

居民财产分配差异测度研究 ——基于基尼系数群组 and 来源的 二维分解方法



刘学良*

摘要：本文提出一个新的基尼系数二维分解方法，该方法可以同时进行基尼系数的按群组和按来源分解操作，从而避免了不平等的群组分解和来源分解因使用方法不同而难以融合的缺陷。基于此方法，本文使用 CFPS 数据测算了 2012—2020 年我国城乡居民财产分配差距，发现这一时期居民财产分配差距呈波动上升态势。其中，城乡内部特别是城镇内部居民持有住房资产的差异是这一时期财产分配不平等扩大的重要原因，但如果考虑扣除房贷后的净资产，则其对差距扩大贡献明显缩小；而城乡间差异不是导致居民总体财产差距扩大的原因，其中，土地资产在这一时期起到扩大城乡组间差距的效果，而住房资产及住房净资产起到缩小组间差距的作用。因此，十八大以来，农村居民住房条件和住房财产明显提高对于缓解城乡差异起到积极作用，但不可忽视的是，城乡间住房资产和住房条件的绝对差距仍然很大，缩小城乡间居民住房条件差距仍然任重道远。

关键词：基尼系数；群组分解；来源分解；二维分解；财产分配差距

一、引言

在收入分配的相关研究中，测算不平等的高低和变化趋势，并寻找影响不平等的原因是基础而又关键的工作。测量不平等的高低需要借助一些指标，常用的符合匿名性、人口无关性、Dalton-Pigou 转移性原则、规模不变性等公理化原则的指标，包括基尼系数、泰尔指数、变异系数、阿特金森指数等。其中，除了满足上面这些公理化原则，基尼系数还有一些特性，使其格外易于使用，包括：①基尼系数允许样本取值(收入)为零乃至为负；②基于洛伦兹曲线定义的基尼系数具有清晰直观的几何解释；③基尼系数有明确的社会福利含义；④基尼系数的取值范围为 0~1，比较起来十分方便，并且其最小值和最大值有明确的经济学含义；⑤基尼系数能很好地实现按收入来源的分解，从而分析不平等的来源特征。这些特点使得基尼系数成为最为常用的不平等测度指标。

* 刘学良，中国社会科学院经济研究所(邮编：100836)、中国社会科学院大学经济学院(邮编：102488)，E-mail: lxl@cass.org.cn。本研究得到中国社会科学院创新工程项目“扎实推动共同富裕的评价监测体系和政策体系研究”，国家高端智库基础研究项目“完善收入分配体制推进共同富裕研究”的资助。感谢匿名审稿人的评论和建议。

基尼系数的不足之处是其群组分解方法存在缺陷。已有研究指出,基尼系数不满足加和可分解性(Bourguignon, 1979; Shorrocks, 1980, 1984; Cowell, 1980),常见于文献的基尼系数群组分解方法会留有一个因群组之间分布重叠而形成的交叉项,因此,过去的文献认为基尼系数不是完全可分解的。因为基尼系数在群组分解上存在不足,所以当研究者要对收入差距进行按来源分解时,多使用基尼系数进行,但要对差距进行群组分解时,则经常换用泰尔指数等工具。这制造了许多麻烦,如由于群组分解和来源分解使用不同的工具,群组分解和来源分解的结果难以比较对照,当我们想知道群组分解的组间差距和组内差距中各项收入来源的贡献比例时,这通常无法实现;反过来,当我们想知道来源分解下某项收入来源带来的差距中有多少是群组间差距造成的,而多少是组内差距造成的,也是无法实现的,也就是说,未能同时实现群组和来源的二维分解。过去虽有学者提出过基尼系数的二维分解方法(Mussard, 2004; Mussard 和 Savard, 2012),但由于其方法的基础仍然是传统带有交叉项的基尼系数群组分解式,因此其将传统基尼系数群组分解的缺陷也一并继承了。

对基尼系数分解方法的研究仍在继续,其中,程永宏(2008)和 Okamoto(2009)放松了对加和可分解性的要求,提出了一个新的基尼系数子组分解式,其组间差距是子组收入分布的函数(分布距离),而不是子组收入均值的函数,在一定程度上克服了过去基尼系数群组分解的缺陷,是对传统基尼系数群组分解法的改进。在他们研究的基础上,本文提出了新群组分解方法下的按收入来源分解式,从而在一个框架下实现了基尼系数按群组和按来源的二维分解。此分解方法要优于 Mussard(2004)、Mussard 和 Savard(2012)基于传统基尼系数群组分解实现的二维分解方法。

本文结构安排如下:第二节回顾基尼系数按群组和按来源分解方法,指出已有方法的问题;第三节提出新的基尼系数按群组和来源的二维分解方法,并讨论新分解方法的一些统计性质;第四部分是新方法的具体应用,我们使用 2012—2020 年 CFPS 数据对中国城乡居民的财产分配差距进行分析;第 5 部分总结全文。

二、基尼系数按群组分解和按来源分解方法的回顾

(一) 基尼系数的不同定义

基尼系数存在多种定义方式,包括利用洛伦兹曲线的几何方法、基尼平均差方法、分布函数方法、协方差方法等,本质上都是等价的。使用基尼平均差的定义公式如下:

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n |y_i - y_j|}{2n^2 \mu} \quad (1)$$

其中, y_j 和 y_i 是来自同一总体的不同样本, n 是样本量, μ 是收入均值。令排序后的样本向量为 $\mathbf{y} = [y_1, y_2, \dots, y_n]'$, 其中 $y_1 \leq y_2 \leq \dots \leq y_n$, 基尼系数又可以写成收入 y_i 与其序数的协方差的形式

$$G = \frac{1}{n^2 \mu} \sum_{i=1}^n (2i - n - 1) y_i \quad (2)$$

协方差形式的基尼系数还有一些等价变换, 比如式(2)可变为:

$$\begin{aligned} G &= \frac{1}{n^2 \mu} \sum_{i=1}^n (2i-1)y_i - 1 \\ &= \frac{n-1}{n} - \frac{2}{n^2 \mu} \sum_{i=1}^n (n-i)y_i \\ &= \frac{2}{n^2 \mu} \sum_{i=1}^n i y_i - \frac{n+1}{n} \end{aligned} \quad (3)$$

以上几个式子都是等价的, 见于不同文献中(徐宽, 2003; Cowell 和 Flachaire, 2015)。构造序数向量 $\mathbf{A}^{(n)} = [1, 3, 5, \dots, 2n-1]_{1 \times n}$, 则以上结果可改写为以矩阵向量形式表达:

$$G = \frac{1}{n^2 \mu} \mathbf{A}^{(n)} \mathbf{y} - 1 \quad (4)$$

基于分布函数的基尼系数表达式如下:

$$\begin{aligned} G &= 1 - 2 \int_0^T L(x) dF(x) \\ &= \frac{1}{\mu} \int_0^T F(x) [1 - F(x)] dx \end{aligned} \quad (5)$$

其中, T 是收入分布 $F(x)$ 的最大值, 同时假设收入分布的最小值为 0, $L(x)$ 是洛伦兹曲线。累积分布函数其他形式的等价定义此处不再赘述。

(二) 基尼系数按来源分解方法

假设收入可分成 K 种不同的来源, 有 $y_i = y_i^{(1)} + y_i^{(2)} + \dots + y_i^{(K)}$ 。同时, 令所有样本的第 k 种收入来源向量为 $\mathbf{y}^{(k)} = [y_1^{(k)}, y_2^{(k)}, \dots, y_n^{(k)}]'$, 则有 $\mathbf{y} = \sum_{k=1}^K \mathbf{y}^{(k)}$ 。那么, 式(2)就可改写为:

$$\begin{aligned} G &= \frac{1}{n^2 \mu} \sum_{i=1}^n (2i-n-1)y_i \\ &= \sum_{k=1}^K \frac{\mu^{(k)}}{\mu} \sum_{i=1}^n \frac{1}{n^2 \mu^{(k)}} (2i-n-1)y_i^{(k)} \\ &= \sum_{k=1}^K c^{(k)} C(\mathbf{y}^{(k)}) \end{aligned} \quad (6)$$

