

税收累进性 M 指数的方差估计及其应用^①

戴平生

(厦门大学经济学院统计系)

研究目标：推断税收累进性 M 指数的渐近正态分布性质，寻求税收累进性指数的方差估计。**研究方法：**利用 U—统计量构造税收累进性指数的渐近正态分布，并由该 U—统计量提炼辅助分布列获得 M 指数的方差估计。**研究发现：**几个常用的税收累进性指数都可以用 U—统计量的函数来表达，因而具有渐近正态分布性质，并获得方差估计。**研究创新：**给出税收累进性 M 指数的渐近分布性质和方差估计，还简化了税收累进性 K 指数、S 指数的方差估计。**研究价值：**可以用于政府的财税政策评价。将 M 指数的方差估计应用于我国分税制改革以来若干财税政策对省域收入公平性影响的评估发现实施分税制改革、取消农业税、营业税改增值税、提高个税起征点等都能显著增进中国大陆省域收入的公平性。

关键词 U—统计量 税收累进性 M 指数 方差估计 税收政策评估

中图分类号 F224.0 **文献标识码** A

DOI:10.13653/j.cnki.jqte.2019.03.008

引 言

税收累进性 M 指数是由国内学者根据税收关于税前收入基尼系数的相对边际效应提出的一种税收累进性测度（戴平生，2014）。M 指数从经济学角度解释了税收累进性为税收对税前收入基尼系数的贡献率与平均税率之差。即当该贡献率大于平均税率时税收是累进的，当其小于平均税率时税收是累退的，当其差值等于 0 时税收是成比例的。同时 M 指数对不同的税项具有可加性，对总体税收累进性进行结构性分析十分方便，可以给出不同税种在累进性中的贡献程度。M 指数大于 0 意味着相对边际效应大于 0，即税收有利于收入不平等的改善。由于 M 指数与其他传统的税收累进性 K 指数（Kakwani，1977）和 S 指数（Suits，1977）等测度值通常都为纯小数，实证分析中需要知道它们在统计意义上是否真正异于 0，以及不同时期或不同区域的税收累进性之间是否存在显著差异，因此知道这些税收累进性指数的分布是以正态分布为极限的，在大样本情况下就可以用正态分布来近似，这对区域的财税政策、税收改革评价有着十分重要的意义。

由于 M 指数本质上是一种基尼系数的相对边际效应，从理论上得到它的渐近正态分布性质就可以在若干领域对基尼系数的相对边际效应进行统计推断，对宏观或者微观社会经济现象不同来源、不同组群在区域收入或者资源配置不平等方面的影响进行分析。例如，对产业集中度、转移支付财力均等化、旅游改善收入不平等诸多研究都提供了一种推断检验方

^① 本文获得基金项目“拓展基尼系数的理论创新及其在我国社会经济共享发展中的应用研究”（16BTJ014）的资助。

法。研究基尼系数相对边际效应的国内外学者屈指可数，从 Lerman 和 Yitzhaki 于 1983 年在其工作论文中提出以来更多的是具体的应用，他们将某一收入来源增长所引起收入基尼系数的相对变化率定义为该项来源关于收入基尼系数的相对边际效应。若其相对边际效应大于 0，则该项来源的增长趋向于缩小收入差距；若其相对边际效应小于 0，则趋向于扩大收入差距。不仅各来源的相对边际效应之和等于 0，国内学者后来由于对税收累进性测度的改进又发现相对边际效应具有可加性（戴平生，2014）。因此，对 M 指数的方差估计可以推进基尼系数相对边际效应统计推断的研究。

本文的主要贡献体现在以下方面。一是给出了 M 指数的渐近正态分布性质，获得基尼系数相对边际效应的推断检验；二是极大地简化了 Arcarons 和 Calonge (2015) 在税收累进性 S 指数中对 U—统计量的探寻，直接给出 M 指数、K 指数和 S 指数辅助向量公式，使它们的方差估计变得十分便捷；三是给出计算 K 指数和 S 指数的协方差表达式，利用回归系数法估计它们的标准差；四是将文献中限于个体数据应用的方法拓展到分组数据，从而有助于使用统计年鉴的宏观数据进行财税政策的分析。本文的其他部分这样安排：第一部分对相关研究的文献进行综述，第二部分推导税收累进性 M 指数的渐近分布估计其方差，第三部分使用《中国税务年鉴》数据进行实证分析，最后给出结论。

一、相关文献综述

税收累进性 M 指数、K 指数和 S 指数都脱胎于以基尼系数为特例的集中度指数。基尼系数对应于以累计人口份额为横坐标、累计收入份额为纵坐标的洛伦兹曲线（根据人均收入递增排序），而 S 指数是在此基础上，以累计税前收入份额为横坐标、累计税收份额为纵坐标类比洛伦兹曲线的集中度指数；K 指数是用税收集中度指数与税前收入基尼系数之差定义的，M 指数则是税收关于税前收入基尼系数的相对边际效应。因此，对税收累进性指数的方差估计通常会从基尼系数或集中度指数入手。

1. U—统计量法

基尼系数源于测度数据离散程度的基尼平均差，是基尼平均差除以两倍的收入均值。若通过样本可以分别估计基尼平均差和收入均值两者的方差，那么通过一阶泰勒展开在大样本情况下就可以估计基尼系数的方差。Langel 和 Tillé (2013) 对数十种估计基尼系数或基尼平均差方差的做法进行了分析评论，他们把这些研究分为三个大的主要方向：调查抽样方法、稳健统计方法和经济学方法。这里按本文的思路仅对那些与税收累进性指数的方差估计关系更为密切的做法，如基尼平均差、U—统计量、回归系数和重抽样等方法进行必要的评述和讨论。对基尼平均差的方差估计可以追溯到 Nair (1936) 的研究，Glasser (1962) 也给出了基尼平均差和集中度指数的方差公式；基尼平均差、均值关于样本都满足对称性分别为一阶、二阶的 U—统计量，可以利用 U—统计量函数的渐近正态分布进行方差的估计 (Hoeffding, 1948)。Schechtman 和 Yitzhaki (1987) 利用 U—统计量讨论了基尼相关系数的渐近正态分布性质，其中的基尼相关系数等于基尼系数 G_x 与集中度指数 C_y 的比值。Bishop 等人 (1998) 在该研究的基础上归纳出了一类基于基尼税收累进性指数的推断检验，他们把税收累进性 K 指数作为该类指数的一种特例，因为它等于税收集中度指数 C_y 与税前收入基尼系数 G_x 之差。Arcarons 和 Calonge (2015) 构造了一个税收累进性 S 指数的 U—统计量，从而获得其渐近正态分布的性质及其方差估计。

