

中国收入差距的变动及其原因分析： 1985～2012 年

张 涛

（北京大学经济学院）

【摘要】本文对基于最大熵密度的 GMM 方法予以改进，利用收入分组数据考查中国收入差距在 1985～2012 年的变动趋势，并使用一种新的基于回归方程的分解方法对收入差距及其上升的原因进行分析。实证结果表明，中国总体收入差距在 1985～2008 年迅速扩大，但随后有所下降，隐约呈倒“U”形。占总体 50% 的中等和上中等收入群体一直稳定获得总收入中略高于 50% 的份额，符合帕尔玛“50/50 法则”。城乡内部差距和城乡之间差距均呈上升趋势，而后者增长更快，对总体差距的贡献比重上升。城乡二元分割始终是全国收入差距的首要来源，而家庭规模是全国收入差距不断上升的最重要解释因素。

关键词 广义矩估计 回归分析 收入差距 50/50 法则

中图分类号 F224.0 **文献标识码** A

DOI:10.13653/j.cnki.jqte.2016.12.001

引 言

对各国发展水平的比较和民众幸福程度的关注往往离不开对收入分配问题的考查。收入分配不均通常意味着社会中的一部分人无法选择更有价值的生活方式，只能享有远低于其他人的福利水平。在中国，收入差距扩大是造成人们对改革抱有质疑态度的重要原因，因为人们在看到“蛋糕”不断做大的同时，却发现其分配越来越不均等。

本文主要对 1985～2012 年我国收入分配的均等状况进行估算和解释。对中国收入分配问题进行研究的最大挑战是数据的缺乏。由于时间跨度最长、最具代表性的国家统计局个体数据无法获取，相关研究受到极大限制，大多数学者只能依赖于其他途径获取的数据集中于对少数年份的考查，或者分别对城镇或农村进行研究，但也有部分学者对构建中国总体收入差距指数的连续序列进行了尝试。本文主要对后者中最具代表性的方法进行改进，并结合相关信息提高估计的准确度。在此基础上，本文计算了农村、城镇和全国的四种收入差距指数，对广义熵指数进行了分解，并使用基于回归方程的分解方法探讨我国收入差距及其上升的主要决定因素。此外，本文将尝试澄清关于库兹涅茨（Kuznets）曲线的误解，并验证帕尔玛（Palma）“50/50 法则”在中国的适用性^①。

本文的主要贡献有以下四点：

^① 帕尔玛“50/50 法则”是指：一国收入水平处于第 5 到第 9 个十分位的半数人口（中等和上中等收入组）总是趋向于占有 50%（或略高）的全国总收入。详见 Palma（2011、2014）以及联合国开发计划署（2014）。

第一, 根据统计局城乡收入分组数据各自的特点, 本文对 Wu 和 Perloff (2005、2007) 提出的最大熵原理 GMM 方法予以改进, 免除反复抽样带来的沉重负担。对农村地区, 本文使用各收入组家庭比例的对角矩阵作为权重矩阵; 对城镇地区, 本文先计算收入分组数据基尼系数的上下限, 然后根据“经验法则”得到基尼系数的估计值和分组数据区间端点的估计值, 进而对收入分布函数的基准形式进行调整, 不仅避免了反复抽样, 而且得到的估计结果更为准确。

本文的计算结果表明, 1985 年以来全国收入分配状况不断恶化, 基尼系数从 1985 年的 0.321 上升到 2008 年的 0.468, 上升 46%。2009 年以后收入分配状况略有改善, 2012 年基尼系数降为 0.438, 与 2002 年水平相当。农村的收入差距也一直在上升, 基尼系数在 2011 年达到最大值 (0.385)。城镇的收入差距上升更快, 但在 2008 年达到最大值 (0.338) 后有所下降。对平均对数离差和泰尔指数的分解都表明, 组内差距 (城乡内部差距) 对总体差距的贡献比重在下降, 组间差距 (城乡之间差距) 的贡献比重在上升, 但这两种差距的绝对值都一直在上升, 共同推动了全国总体收入差距的上升。

第二, 本文使用一种新的基于回归方程的方法对收入差距及其变化进行分解, 这一方法由 Fields (2003) 发展并应用于对美国工资差异的研究。与其他一些回归分解方法相比, Fields 方法的最大优点在于其分解得到的各解释因素对收入差距的贡献比重不依赖于所选用的收入差距指数。本文的实证结果表明, 城乡二元分割仍是我国总体收入差距的首要贡献因素; 家庭的产业部门特征是早期收入差距的第二位影响因素, 而家庭规模大小是后期收入差距的第二位影响因素。此外, 对收入差距上升的分解表明, 家庭规模是最重要的解释因素, 其次是地理位置; 教育的影响越来越大, 成为收入差距上升的第三位解释因素。

第三, 本文将验证帕尔玛 (Palma) “50/50 法则” 在中国的适用性, 并在收入差距的度量中使用帕尔玛比值 (Palma Ratio) 这一新的指标^①。本文的计算结果显示, 帕尔玛 “50/50 法则” 在中国有相当的适用性: 第 5 到第 9 (D5~D9) 十分位的 50% 总人口占有全国总收入的份额一直维持在 56.6% 左右, 农村和城镇这一份额的均值相应为 53.9%、54.8%, 说明近 30 年以来, 城乡或全国的中等收入层和上中等收入层始终能稳定地获得整体收入略高于 50% 的份额。

第四, 本文将探讨我国收入差距变化中隐约出现的倒 “U” 形和库兹涅茨假说的关系。我们将通过补充计算来阐明库兹涅茨的结论, 并指出库兹涅茨假说的成立并不依赖于农业部门收入不平等程度不高于非农业部门这一前提假设。库兹涅茨假说带给我们的最大启示是: 不同的收入分配政策可以极大影响收入分配结果, 收入分配与经济发展阶段并无必然联系。

一、相关文献

1. 为什么要研究收入差距: 收入差距的影响

对收入差距问题的兴趣首先令人想到收入差距的影响这一主题。在所有的社会中, 人与人幸福程度的差别都有许多种不同体现, 如受教育水平、健康程度、预期寿命、社会地位等, 这其中许多方面比收入差距更重要。但我们为什么更加关注收入差距? 一个简单的回答是, 收入差距与其他方面的差异高度相关, 很大程度上决定了其他方面的差异程度, 因此可

^① 帕尔玛比值 (Palma Ratio) 是指收入总体中最高 10% 人口总收入与最低 40% 人口总收入的比值, 是帕尔玛 “50/50 法则” 的推论。

以视为社会总体差异状况的指标。例如，阿玛蒂亚·森和让·德雷兹（2006）就认为，低收入（或贫困）意味着个人可选择性的减少和社会机会的缺乏，代表着“能力的剥夺”，经济上的较大差距代表着社会总体的失败。世界银行（2009）认为，个人或家庭的收入决定其享有教育和医疗服务的质量，而受教育程度和健康状况又决定收入的前景，从而增加了形成“不平等陷阱”的可能性。Whyte（2014）也表达了类似的观点。

收入差距会通过不完全信贷市场、不完全要素市场或政治经济学因素影响经济增长（Benjamin 等，2011）：一方面，社会收入或财富分配越不均等，经济的增长率就越低（Alesina 和 Rodrik，1994；Persson 和 Tabellini，1994）；另一方面，较低的不平等会使一国或地区受益（如韩国、中国台湾），因为它减少了再分配带来的效率损失（阿玛蒂亚·森和让·德雷兹，2006）^①。

2. 中国的收入差距：改革前与改革后

与改革后相比，改革前的数据更为匮乏。珀金斯等（1992）和 Benjamin 等（2008）证实：改革前农村地区内部差异较小，并主要反映家庭抚养比特征，地区之间由于禀赋、自然条件差异和自给自足政策的影响存在一定差距，但自 1953 年到 1978 年，农村总体收入差距变化不大；城镇地区由于强烈的平均主义，收入差距维持在更低水平。因此，全国总体收入差距基本由城乡之间的差距所决定^②，所以使用城乡人均收入（或消费）比可以较好反映中国总体收入差距状况，李成瑞（1986）、谢伏瞻（1994）、国务院研究室课题组（1997）、赵人伟和李实（1997）都证实了这一结论。

虽然对收入差距到底有多大存在广泛争议，但所有学者对改革后中国收入差距的迅速上升已达成共识。中国总体差距的上升主要是由于城乡内部收入差距迅速扩大的同时，城乡之间的差距也在迅速扩大（Wu 和 Perloff，2005；程永宏，2006、2007；Ravallion 和 Chen，2007；世界银行，2009；王亚峰，2012；Li 和 Sicular，2014）。由于统计局的数据存在一定缺陷（Yang，1999；Bramall，2001；Li 和 Sicular，2014），许多学者尝试对基尼系数的估计值进行修正（Khan 和 Riskin，2008；王小鲁，2010；李实和罗楚亮，2011）。还有很多学者对中国收入差距上升的原因进行了研究（万广华，2004、2013；许庆等，2008；万海远和李实，2013）。这些文献大多发现：地理位置仍是收入差距的最重要解释因素；农村非农就业机会的不均等使农村收入差距状况恶化；教育对收入差距有越来越强的解释力。此外，国家统计局不断对城乡进行重新划分加大了收入差距的上升幅度（Benjamin 等，2008）。

