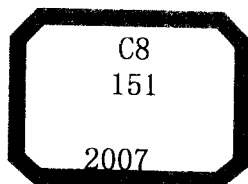


次级统计

孙宪华 编著

数据的应用



次级统计数据的应用

孙宪华 编著



图书在版编目(CIP)数据

次级统计数据的应用/孙宪华编著. —天津:天津大学出版社, 2007.3

ISBN 978-7-5618-2407-8

I. 次... II. 孙... III. 统计数据 - 研究 IV. C8

中国版本图书馆 CIP 数据核字(2007)第 028699 号

出版发行 天津大学出版社
出 版 人 杨欢
地 址 天津市卫津路 92 号天津大学内(邮编:300072)
电 话 发行部:022-27403647 邮购部:022-27402742
网 址 www.tjup.com
短信网址 发送“天大”至 916088
印 刷 廊坊市长虹印刷有限公司
经 销 全国各地新华书店
开 本 169mm × 239mm
印 张 19.5
字 数 361 千
版 次 2007 年 3 月第 1 版
印 次 2007 年 3 月第 1 次
印 数 1 - 3 500
定 价 30.00 元

前 言

到哪里去找用于分析的统计数据,是经济和社会问题研究中的基础问题之一。有时人们并不是为分析的理论框架构建发愁,而是苦于“无数可算”。经济管理部门的官员们经常抱怨“有用的数据找不到,无用的数到处是”,这表明他们虽然能够见到泛泛的统计数据,却不能从中发现出对分析有益处的统计数据来。仔细一想,岂止他们,就是我们这些所谓圈中的统计人,也往往陷入没有数据可用的尴尬境地之中。

数据的使用是如此状况,那么数据的搜集呢?从官方统计来看,数据的搜集或曰统计调查也面临着诸多的困难,如调查方案的设计、审批和备案的繁杂,调查方法组合应用的力度不强,数据提供方变化范围加大和变化频率加快,数据使用方快速增长的需求压力,等等。官方统计作为数据供方和数据需方联系纽带的数据整合商,也存在着硬件上投入相对不足和软件上统计理念相对滞后以及统计理论素养欠缺的局限。由此,官方统计还未能完全解决好被调查者提供数据的积极性下降和数据使用方数据需求急剧上升的矛盾,也没能从根本上将统计调查方法形成一个体系。进一步讲,还基本谈不上将调查体系和统计分类体系、统计指标体系、SNA 及相关的核算标准体系、GDDS 数据公布体系有机地整合为一个大的系统。产生上述问题的原因很多,其中很重要的一个影响因素就是忽视了次级统计数据中蕴藏的巨大资源。不论是数据的使用者,还是数据的制造商(官方统计的别称)都更多地将目光聚焦在原始数据的搜集和利用上,以至于对现实既存的、巨大规模的次级统计数据信息流视而不见,听而不闻。

原始数据指的是统计机构,或者统计人员按照定义的统计问题为具体的目的而专门搜集得到的数据。所谓次级统计数据也叫间接统计数据或二手统计数据,是相对于原始数据而言的数据来源。次级统计数据是指由数据使用者之外的人员和机构所搜集的数据,搜集次级统计数据的主要目的是用于重新界定和设计的研究题目。传统统计调查方法主要侧重于对原始数据的搜集,然而在统计工作中可动用的人财物力资源有限,在社会对统计需求上升和被调查者负担日益加重的背景下,对原始数据的过多搜集已经受到统计资源的制约以及被调查者对负担的抱怨和抵触。因此,在普查、抽样调查和统计报表制度之外,另辟蹊径开发新的更为经济有用的统计调查方式是十分必要的,其中的重要一环就

是搜集和利用次级统计数据的资源。搜集、整合与分析次级统计数据的方法和过程,在国际上被称为次级统计。

研究可能得到的次级统计数据是搜集原始资料的先决条件,一般应从次级统计数据开始分析。只有次级统计数据的来源已经全用完了,或者有了一定的剩余以后,才能考虑进行调研,搜集原始资料。发达国家官方统计机构带头在教学和科研中实施“次级统计数据工程”取得了显著的效果。首先,是教学科研中通过对经济和社会数据有效整合与分析,使得学生对专业知识点和统计数据量化的结合有了进一步的认知,从而也就提高了民众的统计素养;其次,统计学专业教学中通过真实次级统计数据的使用,使得对统计方法和程序的理解更为形象化。在次级统计数据使用中,不同统计方法及结论的反差也会促进对统计方法的进一步创新,同时众多研究者和研究机构对次级统计数据的广泛使用,搭建了数据平台基础之上的研究平台。这个研究平台有助于研究观点的碰撞,研究方法的借鉴以及研究结论的比较和整合。

