 倍。因而对应的面积关系为 $S_1=\nu_1\theta_1 S_D$，$S_2=\nu_2\theta_2 S_U$。又由于两个子群的基尼系数分别为 $G_1=2S_D$，$G_2=2S_U$，总体基尼系数 $G=2(S_1+S_2+S_{\triangle OBC})$，所以 $G=\nu_1\theta_1 G_1+\nu_2\theta_2 G_2+2S_{\triangle OBC}$，从而表明两个子群内部不平等对总体不平等的贡献额分别为 $\nu_1\theta_1 G_1$ 和 $\nu_2\theta_2 G_2$（此时的 $2S_{\triangle OBC}$ 反映两个子群之间的不平等）。

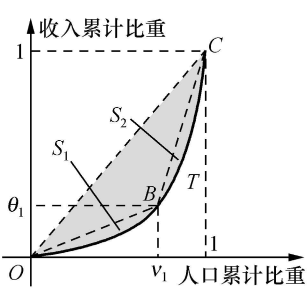

图 3 总体洛伦茨曲线（子群无交错）

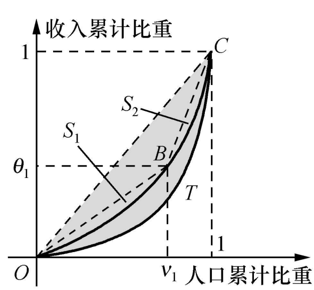

图 4 总体洛伦茨曲线（子群交错）

其次，如果两个子群之间存在交错现象，即低收入群中部分高收入者的收入大于高收入群中部分低收入者的收入。如果总体不是严格按照从低到高

排序，而是先对低收入群进行排序，接着再对高收入群排序，这样得到图 4 中的曲线 OBC，当然实际的总体洛伦茨曲线仍然为 T。设 D 上任意一点的坐标为 $\left(\frac{i}{m},\frac{\sum_{j=1}^{i}x_j'}{\sum_{j=1}^{m}x_j'}\right)$，那么从总体的角度（曲线 OBC）看，该点坐标为：

$$\left(\frac{i}{n},\frac{\sum_{j=1}^{i}x_j'}{\sum_{j=1}^{n}x_j}\right)=\left(\frac{m}{n}\cdot\frac{i}{m},\frac{\sum_{j=1}^{m}x_j'}{\sum_{j=1}^{n}x_j}\cdot\frac{\sum_{j=1}^{i}x_j'}{\sum_{j=1}^{m}x_j'}\right)=\left(\nu_1\cdot\frac{i}{m},\theta_1\cdot\frac{\sum_{j=1}^{i}x_j'}{\sum_{j=1}^{m}x_j'}\right). \tag{9}$$

同理，对于 U 上任意一点的坐标 $\left(\frac{i}{n-m},\frac{\sum_{j=1}^{i}x_j''}{\sum_{j=1}^{n-m}x_j''}\right)$，从总体看的坐标应为：

$$\begin{aligned}\left(\frac{m+i}{n},\frac{\sum_{j=1}^{m}x_j'+\sum_{j=1}^{n-m}x_j''}{\sum_{j=1}^{n}x_j}\right)&=\left(\frac{m}{n}+\frac{n-m}{n}\cdot\frac{i}{n-m},\frac{\sum_{j=1}^{m}x_j'}{\sum_{j=1}^{n}x_j}+\frac{\sum_{j=1}^{n-m}x_j''}{\sum_{j=1}^{n}x_j}\cdot\frac{\sum_{j=1}^{n-m}x_j''}{\sum_{j=1}^{n-m}x_j''}\right)\\&=\left(\nu_1+\nu_2\cdot\frac{i}{n-m},\theta_1+\theta_2\cdot\frac{\sum_{j=1}^{n-m}x_j''}{\sum_{j=1}^{n-m}x_j''}\right).\end{aligned} \tag{10}$$

不难看出，式（9）、式（10）与式（7）、式（8）在本质上是一致的，因而均有 $S_1=\nu_1\theta_1 S_D$，$S_2=\nu_2\theta_2 S_U$。所不同的是，当子群之间不存在交错时，图 1 中的 D 和图 2 中的 U 按一定比例缩小后恰好为图 3 中的总体洛伦茨曲线 T 的一部分；而当两个子群之间存在交错时，D 和 U 按一定比例缩小后将不再是 T 的一部分（如图 4 所示）[3]。

虽然上面只是针对两个子群进行了分析，但对于任意个子群而言，我们完全可以类似地推导出上述结论，即第 k 群的群内不平等对总体基尼系数的贡献为 $\nu_k\theta_k G_k$，从而表明群内基尼系数 G_k 的权数应该为 $\nu_k\theta_k$。

金成武（2007）用矩阵向量说明了基尼系数子群分解的一般形式、意义以及分解的非唯一性，讨论了基尼系数子群分解的数学实质，揭示了各种分解形式的内在联系。金成武（2007）的研究表明，对于基尼系数子群分解可

[3] Lambert and Aronson(1993)作过类似分析，但他们的目的主要是用几何图形解释前面的式(4)，本文则重点说明群内基尼系数的权数，对式(4)并不完全认同。

