

# 中国政府统计数据质量研究

## ——一个文献综述

赵学刚 王学斌 刘康兵\*

**摘要:** 近年来,中国政府数据统计质量问题日益受到国内外学者的关注,相应的研究文献大量涌现,研究主要集中在统计数据质量的含义和衡量标准、统计数据失真的作用机理、统计数据质量控制技术、评估方法和管理理论,同时还有针对中国统计数据质量进行实证方面的评估以及改进统计数据质量的政策建议方面的讨论。本文在回顾现有文献基础上,指出要从根本上解决我国政府统计数据失真问题,还得从改革现行的统计管理体制入手,摆脱地方政府干扰,确保统计独立性,同时认为在现行框架下增加诸如消费价格指数权重、房屋销售价格及房价指数、固定资产投资等指标的统计透明度可以在很大程度上避免当前很多对中国政府统计数据不必要和不公正的质疑。

**关键词:** 统计数据质量 理论研究 经验证据 对策

### 一、引言

2010年2月25日,国家统计局发布的《2009年国民经济和社会发展统计公报》显示,2009年全国70个大中城市房屋销售价格上涨1.5%。<sup>①</sup>而根据2010年1月19日国家统计局公布的2009年全国房地产市场数据,相比前一年均价上涨约24%。<sup>②</sup>2009年广大居民普遍感觉房价疯涨,出于与实际感觉的巨大反差,网民对当年“70个大中城市房屋销售价格上涨1.5%”的数据提出了大量质疑。国家统计局城市司2010年2月28日不得不公布了该数据的统计调查方案予以回应,以说明两套数据计算方法的差异。

其实中国政府统计数据质量问题一直存在,早在1988年Peckins就提出了中国经济增长率可能存在上偏误差的问题。2001年美国匹兹堡大学Rawski教授先后发表两篇文章,几乎是全方位地对官方公布的中国经济增长数据表示高度的质疑(Rawski 2001a 2001b)。这一论断与著名的美国经济学家克鲁格曼(Paul Krugman)此前的疑虑遥相呼应,素有“国际经济乌鸦”之称的他曾经说过,“中国经济数据,是一堆不可置信的符号体系。看中国的经济增长数据,就像看一部科幻小说。”<sup>③</sup>2002年,中国经济改革研究基金会国民经济研究所副所长王小鲁针对我国经济增长统计的准确性指出了两个“正常”和两个“不正常”,即二十世纪八九十年代以来我国经济结构的调整和生产效率的提高导致“经济增长快于能源增长应当是正常的,但经济增长在某些时间段或某些年份与能源增长发生找不到原因的重大偏离甚至反向变动,是不正常的”;经济体

\* 赵学刚,上海大学经济学院,邮政编码:200444 电子信箱: xuegangzha@126.com; 王学斌,上海大学经济学院,邮政编码:200444 复旦大学城市规划与发展研究中心,邮政编码:200433; 刘康兵,上海大学经济学院,邮政编码:200444

本文在上海大学Seminar“当代中国经济论坛”(No 42)上进行了讨论,感谢上海大学经济学院倪中新副教授、张国平教授、韩太祥副教授和顾标博士的评论,同时特别感谢匿名审稿人的建设性意见,文责自负。

①其中新建住宅价格上涨1.3%,二手住宅价格上涨2.4%,房屋租赁价格下降0.6%。《2008年国民经济和社会发展统计公报》显示,70个大中城市房屋销售价格上涨6.5%,其中新建住宅价格上涨7.1%。2009年的相应数据要比2008年的都低。

②2009年全年房地产成交量近9.4亿平方米,以总成交额计算平均成交价格约4695元/平方米,相比前一年均价上涨约24%,平均每平方米上涨813元。其中上海、深圳和北京位居房价涨幅前三名,同比分别上涨27.56%、18.32%和16.38%。

③转引自赵晓、高辉清、钟伟、巴曙松,2002《中国经济增长的真实故事》,《财经界》4月28日。

制改革导致“综合要素生产率的一定程度上升是正常的,而没有原因的突然大幅度跃升是不正常的”。关于消费价格指数(CPI)、平均工资、房价、失业数据、城镇化数据、居民收入数据、第三产业数据、房地产投资等统计数据被国家统计局搞得一片纷乱,经常存在不一致性,很难让人信服。最近,中国统计数据体系正在受到越来越多的质疑和拷问。随着中国经济结构的转型,我国的统计体系要从以制造业特别是工业生产为主的统计调整到涵盖商品消费、服务业、城市化水平等指标在内的全方位统计。

改革三十多年来,中国的统计体系已经发生了很大的变化,统计部门也做出了很多努力。但总体看来,统计体系严重滞后于经济发展现实,统计工作显然还有更大的改进余地,而且当下更显迫切。中国下一步要改善经济增长质量、改善宏观经济决策、进一步深入改革和完善市场经济体制,其中一个很重要的环节就是要改革现有的经济统计体系,这就必须要涉及到统计部门的改革。由于现有的关于这个主题的文献很多,评述所有学者的贡献几乎是一件不可能完成的事情,因此,笔者只回顾了其中极具影响力的文献(当然,疏漏也在所难免),根据“理论——实证——政策”这一主线对其进行梳理,希望本文能够引起学界和业界对中国统计体系的关注及其改革方向和对策的讨论。

本文余下部分结构安排如下:第二部分研究了统计数据质量理论方面的内容,第三部分研究了统计数据质量评估的经验证据,第四部分研究了改进统计数据质量的对策,第五部分是一个简单的小结。

## 二、统计数据质量的理论研究

目前国际上针对统计数据质量还没有一个统一的定义,不同的学者也有各自不同的定义。总的来说,从用户、生产者和被调查者三个角度考察,衡量统计数据质量的标准共有 11 个,即适用性、准确性、及时性、可比性、可衔接性、可取得性、可解释性、客观性(或诚信)、方法专业性(或健全性)、有效性、减轻调查负担。其中适用性、准确性、及时性、可取得性、可比性和可衔接性是国际统计界公认的数据质量之基本构成要素和各国政府统计机构对数据进行质量检测、监管的重要内容和依据(余芳东, 2002; Wang and Strong 1996; Naumann and Rolker, 2000)。我们可将上述标准理解为统计数据质量的内涵。绝大部分文献,仍然以准确性评估为核心内容,因此统计数据质量存在问题一般都是指统计数据失真或失真。