3. 最高收入份额之辨：帕尔玛比值和库兹涅茨曲线

随着工业化进程的推进，一国的收入分配状况在长期会如何变化？Kuznets（1955、1963）对英、美、德、荷等国的考查部分回答了这一问题，他的部分结论被后来的学者称为库兹涅茨倒“U”曲线假说。虽然有研究者验证了这一假说（Ahluwalia 等，1979；Anand 和 Kanbur，1993），但近来经济学家使用同样的方法对上述一些国家收入分配的后续情况进行考查发现：库兹涅茨曲线消失了。从 20 世纪早期到第二次世界大战后的数 10 年，最高收入（或财富）组占总体收入的份额确实出现大幅下降，但近年来最高收入组的收入份额已经

^① 对收入差距相关影响的更多了解可参考王弟海（2009）和 Benjamin 等（2011）及其中引用的文献。

^② 这种解释可以看作广义熵指数的分解：组内差距不变的情况下，总体差距由组间差距决定。但组内差距不变还依赖于城乡人口流动受限等条件。

开始重新上升 (Piketty 和 Seaz, 2003; Atkinson 和 Piketty, 2007; Kopczuk 等, 2010)。本文的计算结果表明, 我国的收入分配状况近来隐约出现倒“U”形趋势, 这到底符合库兹涅茨假说吗? 我们将在下文进行讨论。

此外, Palma (2011、2014) 通过对全球 130 多个国家和地区横截面数据的分析得出“50/50 法则”: 收入总体中第 5 到第 9 十分位 (D5~D9) 的 50% 总人口占有的总收入份额始终维持在 50% 或略高。另一半人口, 即最高 10% (D10) 收入组和最低 40% (D1~D4) 收入组只能共同占有剩下的 50% 或略低的收入份额。这一法则意味着最常用来度量收入差距的基尼系数的有用性令人质疑, 因为基尼系数通常被认为最能反映收入分配中部的变化情况, 但按照帕尔玛法则, 这一部分的变化似乎是最难发生的^①。相应的, 根据帕尔玛法则, 可以使用 D10 与 D1~D4 的收入比衡量总体的不平等状况, 联合国开发计划署 (2014) 已采用这一指标度量各国的不平等^②。本文也将这一指标纳入收入差距的参考指标体系, 并检验帕尔玛“50/50 法则”在中国的适用性。

4. 对收入差距的估计: 参数法和非参数法

本文也和对收入分布的估计方法直接相关。根据可获取数据的类型, 文献中的估计方法大致归为三类 (Cowell 和 Flachaire, 2015): 参数法、非参数法和半参数法。Bourguignon 和 Morrison (2002)、SalaiMartin (2006) 使用了非参数法中的核密度技术对世界的总体不平等进行研究。Wu 和 Perloff (2005、2007)、程永宏 (2006、2007)、王亚峰 (2012) 和 Hajargasht 等 (2012) 使用参数法直接估计收入分布函数, 其估计结果的优劣很大程度上取决于所选用收入分布函数的形式^③。由于传统上各国收入统计中较多使用分组数据, 因此也有学者尝试使用洛伦兹曲线或其他数学原理对基尼系数的上下界进行估计 (Gastwirth, 1972; Gastwirth 和 Glauber, 1976)。

二、数据和方法

1. 数据介绍

本文使用的是国家统计局 1986~2013 年《中国统计年鉴》中分城乡的家庭收入分组数据和 CHIP (中国家庭收入调查) 1988 年、2002 年两年度的数据^④。国家统计局数据存在许多令人广为诟病的缺陷: 收入定义狭隘、地理覆盖率存在严重偏差、对生活成本的空间差异未作调整和对最低最高收入组抽样严重不足等。但统计局的数据仍是最适合对中国收入分配进行研究的数据, 因为与其他调查数据相比, 统计局数据覆盖范围最广、样本量最大因而最具全国代表性, 而且其时间跨度和数据一致性也是其他数据无法比拟的^⑤。

① 对各种不平等指数优缺点的讨论可参见 Ahmad 和 Wang (1991)、Krol 和 Miedema (2009)。

② 另一个自然而然的推论是, 可以用最高 10% 收入组的收入份额作为总体不平等的度量指标, 因为按照帕尔玛“50/50 法则”, 帕尔玛比值 (Palma Ratio) = $D10 / (D1 \sim D4) = D10 / (50\% - D10)$, 该函数是 D10 的单调增函数。

③ Wu 和 Perloff (2005、2007)、Park 和 Bera (2009)、Hajargasht 等 (2012)、Cowell 和 Flachaire (2015) 等对收入分布函数的形式进行了讨论, 本文进行了综合考虑。

④ 出于数据前后一致性考虑, 本文没有使用《中国统计年鉴》1982~1985 年和 2014 年、2015 年; 由于国家统计局数据暂未获准公开, CHIP 的 2007 年度数据暂不可用。

⑤ 例如, 学者比较常用的其他数据中: 中国家庭收入调查 (CHIP) 数据作为国家统计局数据的子集, 覆盖省份更少, 而且每年 (目前仅覆盖 1988 年、1995 年、2002 年、2007 年、2013 年五年) 有所变化, 样本量仅为统计局的 1/10~1/5; 中国健康与营养调查 (CHNS) 只覆盖了 8 个省 (1997 加入黑龙江, 2011 年后加入北京、重庆、上海), 样本量更少, 仅为统计局的 1/30; 中国家庭追踪调查 (CFPS) 的调查从 2010 年开始, 仅有 3 年的数据, 样本量也仅为统计局的 1/10。

中国家庭收入调查 (CHIP) 数据是国家统计局数据的子集, 与统计年鉴数据最为接近、最具可比性。最为重要的是, CHIP 包含个体家庭收入数据, 为检验收入分布函数的具体形式提供了重要参考依据, 也为确定真实基尼系数在分组数据基尼系数上下界中的具体位置提供了“经验法则”^①。我们也将 CHIP 的城镇数据按照统计局的方式分组并获取各组端点与前后区间均值的比例关系, 以此作为计算统计局分组数据端点的“经验法则”。此外, CHIP 的个人和家庭数据使本文对收入差距及其上升的原因进行实证分析成为可能。

2. 相关方法

本文收入分布函数的估计方法是对 Wu 和 Perloff (2005、2007) 方法的改进。本文首先基于城乡收入分组数据估计各自的收入分布函数, 然后按城乡人口进行加权获得全国总体的收入分布函数, 并分别计算相关的收入差距指数。随后, 本文将平均对数离差 (MLD, 对应广义熵指数参数为 0 的情形) 和泰尔指数 (Theil Index, 对应广义熵指数参数为 1 的情形) 分解为组内差距与组间差距, 并分析二者的变化趋势和对总体差距的贡献情况^②。最后, 本文使用 Fields (2003) 发展起来的基于回归方程的分解方法对收入差距及其上升的原因进行分析, 该方法尚未应用于中国收入差距的研究^③。

(1) 对分组数据最大熵密度的 GMM 估计。根据 Wu (2003)、Wu 和 Perloff (2005、2007), Jaynes 提出的最大熵原理是指, 对某一随机事件 (或分布) 的概率分布进行赋值 (或选择) 时, 应在与已知信息 (约束) 一致的情况下, 最大化 Shannon 信息熵 (目标函数)^④。具体而言, 分布函数 $f(x)$ 的 Shannon 信息熵表达式为:

$$-\int f(x) \cdot \ln[f(x)] dx \quad (1)$$

满足的约束条件为:

$$\int f(x) dx = 1; \int g_i(x) \cdot f(x) dx = \mu_i \quad i = 1, 2, \dots, N \quad (2)$$

该最大化问题的唯一解为:

$$f(x) = \exp\{-\lambda_0 - \sum_{i=1}^N \lambda_i \cdot g_i(x)\} \quad (3)$$

其中, $g_i(x)$ 为矩函数, λ_i 为拉格朗日乘子。

设收入数据分为 K 组, 且已知分组数据的端点为 $(x_0, x_1, \dots, x_K)$, 各组家庭所占比例为 $(s_1, s_2, \dots, s_K)$, 各组收入的条件均值为 $(m_1, m_2, \dots, m_K)$ 。根据 $f(x)$ 可计算每一区间的对应总体矩, 其中各区间家庭比重为 $\int_{x_{k-1}}^{x_k} f(x) dx$, 各区间条件均值为 $\int_{x_{k-1}}^{x_k} x \cdot f(x) dx$, $k = 1, 2, \dots, K$ 。由此可构造矩条件:

$$h_k = \begin{bmatrix} \int_{x_{k-1}}^{x_k} f(x) dx - s_k \\ \int_{x_{k-1}}^{x_k} x \cdot f(x) dx - m_k \end{bmatrix} \quad k = 1, 2, \dots, K \quad (4)$$

① 这里的“经验法则”是指基尼系数真实值在上下界间的具体位置。将 CHIP 数据按统计局城镇收入分组方法进行分组, 本文发现真实基尼系数距离下界和上界的距离比约为 2:3。