能的多种组合结果,探求合适分解形式是非常必要的。从群内不平等对总体不平等的作用程度看,本文认为群内基尼系数的权数应该是唯一的,即为$\nu_k\theta_k$,其他形式的权数均有失科学性。

四、具有基尼系数内涵的群间不平等测度指标

由于一般认为群间不平等仅仅取决于各个子群的平均收入,因而通常用基尼系数度量的群间不平等就是$G_B=\frac{1}{2\mu}\sum_{k=1}^{s}\sum_{l=1}^{s}\nu_k\nu_l|\mu_k-\mu_l|$,即以子群算术平均数$\mu_k$为代表值计算出的基尼系数。但是$\mu_k$作为子群代表值可能存在一定偏差,因为它的代表性高低与群内各单位的收入分布密切相关,群内不平等越大,其代表性就越差。比如两个子群收入分布分别为(1,1,28)与(10,10,10),从社会福利的角度看,没有任何理由认为两个子群之间的差距为0,但如果利用算术平均数测度差距,两个子群之间就没有差距。实际上,第一个子群平均每个人的福利状况要低于第二个子群,因此在比较子群之间的不平等时,Blackorby、Donaldson and Auerserg(1981)强烈建议用平均分配等值收入(equally distributed equivalent income)代替算术平均数。Atkinson(1970)提出的平均分配等值收入就是幂平均数$\left(\sum_{i=1}^{n}x_i^l/n\right)^{1/q}$(当$q\neq0$时)或几何平均数$\left(\prod_{i=1}^{n}x_i\right)^{1/n}$,由于$q<1$的幂平均数赋予低收入者较大权重($q$越小,低收入者赋予的权重就越大),一般认为它比算术平均数能更好地反映子群之间的不平等。当然参数q的取值不同,群间不平等的具体数值也会有所不同,这在一定程度上影响了它在实证分析中的应用。

应该说如何科学测度群间不平等一直以来也是学者们深入研究的一个议题,除了等值收入外,Dagum(1980)和Yitzhaki(1994)等探讨了经济距离(Economic Distance)测度群间差距的思路,Fossett and South(1983)分析了群间不平等的概念及测度准则,Vinod(1985)and Gastwirth(1985)提出了经济优势(economic advantage)的分析思路,Ebert(1984)利用分布函数距离进行测度。与群内不平等一样,在建立公理的基础上探讨合适的群间不平等指标正成为一种研究范式,这无疑为科学测度群间不平等提供了一个规范和准则。但笔者以为,对任何一个不平等指标进行子群分解时,需要根据该指标的内涵构造群间不平等的测度方法,至于这个群间不平等指标自身的合理性,那不是分解所能解决的问题。由于每个不平等指标的内涵不一样,其对应的群间不平等指标也应有所区别。

若总体收入分布$X=(x_1,x_2,\cdots,x_n)$,根据Runciman(1966)以及Yitzhaki(1979)定义的相对剥夺(Relative Deprivation),第i个单位跟第j

个单位相比而言的相对剥夺为：

$$\delta_{ij}=\begin{cases}x_j-x_i, & \text{若 } x_i<x_j,\\ 0, & \text{若 } x_i\geqslant x_j,\end{cases}\tag{11}$$

则第 i 个单位的平均相对剥夺为 $\delta_i=\dfrac{\sum_{j=1}^{n}\delta_{ij}}{n}$，所有单位平均的相对剥夺为 $\bar{\delta}=\dfrac{\sum_{i=1}^{n}\delta_i}{n}$。由于基尼系数 G 为：

$$\begin{aligned}G&=\frac{1}{2n^2\mu}\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i-x_j|=\frac{1}{n^2\mu}\sum_{i=1}^{n}\sum_{j=1}^{n}\max(0,x_j-x_i)\\&=\frac{1}{n^2\mu}\sum_{i=1}^{n}\sum_{j=1}^{n}\delta_{ij}=\frac{\sum_{i=1}^{n}\delta_i}{n\mu}=\frac{\bar{\delta}}{\mu},\end{aligned}\tag{12}$$

因此基尼系数可以看做所有单位平均的相对剥夺与平均数的比值。应该说，相对剥夺 δ_{ij} 的概念充分反映了基尼系数的本质内涵，$G=\dfrac{\sum_{i=1}^{n}\delta_i}{n\mu}=\dfrac{\bar{\delta}}{\mu}$ 是体现基尼系数内涵的一个计算公式。

若某个收入 x_i 属于 Ω_k，则该收入的平均相对剥夺为：

$$\begin{aligned}\delta_i&=\frac{\sum_{j=1}^{n}\delta_{ij}}{n}=\frac{\sum_{j\in\Omega_1}\delta_{ij}+\sum_{j\in\Omega_2}\delta_{ij}+\cdots+\sum_{j\in\Omega_s}\delta_{ij}}{n}\\&=\frac{n_1}{n}\frac{\sum_{j\in\Omega_1}\delta_{ij}}{n_1}+\frac{n_2}{n}\frac{\sum_{j\in\Omega_2}\delta_{ij}}{n_2}+\cdots+\frac{n_s}{n}\frac{\sum_{j\in\Omega_s}\delta_{ij}}{n_s},\end{aligned}\tag{13}$$

