

· 社会建设研究(学术主持人:谢立中) ·

主持人语:本期“社会建设研究”栏目收录了三篇文章:谢立中的《试析导致当前我国居民收入分配基尼系数数值多样化的主要原因》一文,通过对现有相关文献的梳理,探讨了导致不同研究人员在对我国居民收入分配基尼系数进行计算时得出不同结果的主要原因,以期帮助我们更好地理解现有基尼系数计算结果的含义;田明、孙林的《进城农民工工作流动及其影响因素分析》一文则以问卷调查的资料为依据,分析了当前我国农民工职业流动速度过快的原因,对我们制定与农民工群体相关的政策提供了可资借鉴的信息;崔恒展的《基于唐朝给侍制度的家庭养老支持政策思考》一文则是另辟蹊径,试图借鉴唐朝给侍制度的经验,对我国家庭养老政策提出了有价值的补充建议。这几篇文章的作者具有不同的学科背景,所涉论题之间虽缺乏紧密的逻辑联系,但实际上是从不同方面入手探讨了当前我国社会转型和社会建设过程中的一些重要问题,相信它们的发表会有助于推动人们对当前我国社会转型和社会建设事业的理解和思考。

试析导致当前我国居民 收入分配基尼系数数值多样化的主要原因

谢立中

(北京大学社会学系,北京 100871)

[摘要] 对于“我国当前居民收入分配的基尼系数到底有多大”这个问题,现有相关文献中存在着多种不同的答案。对现有相关文献的梳理表明,导致不同研究人员得出不同基尼系数计算结果的原因是多方面的。由于这些原因的存在,使得我们不应该将我们在相关文献中所看到的任何一个基尼系数数值简单地视为是对某个国家或地区特定时期居民收入分配差距实际状况的客观再现,而应该将其视为该基尼系数的计算者在特定话语系统的引导和约束下所完成的一项有关该国或该地区该时期内居民收入分配差距状况的话语建构。

[关键词] 居民收入分配差距; 基尼系数; 话语建构

[中图分类号]F126.2 **[文献标识码]**A **[文章编号]**1003-4145 [2013]08-0034-10

基尼系数是测量一个国家或地区居民收入分配差距的重要指标。随着改革开放的不断深入,中国社会成员在收入方面的差距逐渐加大,“基尼系数”逐渐开始成为人们热烈讨论的一个重要话题。然而,正如众所周知的那样,对于“我国当前居民收入分配的基尼系数到底有多大”这个问题,不同的研究人员或机构给出了非常不同的答案。

那么,导致不同研究文献作者得出不同基尼系数计算结果的主要原因到底是什么?带着这个问题,笔者对近年来有关我国居民收入分配之基尼系数的诸多研究文献进行了仔细的梳理和分析。结果发现,至少就当前我国居民收入分配基尼系数的研究而言,导致不同研究人员得出不同基尼系数计算结果的原因主要有以下几个方面:1. 不同文献作者所使用的“收入”概念在内涵与外延方面有所不同;2. 数据采集方法和估算

收稿日期:2013-04-07

作者简介:谢立中(1957—),哲学博士,北京大学社会学系教授、博士生导师,主要研究方向为社会学理论、社会发展与现代化、社会政策。

基金项目:本文系北京大学守望社会学基金资助课题“中国当前社会转型所面临的主要问题及出路”的研究成果。

结果上的差别;3.不同文献作者在计算居民收入时所用价格指标不同;4.不同文献作者所用收入分布拟合函数及基尼系数计算方法不同;5.不同研究人员在对居民收入进行入户调查时抽样结果不同。以下简述之。

一、“收入”概念的内涵与外延不同

导致不同研究人员在对当前我国居民收入分配之基尼系数进行计算时得出不同结果的主要原因之一,就是不同研究人员在调查收集我国居民收入分配的状况时,采用了不太相同的“收入”概念。

参照赵人伟等人在《中国居民收入分配研究》等书中的说明和李实、罗楚亮在《中国收入差距究竟有多大?》一文中的介绍,在现有的当前中国收入分配研究文献中所使用的收入定义至少有以下几种:

一是国家统计局主持进行的住户调查中所使用的“城镇可支配收入”或“农村纯收入”定义,也可称为我国的官方收入定义(以下简称定义1)。除了中国社会科学院经济研究所“中国居民收入分配课题组”(CHIP)之外,绝大部分现有中国收入分配问题研究文献采用的都是这一收入定义,因为这些文献使用的收入分配数据都是国家统计局公布的住户调查数据。这一定义下我国居民的收入主要包括以下收入项目:工薪收入、经营性收入、财产性收入和转移性收入等。其中,城镇居民可支配收入具体包括以下项目:(1)工资收入:工资及补贴、其他劳动收入;(2)经营性净收入;(2)财产收入:利息和红利、其他财产租金、出售财物收入;(4)转移性收入:离退休金、救济保险收入、其他;(5)借贷收入:收回借出性收入、贷款收入、其他;(6)其他收入等。农民居民纯收入则具体包括以下项目:(1)农业收入;(2)非农业收入:在本地非企业或企业就业所获工资、外出就业所获工资、非农经营收入;(3)财产性收入:利息、股息、租金、其他;(4)转移性收入:离退休金、救济抚恤金、其他;(5)亲友赠送收入等。

二是美国加州大学河边分校名誉教授卡恩(Khan)的收入定义(以下简称定义2)。简单来说,该收入定义是在国家统计局住户调查收入定义基础上增加了三项收入:一是公有住房的实物性租金补贴,二是私有住房的折算租金(imputed rent),三是各种实物收入(如单位发放的食品、日用品等实物和有价证券)的市场价值。其中,城镇居民的“可支配收入”具体包括以下项目:(1)在职职工的工资性货币收入、离退休人员返聘货币收入、非就业(含离退休和非离退休)人员收入(含退休金、福利收入等);(2)补贴(票证补贴、住房补贴和其他补贴等)和实物收入;(3)自有住房折旧金;(4)个体、私营企业主收入;(5)资产收入(利息、红利、出租房屋或其他物品的收入等);(6)其他收入(如转移收入)。^①农村居民可支配收入具体包括以下项目:(1)家庭生产经营收入(包括出售农业、工业、副业产品所取得的现金收入和自产自销产品的估算收入);(2)个人工资性收入(包括工资、退休金等现金收入,以及单位发放的实物收入);(3)自有房产折旧金;(4)从集体和企业获得的非工资收入(奖金、红利、福利等);(5)财产收入(存款利息、出租土地或其他财产所得收入等);(6)净补贴(救济金和以市场价格计算的救济粮等);(7)其他收入(如转移收入)等。^②中国社会科学院经济研究所中国居民收入分配课题组对我国城乡居民进行的收入分配入户调查,都是按这一收入定义来进行的,课题组成员所完成的大部分研究论文自然也都使用了这一收入定义。^③