其中, $C(\mathbf{y}^{(k)})$ 是收入成分 k 的集中系数, 或称伪基尼系数 (Pseudo-Gini Coefficient), 其计算与总体基尼系数一样, 只不过人口和收入分项变量仍按照收入总量从小到大排序, 而不是按分项收入从小到大排序。权重 $c^{(k)}$ 是第 k 项收入来源占总收入的平均比重 (收入份额)。这样, 总体基尼系数就等于分项收入的集中系数的加权平均, 从而把总体基尼系数分解为各项收入来源的贡献 (Fei 等, 1978; Kakwani, 1977)。

(三) 基尼系数按群组分解方法

然而, 基尼系数的按群组分解一直存在争议。当一个总体 F 被分为 J 个子组时, 子组中的个体也构成 J 个分布, 令 $\Pi = \{N_1, N_2, \dots, N_J\}$ 。差距不平等按群组分解, 就是希望找到一个方式, 将收入差距 $I(F)$ 分解成组内差距和组间差距两部分:

$$I(F) = \sum_{j=1}^J w_j I(F_j) + I(F_{\Pi}) \quad (7)$$

其中,式(7)右边的第一项为组内差距,它等于各组不平等指标的加权平均,权重 $w_j \geq 0$;第二项为组间差距,由 $I(F_{\Pi})$ 决定。Cowell(2000)指出,权重 w_j 应是 j 组的人口权重 p_j 和收入权重 s_j 的函数^①且应大于等于 0,即 $w_j = w(p_j, s_j) \geq 0$ 。此外,一个更加严格的约束条件是:

$$\sum_{j=1}^J w_j = 1 \quad (8)$$

这一约束条件指出,组内差距权重的和应恰好等于 1,如果组间差距等于 0,那么总体不平等应恰好等于各组不平等的加权和。权重之和为 1 的约束条件在理论上是合理的,可以通过一个特别的例子说明:我们要求不平等测度指标应满足人口无关性公理,因此若把一个收入样本复制 n 份,这 n 份样本是完全相同的,因此组间差距应为 0,这时权重之和必须为 1,否则就会违反人口无关性原则。当施加以上的约束条件时,广义熵指数是满足要求的,其中广义熵指数中的平均对数离差(MLD)指数的权重为人口占比 p_j 、泰尔指数其权重为收入占比 s_j (Bourguignon, 1979; Shorrocks, 1980, 1984; Cowell, 2000)。

但是,过去提出的基尼系数分解方法不满足以上要求。广泛见于文献中(Mookherjee 和 Shorrocks, 1982; Silber, 1989; Dagum, 1997)的基尼系数按群组分解式为:

$$\begin{aligned} G(F) &= \sum_{j=1}^J w_j G(F_j) + I_b + \varepsilon \\ &= \sum_{j=1}^J p_j s_j G(F_j) + \frac{1}{\mu} \sum_{i=1}^J \sum_{j>i}^J p_i p_j |\mu_i - \mu_j| + \sum_{i=1}^J \sum_{j>i}^J \varepsilon(F_i, F_j) \end{aligned} \quad (9)$$

式(9)中第一项是组内差距,各组基尼系数在组内差距中的权重为 w_j ,它等于 j 组的人口权重和收入权重之积;第二项是组间差距项,它取决于各组均值的差异和人口权重;第三项则是余项(或称交叉项),余项的大小取决于各组收入分布的重叠程度,当各组收入分布完全不重叠时,余项 ε 为 0,就得到了 Sundrum 算法(Sundrum, 1990)^②。上述分解方法存在一些明显缺陷。首先,它不满足加和可分解性(additive decomposability),在组内和组间差距项外还有余项存在,因此不是一个完全分解形式。有的学者,如 Mookherjee 和 Shorrocks(1982)认为余项难有精确解释,有的学者,如 Silber(1989)等则认为它有明确的显而易见的经济学意义,它反映了不同群组的收入类聚,或者说重叠程度^③。其次,上述分解方法的缺陷不仅存在于组间差距和余项中,在组内差距中同样存在:组内差距权重 w_j 由于是人口权重和收入权重之积,权重之和

① 为方便区分子组的收入权重和收入来源的收入权重,令子组 j 的收入权重为 s_j ,收入来源 k 的权重为 $c^{(k)}$,样本的个体和子组用下标表示,收入来源用上标表示。

② 然而,多数情况下,不同子组的收入经常是重叠的,这导致交叉项的影响经常很大。

③ 交叉项在一些文献中又被命名为超变密度(intensity of transvariation)而得到一定传播和应用(Dagum, 1997),国内学界广泛使用这一方法和超变密度的术语。

必然要小于 1, 因此如果各组之间差距为 0 时, 总的收入差距不平等却不等于各组不平等的加权和, 这显然也是不合理的。

由于基尼系数在群组分解上存在缺陷, 使得基尼系数的应用受到了一定限制。然而, 尽管在群组分解中存在不足, 但作为测度收入差距不平等的众多方法之一, 基尼系数却保持了长久的生命力, 至今仍是最流行的收入差距测度指标。Cowell (2000) 对基尼系数一直如此流行给出了两个可能解释: 一是因为基尼系数的应用历史更加悠久, 而研究存在惯性; 二是因为基尼系数与洛伦兹曲线存在一定关系, 使得人们似乎可以通过基尼系数立刻“看到”不平等的状态一样。以上解释虽是原因之一, 但并不是此问题的全部答案, 在本文看来, 关键问题是其他的替代指标也不是完美的。如果存在其他不平等指标能具有基尼系数的各种优点, 同时又克服了基尼系数的缺陷, 基尼系数自然应早已被束之高阁。然而, 这样的指标并不存在, 泰尔指数等工具也有其各自的不足, 甚至不足之处可能比基尼系数更多, 这恐怕才是基尼系数一直如此流行的关键原因。

针对基尼系数分解研究仍在继续, 值得注意的是, 程永宏 (2008) 和 Okamoto (2009) 放松了对加和可分解性的要求, 分别提出了一个新的基尼系数群组分解式:

$$G = \sum_{i=1}^N s_i G(F_i(x)) + \frac{1}{\mu} \sum_{i=1}^N \sum_{j>i}^N p_i p_j \int_0^T (F_i(x) - F_j(x))^2 dx \quad (10)$$

其中, s_i 和 p_i 分别为 i 子组的收入和人口占总体比重, 组间差距项 $\int_0^T (F_i(x) - F_j(x))^2 dx$ 为 i, j 两组间绝对分布距离 (distribution distance) 的测度。可以看到, 以上分解方法仍不满足狭义的加和可分解性原则, 组间差距不是各组间均值差异的函数, 而是整个收入分布的函数, 但它比旧有的带有余项的分解有一些改进: 第一, 组内差距的权重为各组收入占比 s_i , 权重之和等于 1, 因此当各组间差距为 0 时, 总的基尼系数就等于各组加权平均, 这符合 Cowell (2000) 指出的关于权重的约束条件; 第二, 相比传统方法, 分布距离作为组间差距项也有一定合理性^①。

三、基尼系数按群组和按来源的二维分解的一致框架

客观地说, 程永宏 (2008) 和 Okamoto (2009) 的基尼系数新群组分解方法提出后并没有得到广泛应用。直到最新的研究, 在需要从收入来源视角分析不平等时, 人们使用基尼系数来源分解法, 而当需要进行群组视角分析时, 又切换到泰尔指数等广义熵指数族, 这才引发了关注。出现这一结果的原因在于, 如果只是进行差距的群组分解, 那么不管是新的还是旧的基尼系数群组分解方法, 本质上都没有比泰尔指数等方法更容易理解和使用。因此, 如果研究目的只是进行群组分解, 泰尔指数等工具一般已经够用, 除非遇上样本中取值可能小于等于 0 的情况, 否则没必要使用更为复杂的基尼系

① 加和可分解性原则关于组间差距仅为各组间均值差异的函数的要求可能是不必要的甚至是不合理的 (Blackorby 等, 1981; 程永宏, 2008), Okamoto (2009) 提出了组间差距测度的完全同分布条件 (Completely Identical Distribution Condition), 即当且仅当两个子组的收入分布完全相同时, 两个子组间的差距才为 0。