其中 $\dfrac{\sum_{j\in\Omega_k}\delta_{ij}}{n_k}$ 表示 x_i 与本群内部所有单位的平均相对剥夺，$\dfrac{\sum_{j\in\Omega_l,l\neq k}\delta_{ij}}{n_l}$ 反映了 x_i 与子群 Ω_l 所有单位相比的平均相对剥夺。从而 Ω_k 与 $\Omega_l(k\neq l)$ 相比而言的平均相对剥夺为：

$$\bar{\delta}_{kl}=\frac{\sum_{i=1}^{n_k}\sum_{j\in\Omega_l,l\neq k}\delta_{ij}}{n_kn_l}.\tag{14}$$

$\bar{\delta}_{kl}$ 反映了子群 Ω_k 跟子群 Ω_l 相比的绝对不平等，用它除以 Ω_k 的平均收入 μ_k 后，可以反映子群 Ω_k 跟子群 Ω_l 相比而言的相对不平等，记为 G_{kl}，即

$$G_{kl}=\frac{\sum_{i=1}^{n_k}\sum_{j\in\Omega_l,l\neq k}\delta_{ij}}{n_k n_l \mu_k}. \tag{15}$$

由于子群 Ω_k 内所有个体之间的相对不平等为 $G_{kk}=\frac{\sum_{i=1}^{n_k}\sum_{j\in\Omega_k}\delta_{ij}}{n_k n_k \mu_k}=\frac{\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}\delta_{ij}}{n_k^2\mu_k}$，它恰好是子群 Ω_k 的群内基尼系数，而 G_{kl} 与 G_{kk} 的结构完全相同，因此 G_{kl} 是一个具有基尼系数内涵的群间不平等指标。

作为符合基尼系数内涵的一个新的群间不平等指标，G_{kl} 主要有三个方面的特性。首先，一般来说 $\bar{\delta}_{kl}\neq\bar{\delta}_{lk}$，$G_{kl}\neq G_{lk}$。即 G_{kl} 中的 k 和 l 不是对等的关系，G_{kl} 测度的是第 k 个子群同第 l 个子群相比而言的相对不平等，关注的是第 k 个子群的收入不平等，比较对象是第 l 个子群的收入；而 G_{lk} 测度的是第 l 个子群同第 k 个子群相比而言的收入不平等，关注的是第 l 个子群的收入不平等，比较对象是第 k 个子群的收入。不满足对等性有点类似数学中的"A 数比 B 数多 10%"并非"B 数比 A 数少 10%"，两种说法对应的基数不一样。当我们将不平等与相对剥夺联系在一起时，不平等就是与外界相比较后的一种感受，通常低收入者能够感受到不平等，而高收入者并没有这种感受，因此不满足对等性是比较合理的。此外，群间不平等是各子群所有单位相比其他子群所有单位而言的，即便是平均收入最高的子群，只要该子群与其他子群的收入存在交错部分，就存在相对不平等。最后，群间不平等 G_{kl} 的最小值为 0（针对收入最高且与其他子群没有交错的子群来说），但它的最大值可能大于 1，因此它不具备基尼系数的值域[0,1]。

五、一个新的基尼系数子群分解公式

Sen（2003）指出，如果不平等指标是孤立地对待每一个单位（比如方差和一般熵指数等），而不是将每一个单位与其他单位相比较，则该指标一定满足子群可分解性；反之，如果不平等指标是由每一个单位与所有其他单位相比较而构成的函数（比如基尼系数），则该不平等指标不满足子群可分解性。Sen（2003）的概括无疑非常精炼，但他所指的可分解形式显然受了传统理解的约束，即群间不平等就是以子群算术平均数为自变量的某个函数。但是如果群间不平等指标符合基尼系数的内涵，即本文提出的 G_{kl}，情况又会如何呢？

根据式（12）和式（13）有

$$G=\frac{\sum_{i=1}^{n}\delta_i}{n\mu}=\frac{\sum_{k=1}^{s}\sum_{i=1}^{n_k}\delta_i}{n\mu}=\sum_{k=1}^{s}\left(\sum_{i=1}^{n_k}\frac{n_k^2\mu_k}{n^2\mu}\frac{\sum_{j\in\Omega_k}\delta_{ij}}{n_k^2\mu_k}+\sum_{i=1}^{n_k}\frac{n_k n_l\mu_k}{n^2\mu}\frac{\sum_{j\in\Omega_l,l\neq k}\delta_{ij}}{n_k n_l\mu_k}\right)$$

$$=\sum_{k=1}^{s}\nu_k\theta_k G_{kk}+\sum_{k=1}^{s}\sum_{l,l\neq k}\nu_l\theta_k G_{kl}. \tag{16}$$

式（16）表明基尼系数满足子群可分解性，而且群间不平等指标与群内不平等指标的意义一致，均具有基尼系数的内涵。对比现有的基尼系数子群分解公式，式（16）至少有以下三个方面的特点：

