

北京市属高等学校人才强教深化计划项目 (PHR20090505)

北京市属高等学校创新团队建设与教师职业发展计划项目——价格波动研究创新团队：价格波动研究 (IDHT20130505)

科研基地建设-科技创新平台-大数据背景下农产品流通中食品安全政策研究 PXM2015-014213-000039

DIAOCHA YU FENXI

调查 5 分析

第五辑

李朝鲜 主编

中国商业统计学会市场调查与教学研究分会
北 京 工 商 大 学



经济科学出版社
Economic Science Press

北京市属高等学校人才强教深化计划项目 (PHR20090505)

北京市属高等学校创新团队建设与教师职业发展计划项目——价格波动研究创新
团队: 价格波动研究 (IDHT20130505)

科研基地建设—科技创新平台—大数据背景下农产品流通中食品安全政策研究
PXM2015-014213-000039

调查与分析

第五辑

李朝鲜 主编

中国商业统计学会市场调查与教学研究分会
北 京 工 商 大 学

经济科学出版社

图书在版编目 (CIP) 数据

调查与分析. 第5辑/李朝鲜主编. —北京: 经济科学出版社, 2015. 3

ISBN 978 - 7 - 5141 - 5584 - 6

I. ①调… II. ①李… III. ①市场调查②市场分析
IV. ①F713. 52

中国版本图书馆 CIP 数据核字 (2015) 第 058527 号

责任编辑: 刘怡斐 张建光

责任校对: 徐领柱

责任印制: 邱 天

调查与分析

第五辑

李朝鲜 主编

中国商业统计学会市场调查与教学研究分会

北 京 工 商 大 学

经济科学出版社出版、发行 新华书店经销

社址: 北京市海淀区阜成路甲 28 号 邮编: 100142

总编部电话: 010 - 88191217 发行部电话: 010 - 88191522

网址: [www. esp. com. cn](http://www.esp.com.cn)

电子邮件: [esp@ esp. com. cn](mailto:esp@esp.com.cn)

天猫网店: 经济科学出版社旗舰店

网址: [http: //jjkxbs. tmall. com](http://jjkxbs.tmall.com)

北京万友印刷有限公司印装

787 × 1092 16 开 16 印张 320000 字

2015 年 4 月第 1 版 2015 年 4 月第 1 次印刷

ISBN 978 - 7 - 5141 - 5584 - 6 定价: 66.00 元

(图书出现印装问题, 本社负责调换。电话: 010 - 88191502)

(版权所有 侵权必究 举报电话: 010 - 88191586)

电子邮箱: [dbts@ esp. com. cn](mailto:dbts@esp.com.cn)

编 委 会

编委会主任：李朝鲜

编委会委员：金勇进 秦仲云 刘建平
徐映梅 米子川 蒋剑辉
余华银 方 燕 孙际平
蒋 妍 任晓阳 毛炳寰

目 录

移动互联时代抽样调查的设计问题	亢志伟 侯志强 1
北京城镇低收入居民分享经济增长成果状况及影响因素的调查研究	阮 敬 9
基于 SEM 的大学生诚信现状调查分析	张首芳 成高娟 22
关于居民家庭投资理财现状的调查	
——基于烟台市区居民的调查	马 乐 王相懿等 40
关于西安市民环保意识的调查分析与对策研究	颜卫忠 徐婷玉 51
关爱老人，就是关爱明天的自己	
——基于烟台市老年人幸福度的调研报告	张首芳 张 倩等 58
关联规则挖掘技术及其在互联网行业的应用	吴佳佳 侯志强 71
电子商务服务业影响力指数测算	顾军晓 80
火锅店的选址区位分析	陈正伟 88
关于构建物流企业绩效考核指标体系的思考	赵 旭 102
基于主成分分析的中国各省市经济发展质量研究	张 敏 李 勇 109
人均国内生产总值的时间序列分析及预测	赵田丹 117
我国当前宏观经济效益的实证分析	马 艳 126
北京市第三产业与经济增长的实证分析	郑慧慧 135
河南省农业产业结构调整对农业经济的影响研究	马丽荣 144
浅析金义都市区金融发展与经济增长的关系	罗 平 153
新时期我国商业信用影响因素分析	
——基于银行贷款视角的实证研究	李蓓蓓 杜 伟 160