### (一)统计数据失真机理

统计部门的行为会直接影响数据质量,如统计数据搜集整理过程中出现的设计误差、调查误差、汇总误差等(王强, 1990; 颜德纶, 1990)。统计主体间的博弈是产生数据质量问题的重要诱因,如企业和统计部门、上下级、区域间基于利益的博弈(朱秀君, 2000; 卢冶飞, 2003; 朱文兴、张继良, 2005); 官员业绩博弈、干预基层博弈和干预统计博弈(林勇、杨言勇, 2008); 地方官员在经济竞争当中的政治晋升博弈(周黎安, 2004)。同时,袁平德(2004)指出被统计部门(地方、企业)之间的博弈容易造成各种攀比现象,表现在总量上攀“块头”、均值上攀“高低”、纵比上攀“速度”、横比上攀“位次”、“唯上”中攀“宠幸”。

对于统计体系,现行的国内生产总值(GDP)核算体系本身的缺陷则是文献所关注的重点。世界银行(1992)认为中国以样本为基础的生产指数和独立构造的价格平减指数欠缺,没有对未出售的库存进行冲销,导致了官方对 GDP 总量数据的低估和速度的高估;李成瑞(1997)指出经济成份分类方面的科学性和所有制结构统计调查方面的准确性问题导致了含混或不准确和某些虚假不实现象的发生;孟连和王小鲁(2000)总结了七个导致数据偏差的可能机制。<sup>①</sup>许宪春(2002)认为现行的 GDP 核算体系中,产业部门分类和支出项目分类仍然过粗,经常性的服务业统计调查制度欠缺,缺乏 GDP 的分季度生产核算和使用核算,价格指数资料存在缺口,未被观测的经济存在统计遗漏,统计系统的独立性不强,某些统计数据容易受到地方干扰;任若恩(2002)认为中国经济增长率被高估的主要原因是长期以来实行的可比价格法<sup>②</sup>,考虑采用国际上通用的价格指数缩减法和生产指数法来测算经济增长率。岳希明(2005)指出劳动统计方法的差异

<sup>①</sup>分别是上级下达指标,政府及媒体对地区、部门、企业按经济增长的排序诱导的攀比,基层统计业务量大和人员编制少之间的矛盾,规范、科学的估算方法和制度的欠缺,工业统计方面乡镇企业、村及村以下企业统计体系不健全,城乡个体私营经济、外商和港澳台商投资经济在统计体系方面的不合规,政绩考核标准。

<sup>②</sup>理论上可比价格法的正确性与产品种类实际发生的变化之间的矛盾成了这种方法衡量经济增长率的最主要的问题,使得可比价格实际上成了当期价格。

导致了就业人员总数、就业人员行业以及它们的时间变化方面较大的差异。宋小川(2007)通过讨论现行GDP核算体系的缺陷<sup>①</sup>,认为中国GDP及其增长率被严重低估。

同时统计管理体制、统计制度方法、统计基础工作、统计外部干扰等是影响政府统计数据质量的重要因素(政府统计数据质量研究课题组,1999)。肖宜滨(1999)在法学范畴内剖析了影响统计数据质量的原因,主要有行政自由裁量权对主体的干扰、部分统计工作者统计法制观念淡薄、请求权的非强制性。

以上文献主要从统计部门、统计体系和统计外部体制环境三个方面分析了中国统计数据失真的原因。事实上,三个方面是相互影响的,尤其是外部体制环境对统计部门的影响这一因素在中国尤为显著。很多时候,统计部门受到地方政府的干扰以及同级部门之间源于考核指标的竞争,而牺牲了统计数据的真实性。尤其是地方政府源于升迁和争取中央政策支持而产生的数据攀比和赶超,不能不引起公众对其所公布的统计数据真实性的怀疑。很多地区同年GDP增长率和该地区历年增长率惊人的相似以及和上级要求指标惊人的一致即为明证。很多文献也反复提及这一因素。理论上,统计体系本身的缺陷对数据质量的影响应该是系统性的,即在影响方向上应该是不变的。因此,导致目前统计数据存在问题的很大一部分原因可以归结为地方政府催生的统计部门间的博弈。

基于对以上三个方面因素作用机理的理解,国内外学者主要从统计数据质量控制、统计数据质量评估、统计数据质量管理三个方面出发,进行了大量的理论研究,给出了控制、评估和管理方法。

## (二)统计数据质量控制

统计数据质量的控制,主要是数据搜集和整理过程中对误差的控制,针对统计误差所开展的理论研究,涉及测量误差的统计分布和检验、抽样误差以及非抽样误差、缺失数据的统计处理方法、汇总数据的误差等。基于这些研究,产生了抽样调查方法和非抽样误差控制技术(金勇进、陶然,2010)。Neyman(1934)关于分层抽样最优分配及回归、比率估计的研究是统计数据质量研究的一个重要里程碑。Hansen等(1961)提出了调查误差模型,该模型把总误差(均方误差)分解为测量误差、抽样误差、测量误差与抽样离差的协方差以及偏差的平方四个组成部分,这一模型的建立为调查过程中误差的控制奠定了重要基础,随后的样本重复计量、交叉子样本等方法都是在此基础上发展而来。同一时期,Wamer(1965)则对敏感性问题展开专项研究,提出了对敏感性问题调查的随机化方法。Dalenius(1983)认为数据质量的综合控制应包括调查设计控制和调查实施控制两部分,后者又分为预防控制和过程控制。Housman(2003)则将CPI偏差归结为四类:替代性偏差(substitution bias)、新产品偏差(new goods bias)、质量改进偏差(quality change bias)、购买途径偏差(outlet bias)。Costa(2001)和Hamilton(2001)利用恩格尔曲线从总偏差的角度对CPI偏差进行了估计,依据恩格尔定理,他们认为家庭消费支出偏差可能源自对价格的错误估计。这一方法克服了传统方法只能估计单因素偏差的缺陷,同时数据要求较少,可操作性强。

