

统计数据质量与用户满意度： 测评量表设计与实证研究^{*}

王 华 金勇进

内容提要:本文针对政府统计数据质量评估问题,综合若干国家统计局部门与国际组织对统计质量的定义,确立政府统计数据质量维度体系;进而参照国际货币基金组织的“数据质量评估框架”设计了具体的用户满意度测评量表。基于该量表的问卷调查结果显示,量表具有良好的测量信度与效度。研究中提出的八项质量维度可以充分反映政府统计数据质量的不同侧面,可以有效解构用户满意度的特征;用户对各质量维度的重要性认知、用户专业技术职务、对统计数据的依赖程度以及搜集渠道的选择等,与其满意度之间亦表现出明显的关联模式。

关键词:统计数据质量;用户满意度;测评量表

中图分类号:C812 文献标识码:A 文章编号:1002-4565(2010)07-0009-09

Statistical Data Quality and Users' Satisfaction: Designing Evaluation Scale and Empirical Study

Wang Hua & Jin Yongjin

Abstract: Concerning assessment of official statistical data quality, a quality dimension framework has been determined, with reference to definition of statistical quality by several national statistical authorities and international organizations. Referring to DQAF by IMF, a user-satisfaction evaluation scale has been designed. The evaluation scale has adequate reliability and validity, according to test to questionnaire interview data. The paper argues that the eight quality dimensions can reflect various aspects of official statistical service quality, and deconstruct the structure of user satisfaction. In addition, user characteristics, such as importance perception to quality dimensions, the title of a technical post, degree of dependence on statistics and data collection channels, appear evident associate pattern with their satisfaction.

Key words: Statistical data quality; Users' satisfaction; Evaluation scale

一、问题的提出与概念界定

(一) 问题的提出

统计数据信息是国内外各类宏微观经济主体进行科学决策的重要依据。我国政府统计部门把提高统计数据质量摆到了统计事业发展的战略地位,一直致力于加强质量管理。但在关于统计数据质量的评估工作中,除了对数据准确性的评估研究在数据用户(经济学者)中有所开展以外(王华和金勇进,2009),基本都是在统计部门内部自行实施的;评估多采用系统内部自我检查的方式,评估过程缺乏必要的公众参与和信息反馈,透明度不高,加之缺乏有效的定量评估手段,导致评估结果难以获得社会各界的认可。对于政府统计数据质量的现实水平一直

存在极大争议,甚至影响到政府统计部门的社会声誉与我国政府的国际形象。

随着统计质量概念内涵的不断扩展,各国政府统计部门和有关国际组织充分注重统计数据信息对用户需求的满足程度,开始基于用户角度来评判统计数据质量。目前已经有多个国家的政府统计部门在数据质量管理框架内开展了用户满意度调查(Ehling & Körner 2007),为有效管理和改进统计数据质量提供重要信息;这也应该是我国政府统计数据质量评估的发展方向。通过用户满意度调查来测

^{*} 本文为国家自然科学基金项目“我国统计数据质量的评估系统研究”(项目编号:70671104)阶段性成果。

量用户对统计数据质量的感知状况,符合近十年来新公共管理运动所关注的政府服务的“顾客导向”,可以推动政府统计部门确立对社会公众负责、重视服务质量的管理理念和运行机制,也有助于建立统计部门与用户之间的有效沟通机制,为统计质量评估与控制提供一条重要的信息渠道,切实推动政府统计部门提高质量管理的水平。

从目前开展的统计数据用户满意度调查的情况来看,基本仍处于经验层面的、以实践探索为主的阶段,尚无成熟的方案可供借鉴利用,也没有一般化的理论与方法论研究成果问世。本文在参考相关实践与标准的基础上,采用企业管理层面顾客满意度研究与公共管理层面公众满意度研究的成熟范式,探讨针对政府统计部门所提供的统计数据信息开展用户满意度测评的方法,重点是设计测评量表探查用户满意度的结构与特征,以此从用户角度有效刻画我国政府统计数据质量的现状,并为相关研究与实践奠定一定的方法论基础。

(二)概念界定

统计数据信息的提供涉及统计调查设计与实施、统计数据编制与发布等诸多环节,“统计数据质量”的概念范畴因而也可涵盖统计数据产品、统计数据生产过程以及统计数据生产的制度环境等多个方面的质量,其具体构成将在下一部分专门讨论。

本文所指“用户”是统计数据信息的使用者,与市场营销领域“顾客”这一概念的范畴相当,具体包括在日常活动与行为决策中需要使用统计数据信息的各类政府部门、国际组织、学术机构、社会团体、企业与普通居民户,以及各类组织机构中的具体个体。

“满意”首先是心理学中的概念,在由 Cardozo (1965)引入市场营销领域之后,才产生了“消费者满意度”、“顾客满意度”等可用于从顾客角度表征企业产品和服务质量的主观概念。目前学术界普遍公认“顾客满意感是顾客需求得到满足后的一种心理反应,是顾客对产品和服务的特征或其本身满足自己需求程度的一种判断(Oliver,1997)”。而政府统计数据“用户满意度”则是指统计数据用户对政府统计部门提供的统计数据产品和服务满足自己需求程度的判断,是对统计数据质量的一种主观感知。从测量操作角度出发,此处将用户满意度界定为用户对自己以往使用统计数据经历的总体累积性满意程度;而在讨论具体的统计质量维度时,本文

对用户满意度与质量感知两个概念不作特别区分。

二、测评量表设计与数据收集

(一)统计数据质量维度的确定

统计数据质量是一个多维度的概念,并不简单等同于统计数据的准确性。要有效测评统计数据质量,前提是确立一套合理的统计数据质量概念维度体系。

在统计数据质量的概念界定中,很多国家的统计部门以及有关国际组织采用了 ISO 8402 的质量定义:“质量是产品或服务所具备的满足明确或隐含需求能力的特征和特性的总和”,强调从用户的角度来界定质量的水准,注重统计数据信息对用户需求的满足程度;由此将统计数据质量逐渐扩展为一个内涵较为宽泛的概念,确定了包括适用性、准确性、及时性、可取得性、可解释性和一致性等特征在内的统计数据质量概念体系,如表 1 所示。余芳东(2002)则从统计数据用户、生产者和被调查者三个角度,总结出了适用性、准确性、及时性、可比性、可衔接性、可取得性、可解释性、客观性、方法健全性、有效性以及减轻调查负担共 11 个统计数据质量的维度,涵盖了统计数据生产的制度、过程与产品等各个方面。

