

DOI: 10.16538/j.cnki.fem.20220716.301

二手数据用于多方法研究的建议

——以消费者行为研究为例

柳武妹, 杨巧英, 康娜

(兰州大学管理学院, 甘肃兰州 730000)

摘要:近年来,对研究结论外部影响力和可复制性的寻求引发了管理学者对多方法研究的关注。而二手数据法在增强研究结论的外部效度和现实影响力方面发挥着不可替代的作用。因此,在多方法研究中使用二手数据成为一种主要趋势。本文以2015年1月—2022年6月期间发表在JCR、《管理世界》等十余种国内外管理学期刊中的213篇(197篇英文,16篇中文)开展了二手数据用于研究的消费者行为领域多方法文章为研究样本,总结并归纳了二手数据在多方法研究中的用途、优势与运用挑战。随后,本文对多方法研究中二手数据法的独特开展流程进行了总结和说明。在此基础上,本文针对具体开展含二手数据的多方法研究时可能遇到的一系列问题给出了思考和建议。我们希望这些工作能够为拟借助二手数据开展多方法研究的学者提供参考。

关键词:二手数据;多方法研究;操作流程;消费者行为

中图分类号:F270 **文献标识码:**A **文章编号:**1001-4950(2022)11-0121-17

一、引言

2015年学术界出现了社会科学研究的可复制性危机。对此,学者们进行了多方面的归因(Fiske, 2016; Maner, 2016; Schaller, 2016)。其中,从研究方法层面来看,无法复制危机与单一方法(尤其是实验室实验)的重度依赖有关(Maner, 2016)。自此,多方法研究开始受到越来越多学者的关注。二手数据研究方法使用的是大范围的外部数据以及基于真实世界的历史数据,这种数据来源的可追溯性以及真实性提升了研究结论的外部效度和可复制性。因此,使用二手数据在进行多方法研究时更加受到学者们的鼓励和推崇。其中,管理学科下的消费者行为领域使用二手数据开展多方法研究的势头尤其明显。

自2015年以来,消费者行为学的顶级期刊之一*Journal of Consumer Research* (JCR)上刊发

收稿日期: 2022-01-23

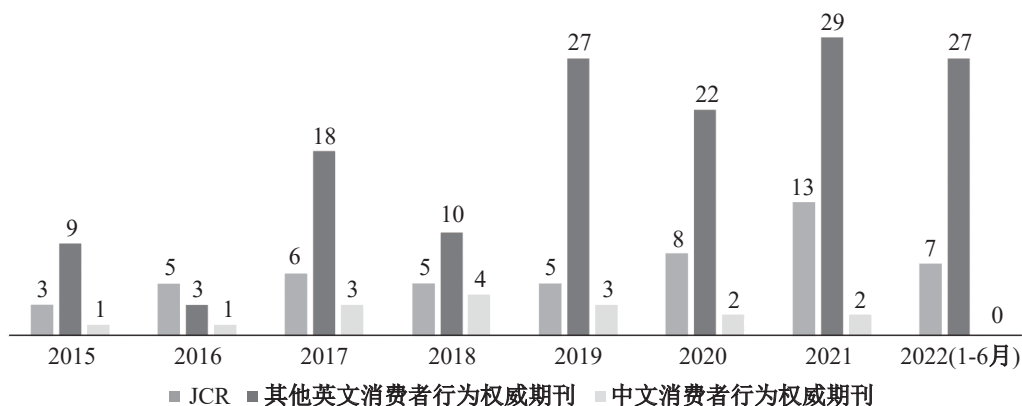
基金项目: 国家自然科学基金面上项目(71972092)

作者简介: 柳武妹(1985—),女,兰州大学管理学院教授,博士生导师;

杨巧英(1997—),女,兰州大学管理学院博士研究生(通讯作者,yqy2513842008@163.com);

康娜(1991—),女,兰州大学管理学院博士研究生。

的借助二手数据开展多方法研究的文章数量逐年增加(参见图1),累计达到52篇,占2015年1月—2022年6月JCR刊文总数的11.8%(52/439)。从发表年份来看,2015年全年仅有3篇多方法文章采用了二手数据法,而2022年上半年就有7篇,占2022年上半年JCR总发文数的29.2%(7/24)。除JCR外,我们还检索了*Journal of Marketing Research*(JMR)、*Journal of Marketing*(JM)等九种消费者行为权威期刊上2015年以来刊发的涉及二手数据的多方法文章。我们也发现,这些期刊上使用二手数据开展多方法研究的文章数量呈上升趋势(参见图1)。然而,当我们检索国内管理学权威期刊(如,《管理世界》《南开管理评论》)2015年以来刊发的消费者行为领域的文章时,我们发现仅有16篇多方法文章使用了二手数据法(参见图1)。



资料来源:本文作者整理。

图1 2015年1月—2022年6月消费者行为领域权威期刊刊发的包含二手数据法的多方法文章数量^①

上述文献计量分析表明,借助二手数据开展多方法研究已经成为国际消费者行为研究中的重要趋势。然而,国内消费者行为领域学者对二手数据子研究的运用非常有限。出现这一现象的原因可能是多数学者只擅长某一种研究方法,对于多方法研究以及如何使用二手数据开展多方法研究存在以下几个方面的困惑:(1)多方法研究中的二手数据子研究用来做什么,有何优势和挑战?(2)多方法研究中的二手数据法和一般二手数据法在操作流程上有何差异?(3)研究者何时、何地以及如何多方法研究中开展二手数据子研究?(4)多方法研究中的二手数据法可以和哪些其他研究方法进行搭配和组合,在文中如何安排位置?(5)借助二手数据开展多方法研究时需要注意哪些问题?

作为管理学研究的一个重要分支,消费者行为领域对多方法的运用,尤其是对二手数据子研究的运用已成为一种趋势。这一重要研究趋势预示着管理学研究未来在方法使用层面的重要变化。因此,我们选用消费者行为领域的文献样本进行系统分析,具体回答用于多方法研究的二手数据子研究应如何开展等问题。本文内容安排如下:首先,介绍用于多方法研究的二手数据法对于消费者行为研究的用途、优势与挑战。随后,以2015年以来发表在JCR、《管理世界》等期刊上的213篇(197篇英文,16篇中文)包含二手数据法的消费者行为领域的多方法文章为例,系统梳理多方法研究中二手数据法的具体开展流程。并且着重阐述和一般二手数据法相比,用于多方法研究的二手数据法在使用方面存在哪些独特之处,以及具体开展包含二手数据的多方法研究时会出现哪些疑问与难题。最后,针对学者在借助二手数据开展多方法研究时可能遇到的一系列疑问与难题给出思考和建议。

^①图中涉及的其他英文权威期刊包括*Journal of Consumer Psychology*、*Journal of Marketing*、*Journal of Marketing Research*、*Journal of International Research in Marketing*、*Journal of Academy of Marketing Science*、*Journal of Retailing*、*Psychology & Marketing*、*Journal of Consumer Affairs*、*Marketing Letters*;中文权威期刊包括《管理世界》《南开管理评论》《中国工业经济》《心理学报》《营销科学学报》。

