

doi:10.19920/j.cnki.jmsc.2025.09.007

前作电影口碑多样性对续集电影票房的影响^①

——基于电影评论文本分析的实证研究

王昊男, 孙 琦*, 陈隽旒, 吴 芳

(上海财经大学商学院, 上海 200433)

摘要: 由于电影市场具有高风险与高投资的特性, 制作公司往往通过拍摄续集电影的方式来稳定票房收入。但相关研究对前作电影的影评内容, 特别是内容所呈现出的口碑差异对续集电影的影响关注较少。本研究对 2014 年至 2018 年在中国上映的 150 部续集电影真实数据进行实证分析, 使用基于机器学习的文本分析方法对前作电影影评所呈现的口碑多样性进行测量, 研究发现前作电影口碑多样性对续集电影票房具有显著的正向影响。同时, 本研究验证了信息丰富度在主效应中发挥的机制作用。具体而言, 前作电影口碑数量越多, 前作电影口碑多样性对续集电影票房的积极影响越强。续集电影新闻媒体报道数量和续集电影上映周数则对主效应起到了负向调节作用。文章拓展了电影产业研究的相关理论, 为企业灵活使用在线口碑因素进行营销提供了新的视角, 具有管理借鉴意义与实践价值。

关键词: 口碑多样性; 续集电影; 票房; 信息丰富度

中图分类号: F713.5 **文献标识码:** A **文章编号:** 1007-9807(2025)09-0103-16

0 引 言

自 2002 年电影产业改革以来, 我国电影市场无论是规模还是数量方面都取得了巨大成就。为进一步加大对电影产业的扶持力度, 服务于广大人民群众日益增长的精神文化需求, 《中华人民共和国电影产品促进法》于 2016 年 11 月 7 日正式颁布, 这再次证明了电影产品在发展高质量文化产业中所处的重要地位。但随着电影发行数量激增和流媒体等技术的发展, 新电影的发行风险不断增强^[1]。在电影行业中, 高投资并不意味着高回报。例如, 中美合拍电影《长城》的投资成本超过 10 亿, 但国内累计票房仅有 11.73 亿元, 保守估计亏损约 6 亿元。近年来, 不少公司通过延续市场已有的成功电影资产, 即拍摄续集电影, 来降低票房收入的风险^[2]。2022 全球票房排行榜前十

名的电影中有七部是续集, 且排名前三的《壮志凌云 2》、《阿凡达: 水之道》和《奇异博士 2》都是续集电影, 而在 2000 年前十名中仅有两部为续集电影。基于前作电影的口碑与受众基础, 续集电影往往会收获更多的收入, 尽管它们的评分可能低于前作电影^[3]。针对当前的电影市场环境, 本研究对影响续集电影成功因素的探究具有现实意义与价值。

目前, 已有不少学者对影响续集电影票房收入的因素进行了探究, 例如侯永等^[4]指出前作电影的观影经历会增加消费者的品牌记忆, 从而提高续集电影的票房收入。王海忠等^[5]研究指出续集电影采用文字命名策略相比于数字命名策略, 可以增强消费者的观影意愿。何双男和徐晓庚^[6]认为续集电影是前作电影的延续, 只需要较少的广告宣传费用, 便能实现良好的营销效果。在已

① 收稿日期: 2023-04-26; 修订日期: 2024-09-05。

基金项目: 国家自然科学基金资助项目(72072109; 72272091; 72350610275); 上海财经大学创新团队计划(2022110919); 上海财经大学研究生创新基金资助项目(CXJJ-2024-423)。

通讯作者: 孙 琦(1982—), 男, 北京人, 教授, 博士生导师。Email: sun.qi@mail.shufe.edu.cn

有文献的基础上,进一步研究了前作电影的口碑对续集电影市场表现的影响机制,而影评则是电影口碑信息获取的重要来源,消费者通过阅读影评来获取有价值的信息,以支持其消费决策^[7].与此同时,本研究发现当前学者主要从口碑数量与口碑效价两个角度来探究电影票房的影响因素,有关口碑差异的研究较少^[8-10].随着社交媒体平台覆盖人群规模增大,评论者倾向于生产与现有内容相似或冗余的观点,这给消费者获取有价值的信息造成了困难^[11].在这种情况下,影评内容的多样性一定程度上代表着信息的丰富程度^[12],而消费者更易被信息量丰富的评论所影响^[13].综上所述,本研究证明了前作电影口碑多样性是影响续集电影票房的重要因素,这一发现不仅在理论上具有重要意义,也为电影票房预测和电影产业发展提供了实践指导.

以续集电影为研究对象,本研究基于机器学习文本分析的方法(Word2vec模型和LDA模型)测量前作电影口碑多样性,通过对2014年到2018年在中国大陆上映的258部电影的真实数据进行实证分析,结果表明前作电影的口碑多样性对续集电影票房表现具有积极影响.进一步分析表明,前作电影口碑数量、续集电影新闻媒体报道数量和续集电影上映周数分别产生了正向、负向与负向的调节效果,这进一步验证了信息丰富度在前作电影口碑多样性对续集电影票房的影响中发挥了机制作用.最后,本研究进行了一系列的稳健性检验来确保实证结果的有效性.

本研究的贡献体现在以下三个方面:第一,本研究探究前作电影口碑多样性对续集电影票房的影响补充了目前对口碑差异的研究.当前研究中学者更加关注口碑数量与口碑效价,而对口碑差异的关注度较弱^[14].希望在已有文献的研究基础上增强对影评内容差异的探究,通过文本分析的方法来测量影评内容的多样性以获得更多的非结构化信息.同时,从电影票房的影响因素来看,本研究发现前作电影口碑差异可以对续集电影市场表现产生影响,这为丰富电影市场研究及电影票房预测做出了理论贡献.

第二,本文所研究的口碑差异,重点集中在口碑的多样性(口碑内容丰富度),而非口碑的极端

性(口碑的正负情感倾向).口碑的丰富度在于其中所包含的信息内容的多样性和广度,这些信息可以涵盖影片的各个方面,如情节、表演、制作等.为了对前作电影口碑的多样性进行量化分析,本研究采用了机器学习文本分析的方法,通过对大量口碑评论文本的处理和挖掘,实现了对前作电影口碑的内容丰富度的衡量.同时,验证了信息丰富度所发挥的机制作用.由于对前作电影信息丰富度的影响不同,前作电影的口碑数量、续集电影新闻媒体报道数量与续集电影上映周数对主效应产生了不同的调节效果.上述研究结论丰富了前作电影口碑多样性对电影票房的作用机制,针对不同情景下影响的研究是对相关理论的补充.

第三,从管理与实践意义来看,本文对前作电影口碑与续集电影票房的研究,可以拓展到其他体验型产品与服务上,从而帮助企业在进行品牌延伸时,综合考虑母品牌的口碑内容多样性为新产品销售所带来的影响.最后,本研究顺应互联网发展趋势,强调社交媒体中在线口碑的重要性.这为企业灵活使用社交媒体上用户评论进行网络营销,塑造良好的产品形象,打造高质量文化产业品牌提供了理论依据,具有管理借鉴价值.