2. 回归系数法

另一条思路是构造线性回归方程，通过估计回归方程的系数获得一类集中度指数及相应标准差（Kakwani 等，1997），这一方法对基尼系数也适用。税收累进性 K 指数等于集中度指数与基尼系数之差，可以构造一个线性回归方程对其参数进行估计并获得相应方差。

不妨设各研究单元（如中国大陆的 31 个省份）的人均收入按递增排序为 x_1, x_2, \dots, x_n ，其对应单元的人口数为 q_1, q_2, \dots, q_n ；记 q, S_x 分别为总人口数、总税前收入，即 $q = q_1 + \dots + q_n$ 和 $S_x = x_1 q_1 + \dots + x_n q_n$ ；记 $p_i = q_i/q$ 为第 i 个单元在总人口中的人口份额， F_i 是至第 i 组的累计人口份额（ $F_i = p_1 + \dots + p_i, i=1, \dots, n, F_0=0$ ）；记 $x_i q_i/S_x$ 为第 i 个单元在税前总收入中的收入份额， L_i 是至第 i 组的累计收入份额（ $L_i = x_1 q_1 + \dots + x_i q_i, i=1, \dots, n, L_0=0$ ）；而对应人均税收记为 t_1, t_2, \dots, t_n ，总税收收入记为 $S_t = t_1 q_1 + \dots + t_n q_n$ 。由收入份额法可以得到税前收入基尼系数以及税收集中度指数的协方差表达式（戴平生，2013）：

$$G_x = \text{Cov}\left(\frac{x}{\bar{x}}, \omega\right) \quad C_t = \text{Cov}\left(\frac{t}{\bar{t}}, \omega\right) \quad \bar{x} = \frac{S_x}{q} \quad \bar{t} = \frac{S_t}{q}$$

其中 $\omega_i = F_i + F_{i-1} - 1, i=1, \dots, n$ ，满足 $\omega_1 p_1 + \dots + \omega_n p_n = 0$ 。构造以下回归方程：

$$\sigma_\omega^2 \left(\frac{t_i}{\bar{t}} - \frac{x_i}{\bar{x}} \right) \sqrt{P_i} = \alpha \sqrt{P_i} + \beta \omega_i \sqrt{P_i} + \varepsilon_i$$

这里 σ_ω^2 为 $\{\omega_i, i=1, \dots, n\}$ 的方差。由残差平方和最小可以得到参数估计：

$$\beta = \text{cov}\left(\frac{t}{\bar{t}} - \frac{x}{\bar{x}}, \omega\right) = G_t - G_x = K$$

对于税收累进性 S 指数，由以累计税前收入份额为横坐标、累计税收收入份额为纵坐标的类比洛伦兹曲线，通过几何算法可以得到：

$$\begin{aligned} S &= \sum_{i=1}^n \frac{q_i t_i}{S_t} (L_i + L_{i-1} - 1) \\ &= \sum_{i=1}^n \frac{q_i x_i}{S_x} \left(\frac{t_i/x_i}{S_t/S_x} \right) (L_i + L_{i-1} - 1) = \text{cov}\left(\frac{t/x}{S_t/S_x}, \omega^* \right) \end{aligned}$$

其中 $\omega_i^* = L_i + L_{i-1} - 1, i=1, \dots, n$ ，它满足 $\omega_1^* x_1 q_1 + \dots + \omega_n^* x_n q_n = 0$ 。我们注意到 S 指数的协方差公式由两个因素构成，一个是各研究单元的平均税率与总体平均税率的比值，另一个是累计税前收入份额形成的组合系数。因此构造以下回归方程：

$$\sigma_{\omega^*}^2 \left(\frac{t_i/x_i}{S_t/S_x} \right) \sqrt{\frac{q_i x_i}{S_x}} = \alpha \sqrt{\frac{q_i x_i}{S_x}} + \beta \omega_i^* \sqrt{\frac{q_i x_i}{S_x}} + \varepsilon_i$$

这里 $\sigma_{\omega^*}^2$ 为 $\{\omega_i^*, i=1, \dots, n\}$ 的方差。由残差平方和最小就可以得到税收累进性 S 指数的参数估计，并获得标准差的估计值。通过构造相应回归方程得到的税收累进性指数的方差估计，因为回归模型中含有数据的求和、平均使得各观测值的独立性受到一定程度的影响，估计值可能出现一定程度的偏离。为弥补这一缺陷，Kakwani 等（1997）在其论文的附录中提供了一种集中度指数按一阶泰勒公式展开的方差估计方法。

随着计算机技术的发展，一些学者利用 Quenouille（1949）提出的刀切法（Jackknife）和 Efron（1979）提出的自助法（Bootstrap），通过重抽样估计税收累进性指数的方差（如

Yitzhaki, 1991; Mills 和 Zandvakili, 1997; Xu, 2000)。

税收累进性 M 指数没有单纯的协方差表达形式, 无法像 K 指数、S 指数那样通过构建回归方程获得方差估计。但 M 指数也是一种基于基尼系数的税收累进性指数, 它可转化为集中度指数与均值的函数, 而集中度指数和均值都可以用 U-统计量来表示 (集中度指数包含了基尼系数)。因此 M 指数的方差估计可以使用 U-统计量法, 或者使用重抽样方法。

二、税收累进性 M 指数的方差估计

税收累进性指数的方差估计有非参数方法和参数方法。非参数方法目前人们更多地采用更适合非线性变化的 Bootstrap 重抽样技术, 参数方法则需要找出税收累进性指数的分布或渐近分布。M 指数、K 指数、S 指数如果可以获得相应的渐近正态分布, 这样就便于对不同点点和不同区域的测度进行比较, 因此选择参数方法有其独特的优势。

沿用前文的变量及指标符号, 下面先给出 M 指数的计算公式:

$$M = s(t) - \frac{S_t}{S_x} \quad s(t) = \frac{\sum q_i t_i \omega_i}{\sum q_i x_i \omega_i} \quad (1)$$