本文的理论贡献主要有三点:(1)本文是对消费者行为领域多方法研究以及二手数据子研究的首次介绍和梳理,弥补了目前消费者行为研究对二手数据子研究介绍和倡导的不足。(2)本文提出了二手数据子研究有别于一般二手数据研究的地方并提供了开展建议层面的独特性学术思考。尽管已有书籍(如陈晓萍等,2008)和文章(如Ellram和Tate,2016)涉及对二手数据的介绍,但是所针对的是单一二手数据研究如何开展和使用等问题,无法回答一般的二手数据研究和多方法研究中的二手数据子研究究竟在操作和运用上存在何种区别,以及实际开展含二手数据的多方法研究时应注意哪些问题。本文则针对这些问题进行了思考与解答。(3)本文可以起到抛砖引玉的作用。通过详细介绍多方法研究中二手数据子研究的具体运用和开展建议,本文能促使更多的国内学者关注和运用二手数据子研究,进而催生出更多具有实践价值和现实影响力的顶天立地的研究成果。

二、二手数据在多方法研究中的用途、优势与运用挑战

(一)多方法研究中二手数据的用途

二手数据(secondary data)通常被界定为由研究人员以外的人为了不同的目的而收集的定量或定性数据(Ellram和Tate,2016)。因此,二手数据在数据来源上指通过公共及公开渠道取得的已存数据,这使得二手数据具备历史性、间接性、被试不可及性等特征。消费者行为研究中不乏使用二手数据的研究,且二手数据可与实验法、调查法等其他方法共同使用以实现研究者想要开展多方法研究的目的。具体而言,学者们在开展多方法研究时可以借助二手数据实现下列三种用途:

1.辅助检验理论观点

通常情况下研究者会借助二手数据分析来检验感兴趣的变量之间的关系,并通过真实历史数据来验证所观察到的社会现象。例如,Fan等(2020)借助从DemoHour上收集的902个众筹项目和116 153项捐款数据,揭示了项目发起首日的小额捐款比例能够正向预测筹资结果。通过二手数据来揭示真实市场中的一些消费现象,能够为研究的主要假设提供关键性的现实证据。值得注意的是,有个别研究也使用二手数据来检验变量间关系的内在机制和边界,以及排除一些可能的替代解释(Aspara和Wittkowski,2019;Berger等,2021;Xu等,2021)。对于那些与研究匹配性较好的二手数据而言,若能找到合理的相关代理指标,则也可以像实验研究一样进行中介效应和调节效应检验。例如,Gullo等(2019)利用一家加州大型食品连锁店的扫描仪面板数据,指出人们在一天中的早上时间段多样化寻求行为(variety-seeking)较低。并且,Gullo等人还将日照程度和日出后的时长作为生理唤醒的代理指标,间接验证了生理唤醒程度的中介作用。

2.辅助理论构建

二手数据研究不仅可以用于理论的检验,在理论构建中同样可以起到至关重要的作用。具体来说,二手数据子研究可以帮助研究者初步挖掘理论观点,为随后的理论构建工作奠定重要基础。以Chung等(2022)的研究为例,该研究旨在阐明消费者参与共享经济的动机,并讨论了这些动机可能导致的一些下游结果。作者首先选取了Airbnb网站(一个提供全球民宿短租公寓预定的平台)上43 343名民宿主人对“Why did you start hosting?”这一问题的文本回答,并借助机器学习(machine learning)的方式提取了这些民宿主人选择出租公寓的动机。借助这一分析结果,作者建立了初步的理论观点。随后在这一发现的基础上,作者又开展了两项实验研究,分别检验了民宿主人进行房屋出租的动机与租客满意度、房屋定价之间的因果关系,从而丰富了研究的理论发现(Chung等,2022)。

3. 辅助量表开发

除上述用途外,二手数据还可以用于量表的辅助开发。先行借助二手数据进行消费现象观测,对于后续相关量表的开发具有重要指导意义。例如,Kappes等(2021)在研究一中就使用来自英国一个货币管理App上2 000多名消费者的账户交易记录,实证分析了消费者在花钱意味着财富(spending implies wealth, SIW)这一信念方面的差异会如何影响其投资和储蓄行为。随后基于现实影响的普遍性,作者进行了相关的SIW量表开发(Kappes等,2021)。从研究形式来看,二手数据的使用使得这一研究区别于传统的量表开发,通过二手数据分析得到的结果为后续的量表开发奠定了重要的现实基础,从而大大提升了研究的理论和实践价值。

(二)多方法研究中二手数据的优势

相比通过其他研究方法(如实验法、调查法、访谈法等)所获取的数据,二手数据有着自身独特的优势,具体包括以下几个方面:

第一,为研究观点提供具有时间跨度的大样本支撑证据。通常采用一手数据的研究方法会由于研究者个人资源(时间、经费和人手等)的限制而难以获取大样本,尤其对于跨时段样本和纵向数据分析而言,只通过获取一手数据很难达到研究者的研究目的。相比之下,二手数据的样本量通常很大,采用二手数据无论是在样本大小还是时间跨度上均具有明显的优势。通过使用特定方式可以轻松获取存在于外部世界的大量数据,弥补了实验法以及调查法等一手数据获取方式在样本量方面的重要不足。例如,Consiglio等(2018)通过使用来自意大利1 421个城市203 071名推特活跃用户所发布的2 194 856条带地理定位的推文,为社会拥挤和消费者口碑(word of mouth)之间的关系提供了大样本的证据。

第二,增加研究结论的客观性与真实性。一些外部数据,尤其是企业的销售数据,是基于市场真实运作以及客观的销售来源的,基本上不包含数据来源者以及数据接收者的主观臆断,或者受主观臆断的影响较小。数据中所包含的信息通常具有较高的客观性,因而在用于检验变量间关系时,可以发挥客观实际的作用。除此之外,由于数据来源者和数据接收者存在时空上的分离,数据获取少了很多人工因素的干扰,并且也可以避免被试猜测实验目的或是出于社会称许性而刻意美化数据。这对于保证数据结果的真实性尤为重要。

第三,增加研究结论的可复制性,使实证研究结论更加可靠。正如本文开头所讲,研究结果的可复制性是学术研究中的一大难点。通常一些一手数据获取方法(如实验法)会由于获取情境的不可复制而使得研究结果无法完全复制,而二手数据研究则避免了这一问题。二手数据在数据来源方面的可追溯性保证了他人可以根据同样的数据源并按照相同的步骤去创建样本,从而保证研究结论的高度可复制性。此外,由于数据的获取避免了很多主观因素的影响,数据的重复检验过程不会受到太多不一致因素的影响,这在一定程度上保证了结果的可复制性。

(三)多方法研究中运用二手数据的挑战

尽管多方法研究中二手数据子研究存在上述优点,然而具体运用二手数据子研究检验研究观点会面临以下挑战:

首先,数据搜索难度高。在寻找用于分析的二手数据时,可能会花费大量时间来搜索适当的数据来源和解释数据(Ellram和Tate,2016)。当研究者决定使用二手数据来佐证文章观点时,少数情况下是已经发现了适合自身研究主题的二手数据,并且这些二手数据可以和自身研究目的完美匹配。对于这类情况而言,搜索数据会较为简单,只需要研究者根据自身研究目的来对数据进行针对性解释。相比之下,多数情况下研究者在想要运用二手数据方法时首先遇到的问题就是从哪里获取合适的的数据。这其中会涉及大范围的数据搜索以及数据匹配工作,可能会致使很多学者在前期搜寻无果后放弃使用二手数据的想法。