表 1 若干国家统计部门与国际组织定义的统计数据质量维度

质量维度	加拿大	欧盟	瑞典	澳大利亚	韩国	IMF	OECD
适用性/内容	√	√	√	√	√	√	√
准确性/可靠性	√	√	√	√	√	√	√
及时性与准时性	√	√	√	√	√	√	√
可得性/可用性	√	√	√	√	√	√	√
可解释性/清晰性	√	√	√	√	√	√	√
一致性/可比性	√	√	√	√	√	√	√
方法健全性		√				√	
制度环境/可信性/客观性		√		√		√	√
(成本)有效性		√			√		√

注:各国对不同统计数据质量维度的划分和定义存在差异,根据意义的相近性予以合并或拆分表述。

资料来源:各国统计部门与国际组织的官方网站;亦可参见文献余芳东(2002)、ABS(2009)、Eurostat(2009)、IMF(2003)、OECD(2003)、Rosen & Elvers(1999)、Statistics Canada(2003)及 Statistics Korea(2009)。

上述统计数据质量维度大体可以归纳为四类:第一类是适用性、准确性与一致性,涉及统计数据产品本身的特性,可定义为统计数据的产品质量维度;第二类是及时性、可得性与可解释性,涉及统计数据

发布环节的特性,可定义为统计数据的发布质量维度;第三类是方法健全性与制度环境,涉及统计数据生产的方法与制度保障,可定义为统计生产的环境质量维度;第四类是(成本)有效性、减轻调查负担等,涉及统计生产的资源约束条件,可定义为统计数据生产的投入质量维度。

严格来讲,从用户角度定义的统计数据质量仅包括统计数据的产品质量与发布质量,因为只有这两类质量维度关乎用户需求,用户也仅能对这两类属性有直接感知。因此,本文重点将统计数据质量维度确定为统计数据的产品质量与发布质量维度;表 1 也充分表明,这两类内容是得到各国统计部门与国际组织普遍公认的统计数据质量维度。进一步,考虑到提升社会形象对于当前我国统计部门的日常运作与质量管理有着非常重要的意义,用户对统计方法与制度环境质量的主观感知又与政府统计的社会形象紧密相关,因此也将其纳入测评体系。

(二) 测评量表设计

针对统计数据的产品、发布和环境三类质量维度,本研究参照国际货币基金组织(IMF)的“数据质

量评估框架(DQAF)”设置具体的测评项目。我国加入 IMF 出台的“通用数据公布系统(GDDS)”以后,遵照同样由 IMF 发布的“数据质量评估框架(DQAF)”来进行统计数据质量管理与评估,无疑有助于达成 GDDS 的各项要求。DQAF 涵盖了与数据质量相关的统计体系管理、统计程序和统计产品的各个方面;此框架根植于联合国官方统计基本准则,并在 IMF 关于数据公布的动议 SDDS(专用数据公布系统)和 GDDS 下完善成形。DQAF 综合了数据公布动议中的良好作法,并且是经过与多方进行深入细致磋商的结果,为进行现行作法与最佳作法的对比评估提供了一个良好框架。

“数据质量评估框架(DQAF)”中有关统计数据质量维度的划分,与本研究中的划分标准有很大不同,但框架内各维度下的具体评估要素和指标则为本研究中各质量维度的测评提供了良好的参照标准。根据本研究中的质量维度架构,从 DAQF 中筛选提炼评估要素和指标,并设计具体的测评项目,最终形成 8 项质量维度与 25 个对应的测评项目(如表 2 所示),由其构成调查问卷的主体内容。在量表设

表 2 统计数据质量用户满意度测评量表

质量维度	测评项目	参照标准
ξ_1 制度规范性	X_{11} 现有统计法规对统计数据生产过程的规范效力	DQAF 0.1
	X_{12} 统计部门的数据质量控制意识与措施	DQAF 0.4
	X_{13} 统计数据生产过程的客观性与中立性	DQAF 1.1
ξ_2 方法健全性	X_{21} 统计调查及数据编制方法的合理性与健全性	DQAF 2/3.3
	X_{22} 统计概念、范围及分类体系的科学性	DQAF 2.1-2.3
	X_{23} 统计数据生产过程的公开性与透明度	DQAF 1.2
ξ_3 适用性	X_{31} 统计数据的信息量、内容涵盖的全面性	DQAF 0.3
	X_{32} 统计数据的发布形式(图、表、电子文档格式等)	DQAF 5.1.1-5.1.2
	X_{33} 统计部门的信息咨询服务	DQAF 5.1.5/5.3
ξ_4 准确性	X_{41} 统计数据可以准确反映实际情况的程度	DQAF 3.1
	X_{42} 统计部门关于统计数据质量的评估结果的可信性	DQAF 3.2/3.4
	X_{43} 统计数据修正规则和程序的合理性与科学性	DQAF 3.5/4.3
ξ_5 及时性	X_{51} 统计数据的发布速度	DQAF 4.1
	X_{52} 统计数据的发布频率(年度、季度或月度)的合理性	DQAF 4.1
	X_{53} 统计数据发布时间的事先告知程度	DQAF 5.1.3
ξ_6 可得性	X_{61} 统计数据获取途径和方式的多样性与便利性	DQAF 5.1.2/5.3
	X_{62} 统计数据产品和信息咨询服务的收费标准	DQAF 5.3.2
	X_{63} 统计数据获取的公平程度	DQAF 5.1.4
	X_{64} 统计部门对用户特殊需求的满足程度	DQAF 5.1.5/5.3
ξ_7 可解释性	X_{71} 关于统计数据概念、范围的解释性文字的充分性	DQAF 5.2.1
	X_{72} 关于统计调查和估算方法的说明文字的充分性	DQAF 5.2.1
	X_{73} 关于统计口径调整和数据修正的说明文字的充分性	DQAF 4.3
ξ_8 一致性	X_{81} 统计数据的关联项目之间的一致性	DQAF 4.2.1
	X_{82} 统计数据在时间维度上的连续性	DQAF 4.2.2
	X_{83} 不同来源、部门、项目的统计数据的可比性	DQAF 4.2.3