中小企业板市场推出对我国证券市场资本配置效率影响的

实证研究 刘倚铭 170

黄金价格与消费者信心指数的互动关系研究 李坤格 181

股票价格指数编制的必要性与方法 李 智 189

我国棕榈油期货价格发现功能探索 李婷婷 197

国际大米市场价格波动对国内市场的传递效应 袁 茵 209

中国与美国玉米期货市场价格发现效率对比分析 付京亚 217

加强多元统计分析实验教学的案例应用

——基于市场研究 董寒青 225

立体化统计学实践教学体系的探索与构建 李玉萍 231

创新型课程统计模型与统计实验的建设和实践 王 庚 236

“任务导向实践课程开发模式”的探索与实践

——统计分析报告与论文写作项目教学研究 葛联迎 242

移动互联时代抽样调查的设计问题

亢志伟 侯志强

摘要：本文分析了基于网络社区和手机互联网及应用发展背景下的抽样调查设计与实施的新构想。本文具体讨论在网络社区中怎样选用调查方式、问卷设计和控制网络调查的质量及提高样本的代表性问题；本文也对近年来手机网络的爆炸式发展提出了一些抽样调查的见解；在本文的最后，单独列出在手机互联迅猛发展背景下的抽样调查发展。文章从新的手机浏览标准和 APP 应用两方面来展开，展示移动应用对于抽样调查的冲击和抽样调查的新方式。

关键词：网络调查；手机调查；抽样设计

一、引言

在当今社会，抽样调查作为一种为种类繁多的项目提供统计数据的手段，早已被人们定为可行性的理论和实践方法。抽样调查的适用领域有社会学、政治学、人口学、经济学、社会心理学、教育和公共卫生等，在理论和实践不断发展的基础上，抽样调查进行不同规模的形成、检验和改进研究假设。^①大多数国家政府部门会乐于采用抽样调查来了解经济、社会等各种参数，从而达到做出相应决策的目的，尤其是对于教育、就业和失业、营养、收入和支出、旅行方式、经济决策产生的结果等许多统计方面投入大量的人力物力；政府部门还对各大行业 and 部门进行抽样调查，比如原材料市场，以获取重要的市场和经济信息。除了政府以外，地方政府甚至单个市场研究人员都会采用抽样调查理论方法来进行地方性规划统计，研究消费者心理。在政治领域，抽样调查仍然发挥着重要的作用，各种民意调查实时地观测着公众舆论的走向和趋势。

随着互联网的发展，传统抽样调查的方式由面访式调查、电话调查、邮寄调查、互动式语音应答（IVR）调查等随着网络调查具有自愿性、经济性、时效性、方便性、匿名性等诸多特征演变成了节省成本、提高效率、受众广泛等特征的网络抽样调查方式，但与此同时，网络也带来了更多的调查质量不可控因素，所以也存在很多不

① 李政忠. 网络调查所面临的问题与解决建议 [J]. 资讯社会研究, 2004 (1).

足。在国外,近年来的网络随机化调查方法研究不论从抽样设计还是估计量的构造,都开始深入到复杂网络抽样调查中,主要包括郑大世敏等人讨论了网络分层抽样下属性特征随机化调查理论与方法;索罗·凯西等人利用二级连续抽样方案改进了 Warner RRT;张纭炬讨论了利用两个独立的子样本同时顾及敏感属性比例。虽然关于随机化调查的研究有很多,但是这方面的研究还远不够全面和系统,主要体现在这些调查方法没有统一的评价标准并且与抽样设计有关。在随机化调查中,关于属性特征和数量特征具有许多具体方法,其方法中抽样设计多见于 SRSWR、SRSWOR、PPS 等方案。李政忠(2004 年)在国内第一次探讨了网络抽样调查工具的可行性和局限性,并说明为了解决网络样本误差必须对样本进行加权处理,这包括诸多变量,其中就包括人口统计变量、生活方式变量和人对于其态度等变量。

近年来,互联网的发展非常迅猛,直到现在,已经发展到渗透到每一个人的生活当中。前几年,网络人群主要聚集在各种主题论坛当中,但现今上网人群已经聚集在各种社交网络社区当中,过去 10 年间,随着科技快速发展,作为时代前沿群体的年轻人越来越多地将他们的生活和手机、电脑、微博等数字产品及平台连接起来。据统计,截至 2012 年 9 月底,中国网民规模达到 5.5 亿人,微博实际用户规模已经达到 2.8 亿人。对于上述发展,统计部门不能不着重考虑这类新兴的抽样调查方法。实际上这种网络调查方法已经发展了几年,网络调查有许多优点是传统调查所没有的,最主要的是节省成本和统计速度快,节省了许多中间工作,是提高统计效率的一种有效方式。