三是李实、罗楚亮等人在《中国收入差距究竟有多大?》一文中所建议使用的收入定义(以下简称定义3)。李实、罗楚亮称这一定义为“福祉含义的收入定义”,它是在卡恩的收入定义上进一步做两方面的调整:一是增加了给城乡居民带来实际福祉的社会保障和社会福利的市场价值,二是将所有的名义收入折算成统一可比的实际收入。^④

实际上,除了上述三种收入定义之外,在现有研究我国收入分配差距的文献中,还有其他一些收入定义

^①赵人伟、格里芬《中国居民收入分配研究》,中国社会科学出版社1994年版,第9页。原文中“补贴和实物收入”一句为“补贴和货币收入”疑为有误。更详细的说明见第45页。

^②赵人伟、格里芬《中国居民收入分配研究》,中国社会科学出版社1994年版,第21页。更详细的说明见第43-44页。

^③赵人伟等人对此有过明确说明“国家统计局住户收入调查所使用的收入定义与我们所用的定义有些不同。国家统计局的城市居民收入主要是货币收入。我们所用的收入还包括了实物收入和根据市场价格估算的住房补贴在内的各种补贴。我们与国家统计局在农村居民收入定义上的不同主要是,我们的收入定义包括了农村居民自有住房的折旧金;此外,我们对自我消费的那部分农产品的价值是按市场价格估算的。”见赵人伟、格里芬《中国居民收入分配研究》,中国社会科学出版社1994年版,第6页,注释2。

^④卡恩等人的补充说明“虽然我们的收入定义比国家统计局运用的定义更全面,但重要的是我们的定义也有局限性。它主要表现在衡量农村生产活动的收入方面,我们已经注意到了这样一个问题:我们的方法不能直接估算来自生产活动的农业净收入的部门构成情况。另一个问题是没有考虑到存量的变化(销售产品的价值和自我消费产品的价值相加,得到生产性收入。这就是说,我们假设库存保持不变)。第三个问题是忽略了作为成本因素的资本的折旧问题。”赵人伟、格里芬《中国居民收入分配研究》,中国社会科学出版社1994年版,第46-47页。

(以下简称定义4)。例如,古斯塔夫森、李实等人在《中国收入不平等及其地区差异》一文中就使用过一个不同于上述三个定义的收入定义,即在国家统计局住户调查收入定义的基础上加上住房补贴和自有住房折算租金(但没有包括实物收入,也没有将名义收入折算为实际购买力)。^①

收入数据估算方面的差异对居民收入分配差距譬如基尼系数的测算可能会有很大影响。当然,收入数据估算方面的差异并非必然导致基尼系数等收入分配差距计算方面的差异。如果两个收入定义之间相互有别的那些收入项目(或者说:定义A包含但定义B不包含的那些收入项目,例如CHIP课题组的收入定义中包含但国家统计局的收入定义不包含的“自有房屋租金价值”和“实物补贴”两个项目),其在居民之间的分配状况与两个收入定义(如CHIP课题组的收入定义和国家统计局的收入定义)相互一致的那些收入项目在居民之间的分配状况大体相同,那么收入定义及其数值上的差别并不一定会导致按这两种数据来计算的收入分配差距(如基尼系数)的差异。但是,如果两个收入定义之间相互有别的那些收入项目在居民之间的分配状况与两个收入定义相互一致的那些收入项目在居民之间的分配状况不同,那么按这两种数据来分别计算的基尼系数等收入分配差距指数就一定会不同。^②毫无疑问,前一种可能性更多地只是存在于理论上,在现实生活中,实际存在的情况应该常常是后者。这就是现有相关文献中的那些基尼系数数值为什么会产生差异的主要原因之一。正如CHIP课题后期主持人李实教授所总结的那样“对基尼系数不同的估计结果说明了在估算收入差距时最为重要的是选择合理的收入定义,所谓合理的收入定义是指收入能够更加真实地反映居民的实际福祉。”^③

二、数据估算结果方面的差别

迄今为止,无论国内还是国外,计算家庭人均收入基尼系数所需要的所有收入信息都是通过入户调查时直接向被调查对象进行询问的方式获得的。在入户调查过程中,调查者以问卷形式向被调查对象询问其收入及其结构。在中国,除了国家统计局的住户调查采用了让调查对象进行日常记账的方式来帮助形成问卷所需要的收支数据,以及CHIP课题组的住户调查也在一定程度上利用了国家统计局的调查对象和记账数据外,在很多情况下,调查人员都只是或只能让被调查对象凭自己的记忆和大致估计的收入数据加以应答。

由于以下几方面的原因,这种主要由被调查对象自己对问卷中相关问题的回答得来的居民收入信息会有一定程度甚至相当程度的不一致性:

1. 数据回忆或估计结果的不一致性

首先是数据回忆结果的不一致性这一问题。前面提到,除了国家统计局以及CHIP课题组的入户调查以外,很多有关中国居民收入分配的住户调查(例如西南财大中国家庭金融调查)都是让被调查对象凭自己记忆的收入数据来对相关的问卷项目加以应答。很多人认为,这种依赖被调查对象的记忆来获得的收入数据,与实际情况很难完全相符。岳希明、李实在对西南财大中国家庭金融调查的结果进行评论时就曾经提到这一点。他们明确认为“国家统计局通过日常记账方式收集样本户收入和支出信息,比西南财大采取的一次性回忆的数据收集方法更为准确。”^④尽管对这种看法仍存有争议,但有一点应该是多数人都会认同的,这就是采用凭被调查对象的记忆这种方式来获得收入数据,其唯一性是很难保证的。

