

顾客至上:消费者在线投诉的基本面预测能力研究*

蔡卫星 蒲雨琦 李浩民

摘要:随着数字经济时代的到来,数据要素成为这一新经济形态的核心引擎,催生了更多另类数据。作为一种典型的另类数据,消费者在线投诉信息对企业发展具有至关重要的影响。本文基于千万级别的消费者在线投诉数据,系统研究了消费者在线投诉的信息含量问题。研究发现:首先,消费者在线投诉数据含有与公司基本面相关的信息,对公司未来营业收入增长率具有良好的预测效果。其次,从不同类型的消费者在线投诉的预测能力来看,针对质量的“结果失败”投诉和企业无回应的投诉具有更强的基本面预测效果。再次,我们针对消费者在线投诉的预测能力开展了进一步讨论,发现消费者在线投诉具有行业溢出效应,对同行业其他企业同样具有一定的基本面预测能力;从预测窗口来看,具有一定的持续预测能力;从预测类型来看,对企业财务危机也具有较好的预测能力。此外,消费者在线投诉的基本面预测能力具有显著的异质性,在低信息透明度和消费者业务重要性更高的样本中更加显著。本文的研究结论对于监管机构、上市公司、专业投资者等相关主体具有重要的启发价值。

关键词:消费者投诉 信息含量 另类数据 基本面预测

一、引言

进入21世纪之后,日益显著的买方市场特征使企业之间的竞争在很大程度上集中在消费者身上(范秀成等,2009),通过口碑营销不断提高消费者满意度逐渐成为企业在营销实践中的重要目标(李研等,2018)。尤其是随着社交网络的快速发展,消费者在购买产品和服务前越来越依赖于其他消费者的评价和意见,口碑对消费者的购买决策具有重要的影响(孙瑾、苗盼,2021),具有鼓励性的正面口碑与劝阻性的负面口碑形成口碑效价(潘诺利文等,2012)。负面口碑的形成往往和消费者投诉联系在一起,消费者个人投诉被定义为对产品或服务不满的消费者在传播负面口碑(黄敏学等,2010)。从现有文献来看,以往关于消费者投诉的研究大多数基于问卷或实验的小样本数据(黄敏学等,2010;孙瑾、苗盼,2021),缺乏来自真实世界的大数据,这主要是因为以往大规模的消费者投诉数据很难获取,从而使得相关领域在研究主题和研究内容上受到很大限制。

数字经济时代的到来为相关研究提供了新的契机,越来越多的人类活动转向线上形式,消费者投诉行为也不例外。2018年“3.15”前夕,新浪推出了首个公开性在线投诉平台“黑猫投诉”。相对于“3.15”消费者投诉热线等传统平台,“黑猫投诉”平台具有投诉门槛低、投诉信息全公开、投诉记录可回溯等优点,一经推出就受到了广大消费者关注,迅速成为国内规模最大、信息量最全、影响力最广的消费者在线投诉平台之一。截至2022年9月30日,“黑猫投诉”平台一共收录了大约1020万条消费者在线投诉信息,这为我们开展消费者投诉相关问题研究提供了数据支撑。消费者在线投诉数据属于近年来广泛讨论的另类数据,相对于传统数据而言,具有数据体量更大、时效性更好、真实性更强、颗粒度更细等优点(笪等,2011;弗鲁特等,2017;廖理等,2021)。近年来,另类数据的信息含量问题成为金融、财务、会计等学科最前沿的研究领域之一,受到学术界和实务界越来越多的重视(伯格等,2020)。

收稿时间:2023-5-29;反馈外审意见时间:2023-9-1、2023-12-4;拟录用时间:2024-3-27。

*本项研究得到国家社科基金重点项目“大数据驱动下中国地方金融风险监测与防控研究”(基金号:19AJY027)的资助。感谢匿名评审专家的宝贵意见和建议。文责自负。李浩民为本文通讯作者。

在此背景下,我们基于“黑猫投诉”平台并采用Python爬虫技术方法获得了超过1000万条真实的消费者在线投诉信息,并将其与上市公司数据进行匹配,在此基础上系统研究了消费者在线投诉的信息含量。本文数据具有一些明显的优势:第一,数据真实性。在线投诉数据是来自于真实世界的消费者行为,受各类虚假信息等噪音的影响较小(布道克等,2019),特别是在“黑猫投诉”平台投诉规则设计下可以确保投诉信息源自真实的消费者购买行为。第二,数据实时性。在线投诉数据可以追溯到日度层面实时更新,相较于传统公司财务数据以季度为单位的披露具有明显的时效性(廖理等,2021),从而可以对公司基本面进行更及时的预测。第三,数据相关性。在线投诉直接影响公司口碑进而影响潜在销售量,与公司基本面有直接相关性,是一种更为直接有效的先行指标。