计过程中,结合个人访谈来判断具体问项的适宜性,对问项中的论述模糊、意义交叉等问题进行了识别和调整。问卷采用10级Likert量表来反映用户对各问项的质量感知(满意度)。

在表2中,制度规范性维度反映统计数据生产过程中的法规效力、质量控制与立场客观性,是对表1中制度环境维度的重新表述,用户对该维度的质量感知实际反映了统计部门的社会形象;方法健全性维度反映了统计调查与数据编制方法的科学合理性,对国际通用的统计概念、标准、范围和分类体系的采纳程度,以及与此相适应的统计生产过程的公开透明性;适用性维度反映统计数据信息的完整性、有用性,信息发布方式的适宜性以及对用户信息咨询的满足程度;准确性维度反映统计估算值与目标特征真值的接近程度、以及相关数据质量评估与修正操作的科学可信性;及时性维度反映统计数据发布的速度、频率的合理性以及数据发布时间表的透明度;可得性维度反映统计部门便利、公平地向用户提供统计数据及满足其特殊信息需求的程度;可解释性维度反映关于统计数据及其调查估算和调整修正方法的解释说明性信息的提供程度;一致性维度反映统计数据在空间、时间及项目间的一致可比性。

(三) 数据收集与预测试

受研究条件所限,从实施可行性角度考虑,本研究仅针对高校学者(经济管理类专业的教师与研究生)进行调查,作为此领域研究的一次局部尝试。经济管理类学术研究的严谨性与经验依赖特征,决定了经济管理类学者对统计数据有更为严格的质量要求与更为丰富的使用经历,问卷调查的信息因而具有更高的可参考性。自然,此研究模式也可容易地推广到其他用户群体。

在正式调查之前,首先在较小范围内对问卷结构进行预测试,以进一步消除问卷的歧义和不明确之处。试调查中发放问卷60份,回收53份。基于Cronbach's α 系数的问卷信度检验结果表明,除适用性维度的 α 系数为0.747外,其他质量维度的 α 系数都超过了0.8,问卷结构具有较高的内部一致性。基于探索性因子分析的效度检验结果表明,在按照累计方差贡献度80%的原则提取公共因子,并对因子载荷矩阵进行方差最大化旋转后,各测评项目基本按照预期归并为相应的因子(质量维度),载荷系数基本都大于0.5,测量量表具有必要的结构效度;

只有 X_{23} 的载荷小于0.4, X_{33} 错归入准确性因子,将二者予以删除,以提高问卷的聚合效度与区分效度。调整后的8个维度的23个问项构成正式调查问卷的核心部分。

正式调查在我国五省市的七所高校发放问卷340份,回收319份,剔除受访者回答“从不使用统计数据”的问卷与关键数据整体缺失的问卷16份,得到有效问卷303份,有效回收率为89.1%。在正式检验与数据分析之前,对于问卷中的项目数据缺失进行插补——考虑到同一用户对不同问项的质量感知具有较高的相关性,具体插补方法是以同一份问卷中无缺失问项的回答值的均值来替代缺失值。

三、测评量表的正式检验

根据正式调查问卷的数据结果来检验测评量表的信度与效度,以确保所研究的概念得到可靠和正确的测量。信度是指问卷问项测量的一致性与稳定性,本研究主要采用Cronbach's α 系数与组合信度指标来进行问卷信度检验。 α 系数通过计算测评项目间的平均相关系数来判定问卷的内部一致性(Cronbach, 1951);根据Nunnally(1978)的建议, α 系数大于0.7时信度是可接受的。由于Cronbach's α 系数只是测量信度的一个下限估计,会造成对真实信度的低估;而当存在相关测量误差或存在多个潜变量时, α 系数是不适合的(Bollen, 1989),为此研究人员建议采用组合信度(Fornell & Larcker, 1981)来替代之。组合信度也用于评估测量的内部一致性,当大于0.7的标杆值时被认为是适当的。

效度指问卷能够测量出研究者想要测量的主题的程度,主要包括内容效度与结构效度。本研究中问卷问项的设计广泛参考了有关统计机构的良好实践,且经过预先访谈和测试,可认为具有相当的内容效度。对于结构效度的检验则可考察聚合效度与区分效度两项标准。聚合效度可由潜变量的标准化因子载荷以及平均方差提取量(AVE)来说明,标准化载荷系数显著大于0.5,或者AVE在0.5以上,表示对潜变量的测量有足够的聚合效度(Fornell & Larcker, 1981; Steenkamp & van Trijp, 1991);而区分效度检验则要求任一潜变量的AVE大于该潜变量和其他各潜变量之间相关系数的平方。

本研究的信度与效度检验都基于确证性因子分析结果进行。应用SAS软件的CALIS过程得到确证性因子分析的结果如表3所示。

表 3 量表问项的确定性因子分析结果

质量维度	测评项目	载荷系数	t 统计量	测量误差	α 系数	组合信度	AVE
ξ ₁ 制度规范性	X ₁₁	0.700		0.715	0.835	0.838	0.636
	X ₁₂	0.851	5.34	0.526			
	X ₁₃	0.833	5.11	0.554			
ξ ₂ 方法健全性	X ₂₁	0.866		0.500	0.813	0.815	0.689
	X ₂₂	0.792	5.36	0.611			
ξ ₃ 适用性	X ₃₁	0.895		0.446	0.723	0.749	0.605
	X ₃₂	0.640	2.21	0.768			
ξ ₄ 准确性	X ₄₁	0.868		0.497	0.870	0.878	0.707
	X ₄₂	0.906	9.03	0.423			
	X ₄₃	0.739	4.86	0.674			
ξ ₅ 及时性	X ₅₁	0.886		0.463	0.829	0.842	0.643
	X ₅₂	0.846	7.10	0.533			
	X ₅₃	0.655	2.94	0.756			
ξ ₆ 可得性	X ₆₁	0.813		0.582	0.840	0.842	0.584
	X ₆₂	0.676	3.17	0.737			
	X ₆₃	0.789	4.96	0.674			
	X ₆₄	0.771	5.01	0.637			
ξ ₇ 可解释性	X ₇₁	0.833		0.553	0.885	0.886	0.720
	X ₇₂	0.867	7.50	0.498			
	X ₇₃	0.846	7.03	0.533			
ξ ₈ 一致性	X ₈₁	0.763		0.646	0.803	0.807	0.583
	X ₈₂	0.720	3.75	0.694			
	X ₈₃	0.805	5.22	0.593			
模型拟合指数	$\chi^2/df = 2.392$ RMSEA = 0.068 GFI = 0.878 AGFI = 0.833 CFI = 0.935 NFI = 0.894 NNFI = 0.918						