搜集和利用次级统计数据的意义和思路可以从以下几个方面来说明。

(1)首先是经济财力上的考虑,因为原始数据搜集的成本是异常高的。以2000年第五次全国人口普查为例,仅中央政府投入的经费就达到6.5亿元人民币之多,虽然按13亿的人口大国规模来讲,人均不过区区的0.5元,但是如此大的总量支出,对于并不富足的中国确实有一个负担过重的问题。此次人口普查若加上地方政府的投入以及按惯例免费使用的近百万人口普查调查员的工资补助等开支和电脑设备折旧的成本,则实际总支出会比6.5亿元成倍地增加。发达国家的普查以及其他统计调查项目基本上要求议会三读表决通过后才能实施,其中最主要的考量标准是可动用的经费的多少,并以此来决定上普查,还是上抽样调查,是上大样本还是上小样本抽样调查,甚至会责成官方统计以警察局、税务局、教育局等各种官方政府机构的次级统计数据的方式整合成人口普查数据,此举极大地降低了普查的总负担,且收到和普查基本上类似的效果。国务院已经对2004年年末时点和2004年全年时期的数据特征状况,在全国开展了第一次(第一产业以外的)经济普查,对此应该加强数据搜集过程中次级统计数据的作用。至少可将经济普查数据和相应的次级统计数据进行对比分析,为判断经济普查数据的质量状况提供一个参照系。

(2)科学的永续发展是在不断的发现创新和缜密的质疑当中得以前进的,次级统计数据的使用恰恰使得科学发现得以深化。对数据的分析和使用,无论是数据搜集目的的确定、数据整合结构的建立、数据分析方法的使用以及诠释数据的视野角度,都会受到时代特征的影响和认识是不断深化的规律制约。随着统

计手段中计算机能力的快速提高和统计新方法的层出不穷,对次级统计数据进行较深层面的重新整合会深化对原有问题的质疑和对新问题的发现。这方面既有伟人,在近百年前的列宁利用沙皇俄国官方统计公布的农民数据进行重新分组后,对农民分层的再认知和对农民阶层的新特征展现;又有近代探索性数据分析方法中众多学者提出的不同分类界限标准,即针对同一批数据的不同分类传达了不同的信息,有些真实的反映了信息,有些则歪曲了信息,有些反映得充分,有些则反映得不充分,因为组距式分组中会存在一定程度的信息量损失,问题的关键在于如何减少此类损失,或者减轻此类损失的危害。为了达到一定程度上完美的效果,对次级统计数据的不断反复使用和探索有助于发现真理。众多的次级统计数据使用者的经验表明,对次级统计数据的重构还会影响到对研究假设的重新定位,从而进一步发现和检验新的理论。所以分析现有的次级统计数据是界定问题的重要步骤,只有充分地分析和利用了次级统计数据以及次级统计数据的边际回报下降之后,才应开始讨论是否值得搜集原始数据的问题。

(3)对次级统计数据的反复多次使用具有数据实验室的意义。通常经济、管理、人文和社会科学中的数据是由统计调查得到的,是由人类行为特征数据所组成,而不是像自然科学中是由对实验数据的观察而来。但是,近年来社会科学研究中,定量分析的强化态势,使得次级统计数据的实验室意义得以展现。严谨的科学理论对社会经济实践有着重要的导向意义,并且能被实践活动所重现。在社会科学对理论验证的实践中,由不确定性背景造成的可能风险损失,要远远大于自然科学实践的同类损失,这使得经济学实验室或数据模拟实验室成为社会科学研究中的新宠。数据实验室中对次级统计数据的反复比较、重组和引申,具有和自然科学实验中的同样功能,都可以验证某种理论的科学性和发现新的行为规律以及理论模式。例如,利用投入产出矩阵 SNA 的数据模型可以模拟出价格政策、产业政策、进出口政策和消费政策等各种外部因素对国民经济产出总量的影响状况和转折边界。不断地模拟是在理论的新发现和方法的新改进基础上进行的,从而次级统计数据的价值也就被不断地在实验室中发现出来,并最终会在实践中放出光彩。微观层面的机构单位出于各自需要,在行业和国民经济背景的量化分析中使用次级统计数据,不仅具有次级统计数据反复使用的特征,同时对于有些属于政府统计部门独有权限提供的数据,其他单位根本无法,也没有权力去搜集原始数据。