第一，式（16）右边的前一项 $\sum_{k=1}^{s}\nu_k\theta_k G_{kk}$ 表示各群群内基尼系数的加权平均，它跟 G 的比值可以反映群内不平等对总体基尼系数的贡献；后一项 $\sum_{k=1}^{s}\sum_{l,l\neq k}\nu_l\theta_k G_{kl}$ 表示的是每个子群跟其他子群相比的群间不平等的加权平均，该项跟 G 的比值反映了群间不平等对总体基尼系数的贡献。此外，$\nu_k\theta_k G_{kk}/G$ 与 $\nu_l\theta_k G_{kl}/G$ 分别表示第 k 群的群内不平等以及第 k 群与第 l 群相比的群间不平等对总体不平等的贡献率，且 $\sum_{l,l\neq k}\nu_l\theta_k G_{kl}/G$ 测度了第 k 群与所有其他子群相比而言的群间不平等对总体不平等的贡献率。

第二，式（16）中所有群内不平等指标与群间不平等指标的权数之和为 1，即 $\sum_{k=1}^{s}\nu_k\theta_k+\sum_{k=1}^{s}\sum_{l,l\neq k}\nu_l\theta_k=\sum_{k=1}^{s}\nu_k\theta_k+\sum_{k=1}^{s}(1-\nu_k)\theta_k=1$。也就是说，这个分解公式是真正意义上的加权平均。[4]

第三，与基尼系数的其他分解公式相比，式（16）的科学性比较突出，应用性也更强。式（16）表明，如果按照基尼系数的内涵定义群间不平等，那么式（4）中的剩余项 R 就不复存在，而且能区分每个子群的群间不平等对总体不平等的贡献，解释能力大大增强。虽然式（5）与式（16）的外形非常相似，但两个分解公式第二项的构造和解释完全不同。应该说，就绝对差距 $\Delta_{kl}=\sum_{i=1}^{n_k}\sum_{r=1}^{n_l}\frac{|x_{ki}-x_{lr}|}{n_k n_l}$ 而言，它测度的是第 k 群任一单位与第 l 群任一单位之间的平均绝对差距，这是非常合理的，但作为反映平均相对差距的 G'_{kl}，等于绝对差距除以两个子群平均收入之和，即为 $G'_{kl}=\Delta_{kl}/(\mu_k+\mu_l)$，如此定义应该不符合基尼系数的内涵。更重要的是从实际应用角度看，式（16）能够分析每个子群的群间不平等的大小及其贡献，而式（5）只能笼统说明两个子

---

[4] 可能是受泰尔指数分解公式的影响，多数学者认为群内不平等的权数之和应该等于 1（Cowell，2000）。笔者认为，当把群内不平等与群间不平等同等看待时，应该是所有权数之和等于 1，从某种意义上看，这才是真正的加权平均。当然每个不平等指标均有着特定内涵及分解形式，并不能强求所有不平等指标分解时权数之和均等于 1。

群之间的不平等，不能深入分析这个不平等的内在结构。式（6）将群间差距分解为两个部分，从经济富裕的角度解释群间不平等具有一定新意，但这两个部分在数值上等同于式（4）的后两项，解释效果是基本类似的。至于其他一些群内基尼系数权数不为 $\nu_k\theta_k$ 的分解公式，它们的经济意义和解释能力也大大差于式（16），而且群内基尼系数的权数不是非常合理。

六、中国总体基尼系数的城乡分解

在对收入不平等进行实证分析时，基尼系数的收入来源分解与泰尔指数的子群分解是常用的两种分析方法。由于一般认为基尼系数不满足子群可分解性，因此分析中国城乡不平等对总体不平等的贡献时，绝大部分文献采用了泰尔指数的分解结果，还有一部分文献利用忽略剩余项 R 的式（4）进行分析。与现有实证分析不同的是，本文将利用式（16）对中国总体基尼系数按城乡进行分解，分析城乡各自的群内不平等以及群间不平等对总体基尼系数的贡献。

对城乡两个子群而言，式（16）可以化为：

$$G = \nu_U\theta_U G_{UU} + \nu_R\theta_R G_{RR} + \nu_R\theta_U G_{UR} + \nu_U\theta_R G_{RU}, \tag{17}$$

其中下标 U 代表城镇，下标 R 表示农村。有关年份的收入不平等如表 1 所示。不难发现，城镇内部基尼系数、农村内部基尼系数、农村群间不平等以及总体基尼系数均呈扩大趋势，2005 年与 1990 年相比，四个指标分别增加了 82.21%、22.53%、135.19%和 46.28%。城镇群间不平等却呈现缩小趋势，2005 年与 1990 年相比降低了 52.27%，但由于该指标数值较小，它对总体不平等的影响非常有限，这从表 2 中第 2 到第 5 列的贡献率可以看出。在总体不平等的贡献率中，一半以上来自农村群间不平等项，近期更是达到了 60%以上，这说明增加农村居民尤其是农村中下层居民的收入水平有着重要意义，它不仅有助于缩小农村居民与城镇居民的收入不平等，而且非常有利于降低总体收入不平等。此外，农村内部不平等虽呈现扩大趋势，但由于农村人口比重的大幅度下降[5]，农村内部不平等项 $\nu_R\theta_R G_{RR}$ 的贡献率不断缩小，2005 年的贡献率只有 12.62%。与之相反的是，由于城镇内部不平等的较大幅度增加，以及城镇人口比重的持续上升，城镇内部不平等项 $\nu_U\theta_U G_{UU}$ 的贡献率呈现增长趋势，2005 年达到了 21.69%。目前我国正处于加速发展城市化的进程中，城镇人口比重 ν_U 以及收入比重 θ_U 的不断增加将是必然趋势，因此如果要控制城镇群内不平等项 $\nu_U\theta_U G_{UU}$，唯一有效方法就是降低 G_{UU}，这就需要我们

