

城乡收入差距的变化及其对全国收入差距的影响^{*}

罗楚亮

内容提要 本文在中国居民收入分配课题组（CHIP）2007年和2013年全国住户调查数据的基础上，讨论了这一时期城乡居民收入差距的变化及其对于全国居民收入差距的影响。本文的主要发现表明，从人群组分解来看，城乡之间收入差距的缩小和农村人口份额的下降是导致这一时期全国居民收入差距缩小的主要原因，但城乡之间的组间不平等对于当年的全国收入差距仍具有非常高的解释份额。在基于收入函数的不平等指数分解中，控制教育、家庭结构、地区等因素后，城乡差距对于总体收入差距仍具有非常高的解释份额，也是这一时期收入基尼系数变化最为关键的解释因素。尽管城乡收入差距随着收入分布分位点的上升而下降，但从基于收入分布的分解分析来看，禀赋效应对城乡收入差距的解释作用随着收入分布分位点的上升而上升，教育的禀赋效应尤其如此。

关键词 城乡差距 收入不平等 人群组分解 收入函数分解

一 引言

长期以来，中国居民收入差距都表现出快速上升的趋势。收入差距不断扩大，成为中国在经济转型时期的一个重要特征。李实（2003）回顾了2000年以前收入基尼系数的各种不同估计结果，发现估算的20世纪80年代初期基尼系数通常在0.3左右。Ravallion & Chen（2007）描述了中国收入分配的基尼系数在1981年至2001年期间不断攀升的趋势。根据中国居民收入分配课题组（CHIP）历次住户调查的数据估算表

* 罗楚亮，北京师范大学经济与工商管理学院、中国收入分配研究院，电子邮箱：luochl@bnu.edu.cn。本文受国家社会科学基金“劳动力市场转型的收入分配效应研究”（13AJY007）、北京师范大学自主科研基金资助。

明，收入基尼系数在 2002 年为 0.46 左右，2007 年上升至 0.48（李实等，2013），而最近的一轮调查结果（2013 年）则显示全国收入基尼系数有所下降。国家统计局公布的 2003 年以来的全国收入基尼系数多在 0.47~0.49 之间，2008 年达到峰值 0.491，此后则有所下降，2013 年为 0.473，2015 年降至 0.462^①。对于收入差距的这种下降倾向，目前仍缺乏相应的解释。本文在 2007 年和 2013 年住户调查数据的基础上，发现城乡收入差距的缩小以及农村人口份额的下降导致了这一期间全国收入差距的下降。

在中国，城镇和农村长期被视为两个分割的系统，城乡之间存在着巨大的发展差距，这既表现在收入方面，也表现在居民消费、公共服务等其他方面，但城乡收入差距仍是其中最为基本的方面，并且对其他福利形式的城乡差距具有制约作用。既有的研究通常都揭示了中国城乡收入差距长期处于非常高的水平，并且是导致全国总体收入差距居高不下、持续扩大的重要因素。根据国家统计局公布的城镇和农村人均收入，城/乡人均收入比率从改革开放初期到 20 世纪 80 年代中期有所下降，但随后总体上持续上升。从 1990 年到 2015 年期间，城/乡人均收入比率从 2.20 上升到 2009 年的 3.33 倍，此后出现了下降倾向，如 2013 年降至 3.03 倍，2015 年进一步降至 2.95 倍，但总体上仍处在比较高的水平^②。如果把全国收入差距按照人群组进行广义熵指数（Generalized Entropy，简称 GE 指数）分解，城乡之间的收入差距对总体收入差距的解释份额从 1988 年的 38% 上升到 1995 年的 43%（别雍·古斯塔夫森、李实，1999），Sicular et al. (2007) 的研究显示 2002 年的解释份额达到 47%^③。这意味着在全国收入差距上升的时期，城乡居民收入差距具有非常重要的贡献。持续存在的巨大城乡差距在社会公平、经济发展等方面的不利影响已经被广泛认识到，在此不再赘述。

根据中国居民收入分配课题组 2007 年和 2013 年的城乡住户调查数据，这一期

-
- ① 当然，对于近年来的收入差距存在着不同的估计结果。例如，西南财经大学中国家庭金融调查（CHFS）报告 2011 年的全国收入基尼系数达到了 0.61。北京大学的中国家庭追踪调查（CFPS）给出的 2010 年和 2012 年的收入基尼系数分别为 0.530 和 0.532（Xie & Zhou, 2014）。但由于这些都是近期才出现的调查数据，难以据此推断一段时期以来的收入差距变动的趋势特征。
 - ② 由于很少有国家采取与中国类似的城乡分割体制，这种国际比较可能是比较困难的。蔡昉和杨涛（2000）搜集的 36 个国家的“农业与非农业标准劳动者的收入比率”通常都低于 1.5，远远低于中国的城乡收入比率。当然这种比较可能存在口径上的不一致性。
 - ③ 利用相同年份的住户调查数据，罗楚亮（2006）发现，按照常用的城镇居民人均可支配收入和农村居民人均纯收入，城乡收入差距对全国收入差距的解释程度分别为 33%、37% 和 40%。

间居民收入差距特征出现了非常明显的变化。城镇和农村内部收入差距仍有所扩大，但城乡收入差距和全国收入差距有一定幅度的缩小。本文在描述城乡居民收入分布、收入差距的基本变动特征基础上，利用人群组分解和回归分析，讨论了城乡差距的基本状况及其变动特征，以及城乡差距对于全国收入差距的影响。本文的其余部分包括：第二部分描述本文所采用的数据，并对城乡收入分布状况进行简要的描述；第三部分对基尼系数、平均对数离差（Mean Log Deviation，简称 MLD 指数）和泰尔指数进行人群组分解；第四部分在收入函数的基础上，讨论城乡差距对总体收入差距的影响；第五部分讨论城乡差距在收入分布不同位置上的表现特征；最后是全文的总结。

二 数据描述与收入分布特征

本文所使用的数据来自于中国居民收入分配课题组 2007 年和 2013 年的城乡居民住户调查^①。两次调查样本都来自于国家统计局的常规住户调查，收入信息来自于住户调查的日记账，调查实施都是由国家统计局的住户调查系统完成。2007 年城乡住户调查沿袭了中国城乡分割的基本特征，在国家统计局内部由不同的部门负责，因此该年的城镇和农村住户调查分别独立完成。由于国家统计局对城乡住户调查进行了一体化改革，城乡住户调查由一个部门统一实施，并且在抽样设计、调查执行和指标核算等方面不再强调城乡之间的差异性，因此 2013 年住户调查没有再明确区分城镇与农村。为了与以往研究的可比性以及研究主题的延续性，CHIP 调查仍基于地域原则区分了城镇与农村问卷^②。

表 1 给出了 2007 年和 2013 年城镇与农村的样本数量及其省份分布状况。2007 年城镇与农村住户数量分别为 10000 户和 13000 户，城乡所覆盖的省份大多是相同的，只有上海和河北分别只进行了城镇和农村住户调查。2013 年数据一共包括 15 个省份的 17417 户，按照地域区分的城镇与农村住户数量分别为 6866 户和 10551 户，城乡住户调查数据所涉及的省份都是相同的。2007 年和 2013 年所覆盖的省份大体上是相同的。

^① 基于前三次调查（1988 年、1995 年和 2002 年）数据对城乡居民收入差距的讨论可参见罗楚亮（2006）。

^② 这种城乡划分的方式也是与以往住户调查的城乡概念是一致的。在住户调查中，城乡主要是按照家庭所在地的地理特征来区分的，而不是户籍。因此，在以往的农村住户调查中，可能包含城镇户口；而在城镇住户调查中，原则上也可能有农村户口。

不同的省份包括，2007 年的浙江、福建和上海（城镇）、河北（农村）没有包括在 2013 年数据中，而 2013 年新增了山东和新疆。总体而言，两年数据都分别覆盖了东中西部不同区域，但不同区域以及城乡之间的人口比重与总体人口结构构成存有差异，因此本文的讨论对个体样本在东中西部区域与城乡之间进行了重新加权。

表 1 样本量及其分布

	2007 年		2013 年	
	城镇	农村	城镇	农村
住户数量	10000	13000	6866	10551
个人数量	29262	51847	20331	39408
省份数量	16	16	15	15
相同省份	北京、山西、辽宁、江苏、安徽、河南、湖北、湖南、广东、重庆、四川、云南、甘肃			
特定省份	上海（城镇）、浙江、福建、河北（农村）		山东、新疆	

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

表 2 给出了城镇和农村住户中人均收入水平及其不平等程度的基本结果。为了便于比较，表 2 同时给出了基于 CHIP 数据和国家统计局（NBS）公布的相应信息。总体来说，CHIP 数据中的人均收入水平要略高于国家统计局公布的数据。2007 年，CHIP 数据中城镇与农村人均收入分别比 NBS 高出 8% 和 4.6%；2013 年，CHIP 数据中城镇与农村①人均收入分别比 NBS 高 5.8% 和 10.2%。这导致 2007 年 CHIP 数据中的城/乡收入比率要略高于 NBS，而 2013 年则相反。但 CHIP 与 NBS 数据总体上差异不大，并且城/乡收入比率在 2007 年到 2013 年期间的下降倾向在两个数据来源中都是相同的。从收入不平等程度来看，2007 年城镇与农村收入基尼系数非常接近。CHIP 数据中 2013 年农村收入基尼系数高于 NBS，但 2013 年城镇基尼系数没有公开的官方结果。