数分解方法了。

在本文看来,基尼系数相对泰尔指数等广义熵指数的真正优势,应是按来源分解与按群组分解充分结合起来。众所周知,泰尔指数除了要求样本取值为正外^①,其函数形式还使其无法进行按来源分解操作,这决定了泰尔指数无法实现按群组和按来源的二维分解。而基尼系数是可以按来源分解的,如果能把基尼系数的群组分解和来源分解结合起来,使之成为一个可以进行二维分解的方法,就可以在一个模型框架内拓展收入差距分析的维度,这对于收入差距的结构和成因的分析大有裨益。实际上,过去已有文献尝试这一工作,Mussard 基于传统基尼系数群组分解式提出了基尼系数的按群组和收入来源的二维分解方法(Mussard, 2004; Mussard 和 Savard, 2012), Mussini (2013) 还给出了相应方法的矩阵形式表达式。但是,这些尝试仍然是基于前文所述的传统的基尼系数群组分解方法,因此把传统方法的缺陷也一并继承了。

因此,为克服已有方法的不足,本文将在程永宏(2008)和 Okamoto(2009)研究的基础上,探索将基尼系数的按群组分解和按来源分解方法融合,以实现其二维分解。

(一) 新群组分解方法的按来源分解

新的群组分解方法中,组内差距为各子组基尼系数的加权平均,而基尼系数本身是可以实现按来源分解的,因此组内差距的按收入来源分解无需额外证明。需要探讨的是组间差距的按来源分解问题。先考虑两子组的情况,令已排序后的样本总体 $\mathbf{y} = [y_1, y_2, \dots, y_n]'$, 并把样本分为 \mathbf{y}_1 和 \mathbf{y}_2 两个子组,两组样本数分别为 n_1 和 n_2 , 两个子组合并的样本总量 $n = n_1 + n_2$, 两子组的人口占比为 p_1 和 p_2 。同时,构造序数向量 $\mathbf{A}^{(n)} = [1, 3, 5, \dots, 2n-1]_{1 \times n}$, 以及 $\mathbf{A}^{(n_1)}$ 和 $\mathbf{A}^{(n_2)}$, 分布距离就可写为(刘学良、田青, 2009):

$$\int_0^T (F_1(t) - F_2(t))^2 dt = \frac{1}{n_1 n_2} \mathbf{A}^{(n)} \mathbf{y} - \frac{n}{n_1^2 n_2} \mathbf{A}^{(n_1)} \mathbf{y}_1 - \frac{n}{n_1 n_2^2} \mathbf{A}^{(n_2)} \mathbf{y}_2 \quad (11)$$

以上分布距离的离散矩阵形式也可以通过基尼系数的矩阵形式定义式(4)并快速得到证明。现假设收入可分为 K 种不同的来源,第 k 种来源收入的向量为 $\mathbf{y}^{(k)}$, 样本的排序仍以总收入从低到高排序,有:

$$\mathbf{y} = \sum_{k=1}^K \mathbf{y}^{(k)} \quad (12)$$

同样,有子组的分收入来源向量 $\mathbf{y}_1^{(k)}$ 和 $\mathbf{y}_2^{(k)}$ 。

那么,两子组间的分布距离就可改写为:

$$\begin{aligned} \int_0^T (F_1(t) - F_2(t))^2 dt &= \frac{1}{n_1 n_2} \mathbf{A}^{(n)} \left(\sum_{k=1}^K \mathbf{y}^{(k)} \right) - \frac{n}{n_1^2 n_2} \mathbf{A}^{(n_1)} \left(\sum_{k=1}^K \mathbf{y}_1^{(k)} \right) - \frac{n}{n_1 n_2^2} \mathbf{A}^{(n_2)} \left(\sum_{k=1}^K \mathbf{y}_2^{(k)} \right) \\ &= \sum_{k=1}^K \left[\frac{1}{n_1 n_2} \mathbf{A}^{(n)} \mathbf{y}^{(k)} - \frac{n}{n_1^2 n_2} \mathbf{A}^{(n_1)} \mathbf{y}_1^{(k)} - \frac{n}{n_1 n_2^2} \mathbf{A}^{(n_2)} \mathbf{y}_2^{(k)} \right] \end{aligned} \quad (13)$$

① 样本值必须大于 0 在现实中经常是难以满足的。其中,收入为 0 是常见现象,为避免样本损失,过去常见的处理方式是对收入加一任意小值(比如 1),但对这种操作方法不乏批评。并且,财产等变量就更不便于这样操作了,因为净财产可能为负且经常数值较大,加上较大数值才能把所有样本变为正值,这会严重改变样本分布,从而使结果变得不可靠。

这样,分布距离就可以被分解成 K 种收入来源的子成分。回顾前文基尼系数按来源分解方法可发现,组间差距项的按来源分解与总体基尼系数按来源分解的方法是一致的,分项收入的组间差距项的计算与总收入的组间差距一致,只不过分项收入变量仍按照收入总量从小到大排序,而不是按分项收入从小到大排序。因此,组间差距项按来源分解的经济含义与总体基尼系数按来源分解的含义是一致的。

以上结果可推广至多子组。类似基尼系数按来源分解中定义的伪基尼系数(集中系数),定义 i, j 两子组分布距离中第 k 种来源收入的伪分布距离 $D_{i,j}^{(k)}$ 为:

$$D_{i,j}^{(k)} = \frac{1}{n_i n_j} \mathbf{A}^{(n_i+n_j)} \mathbf{y}_{i,j}^{(k)} - \frac{(n_i+n_j)}{n_i^2 n_j} \mathbf{A}^{(n_i)} \mathbf{y}_i^{(k)} - \frac{(n_i+n_j)}{n_i n_j^2} \mathbf{A}^{(n_j)} \mathbf{y}_j^{(k)} \quad (14)$$

其中 $\mathbf{y}_{i,j}$ 为 i, j 两子组样本合并后重新排序所形成的样本。

组间差距项就可简写为:

$$\int_0^T (F_i(t) - F_j(t))^2 dt = \sum_{k=1}^K D_{i,j}^{(k)} \quad (15)$$

总体基尼系数就可写为:

$$\begin{aligned} G(F) &= \sum_{j=1}^J s_j G(F_j) + \frac{1}{\mu} \sum_{i=1}^J \sum_{j>i}^J p_i p_j \int_0^T (F_i(t) - F_j(t))^2 dt \\ &= \sum_{j=1}^J s_j \sum_{k=1}^K c_j^{(k)} C(\mathbf{y}_j^{(k)}) + \frac{1}{\mu} \sum_{i=1}^J \sum_{j>i}^J p_i p_j \sum_{k=1}^K D_{i,j}^{(k)} \end{aligned} \quad (16)$$

两子组情况下的特例如下:

$$\begin{aligned} G(F) &= s_1 G(F_1) + s_2 G(F_2) + \frac{p_1 p_2}{\mu} \int_0^T (F_1 - F_2)^2 dx \\ &= s_1 \sum_{k=1}^K c_1^{(k)} C(\mathbf{y}_1^{(k)}) + s_2 \sum_{k=1}^K c_2^{(k)} C(\mathbf{y}_2^{(k)}) + \frac{p_1 p_2}{\mu} \sum_{k=1}^K D_{1,2}^{(k)} \end{aligned} \quad (17)$$

这样,我们就实现了新基尼系数群组分解方法下的按收入来源分解,组内差距项和组间差距项都可进一步分解为 K 种收入来源的贡献。

(二) 新基尼系数分解方法的统计性质

1. 新分解方法与加和可分解性和传统分解方法的关系

本节进一步讨论组间差距项(分布距离)的统计性质^①。如前所述,新分解方法的组间差距项不是子组收入均值的函数,因此不满足加和可分解性。然而,加和可分解性虽不完美,却仍有很大参考意义。虽然从福利经济学意义上看,测度差距应是测度福利或效用的差异,因此理论上应根据收入分配情况来计算所谓的等值收入(Atkinson, 1970; Blackorby 和 Donaldson, 1978)或有效收入(万广华, 2009),但是如何测度这一等值收入(有效收入)比较困难,且结果会依赖特定的福利或效用函数。因此,现实中通过比较群组均值来考察群组间差距大小仍是常用做法。