注:表中载荷系数为标准化估计值,t 统计量值是基于标准化载荷系数等于 0.5 的原假设得出的计算结果。

由表 3 可知,确定性因子分析的整体拟合效果比较理想, χ^2 与其自由度的比值为 2.392,小于允许上限 3;近似误差均方根 RMSEA 的值为 0.068,小于允许上限 0.08;绝对拟合指数 GFI 与 AGFI 都超过 0.8;相对拟合指数 CFI 与 NNFI 都超过了 0.9,NFI 也接近 0.9,表明测量模型是有效的。在信度方面,各维度的 Cronbach's α 系数与组合信度系数均大于 0.7,表明量表具有良好的内部一致性。在效度方面,各问项的标准化因子载荷系数都显著大于 0.5,各潜变量的平均方差提取量(AVE)也均超过了 0.5 的下限,表明各测量项目具有较好的聚合效度;比较潜变量 AVE 与该潜变量和其他各潜变量之间相关系数的平方,结果表明除一致性维度的 AVE 为 0.583,略小于一致性与可解释性二者之间的相关系数 0.781 的平方 0.610 以外,其余质量维度的 AVE 均大于彼此之间相关系数的平方,可知各潜变量的测量区分效度也是可以接受的。

以上分析表明,本研究的测量问项基本可以通

过信度与效度检验,质量维度及测评量表的设计是较为合理的,调查得到的数据可以用于进一步的研究分析。

四、用户满意度比较研究

利用调查结果中各统计数据质量维度的用户满意度信息,结合调查中得到的用户背景信息(如用户专业技术职务、对统计数据的依赖程度、数据搜集渠道的选择等),以及用户对各质量维度的重要性认知与综合满意度,可以进行综合比较分析,全方位考察统计数据质量的用户满意特征,为有效改进统计数据质量提供有益借鉴。

(一) 用户重要性认知与满意度的综合比较

在利用量表测评用户满意度外,问卷中也设置了简单的 10 个问项,用以测评用户的重要性认知。各问项与质量维度保持了一对一或二对一关系,对前者采用加权平均方法进行综合,从而得到用户对各质量维度重要性的认知评价。

确定性因子分析给出了各因子(质量维度的用户满意度)的最终得分,此得分是零均值的。为了实现各因子间的横向比较,采用顾客满意度测评实践中最常用的标准化方法

$$\begin{aligned} \bar{\xi}_j^* &= \frac{\bar{\xi}_j - \min\{\xi_j\}}{\max\{\xi_j\} - \min\{\xi_j\}} \times 10 \\ &= \frac{\bar{\xi}_j - \min\{\xi_j\}}{\max\{\xi_j\} - \min\{\xi_j\}} \times 10 \end{aligned}$$

得到 10 分制的平均因子得分。各统计数据质量维度的用户重要性认知与满意度的计算结果如表 4 所示。

表 4 各统计数据质量维度的用户重要性认知与满意度

质量维度	重要性认知		满意度	
	平均值	标准差	平均值	标准差
ξ ₁ 制度规范性	7.818	2.116	5.104	1.922
ξ ₂ 方法健全性	8.230	1.933	5.451	1.969
ξ ₃ 适用性	7.435	2.053	5.402	1.945
ξ ₄ 准确性	8.447	1.833	4.985	1.905
ξ ₅ 及时性	8.039	1.889	4.944	1.824
ξ ₆ 可得性	8.425	1.820	4.168	1.933
ξ ₇ 可解释性	8.082	1.856	5.273	1.943
ξ ₈ 一致性	8.695	1.524	4.796	1.878

如果将用户对各质量维度的重要性认知分值视为用户对统计数据质量的期望水平,那么用户的实际质量感知(满意度)与期望水平之间还存在明显的差距。为了探查统计数据质量的内部特征,为统

计部门实施有针对性的统计质量改进工作提供指导,可以借助两类评分值进行重要性—满意度对比分析,如图1所示。

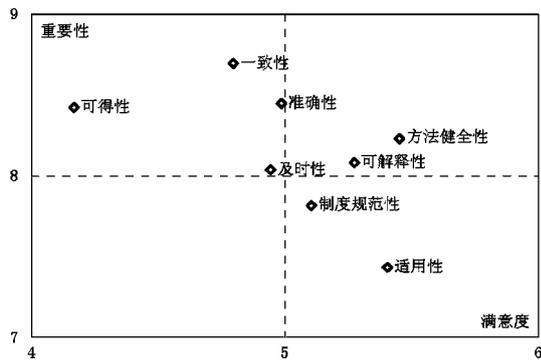


图1 统计数据质量维度的重要性—满意度矩阵

从图1中重要性认知的单一坐标轴方向看,用户对各质量维度的重要性认知水平依次为一致性、准确性、可得性、方法健全性、可解释性、及时性、制度规范性,最后是适用性。高校学者应用统计数据的主要目的多在于对比分析与计量建模,因此对于表征统计数据产品质量的一致性与准确性维度有较高的要求和期望。其次如可得性、可解释性与及时性等表征统计数据发布质量的维度,是影响统计数据应用效率的关键要素,用户重要性认知水平略低。用户对方法健全性维度虽然没有直观感受,但高校学者对统计数据生产过程与方法的了解程度相对较高,也更能体会其在保障统计数据质量方面的重要性。制度规范性维度所反映的统计数据生产的制度环境,由于已经超出统计数据质量的概念范畴,与统计数据质量的对应关系不强,用户的重要性认知相对较低。适用性维度的用户重要性认知水平最低,反映出当前阶段统计数据信息的“格式质量”(如内容完整性与格式适宜性)还不属于用户的关注重点。