(4)次级统计数据的使用使得数据的动态比较成为可能。数据的动态分析是一个延续不断的过程,是在环比和定基比较中实现的,有时也存在着动态对比中基期的转换和跨基期比较的问题。在分析数据的趋势分量、季节分量和周期

分量时,都需要时序数列的期间是足够长的,特别是周期分量的识别需要至少20~30年以上的历史数据才能看清其大致模式。时过境迁,远离现实的历史数据已经不再可能“从头再来一次”式的原始搜集,而只能从次级统计数据的利用中使动态比较成为可能。虽然理论上说,对历史数据可以进行定义和口径上的调整,如同中国的国民经济核算实施SNA之后,对原先MPS体系下的数据进行调整那样。但是一般说来,调整后的数据仍然属于次级统计数据,是更深度的次级统计数据。对这种调整后的次级统计数据使用,从数据利用的整体过程来看,属于次级统计数据的多次重复使用。由于历史数据具有的次级统计数据的特征,使得在其和现实数据的比较中,可比性问题的的重要性进一步凸现。一方面,历史的发展和社会的进步体现在不同时间的数据上,不仅数据的总量有了很大的变化,数据内部的内容结构也会产生较大程度上的重组,因此对历史次级统计数据涉及的概念、口径和范围以及进一步数据得以产生和存在的理论框架的了解就显得分外的重要;另一方面,对历史次级统计数据使用中的可比性问题不要过分地苛求,因为这些数据毕竟是对已经发生的历史事件的记录。现实数据和历史数据之间的动态比较,需要权衡好可比性和现实性之间的利弊,不可偏废其中的任何一方,应找到恰当的平衡点。动态次级统计数据这笔昂贵的数据化文献财富的安全保护问题,也是有效利用次级统计数据的前提条件。特别是随着次级统计数据资源规模的日益庞杂和结构的过分复杂,使得以电子计算机和网络媒体为手段的数据库存储技术得以快速发展,由此也就产生了诸如数据丢失、染毒、磁化等问题,这对次级统计数据的开发和利用是致命的。从量化的角度看,对历史事件的反思,体现在对次级统计数据开发利用上的多角度、深层次地探究历史次级统计数据中体现的事物规律性,不仅能够加深对相关主题的动态演变过程的重新审视,而且从改进和审视统计方法应用的角度来看也是十分必要的。

(5)国际间的数据比较和海量数据的统计分析是现代经济和社会科学研究的重要一环,意义特别重大,它适应了经济全球化的潮流和信息社会的时代特征。在与此相关的研究中,次级统计数据的资源就如同对外经济的“舶来品”特征一样,也要在很大的程度上依靠贸易国或有关的国际组织来提供。因此,也就要求这些次级统计数据在国与国之间应该有基本一致的定义、核算原则、统计口径和范围。国家之间统计数据的相互依赖应是官方统计中应用次级统计数据最多的领域之一,并由此导致了经济社会统计中的广泛国际合作。在开放经济条件下,国际之间经济往来的规模和范围日益拓展,交易方式和交易的结构形式也渐趋复杂,与此相适应的国际间的统计协作也越来越显得重要。在SNA体系的