1 文献综述

1.1 续集电影票房的影响因素

因为续集电影是在前作电影的基础上发行的,所以续集电影的市场表现会受到一般电影特性和前作电影特征两方面的影响.首先,在一般电影特性方面,Hofmann等^[15]实证表明了明星效应对电影票房具有显著地正向影响,相比于艺术明星,商业明星对票房的影响更大.Sood与Drèze^[16]使用实验与二手数据分析的方式指出,采用数字命名策略的续集电影与前作电影的感知相似程度更高,这导致消费者寻求新刺激的需求没有得到满足,从而降低了观影意愿.严建援等^[17]利用豆瓣和中国电影网的数据构建联立方程模型,研究发现剧透效应对电影票房收入具有显著的负向影响;其次,在前作电影特征的影响方面,Moretti^[18]在使用ACNielsen-EDI公司的票房数据分析时指出,续集电影因为有前作电影的信

息做参考,因此消费者对续集电影的偏好比前作电影的偏好更加精确。侯永等^[4]通过实验研究发现,前作电影的评分会发挥信号作用,直接影响续集电影的票房。Basuroy 和 Chatterjee^[19]对 167 部电影实证研究指出,续集电影与前作电影之间发行的时间间隔越短,续集电影的票房表现更好。综上所述,发现已有研究在探索前作电影对续集电影的影响时,重点强调了口碑的影响作用。因此,基于现有文献的研究基础,本研究重点探究了前作电影口碑多样性对续集电影票房收入的影响。

1.2 在线口碑对消费者的影响

互联网的发展为消费者提供了收集多元化信息和为他人提供消费建议的机会^[20]。产品与服务的在线口碑逐步成为影响消费者购买决策的重要因素。已有研究证明影评会对电影的票房表现产生影响^[21]。近年来,不少学者研究了在线口碑与电影市场表现的关系,重点集中在口碑数量和口碑效价两项指标的影响作用^[8-10]。其中,口碑数量代表公众对产品的兴趣强度,口碑效价表示公众对产品的情感^[22]。例如,Liu^[23]利用雅虎电影网站的数据研究认为在线口碑数量将对电影票房产生积极影响。Moon 等^[3]利用 246 部电影的评分信息回归分析,结论为早期获得专业影评人高评分的电影,即具有高口碑效价的电影,它们的票房收入会更高。而口碑差异反映了消费者对产品或服务观点的不一致性^[24]。近年来许多学者将该指标纳入在线口碑模型中,但相比于口碑数量和口碑效价,当前对口碑差异的研究较少^[14],尤其对口碑差异中所展现信息的丰富度关注不足。Basuroy 等^[25]利用动态联立方程模型对 175 部电影分析指出,是否为续集电影可以作为一种质量信号,其对票房收入的影响会随着影评人意见的一致性减弱,因为影评一致性将会降低消费者不确定性决策的风险。Sun^[12]使用亚马逊和 Barnes & Noble 的图书评级数据建立理论模型,研究认为高评价差异会带来的丰富信息,所以产品口碑效价较低时,高差异反而提高了消费者需求。Wang 等^[26]利用数据分析和行为实验方法在电影等多个领域研究发现高方差可能通过增加消费者感知风险来降低消费者购买意愿,但也可

以通过放大消费者感知的独特性来增加销量。综上所述,当前研究对于口碑应该更一致性还是多样性的观点并不统一。

近年来,随着机器学习技术的发展,众多学者通过对口碑内容的文本信息进行量化处理,进一步探究非结构化口碑信息对消费者的影响。例如,郝媛媛等^[27]使用机器学习的方法测量电影在线评论内容的文本情感,来探索其对评论有用性感知的影响。孙春华等^[28]通过对电影预告片评论的文本情感进行分析,研究发现预告片评论中所包含的积极情感词汇越多,电影票房越高。Kim 等^[10]利用 LDA 主题模型对评论家和消费者影评内容进行文本分析,来探究不同来源的影评的一致性对消费者观影意愿的影响。在以往文献的研究基础上,本研究使用机器学习的方法对前作电影口碑多样性进行测量,并重点关注口碑多样性所呈现的内容丰富度对续集电影票房表现的影响。

2 研究假设与研究框架

2.1 前作电影口碑多样性与续集电影票房

在本研究的情景下,认为前作电影影评的多样性展现了该产品口碑信息的丰富度,从而对消费者观看续集电影的意愿产生了积极影响,具体理由如下:首先,对于以电影为代表的体验型产品来说,与搜索型产品相比,体验型产品的属性更主观、更难以进行评估^[29]。消费者需要更多的产品信息来降低消费的不确定性,所以消费者会参考前作电影的口碑做出续集电影的消费决策,其也希望前作电影影评提供更丰富的产品信息。在此基础上,已有学者指出随着平台覆盖更广泛的人群,后来知识生产者更倾向于生产与现有内容相似或者冗余的内容^[11]。相似的内容降低了信息丰富性,造成了消费者的消极情绪反应^[30]。因此,前作电影口碑多样性越高,信息丰富度越高,越有利于消费者观看续集电影。West^[13]通过实验发现,受试者会增加对信息丰富的反馈的注意力,这表明他们在某种程度上能够感知到这种信息的重要性,即使其无法准确报告这种感知。刘伟和徐鹏涛^[31]进一步实证发现,信息丰富的在线

评论提高了消费者对信息的感知有效性,进而对消费者购买决策产生了积极影响.这是因为消费者通常在做出购买决策时面临信息不完整的情况,其需要了解产品质量、卖家信誉以及替代品的信息,但往往信息的获取成本高昂,消费者决策的不确定性无法完全消除^[32].因此,消费者会通过借鉴他人的评论和观点来降低不确定性风险并制定自己的购买决策^[33].多样化的影评可以为消费者提供更丰富的他人观点和体验,帮助消费者确认自己的品味和选择是否与他人一致^[34].综上所述,本研究推断对于电影产品的消费,消费者需要获取更加丰富、多元化的口碑信息来支持其购买决策.

基于上述理论推导,本研究提出以下假设.

假设 1 前作电影口碑多样性越高,续集电影票房表现越好.

2.2 前作电影口碑数量的调节

从数量的角度来看,口碑数量是在线口碑研究中一个重要方向,已有研究发现口碑数量对电影的票房表现具有积极影响^[35, 36].口碑数量反映了影评中所包含的信息总量,影响消费者对产品信息的了解和购买意愿^[37].口碑数量越多,消费者对产品的讨论越多,这意味着生成了更丰富的产品信息,促进了消费者对产品的了解并做出购买决策^[38].综上所述,前作电影的口碑数量反映了消费者对该电影的关注和信息传播的程度,较多的口碑数量提高了信息丰富程度.口碑数量的增加意味着更多的消费者参与讨论,并且可能导致更广泛的产品认知和信息传递,从而提高了前作电影的信息丰富度.如果本研究假设前作电影口碑的信息丰富度在主效应中发挥着机制作用,那么可以发现前作电影的口碑数量通过增加信息丰富度对主效应起到正向调节作用.

基于上述理论推导,本研究提出以下假设.

假设 2 前作电影口碑数量越多,前作电影口碑多样性对续集电影票房的正向影响越强.

2.3 续集电影新闻媒体报道数量的调节

从信息渠道的角度来看,新闻媒体报道对在线平台用户生成的影评口碑具有替代性作用,续

集电影新闻媒体报道数量的增多使消费者接触的信息更加广泛,进而降低了前作电影口碑信息丰富性的价值,因此对主效应具有负向调节作用.通常情况下,新闻媒体会聘请具备电影观赏经验的评论家来审视新发布的电影,其会提供更加多元化和新颖的评论,以便为公众提供信息参考^[10].尽管新闻媒体报道评论家和其他普通消费者的评价都有助于决策^[39],但不同的信息渠道之间存在替代作用^[40].已有学者指出当消费者通过新闻媒体报道获得大量电影的信息时,其对电影口碑信息的依赖性下降^[17].综上所述,由于信息渠道之间的替代性,随着新闻媒体报道数量的增加,专业影评人提供了更丰富的评论信息与多元化观点,这削弱了前作电影用户口碑信息的价值.本研究认为如果信息丰富度在主效应中发挥了机制作用,那么续集电影新闻媒体报道数量将产生负向调节作用.

基于上述理论推导,本研究提出以下假设.

假设 3 续集电影新闻媒体报道数量越多,前作电影口碑多样性对续集电影票房的正向影响越弱.