其次是数据估计结果的不一致性这一问题。即使是像国家统计局住户调查这样主要借助于被调查对象日常收支记录来进行的家庭收入入户调查,以及CHIP这种大量借助于国家统计局的调查对象和记账数据来进行的家庭收入入户调查,也免不了有许多收入项目无法从住户的日常账目中获得,而只能通过估算的办法来形成相关收入数据。例如,对自有房屋租金的折算就是重要一例。卡恩等人曾经说明过CHIP1988年住户调查中城乡居民自有住房租金的估算方法。按照他们的说明,“城市住房的价值,是把各省每平方米的

^①古斯塔夫森、李实、史泰丽等《中国收入不平等及其地区差异》,载李实等主编《中国居民收入分配研究III》,北京师范大学出版社2008年版,第98页。

^②举例说明:假设收入定义I包含a、b两个项目,他们在家庭x和家庭y之间分配的数量状况均为100元和200元。此时家庭x的总收入为 $100a + 100b = 200$ 元,家庭y的总收入为 $200a + 200b = 400$ 元,家庭x和家庭y之间的收入差距为1:2。现设收入定义II包含a、b、c三个项目(即比收入定义I多了一个项目c),收入项目c在家庭x和家庭y之间的分配状况分别为100元和500元。则此时家庭x的总收入为 $100a + 100b + 100c = 300$ 元,家庭y的总收入为 $200a + 200b + 500c = 900$ 元,家庭x和家庭y之间的收入差距为1:3。

^③李实、史泰丽、古斯塔夫森主编《中国居民收入分配研究III》,北京师范大学出版社2008年版,第6页。

^④岳希明、李实《我们更应该相信谁的基尼系数?》,http://www.ciidbnu.org/news/201301/20130123092800706.html.

平均造价乘以房屋面积间接估算出来的。农村住房的价值是调查户直接估算出来的。”^①类似地,在CHIP1995年进行的调查中,城镇居民“公有住房补贴和自有房屋估算租金价值的估计值是根据被调查人对其住房的市场价值的估算得出的。”卡恩等人明确承认这种估算会有偏差。他们写到“中国城镇的房地产市场是否发达到人们可以得出准确的估计仍是一个值得讨论的问题。主观的估计可能会产生很大的误差。”^②

例如,按照卡恩等人的说明,在CHIP1988年的农村住户调查中,对农村居民人均自有住房折算租金的估算值为73.49元人民币。虽然国家统计局没有将自有住房折算租金列入农村居民收入范围之内而是列入了支出范围之内,但他们在当年的农村住户调查中也对农民自有住房的折算租金有过一个估算,其估算值为71.1元人民币,比CHIP的估算值少了2.39元。因此,即使国家统计局事后对收入定义进行修改,将农民自有住房折旧金计入农民收入范围,他们对农民收入的估算值与CHIP的估算值之间也会有一定差异。卡恩等人认为“这种差异肯定是由于国家统计局和我们估算房租价时运用的方法不同而造成的。”^③

除了自有住房租金估算方面可能产生的差异之外,还有一些其他收入项目也存在着估算差异的问题。卡恩等人回顾说,“1995年,我们CHIP对农村家庭生产性收入的估计值比国家统计局的高出15%。它是14%的农业收入差异和19%的非农收入差异的平均值。而且,我们对家庭生产毛收入的估计结果比国家统计局高出5%,这表明我们之间的估计结果的部分不同是由于生产费用计算方法的差异。”^④卡恩等进一步具体解释道:这可能是因为“我们在问卷中问了包括雇用劳动费用在内的生产总成本,然后将其从毛收入中减掉,也许这会造造成对成本的低估。”^⑤

2. 问卷设计方面的差异

岳希明在对西南财大中国家庭金融调查与研究中心最新公布的《中国家庭收入不平等报告》进行评论时就指出:该中心在进行入户调查时所用问卷的第三部分(“三、家庭成员的工作和收入信息”)调查家庭成员的工资等相关收入(或者说能够界定为每个家庭成员的收入)时,在调查表的抬头给出这样的提示“注意:除受访者和配偶,其他家庭成员只询问A3003及A3006”,而问卷A3003是关于工作性质的问项(A3003该工作的性质是?);A3006询问的是家庭从业人员的行业属性(A3006属于什么行业?)。问卷把家庭中“最了解家里财务状况的人”定为受访者,也就是“调查”中回答调查员问题的家庭成员。根据上面的提示,“调查”并没有搜集除了受访者和配偶之外的收入,这就将低估有三人以上从业人员家庭的收入。不仅如此,如果家庭主要劳力(如夫妻)外出打工,而家中只留下年老的父母和年幼子女的话,该家庭的收入可能几乎为零(可能有少量的农业收入),对于外出务工越老越多的农村住户来说,这一提示的影响会相当严重。^⑥这也意味着,即使是对同一户家庭的入户调查,如果调查者在问卷中对类似问题的设计有差别,那么其所得到的信息也会有差别。

3. 刻意瞒报收入造成的数据差异

被访者可能出于各种动机而刻意瞒报自己的收入。这些动机可能是:第一,高收入者出于财产安全或偷税漏税等方面的动机而不想让别人知道自己的真实收入;第二,某些享受着“最低生活保障金”或“失业救济金”等政府福利待遇的低收入者出于维持这些只有低收入者才能享有的特殊待遇而不想让他人知道自己的真实收入。大多数现有研究文献都已经指出了前一种类型的收入瞒报动机,但对后一种类型的收入瞒报动机大多付诸阙如。第三,被访者因担心泄露不法收入或灰色收入而刻意瞒报收入。

有人曾经以浙江省2003年的统计数据为例来说明入户调查所得居民收入数字(以及依此数字为据计算出来的基尼系数)是如何的不可靠。这位作者首先利用该省2003年统计年鉴公布的城镇居民入户调查分组收入数据和农村居民入户调查分组收入数据,计算了当年浙江省全省居民基尼系数,计算结果为0.3749。然而,当他进一步去考察与城乡居民收入、消费和储蓄有关的一些数据时,却获得了一个意外的发现:2003