[5] 农村人口比重由 1990 年的 73.59%逐渐下降到 2005 年的 57.01%，关于各年城乡人口比重数值可参见《2006 年中国统计年鉴》表 4-1。

加大对城镇贫困人员的社会救助，切实提高城镇中下阶层居民的收入水平。

表 2 右边三列分别为依据式（4）和式（5）计算的群间差距贡献率及剩余项贡献率，虽然从数量上看三个式子是相等的，即 $A+B=C+D=E$，但式（17）的经济意义和解释能力明显要强于其他两个式子。式（5）只能说明城乡差距总的贡献率，不能具体分析这个差距中农村和城镇各自的贡献率，发现不了农村相比城镇的巨大差距；式（4）也存在这个问题，而且由于剩余项 R 是一个有别于收入差距的概念，从经济意义上是很难用 R 解释总体收入差距的。

表 1　城乡内部基尼系数、群间不平等及总体基尼系数

年份	城镇内部基尼系数 G_{UU}	农村内部基尼系数 G_{RR}	城镇群间不平等 G_{UR}	农村群间不平等 G_{RU}	总体基尼系数 G
1990	0.1765	0.2938	0.0308	1.0353	0.3133
1995	0.2077	0.3373	0.0247	1.5246	0.3788
1998	0.2261	0.3308	0.0242	1.5400	0.3783
1999	0.2329	0.3360	0.0222	1.6464	0.3881
2000	0.2451	0.3530	0.0201	1.8329	0.4071
2001	0.2557	0.3598	0.0209	1.8873	0.4142
2002	0.3067	0.3644	0.0252	2.0018	0.4372
2003	0.3167	0.3717	0.0228	2.1509	0.4487
2004	0.3239	0.3583	0.0197	2.1956	0.4483
2005	0.3216	0.3610	0.0147	2.4423	0.4583

注：本表城镇基尼系数是根据城镇居民收入的七等分数据计算的，农村基尼系数根据农村居民收入 20 等分数据计算（其中 1990 年为 12 等分数据）。要说明的是，统计年鉴中收入高低是按家庭人均收入排序的，本文依据各个收入等级家庭的平均人口，将各个等级的家庭比重转化为人口比重，以各个等级的人口比重及其人均收入计算基尼系数，关于这个问题的详细讨论可参见洪兴建和李金昌(2005)。

数据来源：根据有关年份《中国统计年鉴》、《中国农村住户调查统计年鉴》整理计算。

表 2　各部分对总体基尼系数的贡献率

年份	城镇内部贡献率(%)	农村内部贡献率(%)	式(17) 城镇群间贡献率(%)A	式(17) 农村群间贡献率(%)B	式(4) 群间差距 G_B 贡献率(%)C	式(4) 剩余项 R 贡献率(%)D	式(5) 群间贡献率(%)E
1990	6.26	40.06	3.03	50.65	49.66	4.02	53.69
1995	8.00	31.47	2.32	58.21	55.89	4.65	60.53
1998	11.05	26.01	2.35	60.59	58.21	4.73	62.94
1999	12.11	23.73	2.16	61.99	59.80	4.35	64.16
2000	13.33	21.46	1.92	63.29	61.32	3.88	65.20
2001	14.65	19.99	1.98	63.38	61.35	4.01	65.35
2002	17.89	17.64	2.29	62.18	59.87	4.60	64.47
2003	19.37	15.89	2.05	62.69	60.62	4.12	64.74
2004	20.88	14.35	1.76	63.01	61.22	3.55	64.77
2005	21.69	12.61	1.31	64.39	63.08	2.62	65.70

注：本表贡献率数据是由式（18）、式（4）以及式（5）右边的每一项除以 G 计算得来，其中三个等式计算的城乡群内贡献率是相同的。

数据来源：根据表 1 以及相关年份《中国统计年鉴》整理计算。

本文测算的结果与程永宏（2007）的结果有一定差异，主要原因是两者的分解公式不一样。程永宏（2007）的群内基尼系数权重为人口份额，而本文权重为人口份额与收入份额之积，由于他的权重大于本文权重，因而其计算的群内不平等贡献率大于本文结果，相应地他测算的群间不平等贡献率远远小于本文结论。当然本文直接根据样本的分组数据计算，而程永宏利用了特定函数拟合收入分布，这也是造成差异的一个次要原因。

下面我们再从基尼系数变化的角度分析有关因素的影响。由于式（17）可以写成：

$$
\begin{aligned}
&G(\nu_U,\nu_R,\mu_U,\mu_R,G_{UW},G_{RW},G_{UB},G_{RB})\\
&\quad=\nu_U^2\frac{\mu_U}{\nu_U\mu_U+\nu_R\mu_R}G_{UW}+\nu_R^2\frac{\mu_R}{\nu_U\mu_U+\nu_R\mu_R}G_{RW}\\
&\qquad+(\nu_U-\nu_U^2)\frac{\mu_U}{\nu_U\mu_U+\nu_R\mu_R}G_{UB}+(\nu_R-\nu_R^2)\frac{\mu_R}{\nu_U\mu_U+\nu_R\mu_R}G_{RB},
\end{aligned}
\tag{18}
$$