② 关于广义熵指数的分解方法和原理, 可参见 Shorrocks (1980)、Cowell 和 Flachaire (2015)。

③ Cai 等 (2010) 在对中国城镇收入差距的研究中部分使用了这一方法。

④ 这一原理听起来类似于极大似然估计, 事实正是如此, 二者在一定条件下存在对偶关系, 见 Wu (2003)。

GMM 的目标函数为：

$$Q(\lambda) = H(\lambda)'WH(\lambda) \quad (5)$$

其中, $H(\lambda) = [h_1, h_2, \dots, h_K]'$, W 为权重矩阵。

使用这一方法的关键在于对未知矩函数 $g_i(x)$ 的选择和对权重矩阵 W 的计算, 此外, 城镇分组数据的端点未知使估计变得复杂。本文对此作如下处理：

第一, 根据相关文献的建议, 本文选取一些对收入分布拟合效果较好的函数形式, 并用 CHIP 数据进行验证和筛选^①。本文最终选择的 $g_i(x)$ 基准形式有: $\ln(x)$, $\ln(1+x^2)$, $\ln(1+x^4)$, $\text{atan}(x)$, 这些矩函数具有第二类广义贝塔函数的优点, 但形式比后者更加灵活^②。

第二, 针对城镇分组数据端点未知的特点, 本文使用 CHIP 数据确定区间端点与前后区间均值的关系、真实基尼系数与分组数据基尼系数上下界的关系。然后将这些结论作为“经验法则”应用于统计局数据, 并对区间端点值和分布函数基准形式进行调整以提高估计准确度^③。

第三, 关于权重矩阵的计算, 本文发现: 对农村地区而言, 权重矩阵不依赖于个体(家庭)数据, 仅取决于各区间家庭比例, 所以可以直接使用各区间家庭比例的对角矩阵作为权重矩阵, 无须根据第一阶段的估计结果进行反复抽样; 对城镇地区, 利用上面提到的“经验法则”, 也避免了反复抽样, 大大减轻了计算负担^④。

(2) 对收入差距及其变化的回归分析。除了将平均对数离差和泰尔指数进行组内、组间分解, 本文还使用一种新的基于回归方程的分解方法 (Fields, 2003) 对收入差距及其上升进行分解^⑤。同 Morduch 和 Sicular (2002)、万广华 (2004) 和 Yue 等 (2008) 使用的回归方法相比, 本文方法的优点是各解释变量对收入不平等的贡献比重对所有收入差距指数而言都是相同的。此外, 该方法可以对收入不平等的上升进行解释, 并可以将解释因素的贡献分解为“价格效应”和“数量效应”^⑥。在本文的分析中, 除了使用性别、教育、经验及其平方、地理位置、产业部门和职业属性等传统解释变量外, 还包含了更符合中国特点的解释因素, 如城镇虚拟变量、家庭规模、党员身份等, 实证结果表明这些解释变量对中国的收入差距及其上升具有显著影响: 城乡二元分割是全国收入差距的首要贡献因素, 而家庭规模是全国收入差距上升的最重要原因。

三、估计结果：城乡与总体收入差距变动趋势

1. 收入分布函数和收入分配的变化

在估计出某一年份农村和城镇的收入分布函数以后, 本文对其进行标准化, 使其面积等于城乡当年各自的总人口数, 条件均值等于根据 CPI 调整后的家庭人均收入的总体均值^⑦。

① 本文对函数形式的验证主要是将估计得到的分布函数与 CHIP 数据的核密度估计结果进行图形比较, 同时使用 Kolmogorov-Smirnov (K-S) 检验。

② 对函数形式的建议可参考 Wu 和 Perloff (2005、2007)、Park 和 Bera (2009)、Hajargasht 等 (2012)、Cowell 和 Flachaire (2015) 等。

③ 就精确度而言, 按照本文“经验法则”获得的城镇基尼系数估计值与王亚峰 (2012) 采用另一方法获得的估计结果极其接近。可对照本文的结果和王亚峰 (2012) 的研究。

④ 本文发现这种做法比 Wu 和 Perloff (2007) 提出的蒙特卡洛反复抽样法更为理想。

⑤ 该方法的具体介绍和应用参见 Fields (2003)。

⑥ 大体上说, “价格效应”是指回归系数变化导致的收入差距变化, “数量效应”是指解释变量变化导致的收入差距变化。

⑦ 如非特别指出, 本文中的计算结果都以 2002 年为基期。

全国总体的收入分布可通过对城乡收入分布函数进行加权得到：设农村和城镇人口比重分别为 s_r, s_u ，收入分布函数分别为 $f_r(x), f_u(x)$ ，则总体收入分布函数为

$$f(x) = s_r \cdot f_r(x) + s_u \cdot f_u(x) \quad (6)$$

图 1、图 2 分别报告了 1985、1995、2005、2012 各年农村和城镇各自的收入分布状况。总的来看，随着时间的推移，城乡各自收入分布的图形一直在向右移动，特别是 1995 年以后，这表明农村和城镇的整体收入水平都有大幅提高。农村收入分布的众数在 1985 年仅为 885 元，到 2012 年增加为 3300 元；城镇收入的众数则由 1985 年的 2400 元增加到 2012 年的 13560 元。此外，城乡各自分布图形由较为狭窄逐渐变宽，表明城乡各自的收入差距在这一时期大为上升，下文将根据收入差距指数对这一问题具体分析。

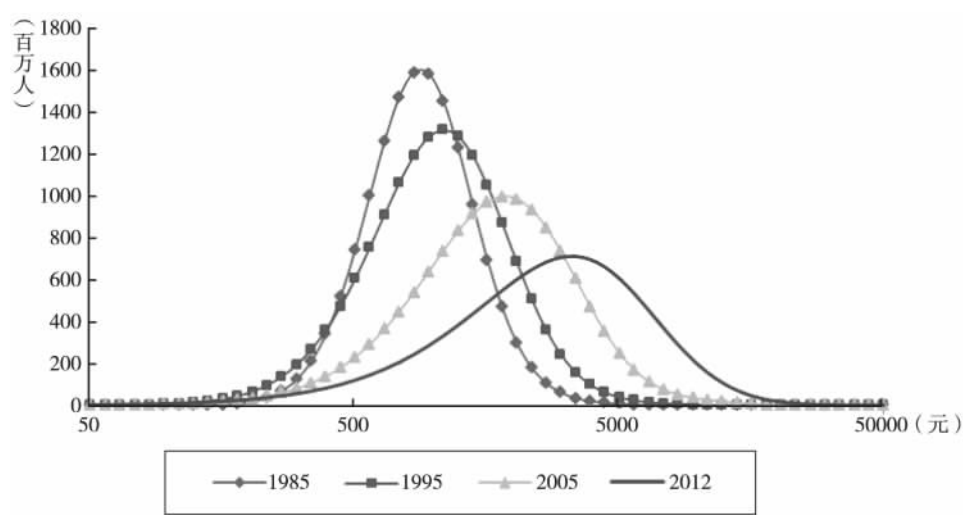


图 1 农村收入分布的变化

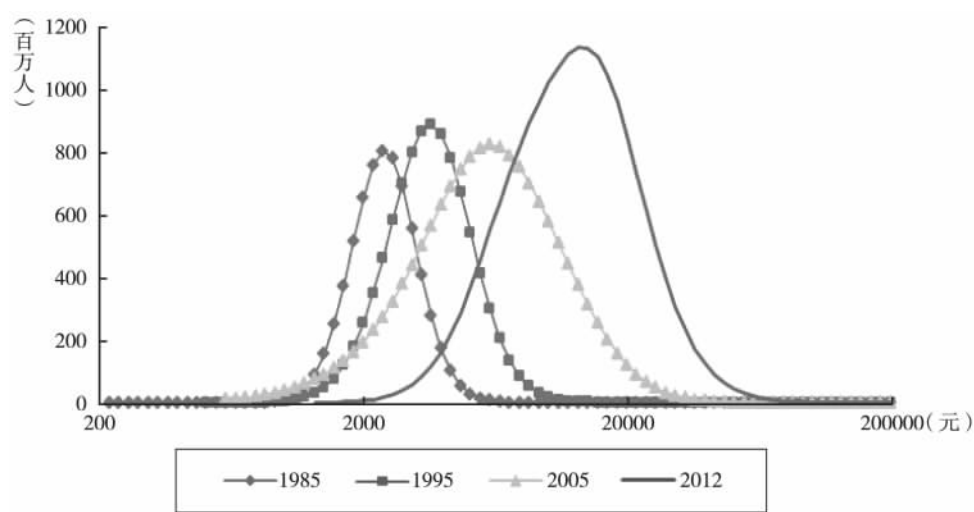


图 2 城镇收入分布的变化

图3报告了全国1985年、1995年、2005年、2012年的总体收入分布变化情况。为了考查收入差距下降的影响,该图也给出了2008年总体收入差距达到最大值时的全国收入分布。同城乡各自的收入变动一致,全国总体收入分布从1985年到2012年也一直在向右移动,表明全国居民总体收入水平的提高,总体收入的众数由1985年的885元上升到2012年的3610元。2012年收入分布图形右侧高收入人群增加的面积远大于1985年左侧低收入人群减小的面积,这从另一方面反映了全国居民福利水平的提高。但是,总体分布图形逐渐变宽,特别是收入(相对)较低人口比重的提高(左侧面积增大),表明全国的收入差距在这一时期恶化了。