结合各质量维度用户满意度的信息则可发现,用户满意度与重要性认知水平之间存在明显的负相关关系,表现出用户重要性认知水平越高,则满意度越低的趋势。如果将重要性—满意度矩阵划分为四个区域,一致性与可得性所在的区域可界定为亟待改进区,这两个质量维度的用户重要性认知水平最高,但满意度却最低,因而亟需统计部门重点加以改进;准确性与及时性处于亟待改进区的边缘,满意度居于中等水平,是其次需要考虑改进的方面;方法健

全性与可解释性所在的区域可界定为质量优势区,这两个质量维度的用户重要性认知程度较高,同时满意度也较高,是统计部门在质量管理中较为先进而需保持的方面;制度规范性与适用性所在区域可界定为“锦上添花”区,属于用户不太注重、但满意度较高的方面。综合而言,统计数据质量的改进重点,首先在于可得性与一致性,其次是准确性与及时性,再次是可解释性与方法健全性,制度规范性与适用性则在条件允许下适度加以保障。

(二) 不同专业技术职务用户的满意度比较

为了发现统计数据质量与用户满意度的更多模式,把握用户特征与满意度之间的关联,本研究也进行了针对不同类型用户的分类比较分析。

在用户专业技术职务方面,本次调查的受访者中有教授17人、副教授28人、讲师72人、博士研究生158人、其他类型28人,比重分别为5.6%、9.2%、23.8%、52.2%和9.2%。区分教授—副教授—讲师—博士研究生四类用户进行方差分析,结果显示四类用户间的满意度差异不显著;但数据结果中揭示了明显的差异模式——职务越高则满意度越高,统计不显著的原因可能在于各组内的回答方差相对较大。进一步观察发现副教授与教授间的满意度差异较大,与讲师则较为接近,为此将副教授与讲师合并为一组,方差分析结果如表5所示。

表5 不同专业技术职务用户的满意度比较(方差分析)

	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8
教授	5.728	6.044	5.995	5.737	5.777	5.115	6.131	5.448
副教授/讲师	5.290	5.575	5.371	5.075	5.042	4.202	5.375	4.943
博士研究生	4.878	5.276	5.267	4.856	4.747	3.954	5.025	4.559
组间均方	9.010	6.124	4.106	6.511	9.357	10.919	11.249	8.890
组内均方	3.634	3.774	3.749	3.560	3.352	3.701	3.677	3.464
F值	2.479*	1.623	1.095	1.829	2.792*	2.951*	3.059**	2.566*
p值	0.086	0.199	0.336	0.163	0.063	0.054	0.049	0.079

注:*、**和***分别表示在10%、5%和1%水平下具有统计显著性;表6—表9同。

由表5可知,各组用户对数据可解释性(ξ_7)的满意度在5%水平下差异显著,对制度规范性(ξ_1)、及时性(ξ_5)、可得性(ξ_6)、一致性(ξ_8)的满意度在10%水平下差异显著;而即使是方法健全性、适用性与准确性($\xi_2 - \xi_4$)这三个差异不显著的质量维度,也同样表现出用户职务越高则满意度越高的模式。究其原因进行分析,用户专业技术职务越高,意味着其从事研究的时间越长、使用统计数据的经验越丰

富, 因此对统计数据生产的制度环境与资源条件更了解, 对统计数据编制规则更熟悉, 在搜集统计数据的过程中得到统计部门支持的能力也更强; 所有这些优势都是保障用户正确有效使用统计数据的前提基础, 从而奠定了用户对统计数据质量较为正面的感知印象。另一方面, 表 5 结果也暗示了统计部门根据用户能力与需求的差异而提供有针对性的解释说明信息、并保证统计信息服务公平性的必要性。

(三) 不同数据依赖程度用户的满意度比较

本次调查样本中, 将统计数据视为自己研究中的“基础性/必备资料”的受访者有 199 人, 视为“辅助性/参考资料”的有 104 人, 比重分别为 65.7% 和 34.3%。对两组用户的满意度均值的差异进行显著性检验, 结果如表 6 所示。

表 6 不同数据依赖程度用户的满意度比较 (均值差异检验)

	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8
基础性/必备资料	5.263	5.601	5.397	5.103	5.005	4.196	5.454	4.847
辅助性/参考资料	4.802	5.163	5.412	4.761	4.826	4.113	4.928	4.699
均值差	0.461	0.438	-0.015	0.341	0.179	0.083	0.526	0.148
标准误	0.231	0.237	0.236	0.230	0.221	0.234	0.234	0.227
t 值	1.992**	1.845*	-0.065	1.483	0.809	0.354	2.252**	0.649
自由度	301	301	301	301	301	301	301	301
p 值	0.047	0.066	0.948	0.139	0.419	0.723	0.025	0.517

注: 均值差异显著性检验分别针对“方差相等”与“方差不等”两种假设进行, 根据方差齐性检验实际结果 (5% 显著性水平) 在表中予以选择性列示; 表 7 和表 8 同。

由表 6 可知, 除适用性维度 (ξ_3) 外, “基础性/必备资料”组用户的满意度普遍高于“辅助性/参考资料”组; 同时对制度规范性 (ξ_1) 与可解释性 (ξ_7) 的满意度在 5% 水平下差异显著, 对方法健全性 (ξ_2) 的满意度在 10% 水平下差异显著。统计数据依赖程度对用户满意度的影响, 与专业技术职务对用户满意度的影响表现出相似的模式。实际上, 交叉分析表明专业技术职务越高的用户越倾向于将统计数据视为基础性/必备资料, 该比重在讲师和博士研究生中约为 64%, 在副教授中为 71.4%, 在教授中则为 82.4%。因此可以认为, 用户对统计数据依赖程度的不同往往意味着其使用统计数据的频率不同、经验有别, 经常使用统计数据的用户对统计数据生产的制度环境、技术方法及概念规则等更为了解, 可以更为有效、正确地使用统计数据, 从而满意度更高。

(四) 不同数据搜集渠道用户的满意度比较

本次调查中列举了六种数据搜集渠道, 即正式的统计出版物、统计部门或专业机构网站、专业电子数据库、统计公报与专题调查报告、自行实施调查以及直接联系统计部门获取数据, 利用各渠道的受访者比重分别为 67%、71.6%、56.1%、65.4%、33.7% 和 17.8%。用户可以同时利用多种渠道搜集统计数据, 因而各种渠道之间并不是互斥的。但对特定的渠道而言, 根据是否利用该渠道将用户划分为“利用”与“未利用”两组, 如果两组用户之间的满意度具有显著差异, 在其他条件相同的情况下, 可以认为满意度差异与该渠道的特性有关。