① 2013 年城乡居民收入数据在《中国统计年鉴》（2015 年）上存在着不同来源。按照“表 6-6 城乡居民人均收入”，城镇居民人均可支配收入和农村居民人均纯收入分别为 26955.1 元和 8895.9 元；但在“表 6-21 城镇居民分地区人均可支配收入”中，全国的城镇人均可支配收入为 26467 元；相应地，在“表 6-25 农村居民分地区人均可支配收入”中，全国农村人均可支配收入为 9429.6 元。按照后一来源，城乡人均收入比率为 2.807，低于表 2 中的 3.03。这种差异反映了两个年份中收入概念的不同。即便是城镇的可支配收入，2013 年的核算口径也与 2007 年有比较大的差异。农村收入的口径差异要更大一些。

表2 城乡收入水平与不平等程度

		2007年		2013年	
		人均收入	基尼系数	人均收入	基尼系数
CHIP	城镇	14890	0.3342	28524	0.3546
	农村	4329	0.3755	9804	0.4102
	城/乡收入比率	3.44		2.91	
NBS	城镇	13786	0.3400	26955	
	农村	4140	0.3740	8896	0.3857
	城/乡收入比率	3.33		3.03	

资料来源：CHIP 收入信息根据 CHIP2007 年和 2013 年调查数据计算得到；“购买力平价调整”基于 Brandt & Holz (2006) 推算的城乡和省份之间的生活费用指数调整计算得到；NBS 的收入信息来自于《中国统计年鉴》(2014 年)；NBS 城镇和农村基尼系数来自于相应年份《中国居民收入分配年度报告》。

相关年份的《中国统计年鉴》也给出了分组收入的基本信息，表 3 给出了 CHIP 与统计年鉴分组收入的比较^①。2007 年 CHIP 城镇各收入组的收入水平略高于统计年鉴的结果，农村则相反，但不同收入组中两种来源的收入差异相对幅度没有明显的不同，

表3 CHIP 与统计年鉴分组收入的比较

			最低 20%	中低 20%	中间 20%	中高 20%	最高 20%
2007 年	城镇	统计年鉴	5357	8901	12042	16386	29509
		CHIP	5455	9146	12427	16880	30558
		统计年鉴/CHIP 比率	0.982	0.973	0.969	0.971	0.966
	农村	统计年鉴	1347	2582	3659	5130	9791
		CHIP	1302	2452	3477	4908	9519
		统计年鉴/CHIP 比率	1.034	1.053	1.052	1.045	1.029
2013 年	城镇	统计年鉴	9896	17628	24173	32614	57762
		CHIP	9136	17043	23752	32673	60216
		统计年鉴/CHIP 比率	1.083	1.034	1.018	0.998	0.959
	农村	统计年鉴	2878	5966	8438	11816	21323
		CHIP	2465	5179	7646	11304	22662
		统计年鉴/CHIP 比率	1.167	1.152	1.104	1.045	0.941

注：《中国统计年鉴》2007 年城镇没有直接公布最低 20% 和最高 20% 收入组的人均收入水平，这里通过分别对最低收入户 10%、低收入户 10%、高收入户 10% 和最高收入户 10% 取简单平均数得到。

资料来源：CHIP 收入信息根据 CHIP2007 年和 2013 年调查数据计算得到；统计年鉴收入信息来自于相应年份《中国统计年鉴》。

① 来自 CHIP 数据的分组收入是加权调整后的结果。

相对差异幅度通常在 3% ~ 5% 左右。因此, CHIP2007 年所得到的城镇和农村内部基尼系数与国家统计局的公布结果非常接近, 尽管城乡差距有所差异。但 2013 年数据中, 城镇和农村数据都表现出, 低收入组中, CHIP 的人均收入要低于统计年鉴上的结果; 而高收入组中, CHIP 人均收入要略高于统计年鉴上的结果。这可能导致 CHIP 数据中城镇和农村内部收入差距要比国家统计局大样本住户调查的结果要略高一些。表 2 中 2013 年农村内部基尼系数在 CHIP 和 NBS 来源上的差异已经说明了这一点, 据此可以推测城镇也可能具有相同的特征。

为了描述城乡居民收入的分布特征, 分城乡收入(对数)核密度分布的特征如图 1 所示。图 1 既给出了同一个年份中城镇与农村收入分布的比较, 也给出了城镇(urban)和农村(rural)内部不同年份收入分布的比较。对于给定的年份(2007 年或 2013 年), 图 1-(a) 和图 1-(b) 表明, 城镇收入分布曲线通常处在农村居民收入分布曲线的右侧, 并且城镇收入分布曲线要更为尖一些, 表明城镇居民平均收入水平会更高一些, 而收入差距则会相对略低。这一基本特征对于 2007 年和 2013 年都成立。图 1-(c) 描述了 2007 年和 2013 年城镇收入的分布特征, 而图 1-(d) 描述了这两个年份中农村收入的分布特征。总体而言, 2013 年的分布曲线向右推移, 并且 2007 年的分布曲线要更为集中一些。图 2 分别给出了同一年份中城镇与农村收入分布的核密度差异($ud - rd$), 以及城镇(ud)与农村(rd)内部不同年份收入分布的核密度差异。

尽管从表 2 中可以看到, 2007 年到 2013 年期间, 无论是城镇还是农村内部居民收入水平都有较大幅度的增长, 但图 3 表明这种增长幅度在收入分布不同位置上是不同的。从图 3 来看, 在任意的收入分位点上, 农村居民收入增长率(以两年收入对数差来衡量, 即 $\ln(RY2013) - \ln(RY2007)$)都要高于城镇居民, 即 $\ln(UY2013) - \ln(UY2007)$, 但城镇和农村的收入增长曲线表现形式存在较为明显的差异性。农村收入增长曲线具有明显的向上倾斜的特征, 收入分布高分位点对应的收入增长率要明显高于低分位点; 而城镇收入分布不同分位点上的收入增长率差异则没有那么明显, 尤其是在 10 到 90 分位点之间, 收入增长曲线总体上非常平缓。这一特征意味着, 农村收入差距的扩大幅度要更大一些, 城乡居民收入差距总体上有所缩小。后一特点也体现在图 4 中。图 4 给出了 2007 年和 2013 年不同收入分位点上的城乡居民收入差距(以城镇收入与农村收入的对数差来衡量 $\ln(UY) - \ln(RY)$)。2013 年曲线整体落在 2007 年下方, 表明城乡收入差距的下降发生在收入分布的不同位置上。两个年份的城乡收入差距曲线都表现出随着收入分位点上升而下降的特征, 意味着农村高收入人群与城镇高收入人群之间的收入差距要低于农村低收入人群与城镇低收入人群之间的收入差距。