2.4 续集电影上映周数的调节

从时间趋势的角度来看,随着续集电影上映时间积累,消费者将获得越来越多的产品信息^[41].由于续集电影信息丰富性提高,前作电影口碑信息对续集电影票房的影响将被削弱.同时,随着产品推出时间增长,消费者的信息获取日益全面,其可能已经从各类信息渠道获取了丰富的产品信息并形成了对产品的看法.总体而言,在产品生命周期的后期阶段,在线口碑对消费者购买行为的影响逐渐减弱^[42].综上所述,本研究认为如果信息丰富度在主效应中发挥了机制作用,那么续集电影上映周数将会对主效应产生负向调节作用.

基于上述理论推导,本研究提出以下假设.

假设 4 续集电影上映周数越多,前作电影口碑多样性对续集电影票房的正向影响越弱.

2.5 研究框架

基于已有文献研究及假设推演,具体的研究框架如图 1 所示.

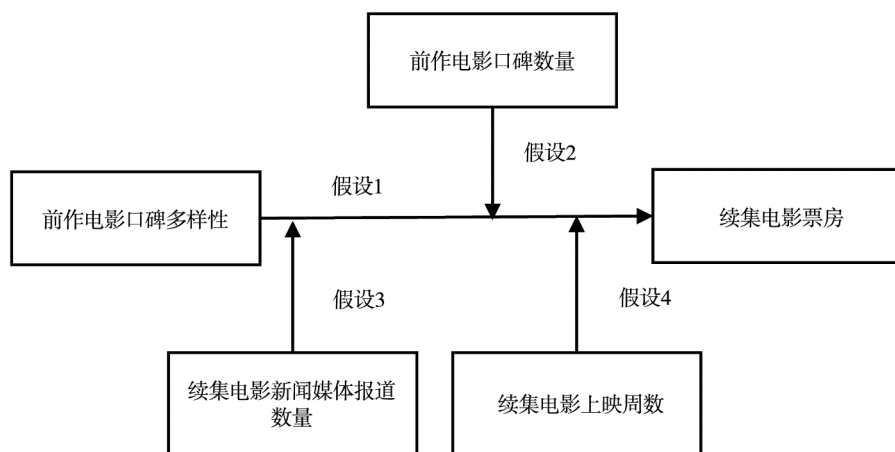


图1 研究框架

Fig. 1 Research framework

3 研究设计

3.1 数据来源

本研究要探究前作电影口碑多样性对续集电影票房的影响。本研究选择2014年至2018年在中国上映的150部续集电影作为研究对象,并收集了其对应的前作电影。因为存在第二部续集电影是第三部续集电影的前作电影的递推关系,最终样本共包含258部前作和续集电影,形成150个匹配对。^② 研究样本中的电影类型包括动作、冒险、喜剧、悬疑、剧情、爱情与动画七大类。本研究从艺恩数据库^③和互联网电影资料库^④收集了每部电影的上映时间、周票房、场均人次、制作成本、演员、获奖成绩和发行公司等特征信息。国内电影的信息来自中国领先娱乐产业信息提供商——艺恩数据,他国电影信息来自互联网电影资料库(IMDb)。IMDb是世界最大的关于电影演员、电影和电影制作的在线数据库,被广泛应用于电影及续集电影的研究^[5]。同时,本研究通过爬虫软件抓取了150部前作电影的全部影评共69 723条。电影影评以及电影评分来自国内著名电影分享与评论网站——豆瓣电影网^⑤。

3.2 变量测量

3.2.1 被解释变量：续集电影票房

本研究的被解释变量为150部续集电影的周票房。电影票房衡量了消费者观影意愿和决策行为,续集电影票房越高,说明消费者有越强的观影意愿^[5]。因为电影票房随着上映时间增加存在明显的“衰退效应”^[43, 44],同时本研究考虑了电影上映周数的调节作用,所以本研究将续集电影的票房按照上映周为单位进行归纳,最后共形成711条观测值。而且续集电影周票房数据存在明显的右偏分布,为缓解异方差性及异常值的影响,本研究对该变量进行了自然对数处理^[17, 45]。

3.2.2 解释变量：前作电影口碑多样性

本研究的解释变量为前作电影的口碑多样性,本研究根据豆瓣电影网150部前作电影的69 723条影评,使用机器学习方法进行文本信息多样性研究。

由于本研究希望衡量影评所包含的内容丰富性,因此在对文本预处理后,本研究使用Word2vec算法将影评文本内容向量化。在此基础上,本研究按照电影ID的分组计算每部电影全部影评的余弦相似度。本研究计算口碑多样性的详细步骤如下:

② 本研究不考虑系列电影中第三部对第一部的影响,仅考虑相邻两部电影之间的关系,例如第一部与第二部,第二部与第三部电影之间的匹配关系,同时本研究在控制变量中对不同部数之间的关系进行了控制,以期排除干扰因素的影响。

③ <https://www.endata.com.cn>.

④ <https://www.imdb.com>.

⑤ <https://www.douban.com>.

首先,本研究对影评内容进行文本清洗,清除特殊符号等干扰内容. 为了避免英文字母对后续步骤的干扰,本研究只保留中文汉字,并将每条影评转换为句子列表;接下来,本研究使用 Jieba 对中文文本进行分词,同时依据中文标准停用词库去除影评内容中的停用词;接下来,本研究使用 Word2vec 对预处理内容进行向量化. Word2vec 是一种用于学习词语之间语义关系的机器学习算法,它基于神经网络将词语转换为高维向量. 通过接收大量文本语料作为输入,该模型为每个词生成一个向量. 这种向量表示可以捕捉词语之间的语义和句法关系,将相似的词语放置在接近的向量空间,而不相似的词语则被分开. 近年来,已有不少学者使用 Word2vec 对文本信息进行处理,以便识别文本数据的潜在语义^[46, 47].

本研究使用中文语料库中训练出的 Word2vec 模型,将每条影评转换为 128 维向量. 该训练模型基于 skip-gram 算法学习词向量,模型语料来自百度百科、搜狐新闻等,最终词典剔除了所有的非中文字符,词典包含约 600 万词汇. 由于训练的 Word2vec 模型对大量丰富的中文语料进行建模,该模型能够较好地捕捉中文语境中的语义和关联信息.

然后,本研究使用余弦相似度度量每部电影所有影评内容之间的文本相似性. 内容的相似程度由向量之间夹角的余弦值表示. 具体计算公式如下

$$S_{lk} = \frac{\mathbf{A}_l \times \mathbf{A}_k}{|\mathbf{A}_l| \times |\mathbf{A}_k|} \quad (1)$$

式中 \mathbf{A}_l 与 \mathbf{A}_k 分别代表第 l 条影评向量与第 k 条影评向量, S_{lk} 表示这两条影评的余弦相似度,衡量了两条影评之间的内容相似程度. 余弦值越大,内容之间的相似性越高.

因为本研究将影评相似度的计算归纳到每部电影中,接下来对每部电影的影评相似度取平均值,具体计算公式如下

$$\bar{S}_i = \frac{1}{N_i} \sum_{n=1}^{N_i} S_{ln} \quad (2)$$

其中 \bar{S}_i 表示第 i 部电影的影评相似度, N_i 表示第 i 部电影中每两条不重复影评匹配的组合个数,即如果一部电影有四条影评,则 N_i 的数值为 6. 至此,本研究已经计算出每部电影的影评相似度.

但本研究想要得到每部电影口碑多样性指标,所以本研究取相似度的负值,为了便于解释,本研究将该值进行归一化.

$$D_i = -\bar{S}_i \quad (3)$$

$$Diversit y_i = \frac{D_i - \min(D_i)}{\max(D_i) - \min(D_i)} \quad (4)$$

最终,本研究计算出了第 i 部电影的影评文本差异,并以此衡量每部前作电影口碑的多样性,最终形成 150 条观测值,前作电影口碑多样性的均值为 0.250.