其次,数据处理工作繁杂。现有的数据集可能包含很多研究者所需之外的变量,需要研究者花费大量的时间来整理数据以及排除噪音(Saunders等,2016)。在运用实验法以及调查法等可以直接获取数据的方法时,研究人员可以根据具体的研究目的去设计实验以及调查问题去获取相应的数据;而当采用二手数据研究方法时,由于只能间接获得来自于其他渠道的数据,且这些数据通常是基于各种不同目的而形成的,因此研究者只能通过个别信息的匹配来筛选可用的二手数据,这在实际工作量方面具有相对难度。

最后,无法得出因果推断。通过二手数据分析得来的结论一般只能为研究者的假设提供初步的相关性证据,而对于变量间更为重要的因果关系,却无法给出说明。这一局限性主要源于二手数据中内在机制的不可观测,大多数情况下研究者只能借助一定的依据从二手数据中提取出预测变量和结果变量的代理指标,而对于一些个体心理变量,很难从其中找到相应的替代指标。对于一些既看重主效应也看重内在机制的研究而言,二手数据无法作为研究人员所提观点的主要支撑证据,因此研究者需要借助更为严谨的实验法去弥补二手数据研究的不足。

以上总结了二手数据在多方法研究中的用途、优势和运用挑战。总体而言,在多方法研究中采用二手数据法有其自身的必要性。合理借助二手数据研究,能够使研究结论摆脱对单一方法的依赖,从而实现研究结果的可复制性,提高研究的外部影响力。接下来,我们将针对多方法研究中二手数据方法的具体操作流程展开说明。

三、在多方法研究中开展二手数据子研究的操作说明

一项二手数据研究的完整流程,主要包括数据的选择、获取、清洗、处理、分析、稳健性检验、内生性分析等多个步骤。该部分将以消费者行为领域涉及二手数据子研究的213篇多方法文章为例,主要针对多方法研究中的二手数据法与一般二手数据法之间的不同之处展开说明,并在此基础上归纳总结出表1。对于一些没有差异的步骤我们将不再赘述。

表1 多方法研究中的二手数据法和一般二手数据法的操作对比

操作步骤	多方法研究中的二手数据子研究	一般二手数据研究	是否存在重要差异
数据来源	以网络数据为主	以公共数据库为主	√
数据获取	以免费获取为主	以付费获取为主	√
数据规模	小到中等	较大	√
数据清洗	较多清理工作	较少清理工作	√
数据处理	灵活提取目标变量	寻找可用的代理指标	√
数据分析	以统计分析为主	计量经济模型	√
稳健性检验	简单说明或不说明	需要进行严谨的稳健性检验	×
内生性分析	简单分析或不分析	需要进行严谨的内生性分析	×

资料来源:本文作者整理。

(一)从哪里获取二手数据?

本文通过梳理发现,用于多方法研究的二手数据法一般存在四个方面的数据来源。一是网络数据。这类数据主要涵盖一些网络平台(如Facebook、Twitter)以及一些公开的网站(如RottenTomatoes.com)上的消费者行为数据,涉及消费者在网络上发布的评论、视频、帖子以及一些搜索数据等(Pancer等,2019;El Hana和Sabri,2021;Nguyen等,2021)。二是企业数据。具体指来自某个公司、零售商或者连锁商店等经营具体业务的组织的内部数据,包含企业销售数据、企业与客户的交互数据以及顾客追踪数据等(Harding等,2019;Kappes等,2021;Packard和

Berger, 2021)。这类数据只能通过特定组织的渠道获取,因此在数据独特性上具有一定的优势。三是公共数据库数据。通常包括由一些组织开展的大范围调查的数据以及一些研究机构整理的大样本数据(Wang等,2021)。这些数据库的数据通常以结构化的形式呈现。四是国家机构数据。指来自一些国家机构或单位的统计数据,一般会在相关机构的官网公布,研究者可以在得到相关单位的授权后获取。

针对上述四种二手数据来源,我们整理了每种来源所对应的常见数据类型及相关研究的例子,详见表2。通过这一表格,我们能清楚地了解目前消费者行为多方法研究中都使用了哪些二手数据类型。此外,我们还在表格中汇总了每种数据来源的优势、劣势及适用的研究主题。最后,基于我们所检索到的213篇涉及二手数据的多方法文章,我们统计了每种二手数据来源的使用占比。结果发现,49.09%的研究采用了来自网络的二手数据,企业数据次之(30.45%),使用公共数据库以及国家机构数据的研究则分别占到14.55%和5.91%。基于这一分布比例可以看出,不同于传统的单一二手数据研究,多方法研究中二手数据的来源有着很强的学科属性。

表2 二手数据子研究中常用的数据类型汇总

二手数据来源	常见数据类型	相应研究举例	优势	劣势	适用的研究主题
网络数据	网络评论(TripAdvisor、Yelp、Amazon、天猫、淘宝)	Rocklage等(2021);朱华伟等(2017)	数量丰富、种类繁多,可支持开展多种主题的研究	需要借助特定的方法获取;半结构化数据居多,需要人工编码处理	消费者评论、消费者体验、消费者说服、网络口碑、品牌传播
	社交平台发布的帖子、视频(Instagram、Quora论坛、Twitter、Facebook、YouTube、新浪微博)	Chung等(2022);徐岚等(2020)			
	购物平台产品数据(Amazon、天猫、唯品会)	Sun等(2021);周元元等(2017)			
	搜索行为数据(谷歌趋势、百度指数)	Xu等(2021);Ordabayeva和Fernandes(2018)			
	筹资项目数据(Kickstart、Demohour)	Wang等(2021a);Fan等(2020)			
	eBay在线拍卖数据	Reich等(2018);Smith等(2016)			
	Billboard单曲排行榜数据	Valsesia等(2016);Nunes等(2015)			
	电影票房数据(Box Office Mojo)	Ryoo等(2021);Hofmann等(2017)			
报纸、杂志、新闻媒体上的文本数据	黄静等(2019);Kopalle等(2017)				
企业数据	企业销售数据	Hodges和Chen(2022);Choe和Kan(2021)	可以获取到系统的结构化数据和纵向数据集,数据独特性较高	只能通过特定渠道获取,需要一定的权限	品牌相关策略的市场影响力、特定促销策略的市场反应、客户满意度
	用户访问数据(网站点击、App使用)	Berger等(2021);Sussman等(2021)			
	客户个人信息数据	Kappes等(2021);Meyners等(2017)			
	客户服务交互数据	Packard和Berger(2021);Singh等(2020)			
	顾客追踪数据	Harding等(2019);Kettle等(2016)			
	企业档案数据	Whitley等(2022);Lawrence等(2019)			

表 2 (续)

二手数据来源	常见数据类型	相应研究举例	优势	劣势	适用的研究主题
公共数据库数据	品牌资产评估(BAV)调查	Gupta和Hagtvedt(2021); Beck等(2020)	以结构化数据的形式呈现,需要较少的处理工作	可支持开展多个主题的研究,缺乏数据独特性	公民道德行为、文化背景以及价值观研究、消费者态度
	中国家庭面板数据(CFPS)	Wang等(2021b)			
	Nielsen消费者面板数据集	Galoni等(2020); Lalwani和Wang(2019)			
	Strategic Business Insights(SBI)调查数据	Mrkva等(2020, 2021)			
	the Lipper Hedge Fund database	Esteky等(2018)			
	Yelp's Academic Dataset	McGraw等(2015)			
国家机构数据	国家疾控中心健康调查数据	Scott等(2020)	可以对其他类型的二手数据进行补充	需要获得相关单位的授权才可获取	社会事件对消费层面的影响、宏观环境及政策因素在个体消费层面的反应
	美国联邦调查局(FBI)县级犯罪率数据	Beck等(2020)			
	芬兰统计局数据(国防军登记数据、车辆注册数据)	Aspara和Wittkowski(2019)			
	美国劳工统计局消费者支出调查数据	Tully和Sharma(2018)			
	美国众议院档案数据	Davidson和Therault(2021)			
	美国农业部国际食品消费模式数据集	Gao等(2017)			
	瑞士劳工组织公司薪酬比率数据	Mohan等(2018)			

资料来源:作者手动整理(基于213篇包含二手数据子研究的多方法文章)。

消费者每天通过手机、电脑等移动设备参与的网络活动数不胜数,这也使得网络尤其是一些社交媒体上存储了丰富的消费者行为数据。丰富多样的网络二手数据来源不仅为消费者行为学者提供了众多有趣的研究视角,同时也使得消费者行为研究中二手数据的运用独特于传统的单一二手数据研究。

(二)如何获取二手数据?

