

# 如何提高会计调查研究质量： 一个基于样本代表性的理论评析框架

曾爱民, 南星恒

(厦门大学 管理学院, 福建 厦门 361005)

[摘要] 调查研究是我国会计研究中最先使用的实证研究方法, 但如何提高调查研究质量却尚未得到应有的重视。本文从提高样本代表性, 进而提高调查研究质量的视角出发, 提出一个理论评析框架, 并对该框架进行全面、系统的阐释, 为提高和评价调查研究的质量提供客观、科学的理论依据。

[关键词] 会计调查; 质量; 样本代表性; 理论评析框架

[中图分类号] F234.3 [文献标识码] A [文章编号] 1673-5609(2009)04-0094-05

## How to Improve the Quality of Accounting Survey Research A Theoretical Analytical Framework Based on the Representativeness of Sample

ZENG Ai-m in NAN Xing-heng

(School of Management, Xiamen University, Xiamen 361005, China)

**Abstract** Survey research method is the earliest empirical method used in accounting research in China but little importance has been attached to how to improve the research quality. In order to improve the representativeness of the samples and the quality of the survey research, this paper presented a theoretical analytical framework, and made comprehensive and systematic analysis. The framework offered an objective and scientific theoretical basis to improve and assess the quality of survey research.

**Key words** accounting survey; quality; representativeness of sample; theoretical analytical framework

### 一、引言

调查研究 (survey research) 是一种以自填问卷或结构访问的方法, 系统地、直接地从一个取自总体的样本中收集有关数据, 并通过对数据的统计分析来认识社会现象及其规律的社会研究方法。由于问卷、统计分析是调查研究这种研究方法本质的特征, 因此人们也将调查研究称为“问卷调查”或“统计调查”。

调查研究法是我国会计研究中最先使用的实证研究 (positive research) 方法。1988年以前, 我国会计研究基本上是传统的规范研究方法一统天下。1988年, 裘宗舜、王平在《会计改革若干问题——一张有意义的社会问卷调查表》一文中率先使用调查研究法, 开

我国实证会计研究之先河 (刘峰、雷科罗, 1993; 刘玉廷, 2000)。虽然后来由于我国资本市场的逐渐兴起, 经验研究 (empirical research) 后来居上, 蓬勃发展, 占据了我国会计实证研究的绝对主流地位。但由于调查研究便于大范围了解有关社会公众或组织管理者的观点和态度, 获取组织内部会计决策和运行控制等方面的重要信息, 而这些都是资本市场数据库所无法提供的, 因而调查研究法在我国会计研究的某些领域 (比如, 管理会计、会计准则制定、会计教育和会计人才市场需求等) 中一直发挥着不可替代的作用。

但在很多时候, 研究者使用调查研究法的内在冲动并不因其在获取数据方面的特殊地位, 而源于其表面上的简单性——打印一些问题, 复印问卷, 邮寄给被

\* [收稿日期] 2009-03-23

[作者简介] 曾爱民 (1971-), 男, 湖南永州人, 厦门市厦门大学会计系 2007级博士生。

调查者, 等待回函, 然后分析答卷。但是, 任何调查都包括了一些重要步骤, 需要认真考虑并按特定的顺序予以执行。如果不知道这些步骤则容易得出一些低质量的调查结果 (Mangione 1995)。因而, 调查研究法常受到不够严谨、科学和不具有公信力之类的抨击 (徐政旦等, 2002)。早期有关研究亦发现调查研究的核心缺陷是, 许多研究都未遵循调查设计和管理的基本原则 (Mangione 1995)。如果正确设计和管理, 调查研究可以是一种大规模、高质量数据的来源 (Dirmond 2000 Dillman 1978 1999)。因而提高调查研究质量关键在于方法的运用, 而不是方法本身。那么, 如何运用才能尽可能地提高调查研究的质量呢? 接下来, 我们将从提高有效样本代表性的视角出发, 提出一个理论评析框架, 对此展开较为系统的研究。

## 二、理论评析框架的提出

由定义可知, 调查研究从总体中抽取样本, 收集定性或定量的有关数据。因而, 一个代表性样本 (representative sample) 应是总体的一个缩影, 样本各种平均性指标应与总体相应的指标十分接近。这样, 通过研究总体中的部分个体, 得到的却不仅仅是这部分个体的情况, 而是渗透、折射、体现在这一部分个体身上的总体的情况, 从而可以顺利地实现调研目的。否则, 如果样本缺少代表性, 即使花费再多的人力、物力、时间, 调查数以万计的对象个体, 进行高深莫测的数学统计和模型建构, 结果也是徒劳无功。