李成瑞(1984)从应用层面,将分阶段实施控制技术分别应用于我国第三次、第四次人口普查,取得了良好效果。傅德印(2000)系统讨论了政府统计数据控制技术,涉及调查变量和调查项目、调查误差类型以及建立和实施政府统计数据质量文档及监测等方面。杨清(2000)分析了系统性误差的特点和分解、产生原因和解决方法以及测定方法。同时,杨清和吴伟霞(2000)指出运用方差分析、多重比较的q检验法(或HSD检验法)和交叉样本技术对原始资料的偏误进行监控。方差分析和多重比较检验二者相互补充,前者分析了控制变量是否对观测变量产生显著影响,后者则进一步确定控制变量的不同水平对观测变量的影响程度。在进行了交叉样本分析之后,进一步运用方差分析和多重比较,则可减弱或消除系统性误差的影响。

## (三)统计数据质量评估

统计数据质量评估一方面可以检验所搜集到的数据本身的质量,另一方面还可以检验数据搜集过程中质量是否得到有效控制。

刘孝新等(1997)运用事后预测和反常结果判断法对统计数据质量进行了评估和推断。成邦文和董丽娅(2002)从一致性误差和非一致性误差两个方面分析了社会经济统计数据误差,关于前者的分析内容有趋

<sup>①</sup>主要是没有包括非市场经济、地下经济和闲暇活动,低估了劳务产业创造的价值和增长潜力,无法追踪企业在技术革新、产品设计、品牌打造和员工培训等无形资产方面的投资,忽略了我国经济中迅速增加的知识经济成分,特别是教育投资成分。

势性分析、机构对比分析和相关指标关联分析,而对后者的分析则包括分布形态一致性检验和异常数据的识别。刘洪和黄燕(2007, 2009)先后运用趋势模拟评估法和相关指标分析法,对统计数据质量进行了定量评估。张德然(2003)给出了根据各数据点的跳跃度所提供的信息来检验异常值的方法。而 Rawski(2001)则认为由于现行的各类评估方法在逻辑思路和技术层面都有一定的局限性,因此在讨论中国统计数据质量时,应该更多关注整体经济数据的精确度、范围、覆盖面、统计方法以及其他技术问题。

管于华和尚卫平(2003)根据 GDDS 国际标准对中国国民核算账户进行评估,从实际部门、财政部门、金融部门、对外部门和社会人口部门分别指出了中国统计数据的缺陷。常宁(2004)认为国际货币基金组织(MF)的数据质量评估框架 DQAF 是对 SDDS、GDDS 的补充和完善,并从评估对象、评估机构、评估方法、评估表、评估结论五个要点提出了在我国实施 DQAF 的基本思路。<sup>①</sup>王华和金勇进(2009)根据辅助资料信息的来源及其表现形态、构造参照标准的技术假定及方法范畴、实际统计数据与参照标准的比较逻辑对统计数据准确性评估方法进行了归类。<sup>②</sup>

国外关于数据质量评估方法,有 AMQ<sup>③</sup>,它由 PSP/IQ 模型(一个 2×2 表格,行决定于将数据看成产品还是服务,列决定于评估是依据正规的质量规范还是消费者期望)、QA 工具(一套测量数据质量的调查问卷)和质量差距分析技术三部分构成(Wang, Strong and Kahn, 1999; Yang, 2002);建立数据质量标准列表的一种在组织内部测量数据质量的方法,标准分为直接评估标准和间接评估标准,前者采用传统的软件测量方法即问卷调查方式,而后的分数由直接评估标准计算得到(Bobrowski, Marre and Yankelevic, 1999);Naumann 和 Rolker(2000)采用面向评估的方法对 Web 数据质量标准进行了分类;Pipino, Lee 和 Wang(2002)给出了简单比率、最大-最小运算和加权平均三种数据质量客观评价的算法;Cappiella, Francalanci 和 Pemic(2004)建立了由选择模块、质量评估模块和简表模块构成的数据质量评估框架,将评估过程得到的数据质量值和用户要求进行比较来判断质量是否符合要求。

可以看出,在数据质量的评估方面,国内学者的关注点是(政府)统计数据(statistics),而国外学者提出的方法则针对的是一般意义上的数据(data)。因此,将国外的研究方法引入国内来评估政府统计数据质量是可行的,同时还需结合数据体系的特点和兼容性进行具体的评估分析。

#### (四)统计数据质量管理

在对统计数据质量进行有效控制和评估的同时,建立统计数据质量管理体系是实现统计数据质量持续改进和统计工作全过程质量提高的有效途径(金勇进、陶然,2010)。颜德纶(1990)将质量管理学中的 PDCA<sup>④</sup>循环理论与统计工作相结合,提出了覆盖统计工作全过程的统计数字质量全面管理的思想;傅德印和刘晓梅(1994)分析了在我国贯彻 ISO 9000 系列标准,建立数据质量体系的可行性,给出了统计数据质量体系的结构和要素及其建立和实施步骤。同时,傅德印(2000)指出政府统计数据质量管理体系,在横向上包括统计数据质量控制技术体系和政府统计数据质量保证体系,纵向上包括政府统计数据质量管理体系的建立、实施、评审及认证等活动。通过回顾和借鉴国外数据质量评价和管理的方法和经验,余芳东(2002)提出了加强我国统计数据质量管理的建议。邱东和陈梦根(2008)从统计能力建设的层面,建议从五个方面提升数据统计质量。<sup>⑤</sup>余根钱(2008)认为应该在统计设计、数据采集、数据加工整理和数据发布等统计工作的各个阶段,对数据质量层层把关,确保数据的准确性和真实性。高敏雪(2009)通过比较政府统计与企业会计运行机制,强调就整个统计数据生产过程进行外部监督来解决统计数据质量问题。

国外关于数据质量管理体系的研究,很多基于全面质量管理(TQM)的原则、方法、指南,也继承了

<sup>①</sup>DQAF、SDDS、GDDS 分别是 Data Quality Assessment Framework(数据质量评估框架)、Special Data Dissemination Standards(数据公布特别标准)、General Data Dissemination System(数据公布通用系统)的缩写。

<sup>②</sup>针对总量统计数据,主要在纵向时间维度上开展,根据统计指标与其相关联指标变动趋势的偏离程度加以评判;针对个体或分类统计数据则主要在横向空间维度上开展,对数据的统计分布形态加以检验,或者利用重复调查或随机实验方法对事先假定的误差参数进行估计。