二、网络调查区别于传统调查应注意的问题

在传统调查当中,调查费用、回答率和数据质量是最关键的问题,这也是网络调查所要面对的问题之一;但基于互联网的特点来说,这三大问题出现得并不如传统调查那么明显。对于调查费用问题,因为互联网的特点,比传统调查大力组织人力消耗资金要少得多;对于回答率问题,网络本身的匿名性和方便性会使无回答率变得低很多;对于数据质量和进一步提高回答率方面,对于数据质量还有待进一步提高,而回答率方面,要面临一些问题但却不是主要问题,这些问题的关键是网络调查方式的选择。提高调查者的回应率和回收率问题,相对于传统调查的网络调查与传统调查计算抽样和非抽样误差时有何区别,以及网络抽样的质量控制问题,虽然说在传统抽样中这些问题是抽样的主要问题,但是在网络抽样时代,这些问题却不是那么难办了。在中国各个社区中,以新浪微博和人人网的用户最多,黏性用户同样如此,我们主要以他们为主要研究对象。

(一) 调查方式的选择

大多数研究都是从回答率和具体抽样调查设计等方面来评价不同调查方式的

调查效果,调查方法其实各有优点和缺点,不能一概而论地肯定和否定它们的用处。比如网络社区发布一项调查,这里要分两种情况:第一种是大众比较关心的话题调查,这种情况是大众都会乐于去回答这些问题并且通过网络社区分享,然后看到的人有人也会回答问题并分享,这种方式不会带来明显的调查费用,而且效果非常好;第二种情况是大众不太关心的话题,这通常需要一定的激励方式,比如参加调查可以参加抽奖等,这样的抽样调查费用需要为: $C = \sum_{i=1} c_i + \varepsilon$, 期中 c_i 为付给每个人的奖励成本, ε 为其他交易费用,当消息传递快速顺畅时, ε 趋于0。^① 此外,网络调查可以与传统调查相结合。比方说,高校联盟组织高校进行教学调查,需要抽取一些学生,抽取学生的方式采用传统抽样方式进行,有选择性地选择学生进行抽样调查,激励可以采用多种方式,然后集中进行网络抽样调查。具体来说,考虑一个网络抽样调查的方式要根据所在社区人群的偏好以及整体活跃度制定相应的调查方式。

为了强调对于传统抽样调查的移动抽样的费用优势问题,我们首先要确定网络抽样调查属于什么抽样调查方式,如果显然,其特性属于分层随机抽样,那么我们以传统分层随机抽样和移动网络分层随机抽样为标准进行费用计算。需要说明的是,分层随机抽样中,层越大,层内变异越大,而在该层平均每个单元的抽样费用越小,则在该层中的抽样应越多,所以首先我们要计算最优分层和费用。

我们主要考虑简单的线性费用函数:

$$C_T = c_0 + \sum_{h=1}^L c_h n_h$$

式中, C_T 是总费用, c_0 是与样本量无关的固定费用, c_h 而是在第 h 层中抽取一个单元的平均费用, n_h 是第 h 层的样本量。

已知估计量的方差为

$$V(\bar{y}_{st}) = V\left(\sum_{h=1}^L W_h \bar{y}_h\right) = \sum_{h=1}^L \frac{W_h^2 S_h^2}{n_h} - \sum_{h=1}^L \frac{W_h S_h^2}{N}$$

其中, $W_h = \frac{N_h}{N}$ 为权层, N_h 为第 h 层的单元数, S_h^2 表示第 h 层总体方差。

下面计算最优分层和费用,这是一个在给定约束条件下求多元函数极值的问题,应该使用拉格朗日乘数法。这里,约束条件为 $C' = \sum_{h=1}^L c_h n_h$, 多元函数为

$V' = \sum_{h=1}^L \frac{W_h^2 S_h^2}{n_h}$ 。首先构造拉格朗日函数

$$f(n_1, n_2, \dots, n_L; \lambda) = \sum_{h=1}^L \frac{W_h^2 S_h^2}{n_h} + \lambda \left(\sum_{h=1}^L c_h n_h - C' \right)$$

① 方佳明. 网络调查中激励的有效性: 基于荟萃分析的结果 [J]. 管理评论, 2013 (1).