本文将消费者在线投诉信息含量的研究聚焦在检验能否根据消费者在线投诉数据构建领先指标,从而对公司的基本面信息进行有效预测。具体来说,首先,本文检验了消费者在线投诉数据是否包含与基本面有关的信息,为此我们重点考察了消费者在线投诉数量与上市公司未来营业收入增长率之间的关系;其次,在此基础上,我们对消费者在线投诉进行细分,分别是针对质量的“结果失败”投诉和针对消费体验的“过程失败”投诉,以及针对消费者在线投诉是否获得企业回应,考察上述不同类型的消费者在线投诉是否具有不同的信息增量;再次,本文就消费者在线投诉数量的预测能力展开了进一步讨论,包括预测能力的行业溢出效应、预测窗口长度和企业财务危机预测等多维分析,以及不同角度的异质性讨论。

本文的研究具有以下几个方面可能的贡献。第一,本文首次基于大规模的消费者在线投诉数据进行分析,进一步丰富了目前快速增长的另类数据类型。近年来,另类数据相关研究获得了广泛关注,但是由于挖掘成本高昂(伯格等,2020;廖理等,2021),大量较难获取但是可能含有更高信息含量的另类数据有待进一步探究。本研究中使用的消费者在线投诉数据正是这样一类数据,根据我们掌握的文献,国内研究还很少涉及,因此研究消费者在线投诉数据的信息含量和投资价值有助于从数据类型和应用场景等多个视角来丰富另类数据研究文献。

第二,本文研究消费者在线投诉对基本面预测的影响,进一步补充了基本面预测影响因素的研究文献。寻找预测基本面的有效指标是金融领域关注的核心命题之一(法玛、弗伦奇,2015),既有研究发现了一系列潜在的影响因素(侯等,2015)。其中,罗(2007)、罗和洪堡(2008)利用美国航空公司顾客投诉信息,开创性地研究了消费者投诉对基本面预测的影响。本研究利用中国数据和更大范围的消费类公司数据进行了广泛扩展,并从信息透明度和消费者业务重要性等角度探讨了异质性影响。

第三,本文聚焦消费者在线投诉行为,进一步拓展了有关企业口碑的研究领域。在内容层面,消费者投诉反映了公司产品或服务方面的缺陷,在很大程度上导致了企业负面口碑的形成与传播(艾拉德等,2020;施文、赵雪,2023)。以往关于企业口碑的研究大多采用小样本实验或问卷调研的形式(黄敏学等,2010;孙瑾、苗盼,2021),缺乏来自真实世界的大数据。本研究利用来自领先平台“黑猫投诉”的千万级别真实数据来分析消费者投诉行为,有助于为企业口碑的相关研究提供更加丰富的数据支撑。

本文的余下部分结构如下:第二部分是制度背景,回顾中国消费者权益保护制度发展历程,重点介绍“黑猫投诉”平台及其影响;第三部分是理论分析与研究假说,在回顾既有研究的基础上提出本文的理论框架并发展研究假说;第四部分是研究设计,对样本和数据来源、模型设定、变量进行说明;第五部分是对消费者在线投诉的信息含量进行实证检验;第六部分是进一步研究,对消费者在线投诉的预测能力进行拓展性研究和异质性考察;最后总结全文。

二、制度背景

改革开放之后,消费者权益保护体系建设成为中国特色社会主义市场经济制度的重要内容。党的十八大以来,在以人民为中心的发展思想指导下,消费者权益保护工作被提到更为重要的位置。习近平总书记高度重视消费者权益保护,在2021年中央全面深化改革委员会第二十一次会议上着重强调“更好保护消费者权