国际交往核算中,官方统计通过两种方式直接地搜集相应国家的原始数据:一是在生产账户、分配账户、资本账户、金融账户等专项账户中通过国外部门和国外项目的设置,诸如用进出口、国外要素收入、国外转移收入等指标来反映国家之间的经济关联和相互经济地位的比较;二是通过系统的国外账户或国际收支平衡表来全面反映一国的对外经济关系全貌。在其他诸如海关、国际金融、外经贸部门的业务核算中也充斥着大量的涉外原始统计数据。然而,国际之间的比较,特别是国际竞争力和国家综合实力的比较,不仅仅体现在涉外经济上,在内部的生产总量、生产结构、人口特征和质量、分配关系、消费投资相互关系以及金融机构的作用等经济社会生活的方方面面也都存在着国际比较的需要。很显然,一国没有能力也没有必要去搜集其他国家的所有原始数据,而在其中涉及的对外经济关系中的原始数据也仅仅限于和本国有关的有限信息,从而使得国家之间次级统计数据的系统和相互利用,在国际组织的协调下得到了空前的发展。主要体现在国际间次级统计数据的一致性和较高的数据质量要求方面,如 SNA、GDDS、财政统计指南和金融统计手册等国际经济统计标准。实际上国际之间次级统计数据的间接利用,伴随着国际经济统计体系的完善和发展,得到了不断进步。以美国和加拿大之间为例,不仅有范围广泛的非外经数据之间的交流和相互利用,就是有些直接贸易的统计数据也不必再要求美加双方各自单独作战或各自为政自搞一套,而是在很多指标上只指定某一方进行原始数据的统计,另一方则直接使用间接的次级统计数据即可。这样就进一步节省了相当紧张的统计资源。

(6)教学和科研中使用大量的次级统计数据不仅促进了相应学科的知识发现,而且在统计方法的意义上得到了提高和完善。以往存在的官方统计数据档案化和统计科研教学数据虚拟化的现象还没有从根本上得到改善,还没有成功地为教学科研搭建一个真正的现实的数据平台。教学研究人员到统计局去索取数据也有畏惧的感觉,尽管有些数据并非是在保密之中,也不是官方统计机构的公务人员有意为难对方。问题的核心是,官方统计数据作为社会公民,特别是教学研究人员最大次级统计数据信息来源的重要价值还没有得到应有的重视。这方面发达国家的政府,及他们的附属机构都通过网络整合出大量的次级统计数据供研究者免费使用,同时具有在线对话的功能,界面也体现了用户友好的设计理念。

在教学和科研实践中,深感统计数据“本身”问题的重要性在于它是数据整合和数据诠释的基本前提。而国内这方面的文献相当匮乏,也不系统。对我们而言,次级统计数据利用问题的研究是一个全新的探索性课题,从框架的设计、

内容的安排,甚至遣词造句和行文风格方面肯定存在很多的瑕疵,敬请读者提出宝贵的意见。天津市社会科学院杨玉秀,天津财经大学龚风乾、孙炜、袁铭,天津工业大学曲岩,天津市统计局杨玉存、潘向友参加了相关研究,国家统计局宋跃征提出了很多建设性意见。写作过程中还参阅了大量的网络文献。在此一并致谢。确信统计方法论学科的发展,一定程度上,会在以次级统计数据使用为媒介的推动下,进一步凸现统计的外在型服务价值。

孙宪华

2006年8月10日

目 录

第一章 次级统计数据的基本问题	(1)
第一节 官方统计基本原则:不要加重调查对象负担	(1)
一、统计数据可以取自各种渠道	(1)
二、各种来源的统计数据之间要协调和统一	(2)
第二节 获取信息权利和提供信息义务的平衡	(3)
一、公民获取信息权利及其保障机制	(3)
二、公民提供信息的义务及其履行	(4)
三、对几个相关问题的思考	(4)
四、世界银行信息公开政策	(6)
第三节 联合国对统计数据搜集的国际协调	(14)
一、联合国统计机构消除统计数据重复搜集的努力	(14)
二、联合国数据重复搜集和质量程序	(15)
三、联合国统计机构的次级统计数据服务	(17)
第四节 GDDS 对转发统计数据的有关规定	(22)
一、GDDS 的数据类别和指标	(22)
二、GDDS 评估统计数据质量的规则	(23)
第五节 为专项研究提供的次级统计数据服务	(24)
一、信息服务公司提供的次级统计数据服务	(24)
二、经济学家提供的次级统计数据服务	(25)
三、次级统计数据的双向意义	(26)
第二章 次级统计数据的基本来源和方法程序	(28)
第一节 次级统计数据分析法	(28)
一、使用次级统计数据的意义	(28)
二、次级统计数据的优缺点	(32)
三、次级统计数据的基本来源	(34)
四、出版物中次级统计数据资源的搜集方法和程序	(36)
第二节 统计名录库是有效利用次级统计数据的前提	(37)
一、全面完善的基本单位名录库的重要性	(37)
二、欧洲国家对名录库等行政数据用于统计目的的质量管理	(39)