最经典的例子是, 1936年, 美国著名的《文学摘要》杂志社为了预测总统候选人罗斯福与兰登两人谁能当选。他们按电话簿和俱乐部成员名单上的地址发出 1000 万封调查信, 回收 240 多万封。这在统计史上都是少有的样本容量, 花费了大量的人力、物力和财力。《文学摘要》深信自己的统计结果, 即兰登将以 57% 对 43% 的比例获胜, 并大张旗鼓地进行宣传。但选举结果却是罗斯福以 62% 对 38% 的巨大优势获胜! 这次调查断送了这家原本颇有名气的杂志社的前程, 不久便关门停刊。是什么原因造成这次调查研究的失败呢? 后来统计学家分析总结的原因主要是, 调查的抽样方法不公平, 样本不具有代表性。因为样本不是从总体 (全体美国公民) 之中随机抽取的。当时有私人电话和能参加俱乐部的都是比较富裕的家庭, 仅从这些家庭中抽取的样本严重偏离了总体。而事实上, 1929—1933 年的世界经济危机使美国经济遭受重创, 时任总统的罗斯福实施了“罗斯福新政”, 动用行政手段干预经济, 经济政策由自由资本主义转为国家垄断资本主义, 损害了部分富人的利益, 所以罗斯福在富人选票中百分率较低在情理之中。但大多数的美国人从中得到了好处, 因此罗斯福赢得大多数选民的支持又是必然结果。从这个历史事件可知, 样本代表性是决定调查研究质量的“生命线”。

为了更清楚地理解, 我们根据调查研究的数据收集过程提出了一个基于样本代表性的理论评析框架, 具体如图 1 所示。

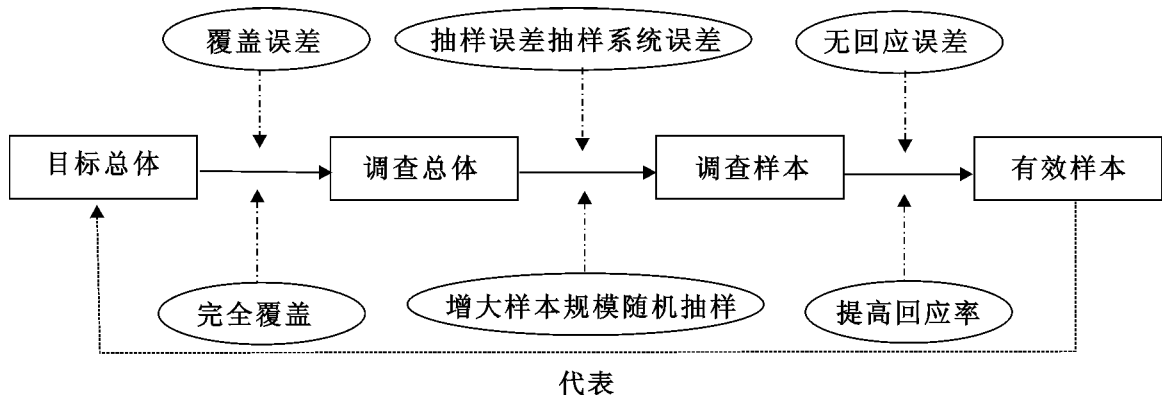


图 1 基于样本代表性的理论评析框架

由该理论评析框架可知, 调研数据的获取是一个根据目标总体获取相应的调查总体 (抽样框), 再从调查总体抽取调查样本, 最后从有效样本获得有关数据的过程。调查研究的目的在于探索、描述或解释存在于目标总体中的现象或规律, 但分析、研究的数据却从有效样本中获得, 因而, 有效样本的代表性直接决定了调研数据的质量, 进而决定了调查研究的质量。当样

本缺乏代表性, 调查结果 (样本数据) 与事实 (总体数据) 出现差异, 便产生了调查误差 (survey error)。调查的数据质量与调查误差成反比关系。调查误差主要包括: 抽样误差 (sampling error) 和非抽样误差 (non-sampling error)。其中非抽样误差可以归为三种类型: 抽样框误差<sup>①</sup> (frame error)、无回应误差 (non-response