则结果是每个层分配样本量为 $n_h = n \cdot \frac{\frac{W_h S_h}{\sqrt{c_h}}}{\sum_{h=1}^L \frac{W_h S_h}{\sqrt{c_h}}}$ 时, 总费用最小。

再考虑移动网络分层随机抽样, 由于移动互联网络的特性使每个层样本量是可控的, 只需在代码上进行控制即可, 所以上述公式仍然可用, 有总费用公式为:

$$C_T = c_0 + \sum_{h=1}^L c_h \frac{n \cdot \frac{W_h S_h}{\sqrt{c_h}}}{\sum_{h=1}^L \frac{W_h S_h}{\sqrt{c_h}}}$$

其中 c_0 为固定费用, 相对于传统分层随机抽样调查, 移动网络分层抽样调查的固定费用可以忽略不计, 所以其费用简化为:

$$C_T' = \sum_{h=1}^L c_h' \frac{n \cdot \frac{W_h S_h}{\sqrt{c_h'}}}{\sum_{h=1}^L \frac{W_h S_h}{\sqrt{c_h'}}} = \sum_{h=1}^L \frac{n W_h S_h \sqrt{c_h'}}{\sum_{h=1}^L \frac{W_h S_h}{\sqrt{c_h'}}}$$

所以 c_h 越高, 则总费用越高。传统调查的费用 c_h 通常为纸质、社区组织人力等费用的总和; 而移动网络抽样调查则不需要纸质和大量的人力成本, 只需要平台费用 (如果靠自己的平台组织调查则无需费用或者奖励激励费用), 所以总体上 $C_T' < C_T$ 。

(二) 关于调查问卷设计

把问卷投入到相应的网络社区之前, 先要对调查问卷进行设计。网络调查问卷设计涉及的方面非常多, 用多种手段使调查问卷多样化, 更贴近所受群体的体验需求, 这都是需要注意的问题。做这些前期工作也是为了让网络社区的参与者更能积极互动, 贴近参与者的生活和所关注的趋势、潮流以及民生。网络抽样调查设计的主要因素有: 图标布局、问题陈述和每页的问题数目。另一种说法是, 要使网络调查成功实施, 需要注意调查问卷设计问题、群体选择、提高回答率的激励、调查结果分类。从技术上来说, 需要多浏览器支持, 避免重复提交, 提出问题要逻辑化, 较长问卷要有保存功能。其实, 这些问题只要稍加注意就可以了, 网络科技日益发展的今天, 许多技术上的问题已经不是问题了。

基于微博和人人网等 SNS 社区的抽样调查与传统纸质调查虽然媒介不同, 但是就问卷设计来说, 框架是相似的。大体框架包括在进行调查前需要进行词汇使用频度的统计, 然后以消除认知偏差。网页调查具有成本低并且传播快速的优点, 但是这是以调查方式所依附的主体背景是不是足够吸引人或者奖品激励是不

是足够为前提的。网页调查可以放进不可被被调查者更改的数据库,其实这些也不必太过在意,现今大多调查应依附于第三方平台的支持,安全性还是比较有保障的,而且调查对象很少有动机这样做。另外一点,调查设计尽量采用非开放型答案的方式,否则会带来调查结果统计的困难,增加人力成本,也可能带来更多的无用信息,而且也带来了更多的无回答率。

根据统计表明,一份简单明了的调查设计页面或者华丽页面对回答率是有影响的,两者的回答率分别为41.1%和36.29%,网页界面设计的复杂性增加了受调查者的反感,因而更多的无回答出现。另外,简单页面也比华丽页面有更高的完成率。还有一份统计表明,一屏设计多个项目比设计成单个项目完成时间更快。最后,命令按钮和文本框同样对调查结果的回答率造成影响。

就上面的问题,我们都需要给予足够的重视,因为回答率是非常重要的一个因素。对于纸质调查来说,我们能从制度上保证一些回答率,比如教育部向学校下达一项抽样调查任务,学校可以以样本轮换的方式抽取一些学生进行这次调查,这带有一些制度上的强制性,有效地保证了调查结果的回收。但是网络调查则不同,在保证调查问卷设计的同时,还需要考虑受调查者买账不买账的问题,网络是自由的,网民只会关注他所关注的东西,问卷调查不能空洞洞,必须尽量与潮流趋势和民生结合才能最大限度地使网民作答并且转发传播,必要时,要有参与奖品激励。