第三节 地理信息系统次级统计数据的获取	(45)
一、地理信息系统的含义	(45)
二、美国企业 GIS 案例分析	(47)
三、区域商业和经济角度的地理信息系统	(48)
第四节 OECD 提供的次级统计数据	(51)
一、OECD 出版物	(51)
二、OECD 科技统计与科技指标的次级统计数据	(53)
第五节 网络信息的整合	(54)
一、世界各国政府统计局互联网网址	(54)
二、网络信息集成的几个案例	(60)
第三章 企业次级统计数据应用分析	(67)
第一节 企业数据的来源	(67)
一、企业数据的来源途径	(67)
二、企业数据的来源形式	(78)
第二节 不同来源的企业数据之间的比较	(83)
一、不同来源途径的企业数据之间的比较	(84)
二、不同来源形式的企业数据之间的比较	(93)
第三节 企业次级统计数据检索加工和整理	(99)
一、企业次级统计数据检索加工和整理的意义	(99)
二、次级统计数据检索和检索工具	(100)
三、企业次级统计数据信息检索加工和整理方法	(101)
四、企业进行次级统计数据资料分类的特点	(102)
第四节 企业数据挖掘和 GIS 中的次级统计数据	(103)
一、数据挖掘的含义	(103)
二、数据挖掘与传统的数据搜集方法的区别	(103)
三、数据爆炸和知识贫乏	(104)
四、支持数据挖掘的技术基础	(104)
第五节 案例分析:天津医药制造业竞争力研究的次级统计数据分析	(105)
一、企业次级统计数据案例分析的简要说明	(105)
二、需要搜集的指标数据	(105)
三、如何获得次级统计数据	(106)
第四章 次级统计数据在经济统计中的应用	(112)
第一节 次级统计数据在经济学中的应用	(112)

一、官方统计网站次级统计数据在经济学研究中的应用	(112)
二、生产函数中的次级统计数据	(114)
三、关于计算 GDP 的次级统计数据来源	(115)
四、GDP 数据的补充、调整和修订	(116)
第二节 国际机构中的主要次级统计数据	(117)
一、国际劳工组织统计局对次级统计数据的有效利用	(117)
二、劳动力市场主要次级统计数据的界定、来源和局限性	(122)
三、“透明国际”关于腐败的次级统计数据	(152)
第三节 经济统计数据体系中的次级统计数据	(155)
一、中国统计数据库	(155)
二、AREMOS 经济统计数据库提供的次级统计数据	(157)
三、第一时间查询使用海关统计数据	(158)
四、财政数据来源、口径变化以及和其他数据体系的整合	(159)
五、科学推算和次级统计数据	(164)
第四节 网络次级统计数据和服务业统计调查	(165)
一、利用网络搜集服务业次级统计数据	(165)
二、将网络搜集法用于服务业的统计调查的具体方式方法	(166)
第五节 电子政务构建中的统计信息资源库	(168)
一、政务统计信息资源库构建的主要内容	(168)
二、政务统计信息资源库建设中的困难和障碍	(169)
三、加强政务统计信息资源库的管理和对策	(170)
四、房地产交易次级统计数据与统计局的共享	(173)
五、统计局与地税局等部门的网上实时统计信息共享	(174)
六、数字环保统计信息化系统方案促进次级统计数据的应用	(176)
七、中国行政机构次级统计数据来源	(178)
第六节 次级统计数据和发展成果统计测量	(179)
一、对发展成果统计测量予以完善的必要性	(179)
二、统计的需求及其变化	(180)
三、完善国际统计机构系统	(180)
四、对国家统计竞争力的评估	(181)
五、整合次级统计数据的一些新的动议	(182)
第七节 德国官方统计对次级统计数据的应用	(185)
一、统计结果的继续加工	(185)
二、次级统计数据用于统计的限制及解决方法	(188)