即税收累进性 M 指数等于税收对税前收入基尼系数的贡献率减去平均税率。M 指数对各个税项具有可加性, 且与 K 指数 (Kakwani, 1977) 具有相同符号。一些学者发现 K 指数和 S 指数在部分税前税后收入基尼系数很小的情况下变化十分敏感, 有时还出现符号相反的情形 (李香菊和祝丹枫, 2016)。而 M 指数的符号与税收关于税前收入基尼系数的相对边际效应保持一致, 具有更为稳健的性质。

下面具体讨论税收累进性 M 指数的方差估计问题。将式 (1) 略作变形就可以得到以下的表达形式:

$$M = \frac{S_t C_t}{S_x G_x} - \frac{S_t}{S_x} = \frac{\bar{t} C_t}{\bar{x} G_x} - \frac{\bar{t}}{\bar{x}} = \frac{\bar{t}}{\bar{x} G_x} K \quad (2)$$

它表明, 税收累进性 M 指数是均值与集中度指数的函数。下面先引进两个 U-统计量 $\theta_x = \sum_{i,j} (x_i - x_j) I_{(x_i > x_j)} p_i p_j$ 和 $\theta_t = \sum_{i,j} (t_i - t_j) I_{(x_i > x_j)} p_i p_j$, 其中

$$I_{(x_i > x_j)} = \begin{cases} 1 & \text{如果 } x_i > x_j \\ 0 & \text{如果 } x_i = x_j \\ -1 & \text{如果 } x_i < x_j \end{cases}$$

这里是将 Bishop 等 (1998) 研究中的个体数据替换为分组数据, 把 $q (q-1)$ 变为 q^2 (不会影响渐近性质)。显然 θ_x 就是组数据的基尼平均差, θ_t 是在基尼平均差的基础上由成对随机变量定义的。于是基尼系数和集中度指数可以分别表示为 $G_x = \theta_x / 2\bar{x}$ 、 $C_t = \theta_t / 2\bar{t}$, 税收累进性指数 $K = \theta_t / 2\bar{t} - \theta_x / 2\bar{x}$ 和 $M = \theta_t / \theta_x - \bar{t} / \bar{x}$ 。在 $x_1 \leq x_2 \leq \dots \leq x_n$ 的递增排序假定下, 以上两个 U-统计量可以表示为:

$$\theta_x = 2 \sum_{i=1}^n \sum_{j=i}^n (x_j - x_i) p_i p_j \quad (3)$$

$$\theta_t = 2 \sum_{i=1}^n \sum_{j=i}^n (t_j - t_i) p_i p_j \tag{4}$$

对于税收累进性 S 指数，Arcarons 和 Calonge (2015) 给出的 U-统计量过程相当复杂，其实根据税收累进性 S 指数的定义，利用几何算法可以大大简化其中的过程。即从：

$$S = \sum_{i=1}^n \frac{q_i t_i}{S_t} (L_i + L_{i-1} - 1)$$

可以导出：

$$S = \sum_{i=1}^n \frac{q_i t_i}{S_t} \left(2L_i - \frac{q_i x_i}{S_x} \right) - 1 = \sum_{i=1}^n (2t_i \sum_{j=1}^i x_j p_i p_j - t_i x_i p_i p_i) / (\bar{x} \times \bar{t}) - 1$$

这样就引入了第三个 U-统计量

$$\theta_z = \sum_{i=1}^n (2t_i \sum_{j=1}^i x_j p_i p_j - t_i x_i p_i^2) \tag{5}$$

于是 $S = \theta_z / (\bar{x} \times \bar{t}) - 1$ 。由于 \bar{x} 和 \bar{t} 本身就是 U-统计量，因此 M 函数、K 函数、S 函数都是 U-统计量的函数，它们收敛于正态分布函数 (Bishop 等, 1998; Arcarons 和 Calonge, 2015)。这样税收累进性指数就可以通过渐近正态分布，获得各自方差的近似估计。但由于上述引入的三个 U-统计量都是二阶的，估计这些指数的协方差矩阵需要计算繁杂的三阶的样本 U-统计量。为了避免类似于 Bishop 等 (1998) 论文中的繁琐计算，下面通过一阶泰勒展开式计算各个税收累进性关于 U-统计量函数的方差。离散总体税收累进性 M 指数的微分形式为：

$$dM = -\frac{1}{\mu_x} d\mu_t + \frac{\mu_t}{\mu_x^2} d\mu_x + \frac{1}{\theta_x} d\theta_t - \frac{\theta_t}{\theta_x^2} d\theta_x$$

根据 Rao (1973) 提出的 δ 方法，我们有：

$$\begin{aligned} q\hat{\text{var}}(M) &= \frac{1}{x^2} \hat{\text{var}}(\bar{t}) + \frac{\bar{t}^2}{x^4} \hat{\text{var}}(\bar{x}) + \frac{1}{\theta_x^2} \hat{\text{var}}(\theta_t) + \frac{\theta_t^2}{\theta_x^4} \hat{\text{var}}(\theta_x) \\ &\quad - \frac{2\mu_t}{\mu_x^3} \hat{\text{cov}}(\bar{t}, \bar{x}) - \frac{2}{\mu_x \theta_x} \hat{\text{cov}}(\bar{t}, \theta_t) + \frac{2\theta_t}{\mu_x \theta_x^2} \hat{\text{cov}}(\bar{t}, \theta_x) + \frac{2\mu_t}{\mu_x^2 \theta_x} \hat{\text{cov}}(\bar{x}, \theta_t) \\ &\quad - \frac{2\mu_t \theta_t}{\mu_x^2 \theta_x^2} \hat{\text{cov}}(\bar{x}, \theta_x) - \frac{2\theta_t}{\theta_x^3} \hat{\text{cov}}(\theta_t, \theta_x) \end{aligned}$$

等价于：

$$q\hat{\text{var}}(M) = \text{var}(\omega_M) = \sum_{i=1}^n \omega_i^2 p_i - \left(\sum_{i=1}^n \omega_i p_i \right)^2 \tag{6}$$

其中：

$$\omega_M^i = -\frac{1}{x} t_i + \frac{\bar{t}}{x^2} x_i + \frac{1}{\theta_x} \theta_t^i - \frac{\theta_t}{\theta_x^2} \theta_x^i \tag{7}$$