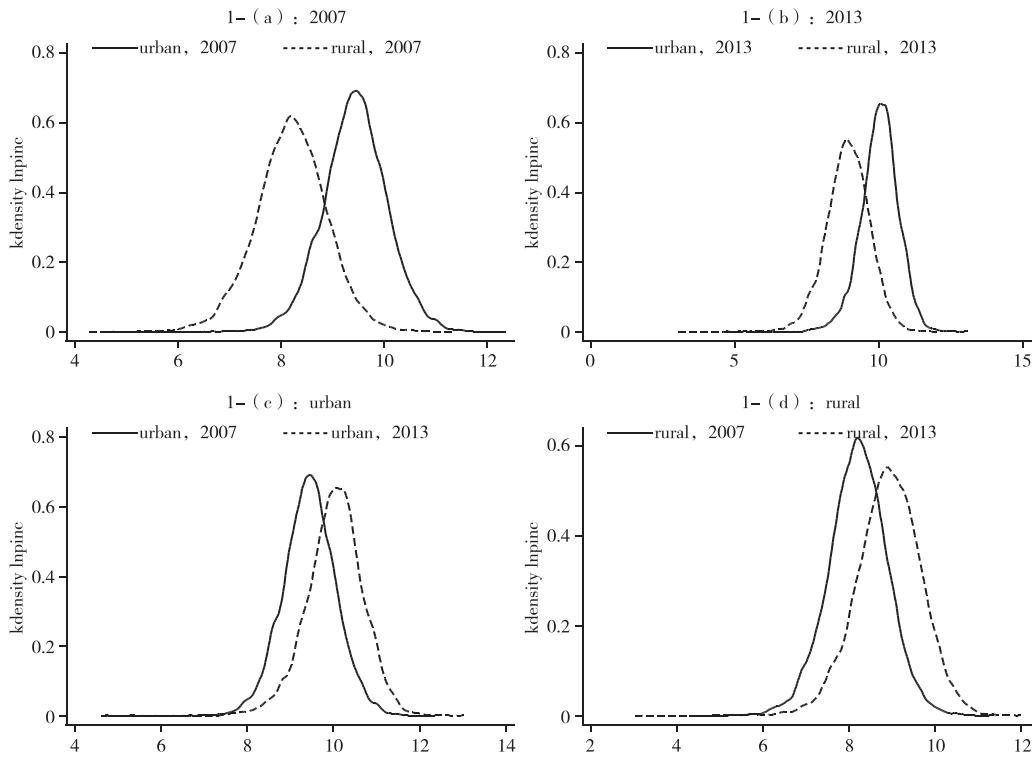


图1 收入（对数）分布的核密度

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

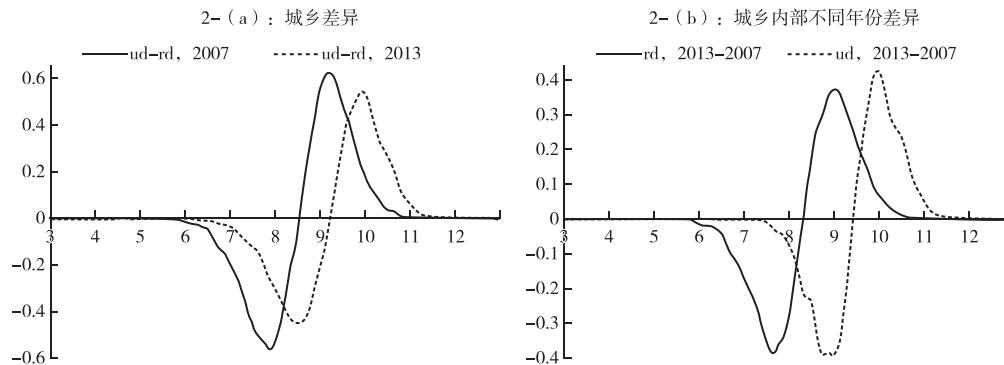


图2 收入（对数）核密度差异

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

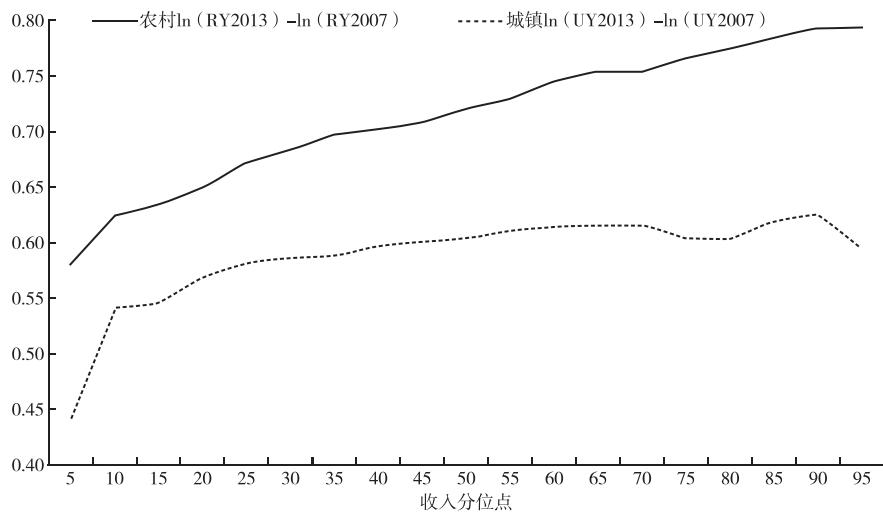


图3 城乡收入增长曲线

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

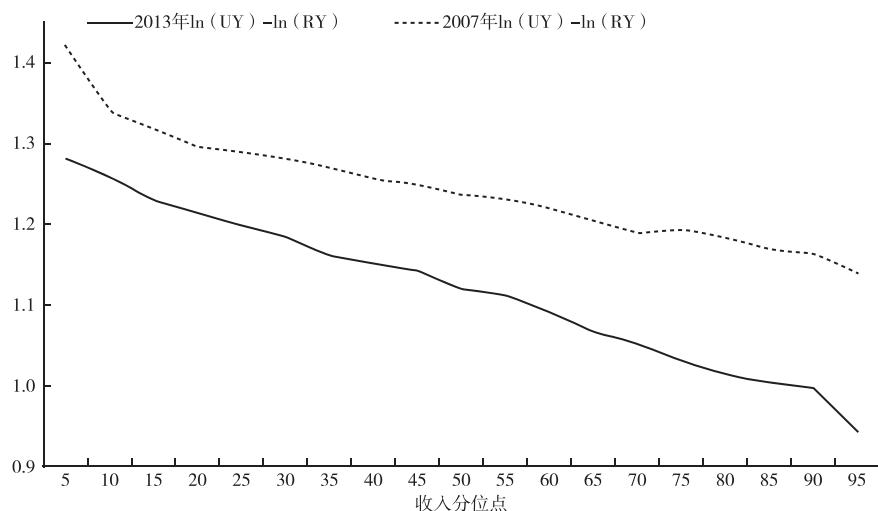


图4 收入（对数）分位点城乡差距

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

三 基于人群组的分解

(一) 城乡分解的总体特征

为了考察城乡差距对于全国收入差距的影响，我们可以将总体不平等分解为组内（城镇内部和农村内部）不平等和组间（城乡之间）不平等。被分解的总体不平等程度，主要以两个（类）指标来衡量，一是基尼系数，这是最为常用的收入不平等程度衡量指标；二是 GE 指数，这一类指数具有能够将总体不平等程度完全分解为组间差距和组内差距的特征。

然而，对于人们常用的基尼系数，这种组内差距和组间差距的分解方式并不能得到完全的分解。一般说来，总体基尼系数与人群组基尼系数之间存在如下关系（徐宽，2003）：

$$Gini_{whole} = \sum_{k=1}^K \omega_k^{INC} \omega_k^{POP} Gini_k + Gini_{between} + R \quad (1)$$

其中， $Gini_{whole}$ 表示整个人群的总体基尼系数， $Gini_k$ 表示人群组 k 的组内基尼系数， $Gini_{between}$ 表示人群组间基尼系数，衡量组间不平等。 ω_k^{INC} 和 ω_k^{POP} 分别表示人群组 k 的收入和人口在总体人群中所占份额。 $\sum_{k=1}^K \omega_k^{INC} \omega_k^{POP} Gini_k$ 衡量了组内不平等， $\omega_k^{INC} \omega_k^{POP} Gini_k$ 则为人群组 k 对总体人群基尼系数的贡献数量。 R 表示未能分解的剩余项^①。表 4 给出了基尼系数的城乡分解结果。

表 4 基尼系数的城乡分解

		城镇人口	城镇收入	组内不平等	城镇	农村	组间不平等	重叠因素	总体不平等
2007 年	绝对量			0.1613	0.0892	0.0721	0.2975	0.0176	0.4765
	份额(%)	38.89	68.64	33.86	18.72	15.14	62.44	3.70	100.00
2013 年	绝对量			0.1807	0.1209	0.0598	0.2496	0.0319	0.4622
	份额(%)	47.21	72.17	39.09	26.15	12.94	54.00	6.91	100.00
变动(2013 - 2007 年)	绝对量			0.0194	0.0317	-0.0123	-0.0479	0.0143	-0.0143
	份额(%)			-135.66	-221.68	86.01	334.97	-100.00	100.00

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

① 这一剩余项意味着基尼系数不能被“组内 - 组间”完全分解，其数值大小通常也是不确定的。一般认为，与不同人群组收入分布的重叠程度相关。

与此不同的是，GE 指数能够将总体不平等完全分解为组间不平等和组内不平等，如：

$$GE_{whole} = \sum_{k=1}^K \omega_k^{POP} (\mu_k/\mu)^\theta GE_k + GE_{between} \quad (2)$$

μ_k 和 μ 分别表示人群组 k 和全部人口的平均收入水平， θ 则为 GE 指数中的一个参数， $\theta = 0, 1$ 则分别为通常的 MLD 指数和泰尔指数。其余各符号含义与基尼系数分解类似。其中，泰尔指数和 MLD 指数是最为常用的两个 GE 指数，本文对于 GE 指数的分城乡分解，将以这两个指数为基础，相应的分解结果可见表 5。

表 5 GE 指数的城乡分解

		2007 年		2013 年		变动	
		绝对量	份额(%)	绝对量	份额(%)	绝对量	份额(%)
MLD 指数: GE(0)	城镇	0.0738	18.22	0.1027	26.46	0.0289	-173.17
	农村	0.1459	36.03	0.1491	38.43	0.0033	-19.66
	组内	0.2196	54.25	0.2518	64.89	0.0322	-192.83
	组间	0.1852	45.75	0.1363	35.11	-0.0489	292.83
	总计	0.4048	100.00	0.3881	100.00	-0.0167	100.00
泰尔指数: GE(1)	城镇	0.1319	33.96	0.1568	43.36	0.0249	-92.99
	农村	0.0772	19.88	0.0778	21.53	0.0006	-2.41
	组内	0.2091	53.84	0.2346	64.89	0.0255	-95.40
	组间	0.1792	46.16	0.1269	35.11	-0.0523	195.41
	总计	0.3883	100.00	0.3615	100.00	-0.0268	100.00

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

根据表 2 可以发现，无论是按照国家统计局公布的结果还是基于 CHIP 数据的计算结果，城镇与农村内部不平等程度在 2007 年到 2013 年期间有所下降。由于长期以来，城乡差距都是中国总体收入差距的重要原因，因此这种下降对于总体不平等程度变化将具有十分重要的影响。表 4 的分解结果也揭示了这一现象。对基尼系数的人群组分解表明，组间不平等对两个年份的总体不平等程度都具有非常重要的解释作用。2007 年组间不平等可以解释总体不平等的 62%，这一份额在 2013 年有所下降，但也仍为 54%，占总体不平

等的一半以上。因此，城乡差距仍是总体差距中最为重要的解释因素^①。

在 2007 年到 2013 年期间，组间不平等对于总体基尼系数贡献的绝对数量以及相对贡献份额都有明显下降。从绝对量来看，组间基尼系数从 2007 年的 0.2975 下降到 2013 年的 0.2496；从相对份额来看，组间不平等的贡献份额下降了 8 个百分点。从两个年份基尼系数的变动来看，城乡差距所体现出的组间不平等对于总体基尼系数的下降具有最为重要的解释份额。