三、统计数据信息的传播方便了公众对次级统计数据的使用	(190)
四、公民友好性和对统计报告人的爱护	(190)
五、利用管理资料作为次级统计数据具有法律依据	(192)
六、几个典型的利用次级统计数据的统计领域	(193)
第五章 次级统计数据在 SNA 中的应用	(197)
第一节 国民账户体系的概念和历史发展	(197)
第二节 国民经济核算体系的框架结构	(198)
第三节 SNA 的数据来源和数据整合的方法	(201)
一、涉及原始数据的投入产出调查	(201)
二、编制投入产出表阶段需要大量的部门次级统计数据支持	(202)
三、部分涉及原始数据的 SNA 调查要结合使用次级统计数据	(204)
四、使用统计部门次级统计数据成果的 SNA 调查	(204)
五、SNA 统计调查中的次级统计数据特点	(205)
第四节 GDP 核算中对次级统计数据的利用及其与原始数据的整合	(208)
一、生产法	(208)
二、收入法	(209)
三、支出法	(210)
第五节 资金流量实物核算表对原始和次级数据的结合应用	(212)
一、按照交易项目和机构部门的类别搜集原始数据	(212)
二、资料的整合衔接	(212)
第六节 资金流量金融表中对原始和次级数据的结合应用	(213)
一、资金流量金融交易表的数据来源	(213)
二、金融交易表的编制方法取决于数据的来源情况	(215)
三、金融交易表核算的汇总、合并和轧差原则	(215)
第七节 国际收支平衡表中对原始和次级数据的结合应用	(216)
一、编制国际收支平衡表的一般方法	(216)
二、我国国际收支平衡表的编制方法	(216)
三、国家外汇管理局编制核心指标对次级统计数据的说明	(217)
第八节 资产负债表中对原始和次级数据的结合应用	(219)
第六章 市场营销调研中的次级统计数据	(221)
第一节 市场营销中的次级数据的分类和来源	(221)
一、市场营销中的次级统计数据的分类	(221)
二、市场营销中的次级统计数据的来源	(221)

三、对市场营销中次级统计数据来源的评价	(228)
第二节 电子网络技术在市场营销调研中的应用	(231)
一、询问法在网上调查中的应用	(231)
二、实验调研法在网上调查中的应用	(233)
三、专家调研法在网上调查中的应用	(233)
第三节 次级统计数据的营销用途和应用案例	(234)
一、需求估计和监督经营环境变化	(234)
二、市场细分和选择目标市场	(235)
三、中国台湾市场调查建立文档次级统计数据信息系统	(236)
四、美国佛罗里达州使用次级统计数据评估需求	(238)
第四节 次级统计数据在国际营销中的应用	(240)
一、两个基本问题	(240)
二、国际营销中与次级统计数据有关问题的争论	(241)
三、开发国际商业情报系统	(245)
第七章 次级统计数据系统的创举:美国的经验介绍	(246)
第一节 构建美国国际贸易数据系统	(246)
一、美国国际贸易数据系统的提出	(246)
二、国际贸易的转型促使政府统计系统变革	(248)
第二节 国际贸易数据系统的发展过程	(253)
一、为 ITDS 作为一个跨机构的项目提供组织保证	(253)
二、ITDS 系统的试运行	(254)
三、机构间的整合是实施 ITDS 的关键措施	(255)
四、ITDS 的全面推进和完善	(256)
第三节 美国交通和出行调查数据库	(259)
一、交通次级统计数据是研究和决策的基础	(259)
二、使用次级统计数据要注意抽样方法和样本量的变化	(260)
三、调查项目的变化及其对次级统计数据的利用和整合	(261)
四、普查结果作为次级统计数据在交通和出行调查中使用	(262)
五、作为次级统计数据的交通数据库对我国的借鉴意义	(263)
第八章 次级统计数据存储管理与整合展示	(264)
第一节 次级统计数据存储与管理	(264)
一、使用传统数据库存储次级统计数据的局限性	(264)
二、数据仓库及相关概念	(266)
三、数据仓库结构	(268)

四、次级统计数据仓库结构设计	(270)
五、当前数据仓库产品对比分析	(273)
六、统计部门数据仓库应用案例	(277)
第二节 次级统计数据清洗与整合	(280)
一、次级统计数据清洗与整合的必要性与实施困难	(281)
二、次级统计数据清洗与整合技术的传统方法	(283)
三、次级统计数据清洗与整合技术基础:XML 语言	(285)
第三节 次级统计数据检索与展示	(289)
一、专业数据查询和元搜索引擎	(289)
二、查询语言与查询操作	(290)
三、基于 XML 实现数据可视化	(291)
四、数据挖掘支持工具	(292)
参考文献	(295)