根据研究主题选择合适的二手数据来源之后,需要进行相应的数据获取工作。通常二手数据的获取有免费获取和付费获取两种方式。免费获取主要针对一些网络上的数据,需要研究人员自行在网站或网络平台上抓取和下载,有时需要借助一些媒介或工具才能抓取到研究者所需要的数据。例如Wang等(2021a)就借助Python软件爬取了Kickstart平台上的音乐项目信息,并通过音频挖掘技术(audio mining technology)分析了项目发起人的声音特征对说服结果的影响。此外,通过一些社交媒体分析平台(例如Infegy公司所创建的Atlas平台),也能实现对消费者在社交媒体上的行为数据的实时监测和获取(Nguyen和Chaudhuri, 2019)。付费获取主要针对一些现成的数据库或数据集,需要研究人员购买使用(Lalwani和Wang, 2019)。对于企业数据和一些国家机构数据,则需要通过与目标企业合作或事先取得相关组织或单位的许可才能获得(如,Cai等, 2016; Choe等, 2021)。需要说明的是,一些二手数据(如网络数据)储存方式和呈现形式具有独特性,研究者需要借助特定的工具进行获取。因此,研究者想要借助二手数据来开展多方法研究,一定程度上需要有相关数据获取技术的支持。

(三)如何确定二手数据规模?

二手数据研究中对数据规模的考虑主要包括两个方面,一是时间跨度,二是样本大小。在

数据收集的时间跨度上,短则一个月(Kwan等,2017),长则可达数十年(Ordabayeva和Fernandes,2018;Rocklage等,2021)。而对于样本大小,少则过百(Goor等,2021),多则几十万(Gullo等,2019;Beck等,2020)。具体时间跨度和样本大小则需要根据数据来源、数据特征以及数据获取难易程度进行考量。如来自于网络的二手数据由于其非结构化特征,需要研究人员在获取数据后进行大量的整理和筛选,因此过度追求大样本量反而可操作性不高,而对于一些已经由某些机构或组织整理好的结构化数据集,则遵循应用尽用的原则。虽然学术界公认的一点是,二手数据具有绝对的样本量优势,但是鉴于多方法研究中二手数据的类型并不完全等同于一般的二手数据研究,当研究人员确定了研究主题以及对应的二手数据来源时,就需要注意在样本量与可操作性之间进行权衡。网络二手数据虽然无法像其他结构化数据一样进行很大样本的获取与分析,但是其存在于真实消费环境下的特点完全弥补了样本规模的不足。

(四)如何清洗二手数据?

在本文所分析的213篇消费者行为多方法研究中,二手数据主要以原始的文本形式呈现,并且数据分布较为散乱,因此研究人员在取得相关二手数据后并不能直接用于研究分析,需要根据研究目的进行大量的清理工作。常见的筛选标准包括数据完整性,即排除掉信息缺失或内容不完整的数据(Mrkva等,2020;Gupta和Hagtvedt,2021);分析匹配性,即依据后续分析目的和方法排除掉不匹配目标变量或无法进行分析的数据(Spiller和Belogolova,2017;Kalra等,2020);数据可处理性,即根据数据处理方式排除掉无法进行处理的数据。那些需要用到文本挖掘或语言计算工具的研究通常需要根据特定的词汇来对文本内容进行编码,因而需要排除不符合编码要求的数据(Rocklage等,2021)。除上述较常采用的筛选方式,还需要根据研究主题排除掉一些异常值或极端数据。例如,Consiglio等(2018)在利用推特内容探究社会密度和口碑传播之间的关系时就排除了14个存在异常推特活动城市的数据(推特名人分享了异常数量的推文)。

(五)如何提取目标变量?

在利用二手数据进行分析时,主要会涉及一些关键变量的选取。由于多方法研究中二手数据来源具有广泛性,研究人员在取得数据后需要灵活提取所需变量,具体变量提取方式有别于一般的二手数据研究。

首先,对于预测变量(或自变量),常见的提取方式有四种。其一,使用人工编码的方式进行分组处理。即根据研究目的对二手数据中的原有信息进行人为分组,从而实现对比自变量不同水平下因变量结果差异的目的(例如,Rocklage等,2021)。其二,借助一些文本分析工具或音频挖掘技术对数据进行数值化转换。一些现有的文本处理工具(如Linguistic Inquiry and Word Count,LIWC)(van Laer等,2019)、音频分析软件(如QA5软件)(Wang等,2021a)能够实现按照一定的计算标准对文本或音频进行相应的数值转换,从而得到可用于统计分析的具体变量。其三,使用原始数据中已有的项目或信息作为自变量的代理指标。其四,将其他来源的数据作为自变量代理指标。研究者通常会采用国家层面的调查数据或统计数据来操作化自变量。如Beck等(2020)就借助美国县级犯罪率数据来衡量消费者的个人控制,从而分析了控制感和领导者品牌偏好之间的关系。

其次,对于结果变量(或因变量),常见的衡量方式有两种。其一,用原始数据中已有的行为指标充当因变量。如产品评级(Schlosser等,2016;Yang和Aggarwal,2019)、评论或帖子的点赞及转发数(Kwan等,2017;Pancer等,2019)、销售数量(Cai等,2016;Mohan等,2018)、交易金额(Smith等,2016)等,此外还包括一些离散指标,如品牌选择(Winterich等,2015)、获取方式选择(Aspara和Wittkowski,2019)以及优惠券的使用(Lalwani和Wang,2019)等。值得说明的是,有

时研究人员也会结合数据集中的已有指标,通过不同的计算方式,形成最适合数据分析方案的代理变量(如Nguyen等,2021)。其二,借助一些语言计算工具对文本数据进行转换。较常使用的语言计算工具有The Evaluative Lexicon(评价词典,用于测量文本中个人反应的效价、极端性和情感性)(Rocklage等,2021)、part-of-speech (POS) tagging(一种词性标注工具)(Nguyen等,2021)、LIWC database(语言查询和字数统计程序)(Schlager等,2021)等。借助这些工具就可以衡量文本内容的情绪效价和得分,从而用于后续的统计分析。

最后,对于一些重要的控制变量也需要加以分析。常用作控制变量的内容有人口统计学信息(数据所涉及对象的年龄、性别、种族、收入等变量)(Scott等,2020;Kappes等,2021;Smith等,2021)、数据类型特征(包括评论长度、字数;视频时长、亮度;电影类型、时长等)(Berger等,2021;Rocklage等,2021;Wang等,2021a)、地区层面特征(如GDP、人均收入状况、城市人口规模)(Ordabayeva和Fernandes,2018;Davidson和Theriault,2021)等。除此之外,许多学者也经常将时间因素(如年份、月份、日期等)作为控制变量纳入模型,这对于运用一些时间序列数据或面板数据的研究来说更为普遍(Kettle等,2016;Galoni等,2020)。一般来说,除了上述一般性的选取标准,研究者还需要结合数据特征、研究目的和分析目的选取相应的控制变量,以使得分析过程更加严谨、研究结论更加可靠。

(六)如何分析二手数据?