①赵人伟、格里芬《中国居民收入分配研究》,中国社会科学出版社1994年版,第52页,注释2。

②赵人伟、李实、卡尔·李思勤《中国居民收入分配再研究》,中国财政经济出版社1999年版,第98页。

③卡恩、格里芬、李思勤等《中国居民户的收入及其分配》,载赵人伟、格里芬《中国居民收入分配研究》,中国社会科学出版社1994年版,第56页。

④赵人伟、李实、卡尔·李思勤《中国居民收入分配再研究》,中国财政经济出版社1999年版,第86页。

⑤赵人伟、李实、卡尔·李思勤《中国居民收入分配再研究》,中国财政经济出版社1999年版,第86页。

⑥岳希明《西南财经大学0.61基尼系数不可信!》,http://www.ciidbnu.org/news/201212/20121228134128706.html。

年该省城镇居民人均收入为 14295 元,人均支出为 13692 元,人均当有节余 603 元;农村居民人均收入为 6709.24 元,人均支出为 5908.32 元,人均当有节余 800.92 元;分别以当年该省城镇人口总数和农村人口总数乘以城、乡各自的人均节余额,该省居民当年节余总额应为 341 亿元人民币。可是,按照同一本统计年鉴公布的数据,该省 2003 年底拥有的城乡居民储蓄总额为 6452.21 亿元,减去 2002 年底该省城乡居民拥有的储蓄余额 5233.73 亿元,可以看到 2003 年该省城乡居民新增储蓄额应是 1218.48 亿元。与“341 亿元”这个数字相比,多了 877.48 亿元。该作者假定,这多出来的 877.48 亿元储蓄款应该都是城镇居民所有(因为按前面的数据该省 2003 年城镇居民人均节余额比农村居民人均节余额还少,作者认为这显然不合理),用当年该省城镇居民总数(1157.5 万)除以 877.48,得到的结果是:2003 年该省城镇人均少计算的收入达 7581 元人民币。这也就意味着,城镇人口在入户调查中人均隐瞒了 7581 元人民币的收入。再假定城镇人口中收入最低的 20% 人口完全没有必要隐瞒自己的收入,因此这些被隐瞒的收入当都为城镇中其他 80% 的人口所有。据此重新计算该省 2003 年全省人口的收入分配基尼系数,结果是高达 0.4783。^①

因此,在对调查对象的收入状况进行调查时,有没有考虑到被调查对象有可能瞒报收入,以及在考虑到这一情况后采用什么样的方法来消除这一问题,都可能会导致所得收入数据的差异。

三、价格指标不同

这里也至少存在着两个问题。

第一个问题是:在把农村居民自产自销的那部分农产品作为收入来进行计算时,到底是以市场价格来进行计算呢,还是以国家确定的合同收购价格进行计算?在 CHIP1988 年进行的农村居民收入调查中,CHIP “仅从生产活动中所得农村收入一项就比国家统计局对各种来源的农村收入的估算高出 3.5%。”^②按照卡恩等人的解释,导致这一差异的“一个原因是衡量自产自销农产品价格的方法不同。”CHIP 的计算是基于当年农产品的市场价格,而国家统计局的计算是基于当年农产品的合同收购价格(包括粮食),因为大部分农产品是由国家按合同收购的。在 1988 年,合同价格低于市场价格。这就使得国家统计局对农业生产收入的计算要少于 CHIP 的计算值。不过,卡恩等人也承认:“在像中国这样一个市场受约束的经济中,很难找到哪种价格‘正确’这个由来已久的问题的理想答案,我们认为使用市场价格比使用完全任意确定的国家收购价格为好。然而,我们必须承认自留的产品与出售的产品很可能在质量上不同,因此运用市场价格也可能造成误差。”^③显然,在到底是应该按照市场价格还是应该按照国家合同收购价格来计算国家合同价存在时期农民的农产品收入这一问题上,人们将难以取得一致意见。这就势必使得持不同观点的人,在对同一户农民家庭的同一收入项目进行计算时,会得出不同的计算结果。

第二个问题是:在计算居民收入时到底是应该按名义收入(即未以各地消费品价格指数调整过的收入数字)来进行呢,还是应该按实际收入(即以各地消费品价格指数调整过的收入数字)来进行?

如果说,前面一个问题更多只是涉及到中国农村居民这样一个特定人群在价格“双轨制”存在的特定期限内收入计算方面的问题的话,那么后面这个问题则是一个普遍性程度更高的问题。

一份世界银行的研究报告显示,在考察中国的收入差距问题时,“如果对城乡之间和地区之间的生活费用差异加以调整,那么全国的收入差距会有不同程度的降低。例如,在不对生活费用差异加以调整的情况下,2001 年全国的基尼系数被估计为 0.447,而在对生活费用差异加以调整以后,该年的基尼系数则降低为 0.395。也就是说,对城乡之间生活费用差异调整后估计出来的城乡居民的实际收入差距并没有超过 0.4。”^④李实等人也“利用加拿大多伦多大学的 Brandt 和 Holz 对中国不同省份城乡生活费用的估计结果,对城乡之间收入差距进行了重新估计”结果发现,“在考虑到不同地区和城乡之间生活费用差异后,2002 年城乡居民的实际收入比值降低到 2.27:1,而该年的城乡之间的名义收入比值为 3.18:1。而且,经过城乡生活费用指

^①胡祖光《基尼系数和统计数据:以浙江省为例》,《浙江社会科学》2005 年第 4 期。

^②卡恩、格里芬、李思勤等《中国居民户的收入及其分配》,载赵人伟、格里芬《中国居民收入分配研究》,中国社会科学出版社 1994 年版,第 49 页。

^③赵人伟、格里芬《中国居民收入分配研究》,中国社会科学出版社 1994 年版,第 50 页。

^④Martin Ravallion, Shaohua Chen, *China's (Uneven) Progress Against Poverty*. World Bank, June 16, 2004. 转引自李实、史泰丽、古斯塔夫森主编《中国居民收入分配研究 III》,北京师范大学出版社 2008 年版,第 14 页。