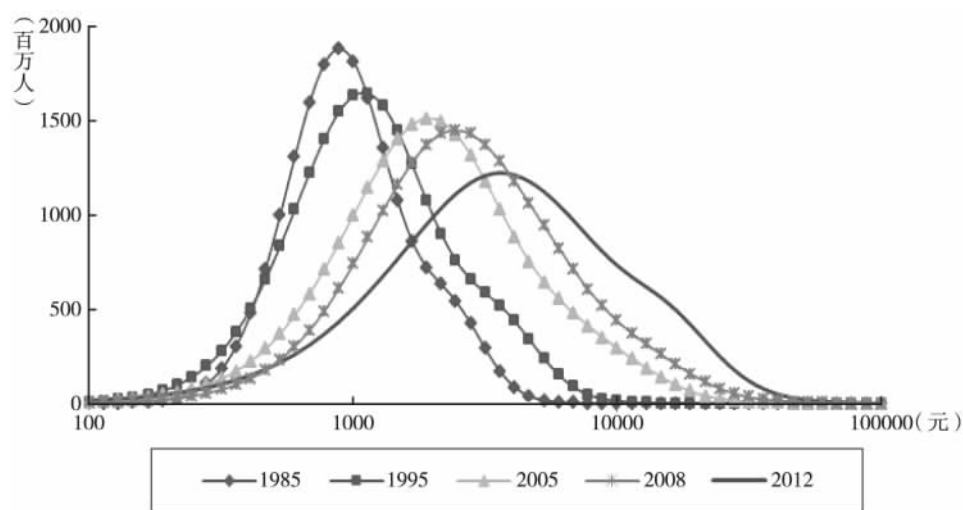


图3 全国收入分布的变化

2. 收入差距的变动趋势

(1) 城乡收入差距。对收入分布函数进行估计之后,本文按照惯例计算了3种常用的收入差距指数:基尼系数、泰尔指数、平均对数离差^①。此外,本文也计算了帕尔玛比值这一最新的收入差距指数。

表1报告了1985~2012年农村、城镇的上述4种收入差距指数。表数据的第一、五列为基尼系数(参见图4),根据这一指数,农村地区的收入差距一直在上升,基尼系数从1985年的0.278上升到2011年的0.385,上升了38%,并在2012年稳定在0.385左右。城镇地区的收入差距上升速度更快,基尼系数从1985年的0.169上升到2008年的0.338,上升100%。2008年后城镇收入分配状况有所改善,但2012年基尼系数仍高达0.313。表数据第二、六列的帕尔玛比值(也可参见图5)和其他列的收入差距指数证实了城乡收入差距类似的变动趋势。

^① 常用收入差距指数还有对数收入方差、最高10%与最低10%人口收入比(10/10比值)、大岛指数(收入中最高20%人口与最低20%人口的收入比,20/20比值)和阿特金森指数(Atkinson, 1970)等。但阿特金森指数与广义熵指数存在单调变换关系(Cowell和Flachaire, 2015),10/10比值、大岛指数与帕尔玛比值存在一定的替代性。为了与传统一致,我们也计算了这些指标,有兴趣的读者可向作者索取。

表 1 农村和城镇的收入差距指数

年 份	农村收入差距指数				城镇收入差距指数			
	基尼系数	帕尔玛比值	泰尔指数	平均对数离差	基尼系数	帕尔玛比值	泰尔指数	平均对数离差
1985	0.278	0.992	0.132	0.126	0.169	0.575	0.049	0.047
1986	0.298	1.092	0.153	0.148	0.171	0.583	0.057	0.051
1987	0.301	1.119	0.158	0.148	0.172	0.590	0.058	0.051
1988	0.321	1.238	0.184	0.171	0.181	0.614	0.063	0.056
1989	0.326	1.271	0.186	0.176	0.187	0.634	0.069	0.060
1990	0.315	1.200	0.175	0.164	0.183	0.620	0.064	0.058
1991	0.331	1.310	0.206	0.184	0.173	0.595	0.059	0.052
1992	0.335	1.349	0.209	0.187	0.191	0.645	0.067	0.062
1993	0.338	1.360	0.204	0.188	0.212	0.712	0.080	0.074
1994	0.336	1.332	0.196	0.186	0.223	0.749	0.086	0.082
1995	0.345	1.392	0.209	0.199	0.216	0.724	0.080	0.076
1996	0.326	1.256	0.179	0.176	0.216	0.724	0.080	0.076
1997	0.335	1.321	0.197	0.190	0.227	0.762	0.094	0.086
1998	0.333	1.309	0.195	0.185	0.234	0.789	0.097	0.092
1999	0.344	1.390	0.212	0.198	0.242	0.822	0.101	0.097
2000	0.357	1.487	0.223	0.217	0.253	0.864	0.109	0.107
2001	0.365	1.549	0.233	0.225	0.265	0.917	0.119	0.117
2002	0.369	1.586	0.239	0.230	0.317	1.188	0.170	0.171
2003	0.379	1.667	0.250	0.244	0.326	1.242	0.179	0.179
2004	0.367	1.562	0.233	0.229	0.333	1.291	0.188	0.188
2005	0.375	1.629	0.242	0.239	0.338	1.324	0.193	0.194
2006	0.373	1.606	0.243	0.238	0.335	1.307	0.190	0.189
2007	0.372	1.598	0.241	0.238	0.332	1.282	0.186	0.185
2008	0.376	1.637	0.274	0.247	0.338	1.331	0.194	0.193
2009	0.383	1.693	0.268	0.255	0.333	1.296	0.188	0.187
2010	0.377	1.633	0.250	0.248	0.327	1.259	0.183	0.181
2011	0.385	1.710	0.261	0.265	0.327	1.259	0.183	0.178
2012	0.385	1.696	0.256	0.263	0.313	1.174	0.167	0.164
1985~2008 年变化	35%	65%	108%	95%	100%	132%	294%	313%
2008~2012 年变化	2%	4%	-6%	7%	-8%	-12%	-14%	-15%
1985~2012 年变化	38%	71%	94%	109%	85%	104%	239%	250%

注：2002 年城乡划分有重大调整，因此 2002 年以后的数据并不严格可比。

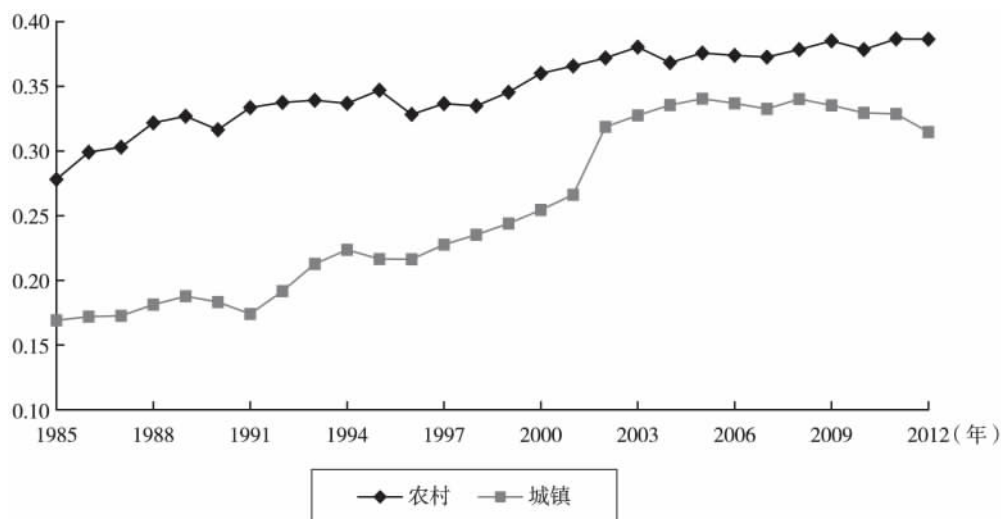


图4 城乡基尼系数的变动趋势：1985~2012年

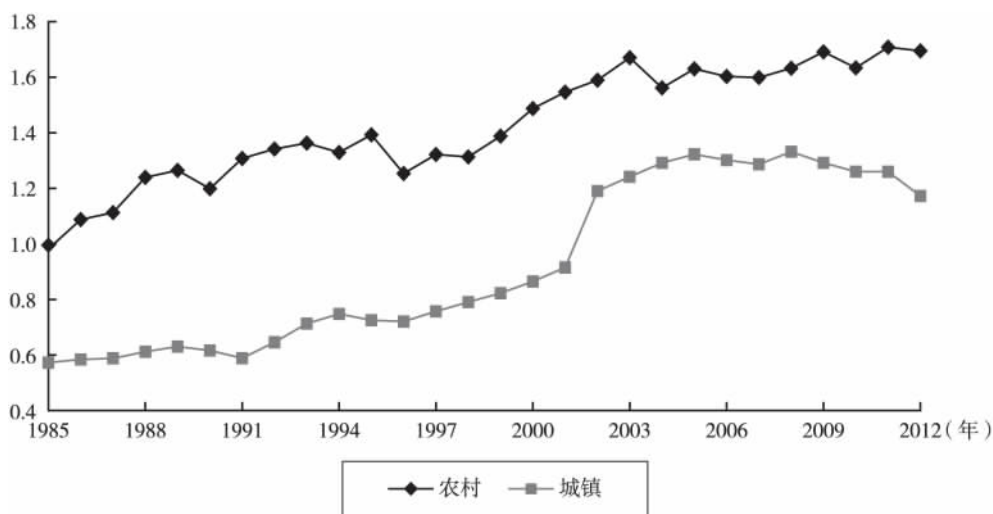


图5 城乡帕尔玛比值的变动趋势：1985~2012年

从图4和图5可以看出，城镇的收入差距上升速度更快，并越来越接近农村地区。图4（和图5）还直观地表明，城乡各自收入差距的上升并不是单调的。例如，农村基尼系数（以及帕尔玛比值）在1996年和2004年都有所下降。前者的下降显然与当时农产品价格的大幅上涨有关，而2004年收入差距的下降除了受农业生产价格大幅上升的影响外，还可能与农业税减征或免征相关。城镇地区的基尼系数（以及帕尔玛比值）在2008年后出现明显的下降，这可能与城镇最低工资的提高、社会保障项目的扩大、金融危机等原因有关^①。