在传统分层抽样调查中,采用西蒙斯模型进行层数分配:

$$Z_i = P_i \pi_{A_i} + (1 - P_i) \pi_{B_i}, i = 1, 2, \dots, k$$

其中, Z_i 为第 i 层中正确回答人的比例,而 π_{A_i} 是第 i 层中有敏感性特征的人群比例, P_i 是第 i 层中被抽样者中写有敏感性问题卡片的比例,则 π_{A_i} 极大似然估计为:

$$\hat{\pi}_{A_i} = \frac{\hat{Z}_i - (1 - P_i) \pi_{B_i}}{P_i}, i = 1, 2, \dots, k$$

由于已知 $\hat{Z}_i$ 服从 $B(n_i, Z_i)$, 而且各层回答相互独立, 所以 π_A 的极大似然估计为:

$$\hat{\pi}_A = \sum_{i=1}^k w_i \hat{\pi}_{A_i} = \sum_{i=1}^k w_i \left[\frac{\hat{Z}_i - (1 - P_i) \pi_{B_i}}{P_i} \right]$$

其中, $w_i = \frac{N_i}{n}$ 为各层权层, 有 $w = \sum_{i=1}^k w_i = 1$; 又因为 $\hat{\pi}_{A_i}$ 是 π_{A_i} 的无偏估计, 则 $E(\hat{\pi}_{A_i}) = \pi_{A_i}$, 于是有

$$\begin{aligned} \text{var}(\hat{\pi}_A) &= \text{var}\left[\sum_{i=1}^k w_i \hat{\pi}_{A_i}\right] = \sum_{i=1}^k w_i^2 \text{var}(\hat{\pi}_{A_i}) \\ &= \sum_{i=1}^k \frac{w_i^2}{n_i} \left[\pi_{A_i}(1 - \pi_{A_i}) + \frac{(1 - P_i)^2 \pi_{B_i}(1 - \pi_{B_i})}{P_i^2} + \frac{P_i(1 - P_i)(\pi_{A_i} + \pi_{B_i} - 2\pi_{A_i}\pi_{B_i})}{nP_i^2} \right] \end{aligned}$$

这里 π_A 通常未知,但是可以通过先验信息获得, π_{B_i} 设计为已知,所以在不考虑费用的情况下,各层样本量最优分配公式为:

$$\frac{n_i}{n} = \frac{\frac{w_i}{n_i} \left[\pi_{A_i}(1 - \pi_{A_i}) + \frac{(1 - P_i)^2 \pi_{B_i}(1 - \pi_{B_i})}{P_i^2} + \frac{P_i(1 - P_i)(\pi_{A_i} + \pi_{B_i} - 2\pi_{A_i}\pi_{B_i})}{nP_i^2} \right]^{\frac{1}{2}}}{\sum_{i=1}^k \frac{w_i^2}{n_i} \left[\pi_{A_i}(1 - \pi_{A_i}) + \frac{(1 - P_i)^2 \pi_{B_i}(1 - \pi_{B_i})}{P_i^2} + \frac{P_i(1 - P_i)(\pi_{A_i} + \pi_{B_i} - 2\pi_{A_i}\pi_{B_i})}{nP_i^2} \right]^{\frac{1}{2}}}$$

由于移动网络分层随机抽样相对于传统调查方法的特殊性,无回答问题其实可以避免好多隐私问题,如果给予相应的奖励机制,是可以让被调查者乐于回答并回答敏感性问题的。

(三) 网络抽样的质量控制

为了保证网络抽样调查的高回答率,我们采用了奖品激励的方式来提高回答率,但是这也衍生出一个问题,那就是为了奖品而奖品。我们先从整体上来说一下质量控制的问题,其实质量控制问题出现的原因仅仅是因为只注重部分不注重整体而出现的误差,但是这个问题仅仅是传统抽样调查出现的问题之一,早已被谈论了很长时间,有经验的问卷调查者不会出现这些错误。网络中对于问卷的随便选择,目的是为奖品的获得,这些才是网络调查质量控制的重点所在。

为了避免不真实回答,传统分层随机抽样仍然采用西蒙斯模型,所不同的是受访者回答真实敏感性问题变成了一个概率问题,所以设 T_i 为第 i 层受访者在回答敏感问题时是真实回答的概率,则 $T = \sum_{i=1}^k w_i T_i$, 所以西蒙斯模型变为:

$$Z_i = P_i \pi_{A_i} T_i + (1 - P_i) \pi_{B_i}$$

但此时 $\hat{\pi}_A$ 不再是 π_A 的无偏估计量,所以我们用均方误差 MSE 作为衡量标准:

$$\text{MSE}(\hat{\pi}_A) = \frac{1}{n} \left\{ \sum_{i=1}^k w_i \left[\pi_{A_i} T_i (1 - \pi_{A_i} T_i) + \frac{(1 - P_i)^2 \pi_{B_i} (1 - \pi_{B_i})}{P_i^2} + \frac{P_i(1 - P_i)(\pi_{A_i} T_i + \pi_{B_i} - 2\pi_{A_i}\pi_{B_i} T_i)}{nP_i^2} \right]^{\frac{1}{2}} \right\}^2$$