这里,我们进一步探讨新分解方法与加和可分解性的关系。一个自然的问题是,组

^① 程永宏(2008)论证了新分解方法的组间差距项满足对称性、人口无关性、零次齐次性等公理性原则,本文对这些公理性原则问题不再赘述。

间差距项中有多少是由子组收入平均水平的差异带来(即符合加和可分解性的那部分),又有多少是因为子组内部的不平等造成的?为回答这一问题,现假设子组内进行收入再分配,使所有个体都获得子组均值收入,从而消除子组内的不平等。令 i 子组的组内均值为 μ_i ,那么消除了组内差异的子组累积分布函数 $\Phi_i(x)$ 就变成:

$$\begin{cases} \Phi_i(x) = 1, & \text{if } x \geq \mu_i \\ \Phi_i(x) = 0, & \text{if } x < \mu_i \end{cases} \quad (18)$$

假设有 i 和 j 两子组,其中 $\mu_j > \mu_i$,消除了组内差异的组间差距项就可写成:

$$\int_0^T (\Phi_i(x) - \Phi_j(x))^2 dx = \int_{\mu_i}^{\mu_j} (1-0)^2 dx = \mu_j - \mu_i \quad (19)$$

因此,如果令子组内部所有人都获得子组均值收入,那么分布距离 $D_{i,j}$ 就恰好等于两子组收入均值之差的绝对值。相应的,这一反事实状态下的组间不平等与事实状态下的不平等的差异,就可以归结为由组内不平等造成的子组间的差异。因此,定义余项 $R_{i,j}$ 为^①:

$$\begin{aligned} R_{i,j} &= \int_0^T (F_i - F_j)^2 dx - \int_0^T (\Phi_i - \Phi_j)^2 dx \\ &= \int_0^T (F_i - F_j)^2 dx - |\mu_i - \mu_j| \end{aligned} \quad (20)$$

这样,分布距离 $D_{i,j}$ 就可以被分成两部分,其中一部分是两子组收入均值不同带来的贡献 $|\mu_j - \mu_i|$,另一部分是子组内部分配不平等所带来的贡献 $R_{i,j}$ 。因此,新分解方法也可以写成与传统分解方法类似的形式,两子组情况下基尼系数的分解可写为:

$$G = (s_1 G_1 + s_2 G_2) + p_1 p_2 \frac{|\mu_1 - \mu_2|}{\mu} + p_1 p_2 \frac{R_{1,2}}{\mu} \quad (21)$$

与传统分解方法相比较,可以发现第二项的定义是完全相同的, $p_1 p_2 |\mu_1 - \mu_2| / \mu$ 是子组的收入均值的函数。因此,新分解方法与传统方法的差别实际上体现在第一项组内差距项和第三项上,由于新分解方法的组内差距中权重为子组的收入占比 s_i ,而传统分解方法中权重为收入权重和人口权重之积 $s_i p_i$,新分解方法得到的组内差距必然大于传统分解方法,而新分解方法的余项则必然小于传统分解方法的余项。因此,在某些情况下,如果余项足够小,那么甚至可以认为新的分解方法是近似满足加和可分解性的^②。当然,这依赖于具体的分布和数据的情况。

2. 新分解方法中群组分解和来源分解的对称性

在不平等分解研究中,经常存在分解的对称性问题,即分解的先后次序是否会影响分解的结果。本文提出的基尼系数群组和来源二重分解也面临对称性问题的挑战,即先进行基尼系数的群组分解再进行来源分解,与先做来源分解再做群组分解的结果是否是一致的。多子组情况下,基尼系数的群组和来源二维分解可改写作先来源再群

① 这里称 $R_{i,j}$ 为余项而不是交叉项,是因为余项不是因为两个子组存在样本交叠,而是因为我们构造了一个假想的反事实状态。

② 不过,相比传统方法的余项被认为可以衡量组间分布的重叠程度,这里的余项目前只是作为一个假想反事实状态出现,尚缺乏明确的经济含义。

组的形式：

$$G(F) = \sum_{k=1}^K \left[\sum_{j=1}^J s_j c_j^{(k)} C(\mathbf{y}_j^{(k)}) + \frac{1}{\mu} \sum_{i=1}^J \sum_{j>i}^J p_i p_j D_{i,j}^{(k)} \right] \quad (22)$$

两子组的情况下，先收入来源再群组的基尼系数二维分解可写作：

$$G(F) = \sum_{k=1}^K \left[s_1 c_1^{(k)} C(\mathbf{y}_1^{(k)}) + s_2 c_2^{(k)} C(\mathbf{y}_2^{(k)}) + \frac{p_1 p_2}{\mu} D_{1,2}^{(k)} \right] \quad (23)$$

若要验证基尼系数的二维分解方法是否满足对称性，只需验证群组 and 来源二维分解中，任一项收入来源 k 的结果与总的基尼系数的按来源分解中收入 k 的结果是否一致即可。

两子组情况下的二维分解中收入来源 k 的贡献为：

$$\begin{aligned} G(F)^k &= \sum_{i=1}^2 s_i c_i^{(k)} C(\mathbf{y}_i^{(k)}) + \frac{p_1 p_2}{\mu} D_{1,2}^{(k)} \\ &= \sum_{i=1}^2 s_i \frac{\mu_i^{(k)}}{\mu_i} \left[\frac{\mathbf{A}^{(n_i)} \mathbf{y}_i^{(k)}}{n_i^2 \mu_i^{(k)}} - 1 \right] + \frac{p_1 p_2}{\mu} \left[\frac{\mathbf{A}^{(n_1+n_2)} \mathbf{y}^{(k)}}{n_1 n_2} - \frac{n \mathbf{A}^{(n_1)} \mathbf{y}_1^{(k)}}{n_1^2 n_2} - \frac{n \mathbf{A}^{(n_2)} \mathbf{y}_2^{(k)}}{n_1 n_2^2} \right] \\ &= \frac{\mathbf{A}^{(n_1)} \mathbf{y}_1^{(k)}}{n n_1 \mu} + \frac{\mathbf{A}^{(n_2)} \mathbf{y}_2^{(k)}}{n n_2 \mu} - \frac{\mu^{(k)}}{\mu} + \frac{1}{n^2 \mu} \mathbf{A}^{(n_1+n_2)} \mathbf{y}^{(k)} - \frac{\mathbf{A}^{(n_1)} \mathbf{y}_1^{(k)}}{n n_1 \mu} - \frac{\mathbf{A}^{(n_2)} \mathbf{y}_2^{(k)}}{n n_2 \mu} \\ &= \frac{1}{n^2 \mu} \mathbf{A}^{(n)} \mathbf{y}^{(k)} - \frac{\mu_k}{\mu} \\ &= c^{(k)} C(\mathbf{y}^{(k)}) \end{aligned} \quad (24)$$

式(24)表明，两子组的按群组 and 来源二维分解中收入成分 k 对不平等的贡献等于总体基尼系数的收入来源分解中收入成分 k 的结果。因此，基尼系数群组 and 来源的二维分解满足对称性。此结论可以推及多子组的情况。

以上为本文提出的基尼系数按群组 and 来源的二维分解方法，下面，我们应用新方法分析我国的城乡居民财产分配差距，以考察新方法在实践中的应用方法和效果。

四、我国 2012—2020 年城乡居民财产分配不平等的 基尼系数二维分解

(一) 我国城乡居民家庭人均净财产的基尼系数二维分解

1. 各年份城乡居民家庭人均净财产分配差距的群组分解

城乡差距是我国居民收入分配差距中的重要组成部分，相关研究文献不可胜数，但财产分配差距的研究还相对不足(李实和万海远, 2015)，因此本文选取财产分配差距为研究对象，以城乡和财产来源为分析视角，来考察我国居民财产分配差异的具体情况。本文所使用的数据为中国家庭追踪调查(CFPS)2012—2020年5期数据^①。参考

^① 本文的测算涵盖了2020年，但需提醒由于受到新型冠状病毒疫情影响，当年情况可能会有些特殊。

靳永爱和谢宇(2014)的研究,城乡家庭的资产按照来源性质被分为土地资产、房产、金融资产、生产性固定资产、耐用消费品资产五项,负债按照来源性质分为房贷和非住房负债两项。所有资产之和减去负债就得到家庭净资产。由于基尼系数允许样本取值为负,我们把负债取负后与资产一并纳入计算,这样最终计算的是家庭净财产的基尼系数。城乡分类变量使用基于国家统计局资料的城乡分类,并剔除了个别信息缺失的样本^①。此外,为了得到反映全国层面财富不平等更准确的估计,计算时还应用了CFPS的历年全国总样本家庭横截面抽样权数^②。家庭净财产的基尼系数及城乡分解的结果见表1。