（2）全国收入差距。表2报告了1985~2012年中国总体收入差距指数和部分收入组所占总体收入份额的估计结果。根据表2数据第一、二列和图6，基尼系数和帕尔玛比值的变化情况极其类似：基尼系数从1985年的0.321上升到2008年的最大值0.468，上升46%，

^① 这些说法有赖于严格的实证分析。例如，Li和Sicular（2014）认为，农业税的取消对不平等的影响作用很小，对减贫有积极作用；万广华（2013）认为，经济周期与收入差距存在一定的反向相关关系。

帕尔玛比值则从 1.175 (1985 年) 上升到 2.710 (2008 年), 上升超过 130%; 随后二者逐渐下降, 基尼系数在 2012 年降为 0.438, 帕尔玛比值降为 2.321, 收入差距状况大致与 2002 年相当。表 2 的其他收入差距指数虽然上升幅度不同, 但都表明了同样的趋势: 从 1985 年到 2012 年, 平均对数离差上升幅度最大, 上升 116%; 帕尔玛比值和泰尔指数分别上升 98%、94%; 基尼系数上升幅度最小, 仅为 36%。

表 2 全国收入差距指数和相关收入组收入份额：1985~2012 年

年 份	全国收入差距指数				相关收入组收入份额			
	基尼系数	帕尔玛比值	泰尔指数	平均对数离差	D10	D1~D4	D10 加 D1~D4	D5~D9
1985	0.321	1.175	0.166	0.169	0.233	0.198	0.431	0.569
1986	0.325	1.200	0.174	0.178	0.232	0.193	0.425	0.575
1987	0.327	1.218	0.176	0.178	0.235	0.193	0.428	0.572
1988	0.345	1.336	0.197	0.201	0.245	0.183	0.428	0.572
1989	0.355	1.401	0.208	0.213	0.248	0.177	0.425	0.575
1990	0.346	1.342	0.198	0.202	0.245	0.183	0.427	0.573
1991	0.361	1.453	0.219	0.223	0.253	0.174	0.427	0.573
1992	0.372	1.534	0.229	0.233	0.261	0.170	0.431	0.569
1993	0.386	1.655	0.245	0.251	0.269	0.163	0.432	0.568
1994	0.390	1.691	0.249	0.258	0.271	0.160	0.431	0.569
1995	0.384	1.648	0.241	0.255	0.266	0.161	0.427	0.573
1996	0.357	1.429	0.207	0.218	0.249	0.174	0.423	0.577
1997	0.375	1.576	0.234	0.244	0.262	0.166	0.428	0.572
1998	0.376	1.588	0.234	0.243	0.263	0.166	0.429	0.571
1999	0.390	1.716	0.252	0.264	0.271	0.158	0.430	0.571
2000	0.404	1.844	0.268	0.289	0.276	0.150	0.425	0.575
2001	0.414	1.966	0.283	0.305	0.283	0.144	0.427	0.573
2002	0.441	2.281	0.327	0.341	0.309	0.135	0.444	0.556
2003	0.453	2.457	0.346	0.366	0.316	0.129	0.445	0.555
2004	0.456	2.494	0.352	0.369	0.321	0.129	0.449	0.551
2005	0.460	2.570	0.359	0.380	0.322	0.125	0.447	0.553
2006	0.461	2.594	0.361	0.383	0.323	0.124	0.447	0.553
2007	0.460	2.578	0.357	0.384	0.319	0.124	0.443	0.557
2008	0.468	2.710	0.374	0.397	0.330	0.122	0.451	0.549
2009	0.465	2.696	0.367	0.399	0.322	0.119	0.441	0.559
2010	0.454	2.541	0.349	0.383	0.313	0.123	0.437	0.563
2011	0.450	2.468	0.343	0.380	0.310	0.126	0.436	0.564
2012	0.438	2.321	0.323	0.364	0.299	0.129	0.428	0.572
1985~2008 年变化	46%	131%	125%	135%	42%	-39%	5%	-4%
2008~2012 年变化	-7%	-14%	-14%	-8%	28%	-35%	-1%	1%
1985~2012 年变化	36%	98%	94%	116%	-9%	6%	-5%	4%

注：D1、D10 分别表示收入总体中第 1 个（最低）、第 10 个（最高）10% 人口占有总体收入的份额，其余依此类推。

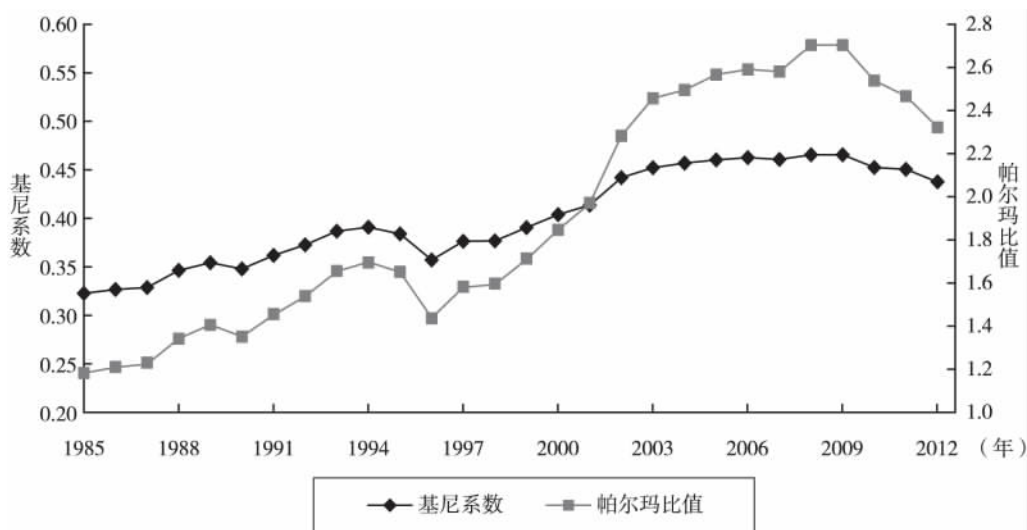


图6 全国收入差距的变化情况：1985~2012年

表2数据和图6同样表明收入差距的上升并非单调的。在1995~1996年，收入差距急剧下降，下文对收入差距指数的分解表明，这一时期组内差距（1995~1996年）和组间差距（1994~1996年）都明显下降了。农产品价格的上涨很可能起到了关键作用：一方面，农业收入的提高使农业收入的均等化效应加强从而减小了农村内部的收入差距；另一方面，农村整体收入的提高也减小了城乡之间的差距。在2008年以后，全国总体收入差距再次出现明显下降并一直持续到2012年。收入差距指数的分解同样表明，这一时期组内差距和组间差距都逐渐下降了，而前文中关于农业生产价格上升、农业税取消、城镇最低工资和金融危机等影响的解释在此同样适用。

3. 帕尔玛“50/50法则”和库兹涅茨假说

(1) 帕尔玛“50/50法则”。本文的估计方法为验证帕尔玛“50/50法则”提供了有利条件。表2的五至八列报告了全国总体收入中最高10%（D10）收入组、最低40%（D1~D4）收入组和其余50%（D5~D9）收入组各自占有总体收入份额的变化情况（也可参见图7）。

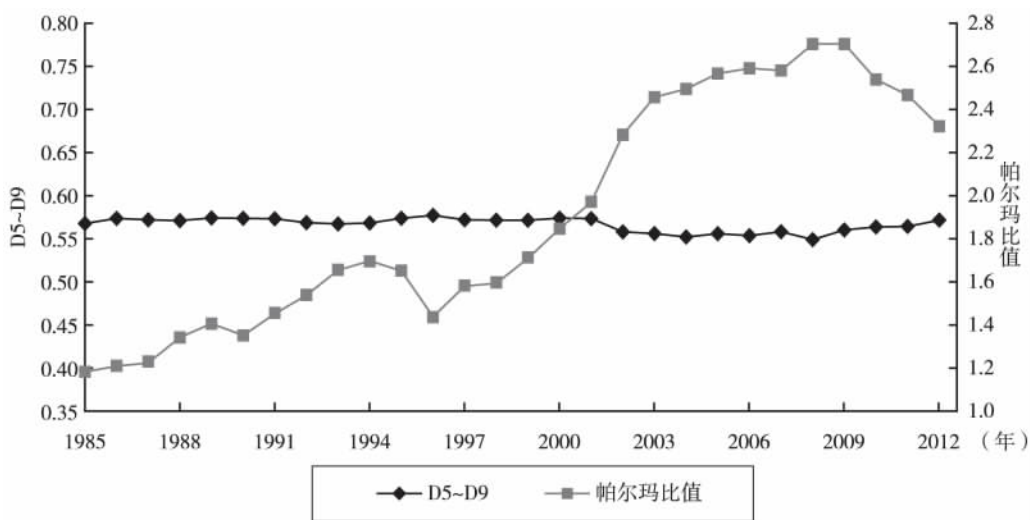


图7 帕尔玛比值与D5~D9人口占总收入的份额：1985~2012年

根据表 2 和图 7，与帕尔玛比值的大幅上升和剧烈波动相比，D5~D9 人口占有的总收入份额始终稳定在 56% 左右，其算术平均值为 56.6%，调和平均值为 56.6%，中位数为 57.0%。这一结论与帕尔玛对 130 多个国家的实证分析结果非常接近（算术平均值为 52.1%，调和平均值为 51.8%，中位数为 52.5%（Palma, 2011）；算术平均值为 51.5%，调和平均值为 51.2%，中位数为 52.0%（Palma, 2014）），说明帕尔玛“50/50 法则”在中国有相当的适用性：中国中层和上中层的 50% 收入群体（D5~D9）倾向于稳定获得总收入中 50% 或略高的份额^①。