与传统二手数据研究不同,多方法研究中的二手数据法多与实验法、调查法等配合使用,因此这类研究中对二手数据的分析方法也基本沿用一些常规的统计分析方式,如方差分析(Gupta和Hagtvedt,2021)、回归分析(Xu等,2021)、相关分析(Esteky等,2018)、T检验(Tully和Sharma,2018)、卡方分析(Goor等,2021)等。除此之外,一些学者也借鉴了部分计量经济学方面的分析模型来拟合所取得的数据,并借助机器学习技术和R语言来实现对模型的估计(Harding等,2019;Matz等,2019)。总体而言,在多方法研究中出现的二手数据法在具体分析方面会更趋于简单,较少采用计量经济学的分析范式。这一现象可以解释为二手数据法在消费者行为领域的本土演化。

(七)如何检验分析结果的有效性?

二手数据研究非常必要的环节就是重复检验分析结果的有效性,包括稳健性检验以及内生性分析。用于多方法研究的二手数据研究其具体实现方式和一般二手数据研究并无差异,因此不再赘述。值得说明的是,在多方法研究中二手数据法通常以子研究的形式呈现,起辅助支持研究结论的作用,因此研究者在借助二手数据开展多方法研究时并不会进行很多的稳健性检验以及内生性分析工作。

以消费者行为领域涉及二手数据的多方法研究为分析样本,我们识别出几点在多方法研究中开展二手数据子研究的独特之处。其一,就数据来源而言,多方法研究中二手数据的来源明显区分于一般二手数据研究,且具有较强的学科匹配性。例如消费者行为领域丰富的网络消费数据使得研究者可以开展多种多样的主题研究,而不用局限于数据类型的单一性。其二,多方法研究中的二手数据来源较广,因此在具体的数据获取阶段也存在特定的差异。其三,在多方法研究中运用二手数据时,需要针对特定数据类型合理选择样本规模。其四,研究者需要针对研究目的灵活提取所需变量。当前数据集中无法获取相关变量时,可以考虑使用其他来源的二手数据作为补充。另外,针对已获取二手数据集中的样本群体精准投送调查问卷可以作为另一种获取目标变量的方式。最后,在进行数据分析时,通过简单的相关性分析、方差分析等方法就可以实现研究目的。总体来看,用于多方法研究的二手数据法在具体开展方面可以形成一套区别于传统二手数据研究的流程。

尽管我们已总结了多方法研究中开展二手数据子研究的详细流程,但是在具体实施方面还需要研究者仔细斟酌与考量。例如,外部世界中虽然存在丰富多样的二手数据,但是否一定存在符合研究者当前研究主题的数据,我们并不能给出肯定的答复。关于寻找数据和确定研究主题二者之间谁先谁后的问题,我们也无法做出确定的回答。诚然,先有合适的二手数据再针对性确定研究主题可能是较为容易的一种方式,但是这种做法会引致学术界关于研究动机(“数据驱动”还是“理论驱动”)的争议。关于二手数据在不同的方法组合中具体扮演什么角色,以及如何实现二手数据分析和和其他研究设计之间的优势互补,我们将在下文进行讨论。

四、多方法研究中二手数据的运用建议

上述我们以消费者行为领域213篇包含二手数据的多方法文章为分析样本,详细阐述了二手数据在多方法研究中扮演的角色及在多方法研究中运用二手数据的详细操作流程。在明确了具体如何使用后,研究者需要考虑该怎样借助这一方法去实现开展多方法研究的目的。什么时候适合采用二手数据子研究?二手数据与什么样的研究主题最匹配?二手数据可以与哪些方法组合搭配?二手数据子研究适合出现在文章的什么位置?哪些情形不适合开展二手数据子研究?这些问题都是研究者在实际开展含二手数据的多方法研究时需要重点考虑的。接下来我们将针对这几个方面给出我们的思考和建议。值得说明的是,借助二手数据来开展多方法研究并非消费者行为领域专有,其他管理学研究分支,如人力资源管理、战略管理、运营管理等领域也可以采用这种多方法研究来夯实研究结论。本文以更具代表性的消费者行为领域为出发点,旨在提供一些具体详实的参考建议。

(一)什么时候适合采用二手数据子研究?

二手数据法具有样本量大、可复制性高、外部效度强等优势,同时也具有数据获取难等劣势。这些特点决定了二手数据有其特定的适用情形。

1.理论探索初期。当研究者想要初步验证感兴趣变量之间的关系时,可以考虑采用二手数据来实现。虽然通过设计实验或者调查问卷也可以初步检验主效应,但是这些方法都需要研究者对变量有清晰的界定和测量。相比之下,使用二手数据能够帮助研究者在形成初步研究想法时获取支撑证据。研究者只需要根据现有变量的概念在二手数据中寻找相应的代理指标,并通过简单的相关分析或者回归分析就可以初步检验预期的研究想法。此外,二手数据具有的大样本优势也能够帮助研究者在研究初期建立对理论观点的信心。

2.前期已经开展了在外部效度上存在局限性的单/多方法研究。当研究者已经开展了一系列具有较强内部效度的研究,并且想要寻求研究结论的外部影响力时,可以运用二手数据法来实现。一些研究方法(例如实验法)的优势在于能够严谨地检验变量之间的因果关系和内在机制,但是人为控制的实验环境也大大影响了研究结论的现实普适性,因此基于实验研究的结论去寻找外部世界的支持证据是一种行之有效的方式。相比田野实验以及调查法等同样具备外部效度的研究方法而言,使用二手数据可以借助其大样本优势以及完全来自于现实世界的特点,去弥补研究缺乏外部影响力的不足。

3.想找寻研究结论的历史数据的证据。当研究结论受到可复制性质疑时,采用二手数据是一种不错的解决方式。二手数据本质上是一种历史数据,并且客观存在于特定的渠道。这种历史性和客观性决定了二手数据在来源上可以追溯,在使用上避免了人为因素的干扰。这些优势是其他研究方法无法满足的。因此,研究者只需要给出数据的来源和具体处理步骤,其他人按照相同的方式获取并处理数据,就可以完全复制研究的结论。

4.当对研究主题有较好的把控时。从表2可以看出,消费者行为研究中常用的二手数据虽

为并无直接关系,但是在很多情况下可以成为一些社会因素、环境因素的代理指标,从而可以方便消费者行为研究者聚焦于社会事件(如新冠疫情、犯罪信息披露)以及国家相关政策在个体消费层面产生的影响。

以上分析了消费者行为领域适合采用二手数据子研究的研究主题,以及研究主题和不同类型二手数据之间的匹配性问题。无论是从数据可及性还是实践现象出发,理解不同二手数据的来源及其用处对于开展实际研究都有着一定的指导作用。研究者可以依据可以获取的二手数据来选定感兴趣的相关研究主题,同时也可以根据既定的主题去寻找对应来源的二手数据。从上述分析也可以看出,可用二手数据类型和研究领域之间存在一定的匹配性。例如,消费者众多的网络行为催生了丰富的网络数据,而这些类型的数据又为消费者行为研究提供了众多有趣的新研究视角。

(三)二手数据可以与哪些研究方法组合搭配?