<sup>③</sup>Assistance in Maintenance & Quality Management 的缩写。

<sup>④</sup>Plan(计划)、Do(执行)、Check(检查)、Action(纠正)的首字母。

<sup>⑤</sup>将数据质量管理纳入国家统计发展战略(NSDS),采纳全面质量管理方案,采取外部评估和内部评估相结合的数据评估和控制方式,建立统计生产者和使用者的有效沟通和协调机制,并建立全国性的统计数据库网络体系。

ISO 9000系列标准的框架 (Wang Storey and Firth 1995; Wang 1998; Wang Allen and Harris 2003)。Ryu和Park (2006)构建了数据质量管理成熟度模型来评估组织数据质量管理能力,共分为初始、已定义、已管理和优化数据管理级四个等级。如同数据质量评估,国外关于数据质量管理的研究也是针对一般意义上的数据 (data)。

### 三、统计数据质量评估的经验证据

理论层面上关于统计数据包括失真机理、质量控制、质量评估和质量管理的研究给出了一般性的研究基础、研究方法和研究框架,针对中国官方统计数据质量问题,国内外学者也进行大量的实证层面的研究,而且这些研究绝大多数直指国际上争议颇大的GDP及其增长率的准确性问题。

#### (一)来自国外的证据

国外关于中国经济增长统计数据质量问题的正式研究成果,大多数认为中国政府公布的经济增长率一定程度上偏高。这些研究从统计口径和统计方法的角度提出了对中国政府统计数据质量的质疑。

1997年的世界银行研究报告认为,中国在1978-1995年期间经济增长率实际应为8.2%。Maddison (1998)研究表明,这一时期中国经济年均增长率应为7.5%,而同期国家统计局公布的增长率为9.9%。不过,余芳东 (2004)指出世界银行按汇率转换的中国GDP数据一定程度上低估了中国实际经济规模,而由于受资料来源和计算方法的限制,世界银行公布的购买力评价结果明显高估中国的实际经济规模,不能反映现实的支出结构、消费结构和商品价格水平。

Ravskij (2001)通过分析1998-2001年中国的官方数据,指出中国经济增长率与能源消耗、生产与投资、消费与收入等数据之间都存在不一致性。认为从1998年开始,中国的GDP数据就存在夸大成份,1998年增长率至多为2.2%,而官方数据为7.8%;1998-2001年期间的中国GDP累计增长实际上只有0.4%~11.4%,而官方统计为34.5%。不过,Ravskij得出这一结论的逻辑前提是GDP与能源消耗应该保持一致增长,而这一前提从其他国家如美国、英国、日本和韩国的发展实践来看,有待斟酌。

Sinton (2001)分析了1990-2000年中国的能源统计数据的内在一致性,认为90年代初期的能源数据相对较真实,但从90年代中期开始质量便出现下降趋势。Huenemann (2001)研究发现90年代以来,中国的官方运输数据与年度石油消耗量和客流量存在不一致性,而且这一问题变得逐渐严重。两位学者都指出从90年代开始,数据质量问题凸显,表现为相关性很高的(宏观)变量间变化幅度和趋势的不一致性。

根据中国官方数据,农村贫困人口从1978年的31%下降到了2000年的3%,Park和Wang (2001)认为这一数据低估了农村贫困率,高估了贫困减少速度。Schaping (2001)也分析了人口数据,指出计划生育政策和人口流动性的增强导致中国的人口统计数据存在严重问题。这一分析是符合中国实际的,由于中国人口统计体系统计项目和统计标准的不完善,导致了农村人口和贫困人口的漏统。

Klein和Oznuur (2002)利用中国1980-2000年的年度数据,根据包括电力、煤炭、货运等在内的15个基本经济变量来解释GDP的增长,结果表明这些变量与中国官方发布的GDP增长之间的相关关系是符合经济规律的。

中国官方数据显示,1997-2001年中国进口额增长了70%,财政收入增长了90%,Lardy (2002)通过解释这两个经济指标数据不可能被高估的原因,认为中国官方发布的1997-2001年GDP增长率为7%以上是可信的,并没有被高估。

利用价格指数法,Szirmai和Ren (2005)发现1980-1987年间工业总产出高估了1个百分点;在1986-1994年间,GDP的增长率高估了3.8个百分点(Ren 1997)。

利用生产指数法,Wu (1997)发现1978-1994年间工业增长高估了3.4个百分点,1978-1997年间高估了3.3个百分点;Maddison (1998, 2001)发现1978-1994年间GDP增长率高估了2.4个百分点,而1995-1998年间高估了2.1个百分点(官方数据为8.7%)。

需要指出的是,价格指数法和生产指数法是国际上通用的计算增长率的两种方法,前者对于工业增长率和服务业增长率的计算都是适用的。而后者则主要适用于工业增长率的计算,对绝大多数的服务业增长率并不适用。

## (二)来自国内的证据

近年来,关于中国宏观经济统计数据质量的讨论在国内也相当热烈。国内学者主要从微观角度进行了评估和检验。

孟连和王小鲁(2000)通过168种工业产品产量指数,货物运输业增长、电力和能源消费量增长各自与工业增长之间的相关分析,以及生产函数分析三种方法,得出结论:改革前的1953-1977年间和改革后的1978-1991年间,工业增长率没有明显虚增,但1992-1997年增长率明显过高,应从16%调整到12%。GDP增长率在1953-1977年期间可能有2.2个百分点的虚增,在1978-1991年期间可能有1.0个百分点的虚增,在1992-1997年期间可能虚增了2.5个百分点。<sup>①</sup>

成邦文、董丽娅和杨峻(2000)采用对数正态分布检验法,对1998年研究与开发机构的人员、经费、资产、活动、成果、技术交易六个方面共86个指标数据进行评估,结论表明在5%的显著性水平上数据质量是比较好的。同时,在对全国4806个科研院所的统计数据评估中,86个指标有65个在1%的显著性水平上通过了检验(占76%,包括主要指标),表明全国研究与开发机构主要调查指标数据的质量也是比较好的。