数调整以后,1995年至2002年期间的城乡之间收入差距的扩大幅度也有所下降,仅仅扩大了3个百分点。”^①此外,李实等人的研究还表明,在考虑公共医疗、养老、教育补贴等社会福利项目后,2002年城镇居民的人均“暗收入”大约为4200元,而农村居民为250元。如果将这部分收入分别记入到城乡居民的平均收入中,那么城乡之间的收入比率会上升到4.35:1,全国的基尼系数将是0.5左右。但如果按照城乡生活费用指数的调整办法,上述估计结果将分别被修正为:城乡之间的收入比率为3.1:1;全国的基尼系数将是0.44。^②

国内也有其他学者试图运用地区之间可比的价格指数来估计由于生活费用差异引起的地区之间收入差距的高估程度。江小涓和李辉以国家发改委价格监测中心网站发布的国内36个城市2005年2月各自的价格数据为依据,对这36个城市该月各自的价格指数进行了计算,然后以这些价格指数对36个城市居民的2003年的人均收入数字进行了调整。结果发现,调整后36个城市居民之间的收入差距明显缩小。例如,按人均名义收入计算时,收入最高的深圳市分别是人均收入最低的西宁市、兰州市和银川市的3.39倍、3.37倍和3.24倍。按人均实际收入水平计算,深圳市与西宁、兰州和银川三市的收入差距分别缩小到1.75倍、1.40倍和1.67倍。^③北京大学中国研究中心宏观组的研究人员也曾经依照购买力平价的思想,编制了我国国内36个城市和25个省、自治区农村地区之间的价格指数,并运用这些价格指数将各地居民的名义收入转换为实际收入,重新测算36个城市之间和25个省、自治区之间收入分配的基尼系数。结果同样显示,中国国内地区之间价格水平差异较大,因此,各地居民实际收入差距小于名义收入差距,以居民实际收入测算的基尼系数较小。例如,36个城市之间以居民名义收入计算的基尼系数为0.16,而以居民实际收入计算的基尼系数为0.11,降低了31%;25个省、自治区农村地区之间以居民名义收入计算的基尼系数为0.13,而以居民实际收入计算的基尼系数则也为0.11,降低了15%。由于该项考察所用样本区域覆盖范围广(36个城市分布于中国各个省区,所考察的农村地区也涵盖除了西藏之外的所有省区),同时统计分析也表明无论是在城市地区还是在农村地区,物价水平均与居民收入成正比,因此可以推断,全国各地的实际收入差距都应该小于名义收入差距。剔除了各地价格水平方面的差异后所计算出来的全国居民基尼系数也会比未剔除价格水平差异前的计算结果小。^④

这些研究表明,在对居民收入及其差距(如基尼系数)进行计算时,即使人们在所有收入项目的原始(既名义收入)数据方面都没有争议,但在各地价格水平存在较大差异的情况下,如果有人坚持按名义收入的数据信息来对收入差距基尼系数进行计算,而另一些人则坚持按实际收入信息来进行计算的话,那么,两类人最终计算出来的基尼系数结果自然也不一样。

四、收入分布拟合函数选择及基尼系数计算方法选择方面的不同

基尼系数的原理虽然简单明了,但计算基尼系数的具体方法则是多种多样。^⑤虽然对于离散数据或连续数据来说,各种计算公式本质上都是等价的(即在不考虑精确度的情况下,按这些公式计算出来的基尼系数值应该都是相同的),但由于在很多情况下,人们都只能是通过一定的样本数据来推论总体收入分布及其基尼系数,而在这一过程中,假如人们是采用连续方法来计算基尼系数的话,需要选择特定形式的函数来对收入数据的总体分布情况进行拟合。由于不同研究人员在选择拟合函数时会有不同主观偏好,这样,在研究人员所使用的收入分配原始数据和基尼系数计算方法完全相同的情况下,计算出来的基尼系数数据也可能会有所不同。因此,洪兴建明确地将“对已有数据进行参数分布拟合时的选择偏差”确定为是“造成中国基尼系数数值不一的另外一个重要原因”。他具体解释说“问题的重要根源是,依据样本数据拟合总体收入分布进而计算总体基尼系数产生的偏差,因为该种采用连续方法的最大不确定性是采用何种形式的函数拟合总体分布或洛伦茨曲线。比如对洛伦茨曲线而言,由于满足洛伦茨曲线的函数 $L(p)$ 只要具备四个特点,即 $L(0)=0$; $L(1)=1$; $dL(p)/$

①李实、史泰丽、古斯塔夫森主编《中国居民收入分配研究Ⅲ》,北京师范大学出版社2008年版,第14-15页。

②李实、史泰丽、古斯塔夫森主编《中国居民收入分配研究Ⅲ》,北京师范大学出版社2008年版,第16页。

③江小涓、李辉《我国地区之间实际收入差距小雨名义收入差距——加入地区间价格差异后的一项研究》,《经济研究》2005年第9期,第15页。

④北京大学中国经济研究中心宏观组《以购买力平价测算基尼系数》,《经济学(季刊)》2006年第10期。

⑤参见徐宽《基尼系数的研究文献在过去八十年是如何拓展的》,《经济学(季刊)》2003年第4期;梁纪尧、宋青梅《基尼系数估算方法述评及科学估算方法的选择》,《山东财政学院学报》2007年第1期;洪兴建《基尼系数理论研究》,经济科学出版社2008年版,第11-37页;周云波、覃晏《中国居民收入分配差距实证分析》,南开大学出版社2008年版,第31-42页。

$dp \geq 0; d^2L(p) / dp^2 \geq 0$ 而满足这四个特征的待选函数又有很多 采用不同函数估计基尼系数肯定有出入 而且有时还相差较大。可能是每个学者对拟合函数的偏好不同 国内经常看到只用某个不加说明的特定函数测算基尼系数的论文 应该说这是造成我国居民收入基尼系数值泛滥的一个作用原因。”^①

具体到中国当前的情境当中 ,一个更多引起争议的问题是如何利用国家统计局公布的城乡收入分组数据来计算全国居民的总体性收入差距基尼系数?