消费者行为实证研究涉及的研究方法主要包括二手数据法、实验法(包括实验室实验与田野实验)、调查法、访谈法等,这些研究方法各有利弊。多方法研究的优势就在于通过不同方法之间的组合搭配来克服单一研究方法的局限,并实现优势互补的效果。具体而言,多方法的组合搭配可以结合特定的研究领域以及领域主流研究方法的局限性进行选择。消费者行为领域主流的研究方法为实验法,因此常见的多方法组合就有实验法+二手数据、实验法+调查法、实验法+访谈法等。这些多方法组合的共同点都在于借助其他研究方法来弥补实验研究缺乏外部效度的不足。由此可以看出,多方法之间的匹配性问题需要结合各方法的优劣势进行分析。

二手数据虽然具有样本量大、可复制性高、外部效度强等优势,但不可避免也存在数据匹配性弱、内部效度不足、数据处理繁杂、耗时更长等劣势,因此尤其需要和其他研究方法共同使用来达到优势互补。通常可与二手数据匹配的研究方法有实验法、调查法、访谈法三种。接下来我们将就这几种组合搭配给出具体的开展建议。

第一,二手数据+实验法。这一方法组合在营销以及消费者行为领域使用最为广泛。在这种组合形式中,二手数据可以同时起到弥补实验室实验缺乏外部效度和增强研究结论可复制性的作用,因此二手数据的优势可以通过这一组合搭配得到最大化的利用。同样,实验法具备的高内部效度优势可以有效弥补二手数据内部效度不足的劣势。通过这一组合搭配可以实现研究内外效度的优势互补,并提升研究结论的可信度。值得注意的是,有部分消费者行为研究会采用二手数据+实验室实验+田野实验的多方法组合,出现这一现象的原因在于田野实验在一定程度上可以补充研究的外部效度。因此在这一组合形式中二手数据更多地被用来增强研究结果的可复制性。此外,二手数据和田野实验的同时使用,无疑会扩大研究结果在现实世界的普适性,并且最大限度地保证研究的严谨性和吸引力。但是这三种方法的组合在实际操作上也给研究者带来一定的挑战。

第二,二手数据+调查法。在这一组合中,二手数据更多地发挥客观性强、样本量大、外部影响力高等优势。这些优势可以很好地弥补调查法主观性较强以及可靠性低等缺点。另外,调查法对问卷设计要求较高,同一变量的测量可能会因研究者偏好的不同而选用不同的衡量工具,因此使用问卷调查得来的结果无法很好地被其他研究者复制。这一重要不足同样可以借助二手数据进行弥补。值得注意的是,二手数据法和问卷调查法可以同时满足外部效度,但欠缺对研究结论内部效度的支持。因此,这两种方法通常还需要结合其他可以满足内部效度的研究方法(如实验法)共同使用。并且从实际开展角度来说,选用这三种方法组合来开展多方法研究更能说服读者,并且借助更多的研究方法来检验理论观点,会提升研究的严谨性。

第三,二手数据+访谈法。这类方法组合多见于集理论发展与理论检验于一体的研究。研

究者可以通过访谈法来探索先验理论知识,即通过了解受访者对复杂事物、新发事物的想法,来构建新的理论。二手数据则扮演了检验理论观点的作用,并且为研究结论提供现实世界的大样本证据。这类集概念开发和概念检验于一体的研究并不常见,且实际开展时也需要研究者具有较强的理论洞察能力,因此在具备新颖性的同时也存在可行性较低的问题。

总体而言,消费者行为学者在实际开展包含二手数据的多方法研究时,可以遵循三个原则。一是结合研究主题的需求和自身的方法优势来选择可以与二手数据搭配的方法。二是可以基于对二手数据的作用要求来选择适合的方法进行搭配。三是可以根据研究想要具体达到的目的和层次来考虑多方法的数量和具体组合。

(四)二手数据子研究适合出现在文章的什么位置?

在做好前期的研究主题确定和方法组合选择后,就需要考虑二手数据子研究在整体文章中的具体位置。对不同子研究之间的顺序进行合理安排,可以提升文章的系统性和说服力。研究者可以根据想要借助二手数据实现的目的来安排二手数据子研究在文章中的位置。具体而言,可以有以下几种选择:

其一,作为实证部分的第一个子研究。当旨在利用二手数据初步检验理论观点时,需要将二手数据作为第一个子研究。首先通过二手数据验证文章的主要假设能够为随后开展其他分析奠定良好的信心和基础,并且可以先行提供目标变量间存在现实影响关系的关键证据。例如Gupta和Hagtvedt(2021)先借助来自一项品牌资产评估调查的数据集,初步检验了品牌标识设计的宽松(vs.紧凑)对消费者品牌态度的负面影响作用,这为品牌文本标识设计所导致的营销后果提供了初步的市场证据。随后作者就在这一研究结果的基础上设计了更具针对性的实验研究,以探究更加具体的因果关系和内在机制。

其二,作为实证部分的最后一个子研究。当旨在利用二手数据提升文章的高度和价值时,需要将二手数据作为最后一个子研究。在实证研究的最后进行二手数据分析主要是为了增加研究结论的外部效度,进而强化研究的现实意义和实践贡献。例如Shah等(2016)就先通过三项实验研究验证了支付方式(现金vs.卡)会影响消费者的购后行为承诺,并提出支付痛苦与心理价值和承诺会起到连续中介作用。随后,他们借助来自一所大学商学院的校友捐赠档案数据集,检验了捐赠者的支付方式对其未来捐款行为的影响,从而在现实世界中复制了支付方式与行为承诺之间的关系(Shah等,2016)。

其三,出现在实证部分的中间位置。当旨在利用二手数据辅助理论构建或者二手数据本身的吸引力不强时,需要将二手数据子研究放置在文章的中间位置。第二部分中提到,二手数据子研究的一个重要用途就是辅助理论的构建。这类文章通常会使用定性研究方法作为首要方法,以此来挖掘理论观点,二手数据则在其后充当新理论的检验手段(如Spiller和Belogolova,2017)。除此之外,当用于分析的二手数据样本量较小或者对研究的代表性不高时,则需要将其放置在中间位置,既达到采用多方法的目的,也避免拉低文章的吸引力。

其四,出现在实证部分的第一个子研究和最后一个子研究。当有多个可用的二手数据资源时,可以考虑使用首尾呼应的方式来运用二手数据。即第一个子研究先使用二手数据表明文章所讨论的现象在现实生活中真实存在,然后通过其他研究方法来检验具体的因果关系和机制,最后再次通过二手数据研究重复验证结论(如,Galoni等,2020;Rocklage等,2021)。这种研究设计全面地利用了二手数据研究的优势,既凸显了研究的现实意义,又能够夯实研究结论的外部效度。同时,借助多来源二手数据也能够彰显研究结论的普适性。

(五)什么时候不适合采用二手数据子研究?