（2）消除对库兹涅茨的误解：中国真的符合“Kuznets 曲线”吗？中国的收入差距在经历了长期的上升以后，在 2008 年达到顶点并在 2009 年出现短暂的稳定，随后呈现下降趋势，这种隐约可见的倒“U”形不禁令人回想起库兹涅茨假说。Wang 等（2014）和 Whyte（2014）认为这一假说对中国的解释并不可信：前者认为库兹涅茨的结论强烈依赖于人口从收入不平等程度较低的农业部门向不平等程度较高的非农业部门流动这一前提假设；后者则举出许多不存在倒“U”曲线的例子，如日本、韩国，它们的收入不平等与经济发展的关系更像是“U”形，而非倒“U”形。

但真实的情况到底是怎样的，中国的情形符合库兹涅茨假说吗？本文从以下两方面予以说明：

第一，库兹涅茨曲线的出现并不依赖于农业部门的不平等必须不高于非农业部门这一假设。库兹涅茨在分析中强调的另两条假设是：非农业部门的人均收入始终高于农业部门，（城镇化或工业化进程中）农业部门的人口比重不断下降^②。对于农业部门收入不平等高于非农业部门的例外情形，一方面，Kuznets（1963）给出了一些例子^③，另一方面，我们可以按照库兹涅茨的分析方法补充计算这一情形下的结果^④。这一简单的计算表明，在农业部门收入不平等高于非农业部门的情形下：1）如果两部门的人均收入比保持不变，总体收入不平等变化仍然会出现倒“U”形，但拐点出现在城镇化初期（城镇化率在 10%~30% 之间）；2）如果在城镇化过程中，两部门人均收入比上升（由 1:2 上升到 1:4），则收入差距仍会出现先上升和后下降的变化趋势。

第二，库兹涅茨曲线的出现与否以及是否反转为“U”形与 Kuznets（1955）的结论并不矛盾。Kuznets（1955、1963）的解释和各国的经验清楚地说明了这一结论：总体收入差距除了取决于农业部门和非农业部门的各自不平等状况，还取决于两部门的人均收入之比，而且更为重要的是，随着城镇化的推进和接近完成，越来越取决于非农业部门的收入差距变化情况，而这一切自然与公众对不平等的态度以及政府的收入分配政策密切相关^⑤。

① 城镇（农村）D5~D9 收入组占有总体收入份额的算术平均值为 54.8%（53.9%），调和平均值为 54.8%（53.9%），中位数为 55.0%（53.9%）。

② 参见 Kuznets（1955）。

③ 1948 年左右的意大利和 1950~1953 年的美国，农业部门的收入不平等均高于非农业部门，见 Kuznets（1963）。

④ 计算方法见 Kuznets（1955）。

⑤ Kuznets（1955）指出，城镇中底层人口有越来越大的话语权和影响力、收入税累进性的加强等因素都会促使非农业部门收入差距减小。

四、对收入差距及其上升的原因分析

1. 广义熵指数的分解：组内差距和组间差距

本文首先将总体收入差距分解为组内差距（城乡各自收入差距的加权和，下称城乡内部差距）和组间差距（城乡之间差距），表3给出了平均对数离差和泰尔指数的分解结果^①。

表3 总体收入差距的分解

年 份	平均对数离差					泰尔指数				
	总体 差距	组内 差距	%组内 差距	组间 差距	%组间 差距	总体 差距	组内 差距	%组内 差距	组间 差距	%组间 差距
1985	0.169	0.107	63.6%	0.062	36.4%	0.166	0.099	59.3%	0.068	40.7%
1986	0.178	0.124	69.5%	0.054	30.5%	0.174	0.115	65.9%	0.059	34.1%
1987	0.178	0.124	69.3%	0.055	30.7%	0.176	0.117	66.3%	0.059	33.7%
1988	0.201	0.141	70.3%	0.060	29.7%	0.197	0.133	67.2%	0.065	32.8%
1989	0.213	0.146	68.2%	0.068	31.8%	0.208	0.135	64.7%	0.073	35.3%
1990	0.202	0.136	67.2%	0.066	32.8%	0.198	0.126	63.8%	0.072	36.2%
1991	0.223	0.149	66.7%	0.074	33.3%	0.219	0.139	63.5%	0.080	36.5%
1992	0.233	0.153	65.6%	0.080	34.4%	0.229	0.143	62.3%	0.086	37.7%
1993	0.251	0.156	62.0%	0.095	38.0%	0.245	0.142	58.2%	0.102	41.8%
1994	0.258	0.156	60.5%	0.102	39.5%	0.249	0.140	56.2%	0.109	43.8%
1995	0.255	0.164	64.2%	0.091	35.8%	0.241	0.144	59.8%	0.097	40.2%
1996	0.218	0.145	66.7%	0.073	33.3%	0.207	0.130	63.0%	0.077	37.0%
1997	0.244	0.157	64.2%	0.087	35.8%	0.234	0.143	60.9%	0.091	39.1%
1998	0.243	0.154	63.3%	0.089	36.7%	0.234	0.142	60.6%	0.092	39.4%
1999	0.264	0.163	61.9%	0.100	38.1%	0.252	0.149	59.1%	0.103	40.9%
2000	0.289	0.177	61.5%	0.111	38.5%	0.268	0.155	57.9%	0.113	42.1%
2001	0.305	0.184	60.5%	0.120	39.5%	0.283	0.162	57.3%	0.121	42.7%
2002	0.341	0.207	60.6%	0.134	39.4%	0.327	0.194	59.5%	0.133	40.5%
2003	0.366	0.218	59.6%	0.148	40.4%	0.346	0.203	58.5%	0.144	41.5%
2004	0.369	0.212	57.4%	0.157	42.6%	0.352	0.202	57.2%	0.151	42.8%
2005	0.380	0.220	57.8%	0.160	42.2%	0.359	0.207	57.7%	0.152	42.3%
2006	0.383	0.217	56.6%	0.166	43.4%	0.361	0.205	56.8%	0.156	43.2%
2007	0.384	0.214	55.7%	0.170	44.3%	0.357	0.201	56.1%	0.157	43.9%
2008	0.397	0.221	55.8%	0.175	44.2%	0.374	0.214	57.2%	0.160	42.8%
2009	0.399	0.222	55.6%	0.177	44.4%	0.367	0.207	56.5%	0.160	43.5%
2010	0.383	0.214	56.0%	0.169	44.0%	0.349	0.198	56.8%	0.151	43.2%
2011	0.380	0.221	58.0%	0.159	42.0%	0.343	0.201	58.5%	0.142	41.5%
2012	0.364	0.211	57.9%	0.153	42.1%	0.323	0.187	57.9%	0.136	42.1%
1985~2009年变化	136%	107%	-7.9%*	188%	7.9%*	121%	110%	-2.8%*	136%	2.8%*
1985~2012年变化	116%	96%	-5.7%*	149%	5.7%*	94%	90%	-1.4%*	101%	1.4%*

注：*表示大小变化。

^① 基尼系数作为最常用的收入差距指数不能完美分解，因此最常用来分解的指数是广义熵指数，对此的讨论可参见 Shorrocks (1980)、Yang (1999)、Wu 和 Perloff (2005) 等。

从表 3 平均对数离差的相关分解结果来看, 城乡内部差距和城乡之间差距在近 30 年都大为上升了, 并且均在 2009 年达到最大值, 随后开始下降。在 1985 年, 组内差距为 0.107, 组间差距为 0.062。到 2009 年, 组内差距上升 107%, 达到 0.222, 组间差距上升 188%, 达到 0.177。由于后者的上升速度更快, 所以组内差距对总体差距的贡献度在减小, 组间差距的贡献度在上升: 组内差距的贡献比重由 63.6% (1985 年) 降低为 55.6% (2009 年), 组间差距的贡献比重由 36.4% (1985 年) 上升为 44.4% (2009 年)。2009 年以后, 组内差距和组间差距都有不同程度的下降, 并且组内差距的贡献比重有所回升, 由 55.6% (2009 年) 上升至 57.9% (2012 年)。这些结果表明组内差距的重要性虽在下降, 但目前仍是总体差距的主要来源, 而二者的同时上升 (组间差距上升更快) 表明, 全国总体收入差距扩大是二者共同作用的结果。

对泰尔指数的分解得到非常类似的结论。1985 年以来组内差距和组间差距都迅速扩大, 组间差距上升更快, 但组内差距对总体差距的贡献度始终大于组间差距: 组内差距在 1985 年对总体差距的贡献度为 59.3%, 到 2007 年降为 56.1%, 但随后开始回升, 并在 2012 年达到 57.9%。

2. 对收入差距及其上升的回归分解

这一部分使用的数据是 1988、2002 年 CHIP 个体 (家庭) 数据, 根据 Fields (2003) 和 Yue 等 (2006), 本文的回归方程设为:

$$\ln Y_{it} = \alpha_t + \sum_{j=1}^k \beta_{jt} \cdot x_{jit} + \epsilon_{it} \quad (7)$$