对于总体基尼系数下降的另一个解释因素来自于农村人群。农村人群对于总体基尼系数的贡献绝对量下降了 1.23 个百分点，从而对于总体不平等程度下降有较高的解释作用。值得注意的是，从表 2 可以看到，农村内部基尼系数从 2007 年到 2013 年是有较大幅度上升的，从 0.3755 增加到 0.4102，也就是说农村内部不平等程度也是在上升的。从基尼系数分解公式中可以看到，人群组内对于总体不平等程度的贡献是由 $\omega_k^{ING} \omega_k^{POP} Gini_k$ 决定的，除了组内不平等程度，相应的人口份额与收入份额也是非常重要的影响因素。而 2007 年到 2013 年期间农村人群对于总体基尼系数解释程度的下降，主要是由于农村人口份额的下降所导致的。从表 4 可以看出，这一期间农村人口份额下降了近 9 个百分点，农村居民总体收入份额下降了将近 4 个百分点^②。这意味着，农村人口份额的下降、城镇人口份额的上升，成为总体收入不平等程度下降的另一个重要解释因素。

^① 按照相同的方式，我们也对前几轮 CHIP 数据的城乡居民收入基尼系数进行了分解，相应结果可见下表。从中可以看到，城乡的组间不平等对于总体不平等（基尼系数）的解释份额一直处于 60% 左右。

		城镇 基尼	农村 基尼	城镇 人口	城镇 收入	组内 不平等	城镇	农村	组间 不平等	重叠 因素	总体 不平等
1988 年	绝对量	0.2240	0.3319			0.1301	0.0538	0.0763	0.2452	0.0146	0.3899
	份额(%)			38.25	62.77	33.36	13.79	19.57	62.88	3.76	100.00
1995 年	绝对量	0.2820	0.3807			0.1531	0.0699	0.0832	0.2606	0.0222	0.4359
	份额(%)			38.44	64.50	35.13	16.05	19.08	59.79	5.08	100.00
2002 年	绝对量	0.3197	0.3670			0.1590	0.0708	0.0882	0.2768	0.0209	0.4568
	份额(%)			35.21	62.89	34.82	0.1590	19.32	60.60	4.58	100.00

注：本表所给出的城镇人口份额是基于样本结构直接计算得到的，没有根据相应年份的城乡人口分布进行加权调整。

资料来源：根据 CHIP1988 年、1995 年和 2002 年数据计算得到。

^② 由于城镇与农村人均收入比率是在下降的，这意味着农村人均收入增长速度要高于城镇。如果城乡相对人口份额保持不变，则农村人群收入份额将会上升。因此，农村人群收入份额的下降，主要的原因仍在于农村人口份额的下降。

从 2007 年到 2013 年，总体人群的 MLD 指数和泰尔指数都有所下降。MLD 指数从 2007 年的 0.4048 下降到 2013 年的 0.3881；泰尔指数从 0.3883 下降到 0.3615。在表 5 对 GE 指数的城乡分解中可以看到，城乡的组间不平等对于 MLD 指数和泰尔指数的下降都具有正向解释作用，这一点与表 4 对基尼系数的分解构成相同。城乡组间 MLD 指数和泰尔指数均下降了 5 个百分点左右，对于 2007 年到 2013 年期间的 GE 指数下降具有关键性的解释作用。这一变化特征与 CHIP 数据的前几轮调查结果有明显的不同。根据泰尔指数的分解结果，1988 年、1995 年和 2002 年城乡差距对于总体差距（泰尔指数）的解释份额分别为 33.48%、37.41% 和 40.20%（罗楚亮，2006）。从表 5 可以看到，2007 年进一步上升至 46.16%。因此，2007 年以前的几个年份中，城乡差距对于总体差距的贡献份额都是逐步上升的，并且在 2002 年到 2007 年期间的上升速度比以往时期要更快一些。

从各人群组的贡献份额变化来看，组间不平等的变化也是唯一的正向解释因素，这一特征与表 4 不同。农村居民对于总体 MLD 指数和泰尔指数的绝对贡献数量也是在上升的，尽管这一上升的幅度非常小，对于总体差距变动的解释份额也非常低。但从 $\omega_k^{POP} (\mu_k/\mu)^{\theta} GE_k$ 中可以看到，某个群体对于总体 GE 指数的影响，除了其人群组内部 GE 指数变化以外，还受到人口构成份额的影响。尽管农村居民对总体 GE 指数的贡献数量在 2013 年略有上升，但其上升的幅度要大大低于农村组内 GE 指数的上升幅度。从表 6 可以看出，城镇和农村内部的组内 GE 指数在 2007 年到 2013 年期间都是在上升的。农村内部的 MLD 和泰尔指数都分别上升了 4.5 个百分点，但由于农村人口份额的下降，导致农村对全国 GE 贡献的绝对量只有非常微弱的上升幅度。

表 6 组内 GE 指数

	年份	组内 GE 指数		城镇人口 份额 (%) ^①	$(\mu_k/\mu)^{\theta}$		对全国 GE 贡献的绝对量	
		城镇	农村		城镇	农村	城镇	农村
MLD 指数： GE (0)	2007 年	0.1893	0.2390	0.3897	1	1	0.0738	0.1459
	2013 年	0.2171	0.2830	0.5632	1	1	0.1027	0.1491
泰尔指数： GE (1)	2007 年	0.1922	0.2458	0.3897	1.7603	0.5145	0.1319	0.0772
	2013 年	0.2173	0.2794	0.4730	1.5251	0.5286	0.1568	0.0778

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

① 这里的城镇人口份额与表 4 的略有差异，因为 GE 指数的计算只限于收入为正的样本。对于收入为 0 或负的情形，在计算 GE 指数时都会被舍去，这会导致样本结构略有差异。

总而言之，对基尼系数和 GE 指数按照城乡人群组的分解表明，城乡组间差距的缩小成为总体不平等程度下降的主要解释因素。除此而外，农村人口比重的下降，也对总体不平等程度的下降具有重要解释作用。

（二）地区生活费用指数调整对城乡差距的影响^①

不同地区之间存在着比较严重的生活费用指数差异，或者说不同地区之间的货币购买力水平不同，这会影响到实际收入差距的测量结果。特别是，城乡之间所存在的生活费用指数差异对于城乡居民收入差距所造成的偏离程度可能会更为严重。因此，是否进行地区生活费用指数差异调整，对于收入差距及其来源解释可能具有比较重要的影响。

表 7 给出了经过生活费用指数调整的城镇与农村内部收入不平等程度，并按照人群组对全国基尼系数和 MLD 指数、泰尔指数进行了分解。从中可以看出，经过生活费用指数调整，城镇、农村内部以及全国的收入不平等程度相对于未进行生活费用指数调整的情形都有不同程度的下降。但从不同年份之间的比较来看，无论是城镇、农村内部还是全国的收入不平等程度，都表现出轻微的上升倾向，全国基尼系数从 2007 年的 0.4216 上升至 2013 年的 0.4230。而在未进行生活费用指数调整的情形中，全国收入不平等程度有所下降。经过生活费用指数调整后，无论是基尼系数还是 MLD、泰尔指数，组内不平等从 2007 年到 2013 年都略有上升，但是组间不平等都有所下降。由此也可以看出，生活费用指数调整主要影响的是城乡之间的不平等程度。这一特征也表现在表 2 中，经过生活费用指数调整后，2007 年城乡居民收入比率从 3.44 下降到 2.42，2013 年则下降至 2.18。生活费用指数调整对于城乡居民收入差距比率的下降具有非常明显的影响。

从基尼系数分解来看，经过生活费用指数调整的组内不平等与表 4 的组内不平等没有明显的差异，但两个年份组间基尼系数比表 4 的结果分别下降 6~7 个百分点。这导致组间不平等对于基尼系数的解释份额也较大幅度地下降。但在基尼系数分解结果中，组间不平等仍是总体不平等最为重要的解释因素。2007 年，组间不平等对于总体基尼系数的贡献份额达 53.87%，2013 年降低了将近 9 个百分点，为 45.20%，仍高于组内不平等的贡献份额。MLD 和泰尔指数的分解结果也显示，生活费用指数调整会降低组间不平等对于总体不平等的贡献份额，但是组间不平等对于总体不平等贡献份额呈现下降倾向的特征并没有改变。

^① 这里的生活费用指数调整以 Brandt & Holz (2006) 为基础，并依据各省份城镇和农村居民消费价格指数 (CPI) 推算到 2007 年和 2013 年。

表 7 生活费用指数调整与收入差距

	2007 年			2013 年		
	基尼系数	MLD 指数	泰尔指数	基尼系数	MLD 指数	泰尔指数
城镇内部	0.3145	0.1686	0.1704	0.3413	0.2037	0.2016
农村内部	0.3724	0.2372	0.2383	0.4073	0.2871	0.2811
组内不平等	0.1627 (38.59)	0.2105 (66.95)	0.1965 (65.46)	0.1793 (42.39)	0.2476 (76.54)	0.2284 (75.73)
组间不平等	0.2271 (53.87)	0.1039 (33.05)	0.1037 (34.54)	0.1912 (45.20)	0.0759 (23.46)	0.0732 (24.27)
交叉项	0.0317			0.0525		
合计	0.4216	0.3144	0.3002	0.4230	0.3235	0.3016