① 抽样框误差是非抽样误差的一个重要部分。如果目标总体中的单元与抽样总体中的单元不一致, 就会产生抽样框误差 (冯士雍等, 1998 金勇进, 1996)。

error) 和计量误差<sup>①</sup> (measurement error) (Judith T. Lessler & William D. Kalsbeek, 1992; Dillman, 2000)。所以, 调查质量与调查误差存在以下关系 (郑丽, 2004):

$$\begin{aligned} \text{调查质量} &= \frac{1}{\text{调查误差}} = \frac{1}{\text{抽样误差} + \text{非抽样误差}} \\ &= \frac{1}{\text{抽样误差} + \text{覆盖误差} + \text{无回应误差} + \text{计量误差}} \end{aligned}$$

可见, 抽样误差和非抽样误差只要有一方或者双方同时增大, 就会使得调查质量下降, 因而减少抽样误差和各类非抽样误差成为提高调查质量的关键所在。

### 三、理论评析框架的阐释

我们将按图 1 中理论评析框架的顺序对各类误差逐一展开分析, 以明确提高调查研究质量的理论机制与具体措施。

#### (一) 总体与覆盖误差

总体是调查研究者希望推广其研究结论于其中的全体要素的集合 (Diamond, 2000)。根据其在调查研究中的地位和作用不同, 我们又可从两个层次上对总体作进一步的考察 (如图 1)。一个层次是目标总体, 目标总体 (target population) 是研究者意欲研究的全体回应者 (respondents) 的集合。这一层次的整体只是概念上的定义, 比如, 家族企业的财务总监; 另一层次则为调查总体, 而调查总体 (survey population) 则是研究者所能获得的用于实际抽样的全体回应者集合。在这个层次上, 总体不再是一种概念上的定义, 而应是一份具有可操作性的名单, 包括目标总体中合格的 (eligible)、可确认的 (identifiable) 并能联系上的个体, 比如, 一份浙江省家族企业财务总监的名单。

如果调查总体与目标总体的覆盖范围不一致, 则会产生所谓的覆盖误差 (coverage error)。覆盖误差包括覆盖不足误差 (undercoverage error) 和覆盖过度误差 (overcoverage error) 两种类型。其中覆盖不足误差是指调查总体因为不完全而小于目标总体所产生了偏差。造成偏差的原因, 即调查总体小于目标总体就被称为覆盖不足或丢失目标总体单元 (Lessler & Kalsbeek, 1992)。这样的话调查总体就没有穷尽目标总体, 一些单元从而成为遗漏单元或缺失单元 (missing elements)。这种误差是覆盖误差中最严重的问题, 因为无法从样本或调查总体中发现丢失单元所造成的误差。与此相反, 调查总体包含非目标总体单元则称为覆盖过度 (overcoverage)。这时调查总体不是小于目标总体, 而是包含目标总体所不包含的单元。这种偏

差的影响很大, 但是潜在威胁却通常会小一些。因为可以在调查中辨认出非目标元素并把它们剔除。

我们可以进一步用例子来加深对覆盖误差的理解。如前所述, 所谓覆盖过度是指调查总体大于目标总体, 比如, 在进行以家族企业财务总监为目标总体的调查时, 调查总体包括了国有企业的财务总监; 而覆盖不足则指调查总体小于目标总体, 比如, 没能包括浙江省外其他地区家族企业财务总监, 则研究结论难免产生偏差。虽然目标总体确定为“家族企业财务总监”, 但实际调查仅限于“浙江省家族企业财务总监”, 则调查研究结论自然难以在全国 (其他省的) 财务总监中推广, 可见覆盖误差是一种系统性偏差 (systematic bias), 直接决定了调查总体对目标总体的代表性。为了减少覆盖误差, 应保证目标总体与调查总体之间的一致性 (Diamond, 2000)。

#### (二) 调查样本与两类抽样相关误差

样本是总体的一个子集, 如果样本是总体的一个代表, 则对样本为真的命题在可计算的误差边际内对总体也为真 (Sapsford, 1999)。如图 1 的理论评析框架所示, 样本又可进一步细分为调查样本和有效样本。其中采用一定的抽样方法从调查总体中抽取, 作为调查对象的个体的集合称为调查样本, 比如, 采用问卷调查方式中的全体问卷发放对象; 但由于种种原因, 发放的问卷并非全部能够收回, 即便收回的问卷也未必全部有效, 因而我们真正用于代表总体进行研究分析的仅仅是由有效问卷的应答者所组成的有效样本。