3.2.3 调节变量: 前作电影口碑数量、续集电影新闻媒体报道数量、续集电影上映周数

前作电影口碑数量为本研究的第一个调节变量,本研究使用从豆瓣电影网爬取的每部电影的影评数量来测量该变量. 根据研究假设中的推理,从数量的角度来看,前作电影评分将会正向调节前作电影口碑多样性对续集电影票房的积极影响.

本研究的第二个调节变量为续集电影新闻媒体报道数量. 本研究从艺恩数据库获取了 150 部续集电影在主流新闻媒体上出现的新闻文章数量. 从信息渠道的角度来看,续集电影新闻媒体报道数量对前作电影信息丰富度产生了替代作用,因此削弱了主效应.

本研究的最后一个调节变量为续集电影上映周数,随着上映周数增加,消费者更容易获得全面的产品信息,从而降低对口碑的依赖,因此认为该变量同样对主效应起到负向调节作用.

3.2.4 控制变量

因为电影票房会受自身特征因素的影响,为了避免其他因素对本研究的干扰,本研究参考已有文献,选取以下控制变量来控制前作电影特征、续集电影的特征、两部电影之间的关系特征^[18, 44, 48, 49]. 首先,从续集电影特征方面,本研究选取新闻媒体报道数量、上映周数、周排片量、电影产地、获奖情况、发行公司、热门档期、制作成本与电影类型作为控制变量. 本研究中热门档期包括了贺岁档、暑假档、五一档与十一档^[17, 50]. 电影分类为动作、冒险、喜剧、悬疑、剧情、爱情六类^[51, 52]. 为符合实际研究情景,本研究在六类的基础上增添了动画类别. 本研究也增加了前作电影特征作为控制变量,例如前作电影的口碑数量、

前作电影豆瓣评分、前作电影总票房、前作电影影评情感、前作电影获奖情况。本研究使用 SnowNLP 库计算每条影评的情感倾向。该指标计算返回结果为积极情感的概率,越接近于 1,则影评的积极情感越强;越接近于 0,则影评的消极情感越强。在上述两类电影特征基础上,本研究也增加了续集关系、命名策略和上映间隔三个变量来对前作电影与续集电影的关系进行控制,进一步排除干扰因素。由于系列电影中不同部数之间的影响关系并不相同,本研究在控制变量中增加了续集关系变量。当续集关系取值为 0 时,前作电影与

续集电影的匹配对为系列电影的第一部与第二部。最后,由于电影市场具有明显的季节性特征,本研究引入了年份固定效应与周固定效应对时间层面上不随个体变化的因素进行控制。为进一步缩小数据之间的绝对差异,缓解异常值的影响,本研究对续集电影周票房、新闻媒体宣传数量、上映周数、周排片量,前作电影的总票房、影评情感与两部电影上映间隔进行取自然对数的线性变换。

上述被解释变量、解释变量、调节变量与控制变量的具体含义及相关信息如表 1 和表 2 所示。

表 1 变量描述
Table 1 Variable descriptions

变量名称	变量含义
被解释变量	
续集电影周票房	第 i 部续集电影第 t 周的票房,取自然对数
解释变量	
前作电影口碑多样性	第 i 部前作电影的影评内容多样性
调节变量	
前作电影口碑数量	第 i 部前作电影的影评数量,取自然对数
续集电影新闻媒体报道数量	第 i 部续集电影主流新闻媒体上出现的文章数量,取自然对数
续集电影上映周数	第 i 部续集电影在第 t 周已经上映的周数,取自然对数
控制变量	
续集电影周排片量	第 i 部续集电影第 t 周在中国大陆的排片数,取自然对数
续集电影产地	第 i 部续集电影是否由外国发售,基准类为中国
续集电影获奖情况	第 i 部续集电影是否获奖,基准类为获奖
续集电影发行公司	第 i 部续集电影是否由中外大型公司发售,基准类为非大型公司发售
续集电影热门档期上映	第 i 部续集电影是否在热门档期上映,基准类为非热门档期
续集电影制作成本	第 i 部续集电影的制作预算,具体分为低预算(3 000 万元及以下)、中预算(3 000 万元至 8 000 万元)、高预算(8 000 万元及以上),基准类为低预算
续集关系	第 i 部续集电影与其前作电影是否为其系列的第二部与第一部,基准类为是
续集电影类型	第 i 部续集电影的类型,具体分为动作、冒险、喜剧、悬疑、剧情、爱情与动画七类
前作电影评分	第 i 部前作电影的豆瓣平均评分
前作电影总票房	第 i 部前作电影的总票房,取自然对数
前作电影影评情感	第 i 部前作电影的正向情感概率,取自然对数
前作电影获奖情况	第 i 部前作电影是否获奖,基准类为获奖
命名策略	第 i 部续集电影是否是数字化命名方式,基准类为非数字化命名方式
上映间隔	第 i 部续集电影与其前作电影之间上映时间的间隔天数,取自然对数

表 2 描述性统计结果

Table 2 Descriptive statistics

变量	均值	标准差	最小值	最大值
续集电影周票房	9 837.903	20 827.221	2	217 126
前作电影口碑多样性	0.250	0.086	0	0.574
前作电影口碑数量	512.712	656.291	3	4 064
续集电影新闻媒体报道数量	4.313	2.578	0	8.640
续集电影上映周数	3.142	1.785	1	8
续集电影周排片量	12.302	16.806	0.007	113.961
续集电影产地	0.468	0.499	0	1
续集电影获奖情况	0.093	0.290	0	1
续集电影发行公司	0.667	0.472	0	1
续集电影热门档期上映	0.516	0.500	0	1
续集关系	0.496	0.500	0	1
前作电影评分	6.557	1.464	2.100	8.700
前作电影总票房	37 332.561	47 203.841	84	24 2431
前作电影影评情感	0.898	0.097	0	1
前作电影获奖情况	0.060	0.239	0	1
命名策略	0.651	0.477	0	1
上映间隔	884.297	517.365	13	2 481

3.3 模型设定

3.3.1 主效应模型

为检验前作电影口碑多样性对续集电影票房的影响,本研究首先构建主效应模型。因为电影的票房表现具有季节性特征,这将导致解释变量与误差项存在相关关系,使回归模型具有内生性问题。所以在回归模型中,本研究在增加控制变量的同时,进一步控制时间效应来控制随时间变动而不随个体变动的不可观测因素。

主效应构建模型如下

$$\ln(sale_{it}) = \beta_0 + \beta_1 Diversit y_i + \theta X_{it} + \mu_t + \varepsilon_{it} \quad (5)$$

其中,被解释变量 $\ln(sale_{it})$ 代表第 i 部续集电影在第 t 周的票房收入,解释变量 $Diversit y_i$ 代表

第 i 部前作电影口碑的多样性, X_{it} 包括表 1 中的 14 个控制变量和前作电影口碑数量、续集电影新闻媒体报道数量和续集电影上映周数三个调节变量。 μ_t 代表续集电影上映的年份与周的时间效应, ε_{it} 为回归模型的残差项。

3.3.2 调节效应模型

进一步为验证前作电影口碑数量、续作电影新闻媒体报道数量与续作电影上映周数对主效应的调节作用,本研究加入调节变量及调节变量与解释变量的交互项。为了避免多重共线性问题,本研究对解释变量和调节变量进行了中心化处理,然后将计算的交互项放入模型中。

调节效应构建模型如下

$$\begin{aligned} \ln(sale_{it}) = & \beta_0 + \beta_1 Diversit y_i + \beta_2 \ln(Num_rating_first_i) \times Diversit y_i + \beta_3 \ln(Index_i) \times Diversit y_i + \\ & \beta_4 \ln(Week_in_theater_{it}) \times Diversity_i + \beta_5 \ln(Num_rating_first_i) + \beta_6 \ln(Index_i) + \\ & \beta_7 \ln(Week_in_theater_{it}) + \theta M_{it} + \mu_t + \varepsilon_{it} \end{aligned} \quad (6)$$

其中 $\ln(Num_rating_first_i)$ 表示第 i 部前作电影的口碑数量, $\ln(Index_i)$ 为第 i 部续集电影的新闻媒体报道数量, $\ln(Week_in_theater_{it})$ 为第 i 部续集电影

在第 t 周的上映周数, M_{it} 包括表 1 中的 14 个控制变量,与主效应模型一致,本研究也对调节效应模型控制了续集电影上映年份与周的时间固定效应。

4 回归结果与讨论

4.1 主效应回归结果

对主效应的回归结果如表3所示,在模型1中本研究只加入前作电影口碑多样性与时间固定效应,本研究发现前作电影口碑多样性正向显著影响续集电影票房($\beta_1=2.881, p<0.05$).