注：括号内为相应来源的不平等程度占总体不平等程度的百分比。

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

四 基于收入函数的收入差距分解

在基于人群组的分解中，“城乡”因素对于总体差距的影响是在没有控制影响收入的其他因素的情形下得到的。为此，本部分依次估计了下述三个方程：

$$\text{方程 I: } \ln Y = \alpha + \theta \times Urb \quad (3)$$

$$\text{方程 II: } \ln Y = \alpha + \theta \times Urb + \beta \times X \quad (4)$$

$$\text{方程 III: } \ln Y = \alpha + \theta \times Urb + \beta \times X + \delta \times (Urb \times X) \quad (5)$$

其中， Urb 表示城镇， X 中所包含的变量有户主年龄、户主受教育程度、家庭劳动年龄（16~65 岁）人口的性别结构和平均受教育年限、家庭规模以及老年人口与劳动年龄人口在家庭成员数量中的比率、所在地区（东部与西部）^①。被解释变量为收入对数，其中收入变量没有进行生活费用指数调整。分城乡及总体样本的相关变量均值可见附表 1。两个年份中方程 II 和方程 III 的估计结果可见附表 2。

基于这三个方程中城镇变量 Urb 的估计系数，可以推算出在控制其他因素的情形下，城乡因素在总体样本中对于收入及其不均等程度的影响。其中，城镇居民的收入

① 需要说明的是，这里的解释变量着重考虑了人力资本特征在收入决定中的作用。一般说来，估计农村居民收入函数时，通常会将生产性固定资产和土地等生产要素投入也作为解释变量，但本文的收入函数回归分析中没有包括它们。主要有两个方面的考虑，一是资本投入和土地投入与当期产出之间具有非常强的内生性，劳动力的人力资本特征对于决定这些投入可能具有重要影响；二是这些影响因素不具有城乡可比性。

优势可以根据 θ 的估计系数推算得到；在控制其他因素的情形下，城乡差距对于总体差距的解释程度可以通过 G·Fields 分解来反映。在收入函数回归结果的基础上，某个因素对于总体不平等程度的解释程度，依照 G·Fields 分解可以表示为：

$$\pi_j = \frac{\hat{\beta}_j \text{cov}(x_j, \ln Y)}{\text{var}(\ln Y)} \quad (6)$$

其中， x_j 为收入函数中的解释变量， $\hat{\beta}_j$ 为相应的估计系数。此外，根据 G·Fields 分解，因素 j 对于不同时期不均等程度变化的贡献份额可以分解为：

$$\frac{\pi_j(t)I(t) - \pi_j(0)I(0)}{I(t) - I(0)} \quad (7)$$

其中， $\pi_j(t)$ 表示变量 j 在时期 t 的 G·Fields 分解份额， $I(t)$ 和 $I(0)$ 分别为时期 t 和时期 0 所对应的不平等指数。

表 8 给出了上述三个方程中城镇变量 Urb 的估计系数及其收入差距含义。如果没有任何控制变量，根据 Urb 的估计系数，推算 2007 年的城镇居民人均收入比农村居民要高出 2.6 倍^①，2013 年有所下降，约为 2.1 倍。按照 G·Fields 分解，城乡差距对于总体不平等程度的贡献份额在 2007 年和 2013 年分别为 37.56% 和 33.04%，也表现出下降趋势。这一贡献份额在两个年份中都要略低于相应年份 GE 指数分解的组间差距贡献份额。

表 8 “城乡”对总体差距的解释

		方程 I	方程 II	方程 III
2007 年	θ 估计值	1.2797	0.7570	0.4621
	exp(θ) - 1 (%)	259.56	113.19	58.74
	G·Fields 分解 (%)	37.56	22.22	13.56 (24.61)
2013 年	θ 估计值	1.1255	0.6929	0.1518
	exp(θ) - 1 (%)	208.18	99.95	16.39
	G·Fields 分解 (%)	33.04	20.34	4.45 (19.06)

注：括号内为所有含有 Urb 的变量 ($Urb \times X$) 的解释份额之和。

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

在方程 II 中，也就是在收入函数估计中控制相应变量并假定城乡人群具有相同的估计系数， θ 的估计值有明显下降，此时城镇变量 Urb 对总体差距的解释程度分别下降

① 基于回归系数推算的城乡收入差距比率要高于表 2 的结果，这有两个方面的原因，一是回归分析中假定“其他条件保持不变”，二是由于因变量设定为对数收入，因此所获得的估计系数度量的是几何平均数差异。

至 22%（2007 年）或 20%（2013 年）。方程 I 与方程 II 中，城镇变量 *Urb* 对总体差距的解释程度的差异可以认为是由于城乡人群组之间具有不同的特征所致。收入函数中所控制的变量的特征差异可以解释城乡差距的 15%（2007 年）或 13%（2013 年）。在方程 III 中，相同解释变量在城镇和农村人群中可能具有不同的估计系数。在这种情形下，收入决定的城乡差异不仅仅体现在要素禀赋（解释变量）的差异，也表现在要素回报（估计系数）的差异。“城乡”因素对于总体差距的解释程度也考虑两种不同情形。一种情形是，与前面的讨论相同，仍只关注“城镇”虚拟变量的解释作用。“城镇”虚拟变量在 2007 年的 G·Fields 分解下降至 13.56%，2013 年进一步降至 4.45%。因此，如果控制城镇与农村收入函数中回归系数的差异，“城乡”因素在 2013 年已经变得不再重要。另一种情形是，如果把系数回报的差异也看作是城乡差异的一部分，也就是说，“城乡”差异既表现在“城镇”虚拟变量，也表现在收入函数估计系数的城乡差异，则“城乡”对于总体差距的解释程度在 2007 年和 2013 年仍分别为 24.61% 和 19.06%，相应份额与基于方程 II 的估计结果在相同年份的分解结果是相近的。比较这两种情形的差异可以发现，收入函数估计系数的城乡差异在 2007 年解释了总体差距的 11%，2013 年这一份额上升到了将近 15%。这意味着城乡差距中，基于要素回报特征差异的解释份额在此期间有所上升。

基于附表 2 估计的收入函数，对于各年份总体收入不平等程度以及两个年份中的基尼系数变化，各因素的 G·Fields 分解结果可见表 9。除了随机扰动项（不可解释因素）以外，“城镇”变量对于根据两个年份中方程 II 和 2007 年方程 III 的总体差距分解中，解释份额是最高的。无论是基于方程 II 还是方程 III 的分解结果，“城镇”变量对于两个年份中基尼系数变动的解释份额也要高于其他解释因素。从表 9 的分解结果来看，“城镇”变量对于两个年份收入差距的变动具有最为重要的解释作用；在多数情形下，“城镇”变量也是当年收入差距最为重要的解释因素。地区因素（东、中、西部地区）虽然对于相应年份收入差距的解释份额并不高，但对于两个年份中基尼系数的下降却具有比较高的解释份额，仅次于“城镇”变量。在基于方程 II 的分解结果中，地区因素可以解释基尼系数下降的 63%；在方程 III 的分解结果中，（农村）地区因素解释了基尼系数下降的 57%，但城镇与地区的交叉项对于基尼系数变动的解释方向则是相反的，也就是说相对于农村而言，地区之间的城镇居民收入差距仍是上升的，但两类因素总体合计仍具有缩小全国基尼系数的效应。地区因素对于农村居民收入差距的缩小，与农村劳动力的跨地区流动具有逻辑上的一致性。总体而言，从 2007 年到 2013 年，全国基尼系数的下降主要是由城乡差距和地区差距的缩小所导致的。

表9 总体样本不平等程度的 G·Fields 分解结果

单位：%

	基于方程 II			基于方程 III		
	2007 年	2013 年	基尼变化	2007 年	2013 年	基尼变化
城镇	22.22	20.34	82.98	13.56	4.45	308.01
教育	10.23	10.35	6.35	8.32	11.13	-82.50
家庭结构	10.23	11.75	-38.90	9.62	11.79	-60.52
地区	5.49	3.70	63.35	6.04	4.46	57.11
城镇教育				4.01	-0.78	158.83
城镇家庭结构				7.50	14.98	-234.27
城镇地区				-0.46	0.41	-28.58
随机扰动项	51.84	53.86	-13.45	51.41	53.57	-18.40

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

教育（包括户主教育程度和家庭劳动年龄人口的平均教育年限）和家庭结构（年龄结构与家庭规模）对于当年收入差距都具有重要的解释作用，但对于两个年份基尼系数的变化则具有不同的影响。教育对于基尼系数的下降具有正向解释作用，尽管在方程 II 的分解结果中可以看到这种份额并不高，只有 6.35%。从方程 III 的分解结果中可以看出，“城镇”与教育的交叉因素对于基尼系数下降的贡献份额达 158.83%。教育对于收入差距的缩小主要来自于城镇^①。家庭结构对于收入差距的解释作用是在上升的，并且从对基尼系数变化的影响来看，具有扩大收入差距的作用。

五 基于收入分布的城乡差距变化特征

为了讨论收入分布不同分位点处城乡差距因素所具有的不同影响，本部分的讨论基于收入分布无条件分位回归（unconditional quantile regression, UQR）结果。无条件分位回归是 Firpo et al. (2007, 2009) 在再中心化影响函数（recentered influence function, RIF）的基础上给出的一种回归方法。相对于普通最小二乘（OLS）回归，