这里 $\theta_x^i = 2 \sum_{j=i}^n (x_j - x_i) p_j$ ， $\theta_t^i = 2 \sum_{j=i}^n (t_j - t_i) p_j$ 和 $\theta_z^i = 2t_i \sum_{j=1}^i x_j p_j - t_i x_i p_i$ ，它们都可以从三个 U-统计量中直接剥离出来。

Arcarons 和 Calonge (2015) 在实证中也推荐了这一辅助向量方法, Kakwani 等 (1997) 在其附录中也给出了类似的集中度指数的方差估计。该辅助向量确实极大地简化了 Bishop 等 (1998) 的计算。对于税收累进性 K 指数、S 指数, 也可以对方差进行类似的估计。对应的辅助向量为:

$$\omega_K^i = -\frac{\theta_t}{2t^2}t_i + \frac{\theta_x}{2x^2}x_i + \frac{1}{2t}\theta_t^i - \frac{1}{2x}\theta_x^i \quad (8)$$

$$\omega_S^i = -\frac{\theta_z}{xt^2}t_i - \frac{\theta_z}{x^2t}x_i + \frac{1}{xt}\theta_z^i \quad (9)$$

税收累进性 M 指数、K 指数和 S 指数都服从渐近正态分布。这表明在大样本条件下, 我们可以用正态分布进行近似。它们近似服从正态分布, 一方面可以检验税收累进性指数是否异于 0 (即税收的效果是成比例的, 不累退也不累进); 另一方面可以对不同时间、空间的税收累进性是否存在显著差异进行比较, 显然这种检验对于讨论不同时期、不同区域税收政策改善区域收入不平等的作用效果是十分重要的。

三、我国分税制以来的财税政策评价

1994 年实行的分税制, 在我国财税体制改革进程中具有重要的历史地位。有专家学者把改革开放 40 年以来的中国税制改革划分为三个阶段: 第一阶段为有计划的商品经济时期, 时间为 1978~1993 年; 第二阶段是社会主义市场经济初期, 时间为 1994~2000 年; 第三阶段是社会主义市场经济完善期, 时间从 2001 年至今。我们可以考察财税改革过程中若干重大事件下的税收累进性变化。

税收作用通常体现在两个方面, 一是产生财政收入为政府运行提供资金来源, 二是实现二次分配增进社会群体与个人的收入公平。税收是否增进社会公平, 税收累进性指数可以对这一作用的方向和大小进行测度。利用《中国税务年鉴》各年的省级税收统计数据, 以各省份行政单元的 GDP 为宏观收入 (即以省级行政区为社会群体单元), 配合相应人口数就可以计算税收累进性 M 指数、K 指数和 S 指数。这里采用 1993~2016 年的数据, 主要讨论 1994 年分税制以来一些重要的税制改革对税收累进性的影响, 因此除综合税收之外, 还将计算分析增值税、消费税、营业税、企业所得税和个人所得税等主要税种的累进性分析它们的变化。其中的 1993 年仍处于分税制改革前, 增值税等税种的统计口径不同 (1993 年外资企业的统一工商税尚未细分), 将该年度税收数据也纳入研究的范围主要是为了观察分税制改革前后综合税收的累进性变化。此外为保持税收统计口径的相对一致, 从 1993~1997 年的总税收收入中扣除教育费附加、国家能源交通重点建设基金、国家预算调节基金及其他收入。表 1 给出了总税收的三种税收累进性指数和标准差, 同时还列出了总税收前后宏观收入基尼系数和标准差。图 1 给出了 1993~2016 年我国总体税收三种税收累进性的测度变化。同时为了便于不同方法的比较, 各年度税前税后基尼系数的相对变化率也标示其中。

显然三种税收累进性指数中, M 指数与税前税后基尼系数相对变化率曲线最为接近。相对变化率等于税前基尼系数减税后基尼系数之差除以税前基尼系数, 它相当于税收关于税前基尼系数的相对边际效应。而 M 指数就是从相对变化率中剥离出来的核心因子, 因此两者的联系自然最为紧密。从图 1 中也可以看出当相对变化率较小时, 税收累进性 K 指数和 S 指数更为敏感。而各种曲线都在横轴上方, 表明 1993 年以来我国税收都能改善省际的收入不平等。

表1 1993~2016年税前税后基尼系数、M指数、K指数和S指数及其标准差(10⁻⁶)

年 度	M	σ_M	K	σ_K	S	σ_S	G_0	σ_0	G_1	σ_1
1993	0.0145	1.6280	0.0359	4.8012	0.0544	43.2578	0.2537	4.0916	0.2508	3.8444
1994	0.0282	1.9603	0.0706	6.7876	0.0955	48.5582	0.2597	4.2048	0.2538	3.8356
1995	0.0148	1.3976	0.0566	6.8418	0.0847	48.1139	0.2569	4.0489	0.2539	3.8258
1996	0.0343	1.8424	0.0904	7.3508	0.1171	49.2226	0.2537	3.9516	0.2455	3.5499
1997	0.0248	2.2559	0.0690	8.4814	0.0989	52.9304	0.2578	4.0757	0.2517	3.6789
1998	0.0371	2.0736	0.0943	8.0861	0.1229	52.3237	0.2624	4.2392	0.2524	3.7357
1999	0.0498	2.4755	0.1188	9.8426	0.1520	57.3604	0.2685	4.4660	0.2546	3.7929
2000	0.0707	2.6953	0.1513	10.8668	0.1868	58.2217	0.2577	4.0244	0.2374	3.2162
2001	0.0836	3.0031	0.1686	11.9012	0.2062	61.6786	0.2770	4.6256	0.2507	3.5626
2002	0.0873	2.9788	0.1705	11.6130	0.2064	60.1551	0.2694	4.3454	0.2425	3.3099
2003	0.0930	3.2136	0.1790	12.4250	0.2172	61.8766	0.2764	4.5091	0.2467	3.3766
2004	0.0974	3.2988	0.1760	11.9554	0.2122	60.2154	0.2716	4.3278	0.2407	3.1853
2005	0.0978	3.3590	0.1681	11.2512	0.2051	57.4142	0.2682	4.1190	0.2377	3.0457
2006	0.1010	3.5476	0.1641	11.0106	0.2004	55.8030	0.2653	3.9578	0.2336	2.9055
2007	0.1296	4.4391	0.1856	12.7217	0.2262	58.4810	0.2569	3.6642	0.2173	2.4920
2008	0.1292	4.0172	0.1788	11.1368	0.2157	53.0822	0.2446	3.2601	0.2081	2.2783
2009	0.1170	3.7599	0.1631	10.1035	0.1975	50.3960	0.2409	3.1578	0.2083	2.2845
2010	0.1241	3.3697	0.1589	8.5258	0.1884	45.2456	0.2267	2.7728	0.1947	1.9964
2011	0.1344	3.4564	0.1576	8.0835	0.1836	42.9171	0.2153	2.4836	0.1834	1.7735
2012	0.1368	3.4490	0.1479	7.2887	0.1709	4.01426	0.2077	2.3005	0.1781	1.6690
2013	0.1353	3.4046	0.1462	7.1747	0.1695	3.95683	0.2045	2.2258	0.1765	1.6380
2014	0.1365	3.5318	0.1466	7.3688	0.1710	3.99868	0.2033	2.1971	0.1762	1.6265
2015	0.1406	3.7810	0.1543	8.1385	0.1802	4.20852	0.2064	2.2638	0.1788	1.6725
2016	0.1374	3.6435	0.1599	8.4796	0.1878	4.34605	0.2097	2.3428	0.1803	1.7004