在模型2中,本研究加入了代表续集电影特征、前作电影特征、前作与续集电影关系的一系列控制变量,同样本研究对年份和周的时间效应进行了控制, R^2 由0.15上升到了0.95,模型解释能力大幅度上升.在排除大量可观测变量的干扰后,前作电影口碑多样性仍然对续集电影票房具有显著的积极影响($\beta_1=0.563, p<0.05$),假设1成立.

表3 主效应估计结果
Table 3 Main effects estimation results

变量	被解释变量: 续集电影周票房(ln)	
	模型1	模型2
前作电影口碑多样性	2.881 ** (0.956)	0.563 ** (0.285)
续集电影上映周数(ln)		-0.662 *** (0.051)
续集电影周排片量(ln)		0.976 *** (0.022)
续集电影产地		-0.133 (0.093)
前作电影评分		0.089 *** (0.027)
续集电影新闻媒体报道数量(ln)		0.116 ** (0.049)
续集电影获奖情况		-0.044 (0.139)
续集电影发行公司		0.004 (0.071)
前作电影总票房(ln)		0.164 *** (0.031)
前作电影口碑数量(ln)		0.030 (0.033)
前作电影获奖情况		-0.106 (0.139)
续集电影热门档期上映		0.147 ** (0.073)
前作电影影评情感(ln)		-0.464 * (0.240)
续集关系		-0.062 (0.053)
上映间隔(ln)		0.125 *** (0.040)
命名策略		0.004 (0.054)
常数项	6.645 *** (0.761)	4.050 *** (0.448)
电影类型虚拟变量	否	是
制作成本虚拟变量	否	是
年份固定效应	是	是
周固定效应	是	是
样本量	711	622
R^2	0.150	0.950

注: * 代表 $p<0.1$; ** 代表 $p<0.05$; *** $p<0.01$.

4.2 主效应稳健性检验

4.2.1 降低向量维度

在计算续集电影口碑多样性时,由于预训练的模型是基于大规模中文语料库文本训练生成的,因此本研究基于 Word2vec 模型将影评转换为 128 维向量,以进行后续的文本相似度分析.为进一步保证回归结果的稳健性,本研究使用 UMAP (uniform manifold approximation and projection) 将向量降低为 20 维度^[53]. 回归结果如表 4 中第一列所示,前作电影口碑多样性仍然正向显著影响续集电影周票房($\beta_1 = 0.454, p < 0.05$),证明了本研究研究结果具有一定的稳健性.

4.2.2 LDA 模型

本研究在上述回归中,以 Word2vec 模型来衡量前作电影的口碑多样性,为验证前作电影口碑多样性对续集电影票房产生了正向影响这一结果的稳健性,本研究将使用 LDA (Latent Dirichlet Allocation) 来测量前作电影口碑多样性. LDA 是一种文档主题生成模型,通过非监督机器学习技术,该模型可以用来识别大规模文档集或语料库中潜藏的主题信息,适用于本研究对影评内容多样性的衡量. Kim 等^[10]在探究主题一致性对票房的影响时,使用 LDA 模型来测量评论家和用户评论内容之间的重叠程度. 本研究同样对影评文本进行了清洗与预处理,为避免模型过拟合情况,本研究抽样选取 70% 的影评内容训练了 5 至 30 主题数的 25 个模型. 本研究参考马鸿佳等^[54]的做法,根据一致性与困惑度来获取最佳主题数量. 从图 2 中可以看到,当主题数为 15 时,一致性得分最高,同时主题数在 15 附近时,困惑度得分出现拐点. 尽管随着主题数目的增加,困惑度得分一直在下降,但主题数 15 之后下降幅度显著降低. 因此,本研究选择 15 个主题数.

最后,本研究选择主题数为 15 的 LDA 模型来对影评进行向量化处理,由于 LDA 生成向量维度不同,本研究也使用 UMAP 算法统一将维度降至 20,在此基础上本研究计算每部电影的口碑多样性指标. 从表 4 第 2 列中可以看到,前作电影口碑多样性对续集电影票房产生了显著的正向影响($\beta_1 = 0.388, p < 0.05$),以 LDA 主题模型测算的解释变量回归结果与以 Word2vec 模型测

算的结果显著性基本一致,因此从方法层面上印证了本研究的解释变量测量方法具有稳健性.

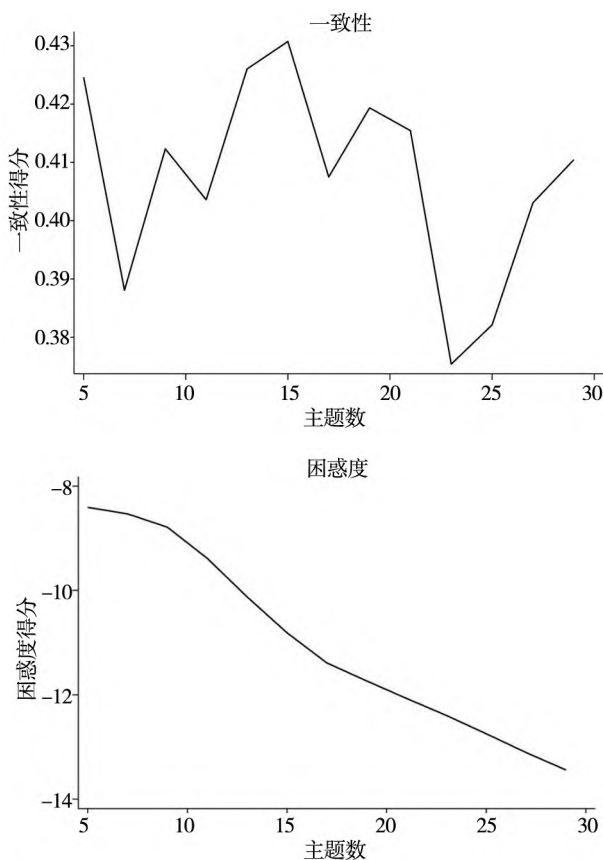


图 2 一致性与困惑度得分

Fig. 2 Consistency and perplexity scores

4.2.3 更改被解释变量

在前面的回归分析中,本研究使用续集电影的票房来衡量续集电影的市场表现,这是因为电影票房代表了观影规模,票房数越高,意味着观看该电影的消费者数量越多. 同时,续集电影评分也可以衡量续集电影的市场表现,电影评分代表了观众满意度^[55, 56]. 在稳健性检验中,本研究使用从豆瓣电影网收集的 150 部续集电影评分作为被解释变量进行回归分析. 豆瓣评分一定程度上还原了消费者群体对一部电影的平均看法. 回归结果如表 4 中第三列所示,前作电影口碑多样性对续集电影评分同样具有显著的积极影响($\beta_1 = 0.934, p < 0.05$). 上述实证结果进一步证明了前作电影口碑多样性有利于续集电影的市场表现,不但对观影规模起到了正向影响,也提高了消费者的观影满意度.