其中 μ_U、μ_R 分别表示城镇居民和农村居民的人均收入。根据式（18）可知，影响全体居民收入基尼系数变动的因素主要有城乡人口比重的变动、城乡人均收入的变动、城乡群内不平等以及群间不平等的变动等，我们可以分析在总体基尼系数的变动中各种因素的影响方向及程度。其中城乡人口比重变动而引起的基尼系数变化可以表示为[6]：

$$
\begin{aligned}
\Delta G_\nu=&G(\nu_{U1},\nu_{R1},\mu_{U0},\mu_{R0},G_{UW0},G_{RW0},G_{UB0},G_{RB0})\\
&-G(\nu_{U0},\nu_{R0},\mu_{U0},\mu_{R0},G_{UW0},G_{RW0},G_{UB0},G_{RB0}),
\end{aligned}
\tag{19}
$$

上式中下标 1 和 0 分别表示报告期和基期，比如 ν_{U1} 表示报告期城镇人口份额。类似地，城乡人均收入变动、城乡内部不平等以及群间不平等变化造成的总体基尼系数变动可以分别如下表示：

$$
\begin{aligned}
\Delta G_\mu=&G(\nu_{U0},\nu_{R0},\mu_{U1},\mu_{R1},G_{UW0},G_{RW0},G_{UB0},G_{RB0})\\
&-G(\nu_{U0},\nu_{R0},\mu_{U0},\mu_{R0},G_{UW0},G_{RW0},G_{UB0},G_{RB0}),
\end{aligned}
\tag{20}
$$

$$
\begin{aligned}
\Delta G_W=&G(\nu_{U0},\nu_{R0},\mu_{U0},\mu_{R0},G_{UW1},G_{RW1},G_{UB0},G_{RB0})\\
&-G(\nu_{U0},\nu_{R0},\mu_{U0},\mu_{R0},G_{UW0},G_{RW0},G_{UB0},G_{RB0}),
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
\Delta G_B=&G(\nu_{U0},\nu_{R0},\mu_{U0},\mu_{R0},G_{UW0},G_{RW0},G_{UB1},G_{RB1})\\
&-G(\nu_{U0},\nu_{R0},\mu_{U0},\mu_{R0},G_{UW0},G_{RW0},G_{UB0},G_{RB0}).
\end{aligned}
\tag{22}
$$

---

[6] 万广华等(2006)注意到选择不同的“参照点”会产生不同的分解结果，因而他们采用了 Shorrocks (1999)提出的 Shapley 值分解法，并认为建立在合作博弈论基础上的这一方法是科学的。虽然这一方法得到了很多学者的青睐，但笔者以为 Shapley 值分解法自身并不完善，其理论基础存有争议。Shapley 值的核心是平均分配合作产生的收益，而不考虑合作各方所面临的风险等因素，这种平均分配方式显然有失公平，因此基于风险因子等一些分配思想被学者们提了出来(戴建华等，2004)。本文选择基期为参照点，其用意就是单纯反映每个因素的影响程度，这也是进行因素分析的最基本方法。当然每一种分解方法可能都有它的不足之处，需要从事物发展变化的机理与经济意义的角度选择最合适的方法。

城乡人口比重反映了总体的人口结构，它的变化对总体基尼系数的影响可以称为结构效应；城乡人均收入变化反映了城乡收入水平的增长差异，由于它导致的总体基尼系数变动可以称为水平效应。在其他因素不变的条件下，城乡群内不平等与群间不平等这些局部不平等的变化也会使得总体基尼系数发生变动，这种变动称为局部不平等效应，具体又可分为群内不平等效应和群间不平等效应。因此式（19）至式（23）这四个公式分别测算了总体基尼系数变动的结构效应、水平效应、群内不平等效应和群间不平等效应。当然两个时期的总体基尼系数变动除了受上述因素独自变动影响外，还与这些因素之间的交互变动（interact）有关。为简洁起见，本文对所有其他交互变化导致的总体基尼系数变动不加细分，统称交互效应，记为 ΔG_I。因此基尼系数的总变动 ΔG 可以分解为：

$$\Delta G = \Delta G_{\nu} + \Delta G_{\mu} + \Delta G_W + \Delta G_B + \Delta G_I. \tag{23}$$

要说明的是，ΔG_W 和 ΔG_B 还可以进一步分解为城镇和农村的群内及群间不平等效应。主要对比年份的总体基尼系数变动以及影响因素的作用程度如表 3 所示。可以发现，在四个对比年份中，农村的群间不平等效应最大，其数值均大于总体基尼系数的变动值，因此这是导致我国全体居民收入基尼系数上升的最重要原因。其次，城镇和农村的内部不平等效应也是引起总体基尼系数上升的原因之一，且城镇内部不平等效应逐渐增强。虽然从 1990—2005 年这一大的对比时段看，农村内部不平等效应要强于城镇内部不平等效应，但从不同对比年份看，城镇内部不平等效应的作用渐渐超过了农村内部不平等效应，而且 2005 年与 2000 年相比，农村内部不平等已经转变为总体基尼系数下降的原因之一。最后，结构效应、水平效应以及交互效应大多导致总体基尼系数的下降，表明这些因素的变动有利于总体基尼系数的缩小。应该说，我国城乡人口结构的变动趋势以及城乡人均收入的上升趋势会持续一段较长时间，因而无论是结构效应还是水平效应，它们降低总体基尼系数的作用还会在一段较长时间内起作用，但从它们降低总体基尼系数的幅度看，远没有增加总体基尼系数的局部不平等效应大。