尽管本文的主要目的在于总结和归纳多方法研究中应如何运用二手数据法,并借此呼吁

广大管理学者去挖掘并使用现实世界中既有的一些数据资源,但是这并不意味着所有的多方法研究就一定要使用二手数据分析方法。首先,当研究主题偏个体认知和态度而非行为变量时,不适合采用二手数据来辅助多方法研究。已有运用二手数据开展子研究的多方法文章多聚焦于一些具体的个体行为,例如消费者的一些在线行为和与社交媒体使用有关的行为。这一现象揭示出二手数据的运用场景有着一定的局限性,即适合来探究具体的个体行为现象,而一些关于个体心理和认知的变量,则无法借助二手数据来观测和研究。其次,当研究重点在于机制的探索而非现象的简单说明时,也不适合开展二手数据子研究。二手数据分析的一大局限性在于只能提供变量间相关关系的初步证据,在多数情况下无法满足对因果推断的说明,以及相应的中介和调节检验。虽然多方法研究的目的在于发挥不同研究设计之间的优势互补,但是基于机制探索的目的而言,应当选择更能检验因果关系的研究方法(例如实验法)来共同开展多方法研究。

五、总结

作为一篇方法研究,本文先总结归纳了二手数据用于多方法研究的具体用途和详细操作流程。随后,本文对一般二手数据法和用于多方法研究的二手数据法之间的区别进行了比较和说明。在此基础上,本文针对具体开展含二手数据的多方法研究时可能会遇到的一些问题进行了说明,并给出了相应的建议。本文仍存在一些不足之处。首先,本文仅以消费者行为领域涉及二手数据的多方法文章为分析样本,对于其他研究领域只能提供一些共通性的开展建议,而无法兼顾特定领域在开展包含二手数据的多方法研究时可能遇到的特定问题。其次,由于2015年前后是学术界对研究可复制性争议比较激烈的时期,因此本文搜索和整理了2015年1月—2022年6月国内外消费者行为权威期刊上刊发的213篇含有二手数据子研究的多方法文章。这就意味着本文总结的操作流程、开展建议也是基于这213篇文章的。实质上,二手数据的使用还可以追溯到更早的时间,对此本文未能关注和分析。最后,消费者行为领域二手数据的使用形式也包括传统的单一二手数据研究和将二手数据作为辅助性数据来源的研究,限于文章篇幅,本文并未涉及对这两类研究的说明。虽然有上述局限,我们仍希望本文能够推动国内管理期刊对二手数据子研究的鼓励和倡导,并且推动国内更多的管理学者在开展多方法研究时关注和使用二手数据。

主要参考文献

- [1]陈晓萍,徐淑英,樊景立.组织与管理研究的实证方法[M].北京:北京大学出版社,2008.
- [2]黄静,文胜雄,童泽林,等.企业家道德不均衡行为对企业家形象评价的探索性研究[J].营销科学学报,2019,15(3):21-35.
- [3]徐岚,赵爽爽,崔楠,等.故事设计模式对消费者品牌态度的影响[J].管理世界,2020,36(10):76-94.
- [4]周元元,胡杨利,张琴,等.时间压力下你想听什么?参照组影响对冲动购买的调节[J].心理学报,2017,49(11):1439-1448.
- [5]朱华伟,张艳艳,龚璇.企业幽默能否化解消费者抱怨:幽默类型与关系范式的匹配[J].心理学报,2017,49(4):526-538.
- [6]Aspara J, Wittkowski K. Sharing-dominant logic? Quantifying the association between consumer intelligence and choice of social access modes[J]. *Journal of Consumer Research*, 2019, 46(2): 201-222.
- [7]Beck J T, Rahinel R, Bleier A. Company worth keeping: Personal control and preferences for brand leaders[J]. *Journal of Consumer Research*, 2020, 46(5): 871-886.
- [8]Berger J, Kim Y D, Meyer R. What makes content engaging? How emotional dynamics shape success[J]. *Journal of Consumer Research*, 2021, 48(2): 235-250.
- [9]Cai F Y, Bagchi R, Gauri D K. Boomerang effects of low price discounts: How low price discounts affect purchase propensity[J]. *Journal of Consumer Research*, 2016, 42(5): 804-816.

- [10]Choe Y, Kan C. Budget depreciation: When budgeting early increases spending[J]. *Journal of Consumer Research*, 2021, 47(6): 937-958.
- [11]Chung J, Johar G V, Li Y Y, et al. Mining consumer minds: Downstream consequences of host motivations for home-sharing platforms[J]. *Journal of Consumer Research*, 2022, 48(5): 817-838.
- [12]Consiglio I, De Angelis M, Costabile M. The effect of social density on word of mouth[J]. *Journal of Consumer Research*, 2018, 45(3): 511-528.
- [13]Davidson A, Theriault D A. How consumer experience is shaped by the political orientation of service providers[J]. *Journal of Consumer Psychology*, 2021, 31(4): 792-800.
- [14]El Hana N, Sabri O. Expressing one's opinions freely on politicians using parodies: Effect of the sources of political parodies (user- vs. media-generated parodies)[J]. *Psychology & Marketing*, 2021, 38(10): 1670-1685.
- [15]Ellram L M, Tate W L. The use of secondary data in purchasing and supply management (P/SM) research[J]. *Journal of Purchasing and Supply Management*, 2016, 22(4): 250-254.
- [16]Esteky S, Wineman J D, Wooten D B. The influence of physical elevation in buildings on risk preferences: Evidence from a pilot and four field studies[J]. *Journal of Consumer Psychology*, 2018, 28(3): 487-494.
- [17]Fan T T, Gao L L, Steinhart Y. The small predicts large effect in crowdfunding[J]. *Journal of Consumer Research*, 2020, 47(4): 544-565.
- [18]Fiske S T. How to publish rigorous experiments in the 21st century[J]. *Journal of Experimental Social Psychology*, 2016, 66: 145-147.
- [19]Galoni C, Carpenter G S, Rao H. Disgusted and afraid: Consumer choices under the threat of contagious disease[J]. *Journal of Consumer Research*, 2020, 47(3): 373-392.
- [20]Gao H C, Zhang Y L, Mittal V. How does local-global identity affect price sensitivity?[J]. *Journal of Marketing*, 2017, 81(3): 62-79.
- [21]Goor D, Keinan A, Ordabayeva N. Status pivoting[J]. *Journal of Consumer Research*, 2021, 47(6): 978-1002.
- [22]Gullo K, Berger J, Etkin J, et al. Does time of day affect variety-seeking?[J]. *Journal of Consumer Research*, 2019, 46(1): 20-35.
- [23]Gupta T, Hagtvedt H. Safe together, vulnerable apart: How interstitial space in text logos impacts brand attitudes in tight versus loose cultures[J]. *Journal of Consumer Research*, 2021, 48(3): 474-491.
- [24]Harding R D, Hildebrand D, Kramer T, et al. The impact of acquisition mode on expected speed of product mastery and subsequent consumer behavior[J]. *Journal of Consumer Research*, 2019, 46(1): 140-158.
- [25]Hodges B T, Chen H P. In the eye of the beholder: The interplay of numeracy and fluency in consumer response to 99-ending prices[J]. *Journal of Consumer Research*, 2022, 48(6): 1050-1072.
- [26]Kalra A, Liu X, Zhang W. The zero bias in target retirement fund choice[J]. *Journal of Consumer Research*, 2020, 47(4): 500-522.
- [27]Kappes H B, Gladstone J J, Hershfield H E. Beliefs about whether spending implies wealth[J]. *Journal of Consumer Research*, 2021, 48(1): 1-21.
- [28]Kettle K L, Trudel R, Blanchard S J, et al. Repayment concentration and consumer motivation to get out of debt[J]. *Journal of Consumer Research*, 2016, 43(3): 460-477.
- [29]Kwan C M C, Dai X C, Wyer R S Jr. Contextual influences on message persuasion: The effect of empty space[J]. *Journal of Consumer Research*, 2017, 44(2): 448-464.
- [30]Lalwani A K, Wang J J. How do consumers' cultural backgrounds and values influence their coupon proneness? A multimethod investigation[J]. *Journal of Consumer Research*, 2019, 45(5): 1037-1050.
- [31]Maner J K. Into the wild: Field research can increase both replicability and real-world impact[J]. *Journal of Experimental Social Psychology*, 2016, 66: 100-106.
- [32]Matz S C, Segalin C, Stillwell D, et al. Predicting the personal appeal of marketing images using computational methods[J]. *Journal of Consumer Psychology*, 2019, 29(3): 370-390.