张新和蒋殿春(2002)首先利用消费品零售总额、固定资产投资、工业制成品库存和出口等月度数据估算了中国1996-2001年间的GDP增长率,结果同官方数据没有太大差别,同时指出1998年的月度数据存在较大疑点。其次,通过计算1978-2000年间能源消耗、能源产量、钢产量、发电量等重要生产资料与GDP增长的弹性,发现能源总产量和总消费与中国的经济增长之间从来不存在十分密切的关系,而钢铁和电力产量则与GDP关系密切。若以发电量的增长弹性衡量GDP的可信度,1998年是真正的可疑点。

阙里和钟笑寒(2005)利用中国28个地区1984-2001年间10个经济基础变量共504个样本数据<sup>②</sup>,运用主成分分析和固定效应变截距模型对中国GDP增长的真实性的地区层次进行了检验,没有发现GDP统计数据存在系统的、长期的错误的证据。同时指出地区影响显著存在,不能排除地区数据造假的嫌疑。而年度特殊影响在大多数年份不显著,不支持1998年GDP7.8%的增长率是严重作假的观点,但也不排除1996年地区GDP统计数据存在一定程度高估的可能。<sup>③</sup>

李子奈和周建(2005)从经济系统角度使用联合估计诊断方法对我国主要36个宏观经济序列进行比较分析,研究表明我国宏观经济统计数据中存在多于10%的异常点,大部分异常点或多或少聚集成堆出现,孤立的异常点不是我国宏观经济的主要特征,这些异常数据基本上是真实的。

刘洪和黄燕(2007,2009)将1978-2003年间的GDP数据以1978年为不变价格换算后,在这些年份数据真实可靠的假定下,运用趋势模拟法认为我国2004年公布的GDP数据是准确的。并根据我国某地区1978-2004年产出、资本和劳动投入数据,研究表明1978年、1984年、1985年、1986年、1991年的GDP数据较为可疑,同时认为应结合相应年份的宏观经济环境、重大事件来评估可疑样本点的质量。

王静(2009)选用1989-2007年间能源消费、全社会货运周转量、居民消费指标来检验GDP数据可信度。针对能源消费指标,结果显示1998-2000年官方GDP数据低于模型的置信区间,而2002-2005年GDP数据高于置信区间,其余年度GDP水平均在置信区间内;对于货运周转量,1998-2000年的官方GDP数据低于置信下限,2002年和2004年略高于置信上限;1996-1998年官方社会消费品零售总额略低于置信下限,1995年和2006年略高于置信上限,此外没有偏离置信区间。最终认为我国GDP数据并不存在显著问题。

许永洪和曾五一(2009)利用1997-2006年的城市价格和收支统计数据,运用Hamilton-Costa方法估计了我国CPI偏差。研究发现,以1997年为基年,1998年和2000-2006年中国公布的定基CPI与城镇居民生活成本指数不存在偏差,只有1999年存在5.83%的累积偏差。

不同文献使用的方法的不同可能会导致评估结果的差异,当然我们也应看到这些方法的适用性及优缺点,但不可否认的是,国内外学者大量的实证研究从多个角度证明了中国政府统计数据质量问题的确存在。

<sup>①</sup>据此,改革前(1953-1977年)经济增长率应由5.9%调整到3.7%;改革期间(1978-1997年)应由9.9%调整到8.4%;1992-1997年期间虚增明显,应由11.5%调整到9.0%。

<sup>②</sup>实际利用了400个。

<sup>③</sup>这一结论与孟连和王小鲁(2000)认为1996-1998年经济增长与通货膨胀背离的观点基本吻合。

那么如何解决现有问题,改革现有统计体系?

#### 四、统计数据质量改进的对策探讨

针对统计数据质量问题,中国政府统计部门已经采取了一系列措施,并取得了初步成效。如改革统计调查方法、统计汇总方法、工业不变价增加值计算方法及扩充统计调查范围等,逐步细化统计的产业部门分类(特别是规模以下工业)和支出项目分类,建立经常性的服务业统计调查制度(特别是新兴服务业),建立各专业分季度统计调查制度,完善价格统计,弥补未被观测的统计缺口,建立相对独立的统计管理体制来完善GDP核算(许宪春,2002)。2002年11月21日,时任国家统计局局长的李德水在由国家统计局和世界银行联合举办的中国统计体系改革发展国际研讨会上指出,中国现行统计体系本身也还亟待完善,包括统计管理体制还存在一些弊端:统计指标体系设置不够合理;基层统计机构和统计基础工作比较薄弱;统计部门内部机构设置和 workflows 也还需要进一步完善等。

基于对统计数据失真机理的理解,国内外学者从影响统计数据质量的各个因素出发,主要从宏观战略层面、统计体系本身、统计管理以及指标数据的准确性等方面提出改进政府统计数据质量的建议。

总体而言,李成瑞(1994)认为统计改革就是围绕统计认识整体功能的完善和加强,对统计组织、统计调查对象和统计资料的使用者及其相互关系进行变革。《统计数字质量研究》课题组(1995)建议在统计调查方法、统计体制、统计协调、防范制度、统计法建设、考核机制、统计队伍和统计基础等方面进行深入改革和完善。

就统计体系本身,应建立与SNA协调配套的BOP、MFS和GFS体系<sup>①</sup>,以充分发挥中国国民经济核算的作用(庞皓,1999)。朱启贵(2009)在总结我国统计学界对国民经济核算理论与方法进行的创新的基础上<sup>②</sup>,指出未来应建立涵盖社会资本、资源环境、隐性经济、人力资本、国民卫生、代际公平、科教等要素在内的新国民经济核算体系。孙颖等(1996)建议进一步改革和完善现金统计制度,使其能够真实、全面地反映现金投放回笼总量及分布结构,更好地发挥金融统计的功能。李金昌(2002)将我国现行的统计调查方法体系存在的问题概括为“一高一低一重”<sup>③</sup>,建议在统计报表、普查制度、网络统计调查系统、统计估算方法和统计调查技术方面进行深入改革和完善。在新旧体制转化过程中要保证统计数字的准确性和及时性(《统计数字质量研究》课题组,1995)。

统计管理方面,当务之急是改现行的统计管理体制为垂直领导体制,使统计部门摆脱地方政府的干预(政府统计数据质量研究课题组,1999)。朱文兴和张继良(2005)认为目前最佳的方案是建立国家支持下的省以下垂直的管理体制模式,最终模式应该是国家垂直管理到基层。