分别针对每种渠道进行两组用户之间满意度均值的差异检验, 结果表明, 统计出版物、统计部门网站、专业电子数据库以及直接联系统计部门四种渠道的用户满意度差异不显著。

表 7 给出利用“统计公报与专题调查报告”渠道与否的两组用户的比较结果, 其中对于一致性维度 (ξ_8) 的满意度, “利用”组在 5% 水平下显著低于“未利用”组; 对于准确性 (ξ_4) 与可解释性 (ξ_7) 的满意度, “利用”组在 10% 水平下显著低于“未利用”组。这在一定程度上显示, 统计公报与专题调查报告虽然保障了数据发布的及时性, 但由于未经必要的统计核对与修正, 不同时期公报或报告中的数据易出现不一致、不可比的现象, 准确性也难以保证; 同时, 统计公报等注重行文连贯性与版式紧凑感, 通常不提供关于统计数据概念范围与编制方法的必要解释说明, 因而不利于读者准确把握统计数据的内涵和外延。

表 7 利用“统计公报”渠道与否两组用户的满意度比较 (均值差异检验)

	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8
利用	5.037	5.437	5.310	4.847	4.901	4.074	5.127	4.606
未利用	5.231	5.478	5.575	5.247	5.024	4.345	5.549	5.155
均值差	-0.194	-0.041	-0.265	-0.400	-0.123	-0.272	-0.421	-0.549
标准误	0.232	0.238	0.235	0.229	0.221	0.233	0.234	0.225
t 值	-0.837	-0.172	-1.129	-1.743*	-0.556	-1.166	-1.802*	-2.440**
自由度	301	301	301	301	301	301	301	301
p 值	0.403	0.864	0.260	0.082	0.579	0.245	0.073	0.015

表 8 给出利用“自行实施调查”渠道与否的两组用户的比较结果, 其中对于各维度的满意度, “利用”组都在 5% 甚至 1% 水平下显著低于“未利用”组。该结果可从两个方面加以解读: 一是对政府统计数据质量比较不满的用户倾向于选择自行搜集数

表 8 “自行实施调查”与否两组用户的满意度比较(均值差异检验)

	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8
利用	4.731	4.975	4.924	4.461	4.526	3.700	4.627	4.222
未利用	5.291	5.687	5.636	5.247	5.150	4.396	5.591	5.079
均值差	-0.560	-0.713	-0.711	-0.787	-0.624	-0.696	-0.965	-0.856
标准误	0.232	0.237	0.233	0.228	0.220	0.232	0.230	0.223
t 值	-2.410 **	-3.012 ***	-3.047 ***	-3.451 ***	-2.841 ***	-2.999 ***	-4.195 ***	-3.834 ***
自由度	300	300	300	300	300	300	300	300
p 值	0.017	0.003	0.003	0.001	0.005	0.003	0.000	0.000

表 9 统计数据质量维度的用户满意度与用户综合满意度(方差分析)

	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8
非常不满意	3.468	3.903	3.716	3.132	3.489	2.612	3.490	2.808
比较不满意	4.288	4.798	4.815	4.191	4.442	3.446	4.736	4.008
一般	5.494	5.722	5.639	5.423	5.142	4.370	5.516	5.156
比较满意	6.669	7.040	7.065	6.448	6.391	6.223	6.855	6.660
组间均方	79.536	66.879	70.140	83.041	52.387	88.722	68.260	99.986
组内均方	2.932	3.245	3.117	2.833	2.836	2.883	3.128	2.559
F 值	27.125 ***	20.613 ***	22.503 ***	29.311 ***	18.470 ***	30.774 ***	21.819 ***	39.074 ***
p 值	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

据,二是自行实施调查的用户对自己的数据质量较有把握,更反衬出对政府统计数据质量的不满;二者之间存在一定的相互因果关系。不论从哪一方面来看,都意味着因统计数据质量不佳而导致统计数据的利用效率低下与统计资源的浪费。从某种意义上讲,这也显示了发展民间统计来弥补官方统计不足的必要性。

(五) 关联分析

在上述测评项目之外,本研究亦设置问项探查用户对统计数据质量的综合满意度,结果显示无人认为“非常满意”,认为“比较满意”的有 36 人,认为“一般”的有 151 人,“比较不满意”的有 91 人,“非常不满意”的有 25 人,比重分别为 11.9%、49.8%、30% 和 8.3%。

作为政府统计数据质量的构成要素,各质量维度的用户满意度应与用户综合满意度保持较高的相关性。表 9 的方差分析结果表明,不同综合满意度的 4 组用户,对于各质量维度的满意水平都在 1% 水平下具有显著差异;且综合满意度越高的用户,对各质量维度的满意度也越高。进一步证明本研究提出的 8 项质量维度可以充分反映政府统计数据质量的不同侧面,可以有效解构用户满意度的特征,从中识别重要的用户需求。

五、研究结论及展望

(一) 研究结论与建议

本研究针对政府统计数据质量的评估问题,综

合若干国家统计局部门与国际组织采用的统计质量概念体系,确立了 8 项政府统计数据质量维度;进而参照国际货币基金组织的“数据质量评估框架”,设计了具体的用户满意度测评量表。基于确证性因子分析的信度与效度检验表明,该测评量表具有良好的内部一致性,且可以有效测量政府统计数据质量的用户满意度。进一步的对比研究也表明,本研究提出的 8 项质量维度可以充分反映政府统计数据质量的不同侧面,可以有效解构用户满意度的特征,从中识别重要的用户需求。

本文的相关研究结论对于政府统计部门有效实施统计数据质量管理具有较高的参考价值。首先,当前改进统计数据质量的重点在于保障统计数据的可得性与一致性。在本研究所提出的 8 项质量维度中,用户对于可得性与一致性的重要性认知水平最高,满意度却最低;对这两项维度加以重点改进,可以有效提高政府统计数据的用户感知质量,切实满足用户的关键需求。具体而言,统计数据可得性的改进主要在于加强统计部门与用户交互渠道的建设,实现统计数据信息提供方式的多样化、便利化与制度化,尤其要保障数据提供的公平性;一致性的改进则要求各统计数据编制单位(综合统计部门以及财政部、人民银行、海关总署等职能部门的下设统计机构)之间保持职能与方法协调,加强数据修订的制度化与透明度。