4.2.4 去除票房极端值

由于电影票房数据分布广泛,存在一些极端值,为检验本研究的实证结果是否是由于极端票房所导致的,本研究参考已有研究排除了在95分位数以上和5分位数以下的续集电影

周票房数据^[44]。回归结果表明,在去除极端值的情况下,前作电影口碑多样性仍然正向显著地影响续集电影票房($\beta_1 = 0.572, p < 0.05$),这进一步验证了本研究的回归结果具有稳健性。

表4 稳健性检验结果

Table 4 Robustness test results

	Word2vec(20 维度)	LDA 模型	续作电影评分	去除票房极端值
前作电影口碑多样性	0.454 ** (0.176)	0.388 ** (0.176)	0.934 ** (0.390)	0.572 ** (0.287)
续集电影上映周数(ln)	-0.666 *** (0.051)	-0.664 *** (0.051)	0.254 *** (0.070)	-0.614 *** (0.055)
续集电影周排片量(ln)	0.975 *** (0.022)	0.976 *** (0.022)	0.100 *** (0.030)	0.948 *** (0.024)
续集电影产地	-0.116 (0.092)	-0.118 (0.092)	0.202 (0.128)	-0.074 (0.094)
前作电影评分	0.078 *** (0.027)	0.080 *** (0.027)	0.431 *** (0.037)	0.106 *** (0.027)
续集电影新闻媒体报道数量(ln)	0.090 * (0.049)	0.094 * (0.049)	0.173 *** (0.067)	0.097 ** (0.049)
续集电影获奖情况	-0.057 (0.139)	-0.063 (0.139)	-1.081 *** (0.190)	-0.003 (0.143)
续集电影发行公司	0.061 (0.077)	0.049 (0.077)	0.007 (0.098)	0.004 (0.073)
前作电影总票房(ln)	0.172 *** (0.031)	0.169 *** (0.031)	0.040 *** (0.042)	0.159 *** (0.032)
前作电影口碑数量(ln)	0.029 (0.033)	0.030 (0.033)	0.216 *** (0.045)	0.006 (0.034)
前作电影获奖情况	-0.075 (0.140)	-0.074 (0.141)	0.150 (0.190)	-0.230 (0.144)
续集电影热门档期上映	0.155 ** (0.073)	0.151 ** (0.073)	-0.034 (0.099)	0.199 ** * (0.073)
前作电影影评情感(ln)	-0.425 * (0.240)	-0.436 * (0.240)	-0.910 * ** (0.328)	-0.378 (0.252)
续集关系	-0.087 (0.055)	-0.079 (0.055)	-0.288 * ** (0.073)	-0.055 (0.054)
上映间隔(ln)	0.127 *** (0.040)	0.126 *** (0.040)	-0.090 * (0.055)	0.100 *** (0.042)
命名策略	-0.014 (0.054)	-0.009 (0.054)	-0.011 (0.074)	0.000 5 (0.054)
常数项	4.112 *** (0.440)	4.134 *** (0.440)	3.006 *** (0.613)	4.088 *** (0.473)
电影类型虚拟变量	是	是	是	是
制作成本虚拟变量	是	是	是	是
年份固定效应	是	是	是	是
周固定效应	是	是	是	是
样本量	622	622	620	560
R ²	0.950	0.950	0.765	0.938

注：* 代表 $p < 0.1$ ；** 代表 $p < 0.05$ ；*** $p < 0.01$ 。

4.3 调节效应回归结果

本研究在主效应回归的基础上,加入前作电影口碑数量、续集电影新闻媒体与续集电影上映周数三个调节变量交互项后的回归结果如表 6 所示. 根据回归结果,前作电影口碑数量对主效应起到了正向调节作用($\beta_2 = 0.545, p < 0.05$),而续集电影新闻报道数量和续集电影上映周数则削弱了前作电影口碑多样性对续

集电影票房的正向影响($\beta_3 = -2.327, p < 0.01; \beta_4 = -0.796, p < 0.1$). 假设 2、假设 3 与假设 4 假设成立,调节效应回归结果从数量、信息渠道和时间趋势的角度证明了信息丰富度在前作电影口碑多样性对续集电影票房的影响中发挥了机制作用. 最后,为减少内生性影响,本研究同样对调节效应模型控制了年份与周的时间固定效应.

表 5 调节效应估计结果

Table 5 Moderation effect estimation results

变量	被解释变量:续集电影周票房(ln)
前作电影口碑多样性	0.844 *** (0.295)
前作电影口碑多样性 × 前作电影口碑数量	0.545 ** (0.226)
前作电影口碑多样性 × 续集电影新闻报道数量	-2.327 *** (0.693)
前作电影口碑多样性 × 续集电影上映周数	-0.796 * (0.411)
前作电影口碑数量(ln)	0.031 (0.033)
续集电影新闻报道数量(ln)	0.098 ** (0.048)
续集电影上映周数(ln)	-0.680 *** (0.051)
常数项	3.656 *** (0.455)
控制变量	是
电影类型虚拟变量	是
制作成本虚拟变量	是
年份固定效应	是
周固定效应	是
样本量	622
R^2	0.951

注: * 代表 $p < 0.1$; ** 代表 $p < 0.05$; *** $p < 0.01$.

5 结束语

5.1 研究结论与贡献

随着互联网与社交媒体的快速发展,当今消费者在观影之前,往往会通过互联网来了解电影的基本信息,以此来降低不确定性决策的风险. 因此,诸如豆瓣等电影社群成为消费者了解电影口碑的重要渠道. 电影作为体验型产品,消费者

在正式观影之前,只能通过预告片、简介或演员等基本信息来对电影进行预判,可对于续集电影而言,消费者多了一条了解电影的渠道,那就是续集电影的前作. 消费者会根据前作电影的表现情况来预判续集电影的质量,这意味着前作电影的特征将对续集电影的票房和消费者的观影意愿产生较大的影响. 随着社交媒体的迅速发展,消费者对一部电影质量的了解渠道不再仅仅局限于评分与票房,其更愿意借鉴其他消费者的影评与观点,

这突显出了电影口碑的重要性。以往的研究已经探讨过前作电影的口碑数量、口碑效价等因素的影响,但对口碑差异,尤其是其中所包含的信息丰富度的关注度不足。因此,本研究希望通过基于机器学习的文本分析方法,从影评文本的非结构化信息中计算出每部电影影评的多样性。同时,本研究验证了信息丰富度所发挥的机制作用。最终,研究结果为前作电影的口碑多样性正向显著影响续集电影的票房,前作电影的口碑数量会对该效应产生正向调节作用,续集电影的新闻媒体报道数量和续集电影上映周数则削弱了主效应的积极作用。

本研究的贡献主要体现在以下三个方面。首先,本研究通过探究前作电影口碑多样性对续集电影票房的影响,补充了现有对口碑差异的研究;其次,本研究关注口碑丰富度而非极端性,通过机

器学习文本分析方法量化了前作电影口碑的多样性,并验证了其对续集电影市场表现的影响机制;最后,本研究在管理与实践上具有重要意义,为企业在品牌延伸和网络营销方面提供了理论依据和实践指导。

5.2 研究局限

第一,本研究将情景局限于电影产品本身,而没有进一步向其他体验型产品扩展,缺乏对其他体验型产品市场的实证研究,未来若能获得此类产品的销售数据,可以继续针对上述理论进行研究;第二,本文重点研究了影评内容丰富程度产生的影响,但是对于文本内容的其他特征,例如语句与词语特性等,并没有进一步的探索,因此在对非结构化信息的挖掘方面仍存在不足;第三,本研究基于二手数据进行研究缺乏内部一致性,未来可以通过实验等方式来进行完善。