其他方面,傅德印和陶然(2007)通过探讨政府统计数据质量成本和统计数据质量之间的关系,从成本角度考虑了提升统计数据质量的思路。<sup>④</sup>从博弈论角度,朱秀君(2000)建议从加大惩罚力度和提高查处概率两方面着手保证统计数据质量;刘海清和熊祖轅(2009)认为引入统计信用可以在有限成本下防范数据失真,并指出统计信用惩戒制度是统计信用的制度基础。肖宜滨(1999)在法律责任、法律程序、统计立法三个层面从法学角度提出了提高统计数据质量的建议。Pipin Lee和Wang(2002)在评估算法的基础上介绍了将主客观评价相结合、在实际应用中改进数据质量的步骤。Keitel(2001)认为GDP支出账户是衡量经济产出的最好方法,但中国的GDP支出账户还有很多缺陷,应予以改善,提高GDP数据准确性。

在具体指标数据准确性方面,中国人民银行武汉分行、国家统计局湖北调查总队联合课题组(2006)认为中国应当尽快建立剔除了自然因素、行政控制、垄断定价等影响的CPI指数即核心CPI以反映消费价格的中长期走势,并由国家统计局按月编制,定期发布,更好地为制订中长期经济发展战略和宏观经济调控服

①SNA、BOR、MFS、GFS分别是System of National Accounts(国民经济核算体系)、Balance of Payments Manual(国际收支统计手册)、Manual on Monetary and Financial Statistics(货币与金融统计手册)、Manual on Government Financial Statistics(政府财政统计手册)的简称。

②主要有提出国民经济核算的平衡原则,建立社会劳动价值论,提出国民大核算理论,整合MPS和SNA四个方面。

③调查成本高、数据质量低、基层负担重。

④O'neill在1995年就提出农业普查和抽样调查的数据质量标准应包括最低成本这一特性。

务。任若恩(2002)则认为提高中国经济增长率的准确性,最终要依赖于国家统计局所使用方法的改进和日常工作中测量中国增长率所使用的数据来源的改善。

## 五、结语与展望

本文通过对国内外关于政府统计数据质量的研究文献进行梳理,介绍了统计数据质量的含义和衡量标准、统计数据失真的作用机理,统计数据质量控制技术、评估方法和管理理论以及一些实证研究结果和学者们关于改善统计数据质量的政策建议。

纵观研究我国统计数据质量的大量文献,关于宏观经济数据失真的机理分析已相当充分,指出了问题的症结即制度性因素,切合中国实际;统计数据质量控制技术则偏向于理论和方法的介绍,相对于评估方法和管理理论,在与实际问题的结合方面比较欠缺;在通过实证研究指出失真问题的存在性和失真程度之后,政策建议则侧重于战略角度,缺乏实用性。因此,加强质量控制技术的实际应用需要进一步深入研究,同时结合数据失真的机理而提出的改进措施应更具针对性和可操作性,以切实提高统计数据质量,增强我国政府统计数据可信度。

结合文献观点,笔者认为要从根本上解决我国统计数据失真问题,还需从统计管理体制入手,变目前的“统一领导,分级负责”为“国家统计局垂直管理到底”,各级统计部门的人权、财权、事权全部交由上级统计部门安排,摆脱地方政府干扰,确保统计的独立性。目前的统计管理体制最大的弊端在于各级统计部门受制于同级的地方政府,是地方政府考核的对象。对各级政府实行目标考核这一激励机制固然合理而且必要,但是统计部门不应该被纳入成为地方政府的考核对象。2010年1月1日起实施的新《统计法》已明确表明,统计的基本任务是对经济社会发展情况进行统计调查、统计分析,提供统计资料和统计咨询意见,实行统计监督。简洁地说,统计的任务或职能就是调查、分析、咨询和监督,没有考核,也不参与考核,杜绝与地方政府的利益关系。

当前公众对经济增长、住房、物价、居民收入、就业等与人民福利息息相关的统计数字“怨声四起”,“被增长”、“被就业”则是广大公民心声的真实写照,这都反映出纸上数字与实际感受之间的巨大反差。我们承认某些数字确实存在质量问题,但很多“怨言”和“疑点”相当程度上是由于统计数据的覆盖范围和测算方法不为公众所知所导致。在现有框架下,CPI权重、房屋销售价格及房价指数、固定资产投资和社会零售总额概念的界定等指标都可以增加透明度,如果统计局能够公开更多细节,只会减少不必要和不公正的质疑。当前,对中国政府统计数据质量的质疑日益高涨,这也与我国经济发展所处的阶段有关,统计部门应该理性认识现有问题,认真研究和听取社会公众和专家意见,改革统计管理体制,进一步改进统计工作,增强统计数据真实性、客观性和实用性,还中国经济以本来“面目”。<sup>①</sup>

### 参考文献:

1. 成邦文、董丽娅, 2002《社会经济统计数据的误差分析》,《统计研究》第11期,第61-62页。
2. 成邦文、董丽娅、杨峻, 2000《研究与开发机构统计数据质量与异常点的对数正态分布检验与识别》,《统计研究》第1期,第42-45页。
3. 常宁, 2004《MF的数据质量评估框架及启示》,《统计研究》第1期,第27-30页。
4. 陈振裕, 1988《中国1985年工业普查的质量保障和质量控制》,《统计研究》第1期,第45-48页。
5. 傅德印, 2000《浅论政府统计数据质量控制技术体系》,《统计与信息论坛》第2期,第19-24页。
6. 傅德印、刘晓梅, 1994《贯彻国际标准,建立健全统计数据质量管理与保证体系》,《统计研究》第6期,第47-50页。
7. 傅德印、陶然, 2007《对政府统计数据质量成本的探讨》,《统计研究》第8期,第8-12页。
8. 高敏雪, 2009《从外部监督入手解决统计数据质量问题的努力》,《统计研究》第8期,第50-52页。
9. 管于华、尚卫平, 2003《论GDDS国际标准与中国统计数据缺陷》,《统计研究》第1期,第63页。
10. 金勇进、陶然, 2010《中国统计数据质量理论研究与实践历程》,《统计研究》第1期,第62-67页。
11. 李成瑞, 1997《当前经济成份分类和所有制构成的统计问题》,《经济研究》第7期,第63-67页。
12. 李成瑞, 1984《中国1982年人口普查的质量控制》,《统计研究》第1期,第3-14页。
13. 李成瑞, 1994《关于统计改革的几个问题》,《经济研究》第9期,第38-42页。
14. 刘洪、黄燕, 2007《我国统计数据质量的评估方法研究——趋势模拟评估法及其应用》,《统计研究》第8期,第17-21页。