其次,充分协调各质量维度之间的矛盾冲突。

在此以及时性与准确性的关系为例加以说明。统计数据准确性是统计数据质量的传统核心特征,也是最引人关注的话题,甚至可以说仍是当前影响我国政府统计形象的主要症结所在。对准确性的改进需要提高统计生产全过程的客观性与科学性,制度制约与技术难度都较大。而随着对统计数据发布及时性要求的提高,必要的统计核对与修正的时间难免受到挤压,对准确性的保障难度会进一步加大。如利用统计公报渠道发布统计数据,虽可保障数据发布的及时性,其准确性却难以令用户满意;《统计年鉴》等正式的统计出版物在准确性方面质量更佳,但时滞较长。为此,应考虑各类统计数据发布渠道的衔接机制,使各类渠道的优势得以互补,从整体上提高统计数据的综合质量。

再次,以更富多样化的方式保障统计数据的可解释性,提高元数据(metadata)的“可得性”。元数据被定义为“关于数据的数据”,具体是指关于统计数据调查与估算方法、统计概念与范围、统计修正规则等方面的信息。只有充分掌握了元数据,用户才能正确地使用统计数据。鉴于不同专业技术职务、不同数据依赖程度以及利用“统计公报与专题调查报告”渠道与否的各类用户,对于可解释性维度的满意度都具有显著差异,根据用户能力与需求的差异而提供有针对性的元数据说明,提高元数据的可得性,已经凸显其必要性。如独立于特定的数据提供渠道而提供一般化的统计方法与制度宣传信息,强化统计知识宣传普及,提高统计生产过程的透明度,有助于在日常培养并提高用户正确使用统计数据的能力,避免误读误用,也有助于提高对特定渠道的用户满意度。

最后,鼓励民间统计的发展,实现统计数据信息提供主体的多元化。在当前由政府统计部门“垄断”统计数据的生产与发布业务的格局下,统计部门难免缺乏用户(公众)服务意识,缺乏动力来切实改善统计数据质量与用户满意度;同时也极易遭受来自所属政府部门的人为干扰,导致统计生产缺乏客观公正性。而统计数据质量不佳则必然导致统计数据信息的利用效率低下与统计资源的浪费。因此,发展民间统计来弥补官方统计不足已极具现实必要性。通过引入民间统计,发展在统计数据信息提供方面的市场竞争,在一定程度上可以化解当前困扰政府统计部门客观性、中立性的难题,推进制度

建设,同时也有助于促进政府统计部门树立公共服务意识,在市场竞争中推动政府统计数据质量的全面提升。

(二) 研究局限与展望

本文针对政府统计数据质量的用户满意度测评加以研究,走在了公共服务用户满意度研究的前沿,对于政府统计数据用户满意度研究、统计数据质量评估研究、甚至是一般性公共服务的公众满意度研究都有较高的参考借鉴价值。

本研究存在以下局限:一是受访者群体仅限于高校学者,统计数据用户的涵盖范围不全面,研究结论的代表性不足;二是测评量表的设计还不够成熟,调查过程中受访者反馈认为测评项目过于复杂,而区分效度未完全达标,表明本量表有待于进一步的提炼完善;三是对于各质量维度之间的关系(如彼此之间的影响关系、对于用户综合满意度的影响程度等)未加以深入研究。在今后进一步的研究中,应扩展测评对象的范围,在更广泛基础上提炼测评项目,系统研究质量维度的结构,使之更具实际指导价值。通过定期、制度化地实施用户满意度测评,可以编制统计数据用户满意度指数,提供政府统计数据质量改善的直观信息;还可以与其他部门的实践相衔接,使之成为公共服务公众满意度指数的一部分,进而成为我国国家级顾客满意度指数体系中的一个有机部分。

参考文献

- [1] 王华,金勇进. 统计数据准确性评估:方法分类与适用性分析[J]. 统计研究,2009(1): 32-39.
- [2] 徐芳东. 外国统计数据质量的涵义、管理以及对我国的启示[J]. 统计研究,2002(2): 26-29.
- [3] Australian Bureau of Statistics (ABS). *ABS Data Quality Framework*. Available at <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/1520.0Main%20Features1May+2009>. May, 2009.
- [4] Bollen, Kenneth A. *Structural Equations with Latent Variables*. New York: John Wiley & Sons, 1989.
- [5] Cardozo, R. M. "An Experimental Study of Consumer Effort, Expectation and Satisfaction." *Journal of Marketing Research*, 1965, 2(3): 244-249.
- [6] Cronbach, L. J. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika*, 1951, 16(3): 297-334.
- [7] Ehling, M., and Körner, T. (2007). *Handbook on Data Quality Assessment Methods and Tools*. Available at http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/quality_reporting.

将外国商业存在引入国际贸易统计^{*}

——针对中国对外货物贸易统计的研究与数据重估

高敏雪 许晓娟

内容提要:货物进出口贸易数据是衡量中国经济的重要指标,在全球化背景下,跨国公司在世界范围内布置生产链,通过外国商业存在拓展了传统意义上的跨境贸易,由此突显出传统跨境贸易统计的不适应性。本文分析了直接投资所引起的附属机构的国别属性问题;搭建了从跨境贸易统计到属权贸易统计的基本框架以及调整方法;最后集中针对中国货物进出口贸易进行了分步、分口径的调整测算,并系统发掘了调整过程所包含的信息,为评价中国货物进出口贸易提供了不同的图景。

关键词:货物进出口贸易;跨境贸易;外国商业存在

中图分类号:C812 文献标识码:A 文章编号:1002-4565(2010)07-0018-07

Introducing Foreign Commercial Presence into Statistical System of International Trade

——Study and Reevaluation for China

Gao Minxue & Xu Xiaojuan

Abstract: The data of export and import of goods are important indicators to measure the Chinese economy. Multinational enterprises arrange the chain of production all over the world, so the commercial presences broaden the coverage of international trade which makes the statistical system of international trade does not reflect the real situation. The paper analyzes the ascription problem of foreign affiliates produced by foreign direct investment, constructs the basic framework of the statistical system of international trade which based on ownership and the method to adjust the statistical system of international trade which based on residency to the new system, and finally re-evaluate the export and import of goods for China by different coverage of foreign affiliates step by step.