① 从附表 2 可以看到，城镇与户主大学受教育程度的交叉变量在 2013 年有非常大幅度的下降，因此城镇收入决定中教育因素对收入差距的缩小作用主要来自于高等教育回报率的下降。此外，在方程 III 的估计结果中，无论是否与“城镇”变量交叉，劳动年龄人口平均受教育年限的回归系数从 2007 年到 2013 年期间也是下降的。

RIF 回归能给出相应解释变量在因变量不同分布分位点上的边际效应。

收入分布分位点处再中心化影响函数定义为：

$$RIF(Y; q_\tau) = q_\tau + \frac{\tau - I(Y \leq q_\tau)}{f_Y(q_\tau)} \quad (8)$$

其中 q_τ 满足 $P(Y \leq q_\tau) = \tau$ ($0 < \tau < 1$)， $P(\cdot)$ 为概率分布； $I(\cdot)$ 为指标函数，满足条件取值为 1，否则取 0； f_Y 为 Y 的边际密度函数。由 τ 和指标函数 $I(\cdot)$ 的定义可知 $E(RIF(Y; q_\tau)) = q_\tau$ 。如果 $E(RIF(Y; q_\tau) | X) = X\beta$ ，则 $q_\tau(Y) = E(X)\beta_\tau$ ， β_τ 的含义为，解释变量对分位点 τ 处收入水平的边际效应。基于分位回归结果，根据 Oaxaca – Blinder 分解思路（O – B 分解），城镇与农村居民收入分布第 τ 分位点 $q_\tau(\ln Y)$ 之间的差距可以分解为：

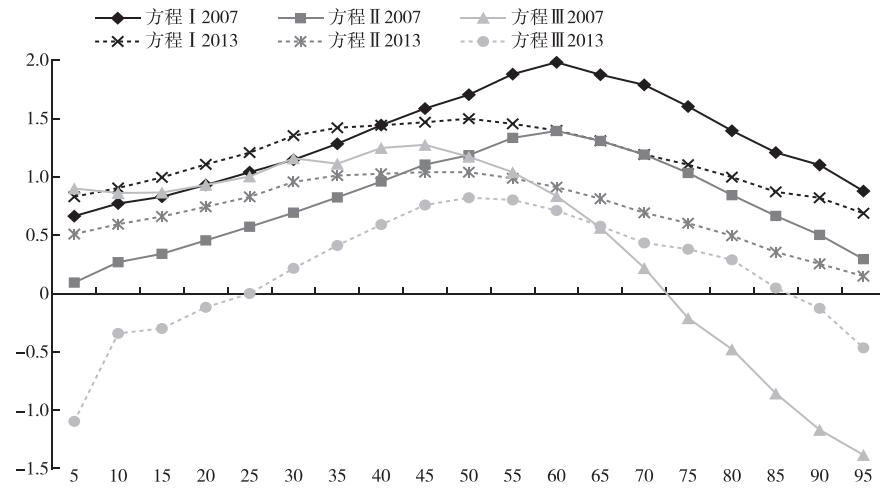
$$q_\tau(\ln Y_M) - q_\tau(\ln Y_F) = (X_M - X_F) \frac{(\beta_M + \beta_F)}{2} + \frac{(X_M + X_F)}{2}(\beta_M - \beta_F) \quad (9)$$

这一等式右边第一项为变量（禀赋）特征变动的效应，第二项则为系数回报变化对给定收入组人群收入的影响。

按照方程 I、II 和 III 的基本设定形式，采用 RIF 回归得到收入分布各分位点上“城镇”变量所对应的估计系数，如图 5 所示。从中可以看出，“城镇”变量的估计系数随收入分布分位点上升而呈倒 U 型变化特征。无论在收入函数中是否控制其他影响因素，这种倒 U 型关系都成立。在城乡样本合并组成的全部人口收入分布的低端与高端中，“城乡”因素对于相应收入的影响程度相对是比较低的；而在收入分布的中间部分，“城镇”变量的估计系数则是比较高的。

为了讨论城乡之间收入差距的影响因素，在附表 3 所估计的分城乡收入函数的基础上，分别对两个年份的城乡收入差距进行了分解，表 10 给出了相应的分解结果，包括基于 OLS 的均值分解和在不同收入分布分位点上的分解结果。

在基于均值的 O-B 分解中，城乡收入差距中的系数效应和禀赋效应在两个年份中的相对贡献份额都是比较接近的，其中禀赋效应的解释份额在 40% 左右，也就是说城镇和农村居民人均收入水平差异中的 40% 可以由收入函数中所包含解释变量的特征差异来解释，其中教育和家庭结构特征分别贡献 21% 和 15% ~ 16% 左右。这种贡献份额在两个年份的均值分解中也是比较稳定的，但两个年份系数贡献份额则有所不同。2007 年的系数效应主要是由常数项所导致的，而 2013 年则主要是由家庭结构特征所造成的。常数项的贡献份额在 2013 年比 2007 年有较大幅度的下降，对城乡收入差距的解释程度从 35.76% 下降到 13.70%。

图 5 *Urb* 在各分位点的估计系数

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

表 10 Oaxaca-Blinder 分解

	均值分解		10% 分位点		50% 分位点		90% 分位点	
	系数	禀赋	系数	禀赋	系数	禀赋	系数	禀赋
2007 年								
合计	0.7620 [59.40]	0.5209 [40.60]	0.8519 [61.96]	0.5230 [38.04]	0.7874 [61.92]	0.4843 [38.08]	0.6608 [55.59]	0.5279 [44.41]
教育	0.0810 [6.32]	0.2631 [20.51]	0.0147 [1.07]	0.2638 [19.19]	0.1112 [8.75]	0.2354 [18.51]	-0.0205 [-1.72]	0.2797 [23.53]
家庭结构	0.2187 [17.05]	0.2071 [16.14]	0.3536 [25.72]	0.2194 [15.96]	0.0912 [7.17]	0.2023 [15.91]	0.1990 [16.74]	0.1910 [16.06]
地区	0.0035 [0.27]	0.0507 [3.96]	0.0455 [3.31]	0.0398 [2.90]	0.0007 [0.05]	0.0466 [3.67]	-0.0439 [-3.70]	0.0573 [4.82]
常数项	0.4588 [35.76]		0.4381 [31.86]		0.5843 [45.95]		0.5262 [44.27]	
2013 年								
合计	0.6782 [60.04]	0.4514 [39.96]	0.8128 [65.06]	0.4364 [34.94]	0.7009 [62.09]	0.4280 [37.91]	0.5090 [50.11]	0.5068 [49.89]
教育	-0.0261 [-2.31]	0.2461 [21.79]	0.1301 [10.41]	0.2031 [16.26]	-0.0566 [-5.02]	0.2213 [19.60]	-0.2494 [-24.55]	0.3323 [32.71]
家庭结构	0.5149 [45.58]	0.1692 [14.98]	0.5725 [45.83]	0.2088 [16.71]	0.5905 [52.31]	0.1724 [15.28]	0.6708 [66.04]	0.1335 [13.14]
地区	0.0347 [3.07]	0.0360 [3.19]	0.0941 [7.53]	0.0245 [1.96]	0.0417 [3.69]	0.0343 [3.04]	0.0052 [0.51]	0.0411 [4.04]
常数项	0.1548 [13.70]		0.0162 [1.29]		0.1253 [11.10]		0.0824 [8.11]	

注：方括号内数字表示百分比。

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

对于收入分布不同分位点城乡居民收入差距，表 10 给出了 10%、50% 和 90% 三种情形下的分解结果。图 4 的结果表明，城镇与农村收入分布的低分位点上，收入差距要相对高于高分位点。表 10 和图 6 的结果则进一步表明，禀赋效应对收入分布低分位点城乡差距的解释份额相对比较低。从表 10 中可以看出，在 10% 和 50% 分位点上，2007 年的禀赋效应解释份额都为 38%，90% 分位点上禀赋效应解释份额上升到 44%；2013 年 10% 分位点上的禀赋效应解释份额为 35%，比 2007 年相应分位点解释份额略有下降，但中位数（50% 分位点）上禀赋效应解释份额与 2007 年相近，高分位点（90%）上的禀赋效应则比 2007 年增加了 5 个百分点。图 6 给出了禀赋效应在各分位点城乡收入差距上的解释份额。尽管这两条曲线总体上都表现出上升倾向，但 2007 年城乡收入分布 80% 分位点以下的差距中，禀赋效应的解释份额都比较平稳，这一部分的曲线比较平坦，而在 80% 分位点以后禀赋效应的解释份额明显上升；2013 年的曲线则表明，禀赋效应的解释份额随着收入分布分位点的上升而表现出更为严格的递增倾向。在收入分布低分位点（60% 以下部分），禀赋效应在 2007 年的解释份额要高于 2013 年，而此后两者关系则相反，即高分位点中，禀赋效应的解释份额在上升。