<sup>①</sup>可幸的是,2010年4月7日国家统计局召开座谈会向社会公众开放CPI的编制方法以及基础数据的采集、加工、汇总过程,迈出了透明统计、阳光统计的重要一步。

15. 刘洪、黄燕, 2009《基于经典计量模型的统计数据质量评估方法》,《统计研究》第3期,第91-96页。
16. 刘海清、熊祖轅, 2009《统计信用与统计数据质量研究》,《统计研究》第12期,第19-22页。
17. 李金昌, 2002《对我国统计调查方法体系改革的回顾与展望》,《统计研究》第7期,第32-35页。
18. 刘孝新、朱慧明、胡先红, 1997《事后预测及反常结果判断法在统计数据质量评估中的应用》,《统计研究》第2期,第69-71页。
19. 林贤郁、方宽, 2008《中外统计规范概览》,中国统计出版社。
20. 林勇、杨言勇, 2008《统计数据质量改进的博弈论研究》,《统计研究》第3期,第106-108页。
21. 卢治飞, 2003《政府统计中造假与治假的博弈分析》,《统计研究》第6期,第55-57页。
22. 李子奈、周建, 2005《宏观经济统计数据结构变化分析及其对中国的实证》,《经济研究》第1期,第15-25页。
23. 孟连、王小鲁, 2000《对中国经济增长统计数据可信度的估计》,《经济研究》第10期,第3-13页。
24. 庞皓, 1999《中国应建立更为完善的宏观经济统计体系》,《经济学家》第2期,第94-99页。
25. 邱东、陈梦根, 2008《基于数据质量观的中国统计能力建设》,《当代财经》第3期,第113-117页。
26. 阙里、钟笑寒, 2005《中国地区GDP增长统计的真实性检验》,《数量经济技术经济研究》第4期,第3-12页。
27. 任若恩, 2002《中国GDP统计水分有多大——评价两个估计中国GDP数据研究的若干方法问题》,《经济学(季刊)》第2卷第1期,第33-52页。
28. 《统计数字质量研究》课题组, 1995《在新旧体制转换过程中如何保证统计数字的准确性和及时性》,《统计研究》第2期,第35-38页。
29. 宋小川, 2007《中国的GDP及其若干统计问题》,《经济研究》第8期,第21-30页。
30. 孙颖等, 1996《现金统计制度存在的问题及改革建议》,《金融研究》第10期,第50-52页。
31. 王华、金勇进, 2009《统计数据准确性评估:方法分类及适用性分析》,《统计研究》第1期,第32-39页。
32. 王静, 2009《我国GDP统计数据质量的量化分析》,《中国统计》第10期,第45-46页。
33. 王强, 1990《关于统计数字质量问题的讨论》,《统计研究》第3期,第59-63页。
34. 王小鲁, 2002《关于中国经济增长率的几点讨论》,《经济学(季刊)》第2卷第1期,第63-76页。
35. 许宪春, 2002《中国国内生产总值核算》,《经济学(季刊)》第2卷第1期,第23-36页。
36. 徐晓海, 2008《中国经济普查的组织机制与数据质量控制》,《统计研究》第11期,第85-91页。
37. 肖宜滨, 1999《提高统计数据质量的法学思考》,《统计研究》第10期,第24-38页。
38. 许永洪、曾五一, 2009《基于中国城市统计数据的CPI偏差估计》,《统计研究》第4期,第3-6页。
39. 颜德伦, 1990《论统计数字的全面质量管理》,《统计研究》第2期,第54-59页。
40. 余芳东, 2002《外国统计数据质量的涵义、管理以及对我国的启示》,《统计研究》第2期,第26-29页。
41. 余芳东, 2004《对世界银行按美元计价的我国GDP数据分析》,《管理世界》第1期,第24-26页。
42. 余根钱, 2008《统计数据全程质量管理初探》,《中国统计》第1期,第7-9页。
43. 袁平德, 2004《虚假统计中的攀比现象》,《中国统计》第3期,第48页。
44. 杨清, 2000《统计数据质量研究新思路——误差研究》,《统计研究》第8期,第33-37页。
45. 杨清、吴伟霞, 2000《对原始资料偏误的监控技术初探:论提高统计数据质量的重要步骤》,《统计研究》第6期,第44-46页。
46. 岳希明, 2005《我国现行劳动统计的问题》,《经济研究》第3期,第46-56页。
47. 张德然, 2003《统计数据中异常值的检验方法》,《统计研究》第5期,第53-55页。
48. 政府统计数据质量研究课题组, 1999《政府统计数据质量问题及对策建议》,《浙江社会科学》第2期,第41-46页。
49. 中国人民银行武汉分行、国家统计局湖北调查总队联合课题组, 2006《关于建立中国核心CPI问题的研究》,《金融研究》第2期,第137-145页。
50. 周黎安, 2004《晋升博弈中政府官员的激励与合作——兼论我国地方保护主义和重复建设问题长期存在的原因》,《经济研究》第6期,第33-40页。
51. 朱启贵, 2009《中国国民经济核算体系改革发展三十年回顾与展望》,《商业经济与管理》第1期,第5-13页。
52. 朱文兴、张继良, 2005《统计管理体制及其模式初探》,《中国统计》第8期,第6-8页。
53. 张新、蔺殿春, 2002《中国经济的增长——GDP数据的可信度以及增长的微观基础》,《经济学(季刊)》第2卷第1期,第1-18页。
54. 朱秀君, 2000《政府统计中造假与治假的博弈分析》,《数量经济技术经济研究》第7期,第63-66页。
55. Bobrow ski M., M. Marre, and D. Yankelevich 1999. "A Homogeneous Framework to Measure Data Quality." Proceedings of the International Conference on Information Quality 115-124
56. Cappiello C., C. Francalanci and B. Pernici 2004. "Data Quality Assessment from the User's Perspective" Proceedings of QIS in Conjunction with the 23rd ACM, SIGMOD International Conference on Management of Data 68-72
57. Costa Dora L. 2001. "Estimating Real Income in the United States from 1888 to 1994 Correcting CPI Bias Using Engel Curves" *Journal of Political Economy*, 109(6): 1288-1310
58. Dalenius T., 1983. *Errors and Other Limitations of Survey, Statistical Methods and the Improvement of Data Quality*. London: Academic Press Inc
59. Hamilton, Bruce W. 2001. "Using Engel's Law to Estimate CPI Bias" *American Economic Review*, 10(3): 619-630
60. Hansen, M. H., W. N. Hurwitz, and M. Bershada 1961. "Measurement Errors in Censuses and Surveys" *Bulletin of the ISI* 38(2).
61. Hausman, Jerry 2003. "Sources of Bias and Solutions to Bias in the Consumer Price Index" *Journal of Economic Perspectives* 5(1): 23-44
62. Huenemann Ralph W. 2001. "Are China's Recent Transport Statistics Plausible?" *China Economic Review*, 12(4): 368-372