虽然对当前中国居民收入分配状况进行入户调查的团队不限于国家统计局 ,很多学者也都曾经组织过这方面的调查。但是 ,正如众所周知的那样 ,连续多年对中国城乡居民收入状况不间断地进行大规模年度入户调查并对主要调查数据加以公布的机构 ,只有国家统计局一家。因此 ,第一 ,对于那些也想对中国收入分配状况进行考察但又没有条件或能力去组织自己的入户调查的人来说 ,利用国家统计局公布的数据资料来进行研究就成为成本最低的一项选择; 第二 ,对于许多想对中国全国居民总体收入分配基尼系数进行时间序列上连续考察的人来说 ,也只能选择利用国家统计局所公开发布的数据。然而 ,正如李实、赵人伟等人曾经描述的那样: 问题是“由于国家统计局的城镇和农村的住户调查是分开进行的 ,分组的数据也是分发表的。至今没有看到他们对全国基尼系数的估计数值。现今所看到的有关全国的基尼系数 ,……大多是根据国家统计局的城镇和农村的收入分组数据 ,应用卡可瓦尼(Kakwani) 的内推法进行估算的。然而 ,从1990年开始 ,国家统计局不再发表城镇住户的收入分组数据 ,只发表等分组数据。而农村住户调查数据仍是收入分组数据。这样就很难将两套数据合成为一套统一的全国收入分组或等分组数据。因而也就无从根据它们对全国的基尼系数进行估计。”^②为了解决这个问题 ,许多研究收入分配的学者不得不想方设法对国家统计局公布的城乡居民收入两套数据进行整合 ,以计算出全国居民收入分配基尼系数 ,而不同学者设计出来的不同数据整合方法 ,也有可能造成全国居民收入分配基尼系数计算结果上的差异。我们以表 1 中的数据来对上述看法进行说明。

表 1 中国居民总体基尼系数

年份	陈宗胜 ^③	向书坚 ^④	程永宏 ^⑤
1981	0.2635	0.2780	<u>0.2927</u>
1982	0.2525	0.2494	<u>0.2769</u>
1983	<u>0.2865</u>	0.2641	0.2709
1984	0.2705	0.2684	0.2773
1985	0.2875	0.2656	<u>0.3073</u>
1986	0.2875	0.2968	<u>0.3239</u>
1987	0.2924	0.3052	<u>0.3247</u>
1988	0.2961(0.3497)	0.3133	<u>0.3384</u>
1989	<u>0.3697</u>	0.3214	0.3529
1990	0.3469	<u>0.3063</u>	0.3587
1991	0.3690	0.3240	N. A.
1992	0.3772	0.3396	<u>0.3993</u>
1993	0.4017	<u>0.3592</u>	0.4183
1994	0.4356	<u>0.3621</u>	0.4300
1995	0.4191	<u>0.3515</u>	0.4169
1996	0.4058		0.3946
1997	0.4027		0.3964
1998	0.4026		0.4001
1999	0.4129		0.4124

从表中的数据我们可以看到 ,三份文献所给出的基尼系数数据在很多年份上都有明显不同(见数字带有下滑线的年份) 。然而 ,需要说明的是 ,在以上三种基尼系数的计算过程中 ,三份文献的作者所用原始收

①洪兴建《基尼系数理论研究》经济科学出版社2008年版,第55页。

②赵人伟、李实、卡尔·李思勤《中国居民收入分配再研究》中国财政经济出版社1999年版,第143页。

③陈宗胜《经济发展中的收入分配》,上海三联书店1994年版,第244页;陈宗胜、周云波《再论改革与发展中的收入分配》,经济科学出版社2002年版,第27页。

④向书坚《全国居民收入分配基尼系数的测算与回归分析》,《财经理论与实践》1998年第1期,第76页。

⑤程永宏《改革以来全国总体基尼系数的演变及其城乡分解》,《中国社会科学》2007年第4期,第52页。

入数据完全相同,都是来自于《中国统计年鉴》等官方资料所公布的国家统计局住户调查数据,但各自在以国家统计局历年公布的城乡住户收入调查两套数据为基础来计算全国居民收入分配基尼系数时所使用的方法不同:

陈宗胜的计算方法是“分层加权法”,其公式可以简单概述如下:

$$\begin{aligned}
 P_i &= P_{Ai} P_A' + P_{Ni} P_N'; i = (1, 2, \dots, n) \\
 P_1 + P_2 + \dots + P_n &= 100; \\
 y_i &= Y_i P_i / (\sum Y_i P_i); \\
 y_1 + y_2 + \dots + y_n &= 100; \\
 Y_i &\in \{Y_1, Y_2, \dots, Y_n\}; \\
 Y_1 \leq Y_2 \leq Y_3 \dots, i \leq Y_n.
 \end{aligned}$$

式中, Y_i 为各阶层平均收入, y_i 为各阶层的收入比重, P_i 为各阶层人口比重, P_{Ai} 为农村各阶层人口比重, P_{Ni} 为城市各阶层人口比重, P_A' 为农村平均人口比重, P_N' 为城市平均人口比重, n 为按 Y_i 排列的阶层数。

具体计算步骤是: (1) 将已知的城乡各收入层次的人口比重按人均收入水平高低顺序列出; (2) 以城乡人口比重为权数算出全国各收入层次的人口比重; (3) 利用各收入层次的人均收入和人口比重计算出全国各层次的收入比重; (4) 利用人口、收入比重计算出中国全国居民的人口/收入基尼系数。^①

向书坚使用的计算方法是“分组加权法”,其公式也可以简单概述如下:

$$G = P_1^2 u_1 / u G_1 + P_2^2 u_2 / u G_2 + P_1 P_2 \text{②}$$

式中, G 表示全国居民的收入分配基尼系数, G_1 和 G_2 分别表示农村居民和城镇居民收入分配基尼系数, P_1 、 P_2 分别表示农村居民和城镇居民占总人口的比重; u_1 和 u_2 分别表示农村居民和城镇居民的人均收入; u 表示全体居民的人均收入。^③ 这些数据都可以根据国家统计局公布的相关数字推算得出。