注： σ_M ， σ_K ， σ_S ， σ_0 ， σ_1 分别表示税收累进性 M，K，S 和税前税后基尼系数 G_0 ， G_1 的标准差。由于省级数据是经过汇总和在很大程度上消除了个体数据的波动性，使得在大的人口数下这些标准差微小保证了各种指数显著异于 0 不能忽略。

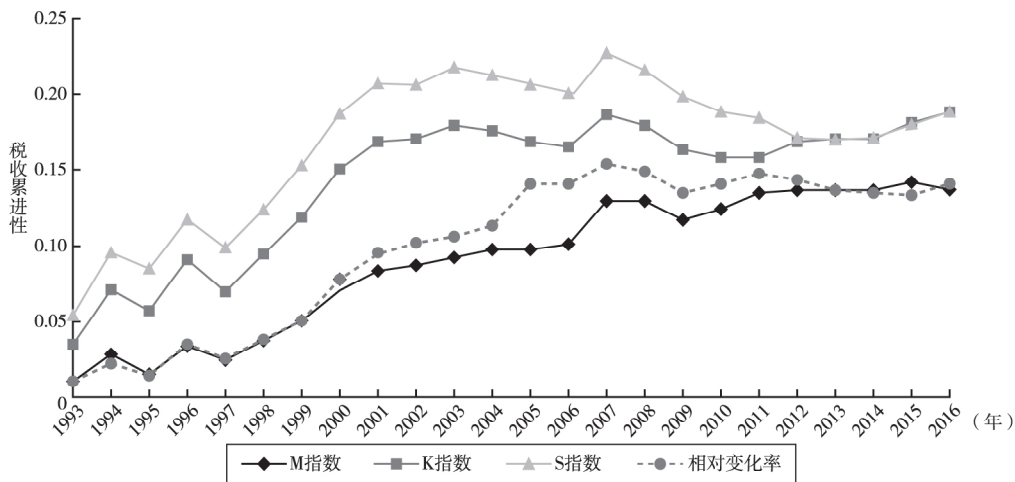


图1 1993~2016年我国税收累进性及税前税后基尼系数相对变化率

1. 分税制改革

根据税收累进性 M 指数的计算结果, 1994 年我国的税收累进性大于分税制改革前的 1993 年, 利用 1993 年和 1994 年税收累进性 M 指数及其标准差的估计容易验证: 不仅两个年度的税收是显著累进的 ($0.0145 \times 10^6 / 1.628 > 3$ 、 $0.0282 \times 10^6 / 1.960 > 3$ 检验统计量远远超出了显著性水平 1% 的临界值); 而且由 $(0.0282 - 0.0145) \times 10^6 / (1.628^2 + 1.960^2)^{1/2} > 3$ 表明两者存在显著的差异。同样, 1995~2016 年的税收累进性都显著大于 0, 也显著大于 1993 年的税收累进性指数。

我国分税制改革的确定, 对于建立公共财政体系和推进市场经济体制具有里程碑意义。分税制是市场经济国家普遍实行的一种财政体制, 实践证明它能较好地解决中央集权与地方分权的问题。十四大将建立社会主义市场经济作为经济体制改革的目标, 分税制的实施不仅推动社会主义市场经济的建立, 理顺中央与地方的分配关系, 调动中央、地方的两个积极性, 同时为当时中国的“复关”及之后的“入世”, 积极融入世界贸易体系奠定了基础。一方面对我国的工商税制进行了全面的改革。建立了以增值税、消费税和营业税为内容的新的流转税课税体系, 外资企业停征工商统一税纳入新的流转税体系, 内资企业实行统一的企业所得税, 个人收入和个体工商户的经营所得实行修订后的《个人所得税法》, 还调整和开征其他的一些税种。另一方面, 在管理体制上国家税务总局在省级及以下行政区设立了国家税务局和地方税务局, 税目则分为中央税、地方税和共享税。其中的中央税和共享税由国税系统负责征收管理, 地方税则由地税系统负责征收管理, 从而加强了我国的税收征管。

2. 农业税改革

农业税是国家对从事农业生产、有农业收入的单位和个人征收的一种税, 俗称“公粮”。《中华人民共和国农业税条例》(1958) 和《关于对农业特产收入征收农业税的规定》(1994) 是我国征收农业税、农业特产税的纲领性文件, 因此这里的农业税改革是指涉农税项的农村税费改革, 其中农业税、屠宰税是我国主要的两大涉农税项。

我国税收统计中并没有各省份的农业税收入数据, 而屠宰税曾是地方财政收入的重要来源之一, 作为单一税项统计。农村税费改革始于 2001 年, 当时中央政府计划五年内完全取消农业税。作为农村税费改革的最终成果, 各种涉农税项于 2006 年完全废止。为了评估取消农业税对改善我国区域收入不平等的影响, 这里只能以各省份第一产业增加值作为农业收入的代理变量, 并根据农业税全国的平均税率为常年产量的 15.5% 估算农业税。时间以农村税费改革的前三年即 1998 年、1999 年和 2000 年作为观察点, 如果农业税显著恶化了省份间的收入不平等即农业税收入具有累退性, 那么农业税的废止就有利于提高收入公平性。表 2 是三年农业税收入累进性 M 指数及标准差的估计结果。