63. Keide, A. Bert 2001. "China's GDP Expenditure Accounts" *China Economic Review*, 12(4): 355-367.
64. Klein, L. R., and S. Ozmucur 2002. "The Estimation of China's Economic Growth Rate" *Journal of Economic and Social Measurement* 28(4): 187-202.
65. Lardy, Nicholas R. 2002. "China Will Keep on Growing" *Asian Wall Street*, 12(6): 43-45.
66. Maddison, Angus. 1998. *Chinese Economic Performance in the Long Run*. OECD Development Centre, Paris.
67. Maddison, A. 2001. *The World Economy - A Millennial Perspective*. OECD Development Centre, Paris.
68. Namann, F., and C. Roker 2000. "Assessment Methods for Information Quality Criteria" Proceedings of 5th International Conference on Information Quality.
69. Neyman, J. 1934. "On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection" *Journal of the Royal Statistical Society*, 97(3): 558-606.
70. Onate, B. T. 1995. "Non-sampling Errors in Agricultural Censuses and Surveys: A Major Concern in the Third World" Bulletin of ISI Proceeding of 50th Session, Book 3, 1001-1003.
71. Park, Albert and Sangui Wang 2001. "China's Poverty Statistics" *China Economic Review*, 12(4): 384-398.
72. Perkins, D. 1988. "Reforming China's Economic System" *Journal of Economic Literature*, 26(2): 601-645.
73. Pipino, L., Y. Lee, and R. Y. Wang 2002. "Data Quality Assessment" *Communications of The ACM*, 10(5): 211-217.
74. Rawski, T. 2001a. "What is Happening to China's GDP Statistics?" *China Economic Review*, 12(4): 347-354.
75. Rawski, T. 2001b. "China's GDP Statistics: A Case of Caveat Lecto?" *China Economic Quarterly*, 12(5): 18-22.
76. Ryu, K. S., J. S. Park, and J. H. Park 2006. "A Data Quality Management Maturity Model" *ETRI Journal* 28(3): 181-201.
77. Schaping, Thomas 2001. "Hide-and-seek China's Elusive Population Data" *China Economic Review*, 12(4): 323-332.
78. Sinton, Jonathan E. 2001. "Accuracy and Reliability of China's Energy Statistics" *China Economic Review*, 12(4): 373-383.
79. Szirmai, A., R. Ren, and M. Bai 2005. "China's Manufacturing Performance in Comparative Perspective, 1980-1992" Research Memorandum No. 581, Institute of Economic Research.
80. Wang, R., T. Allen, and W. Harris 2003. "An Information Product Approach for Total Information Awareness" Proceedings of IEEE Aerospace Conference, 1-15.
81. Wang, R. Y., and D. M. Strong 1996. "Beyond Accuracy: What Data Quality Means to Data Consumers" *Journal of Management Information Systems* 12(4): 5-34.
82. Wang, R. Y., V. Storey, and C. Firth 1995. "A Framework for Analysis of Data Quality Research" *IEEE Transactions on Knowledge and Data Engineering*, 28(4): 623-640.
83. Wang, R. Y. 1998. "A Product Perspective on Total Data Quality Management" *Communications of the ACM*, 10(2): 58-65.
84. Wang, R. Y., D. M. Strong, and B. K. Kahn 1999. "An Information Quality Assessment Methodology" Proceedings of the International Conference on Information Quality.
85. Warner, S. L. 1965. "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias" *Journal of the American Statistical Association*, 309(60): 63-69.
86. Wu, Harry X. 1997. "How Fast Has Chinese GDP according to the National Accounts Concepts of Value Added: The Industry Sector, 1949-1994" COPPA Paper 4.
87. Wu, Harry X. 2002. "How Fast Has Chinese Industry Growth? Measuring the Real Output of Chinese Industry, 1949-1997." *Review of Income and Wealth*, 4(6): 23-27.
88. Xu, Xianchun, Yanfei Ye, and Derek Blades 2000. *National Accounts for China, Sources and Methods*. OECD Publishing.
89. Yang, L., R. Y. Strong, and B. K. Kahn 2002. "AMQ: a Methodology for Information Quality Assessment" *Information & Management*, 15(2): 134-143.

## Research on Quality of the China's Statistics: A Literature Review

Zhao Xuegang<sup>1</sup>, Wang Xuebin<sup>1, 2</sup> and Liu Kangbing<sup>1</sup>

(1: School of Economics, Shanghai University)

(2: Research Center for Urban Planning & Development, Fudan University)

**Abstract** In recent years, the quality of China's statistics has aroused mounting concern by both domestic and foreign scholars and a large number of literatures are springing up, focusing mainly on the definition of data quality and assessment standards, distortion mechanisms, quality control technologies, evaluation methods and management theories of statistical data. Some empirical findings and proposals to improve the data quality are also given. On the basis of these literatures, this paper holds that to solve the problem of data distortion essentially, we must begin with the reform of the current statistics management system to get rid of the interference from local government and ensure the independence of statistics. And we point out that the increase in the degree of openness to statistical process, such as CPI weights, housing sales price and price index, investment in fixed assets etc., may help reduce unnecessary and unjust doubts about the quality of statistics effectively.

**Key Words** Quality of the Statistics; Theory Research; Empirical Evidence; Countermeasure

**JEL Classification** C42, C82, C83

(责任编辑: 孙永平、陈永清)