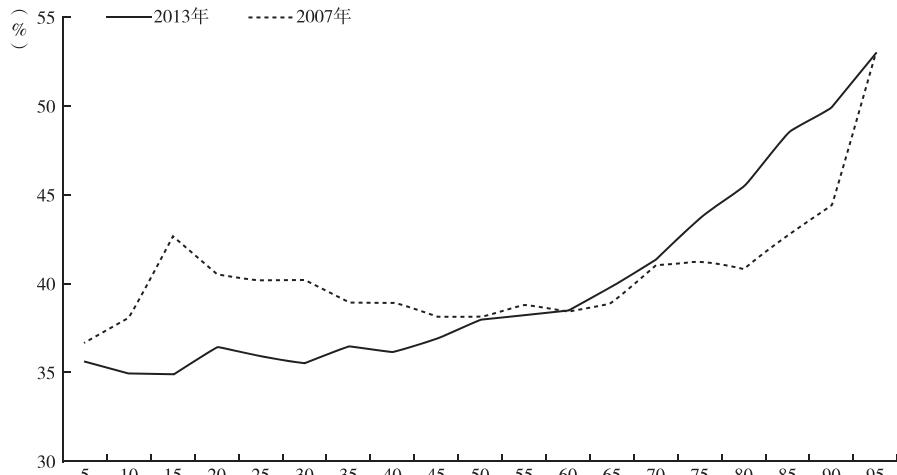


图 6 各分位点城乡差距的禀赋效应

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

就更为具体的解释因素而言，一个非常明显的特征是，2007 年各分位点上常数项的解释份额都是非常高的，而在 2013 年则出现了非常大幅度的下降。如在 10% 分

位点上，常数项的解释份额在 2007 年高达 32%，而 2013 年则只有 1%。与均值分解结果类似，两个年份系数效应解释份额的变化主要表现在家庭结构因素和常数项。2007 年，系数效应主要由常数项贡献；2013 年，家庭结构是系数效应贡献份额的主要来源。

在禀赋效应中，教育和家庭结构特征都具有比较重要的解释份额。这里主要关注教育因素的变化特征。图 7 给出了教育在各分位点城乡差距中禀赋效应的解释份额，其基本特征与图 6 非常相似。2007 年，教育禀赋效应的解释份额在收入分布的绝大多数范围内都是比较稳定的，大约在 20% 左右，但在收入分布的高端，教育禀赋效应的解释份额有较大幅度的上升。2013 年教育禀赋效应对城乡收入差距的解释份额则随着分位点的上升而明显递增，其解释份额从 5% 分位点上的 15% 上升到 95% 分位点上的 35%。事实上，不仅教育禀赋效应的相对贡献份额有较大幅度的提升，从表 10 分解中所给出的绝对贡献也表现出类似的特征。2013 年，90% 分位点上教育禀赋效应大约可以解释城乡收入差距的 33%，而在 10% 和 50% 分位点上这一比率分别为 16% 和 20%，也高于 2007 年 90% 分位点上的解释效应，其比率不到 24%。总体而言，在 2013 年收入分布高端人群中，禀赋特征差异对于城乡收入差距的解释作用越来越大，其中教育禀赋效应的变化特征尤其如此。

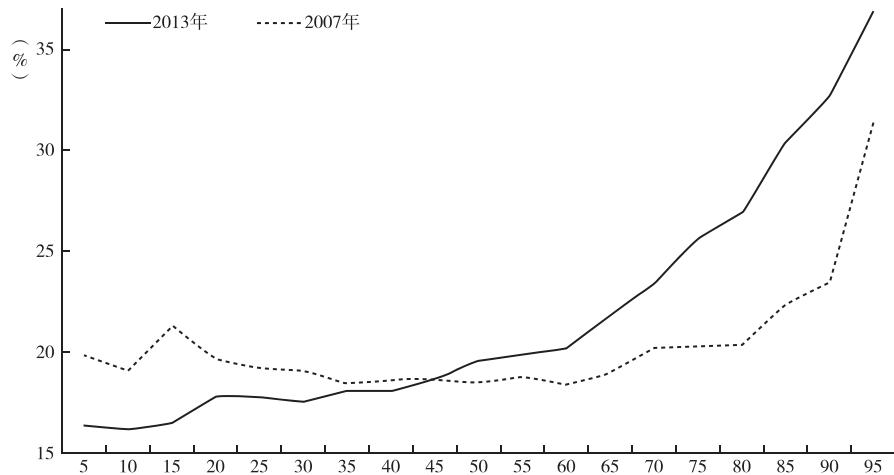


图 7 各分位点城乡差距的教育禀赋效应

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

六 总结

本文在 CHIP2007 年和 2013 年全国住户调查数据的基础上，讨论了这一时期城乡居民收入差距的变化及其对于全国收入差距的影响。本文的主要发现表明，从人群组分解来看，城乡之间收入差距的缩小和农村人口份额的下降是导致这一时期全国居民收入差距缩小的主要原因，但城乡之间的组间不平等对于当年的全国收入差距仍具有非常高的解释份额，因此缩小城乡差距对于缩小总体收入差距仍具有非常重要的作用。回归分析以及基于回归分析基础上的不平等指数分解也表明，城乡差距在总体收入差距中具有重要地位，即便对不同地区之间的生活费用指数进行相应调整，城乡差距依然是非常重要的。从政策含义而言，这一结果表明，缩小城乡差距、加快城镇化进程对于缩小全国收入差距具有非常重要的作用。

在基于收入函数的不平等指数分解中，控制教育、家庭结构、地区等因素后，城乡差距对于总体收入差距仍具有非常高的解释份额，也是这一时期收入基尼系数变化最为关键的解释因素，但收入决定机制中的城乡差距在相当大程度上是通过上述解释因素在城镇和农村收入决定中具有不同的系数效应导致的，也就是说收入决定机制依然存在着非常严重的城乡分割。在 Oaxaca-Blinder 分解中，出现了一个非常有意思的现象是，家庭人口结构对于城乡收入差距的解释作用出现了大幅度的提高，特别是其系数效应在 2013 年的解释份额比 2007 年上升了 20 个百分点以上，在收入分布 90% 分位点的城乡差距中，其解释份额甚至上升到 66%。一个可能的解释是，人口政策可能导致了城乡人口结构差异，并由此改变了劳动力市场上的要素回报特征。对这一现象仍需要进一步更为深入的研究。

从基于收入分布的分解分析来看，尽管城乡收入差距随着收入分布分位点的上升而下降，但禀赋效应对于城乡收入差距的解释作用随着收入分布分位点的上升而上升，教育的禀赋效应尤其如此。也就是说，对于城乡收入分布中相对高端的人群而言，教育程度差异对于城乡差距具有非常重要的解释作用。特别是，从相应的 Oaxaca-Blinder 分解结果来看，教育的系数效应在高收入人群的城乡差距中的解释作用为负。这意味着，农村高收入人群教育回报状况可能要好于城镇高收入人群。这可能与农村劳动力市场上人力资本的总体稀缺状况相关联。因此，提高农村居民教育程度仍将是缩小城乡收入差距的重要选择。

附表1 变量均值

	2007年			2013年		
	城镇	农村	合计	城镇	农村	合计
家庭人均收入对数	9.42	8.14	8.64	10.05	8.92	9.45
户主年龄	48.84	49.01	48.94	50.45	52.07	51.31
户主教育程度：大学	0.3216	0.0067	0.1292	0.2876	0.0105	0.1413
户主教育程度：高中	0.3534	0.1652	0.2384	0.2919	0.1131	0.1976
户主教育程度：初中	0.2603	0.4930	0.4025	0.3028	0.4859	0.3995
劳动年龄人口平均教育年限	10.93	7.48	8.83	10.03	7.41	8.65
劳动年龄人口性别比率	0.4658	0.5068	0.4908	0.4362	0.4728	0.4555
16~65岁成员比率	0.7949	0.7803	0.7860	0.7236	0.7148	0.7190
65岁以上成员比率	0.0780	0.0532	0.0628	0.1075	0.0901	0.0983
家庭人口规模	3.15	4.45	3.94	3.36	4.30	3.86
西部地区	0.2418	0.3252	0.2927	0.2361	0.3191	0.2799
东部地区	0.4233	0.3280	0.3651	0.4415	0.3345	0.3850

资料来源：根据CHIP2007年和2013年数据计算得到。

附表2 全部人群收入函数估计结果

	2007年		2013年		2007年		2013年	
	系数	t值	系数	t值	系数	t值	系数	t值
城镇	0.7570	131.66	0.6929	107.16	0.4621	11.57	0.1518	3.66
户主年龄	0.0019	8.26	0.0029	10.32	-0.0007	-2.44	-0.0003	-0.70
户主教育程度：大学	0.2699	26.38	0.4850	40.22	0.0633	1.98	0.4703	12.98
户主教育程度：高中	0.0643	8.84	0.2127	23.04	0.0036	0.40	0.1085	8.18
户主教育程度：初中	0.0154	2.72	0.1263	17.73	-0.0192	-2.99	0.0697	7.97
劳动年龄人口平均教育年限	0.0511	48.16	0.0307	27.47	0.0571	38.73	0.0397	22.81
劳动年龄人口男性比率	-0.1275	-9.67	-0.2256	-13.91	-0.1330	-7.64	-0.3130	-14.05
16~65岁成员比率	0.3624	26.81	-0.0073	-0.48	0.3994	24.18	-0.0567	-2.69
65岁以上成员比率	0.4891	25.28	-0.1034	-6.01	0.1920	7.89	-0.3562	-14.45
家庭人口规模	-0.1388	-81.19	-0.1789	-82.25	-0.1325	-69.83	-0.1823	-65.84
西部地区	-0.1670	-33.24	-0.0074	-1.06	-0.1940	-31.16	-0.0797	-8.83
东部地区	0.3673	77.95	0.3092	49.93	0.3986	64.85	0.3310	38.07
城镇×户主年龄					0.0046	9.40	0.0058	10.20
城镇×户主教育程度：大学					0.2601	7.21	0.0781	1.96
城镇×户主教育程度：高中					0.1191	6.88	0.1592	8.17
城镇×户主教育程度：初中					0.0645	4.13	0.0631	3.95
城镇×劳动年龄人口平均教育年限					-0.0021	-0.95	-0.0110	-4.85
城镇×劳动年龄人口男性比率					0.1119	4.19	0.2213	6.83