第一章 次级统计数据的基本问题

第一节 官方统计基本原则:不要加重调查对象负担

一、统计数据可以取自各种渠道

联合国统计委员会 1994 年接受国际官方统计基本原则时,引入了欧洲经济委员会关于各成员国接受官方统计基本原则的决议。该决议初稿曾被欧洲委员会的统计学家们以及各国政府统计部门的统计人员仔细推敲。当今的世界就像一个地球村,但要实现政治、经济和社会的全球化,首先必须做到信息处理及其系统的全球化。目前,信息全球化已经是一个看得见的事实,强大的通信系统,卫星电视可以在全世界范围内搜集和传播信息。很显然,信息全球化为世界经济的可持续发展提供了机会,然而它同时也给人类带来了威胁。假如,信息质量不高、信息被错误地使用,那么它就像极端危险的武器和毒品,因此在世界范围内,有关信息及其系统的一些共同的法律规章制度必须被进一步完善、接受并遵从。联合国官方统计基本原则就是这些规章制度中的基本内容,它不仅应用于官方统计,而且应用于官方的其他信息系统及研究领域。不要加重调查对象负担是官方统计的基本原则之一。

用于统计目的的数据可以取自各种渠道,它们有可能是统计调查取得的,也有可能是通过行政记录取得的。统计机构应该在综合考虑数据质量、及时性、费用和调查对象负担等因素后,选择合适的统计数据来源渠道。利用行政记录为统计服务是现代统计的发展方向,将行政记录作为数据源也是各国现代信息基础的发展方向。在各种基础信息系统间开展数据交换必须依法限定和控制,同时还要依据各国特定的法律制度,对其配套的注册登记工作进行明确的说明。必须强调将官方统计与行政记录广泛结合起来是为了减轻调查对象负担,从而节约全社会的统计经费。在许多国家,各种公共组织根据社会公众的意愿存储了大量信息,这些信息是整个社会的财富。对于任何信息系统而言,整个工作过程中花费最多的环节是数据的搜集和加工,因此数据的重复搜集与加工应尽可能避免。在一个良好的社会信息系统中,现有的可利用的文件和数据基础不仅可以被专门负责数据采集的机构所使用;还可用于各种类型的经济社会问题的

分析和研究。

二、各种来源的统计数据之间要协调和统一

官方统计实施统计数据的搜集时必须具备一个前提，就是各种基础信息系统之间要达到一定程度的协调和统一。官方统计机构似乎应该承担这一角色，进行不偏不倚的协调。官方统计机构不是为自己搜集数据，其职责是根据各种用户的权利和需要为之提供统计信息。其他的基础信息系统是为其自身的业务需要开展数据的搜集工作，如税收数据的搜集、社会保险数据的搜集、海关数据的搜集，等等。官方统计基本没有政策决策职能，它只是为他人提供数据。官方统计涉及社会、经济、政治生活的各个领域，而其他基础信息部门仅涉及与其业务有关的领域。官方统计覆盖全国范围，甚至还搜集国外数据进行国际比较，而其他信息部门一般只是覆盖国内的一个特定范围。官方统计运用国际通行的统计标准和科学方法，而其他信息部门是基于一国特定的行政法规开展工作。出于行政管理的目的，行政数据更新较为频繁，行政记录的质量控制也非常具体。但是，行政数据或记录的更新和质量控制标准与统计上使用的标准相比差别很大。基于对几个行政信息系统（社会保障系统、税收系统、海关系统和地籍管理系统等）的比较分析，将行政信息系统和统计系统统一协调起来的关键问题是，各系统有无合作的意愿以及能否相互理解彼此的优势，这些优势可以通过统一协调来实现。例如，根据编制抽样框的目的对税收申报内容进行适当调整就很容易，同样对地籍登记内容进行调整也并非难事。在一些国家，只要对社会保障记录的有关内容和分类进行细微的调整，并采用官方统计的通用识别标准，它就会变成丰富的就业、工资和薪金以及劳动条件等方面的统计数据资源。

使用行政记录满足了统计的要求，以最小的成本和最轻的调查负担及时地获取高质量的统计数据。国际统计机构应该发挥其权威作用，不时地向各国政府解释和宣传统计上利用行政记录的重要性和有效性，并要强调统计上利用行政记录并不意味着评估行政记录数据质量，官方统计不是为自身的需要而是为统计用户的需要来开展统计调查的。还应该强调的是，各基础信息系统不仅要在其所属的机构组织内部发挥作用，同时还要为整个国家和社会的信息基础的发展作出贡献，因此它们应该满足整个信息基础的一般性要求并提供信息服务。