表1 2012—2020年我国城乡居民家庭净财产的基尼系数及其群组分解

年份	数值/ 占比	基尼 系数	组内 差距	农村组 内差距	城镇组 内差距	农村 基尼系数	城镇 基尼系数	组间 差距
2012	数值	0.6652	0.6325	0.1507	0.4818	0.5877	0.6480	0.0327
	占比	100%	95.09%	23.83%	76.17%			4.91%
2014	数值	0.6709	0.6439	0.1530	0.4909	0.6081	0.6559	0.0270
	占比	100%	95.97%	23.76%	76.24%			4.03%
2016	数值	0.7006	0.6764	0.1357	0.5406	0.6174	0.6930	0.0243
	占比	100%	96.54%	20.07%	79.93%			3.46%
2018	数值	0.6896	0.6542	0.1167	0.5374	0.6876	0.6473	0.0354
	占比	100%	94.86%	17.84%	82.16%			5.14%
2020	数值	0.6974	0.6662	0.1103	0.5559	0.6382	0.6721	0.0312
	占比	100%	95.53%	16.56%	83.44%			4.47%

可以发现,基于CFPS数据计算得到的2012—2020年家庭净财产的基尼系数基本接近0.7的水平,居民财产分配差距要明显大于收入分配差距^③。总体来看,除了2016至2018年基尼系数略有下降外,其他年份家庭净财产的基尼系数都在逐步扩大。其中,2012年,家庭净财产基尼系数为0.6652,到2020年基尼系数升至0.6974。这一特定变化轨迹的原因是什么?城乡间居民财产的差距是上述这一变化的原因吗?从群组分解结果看,答案是否定的。城乡组内差距在2012—2016年逐步上升,在2018年有所下降,在2020年又有所上升,因此组内差距与总体差距的变动趋势是一致的。而城乡组间差距在2012—2016年逐步下降,组间差距值从0.0327降至0.0243,但在2018年上升至0.0354,到2020年又回落至0.0312,组间差距对总体差距的贡献占比的趋势也是如此。因此,2012—2020年城乡居民财产分配差距的总体变动趋势不是由城乡间

① 本文还剔除了ID=520413的异常值家庭样本,为2012—2016年出现的一个两口人农村家庭,从家庭收入、持有财产数量等变量来看,其不是富裕的家庭,但在2016年该家庭却突然出现了高达8000万的房贷负债,而其住房资产价值却只有8万元。因此,有理由怀疑此样本可能存在数据错误,鉴于此,本文剔除了该样本。

② 关于CFPS调查抽样设计和权数的使用参见谢宇等(2017)、吕萍(2016)的研究。

③ 由于计算方法、权重设置等差异,已有文献使用CFPS数据测算财产分配差距的结果各有不同,但基本在0.7左右。其中,Xie和Jin(2015)使用未经调整的CFPS数据计算得到的2012年家庭财产基尼系数为0.64,而用胡润富豪榜补足高净值人群财产信息后计算得到的基尼系数达到0.73,结果也见谢宇和靳永爱(2014)的研究;杨灿和王辉(2020)利用CFPS2016数据计算得到的财产基尼系数为0.70;罗楚亮和陈国强(2021)利用福布斯和胡润富豪榜数据补足高净值人群数据,并使用CFPS数据计算得到的2012年和2016年家庭人均财产净值的基尼系数分别是0.712和0.736。

差距导致的,城乡组间差距的变动与总体财产差距变动的趋势恰好相反。

2. 各年份城乡居民家庭财产分配组内和组间差距的来源分解

二维分解方法的优势是可以再对组内差距和组间差距进行来源分解分析。组内差距来源分解的结果见表 2,可以发现,对城乡组内差距贡献最大的财产来源是住房资产(2020 年占比 79.46%),其次是金融资产(2020 年占比 15.19%)、耐用消费品、生产性固定资产和土地资产。与之对应,在负债端,房贷的作用也要明显强于非房贷金融负债,2020 年房贷的组内差距贡献占比为-5.79%,而非房贷金融负债的贡献可忽略不计。由于负债是以负财产的形式被纳入计算的,房贷金额与人均净资产呈正向关系,但其最终呈现对人均净资产差异的负贡献。而非房贷金融负债的组内差距值结果为正,表明非房贷金融负债金额与人均净资产总体呈现负向关系,使得其呈现对人均净资产差异的正贡献。

表 2 2012—2020 年我国城乡居民财产分配组内差距(按来源分解)

年份	变量	人均 净资产	土地 资产	住房 资产	生产性 固定资产	耐用 消费品	金融 资产	房贷 负债	非房贷 金融负债
2012	数值	0.6325	0.0244	0.4708	0.0641	0.0244	0.0597	-0.0123	0.0014
	占比	100%	3.86%	74.42%	10.14%	3.85%	9.44%	-1.95%	0.23%
2014	数值	0.6439	0.0153	0.5221	0.0198	0.0246	0.0702	-0.0138	0.0057
	占比	100%	2.38%	81.08%	3.08%	3.82%	10.90%	-2.14%	0.89%
2016	数值	0.6764	0.0130	0.5293	0.0340	0.0343	0.0776	-0.0175	0.0055
	占比	100%	1.92%	78.26%	5.03%	5.08%	11.48%	-2.59%	0.82%
2018	数值	0.6542	0.0044	0.5622	0.0231	0.0260	0.0584	-0.0203	0.0004
	占比	100%	0.67%	85.94%	3.53%	3.97%	8.92%	-3.10%	0.06%
2020	数值	0.6662	0.0092	0.5294	0.0266	0.0359	0.1012	-0.0386	0.0025
	占比	100%	1.38%	79.46%	3.99%	5.39%	15.19%	-5.79%	0.38%

注:本表中人均净资产列的数值即为表 1 中组内差距值。

组间差距来源分解的结果见表 3,可以发现,对城乡组间差距贡献最高的同样是住房资产,2012 年其贡献达到 104.24%,但住房资产对组间差距的贡献比例在下降,2018 年住房资产对组间差距的贡献占比下降到 95.27%,2020 年进一步降至 95.05%。除住房外,金融资产、耐用消费品和生产性固定资产对城乡组间差距中贡献基本居于第 2~4 位,2020 年贡献占比分别达 21.08%、8.62%和 3.08%。与组内差距中各项资产对总体差距均起到正向贡献不同,城乡组间差距中土地资产的贡献为负,这是因为土地资产主要是农户所持有的农地价值,而农村居民持有的土地资产明显多于城镇居民,这使得土地资产对城乡组间差距的贡献为负^①。从负债端看,房贷负债和非房贷金融负债对组间差距的贡献均为负,即城镇居民的房贷和非房贷负债都要明显高于农村居民,从而起到减小组间差距的效果。

将表 2 组内差距和表 3 组间差距的结果相加,就可以得到总体基尼系数来源分解,结果见表 4。总体差距的来源分解是文献中的常见操作,这里不再具体分析。

^① 在农村内部,土地资产对家庭人均净资产基尼系数的贡献则毫无疑问仍然是正向的,不过其贡献占比在逐步降低,2012 年土地对农村不平等贡献比例达到 14.98%,到 2020 年这一比例已降至 5.84%。

表3 2012—2020年我国城乡居民财产分配组间差距(按来源分解)

年份	变量	人均 净资产	土地 资产	住房 资产	生产性 固定资产	耐用 消费品	金融 资产	房贷 负债	非房贷 金融负债
2012	数值	0.0327	-0.0076	0.0340	0.0005	0.0024	0.0048	-0.0015	-0.0001
	占比	100%	-23.19%	104.24%	1.58%	7.23%	14.83%	-4.48%	-0.21%
2014	数值	0.0270	-0.0053	0.0273	0.0007	0.0020	0.0049	-0.0022	-0.0004
	占比	100%	-19.47%	100.87%	2.77%	7.43%	18.14%	-8.18%	-1.55%
2016	数值	0.0243	-0.0036	0.0233	0.0006	0.0029	0.0049	-0.0032	-0.0005
	占比	100%	-15.03%	95.94%	2.44%	11.81%	20.00%	-13.18%	-1.99%
2018	数值	0.0354	-0.0024	0.0337	0.0008	0.0024	0.0044	-0.0030	-0.0005
	占比	100%	-6.80%	95.27%	2.18%	6.81%	12.34%	-8.48%	-1.31%
2020	数值	0.0312	-0.0028	0.0296	0.0010	0.0027	0.0066	-0.0054	-0.0005
	占比	100%	-9.08%	95.05%	3.08%	8.62%	21.08%	-17.31%	-1.45%