表 2 农村税费改革前涉农税项税收累进性 M 指数及其标准差 (10^{-6})

年 度	农业税		屠宰税		宏观收入			
	M	σ_M	M	σ_M	G_0	σ_0	G_1	σ_1
1998	-0.0185	0.5100	-0.0004	0.0092	0.2624	4.2392	0.2679	4.4391
1999	-0.0176	0.4847	-0.0004	0.0092	0.2685	4.4660	0.2737	4.6582
2000	-0.0185	0.4394	-0.0005	0.0103	0.2577	4.0244	0.2627	4.1964

注: 农业税按第一产业增加值的 15.5% 测算, 符号含义与表 1 相同。实际上 GDP 基尼系数因征收两税项而增大。

测算农产品征收 15.5% 的税收累进性指数表明该税项具有显著的累退性，屠宰税的指数计算也给出了具有显著累退性的结果（如 1998 年的农业税 $-0.0185 \times 10^6 / 0.51 > 3$ ，2000 年的屠宰税 $-0.0005 \times 10^6 / 0.0103 < -3$ ）。因此由 M 指数的可加性可以得出取消涉农税项能提高其他税项综合税收的累进性。表 2 的宏观收入一栏还给出了扣除两项税收前后的基尼系数变化情况，税后基尼系数的提高说明涉农税项恶化了宏观收入的公平性。农业税的免除提高了农民纯收入从而改善区域收入不平等。

3. 中央与地方共享企业所得税

1994 年分税制改革后，增值税、消费税、营业税、企业所得税、个人所得税、关税成为我国财税收入的六大税种。在进入 21 世纪之前中央税以消费税、关税为主，地方税以营业税、企业所得税和个人所得税为主，中央与地方的共享税则以增值税为主。当时增值税收入占财税收入的 35% 以上，是我国的第一大税种；然后依次是营业税、企业所得税、关税和个人所得税，其中企业所得税与营业税的收入规模相近。2000 年中央政府开始实施西部大开发政策，要把东部沿海地区剩余的经济能力用于提高西部地区的经济和社会发展水平，以巩固国防。2002 年中央政府决定将企业所得税从地方税转为中央和地方共享税五五分成，后又调整为六四分成。

企业所得税的税收征管权从地税系统转为国税系统，中央政府将企业所得税的收入用于西部大开发等区域发展战略的建设。其作为中央与地方的共享税，通过中央政府的转移支付不仅有利于缩小我国的区域收入差距，同时收入规模极大的企业所得税能够得到更好的税收征管。2001 年以来企业所得税的收入在我国综合税收中的占比稳步提高，目前已成为仅次于增值税收入的第二大税项，约占我国财税收入总量的 20%。表 3 给出了企业所得税在 2002 年前后各三年 M 指数的计算结果。

表 3 企业所得税的 M 指数及其标准差

年 度	M	σ_M	年 度	M	σ_M
1999	0.0118	5.3929×10^{-7}	2003	0.0222	8.7958×10^{-7}
2000	0.0167	6.9700×10^{-7}	2004	0.0249	9.3608×10^{-7}
2001	0.0203	8.7374×10^{-7}	2005	0.0293	1.2630×10^{-6}

企业所得税的 M 指数都显著异于 0，且后三年也都显著大于前三年的税收累进性水平。可以认为作为中央和地方共享税，提高了企业所得税征管水平，企业可获得了一个更为公平、公开、公正的纳税环境，促进区域收入公平的税收累进性有了显著的提高。

4. 内外资企业所得税合并

我国的改革开放初期，为了引进和规范外商来华投资开展对外经济合作中央政府形成了一系列涉外的企业所得税、个人所得税、工商统一税、城市房地产税等税法。1994 年的分税制改革取消了外资企业的工商统一税，实行增值税、消费税等新的流转税制度。在相当长的一段时期外资企业享有各种税收优惠政策，企业所得税的实际税率明显低于内资企业。为适应建立社会主义市场经济体制和扩大改革开放的要求，根据统一税法、简化税制、公平税负、促进竞争的原则，从 2008 年起内外资企业统一所得税税率。

许多研究表明，外商直接投资扩大了我国沿海地区与内陆地区的收入差距。2000 年后随着我国一系列区域发展战略的实施，经济更为发达的沿海地区陆续将剩余生产能力向中部和西部地区转移。内外资企业所得税的统一有利于内外资企业在中西部的公平竞争，为中部

和西部地区的社会经济发展提供了一个更加规范、统一、公平、透明的税收法律制度环境。以 2008 年前后三年企业所得税 M 指数变化为考察对象，表 4 给出了内外资企业所得税税率合并前后 M 指数的变化。各年企业所得税 M 指数都显著异于 0，且后三年的数值显著大于前三年的数值，表明内外资企业所得税的税率合并有利于增进区域收入的公平性。

表 4 企业所得税的 M 指数及其标准差

年 度	税率合并前		年 度	税率合并后	
	M	σ_M		M	σ_M
2005	0.0293	1.2630×10^{-6}	2009	0.0418	1.8981×10^{-6}
2006	0.0319	1.4838×10^{-6}	2010	0.0400	1.3468×10^{-6}
2007	0.0384	1.6476×10^{-6}	2011	0.0451	1.4663×10^{-6}

5. “营改增”能否增进区域收入公平

增值税是对商品生产、流通、劳务服务中各环节的新增价值或商品的附加值征收的一种流转税。我国采用国际上普遍认可的税款抵扣方式，即按规定税率对销售的商品或劳务计算出销售税额，然后扣除取得该商品或劳务时所支付的增值税费（进项税额）即为所交税款。我国从 2009 年 1 月 1 日开始全面实施消费型的增值税。“营改增”是指将以前缴纳营业税的应税项目改成缴纳增值税，可以避免营业税重复征税、不能抵扣、不能退税的弊端。增值税抵扣有利于企业降低税负、提高盈利能力，同时也是引导企业转型升级的重要经济手段。