其中, t 取值为 1988 和 2002 年。 $\ln Y_{it}$ 为家庭 i 在 t 期人均收入的对数, 解释变量 x_{jit} 包括城镇虚拟变量、家庭规模、地理位置变量 (东部、中部、东北部, 西部省略)、家庭 15~64 岁全职工作人口特征变量 (占家庭总人口比重、性别、教育、经验及其平方、职业类别和产业部门等)。表 4 为相应的 OLS 回归结果^①。

表 4 家庭人均收入对数及其解释因素: 1988~2002 年

变量组	解释变量	1988 年	2002 年
地理位置 (西部省略)	东部	0.282 (0.008)***	0.529 (0.011)***
	中部	0.014 (0.008)*	0.080 (0.010)***
	东北	-0.059 (0.014)***	0.166 (0.014)***
城乡二元结构	城镇虚拟变量	0.680 (0.016)***	0.626 (0.014)***
家庭规模	家庭规模	-0.055 (0.002)***	-0.128 (0.004)***
15~64 岁全职工人 (特征变量)	占家庭人口比重	0.411 (0.017)***	0.344 (0.021)***
	教育	0.005 (0.001)***	0.030 (0.002)***
	经验	-0.012 (0.001)***	-0.037 (0.002)***
	经验平方	0.000 (0.000)***	0.001 (0.000)***
	第二产业部门	0.245 (0.018)***	0.129 (0.023)***
	第三产业部门	0.271 (0.020)***	0.198 (0.023)***
	党员	0.171 (0.012)***	0.230 (0.014)***
	性别	-0.147 (0.016)***	-0.099 (0.018)***
	职业类别	0.129 (0.025)***	0.069 (0.023)***

① 根据这一类文献的传统, 本文不考虑内生性等问题, 只是对异方差进行修正。但值得注意的是, 本文使用的回归方程没有包含城镇虚拟变量的交叉项, 这是由于 Fields (2003) 的方法不适于处理这类问题, 这是该方法的局限。

(续)

变量组	解释变量	1988 年	2002 年
常数项		7.418 (0.023)***	8.421 (0.025)***
R ²		0.572	0.590
F 值		3938.99	1832.17
观测值个数 N		27124	17274

注：被解释变量为家庭人均收入的对数。括号中为修正标准误。*、**、*** 分别表示在 10%、5%、1% 水平下显著。

得益于足够大的样本量，表 4 中各解释变量的回归系数几乎全部高度显著。根据表 4 的回归结果，本文首先对 1988 年、2002 年的收入差距进行分解，以确定每个解释变量对各年总体收入差距的贡献比重，然后我们分别使用基尼系数和对数收入方差计算每个解释因素对总体收入差距（从 1988 年到 2002 年）上升的贡献。表 5 报告了相应的计算结果。

根据表 5 数据第一、二列，除残差项外，对 1988 年收入差距贡献最大的是城镇虚拟变量（28.7%），其次是产业部门（10.2%）、职业属性（5.2%）。在 2002 年，城乡二元分割的贡献度有所下降，降为 24.6%，但仍是总体收入差距的最重要解释因素。此外，家庭规模和地理位置的作用上升，对总体差距的贡献比重分别上升为 9.6%、7.4%，成为总体差距的第二、三位解释因素。

对 1988~2002 年收入差距上升的分解表明（表 5 数据第三、四列），家庭规模是总体差距扩大的最重要解释因素，其次是地理位置和教育因素。城乡二元结构的贡献比重为负说明它对总体差距的贡献度在下降。

表 5 各因素对总体收入差距及其上升的贡献比重：1988~2002 年

解释变量	对收入差距的贡献度		对收入差距上升的贡献度	
	1988 年	2002 年	基尼系数 (1988~2002 年)	对数收入方差 (1988~2002 年)
党员身份	1.6	2.6	11.8	9.9
性别（男性）	0.2	0.3	0.9	0.8
职业类别	5.2	1.7	-28.0	-22.1
地理位置	3.4	7.4	41.7	34.8
经验	0.4	3.1	26.0	21.4
产业部门	10.2	4.6	-42.8	-33.3
城镇虚拟变量	28.7	24.6	-10.3	-3.3
家庭规模	4.9	9.6	50.6	42.4
全职工人比例	1.7	0.0	-14.3	-11.5
教育	1.1	5.0	38.5	31.8
残差项	42.8	41.0	25.8	28.9

注：各数值均为百分数。地理位置包括东部、中部、东北，经验包括经验、经验平方，产业部门包括第二产业、第三产业。

表 6

对家庭规模、教育贡献比重变化的分解

	家庭规模				教育			
	相关成分		权重变化的解释		相关成分		权重变化的解释	
	1988 年	2002 年			1988 年	2002 年		
贡献比重	0.049	0.096			0.011	0.050		
回归系数	-0.055	-0.128	1.350	1.227	0.005	0.030	1.279	1.010
标准差	1.635	1.236	-0.248	-0.227	3.643	3.955	0.024	-0.010
与对数收入的相关系数	-0.414	-0.500	0.211		0.420	0.340	-0.053	
对数收入标准差	0.765	0.819	-0.072		0.765	0.819	-0.020	
总和			1.240	1.000			1.230	1.000

注：关于数据第三、四列，第七、八列的差别和各自计算方法，参见 Fields（2003）。

最后，由于家庭规模、教育对收入差距的影响越来越大，本文对二者贡献权重的上升进行分解，表 6 报告了相关结果。根据表 6 的前四列，家庭规模的标准差减小，但回归系数（绝对值）大为上升，因此家庭规模的贡献上升完全是回归系数变大带来的“价格效应”（或称“系数效应”）：各家庭人口“数量”差异变小，而家庭拥有更多人口的“价格”（或成本）上升。对教育贡献权重的分解得到的结果非常类似：虽然全国家庭教育水平的差别有所上升，但与教育回报率的上升幅度比起来微乎其微，因而教育贡献度的上升完全是“价格效应”，而非“数量效应”。

五、总 结

针对中国国家统计局收入分组数据的特点，本文对基于最大熵密度的广义矩估计法予以改进，避免反复抽样带来的沉重负担，并使用分组数据基尼系数的上下限和“经验法则”提高估计准确度。在通过人口加权得到总体收入分布函数后，本文计算了包括基尼系数、帕尔玛比值在内的四种收入差距指数，并对平均对数离差和泰尔指数进行了解析。本文还使用基于回归方程的分解方法对总体收入差距及其上升进行解释。

本文发现，1985 年以来我国总体收入差距大为上升，但近来有所下降，隐约表现为倒“U”形：基尼系数由 1985 年的 0.321 上升到 2008 年的 0.468，上升了 46%；2009 年以后基尼系数有所下降，但到 2012 年为止仍保持在 0.43 以上。农村的收入差距在这一时期也有大幅上升，基尼系数由 1985 年的 0.278 上升到 2012 年的 0.385。城镇收入差距近 30 年的上升更为迅速，基尼系数在 1985 年仅为 0.169，但到 2008 年已接近 0.34，上升超过 100%，与农村的收入差距程度越来越接近，但近几年城镇的收入分配状况略有改善。

对我国总体收入中最高 10%、最低 40%以及其余 50%人口占有总体收入份额的研究表明，帕尔玛提出的“50/50 法则”在中国有较强适用性：全国 50%的总人口（第 5 到第 9 十分位，D5~D9）占有总体收入的份额始终稳定在 50%或略高（56%）水平。

对广义熵指数的分解发现，城乡内部差距和城乡之间差距都在上升，共同推动了总体收入差距的上升。城乡之间差距上升的幅度大于城乡内部差距，因此城乡内部差距（组内差距）的贡献比重在下降，组间差距的贡献比重在上升，但组内差距始终大于组间差距，仍是收入差距的主要来源。基于回归方程的分析表明，城乡二元分割仍是总体收入差距的首要决定因素，产业部门特征在早期发挥重要作用，家庭规模在后期成为第二位影响因素。家庭规