表4 2012—2020年我国城乡居民财产分配基尼系数的来源分解

年份	变量	人均 净资产	土地 资产	住房 资产	生产性 固定资产	耐用 消费品	金融 资产	房贷 负债	非房贷 金融负债
2012	数值	0.6652	0.0169	0.5048	0.0647	0.0267	0.0645	-0.0138	0.0014
	占比	100%	2.53%	75.89%	9.72%	4.02%	9.70%	-2.07%	0.20%
2014	数值	0.6709	0.0101	0.5493	0.0206	0.0266	0.0751	-0.0160	0.0053
	占比	100%	1.50%	81.88%	3.07%	3.97%	11.19%	-2.38%	0.79%
2016	数值	0.7006	0.0094	0.5526	0.0346	0.0372	0.0825	-0.0207	0.0051
	占比	100%	1.34%	78.87%	4.94%	5.31%	11.77%	-2.95%	0.72%
2018	数值	0.6896	0.0020	0.5960	0.0239	0.0284	0.0627	-0.0233	-0.0001
	占比	100%	0.29%	86.42%	3.46%	4.12%	9.10%	-3.38%	-0.01%
2020	数值	0.6974	0.0064	0.5590	0.0276	0.0386	0.1077	-0.0440	0.0021
	占比	100%	0.91%	80.15%	3.95%	5.54%	15.45%	-6.30%	0.30%

3. 2012—2020年基尼系数跨时期变动的群组来源分解

下面进行基尼系数跨时期变动的来源分解,以考察究竟是什么原因带来了总体基尼系数,以及组内差距和组间差距在2012—2020年的变化。先来看总体基尼系数的跨时期变动分解,如表5所示,2020年城乡居民家庭人均净财产的基尼系数比2012年提高了0.0322,其中城乡组内差距提升了0.0337,贡献占比达104.66%,而城乡组间差距降低了0.0015,贡献占比为-4.66%。因此,如前所述,2012—2020年我国居民总体财产分配差距的扩大是由城乡组内差距扩大而不是组间差距导致的。

表5 2012—2020年城乡居民财产分配差距变动的群组分解

	人均净资产	组内差距	组间差距
数值	0.0322	0.0337	-0.0015
占比	100%	104.66%	-4.66%

再来看总体不平等跨时期变动的来源分解,如表6所示,2012—2020年基尼系数扩大,最主要的因素是住房资产分配差距扩大,贡献占比达到168.18%;其次是金融资产分配差距扩大,占比达到134.10%;耐用消费品虽然也起到扩大差距的效果,但其贡献占比相对要小。而土地资产、生产性固定资产、房贷负债和非房贷金融负债起到的作

用都为负。特别是，扣除房贷后的住房净资产对总体差距变动的贡献为 0.0240，贡献占比达到 74.49%。因此，与已有研究观点一致，持有房产差距扩大是过去一段时期以来我国居民财产分配差距扩大的主要原因，但需注意的是，在住房资产扩张的同时不能忽视的是住房负债也在膨胀，若考虑扣除房贷后的住房净资产，则住房对不平等扩大的贡献会明显降低。

表 6 2012—2020 年城乡居民财产分配差距变动的来源分解

	人均 净资产	土地 资产	住房 资产	生产性 固定资产	耐用 消费品	金融 资产	房贷 负债	非房贷 金融负债
数值	0.0322	-0.0105	0.0542	-0.0371	0.0119	0.0432	-0.0302	0.0007
占比	100%	-32.55%	168.18%	-115.15%	36.89%	134.10%	-93.68%	2.22%

此外，虽然总体差距基本呈现扩大趋势，但在 2016—2018 年有过反复，从来源分解看，这与 2016—2018 年国内资本市场的波动有一定关系。中国股市在 2014 年处于熊市，但 2014 年下半年牛市启动，虽然经历了 2015 年的泡沫破裂，但到 2016 年年底，上证指数收盘仍达到 3104 点，远高于 2014 年的低谷。随后，资本市场在 2018 年时又陷入熊市。从金融资产项的差距值看，2016—2018 年分别是 0.0825 和 0.0627，而 2016—2018 年基尼系数下降 0.0110，金融资产对基尼系数下降的贡献占比高达 179.6%。

只对总体差距的变动进行来源分解得到的结果仍然不够细致清晰，这里继续对组内差距和组间差距的跨时期变动进行来源分解。组内差距的跨时期变动分解见表 7，2012—2020 年组内差距扩大了 0.0337，而住房资产的差距扩大程度达到 0.0586，因此这一时期组内差距的扩大主要也是住房资产的差异扩大造成的，贡献比例达 174%，若计算对全体居民财产分配基尼系数变动 (0.0322) 的贡献，则住房资产的贡献可达到 182%。金融资产对组内差距变动的贡献也较大，占比达到 123%，其中 2018—2020 年资本市场又经历了一轮牛市，这或许可以解释金融资产对组内差距变动的作用。耐用消费品虽起到扩大差距的作用，但量级不可与住房和金融资产相比。同时，土地资产、生产性固定资产对组内差距的贡献则出现绝对下降。从负债端看，房贷负债对组内差距的作用进一步加大，但由于其贡献为负，在变动的来源分解中结果也为负。最终，扣除房贷负债后的净房产对组内差距变动的贡献是 0.0324，贡献占比达到 96%。

表 7 2012—2020 年城乡居民财产分配组内差距变动的来源分解

	人均 净资产	土地资产	住房资产	生产性 固定资产	耐用 消费品	金融资产	房贷负债	非房贷 金融负债
数值	0.0337	-0.0152	0.0586	-0.0375	0.0116	0.0415	-0.0263	0.0011
占比	100%	-45%	174%	-111%	34%	123%	-78%	3%

组间差距的跨时期变动分解见表 8。城乡组间差距除了在 2018 年有明显上升，其他年份均呈缓慢下降趋势，2020 年的城乡组间差距比 2012 年下降了 0.0015。来源分解结果显示，在扩大城乡差距的因素中，起到最大作用的是土地资产 (占比达 -316%)，从表 3 也可以发现，土地资产在城乡组间差距中发挥的作用在逐步变小，其组间差距

贡献占比已从 2012 年的-23%降至 2020 年的-9%。在本文看来,这一现象主要有三方面原因:第一,相比其他资产如房产等,农民手中的土地资产缺乏升值空间,使得农民在土地资产上的优势随时间变化相对减弱;第二,由于土地征收、流转等原因,农民手中的土地资产进一步减少,使得土地资产缩小城乡间财产差距的作用变小,当然,农民会从土地征收造成的土地资产减少中相应得到征地补偿,使财富的形式发生转变,体现为其他资产(如金融资产、房产)的增加;第三,随着城市化进程推进,进城农民越来越多,这使得城镇居民手中拥有的土地资产缓慢增多。CFPS 数据显示,2012 年农村家庭人均土地财产为 11182 元,到 2020 年降为 10518 元,而城镇家庭 2012 年人均土地财产为 3310 元,到 2020 年增至 4091 元。显然,城镇家庭土地资产的增加在很大程度上是由越来越多农民迁入城镇所致。其他因素中,金融资产、生产性固定资产、耐用消费品也起到扩大差距的作用,但贡献相对土地资产要小。

表 8 2012—2020 年城乡财产分配组间差距变动的来源分解

	人均 净资产	土地资产	住房资产	生产性 固定资产	耐用 消费品	金融资产	房贷负债	非房贷 金融负债
数值	-0.0015	0.0047	-0.0044	0.0004	0.0003	0.0017	-0.0039	-0.0004
占比	100%	-316%	295%	-30%	-22%	-115%	262%	25%