我国从 2012 年开始进行“营改增”试点，经过近五年的实践“营改增”在降低企业税负、增加企业发展活力方面取得了良好的效果。2017 年 10 月 30 日，中央政府宣布废止营业税暂行条例，营业税在我国正式退出历史舞台。营业税作为地方税曾是地方政府的主要财政收入来源之一，营业税的逐步退出意味着增值税收入在国家财税综合税收总量中的比重将进一步提高。增值税是最大的中央和地方共享税，共享税收入的增加必然通过转移支付对区域收入不平等的改善产生积极的作用。我国消费型增值税的全面实施，以及“营改增”的过程一方面会带来增值税 M 指数自身的变化，另一方面营业税与增值税收入的此消彼长也会带来综合税收累进性变化。由于 M 指数具有可加性，通过观察营业税与增值税两大税项收入叠加后的 M 指数，可以判断“营改增”过程对综合税收累进性的影响。为此我们对比 2012 年“营改增”试点前后三年营业税和增值税收入合体 M 指数的变化（见表 5），观察该税制改革是否增进了区域收入公平性。

表 5 增值税、营业税 M 指数及其标准差

	年 度	增值税		营业税		叠加处理		
		M	σ_M	M	σ_M	M	平均值	对应 σ_M
合并前	2009	0.0540	1.3290×10^{-6}	0.0122	5.0801×10^{-7}	0.0662	0.0696	2.7061×10^{-7}
	2010	0.0590	1.3192×10^{-6}	0.0111	4.4421×10^{-7}	0.0701		
	2011	0.0613	1.3326×10^{-6}	0.0113	4.4025×10^{-7}	0.0726		
合并中	2013	0.0658	1.3202×10^{-6}	0.0072	3.1902×10^{-7}	0.0686	0.0708	2.5760×10^{-7}
	2014	0.0639	1.3414×10^{-6}	0.0069	3.1315×10^{-7}	0.0730		
	2015	0.0582	1.2484×10^{-6}	0.0104	2.7266×10^{-7}	0.0708		

“营改增”的前后三年，增值税、营业税的 M 指数都显著异于 0，前三年增值税的 M 指数略呈上升变化，而营业税的 M 指数略呈下降变化；但后三年两者都略呈反向变化，即没有出现增值税在税收总量中的份额上升而 M 指数上升，营业税的份额下降而 M 指数下降的情形。可以判断增值税与营业税在各省份的分布结构较为相似，即营业税改增值税一方面提高了自身在区域收

入分配中的均衡能力,另一方面成为增值税后略微削弱了增值税的累进性,因此“营改增”是否增进区域收入均衡性的判断过程较为复杂。但通过前后两者叠加值的M指数进行简单平均,计算标准差后(假定各年M指数相互独立)对前后三年的平均数进行统计意义上的是否相等检验,发现前后M指数存在显著差异,即可以认为“营改增”能够显著改善区域收入不平等。

6. 我国东中西部地区的个税累进性

个人所得税是以个人取得的各项应税所得为对象征收的一个税种。新中国的个人所得税始于1980年,当时的主要征税对象是外国人。1986年本国开始征收个人收入调节税,1994年分税制改革将个人收入调节税、个体工商户所得税一起并入个人所得税。至目前为止我国的个人所得税法先后经过了数次修订,起征点也从1980年的800元到2005年的1600元、2008年的2000元、2011年的3500元;2018年10月1日起将调至5000元。个税起征点的变化,一方面反映了我国居民收入水平的逐步提高,另一方面也是中央政府让人民群众共享改革开放成果的重要体现。长期以来许多学者讨论在不同社会经济发展水平下适宜起征点的问题。显然从税收作为一种调节社会收入公平的重要手段来看,个税起征点并非越高越好。

1994年分税制改革以来,个税起征点经历了2005年、2008年和2011年的三次调整。为了观察个税起征点的变化对个人所得税改善区域收入不平等的作用,我们计算比较了个税起征点2005年调整前三年到2011年后三年M指数的变化,同时考虑了我国东部、中部和西部地区个人所得税累进性的情况(见表6)。

表6 起征点调整下个人所得税M指数及其标准差(10^{-7})

年 度	全 国		东 部		中 部		西 部	
	M	σ_M	M	σ_M	M	σ_M	M	σ_M
2002	0.0069	2.4921	0.0155	3.7948	0.0031	0.6196	-0.0013	0.9893
2003	0.0074	2.7522	0.0175	4.1678	0.0028	0.4961	-0.0024	0.8693
2004	0.0077	2.8506	0.0175	4.3431	0.0029	0.4818	-0.0023	0.8258
2005	0.0077	2.9498	0.0196	4.5841	0.0033	0.8435	-0.0023	0.8289
2006	0.0080	3.0189	0.0193	4.7345	0.0022	0.9334	-0.0025	0.8796
2007	0.0091	3.2560	0.0229	5.1372	0.0031	1.0515	-0.0023	1.0381
2008	0.0103	3.4827	0.0274	5.7155	0.0031	1.2356	-0.0013	1.1535
2009	0.0097	3.2674	0.0260	5.3715	0.0024	1.1697	-0.0018	1.1965
2010	0.0106	3.1746	0.0279	5.2700	0.0010	1.1884	-0.0023	1.3143
2011	0.0116	3.2968	0.0294	5.6554	0.0032	1.4403	-0.0018	1.7288
2012	0.0111	3.1431	0.0280	5.4141	0.0041	1.3259	-0.0024	3.0696
2013	0.0119	3.2286	0.0294	5.6591	0.0029	1.5310	-0.0013	1.6520
2014	0.0132	3.4429	0.0310	6.0382	-0.0014	1.2899	-0.0013	1.5311

注:这里将中国大陆的31个省份划分为东、中、西部三个区域,其中北京、天津、河北、辽宁、上海、江苏、浙江、福建、山东、广东、海南11省份划入东部地区;山西、吉林、黑龙江、安徽、江西、河南、湖北、湖南8省份划入中部地区,其余的12省份划入西部地区。

各年全国、东部、中部和西部的个人所得税M指数都显著异于0,同时还可以验证不同M指数的显著差异。从全国来看,历次个税起征点的调整总体上带来了我国个人所得税累进程度的提高,即提高个税起征点可以增进中国大陆省份的收入公平性。从东中西部来看,个税起征点对东部、中部和西部的影响并不相同。东部地区的个人所得税M指数不仅数值上明显大于中部和西部地区,而且总体呈现出上升趋势,表明个税起征点的调整有利于增进东部地区省间的收入公平