Key words: Export and import; International trade in goods; Cross border trade; Foreign commercial presence

* 本文受国家自然科学基金项目“全球化背景下中国对外经济统计计量与分析”(项目编号 70773114)项目资助。

- [8]Eurostat. *ESS Handbook for Quality Reports*. Luxembourg: Office for Official Publications of the European Communities, 2009.
- [9]Fornell, Claes and Larcker, D. F. “Evaluating Structural Equation Models with Unobservable Variables and Measurement Error.” *Journal of Marketing Research*, 1981, 18(1): 39-50.
- [10] IMF. *Data Quality Assessment Framework*. Available at <http://dsbb.imf.org/Applications/web/dqrs/>. 2003.
- [11] Nunnally, J. C. *Psychometric Theory*, 2nd ed. New York: McGraw-Hill, 1978.
- [12] OECD. *Quality Framework and Guidelines for OECD Statistical Activities*. STD/QFS(2003)1. Available at <http://www.oecd.org/dataoecd/26/42/21688835.pdf>. 2003.
- [13] Oliver, R. L. *Satisfaction: A Behavioral Perspective on the Consumer*. New York: Irwin-McGraw-Hill, 1997.
- [14] Rosen, B. and Elvers, E. “Quality Concept for Official Statistics.” in S. Kotz, C. B. Read, and D. L. Banks (eds.), *Encyclopedia of Statistical Science*, Update Vol. 3. New York: John Wiley & Sons, 1999, 621-629.

一、引言

关于中国对外货物贸易,有两组统计数据值得关注。第一组是货物进出口及其差额,表现为对外货物贸易一直呈现大幅度增长,并且出口远远高于进口,并有较大数额的贸易顺差;第二组数据则显示,外商投资企业在中国对外贸易中占有越来越重要的地位,通过外商投资企业实现的货物进出口已经超过全部货物进出口总额的 50%,而这部分企业的进出口也是呈现为顺差的。

在一般层面上,常常只关注前一组数据,由此得出的结论是:到 2007 年,中国已经是世界第二大货物贸易总额国,同时是世界上最大的货物贸易顺差国。但如果把上述两组数据联系起来,实际情况则是:尽管中国在国际货物贸易市场上已经占有比较重要的地位,其贸易不平衡已经构成国际贸易市场失衡的重要力量,但来自中国的影响在很大程度上是由中国境内的外商投资企业推动的。进一步看,这些外资很可能恰恰来自在国际贸易中呈现为逆差的那些国家,与中国对外贸易顺差以及这些国家对外贸易逆差并行存在的,是这些国家对中国的净投资;他们通过在中国投资实现了在中国当地以及对第三国的贸易,并由此获得了显著的投资收益。

以上连环套情景的出现应该是由国际投资与贸易一体化进程所致。在此进程中,母国通过外国直接投资(FDI)在东道国设立企业然后从事市场经营,使国际间投资越来越深刻地渗透到国际贸易之中,导致国际贸易中呈现出你中有我、我中有你的格局。之所以要将两组数据结合起来使用,是因为传统国际贸易统计并没有考虑上述国际投资对国际贸易所产生的影响,无法将国际间因直接投资所形成的外国商业存在放在一国国际贸易统计框架之中予以体现。对于如中国这样持续引进外资、存在大量外商投资企业、外资企业具有显著进出口倾向的国家而言,单纯依赖传统贸易统计,已经难以提供有效

的数据信息阐释一国对外贸易的真实状况,同时也难以从本质上刻画出国际贸易的全球格局。这就向国际贸易统计提出一个非常现实的挑战:如何改进现有统计概念和方法,针对国际投资贸易一体化现象和进程,考虑外国直接投资企业作为现实的商业存在,给出更有现实意义的国际贸易统计结果。

本文的目标即在于此:立足中国,针对外国直接投资(FDI)对国际贸易的影响,结合国际贸易统计改进的国际进展和经验进行讨论。我们的基本思路是,首先立足传统国际贸易统计规则及其所面临的困境,讨论外国商业存在的国别属性及其对于国际贸易统计的意义;进而讨论将外国商业存在包含到国际贸易统计之中的途径和估算方法;最后展示针对中国国际货物贸易状况进行尝试的估算结果,对不同假定下的估算结果做简要分析。

二、对国际贸易统计的冲击

传统意义上的国际货物贸易主要是指跨境贸易:生产者与消费者分属不同国家,通过贸易实现货物的跨国流动。从一国角度看,其国际贸易是指发生在本国居民与非居民之间的交易。如果一个经济实体在该经济体之经济领土(指该经济体政府所管辖的地域范围及其领土飞地)上具有一定场所(如住房、厂房或其他建筑物)、从事一定规模的经济活动并超过一定的时期,即可确认其为该经济体的居民或常住者(resident),否则就是非居民/非常住者(non-resident)而被视为国外。国际贸易统计的贸易额统计数据就是据此定义的^[1],直到现在,国际收支平衡表^[2]上记录的货物交易也是基于此定义的,这些定义与国民经济核算^[3]的基本原则也是一致的。

在很长时期内,上述定义大体符合国际贸易发生的实际情况,因此国际贸易统计所记录的货物进出口是能够较为准确全面地度量国际贸易规模的。但是,伴随资本在国际间流动壁垒的逐步放松,外国

[15] Statistics Canada. *Statistics Canada Quality Guidelines*, 4th ed. Catalogue no. 12-539-XIE. Ottawa, 2003.

[16] Statistics Korea. (2009). *A Guide to Statistics Korea*. Available at <http://kostat.go.kr/eng2009/engbook/KoStatGuide09/EBook.htm>.

[17] Steenkamp, J. E. M. and van Trijp, H. C. M. "The Use of LISREL in Validating Marketing Constructs." *International Journal of Research in Marketing*, 1991, 8(4): 283-299.

作者简介

王华,男,1978年生,山西翼城人,经济学博士,现为厦门大学台湾研究院助理教授、硕士生导师,研究方向为统计数据质量评估、区域经济统计分析。

金勇进,男,1953年生,北京市人,现为中国人民大学应用统计科学研究中心主任,中国人民大学统计学院教授、博士生导师,研究方向为抽样调查、应用统计。

(责任编辑:赵曾琪)