注:2012—2020 年城乡财产分配组间差距虽有缩小但数值不高,因此在计算各成分贡献时,由于分母很小,容易使得占比的数值显得较高。

在缩小城乡组间差距的因素中,起到首要作用的是住房资产,对 2012—2020 年组间差距缩小的贡献占比达 295%。其次,房贷负债也起到缩小差距的作用,贡献占比为 262%。住房资产和房贷负债合并使得 2012—2020 年组间差距项缩小了 0.0084,贡献占比达 556%。最后,非房贷金融负债也起到缩小组间差距的作用,但贡献占比仅有 11%,远小于住房资产和房贷所起到的作用。因此,在 2012 年,我国总体上由于居民持有房产(净房产)的差距不断扩大,居民财富分配差距也不断扩大,但城乡间居民持有房产(净房产)的差距却缩小了,并且起到缩小组间差距和总体差距的效果。

城镇和农村居民家庭人均净资产和其成分来源的描述性统计也可以佐证上文的发现。如表 9 所示,2012 年城镇家庭人均持有的住房资产是农村家庭的 3.66 倍,到 2020 年这一比例降至 3.44 倍;2012 年城镇家庭人均持有的扣除房贷的住房净资产是农村家庭的 3.64 倍,到 2020 年降至 3.52 倍。

表 9 2012—2020 年城镇和农村居民家庭人均净资产(单位:元)和城乡比例

城乡	年份	人均 净资产	土地 资产	住房 资产	生产性 固定资产	耐用 消费品	金融 资产	房贷 负债	非房贷 负债
城镇	2012	146856	3310	117847	9688	7347	16268	-3579	-4025
城镇	2020	338832	4091	279313	12445	24066	57257	-31832	-6508
农村	2012	53309	11182	32165	5525	2752	5312	-821	-2805
农村	2020	107805	10518	81007	4811	9735	17071	-10820	-4517
城镇/农村	2012	2.75	0.30	3.66	1.75	2.67	3.06	4.36	1.44
城镇/农村	2020	3.14	0.39	3.45	2.59	2.47	3.35	2.94	1.44

注:本表所使用样本与前文基尼系数二维分解所使用样本一致,并应用家庭横截面权重作为权重。

如前所述,受疫情影响,2020年的数据或许存在异常之处,对其可靠性存在一些疑虑。因此,为排除疫情的干扰,我们再考察2012—2018年城乡组间差距的变动。计算2012—2018年城乡财产分配组间差距变动的来源分解结果,可以发现的是,这一时期城乡组间差距略有扩大,但住房资产(净资产)依然起到了降低组间差距的效果。其中,剔除房贷后的住房净资产差距项缩小0.0018,对人均净资产差距扩大的贡献占比为-66%。此外,与2012—2020年的结果一致,2012—2018年城乡差距扩大的主要贡献要素也是土地资产。因此,城镇居民住房资产快速增长的同时也带来了房贷负债的快速增长,而农村居民住房资产的增长对房贷的依赖程度低,这使得从住房资产特别是从扣除房贷后的住房净资产看,城乡间差距对总体差距扩大的贡献明显变小。

所以,2012—2020年,住房资产(净资产)对居民财产分配差距扩大的作用主要体现在城镇和乡村内部,而不是城乡之间。与这一时期城乡居民收入差距缩小的趋势一致,不管是城乡居民拥有住房资产(净资产)的绝对差距,还是住房资产对总体差距的贡献实际上都有所缩小,这一结果切实反映了十八大以来农村居民生活水平,特别是居住条件切实提高。当然,尽管城乡间居民住房资产的差异在2012—2020年有所缩小,但不可忽视的是,城乡间住房资产和住房条件的绝对差距水平仍然很大,2020年城乡间家庭人均净资产组间差异中,住房的贡献仍然达到95.05%。

(二) 家庭人均总财产的基尼系数分解结果

现实中,总资产或许比净资产更能反映人们的生活品质,一个家庭虽然由于背负房贷,家庭净资产并不高,但其生活品质却因为有更好的居住条件而明显提高。因此,这里我们也计算了家庭人均总资产的基尼系数及其分解结果。

如表10所示,2020年人均总资产基尼系数为0.6752,小于人均净资产的基尼系数0.6974,因此,人均总资产所反映的家庭间的差距没有人均净资产反映的那么大。我们知道,基尼系数满足规模不变性,因此如果家庭所借债务是家庭总资产的某一固定比例,那么净资产和总资产的基尼系数是相同的。而净资产的基尼系数更高,就表明存在总资产更高或者说更富有的家庭所借债务占比反而更低,而总资产更低或者说更贫穷的家庭所借债务占比反而更高的现象。特别是,考察数据可以发现,不少家庭由于各种原因负债超过了家庭资产,使得家庭人均净资产为负,这使得家庭人均净资产所反映的差距明显大于家庭人均总资产。从群组分解的结果看,人均总资产的基尼系数较人均净资产更低主要是由组内差距更低所致,而人均总资产和人均净资产的组间差距项差别不大,因此,房贷和非房贷金融负债对财产差距的影响更多体现在组内,特别是城镇内部。

表10 2012—2020年居民家庭人均总资产的基尼系数及分解

年份	人均总资产	组内差距	组间差距	组内差距占比	组间差距占比
2012	0.6488	0.6168	0.0321	95.06%	4.94%
2014	0.6458	0.6187	0.0271	95.80%	4.20%
2016	0.6741	0.6485	0.0256	96.20%	3.80%
2018	0.6726	0.6364	0.0362	94.62%	5.38%
2020	0.6752	0.6430	0.0321	95.24%	4.76%

注:为使结果可比,这里所采用的样本与前面的分析保持一致。

五、总 结

基于程永宏(2008)、Okamoto(2009)对基尼系数群组分解的探索,本文提出了一个新的基尼系数二维分解法,同时实现了基尼系数的按群组和按来源分解操作,避免了群组分解和来源分解的结果无法兼容和比较的问题^①。本文提出的方法要优于Mussard(2004)、Mussard和Savard(2012)基于传统的群组分解实现的基尼系数二维分解方法。

利用新分解方法和中国家庭追踪调查数据,本文对2012—2020年我国居民的财产分配差异变动情况进行了分析,研究发现:①我国城乡居民家庭财产分配差异在2012—2020年总体呈波动上升趋势,2020年基尼系数比2012年高0.0322,这一变动趋势主要是由城乡组内差距决定的,而组间差距与总体差距变动的趋势相反。②组内差距的财产来源分解显示,城乡组内差距中贡献最大的财产来源是住房资产(2020年占比79.46%)。2012—2020年组内差距扩大了0.0337,而住房资产的差距扩大程度达0.0586,住房资产对组内差距扩大的贡献比例达174%,对总体基尼系数扩大的贡献比例达182%,而房贷负债则起到缩小组内差距的作用,最终扣除房贷后的住房净资产对组内差距扩大的贡献占比为96%。③组间差距的来源分解结果表明,2020年城乡组间差距较2012年略微下降了0.0015,其中土地资产起到扩大城乡间差距的作用,提升城乡间差距值0.0047,这是由部分农村居民迁入城镇,把土地资产也带入城镇所致;金融资产也起到扩大城乡组间差距的作用,但作用占比小于土地资产;而住房资产与房贷负债在2012—2020年均起到缩小城乡间差距的作用,其中住房资产使组间差距下降0.0044,房贷使组间差距下降0.0039,两者合并得到的住房净资产使城乡间差距下降0.0084。因此,2012年以来,我国总体上由于居民持有房产(净房产)的差距不断扩大,居民财富分配差距也不断扩大,但城乡间居民持有房产(净房产)的差距却缩小了,并且起到缩小组间差距和总体差距的效果。与这一时期城乡居民收入差距缩小的趋势一致,这从财产的角度反映了党的十八大以来农村居民生活水平,特别是居住条件切实提高。当然,尽管城乡间居民住房财产的差异在2012—2020年有所缩小,但不可否认的是,城乡间住房资产和住房条件的绝对差距水平仍然很大,2020年城乡间家庭人均净资产组间差异中,住房的贡献仍然达到95.05%。

综上所述,本文揭示了住房资产在居民财产差距中的重要作用,发现城乡内部(特别是城镇内部)居民持有的住房资产(净资产)差异是2012—2020年这一时期财产分配差异扩大的重要原因,而城乡间差异,特别是城乡间居民住房资产(净资产)的差异在这一时期出现缩小并起到缩小整体财产差距的效果。