参考文献:

- [1] Belvaux B, Mencarelli R. Prevision model and empirical test of box office results for sequels[J]. *Journal of Business Research*, 2021, 130: 38–48.
- [2] Eliashberg J, Elberse A, Leenders M. The motion picture industry: Critical issues in practice, current research, and new research directions[J]. *Marketing Science*, 2006, 25(6): 638–661.
- [3] Moon S, Bergey P K, Iacobucci D. Dynamic effects among movie ratings, movie revenues, and viewer satisfaction[J]. *Journal of Marketing*, 2010, 74: 108–121.
- [4] 侯永, 王铁男, 李向阳. 续集电影品牌溢出效应的形成机理: 从信号理论与品牌延伸理论的视角[J]. *管理评论*, 2014, 26(7): 125–137.
Hou Yong, Wang Tienan, Li Xiangyang. Mechanism of movie sequel spillover effect: Perspective of brand extension and signaling theory[J]. *Management Review*, 2014, 26(7): 125–137. (in Chinese)
- [5] 王海忠, 欧阳建颖, 陈宣臻. 续集电影的片名策略及其市场效应研究[J]. *管理科学学报*, 2019, 22(11): 19–32.
Wang Haizhong, Ouyang Jianying, Chen Xuanzhen. Movie sequel's naming strategies and its effect on market performance[J]. *Journal of Management Sciences in China*, 2019, 22(11): 19–32. (in Chinese)
- [6] 何双男, 徐晓庚. 社会资本与市场绩效——对电影票房的新解释[J]. *管理科学*, 2021, 34(5): 41–52.
He Shuangnan, Xu Xiaogeng. Social capital and market performance: New interpretation of film box office[J]. *Journal of Management Science*, 2021, 34(5): 41–52. (in Chinese)
- [7] Basuroy S, Chatterjee S, Ravid S A. How critical are critical reviews? The box office effects of film critics, star power, and budgets[J]. *Journal of Marketing*, 2003, 67(4): 103–117.
- [8] Zhu F, Zhang X. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics[J]. *Journal of Marketing*, 2010, 74: 133–148.
- [9] 张哲宇, 罗彪, 梁樑. 网络环境下的消费者态度转变——基于在线口碑信息框架交互作用的实验研究[J]. *管理科学学报*, 2018, 21(11): 18–34.
Zhang Zheyu, Luo Biao, Liang Liang. Consumers' attitude change in network environment: An experimental study of the interaction of information frame of OWoM[J]. *Journal of Management Sciences in China*, 2018, 21(11): 18–34. (in Chinese)

- [10] Kim E, Ding M, Wang X, et al. Does topic consistency matter? A study of critic and user reviews in the movie industry [J]. *Journal of Marketing*, 2022, 87: 428 – 450.
- [11] Wang G A, Jiao J, Abrahams A S, et al. ExpertRank: A topic-aware expert finding algorithm for online knowledge communities [J]. *Decision Support Systems*, 2013, 54: 1442 – 1451.
- [12] Sun M. How does the variance of product ratings matter? [J]. *Management Science*, 2012, 58(4): 696 – 707.
- [13] West P M. Predicting preferences: An examination of agent learning [J]. *Journal of Consumer Research*, 1996, 23: 68 – 80.
- [14] 龚诗阳, 李倩, 姜博, 等. 叫座却不叫好: 明星效应对网络口碑的影响 [J]. *管理科学*, 2020, 33(2): 114 – 126.
- Gong Shiyang, Li Qian, Jiang Bo, et al. A loud but harsh voice: The impact of celebrity endorsement on online WOM [J]. *Journal of Management Science*, 2020, 33(2): 114 – 126. (in Chinese)
- [15] Hofmann J, Clement M, Völckner F, et al. Empirical generalizations on the impact of stars on the economic success of movies [J]. *International Journal of Research in Marketing*, 2017, 34(2): 442 – 461.
- [16] Sood S, Drèze X. Brand extensions of experiential goods: Movie sequel evaluations [J]. *Journal of Consumer Research*, 2006, 33: 352 – 360.
- [17] 严建援, 李扬, 冯森, 等. 网络口碑中的剧透效应——来自电影市场的证据 [J]. *南开管理评论*, 2020, 23(4): 37 – 48.
- Yan Jianyuan, Li Yang, Feng Miao, et al. Spoiler effect in online word-of-mouth: Evidence from the movie market [J]. *Nankai Business Review*, 2020, 23(4): 37 – 48. (in Chinese)
- [18] Moretti E. Social learning and peer effects in consumption: Evidence from movie sales [J]. *The Review of Economic Studies*, 2011, 78(1): 356 – 393.
- [19] Basuroy S, Chatterjee S. Fast and frequent: Investigating box office revenues of motion picture sequels [J]. *Journal of Business Research*, 2008, 61: 798 – 803.
- [20] Hennig-Thurau T, Gwinner K P, Walsh G, et al. Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? [J]. *Journal of Interactive Marketing*, 2004, 18: 38 – 52.
- [21] Basuroy S, Chatterjee S, Ravid S A. How critical are critical reviews? The box office effects of film critics, star power, and budgets [J]. *Journal of Marketing*, 2003, 67: 103 – 117.
- [22] Babić Rosario A, Sotgiu F D, de Valck K, et al. The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors [J]. *Journal of Marketing Research*, 2016, 53: 297 – 318.
- [23] Liu Y. Word of mouth for movies: Its dynamics and impact on box office revenue [J]. *Journal of Marketing*, 2006, 70: 74 – 89.
- [24] 李信, 陈毅文. 口碑追加形式对购买意向的影响: 口碑方向的调节作用 [J]. *心理学报*, 2016, 48(6): 722 – 732.
- Li Xin, Chen Yiwen. Effects of different eWOM supplementary forms on purchase intention: The moderating role of eWOM valence [J]. *Acta Psychologica Sinica*, 2016, 48(6): 722 – 732. (in Chinese)
- [25] Basuroy S, Desai K K, Talukdar D. An empirical investigation of signaling in the motion picture industry [J]. *Journal of Marketing Research*, 2006, 43: 287 – 295.
- [26] Wang F, Liu X, Fang E. User reviews variance, critic reviews variance, and product sales: An exploration of customer breadth and depth effects [J]. *Journal of Retailing*, 2015, 91: 372 – 389.
- [27] 郝媛媛, 叶强, 李一军. 基于影评数据的在线评论有用性影响因素研究 [J]. *管理科学学报*, 2010, 13(8): 78 – 88 + 96.
- Hao Yuanyuan, Ye Qiang, Li Yijun. Research on online impact factors of customer reviews' usefulness based on movie reviews data [J]. *Journal of Management Sciences in China*, 2010, 13(8): 78 – 88 + 96. (in Chinese)
- [28] 孙春华, 刘业政. 电影预告片在线投放对票房的影响——基于文本情感分析方法 [J]. *中国管理科学*, 2017, 25(10): 151 – 161.
- Sun Chunhua, Liu Yezheng. The effects of online pre-launch movie trailers on box office revenue: Based on text sentiment