其实问卷回收时,有效减少移动网络分层随机抽样调查问卷填写随意性的方法是进行逻辑程序设计。比如说,如果一份调查问卷每项调查只有四个选项,只敷衍了事的受调查者通常会快速把每项调查只填写相同位置的选项。^① 笔者认为,程序设计需要考虑两点,只要符合这两点之一就可剔除这份问卷进入数据库:一

① 曾鸿.从我国网民特征看网络调查方案设计[J].经济问题探索,2006(1).

是一直重复选择一个位置选项的；二是作答时间过短的。这在技术上不难实现。^①另外，如果奖品丰厚，程序员可以把这些限定放在受调查者作答的时候进行，只要受调查者敷衍进行回答满足了后台的限定条件，受调查者将无法提交问卷，这时候受调查者就会有激励去好好填写自己的调查问卷了。所以在提高调查质量问题上，移动网络分层抽样具有得天独厚的优势。

三、基于移动网络的抽样调查探讨

艾媒咨询公司的最新数据显示，2011 年中国移动互联网市场产值达 389.7 亿元，2012 年产值达 710.4 亿元，到 2013 年产值将超过 1 000 亿元。另有数据显示，截至 2012 年 6 月，我国手机网民规模达到 3.88 亿，网民中用手机接入互联网的用户占比已达到 72.2%，手机首次超过台式电脑，成为我国网民的第一大上网终端，移动互联网成为新的掘金地。在 2012 年广东互联网大会上，移动互联网的话题占据主导地位，三大运营商以及传统的互联网服务商都将讨论的焦点放在移动互联网领域。业内人士和专家都认为，2013 年将是移动互联网的爆发之年。随着网络智能化的普及，手机应用层出不穷，便携性和随时随地性带给用户不同于 PC 的体验。随着 Windowsphone8、Android、苹果 OS 系统的出现，编程人员极大地在增强用户的体验。据各大社区网站以及热门 APP 统计，未来手机互联将超过 PC 成为网络互动主体。

随着手机行业的发展，我们必须看到未来网络抽样调查的趋势，那就是基于移动互联网和手机抽样调查 APP 应用的抽样调查新方向。未来的网络抽样调查将是以智能手机持有者为中心，手机移动应用、第三方平台支持、抽样调查官方的三位一体的网络抽样调查格局。

首先，关于抽样调查主页的建立，在对原有互联网抽样调查需要注意的事项以外，我们需要跟进手机网页的特点进行技术优化。一是要注意手机网页标准的发展进程，现今手机行业网页标准主要是 WAP 网页，对于电脑上的 WWW 标准，手机还不是很有很好的体验，所以抽样调查设计要考虑进行 WAP 网页的优化；此外，将来手机网页标准已经确定采用 HTML5 标准，网络抽样调查设计要看到这一发展趋势。二是要注重手机用户的网页体验，手机工业标准一直在变化，从手机尺寸到网页标准都参差不齐，进行编程的人员一定要考虑这一变化，及时做好网页体验工作，确保抽样调查的高回答率。

其次，就是手机 APP 应用的开发，如果宣传得当或者嵌套在一些社交或其他软件的话，就会得到广泛的传播，所以这种方式一定是独立开发加与第三方合作为前提条件的，但是这种 APP 开发势必导致调查费用的增加。调查制定者一

^① Cihan Cobanoglu, Nesrin Cobanoglu. The effect of incentives in web surveys: application and ethical considerations [J]. International Journal of Market Research. 2003.