在方法层面,值得注意的一点是,过去进行群组分解常用泰尔指数等方法,而本文

^① 我们编写了执行基尼系数二维分解的Stata程序,读者若有需要可邮件向作者索取。

的测算显示,使用基尼系数群组分解方法得到的城乡组间差距占比明显小于泰尔指数分解的结果。然而,使用不同指标的结果是无法直接比较的,不能以一个指数的结果去评判另一个指数结果。为展示方法差异的影响,这里我们又计算了家庭人均总财产的城乡分组平均对数离差、泰尔指数和变异系数分解,表 11 展示了不同方法群组分解的结果。

表 11 2020 年家庭人均总财产的城乡群组分解

	差距值	组内差距占比	组间差距占比
基尼系数	0.6727	95.17%	4.83%
平均对数离差	1.0218	86.91%	13.09%
泰尔指数	0.9306	87.81%	12.19%
变异系数	4.4892	95.48%	4.52%

注: (1) 由于广义熵指数要求样本值为正,为排除使用样本不同造成的干扰,本表中各个不平等指标的计算和分解只囊括了 CFPS 数据中家庭总资产大于 0 的样本,因此计算结果与正文中表格略有差异。

(2) 此处变异系数分解使用的是 Akita 和 Miyata (2010) 的方法,差距值实际是标准变异系数的二次方。

可以发现,四个方法中得到组间占比最高的是平均对数离差,其次是泰尔指数,再次是基尼系数,而最低的是变异系数,基尼系数和变异系数分解得到的结果较为接近,相比之下,平均对数离差和泰尔指数得到的组间差距占比较高。

参考文献

- [1] 程永宏. 基尼系数群组分解新方法研究: 从城乡二亚组到多亚组[J]. 经济研究, 2008(8): 124-135, 144.
- [2] 靳永爱, 谢宇. 中国家庭追踪调查 2012 年和 2010 年财产数据技术报告[R]. 中国家庭追踪调查技术报告 CFPS-29, 2014.
- [3] 李 实, 万海远. 中国居民财产差距研究的回顾与展望[J]. 劳动经济研究, 2015(5): 28-44.
- [4] 刘学良, 田 青. 关于基尼系数按群组分解的进一步研究[J]. 数量经济技术经济研究, 2009(10): 98-111.
- [5] 罗楚亮, 陈国强. 富豪榜与居民财产不平等估算修正[J]. 经济学(季刊), 2021(1): 201-222.
- [6] 吕 萍. 中国家庭追踪调查数据权重使用常见问题[R]. 北京大学中国社会科学调查中心, 2016.
- [7] 万广华. 不平等的度量与分解[J]. 经济学(季刊), 2009(1): 347-368.
- [8] 谢 宇, 靳永爱. 家庭财产[R]//谢宇, 张晓波, 李建新, 等主编《中国民生发展报告 2014》. 北京: 北京大学出版社, 2014.
- [9] 谢 宇, 张晓波, 涂 平, 等. 中国家庭追踪调查用户手册(第三版)[R]. 2017.
- [10] 徐 宽. 基尼系数的研究文献在过去八十年是如何拓展的[J]. 经济学(季刊), 2003(3): 757-778.
- [11] 杨 灿, 王 辉. 社会网络与家庭财产差距——基于中国家庭追踪调查(CFPS)面板数据[J]. 经济与管理研究, 2020(4): 88-101.
- [12] Akita T., Miyata S. The Bi-dimensional Decomposition of Regional Inequality Based on the

- Weighted Coefficient of Variation[J]. *Letters in Spatial and Resource Sciences*, 2010, 3(3): 91-100.
- [13] Atkinson, A. B. On the Measurement of Inequality[J]. *Journal of Economic Theory*, 1970, 2(3): 244-263.
- [14] Blackorby C., Donaldson D. Measures of Relative Equality and Their Meaning in Terms of Social Welfare[J]. *Journal of Economic Theory*, 1978, 18(1): 59-80.
- [15] Blackorby C., Donaldson D., Auersperg M. A New Procedure for the Measurement of Inequality within and Among Population Subgroups[J]. *Canadian Journal of Economics*, 1981, 14(4): 665-685
- [16] Bourguignon F. Decomposable Income Inequality Measures[J]. *Econometrica*, 1979, 47(4): 901-920.
- [17] Cowell F. A. Measurement of Inequality[M]//Atkinson A. B. and Bourguignon F. (ed.) *Handbook of Income Distribution*, Vol 1, Elsevier, 2000.
- [18] Cowell F. A. On the Structure of Additive Inequality Measures[J]. *The Review of Economic Studies*, 1980, 47(3): 521-531.
- [19] Cowell F., Flachaire E. Statistical Methods for Distributional Analysis[M]//Atkinson A. B. and Bourguignon F. (ed.) *Handbook of Income Distribution*, Vol 2, Elsevier, 2015
- [20] Dagum C. A New Approach to the Decomposition of the Gini Income Inequality Ratio[J]. *Empirical Economics*, 1997, 22: 515-531.
- [21] Fei J. C. H., Ranis G., Kuo S. W. Y. Growth and the Family Distribution of Income by Factor Components[J]. *The Quarterly Journal of Economics*, 1978, 92(1): 17-53.
- [22] Kakwani N. C. Applications of Lorenz Curves in Economic Analysis[J]. *Econometrica*, 1977, 45(3): 719-727.
- [23] Mookherjee D., Shorrocks A. F. A Decomposition Analysis of the Trend in UK Income Inequality[J]. *Economic Journal*, 1982, 92(368): 886-902.
- [24] Mussard S. The Bidimensional Decomposition of the Gini Ratio. A Case Study: Italy[J]. *Applied Economic Letters*, 2004, 11(8): 503-505.
- [25] Mussard S., Savard L. The Gini Multi-decomposition and the Role of Gini's Transvariation: Application to Partial Trade Liberalization in the Philippines[J]. *Applied Economics*, 2012, 44: 1235-1249.
- [26] Mussini M. A Matrix Approach to the Gini Index Decomposition by Subgroup and by Income Source[J]. *Applied Economics*, 45: 2457-2468.
- [27] Okamoto M. Decomposition of Gini and Multivariate Gini Indices[J]. *Journal of Economic Inequality*, 2009, 7(2): 153-177.
- [28] Shorrocks A. F. The Class of Additively Decomposable Inequality Measures[J]. *Econometrica*, 1980, 48(3): 613-625.
- [29] Shorrocks A. F. Inequality Decomposition by Population Subgroups[J]. *Econometrica*, 1984, 52(6): 1369-1385.
- [30] Silber J. Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality[J]. *Review of Economics and Statistics*, 1989, 71(1): 107-115.

- [31] Sundrum R. M. *Income Distribution in Less Developed Countries*[M]. London and New York: Routledge, 1990.
- [32] Xie Y., Jin Y., *Household Wealth in China*[J]. *Chinese Sociological Review*, 2015, 47 (3): 203-229.

China's Household Wealth Distribution Inequality in 2012–2020: Investigation with Subgroups and Components Bi-dimensional Decomposition of Gini Coefficient

Liu Xueliang^{1,2}

(1. Institute of Economics, Chinese Academy of Social Sciences, Beijing 100836, China; 2. School of Economics, University of Chinese Academy of Social Sciences, Beijing 102488, China)

Abstract: This paper proposes a new Gini coefficient decomposition method, with which we can perform the Gini coefficient's subgroup decomposition and income components decomposition simultaneously, so as to integrate the inequality analysis of these two dimensions into one framework and avoid the inconsistency deficiency of present decomposition technique. Taking China's rural and urban resident's property distribution inequality as an example, we demonstrate the use of the new Gini Coefficient's bi-dimensional decomposition method, and find that within-group inequality of resident's housing assets is the main contributor of widening property inequality in 2012-2020, while urban-rural between group inequality, in particular inequality of net housing asset, contributed little to the widening property inequality during this period.

Keywords: Gini Coefficient; Subgroup Decomposition; Components Decomposition; Bi-dimensional Decomposition; Wealth Distribution Inequality

JEL Classification: D31 O15 C40

(责任编辑: 杨 光)

(责任校对: 刘 威)