程永宏使用的计算方法是“城乡混合分解法”,其公式可以简单概述如下:

$$G_n = \theta G_1 + (1 - \theta) G_2 + \alpha \beta G_3$$

其中, G_n 、 G_1 、 G_2 分别为全国总体基尼系数、农村基尼系数、城镇基尼系数, α 和 β 分别为农村和城镇人口份额, θ 和 $(1 - \theta)$ 分别为农村和城镇收入份额; G_3 是包含组内差距的相对城乡差距指标,其计算公式为: $G_3 = D / u$ 。其中, μ 为全国平均收入, D 表示绝对城乡差距。^④ 这些数据也都可以根据国家统计局公布的相关数字推算得出。

表 1 中所引三份文献在全国居民总体基尼系数计算方面的差异都是来自于各自使用的基尼系数计算方法的不同。可见,计算方法上的不同也是导致基尼系数计算结果不同的一个重要原因。即使所用收入分配的原始数据完全相同,只要在计算基尼系数时所采用的计算方法不同,其计算结果也会有一定的、甚至很大的差异。^⑤

五、抽样结果不同

从理论上说,在对中国居民的收入差距进行入户调查时,应该严格按照中国总人口在城乡、男女、地区、行业等之间的分布比例来进行抽样。但事实上,由于各种因素(成本费用方面的限制、调查主题方面的相互冲突等)的影响,不同的调查团队在实施调查时对这一原则的遵循也并不完全一致。有的可能比较严格地遵循了这一原则,有的则可能未能严格按照这一原则来进行抽样。其结果,也有可能导致所得收入数据方面的差异。

例如,卡恩等人认为 CHIP 调查中城镇居民人均收入数据与国家统计局住户调查中城镇人均收入数据

①陈宗胜《经济发展中的收入分配》,上海三联书店 1994 年版,第 240-241 页;陈宗胜、周云波《再论改革与发展中的收入分配》,经济科学出版社 2002 年版,第 26-27 页。

② $u_2 - u_1 / u$

③向书坚《全国居民收入分配基尼系数的测算与回归分析》,《财经理论与实践》1998 年第 1 期。

④程永宏《改革以来全国总体基尼系数的演变及其城乡分解》,《中国社会科学》2007 年第 4 期。

⑤这些差异表明上看是计算方法的不同所致,但其实更主要的是数据分组方式方面的差异所致。关于这方面的分析,可参见洪兴建《基尼系数理论研究》第 56-57 页的分析。

之间的差异,部分原因就在于双方抽样方法不同以及测量方面的误差。^①

2012年12月西南财大中国家庭金融调查与研究中心公布的中国2010年全国居民总体收入差距基尼系数为0.61,不论是与国家统计局公布的同一年份同类基尼系数(0.481)相比,还是与有关学者如李实等人以自己的调查数据为基础计算得出的基尼系数0.48(2007年)相比,都显得偏高。李实、岳希明等人在对这一基尼系数计算结果进行评论时,也都曾指出过造成这一差异的主要原因是该中心在抽样方面的误差。例如,岳希明就认为,CHFS报告对收入水平及其分散估计值出现了系统性偏差。而之所以会出现系统性偏差,主要原因又在于抽样样本的偏差。“仔细观察《CHFS抽样调查设计说明》不难发现,‘调查’的抽样过程明显存在向高收入群体倾斜的偏差,以及高估收入差距的偏差。‘调查’采取了三阶段PPS抽样方法,抽样的第一阶段是从全国2585个市县中抽取80个样本市县;第二阶段从样本市县中抽取320个居/村委会;第三阶段从样本居/村委会抽取住户。第一阶段的80个市县来自25个省,诸如西藏、新疆、内蒙古等低收入省份不在25个省之列。80个市县中来自东中西三个地区的个数分别是32、27、21,占比分别为40%、33.75%、26.25%,而全国2585个市县的东中西分布分别是34.3%、27.2%、38.4%。显然,东部县市样本明显偏多,西部样本比重显著偏低。第二阶段抽样中,‘调查’在把80个样本市县按非农人口比重标准五等分组的基础之上,从非农人口比重最低组到最高组,村委会与居委会抽取的个数比率依次为4:0、3:2、2:3、1:3、0:4。村委会和居民会的抽取直接关系到住户样本的城乡比,在下述的第三阶段抽样中,从村委会抽取的住户为农村住户样本,从居委会抽取的住户为城镇样本。这种村/居委会样本的抽取办法最终导致从非农人口比重最低组(16个市县)仅仅抽取农村住户,相反在非农人口最高比重最高组(同样为16个市县)只抽取城镇住户。如此的抽样方法忽视了非农人口比重最低组中的城镇住户和最高组中的农村住户,而这两部分住户的收入大致处于中等程度,对这部分住户抽样不足,以及相应扩大落后地区(非农人口比重较低市县)和发达地区(非农人口比重较高市县)的比重不可避免地高估收入差距。”“第三阶段抽样是从村/居委会样本中抽取住户样本。‘调查’在住户样本上有意识地多抽取了城镇样本,相应地减少了农村住户样本数。具体地说,每个村委会均抽取20户农户。对于居委会,在把居委会按平均住房价格分成四组后,从属于房价最高一组居委会中(每个居委会)抽取50户,次高组抽取45户,最后的两组依次是40户和35户。如此偏重于城镇住户尤其是高房价住户的住户抽取办法,不可避免地高估家庭财产的持有量,由于财产持有量与收入水平有着密切的关系,同时也高估了高收入人群的权重,从而高估了居民收入差距。”据此,岳希明认为,“‘调查’的抽样以及从此得到收入水平及其差距、家庭资产保有量等估计值完全没有全国代表性,报告中0.61的基尼系数不是中国居民收入不平等水平的可信估计值。”^②李实、罗楚亮等人在相应文章中也做出了类似的分析。^③

对于李实等人所提出的批评,西南财大中国家庭金融调查与研究中心甘犁教授撰文进行了争辩,认为自己的抽样设计是没有问题的。^④对于他们之间的这种争论,我们在这里暂不讨论。这里只是要借助于这个案例来说明一点,即抽样设计方面的差异也可以成为导致不同研究人员在收入分配数据及其基尼系数计算结果方面出现差异的重要原因。