从调查总体中抽取调查样本可能会存在两类误差 (图 1)。其一是抽样误差, 是由于从调查总体中随机抽取调查样本所引起, 因而它是一种随机误差。只要总体中个体差异存在, 则抽样误差不可避免, 但可以事先计算并加以控制, 只要选定了抽样方法, 就能根据相应的公式计算出误差值的大小; 只要在经费许可的范围内, 增大样本规模、改变抽样方法就能对它实现有效的控制。其二是抽样系统误差, 仍是由从调查总体中进行抽样所引起的, 但其主要源于有偏误的抽样方法 (biased sampling methods), 比如采用方便抽样 (convenience sample) 和自愿回应样本 (volunteer sample) 等。由于主观因素破坏了随机原则, 导致样本系统性偏离总体, 因而这种误差是一种系统性误差, 纠正的方法是采用随机抽样 (random sampling)。

#### (三) 有效样本与无回应误差

由图 1 可知, 研究者虽然可以通过科学的抽样设

① 计量误差源于不恰当的问题措词 (question wording) 和问卷结构 (Dillman, 2000), 与有效样本对调查总体的代表性程度无关, 因而在理论评析框架中只包括了抽样误差、覆盖误差、无回应误差三类。

计, 从总体中抽取具有代表性的调查样本, 然而, 调查样本的代表性并不能最终反映调查结果的代表性。因为调查样本对总体的代表性仅是一种调查开始之前的衡量指标。在进行调查的过程中, 会有许多因素导致抽样样本中部分个体的缺失或失效(无回应), 使得最终的有效样本只是调查样本的一部分(图1)。且“无回应者(non-respondents)通常与回应者有着相当的差别<sup>[9]</sup>”(Nachmias & Nachmias, 2000)。比如, 在某一次调查研究中, 调查样本包括100家公司, 其中上市公司80家, 非上市公司20家, 共发出100份问卷, 每家公司1份, 回收60份(全部来源于上市公司, 一种比较极端的情况), 回应率为60%。那么, 这次调查结果能代表调查样本进而代表目标总体吗? 显然不能! 因为无回应者(或回应者)与整个样本之间存在着系统性偏差, 这就是所谓的无回应误差(non-response bias)。很明显, 当无回应误差较大时, 调查数据的质量和调查结果对总体代表性就会受到明显的影响和破坏。

因而调查回应率(response rate)的高低则成为影响有效样本代表性程度另一关键指标。调查回应率是成功完成调查的被调查者数量与实际调查的被调查者数量的比值, 它通常包括自填式问卷调查中的问卷回收率、结构化访问调查和电话调查中的访问回答率。由于低回应率(比如低于20%)的样本更有可能存在回应者的自选择<sup>①</sup>(self-selection)问题, 因而低回应率往往是与高无回应误差相伴相依的。可见, 为降低无回应误差则应提高调查研究的回应率水平<sup>②</sup>。当一项调查研究能从其有代表性的样本获得高回应率时, 则无回应误差一般会相应降低, 但并不是说一定会降低。仍以上例说明, 若回收问卷增至80份(亦全部来源于上市公司, 仅为一种极端性假设, 现实中可能性极低), 则回应率提高到80%, 但其无回应误差却并没因此而降低, 无回应者与回应者仍存在着同样的系统性偏差。

那么为了获得较高代表性的调查数据, 需要多高的回应率才能有效地将无回应误差降低至可接受的水平呢? 对此, Diamond (2000)指出75%—90%之间的回应率通常可以得出可靠的结论, 但研究者应对样本的代表性进行检查(即进行无回应误差分析)。Vander Stede等人(2007)也提出了类似的要求, 认为当回

应率低于80%的时候, 期望避免所有潜在来源的无回应误差是不切实际的(unreasonable)。因而一般情况下, 研究者都需进行无回应误差分析<sup>③</sup>以正确判断有效样本对目标总体的代表性。如果存在明显的无回应误差时, 研究者应该采取相应的补救措施, 比如进一步联系无回应者, 或者采用相应的替代程序等, 以降低无回应误差, 增强有效样本的代表性, 进而提高会计调查研究的质量。

#### 四、研究结论

古扎拉蒂(2004)曾指出“研究结果不能比数据的质量更好”<sup>④</sup>。综上对整个理论评析框架的阐释可知, 欲提高所获取的调研数据质量, 进而提高研究质量, 则尽可能地提高有效样本对目标总体的代表性就自然十分重要了。故而研究者应在保证调查总体对目标总体完全覆盖, 减少或消除覆盖误差的基础上, 选择适当的抽样方法和抽取必要的样本规模以减少抽样误差及与抽样相关的系统误差。同时还需采取一定措施提高回应率以减少无回应误差。