表 3　影响总体基尼系数变动的各种效应分析

对比年份	结构效应（%）	局部不平等效应（%）				水平效应（%）	交互效应（%）	总变动（%）
		群内不平等效应		群间不平等效应				
		城镇	农村	城镇	农村			
1990—1995	−0.12	0.35	1.86	−0.19	7.50	−2.09	−0.75	6.55
1995—2000	−0.59	0.55	0.55	−0.16	4.46	−1.81	−0.17	2.83
2000—2005	−0.83	1.7	−0.13	−0.21	8.56	−3.64	−0.33	5.12
1990—2005	−1.38	1.71	1.78	−0.53	16.64	−5.45	1.72	14.5

资料来源：根据表 1 与相关年份《中国统计年鉴》整理计算。

需要说明的是，本文测度的是城乡居民名义收入不平等，虽然城乡居民

可比收入的不平等要更加合理，但如何将城乡（地区）名义收入转化为可比收入是一个非常复杂的问题。由于本文重点探讨基尼系数的子群分解方法，该方法对城乡实际差距的分解也是适用的，因此关于可比收入的研究可以归结为另外一个问题，关于这个问题的论述可以参见 Johnson（2002）、李实（2003）、江小涓等（2005）、易刚等（2006）以及李实等（2008）。

七、结　　语

如何测度子群之间的收入不平等，是一个比较棘手的方法论问题，而关于不平等指标是否满足子群可分解性，很大程度上又与群间不平等的定义密切相关。本文认为，需要依据每个不平等指标的内涵定义相应的群间不平等，以子群算术平均数的某一函数（包括比值或差值）测度群间不平等只是一种方法，但并非任何情况下都是最合适的方法。由于基尼系数与相对剥夺概念的一致性，本文正是在此基础上定义了群间不平等，从而给出了基尼系数一个新的子群分解公式。从群内基尼系数对总体基尼系数的贡献以及群间不平等指标是否具有一致性的意义看，本文提出的分解公式比较合理，而且相比其他分解公式，本文分解公式在实证分析中具有较好的经济意义和较强的解释力。

我国总体基尼系数城乡分解的实证结果表明，农村群间不平等的贡献率最大，城镇群内不平等的贡献率也有较大程度的上升，这两个因素是导致我国全体居民收入不平等扩大的重要原因。从基尼系数的增量分析结果看，农村群间不平等的贡献最大，已经超过总体基尼系数的增量。因此，为了有效控制我国全体居民收入基尼系数，需要想方设法增加农村居民收入，同时着力提高城镇中下阶层居民的收入水平，切实做好贫困人员的社会救助。

参 考 文 献

[1] Atkinson, A., "On the measurement of Inequality", *Journal of Economic Theory*, 1970, 2(4), 244—263.

[2] Bhattacharya, N., and B. Mahalanobis, "Regional Disparities in Household Consumption in India", *Journal of the American Statistical Association*, 1967, 62(317), 143—161.

[3] Blackorby, C., D. Donaldson, and M. Auersperg, "A New Procedure for the Measurement of Inequality within and among Population Subgroups", *Canadian Journal of Economics*, 1981, 14(4), 665—685.