定要选择合适的方式以减少成本。例如,诺基亚开发出一套易于各个手机平台移植的编程标准 QT 平台,这样就极大地减少了各个手机平台开发人员的雇佣,从而减少了费用。

最后还要说明的一个问题就是对于手机用户人群的区分,用于各个品牌和价位的人一般会有相应的人群特征,编程系统可以利用 APP 读取手机的这些信息来达到区分各个层次人群的目的。如果需要其他信息,可以在进行调查之前向用户索取需要的信息。

但现今,还没有进行完全网络调查的土壤,拿着智能手机和上社交网站的人大多数是年轻人,对于有些抽样调查来说,这样的调查会有明显误差,所以还是需要网络调查与传统调查相结合的方式建立科学的抽样调查方式。^①

四、结论

基于现今互联网的发展,网络抽样调查设计和实施过程中必须遵循一些新思路和新方法,传统抽样调查设计和具体实施方法已经不能适应时代的要求,但是总的纲领仍然是不变的。而且对于不同网络社区及应用来说,应该采取不同的调查策略,合理设计抽样调查方案,因为网络调查者的主体是所在社区的使用者,他们有各自的背景,不能一概而论去设计一套调查方法和问卷。所以每个抽样设计必须与相应的用户圈子相关,并考虑以激励方式尽量减少无回答问题。随着互联网的发展,应该推陈出新网络调查的方式,最关键的是时刻关注大众流量集中在哪里,然后用古典理论方法进行创新。^②

作者简介

亢志伟(1985年生),男,北京市丰台区人,北方工业大学理学院在读硕士研究生方向:抽样统计

通信地址:北京市石景山区晋元庄路5号北方工业大学理学院,100144

电子邮箱:nasly1@126.com

侯志强(1970年生),男,山西省、北方工业大学理学院副教授,博士,硕士生导师

研究方向:抽样技术,经济计量分析

通信地址:北京市石景山区晋元庄路5号北方工业大学理学院,100144

电子邮箱:houzq@ncut.edu.cn

① Dan J Kim, Donald L. Ferrin, H. Raghav Rao. A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents [M]. Decision Support. 2008.

② 李珊平. 课程专题网络调查研究: 基于建构主义理论 [J]. 市场周刊(理论研究), 2013(1).

北京城镇低收入居民分享经济增长 成果状况及影响因素的调查研究*

阮 敬

摘要：本文在构建一种基于收入分布的福利损失函数基础上，利用2006～2011年北京社会经济生活指数住户调查数据对北京市经济发展过程中的城镇低收入群体收入分享状况进行了测度，并从经济增长、收入不平等和物价水平三个方面对分享状况进行了因素分解。结果表明，北京市经济的快速增长能够在一定程度上自动增加低收入群体的分享能力和份额，但其分享到的成果近年来有被收入不平等和物价上涨因素抵消的趋势。

关键词：低收入群体；城镇低收入群体；收入分享；收入份额

一、引言

中国经济在近20年来经历了一个高速发展的时期，但同时却出现了收入差距逐步扩大的局面。与此同时，居民收入的基尼系数从1987年的0.299（世界银行，2012年）上升至2009年的0.48（美国中央情报局，2011年），已经濒临警戒线水平。而北京市经济发展水平在全国处于领先水平，收入分配状况不容小觑，因此在北京经济发展的过程中关注收入分配问题，尤其是低收入群体分享经济增长成果的问题，显得尤为重要。对于经济系统中的不同群体而言，其收入是否能够得到相对改善，是经济增长过程中衡量不同群体是否分享增长成果的重要特征。

理论上对低收入群体的研究多集中于在低收入群体如何识别或划分，即低收入线如何确定，以及在此基础上如何考察其规模，并进一步提出解决低收入群体收入分配问题的政策和措施等方面。拉瓦雷（1994年）、程（1996年）、拉瓦雷和陈（2003年）、宋（2004年）、魏等（2011年）、海和约翰（2011年）等就低收入群体中的贫困问题做了详细研究，并对贫困、经济增长和收入不平等之间的关系进行了大量的详细论述和实证分析，为研究低收入群体在经济发展过程中如何通过缓解贫困和降低不平等状况来获得更多的收入分享提供了理论基础和实际经验。

* 基金项目：北京市优秀人才培养资助项目“群体异质性收入分布的统计测度方法及其实证研究”和北京市“青年英才计划”的阶段成果。

拉瓦雷和陈(2003年)定义的GIC曲线和宋(2004年)定义的PGC曲线,为上述理论问题的实证分析过程提供了技术和方法。而国际上普遍认可的“涓滴式”和“瞄准式”扶贫政策,为解决低收入群体在经济发展过程中的福利分享测度问题提供了相关的理论基础和政策框架。