性；但个税起征点的调整在中西部地区的作用并不明显，两者的 M 指数的数值小且总体都出现了上下波动的变化。西部地区个人所得税 M 指数呈现出长期的微弱累退特征，2014 年中部地区的个人所得税也是累退的，表明中西部地区在个人所得税税收征管方面可能因各省份执行力度不同出现了差异。因为个人所得税的征收使用了累进税率，在同等力度下不会出现累退的情形。

四、结 论

本文给出了税收累进性 M 指数的方差估计，理论是建立在分组数据基础之上的，这为大量存在于各级统计年鉴的财税数据提供了一种可以进行统计推断的重要工具。个体数据作为分组数据的一种特例，也可以使用该方法。文中还推导了税收累进性 K 指数和 S 指数的协方差表达式，给出两者方差估计的回归系数法。税收累进性 M 指数是定义在基尼系数的相对边际效应的基础上，上述方法也适用于各领域涉及边际效应的研究。将该方法应用于对我国 1994 年分税制改革后的财税政策评估研究，发现分税制实施后总体税收的累进性呈现显著的上升趋势，取消农业税显著地增进中国大陆的地区收入公平性，企业所得税等原地方税种转变为中央和地方共享税后显著地改善地区收入不平等。合并内外资企业所得税的税率提升了企业所得税改善地区收入不平等的作用。“营改增”总体上有利于增进中国大陆的地区收入公平性，个税起征点的数次调整在总体上能够增强个税在改进地区收入不平等中的作用，但对东中西部地区收入不平等的作用存在显著差异。

参 考 文 献

- [1] Arcarons J. , Calonge S. , 2015, *Inference Tests for Tax Progressivity and Income Redistribution : the Suits Approach* [J], *Journal of Economic Inequality*, 13 (2), 207~223.
- [2] Bishop J. A. , Formby J. P. , Zheng B. , 1998, *Inference Tests for Gini-Based Tax Progressivity Indexes* [J], *Journal of Business & Economic Statistics*, 16 (3), 322~330.
- [3] Efron B. , 1979, *Bootstrap Methods: Another Look at the Jackknife* [J], *Annals of Statistics*, 7 (1), 1~26 .
- [4] Glasser G. , 1962, *Variance Formulas for the Mean Difference and Coefficient of Concentration* [J], *Journal of the American Statistical Association*, 57 (299), 648~654.
- [5] Hoeffding W. , 1948, *A Class of Statistics with Asymptotically Normal Distribution* [J], *Annals of Mathematical Statistics*, 19 (3), 293~325.
- [6] Kakwani N. C. , 1977, *Measurement of Tax Progressivity: An International Comparison* [J], *Economic Journal*, 87 (345), 71~80.
- [7] Kakwani N. C. , Wagstaff A. , Doorslaer E. V. , 1997, *Socioeconomic Inequalities in Health: Measurement, Computation, and Statistical Inference* [J], *Journal of Econometrics*, 77 (1), 87~103.
- [8] Langel M. , Tillé Y. , 2013, *Variance Estimation of the Gini Index: Revisiting A Result Several Times Published* [J], *Journal of the Royal Statistical Society*, 176 (2), 521~540.
- [9] Lerman R. I. , Yitzhaki S. , 1985, *Income Inequality Effects by Income Source: A New Approach and Applications to the United States* [J], *Review of Economics & Statistics*, 67 (1), 151~156.
- [10] Mills J. , Zandvakili S. , 1997, *Statistical Inference Via Bootstrapping for Measures of Inequality* [J], *Journal of Applied Econometrics*, 12 (2), 133~150.
- [11] Nair U. S. , 1936, *The Standard Error of Gini' s Mean Difference* [J], *Biometrika*, 28 (3/4), 428~436.
- [12] Quenouille M. H. , 1949, *Approximate Tests of Correlation in Time-Series* [J], *Journal of the Royal Statistical Society*, 11 (1), 68~84.

- [13] Rao C. R. , 1973, *Linear Statistical Inference and Its Applications* (Second Edition) [M], New York: John Willey & Sons, Inc.
- [14] Schechtman E. , Yitzhaki S. , 1987, *A Measure of Association Based on Gini's Mean Difference* [J], *Communications in Statistics*, 16 (1), 207~231.
- [15] Suits D. B. , 1977, *Measurement of Tax Progressivity* [J], *American Economic Review*, 67 (4), 747~752.
- [16] Xu K. , 2000, *Inference for Generalized Gini Indices Using the Iterated-Bootstrap Method* [J], *Journal of Business & Economic Statistics*, 18 (2), 223~227.
- [17] Yitzhaki S. , 1991, *Calculating Jackknife Variance Estimators for Parameters of the Gini Method* [J], *Journal of Business & Economic Statistics*, 9 (2), 235~239.
- [18] 戴平生:《拓展基尼系数及其应用的拓展研究》[J],《统计研究》2013年第9期。
- [19] 戴平生:《税收累进性测度的改进:方法、比较和应用》[J],《数量经济技术经济研究》2014年第2期。
- [20] 李香菊、祝丹枫:《中国环境税负空间累进性测度及其影响效应分析》[J],《人文地理》2016年第3期。

Variance Estimation of Tax Progressivity M Index and Its Application

Dai Pingsheng

(School of Economics, Xiamen University)

Research Objectives: To infer the asymptotic normal distribution nature of the tax progressive M index, and to seek the variance estimation of tax progressive indexes. **Research Methods:** This paper sets up an asymptotic normal distribution for tax progressivity M index by U-statistics, and abstracts auxiliary variable from the U-statistics to estimate variance of M index. **Research Findings:** Several used tax progressive indices can be expressed by a function of U-statistics, so that asymptotic normal distribution properties and variance estimations are obtained. **Research Innovations:** The asymptotic distribution nature and variance estimation of the tax progressive M index are given, and the variance estimation methods of the tax progressive K index and S index are also simplified. **Research Value:** To evaluate the government's fiscal and taxation policies. The variance estimation of tax progressivity M index is used to evaluate China's tax policies after tax-sharing reform in 1994. It is found that tax-sharing reform, agriculture tax exemption, business tax to VAT and the threshold adjusting of individual income tax can all improve provincial-level income inequality.

Key Words: U-statistics; Tax Progressivity M Index; Variance Estimation; Evaluation of Tax Policies

JEL Classification: C13; E63; H20

(责任编辑:王喜峰)