- [33]Mohan B, Schlager T, Deshpandé R, et al. Consumers avoid buying from firms with higher CEO-to-worker pay ratios[J]. *Journal of Consumer Psychology*, 2018, 28(2): 344-352.
- [34]Mrkva K, Johnson E J, Gächter S, et al. Moderating loss aversion: Loss aversion has moderators, but reports of its death are greatly exaggerated[J]. *Journal of Consumer Psychology*, 2020, 30(3): 407-428.
- [35]Nguyen H T, Chaudhuri M. Making new products go viral and succeed[J]. *International Journal of Research in Marketing*, 2019, 36(1): 39-62.
- [36]Nguyen P, Wang X, Li X, et al. Reviewing experts' restraint from extremes and its impact on service providers[J]. *Journal of Consumer Research*, 2021, 47(5): 654-674.
- [37]Packard G, Berger J. How concrete language shapes customer satisfaction[J]. *Journal of Consumer Research*, 2021, 47(5): 787-806.
- [38]Pancer E, Chandler V, Poole M, et al. How readability shapes social media engagement[J]. *Journal of Consumer Psychology*, 2019, 29(2): 262-270.
- [39]Rocklage M D, Rucker D D, Nordgren L F. Emotionally numb: Expertise dulls consumer experience[J]. *Journal of Consumer Research*, 2021, 48(3): 355-373.
- [40]Schaller J. Booms, busts, and fertility testing the Becker model using gender-specific labor demand[J]. *Journal of Human Resources*, 2016, 51(1): 1-29.
- [41]Schlager T, Mohan B, DeCelles K, et al. Consumers-especially women-avoid buying from firms with higher gender pay gaps[J]. *Journal of Consumer Psychology*, 2021, 31(3): 518-531.
- [42]Schlosser A E, Rikhi R R, Dagogo-Jack S W. The ups and downs of visual orientation: The effects of diagonal orientation on product judgment[J]. *Journal of Consumer Psychology*, 2016, 26(4): 496-509.
- [43]Scott S E, Rozin P, Small D A. Consumers prefer "natural" more for preventatives than for curatives[J]. *Journal of Consumer Research*, 2020, 47(3): 454-471.
- [44]Shah A M, Eisenkraft N, Bettman J R, et al. "Paper or plastic?" How we pay influences post-transaction connection[J]. *Journal of Consumer Research*, 2016, 42(5): 688-708.
- [45]Smith R K, Newman G E, Dhar R. Closer to the creator: Temporal contagion explains the preference for earlier serial numbers[J]. *Journal of Consumer Research*, 2016, 42(5): 653-668.
- [46]Smith R K, vanDellen M R, Ton L A N. Makeup who you are: Self-expression enhances the perceived authenticity and public promotion of beauty work[J]. *Journal of Consumer Research*, 2021, 48(1): 102-122.
- [47]Spiller S A, Belogolova L. On consumer beliefs about quality and taste[J]. *Journal of Consumer Research*, 2017, 43(6): 970-991.
- [48]Tully S M, Sharma E. Context-dependent drivers of discretionary debt decisions: Explaining willingness to borrow for experiential purchases[J]. *Journal of Consumer Research*, 2018, 44(5): 960-973.
- [49]van Laer T, Escalas J E, Ludwig S, et al. What happens in Vegas stays on TripAdvisor? A theory and technique to understand narrativity in consumer reviews[J]. *Journal of Consumer Research*, 2019, 46(2): 267-285.
- [50]Wang X, Lu S J, Li X, et al. Audio mining: The role of vocal tone in persuasion[J]. *Journal of Consumer Research*, 2021a, 48(2): 189-211.
- [51]Wang Y J, Kirmani A, Li X L. Not too far to help: Residential mobility, global identity, and donations to distant beneficiaries[J]. *Journal of Consumer Research*, 2021b, 47(6): 878-889.
- [52]Winterich K P, Carter R E, Barone M J, et al. Tis better to give than receive? How and when gender and residence-based segments predict choice of donation- versus discount-based promotions[J]. *Journal of Consumer Psychology*, 2015, 25(4): 622-634.
- [53]Xu H Y, Bolton L E, Winterich K P. How do consumers react to company moral transgressions? The role of power distance belief and empathy for victims[J]. *Journal of Consumer Research*, 2021, 48(1): 77-101.
- [54]Yang L W, Aggarwal P. No small matter: How company size affects consumer expectations and evaluations[J]. *Journal of Consumer Research*, 2019, 45(6): 1369-1384.

Suggestions on the Application of Secondary Data to a Multi-method Research: Taking the Consumer Behavior Research as an Example

Liu Wumei, Yang Qiaoying, Kang Na

(*School of Management, Lanzhou University, Lanzhou 730000, China*)

Summary: Since the replicability crisis of social science research appeared in academic circles, scholars have explored its causes in many ways. From the perspective of research methods, the inability to replicate the crisis is related to researchers' heavy reliance on a single research method. Therefore, multi-method research began to attract the attention of management scholars. Among the common management research methods, the second-hand data method has advantages in the traceability and authenticity of data sources. Using second-hand data to carry out multi-method research can effectively improve the external validity and replicability of research conclusions. Taking the field of consumer behavior as an example, the bibliometric analysis shows that the use of multiple methods, especially the use of secondary data sub-study, has become a research trend. This important trend indicates important changes in the use of methods in management research in the future. However, management scholars in China do not use secondary data sub-study very frequently. Up to now, no scholars have systematically introduced multi-method research and how to use second-hand data to carry out multi-method research. To address this research gap, this paper uses the literature in the field of consumer behavior as the research sample. It first answers the question of how to implement the secondary data method as multi-method research, and further explains how to use the secondary data to carry out multi-method research in the empirical research of management.

Taking 213 articles (197 in English and 16 in Chinese) published in more than 10 domestic and foreign management journals (e.g., JCR and Management World from January 2015 to June 2022 that carried out secondary data sub-study in the field of consumer behavior) as the research sample, this paper first summarizes the advantages and challenges of using secondary data method for management research. Subsequently, this paper systematically introduces the operation process of secondary data sub-study. Finally, this paper answers a series of questions that management scholars may encounter while carrying out multi-method research with the help of secondary data.

This paper makes several contributions to the literature: First, to the best of our knowledge, it is among the first to systematically introduce the secondary data sub-study in the empirical research of management. Second, it clearly tells the difference between secondary data sub-study and general secondary data research, thereby providing original thoughts on how to conduct secondary data sub-study for management scholars. Third, it introduces the specific application and development suggestions of secondary data sub-study in multi-method research in detail. By doing so, this paper can encourage more domestic scholars to pay attention to and use secondary data sub-study while doing management-related research.

Key words: secondary data; multi-method research; operation process; consumer behavior

(责任编辑:王舒宁)