模也是收入差距上升的最重要贡献因素，地理位置发挥了次要作用。由于回报率的极大提高，教育的作用越来越明显，并成为收入差距扩大的第三位影响因素。

参 考 文 献

- [1] Ahluwalia M. S. , Carter N. , Chenery H. , 1979, *Growth and Poverty in Developing Countries* [J], *Journal of Development Economics*, 6 (3), 299~341.
- [2] Ahmad E. , Wang Y. , 1991, *Inequality and Poverty in China: Institutional Change and Public Policy, 1978 to 1988* [J], *World Bank Economic Review*, 5 (2), 231~257.
- [3] Alesina A. , Rodrik D. , 1994, *Distributive Politics and Economic Growth* [J], *Quarterly Journal of Economics*, 109 (2), 465~490.
- [4] Anand S. , Kanbur S. M. R. , 1993, *The Kuznets Process and the Inequality-Development Relationship* [J], *Journal of Development Economics*, 40 (1), 25~52.
- [5] Atkinson A. B. , Piketty T. , 2007, *Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries* [M], Oxford and New York: Oxford University Press.
- [6] Atkinson A. B. , 1970, *On the Measurement of Inequality* [J], *Journal of Economic Theory*, 2 (3), 244~263.
- [7] Benjamin D. , Brandt L. , Giles J. , 2011, *Did Higher Inequality Impede Growth in Rural China* [J], *Economic Journal*, 121 (557), 1281~1309.
- [8] Benjamin D. , Brandt L. , Giles J. , Wang S. , 2008, *Income Inequality during China's Economic Transition* [A], in Brandt L. , Rawski T. G. (ed.), *China's Great Economic Transformation* [M], Cambridge University Press.
- [9] Bourguignon F. , Morrison C. , 2002, *Inequality among World Citizens: 1820~1992* [J], *American Economic Review*, 92 (4), 727~744.
- [10] Bramall C. , 2001, *The Quality of China's Household Income Surveys* [J], *China Quarterly*, 167, 689~705.
- [11] Cai H. , Chen Y. , Zhou L. , 2010, *Income and Consumption Inequality in Urban China: 1992~2003* [J], *Economic Development and Culture Change*, 58 (3), 385~413.
- [12] Cowell F. A. , Flachaire E. , 2015, *Statistical Methods for Distributional Analysis* [R], *Handbook of Income Distribution*, Vol 2A, Elsevier.
- [13] Fields G. S. , 2003, *Accounting For Income Inequality and Its Change: A New Method, with Application to the Distribution of Earnings in the United States* [J], *Research in Labor Economics*, 22, 1~38.
- [14] Gastwirth J. L. , 1972, *The Estimation of the Lorenz Curve and Gini Index* [J], *Review of Economics and Statistics*, 54 (3), 306~316.
- [15] Gastwirth J. L. , Glauber M. , 1976, *The Interpolation of the Lorenz Curve and Gini Index from Grouped Data* [J], *Econometrica*, 44 (3), 479~483.
- [16] Hajargasht G. , Griffiths W. E. , Brice J. , Rao D. S. P. , Chotikapanich D. , 2012, *Inference for Income Distributions Using Grouped Data* [J], *Journal of Business and Economic Statistics*, 30 (4), 563~575.
- [17] Khan A. , Riskin C. , 2008, *Growth and Distribution of Household Income in China between 1995 and 2002* [J], in Gustafsson B. , Li S. , Sicular T. (ed.), *Inequality and Public policy in China* [M], Cambridge University Press.
- [18] Kopczuk W. , Saez E. , Song J. , 2010, *Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937* [J], *Quarterly Journal of Economics*, 125 (1), 91~128.
- [19] Krol A. , Miedema J. M. , 2009, *Measuring Income Inequality: an Exploratory Review* [R].

Region of Waterloo Public Health, Document No 451158.

- [20] Kuznets S. , 1955, *Economic Growth and Income Inequality* [J], American Economic Review, 45 (1), 1~28.
- [21] Kuznets S. , 1963, *Quantitative Aspects of the Economic Growth of Nations: VIII. Distribution of Income by Size* [J], Economic Development and Culture Change, 11 (2), 1~80.
- [22] Li S. , Sicular T. , 2014, *The Distribution of Household Income in China: Inequality, Poverty and Policies* [J], China Quarterly, 217 (217), 1~41.
- [23] Morduch J. , Sicular T. , 2002, *Rethinking Inequality Decomposition, with Evidence from Rural China* [J], Economic Journal, 112 (476), 93~106.
- [24] Palma J. G. , 2011, *Homogeneous Middles vs. Heterogeneous Tails, and the End of the “Invert-ed-U”*: Its All About the Share of the Rich [J], Development and Change, 42 (1), 87~153.
- [25] Palma J. G. , 2014, *Has the Income Share of the Middle and Upper-middle Been Stable around the “50/50 Rule”, or Has it Converged towards that Level? The ‘Palma Ratio’ Revisited* [J], Development and Change, 45 (6), 1416~1448.
- [26] Park S. Y. , Bera A. K. , 2009, *Maximum Entropy Autoregressive Conditional Heteroskedasticity Model* [J], Journal of Econometrics, 150 (2), 219~230.
- [27] Persson T. , Tabellini G. , 1994, *Is Inequality Harmful for Growth? Theory and Evidence* [J], American Economic Review, 84 (3), 600~621.
- [28] Piketty T. , Seaz E. , 2003, *Income Inequality in the United States, 1913~1998* [J], Quarterly Journal of Economics, 118 (1), 1~39.
- [29] Ravallion M. , Chen S. , 2007, *China’s (uneven) Progress Against Poverty* [J], Journal of Development Economics, 82 (1), 1~42.
- [30] SalaiMartin X. , 2006, *The World Distribution of Income: Falling Poverty and... Convergence, Period* [J], Quarterly Journal of Economics, 121 (121), 351~397.
- [31] Shorrocks A. F. , 1980, *The Class of Additively Decomposable Inequality Measures* [J], Econometrica, 48 (3), 613~625.
- [32] Wang C. , Wan G. , Yang D. , 2014, *Income Inequality in the People’s Republic of China: Trends, Determinants, and Proposed Remedies* [J], Journal of Economic Surveys, 28 (4), 686~708.
- [33] Whyte M. K. , 2014, *Soaring Income Gaps: China in Comparative Perspective* [J], Daedalus, 143 (2), 39~52.
- [34] Wu X. , 2003, *Calculation of Maximum Entropy Densities with Application to Income Distribution* [J], Journal of Econometrics, 115 (3), 347~354.
- [35] Wu X. , Perloff J. M. , 2005, *China’s Income Distribution, 1985~2001* [J], Review of Economics and Statistics, 87 (4), 763~775.
- [36] Wu X. , Perloff J. M. , 2007, *GMM Estimation of a Maximum Entropy Distribution with Interval Data* [J], Journal of Econometrics, 138 (2), 532~546.
- [37] Yang D. T. , 1999, *Urban-Biased Policies and Rising Income Inequality in China* [J], American Economic Review Papers and Proceedings, 89 (2), 306~310.
- [38] Yue X. , Sicular T. , Li S. , Gustafsson B. , 2008, *Explaining Incomes and Inequality in China* [A], in Gustafsson B. , Li S. , Sicular T. , *Inequality and Public policy in China* [M], Cambridge University Press
- [39] 阿玛蒂亚·森、让·德雷兹：《印度：经济发展与社会机会》[M]，社会科学文献出版社，2006。
- [40] 程永宏：《二元经济中城乡混合基尼系数的计算与分解》[J]，《经济研究》2006年第1期。
- [41] 程永宏：《改革以来全国总体基尼系数的演变及其城乡分解》[J]，《中国社会科学》2007年第4期。
- [42] 国务院研究室课题组：《关于城镇居民个人收入差距的分析和建议》[J]，《经济研究》1997年第8期。
- [43] 李成瑞：《关于中国近几年的经济政策对居民收入和消费状况影响的统计报告》[J]，《统计研究》1986年第1期。

- [44] 李实、罗楚亮：《中国收入差距究竟有多大？——对修正样本结构偏差的尝试》[J]，《经济研究》2011年第4期。
- [45] 联合国开发计划署（UNDP）：《2014年人类发展报告：促进人类持续进步：降低脆弱性，增强抗逆力》[R]，联合国开发计划署，2014。
- [46] 珀金斯、德怀特·H.：《中国的经济政策及其贯彻情况》[A]，《剑桥中华人民共和国史：中国革命内部的矛盾（1966~1982年）》[M]，中国社会科学出版社，1992。
- [47] 世界银行：《从贫困地区到贫困人群：中国扶贫议程的演进》[R]，世界银行东亚及太平洋地区扶贫与经济管理局，2009。
- [48] 万广华：《解释中国农村区域间的收入不平等：一种基于回归方程的分解方法》[J]，《经济研究》2004年第8期。
- [49] 万广华：《城镇化与不均等：分析方法和中国案例》[J]，《经济研究》2013年第5期。
- [50] 万海远、李实：《户籍歧视对城乡收入差距的影响》[J]，《经济研究》2013年第9期。
- [51] 王弟海：《收入和财富分配不平等：动态视角》[M]，格致出版社、上海人民出版社，2009。
- [52] 王小鲁：《灰色收入与国民收入分配》[J]，《比较》2010年第3期。
- [53] 王亚峰：《中国1985~2009年城乡居民收入分布的估计》[J]，《数量经济技术经济研究》2012年第6期。
- [54] 谢伏瞻：《中国收入分配的现状与政策分析》[J]，《管理世界》1994年第2期。
- [55] 许庆、田土超、徐志刚、邵挺：《农地制度、土地细碎化与农民收入不平等》[J]，《经济研究》2008年第2期。
- [56] 赵人伟、李实：《中国居民收入差距的扩大及其原因》[J]，《经济研究》1997年第9期。

Fluctuation and Causal Analysis of China's Income Gap: 1985~2012

Zhang Tao

(School of Economics, Peking University)

Abstract: This paper improves on the GMM method, which is based on the maximum entropy density, and uses interval data of income to examine the trend of income gap in China from 1985 to 2012. A new decomposition method based on the regression equation is used to analyze the reasons for income gap and its increase. The empirical results show that the income gap of China's overall income distribution rose rapidly from 1985 to 2008, but then declined, vaguely akin to an inverted U-shape. The income share of the middle and upper-middle group, which account for half the population, has been stable around slightly above 50%, in line with the Palma "50/50 Rule." Both the income gaps within and between urban and rural areas are on the rise. Since the latter grows faster, its contribution to the overall gap is increasing. The urban-rural division is always the primary source of national income gap, and family size is the most important factor in explaining the rising of the gap.

Key Words: GMM; Regression Analysis; Income Gap; 50/50 Rule

JEL Classification: C13; D31; O15

(责任编辑：彭 战)