六、结语

以上对导致当前我国居民收入分配基尼系数计算结果出现差异之原因的梳理并不一定全面。但即使如此,综合上面所述的内容,我们也已经可以在一定程度上体会到,造成当前我国居民收入分配基尼系数计算结果出现差异的原因是多方面的:举凡对“收入”概念的不同定义、收入数据采集和估算方法的差异、计算收

^①卡恩、格里芬、李思勤等《中国居民户的收入及其分配》,载赵人伟、格里芬《中国居民收入分配研究》,中国社会科学出版社1994年版,第56页。

^②岳希明《西南财经大学0.61基尼系数不可信!》<http://www.ciidbnu.org/news/201212/20121228134128706.html>。

^③参见以下文献:李实《为什么基尼系数会高估?》<http://www.ciidbnu.org/news/201212/20121227000408706.html>;岳希明、李实《我们更应该相信谁的基尼系数?》<http://www.ciidbnu.org/news/201301/20130123092800706.html>;罗楚亮《对〈中国家庭收入不平等报告〉的评论》<http://www.ciidbnu.org/news/201212/20121226133704706.html>;长溪岭《不是基尼系数不能说明中国,而是0.61根本不靠谱》,长溪岭09876的博客<http://blog.people.com.cn/article/3/1355731168068.html>;李实、万海远《提高我国基尼系数估算的可信度——与〈中国家庭金融调查报告〉作者商榷》<http://www.ciidbnu.org/news/201302/20130203104343706.html>。

^④甘犁《以公开科学的抽样调查揭示真实的中国》,《华尔街日报》中文网,2013年1月24日。

入时价格指数选择方面的差异、计算基尼系数时所用拟合函数或计算方法的差异,以及实施入户调查时样本数量及结构选择方面的不同等,都有可能对居民收入分配基尼系数的计算结果产生差别。由此,我们可以进一步引申出以下几点看法:

1. 我们在有关中国(或其他国家与地区)特定时期居民收入分配基尼系数研究文献中所看到的任何一个基尼系数值,都不能简单地被视为是对该国或该地区在这一特定时期内居民收入分配差距“实际”情况的“客观反映”,而应该将它视为该基尼系数值的计算者在特定话语系统(特定的“收入”定义、特定的数据采集和估算方法、特定的价格指数、特定的收入分布拟合函数或基尼系数计算方法以及特定的样本集合)的引导和约束下所完成的一项有关该国或该地区在这一特定时期内居民收入分配差距的话语建构。

2. 当我们看到或听到某个研究人员或机构发布他们关于中国(或其他国家与地区)特定时期居民收入分配基尼系数的测算结果时,我们首先要做的事情不是要把它作为一个“既定事实”来加以接受,并立即去追问“和以往的情况(或其他国家或地区)的情况相比它是高还是低”、“导致它如此高或低的原因是什么”、“我们应该采取什么样的措施来解决如此高的收入分配差距”等一类的后续问题,而是应该去考察一下这一基尼系数值的具体建构过程,考察一下它是在什么样的话语系统(什么样的“收入”定义、什么样的数据采集和估算方法、什么样的价格指数、什么样的收入分布拟合函数或基尼系数计算方法以及什么样的样本集合等)的引导和约束下被建构起来的。只有在完成了这项工作之后,我们才可能对这一基尼系数值的意涵有真正适当的理解,才可以去进一步探讨上述那些后续问题。

3. 当我们发现在既有的研究文献中针对同一国家(或地区)同一时期居民收入分配基尼系数有着不同的计算结果,而我们又想对这些不同计算结果的合理性加以判断、以便做出取舍时,我们要做的事情也应该是去对在这些基尼系数值的建构过程背后引导和约束着这一建构过程的那些话语系统(“收入”定义、数据采集和估算方法、价格指数、收入分布拟合函数或基尼系数计算方法以及抽样结果等)加以考察,对这些话语系统的合理性进行分析和判断。只有那些被我们认为具有合理性(或更高程度合理性)的话语系统以及在這一话语系统的引导和约束下所完成的基尼系数计算结果才能够被我们所接受;反之,就可以且也应该被我们所拒绝(假如我们确信自己有能力对不同话语系统之间的合理性进行判别的话)。

(责任编辑:陆影 luying1203@163.com)

(上接第33页)中作用的持续增加,并不表明劳动资本化程度的下降,反倒是劳动资本化程度提高的表现。活劳动尤其是脑力智力劳动,正在越来越多地转化为对象化形式的资本存在,越来越陷入资本增殖的结构之中。“在自反性现代性中,资本积累同时(愈来愈甚)也是信息的积累。因此不变固定资本(硬件)和不变流动资本(软件)作为生产手段被信息化了。与此同时,劳动力和产品(包括消费资料和生产资料)作为可变资本包含着越来越多的信息内容。此中的要点是,在信息和交流结构中信息(和资本)的积累成为自反性现代性的驱动力,正如在早期现代性中生产资本以及与之相关的社会结构的积累是早期现代性的驱动力。”^①无论是从可变资本、不变资本区分的角度看,还是从固定资本、流动资本区分的角度来看,资本的诸种构成要素都不同程度地发生了知识化和信息化。

后工业社会的来临,并不表明科学技术和知识信息对资本动力的取代,而是表明资本积累方式的具体形态的变换,即以“资本实物化”为主的积累方式转向了以“资本信息化”、“资本知识化”为主的积累方式。无论是后现代主义的文化形态,消费社会的日常生活方式,还是自反性现代化的发展趋向,都是由资本所推动的信息、知识生产形式的社会表征。由此,很多当代学者更愿意将当代社会称之为“知识资本主义”、“信息资本主义”,这恰恰切中了当代现代性发展的新特征、新趋势。

(责任编辑:周文升 wszhou66@126.com)

^①[英]斯科特·拉什《自反性及其化身:结构、美学、社群》,见[德]乌尔里希·贝克等《自反性现代化》,赵文书译,商务印书馆2001年版,第162页。