[4] 陈宗胜，"关于总体基尼系数估算方法的一个建议——对李实研究员《答复》的再评论"，《经济研究》，2002 年第 5 期，第 81—83 页。

[5] 陈宗胜、周云波，《再论改革与发展中的收入分配——中国发生两极分化了吗》。北京：经济科学出版社，2002 年，第 85 页。

[6] 程永宏，"二元经济中城乡混合基尼系数的计算与分解"，《经济研究》，2006 年第 1 期，第 109—120 页。

[7] 程永宏,“改革以来全国总体基尼系数的演变及其城乡分解”,《中国社会科学》,2007 年第 4 期,第 45—60 页。
[8] Cowell, F., “Measurement of inequality”, in Atkinson, A., and F. Bourguignon (eds.), *Handbook of Income Distribution*. Amsterdam: North-Holland, 2000.
[9] Dagum, C., “A New Approach to the Decomposition of the Gini Income Inequality Ratio”, *Empirical Economics*, 1997, 22(4), 515—531.
[10] Dagum, C., “Inequality Measures between Income Distributions with Applications”, *Econometrica*, 1980, 48(7), 1791—1803.
[11] 戴建华、薛恒新,“基于 Shapley 值法的动态联盟伙伴企业利益分配策略”,《中国管理科学》,2004 年第 4 期,第 33—36 页。
[12] Das, T., and A. Parikh, “Decomposition of Inequality Measures and a Comparative Analysis”, *Empirical Economics*, 1982, 7(1—2), 23—48.
[13] 董静、李子奈,“修正城乡加权法及其应用——由农村和城镇基尼系数推算全国基尼系数”,《数量经济技术经济研究》,2004 年第 5 期,第 120—123 页。
[14] Ebert, U., “Measures of distance between income distributions”, *Journal of Economic Theory*. 1984, 32(2), 266—274.
[15] Fossett, M., and S. South, “The Measurement of Intergroup Income Inequality: A Conceptual Review”, *Social Forces*, 1983, 61(3), 855—871.
[16] Gastwirth, J., “Comment on ‘Measurement of Economic Distance between Blacks and Whites’ by H. D. Vinod”, *Journal of Business & Economic Statistics*, 1985, 3(4), 405—407.
[17] 洪兴建、李金昌,“如何正确测算我国居民收入基尼系数”,《南开经济研究》,2005 年第 4 期,第 53—57 页。
[18] 胡祖光,“基尼系数的理论最佳值及其简易计算公式研究”,《经济研究》,2004 年第 9 期,第 112—123 页。
[19] 李实,“中国个人收入分配研究回顾与展望”,《经济学(季刊)》,2003 年第 2 卷第 2 期,第 379—403 页。
[20] 李实,“对基尼系数估算与分解的进一步说明——对陈宗胜教授评论的再答复”,《经济研究》,2002 年第 5 期,第 84—87 页。
[21] 李实、罗楚亮,“中国城乡居民收入差距的重新估计”,载李实等主编《中国居民收入分配研究Ⅲ》。北京:北京师范大学出版社,2008 年。
[22] 金成武,“离散分布收入数据基尼系数的矩阵向量形式及其相关问题”,《经济研究》,2007 年第 4 期,第 149—159 页。
[23] 江小涓、李辉,“我国地区之间实际收入差距小于名义收入差距——加入地区间价格差异后的一项研究”,《经济研究》,2005 年第 9 期,第 11—18 页。
[24] Johnson, D.,“1978 年以来,中国的城乡差距拉大了吗?”《经济学(季刊)》,2002 年第 1 卷第 3 期,第 553—562 页。
[25] Lambert, P., and J. Aronson, “Inequality Decomposition Analysis and the Gini Coefficient Revisited”, *Economic Journal*, 1993, 103(420), 1221—1227.
[26] Mangahas, M., “Income Inequality in the Philippines: A Decomposition Analysis”, World Employment Programme Working Paper, No. 12, Geneva, ILO, 1975.
[27] Mookherjee, D. and A. Shorrocks, “A Decomposition Analysis of the Trend in UK Income Inequality”, *Economic Journal*, 1982, 92(368), 886—902.
[28] Pyatt, G., “On the Interpretation and Disaggregation of Gini Coefficients”, *Economic Journal*, 1976, 86(342), 243—255.
[29] Rao, V., “Two Decompositions of Concentration Ratio”, *Journal of the Royal Statistical Society*, Series A, 1969, 132, 418—425.
[30] Runciman, W., *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth-Century England*. Berkeley and Los Angeles: University of California Press, 1966.

[31] Sen, A.,"评估不平等和贫困的概念性挑战",《经济学(季刊)》,2003 年第 2 卷第 2 期,第 257—270 页。
[32] Shorrocks, A., and G. Wan, "Spatial Decomposition of Inequality", *Journal of Comparative Economic Geography*, 2005, 5(1), 59—81.
[33] Shorrocks, A., "Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value", Mimeo, University of Essex, 1999.
[34] Shorrocks, A., "Inequality Decomposition by Population Subgroup", *Econometrica*, 1984, 52(6), 1369—1386.
[35] Silber, J., "Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality", *Review of Economics and Statistics*, 1989, 71(1), 107—115.
[36] Vinod, H., "Measurement of Economic Distance between Blacks and Whites", *Journal of Business & Economic Statistics*, 1985, 3(1), 78—88.
[37] 万广华、张茵,"收入增长与不平等对我国贫困的影响",《经济研究》,2006 年第 6 期,第 112—123 页。
[38] 徐宽,"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)》,2003 年第 2 卷第 4 期,第 757—778 页。
[39] Yao, S., "On the Decomposition of Gini Coefficients by Population Class and Income Source: a Spreadsheet Approach and Application", *Applied Economics*, 1999, 31(10), 1249—1264.
[40] 易刚、张燕姣,"以购买力平价测算基尼系数的尝试",《经济学(季刊)》,2006 年第 6 卷第 1 期,第 91—104 页。
[41] Yitzhaki, S., "Relative Deprivation and the Gini Coefficient", *Quarterly Journal of Economics*, 1979, 93(2), 321—324.
[42] Yitzhaki, S. and R. Lerman, "Income Stratification and Income Inequality", *Review of Income and Wealth*, 1991, 37(3), 313—329.
[43] Yitzhaki, S., "Economic Distance and Overlapping of Distributions", *Journal of Econometrics*, 1994, 61(1), 147—159.

A New Subgroup Decomposition of the Gini Coefficient

XINGJIAN HONG
(*Zhejiang Gongshang University*)

Abstract There are many approaches to subgroup decomposition of the Gini coefficient, and empirical studies based on different approaches lead to different conclusions. This paper studies the weights of subgroups and the Gini type inter-group inequality, and presents a new subgroup decomposition formula. We also study the overall Gini coefficient in China. The empirical results show that the major contributor to China's overall Gini coefficient is inter-group inequality between rural and urban residents, and the contribution of within-urban inequality is rising.

JEL Classification C60, D31, D63