- analysis method[J]. Chinese Journal of Management Science, 2017, 25(10): 151–161. (in Chinese)
- [29] Huang P, Lurie N H, Mitra S. Searching for experience on the web: An empirical examination of consumer behavior for search and experience goods[J]. Journal of Marketing, 2009, 73(2): 55–69.
- [30] 廖觅燕, 方佳明, 杨晶晶, 等. 应对行为视角下算法推荐内容相似性对 App 持续使用的影响[J]. 南开管理评论, 2023, 26(3): 178–190.
- Liao Miyan, Fang Jiaming, Yang Jingjing, et al. The effect of recommendation content similarity on sustained app use from the perspective of user coping[J]. Nankai Business Review, 2023, 26(3): 178–190. (in Chinese)
- [31] 刘伟, 徐鹏涛. O2O 电商平台在线点评有用性影响因素的识别研究——以餐饮行业 O2O 模式为例[J]. 中国管理科学, 2016, 24(5): 168–176.
- Liu Wei, Xu Pengtao. A study on influencing factors of the helpfulness of online reviews in O2O of restaurant industry: Based on Tobit model[J]. Chinese Journal of Management Science, 2016, 24(5): 168–176. (in Chinese)
- [32] Stigler G J. The economics of information[J]. Journal of Political Economy, 1961, 69(3): 213–225.
- [33] 杨铭, 祁巍, 闫相斌, 等. 在线商品评论的效用分析研究[J]. 管理科学学报, 2012, 15(5): 65–75.
- Yang Ming, Qi Wei, Yan Xiangbin, et al. Utility analysis for online product reviews[J]. Journal of Management Sciences in China, 2012, 15(5): 65–75. (in Chinese)
- [34] Mudambi S M, Schuff D. Research note: What makes a helpful online review? A study of customer reviews on amazon.com [J]. MIS Quarterly, 2010, 34(1): 185–200.
- [35] Liu Y. Word of mouth for movies: Its dynamics and impact on box office revenue[J]. Journal of Marketing, 2006, 70(3): 74–89.
- [36] Gopinath S, Chintagunta P K, Venkataraman S. Blogs, advertising, and local-market movie box office performance[J]. Management Science, 2013, 59(12): 2635–2654.
- [37] 杜学美, 丁璟妤, 谢志鸿, 等. 在线评论对消费者购买意愿的影响研究[J]. 管理评论, 2016, 28(3): 173–183.
- Du Xuemei, Ding Jingyu, Xie Zhihong, et al. An empirical study on the impact of online reviews on consumers' purchasing intention[J]. Management Review, 2016, 28(3): 173–183. (in Chinese)
- [38] Godes D, Mayzlin D. Using online conversations to study word-of-mouth communication[J]. Marketing Science, 2004, 23(4): 545–560.
- [39] Moon S, Bergey P K, Iacobucci D. Dynamic effects among movie ratings, movie revenues, and viewer satisfaction[J]. Journal of Marketing, 2010, 74(1): 108–121.
- [40] Goldfarb A, Tucker C. Advertising bans and the substitutability of online and offline advertising[J]. Journal of Marketing Research, 2011, 48(2): 207–227.
- [41] Weathers D, Sharma S, Wood S L. Effects of online communication practices on consumer perceptions of performance uncertainty for search and experience goods[J]. Journal of Retailing, 2007, 83(4): 393–401.
- [42] Chen Y, Wang Q, Xie J. Online social interactions: A natural experiment on word of mouth versus observational learning [J]. Journal of Marketing Research, 2011, 48(2): 238–254.
- [43] Einav L. Seasonality in the U. S. motion picture industry[J]. RAND Journal of Economics, 2007, 38: 127–145.
- [44] Gao W, Ji L, Liu Y, et al. Branding cultural products in international markets: A study of Hollywood movies in China[J]. Journal of Marketing, 2020, 84(3): 86–105.
- [45] Elberse A, Eliashberg J. Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures[J]. Marketing Science, 2003, 22: 329–354.
- [46] Wang P, Xiong G, Yang J. Frontiers: Asymmetric effects of recreational cannabis legalization[J]. Marketing Science, 2019, 38(6): 927–936.
- [47] Wang Y, Qin M S, Luo X, et al. Frontiers: How support for black lives matter impacts consumer responses on social media [J]. Marketing Science, 2022, 41(6): 1029–1044.
- [48] Burmester A, Becker J U, Heerde H J v, et al. The impact of pre- and post-launch publicity and advertising on new product sales[J]. International Journal of Research in Marketing, 2015, 32(4): 408–417.
- [49] Wu C, Weinberg C B, Wang Q, et al. Administrative trade barrier: An empirical analysis of exporting Hollywood movies to

- China[J]. International Journal of Research in Marketing, 2022, 39(4): 1253–1274.
- [50] 张伟, 李晓丹, 郭立宏. 不同微博营销渠道对产品销量的影响研究: 品牌自有媒体 VS 第三方媒体的路径对比[J]. 南开管理评论, 2018, 21(2): 43–51.
- Zhang Wei, Li Xiaodan, Guo Lihong. The impact of different microblog marketing channels on product sales: A comparative study of brand-owned media vs third-party media[J]. Nankai Business Review, 2018, 21(2): 43–51. (in Chinese)
- [51] Liu A, Mazumdar T, Li B. Counterfactual decomposition of movie star effects with star selection[J]. Management Science, 2015, 61(7): 1704–1721.
- [52] Luo L, Chen X, Han J L, et al. Dilution and enhancement of celebrity brands through sequential movie releases[J]. Journal of Marketing Research, 2010, 47: 1114–1128.
- [53] 郑晓龙, 白松冉, 曾大军. 面向复杂决策场景的认知图谱构建与分析[J]. 管理世界, 2023, 39(5): 188–204.
- Zheng Xiaolong, Bai Songran, Zeng Dajun. Construction and analysis of cognitive graphs for complex decision-making scenarios[J]. Journal of Management World, 2023, 39(5): 188–204. (in Chinese)
- [54] 马鸿佳, 肖彬, 韩姝婷. 创业领域动态能力研究综述——基于 LDA 主题模型[J]. 南开管理评论, 2025, 28(1): 163–174.
- Ma Hongjia, Xiao Bin, Han Shuting. Review on dynamic capabilities in the entrepreneurship field: Based on the LDA topic model[J]. Nankai Business Review, 2025, 28(1): 163–174. (in Chinese)
- [55] 王分棉, 任倩宜, 周煊. 生态位宽度、观众感知与市场绩效——来自中国电影市场的证据[J]. 中国工业经济, 2021, (11): 155–173.
- Wang Fenmian, Ren Qianyi, Zhou Xuan. Niche width, audiences' perceptions and market performance: Evidence from the Chinese film market[J]. China Industrial Economics, 2021, (11): 155–173. (in Chinese)
- [56] Hsu G. Jacks of all trades and masters of none: Audiences' reactions to spanning genres in feature film production[J]. Administrative Science Quarterly, 2006, 51(3): 420–450.

The influence of word-of-mouth diversity of original movies on the box office of sequels: An empirical study based on the analysis of movie reviews

WANG Hao-nan, SUN Qi*, CHEN Jun-ni, WU Fang

College of Business, Shanghai University of Finance and Economics, Shanghai 200433, China

Abstract: Due to the risky nature of the movie market, production companies increasingly rely on sequel movies to generate box office revenues. Existing research has examined various driving factors behind the success of sequel movies. However, researchers have paid limited attention to the impact of the content and variation in reviews of the parent movies on the box office success of sequel movies. This paper conducts an empirical analysis based on data from 150 sequel movies released in China from 2014 to 2018. Using machine learning-based text analysis methods, we measure the diversity of reviews for the parent movies and find a significant positive effect of review diversity on the box office performance of sequel movies. Additionally, we verify the role of information richness as the underlying mechanism. Specifically, the more reviews a parent movie receives, the stronger the positive impact of review diversity on the box office performance of its sequel movies. The number of media reports and the duration of sequel movies in theaters negatively moderate the main effect. This study extends the research on the movie market and provides a new perspective for firms to consider word-of-mouth factors for marketing sequel movies, thus offering important managerial implications.

Key words: word-of-mouth diversity; movie sequel; box office; information richness