续表

	2007 年		2013 年		2007 年		2013 年	
	系数	t 值	系数	t 值	系数	t 值	系数	t 值
城镇 \times 16 ~ 65 岁成员比率					-0.0225	-0.78	0.0851	2.81
城镇 \times 65 岁以上成员比率					0.5895	14.04	0.4572	13.25
城镇 \times 家庭人口规模					-0.0214	-4.90	0.0044	0.98
城镇 \times 西部地区					0.0955	9.14	0.1776	12.71
城镇 \times 东部地区					-0.0635	-6.67	-0.0383	-3.11
常数项	7.9509	438.95	9.2422	449.92	8.0242	359.18	9.5099	346.62
样本数	80809		59404		80809		59404	
F 统计量	11027.88		5548.20		5885.72		2978.25	
调整 R ²	0.6209		0.5284		0.6262		0.5355	

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

附表 3 分城乡收入函数回归结果

	OLS		10% 分位点		50% 分位点		90% 分位点	
	城镇	农村	城镇	农村	城镇	农村	城镇	农村
2007 年								
户主年龄	0.0038 ***	-0.0007 **	0.0070 ***	0.0046 ***	0.0031 ***	-0.0010 ***	0.0025 ***	-0.0007
户主教育程度：大学	0.3234 ***	0.0633 *	0.3815 ***	-0.3251 ***	0.3184 ***	0.0651 *	0.3056 ***	0.2734 ***
户主教育程度：高中	0.1227 ***	0.0036	0.2222 ***	-0.0392 **	0.1357 ***	-0.0091	-0.0164	0.0522 ***
户主教育程度：初中	0.0452 ***	-0.0192 ***	0.0987 ***	-0.0125	0.0788 ***	-0.0163 **	-0.0429 **	-0.0335 ***
劳动年龄人口平均教育年限	0.0551 ***	0.0571 ***	0.0603 ***	0.0832 ***	0.0491 ***	0.0495 ***	0.0508 ***	0.0513 ***
劳动年龄人口男性比率	-0.0211	-0.1330 ***	-0.0317	-0.3277 ***	-0.0046	-0.0980 ***	-0.0612 *	-0.0995 ***
16 ~ 65 岁成员比率	0.3769 ***	0.3994 ***	0.3497 ***	0.3052 ***	0.3882 ***	0.4478 ***	0.3118 ***	0.3221 ***
65 岁以上成员比率	0.7816 ***	0.1920 ***	0.7423 ***	-0.0185	0.7511 ***	0.2420 ***	0.7142 ***	0.0572
家庭人口规模	-0.1539 ***	-0.1325 ***	-0.1522 ***	-0.1541 ***	-0.1587 ***	-0.1210 ***	-0.1350 ***	-0.1320 ***
西部地区	-0.0985 ***	-0.1940 ***	-0.1934 ***	-0.3493 ***	-0.1090 ***	-0.1879 ***	0.0030	-0.0310 ***
东部地区	0.3350 ***	0.3986 ***	0.1769 ***	0.1734 ***	0.3227 ***	0.3807 ***	0.5067 ***	0.6497 ***

续表

	OLS		10% 分位点		50% 分位点		90% 分位点	
	城镇	农村	城镇	农村	城镇	农村	城镇	农村
2007 年								
常数项	8.4863 ***	8.0242 ***	7.5418 ***	7.1038 ***	8.5873 ***	8.0030 ***	9.3573 ***	8.8310 ***
样本数	29222	51587	29222	51587	29222	51587	29222	51587
F 统计量	1155.09	1922.70	188.00	265.57	1010.94	1524.95	272.02	541.10
调整 R ²	0.3029	0.2907	0.0961	0.0924	0.2082	0.1938	0.1198	0.1328
2013 年								
户主年龄	0.0055 ***	-0.0003 ***	0.0049 ***	-0.0020 **	0.0064 ***	-0.0006	0.0062 ***	0.0022 ***
户主教育程度：大学	0.5484 ***	0.4703 ***	0.5758 ***	0.1442 ***	0.5329 ***	0.2960 ***	0.5730 ***	1.0098 ***
户主教育程度：高中	0.2677 ***	0.1085 ***	0.4144 ***	0.0276	0.2626 ***	0.1200 ***	0.1351 ***	0.1695 ***
户主教育程度：初中	0.1328 ***	0.0697 ***	0.2483 ***	0.0497 ***	0.1055 ***	0.0772 ***	0.0486 ***	0.0415 ***
劳动年龄人口平均教育年限	0.0287 ***	0.0397 ***	0.0294 ***	0.0398 ***	0.0262 ***	0.0414 ***	0.0253 ***	0.0459 ***
劳动年龄人口男性比率	-0.0917 ***	-0.3130 ***	0.1060 **	-0.2852 ***	-0.0996 ***	-0.3367 ***	-0.0945 **	-0.2744 ***
16~65 岁成员比率	0.0285	-0.0567 ***	0.0368	-0.1720 ***	-0.0052	-0.0260	0.0028	-0.1748 ***
65 岁以上成员比率	0.1011 ***	-0.3562 ***	0.1809 ***	-0.3366 ***	0.0483	-0.3955 ***	0.0648	-0.4020 ***
家庭人口规模	-0.1779 ***	-0.1823 ***	-0.2459 ***	-0.2036 ***	-0.1763 ***	-0.1929 ***	-0.1200 ***	-0.1741 ***
西部地区	0.0980 ***	-0.0797 ***	0.0902 ***	-0.1993 ***	0.0981 ***	-0.1269 ***	0.0541 ***	0.0273 **
东部地区	0.2927 ***	0.3310 ***	0.2040 ***	0.1684 ***	0.2812 ***	0.3345 ***	0.4108 ***	0.4166 ***
常数项	9.6617 ***	9.5099 ***	8.9433 ***	8.9271 ***	9.7117 ***	9.5864 ***	10.3300 ***	10.2477 ***
样本数	20263	39141	20263	39141	20263	39141	20263	39141
F 统计量	820.33	1026.93	170.32	183.36	741.21	1074.83	165.17	249.96
调整 R ²	0.3079	0.2238	0.1185	0.0593	0.2186	0.1814	0.1339	0.0913

注：***、**、* 分别表示在 1%、5% 和 10% 的水平上统计显著。

资料来源：根据 CHIP2007 年和 2013 年数据计算得到。

参考文献：

- 别雍·古斯塔夫森、李实（1999），《中国变得更加不均等吗？》，载于赵人伟、李实、卡尔·李思勤主编《中国居民收入分配再研究》，北京：中国财政经济出版社。
- 蔡昉、杨涛（2000），《城乡收入差距的政治经济学》，《中国社会科学》第 4 期，第 11 – 22 页。
- 李实（2003），《中国个人收入分配研究回顾与展望》，《经济学（季刊）》第 2 期，第 379 – 404 页。
- 李实、佐藤宏、史泰丽（2013），《近期中国收入差距变化的主要特点》，载于李实、佐藤宏、史泰丽主编《中国收入差距变动分析——中国居民收入分配研究 IV》，北京：人民出版社。
- 罗楚亮（2006），《城乡居民收入差距的动态演变：1988 ~ 2002 年》，《财经研究》第 9 期，第 103 – 112 页。
- 徐宽（2003），《基尼系数的研究文献在过去八十年是如何拓展的》，《经济学（季刊）》第 3 期，第 757 – 778 页。
- Brandt, Loren & Carsten Holz (2006). Spatial Price Differences in China: Estimates and Implications. *Economic Development and Cultural Change*, 55(1), 43 – 86.
- Firpo, Sergio, Nicole Fortin & Thomas Lemieux (2007). Decomposing Wage Distributions Using Recentered Influence Function Regressions. University of British Columbia.
- Firpo, Sergio, Nicole Fortin & Thomas Lemieux (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3), 953 – 973.
- Ravallion, Martin & Shaohua Chen (2007). China's (Uneven) Progress Against Poverty. *Journal of Development Economics*, 82(1), 1 – 42.
- Sicular, Terry, Ximing Yue, Björn Gustafsson & Shi Li (2007). The Urban-Rural Income Gap and Inequality in China. *Review of Income and Wealth*, 53(1), 93 – 126.
- Xie, Yu & Xiang Zhou (2014). Income Inequality in Today's China. *Proceedings of the National Academy of Sciences*, 111(19), 6928 – 6933.

Urban-Rural Income Gap and Its Effects on National Income Inequality

Luo Chuliang

(Business School and China Institute for Income Distribution, Beijing Normal University)

Abstract: Based on household surveys conducted by the China Household Income Projects at 2007 and 2013, this paper focuses on the urban-rural income gap and its effects on national income inequality. The inequality decomposition by subpopulations indicates the declined urban-rural income gap and share of rural population reduced the national inequality from 2007 to 2013, while urban-rural income gap still contributes a large proportion of national inequality. Inequality decomposition based on the regressed income functions, with education attainment, demographic structure and regions being controlled, also indicates the urban-rural inequality is the crucial contributor on national inequality (Gini) during the period of 2007 – 2013. By the Oaxaca-Blinder decomposition on quantiles of income distribution, the proportion of endowment effects increases in the higher quantile, especially for education attainment.

Keywords: urban-rural income gap, income inequality, decomposition by subpopulations, decomposition based on regression

JEL Classification: D31, O18, P25

(责任编辑：西 贝)