在上述理论研究基础上,国内学者也对该问题进行过实证分析和考察。马晓河(2003年)针对我国低收入与高收入的不平等程度进行了探讨,分析了不同收入阶层的现实状况;林伯强(2003年)、洪兴建和高鸿祯(2005年)利用国家统计局公布的官方数据对经济增长和贫困减缓问题进行了实证分析;万广华、张茵(2006年)利用两组住户调查数据对贫困群体分享增长成果的影响因素从不平等程度和经济增长程度两个方面进行了分解;冯星光和张晓静(2006年)利用北京市住户调查数据也进行了同样的分解,并且得到近年来北京市经济发展由较强亲富增长转变为较弱亲贫困增长的结论;胡兵、胡宝娣、赖景生(2005年)利用中国农村住户调查数据和CPI进行缩减的贫困线,分析了经济发展是否有利于农村穷人的问题。这些研究均得到收入不平等在近年来有上升的趋势,并在一定程度上研究了贫困缓解和经济发展的互动关系。阮敬(2007年)以及纪宏和阮敬(2007年)在综合考察“涓滴式”和“瞄准式”扶贫政策基础上,提出重新定义的“亲贫困”增长及扶贫政策,强调在经济增长的过程中要注重贫困群体的利益和分享成果能力的提高,并且讨论了用于测度该种情形的收入分配效应曲线。谢东梅(2008年)对我国农村低收入群体的贫困变动及其影响因素利用省份数据进行了FGT类指数测算的实证分析。刘扬等(2010年)在研究城镇低收入问题过程中引入非参数核密度估计方法并利用北京市住户调查数据研究了北京市低收入群体分享增长成果的问题。

在现有对低收入群体的界定和规模测算研究中,往往是在切实获得住户调查数据的基础之上,依据单纯收入分组比例或收入线的划分来界定低收入群体,并据此研究其群体特征和状况,并没有结合经济发展过程来综合考察低收入群体的收入福利分配问题。所以,本文在充分考虑低收入群体用收入衡量的福利分配在经济发展过程中的动态变化及影响因素,反映出北京市低收入群体在经济高速发展过程中是否分好了增长“蛋糕”的过程和结果,为进一步深化收入分配体制改革提供理论依据和数据支持。

二、数据和测算方法

(一) 数据

本文数据源自首都经济贸易大学统计学院自2006年以来持续实施的“北京社

会经济生活指数体系”数据库中的住户收入调查数据。该数据采用问卷发放和计算机辅助电话调查(CATI)两种方式,在北京市城区连续6年对城镇居民社会生活的满意状况进行了调查。调查样本分布于北京市所辖16个区(2009年以前为18个辖区),2006年得到有效样本量10 934个,2007年有效样本量为10 652个,2008年有效样本量为7 900个,2009年有效样本量为8 490个,2010年有效样本量为7 987个,2011年有效样本量为7 512个。有关收入数据的样本信息如表1所示。

表1 北京社会经济生活满意状况调查中收入数据的样本构成(2006~2011年)

年 份 家庭人均月 收入(%)	2006	2007	2008	2009	2010	2011
无收入	3.64	4.20	3.28	4.16	3.28	2.78
低于1 000元	9.85	8.62	8.86	5.50	5.65	4.17
1 000~2 000元	29.18	26.75	28.04	24.11	23.21	18.36
2 000~3 000元	32.94	33.61	19.57	27.71	24.71	28.24
3 000~5 000元	15.66	17.51	28.40	22.55	23.78	27.01
5 000~8 000元	4.86	5.64	7.59	9.89	12.46	11.41
8 000~10 000元	1.97	1.91	2.63	3.52	4.26	4.78
10 000元以上	1.89	1.76	1.63	2.55	2.67	3.24

由表1数据可以看出,随着经济的快速发展,北京市城镇居民中的低收入群体所占比例逐年降低,而中等收入群体的比例稳固提升,高收入群体比例也有所增加,客观上充分体现了北京市经济发展所带来的好处。但是,这种好处是否真正流向了低收入人群,切实提升了他们参与经济发展的积极性并使其从中获得更多利益呢?从主观上看,本文所用调查数据能够在一定程度上通过低收入群体对其自身收入的满意状况来反映上述现实问题,调查结果如表2所示。

表2 低收入群体对收入的满意状况

年份	低收入群体对自身收入的满意状况	全社会的收入满意状况
2006	42.97	60.22
2007	45.96	66.20
2008	47.56	68.76
2009	58.00	67.62
2010	59.41	67.39
2011	59.94	64.32

注:满意状况>50表示满意;满意状况<50表示不满意;满意状况=50表示一般。