#### [参考文献]

- [1] 刘峰, 雷科罗. 对中国会计理论研究的若干认识[J]. 会计研究, 1993(5): 6-11.
- [2] 刘玉廷. 关于会计研究方法问题[J]. 会计研究, 2000(12): 9-13.
- [3] 徐政旦, 谢荣等. 审计研究前沿[M]. 上海: 上海财经大学出版社, 2002.
- [4] 杨雄胜. 提高我国会计研究质量问题的思考[J]. 会计研究, 1997(11): 1-6.
- [5] 郑丽. 从非抽样误差看调查质量[J]. 统计与决策, 2004(3): 94-95.
- [6] 古扎拉蒂. 计量经济学: 第三版[M]. 林少宫译. 北京: 中国人民大学出版社, 2004.
- [7] Diamond S S. 2000 Reference guide on survey research. In Reference Manual on Scientific Evidence, 2nd Edition. Washington, DC: The Federal Judicial

① 自选择(self-selection)是指个体有选择地加入某些群体的现象。通常更多时候暗含着加入某些群体中的个体的特征具有明显异化的意思。比如, 低回应率的调查回应者往往具有与总体明显不同的某些特征。

② 常用的提高调查回应率的方法主要有: 预先提醒、跟踪提醒、采用混合方式调查以及寻求权威的支持等。

③ 我国许多学者常以20%—40%为回应率的可接受标准, 而且并不进行无回应分析。无回应误差分析常用的方法有三种: 比较早回应者与晚回应者; 比较回应者与目标总体; 比较无回应者与目标总体。

④ 资料处理上一个十分著名的理论“garbage in garbage out (GIGO)”, 指的其实也是相同的道理: 输入(电脑当中)的是垃圾(指低质量的数据), 则输出来的(分析结果)必定也是垃圾。

- Center, 229-276
- [8] Dillman, Don A. 2000. Mail and Internet Surveys: The Tailored Design Method. New York: John Wiley & Sons
- [9] Frankfort-Nachmias C. & D. Nachmias. 1999. Research Methods in the social sciences (5th ed.), N. Y., Worth Publishers and ST. Martin's Press
- [10] Judith T. Lessler and William D. Kaltsbeek. 1992. Non-sampling Error In Surveys, New York: John Wiley and Sons
- [11] Mangione T. W. 1995. Mail Surveys: Improving the Quality. Beverly Hills, CA: Sage Publications
- [12] Sapsford, R. 1999. Survey Research. Thousand Oaks, CA: Sage Publications
- [13] Win A. Van der Stede, S Mark Young and Clara Xiaoling Chen. 2007. Doing Management Accounting Survey Research. In Handbook of Management Accounting Research, Volume 1: 445-478

[责任编辑:薛之]  
[责任校对:兰柯]

(上接第 93 页)但由于电信网、有线电视网和计算机网自身业务的特性和需求不同,融合中还存在着许多待解决的问题。所以,三网融合的发展不能一蹴而就,而是一个长期的、渐进的过程。我国应勇于借鉴国外的先进技术和政策经验,抓住机遇,应对挑战,逐步打破行业垄断,推进我国三网融合进程又快又好的发展。

#### [参考文献]

- [1] Anne Bodzinga SusanWhite. Networking IPTV Services with MS[C], Telecommunications Network Strategy and Planning Symposium, 2006 1-5
- [2] 侯宏. 三网融合背景下电信运营战略思考[J]. 中国新通信, 2008 (12).
- [3] 韦乐平. 三网融合与 3+1 行业架构的设想[J], 广播电视信息, 2008 (3).
- [4] 鲁明月. 三网融合的经济分析[J]. 新闻经济, 2007, (7).
- [5] 王峰. 浅谈“三网融合”[J], 中国传媒科技, 2007 (5).
- [6] 徐贵宝, 张德华. 国内外三网融合业务的发展现状分析[J]. 中国多媒体通信, 2007 (5).
- [7] 李勇, 邓鹏, 王晓东, 周兴铭. “三网”融合的现状与技术发展[J]. 长沙通信职业学院学报, 2007, (3).
- [8] 刘志华. 从三网融合趋势看电信运营商的转型[J]. 甘肃科技, 2006 (11).
- [9] 邹文英. 三网融合态势下电信竞争与管制的新取向[J]. 未来与发展, 2005 (3).

[责任编辑:蒋桐媛]  
[责任校对:薛之]