益”。从整体上看,我国逐渐形成了“有法可依、机制健全、渠道畅通、维权有力”的消费者权益保护整体格局。

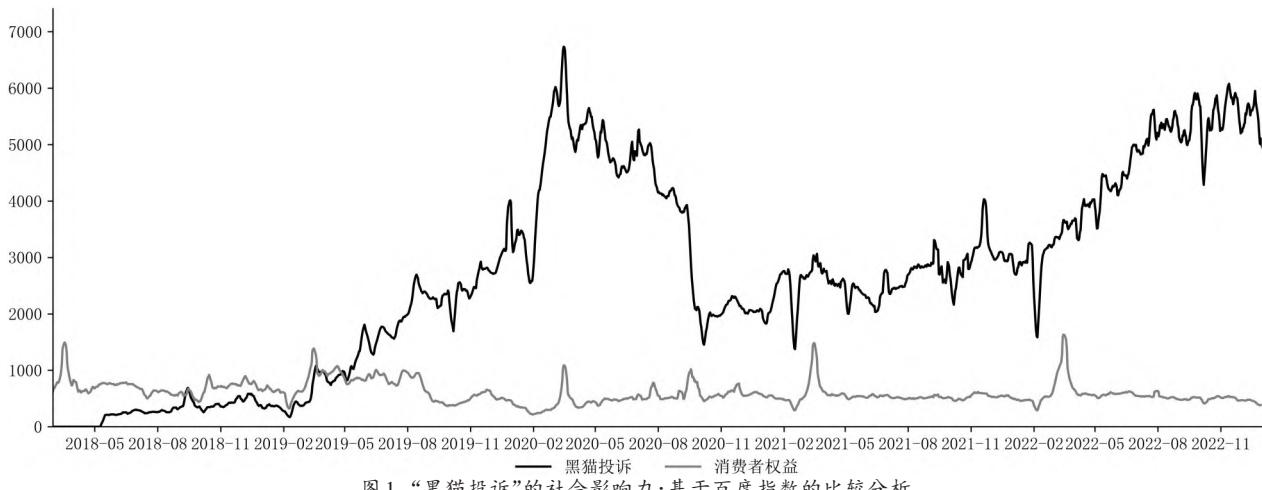
第一,在法律层面,消费者权益保护相关法律不断完善,为消费者保护制度奠定了坚实的法律基础。1993年,《消费者权益保护法》正式颁布,这是新中国历史上第一部专门保护消费者权益的法律,标志着消费者权利保护的法制发展历程进入了一个全新阶段。此后,《消费者权益保护法》经过多次修正,与时俱进地将网络购物等新消费方式纳入其调控范畴,以更为科学有效地保护消费者权益。2024年,《中华人民共和国消费者权益保护法实施条例》正式通过,标志着我国消费者权益保护法律制度体系进一步完善。

第二,在组织层面,“政府监管部门+社会团体组织”相结合共同构成了消费者权益保护工作的组织体系。改革开放后,工商行政管理部门是负责保护消费者合法权益的牵头部门。2018年国务院机构改革之后,国家工商行政管理总局与其他相关部门一起组建了国家市场监督管理总局,消费者权益保护职能随之同时移交,由下设执法稽查局等多个内设机构共同承担消费者权益保护相关职责。除了政府监管部门外,以消费者协会为代表的社会组织也发挥着重要作用。中国消费者协会成立于1984年,履行《中华人民共和国消费者权益保护法》赋予的公益性职责,对商品和服务进行社会监督。截至2021年6月,我国县以上消费者协会已达3138个。

第三,在服务层面,以“12315”为首的官方投诉平台和非官方投诉平台的开发建设为健全消费领域信用体系拓宽了重要渠道。1999年3月15日,国家工商行政管理总局在原国家信息产业部的大力支持下,决定在全国设立专门受理消费者投诉举报的专用电话号码“12315”,并在2017年进一步升级成为全国“12315”互联网平台,标志着我国工商和市场监管部门的消费维权工作跨入“互联网+”时代。与此同时,随着消费者权益保护意识的不断加强,知名新闻网站新浪在2018年“3·15”前夕推出的“黑猫投诉”平台,引起了社会公众的广泛关注,形成了官方投诉平台与非官方投诉平台相互配合、协同发展的整体格局。

作为目前最有影响力的非官方在线投诉平台之一,黑猫投诉平台对消费者的每一次投诉都会生成一个专属投诉页面,内容包括了投诉编号、投诉对象、投诉问题、投诉要求、投诉金额、投诉进度等。具体来说,“黑猫投诉”具有以下优点:第一,信息全公开。“黑猫投诉”平台对所有的消费者投诉记录和企业回复记录全部予以公开,以便社会公众可以随时查看企业和产品的被投诉情况。第二,投诉真实性高。根据“黑猫投诉”规则,消费者需要提供详实资料,包括商品购买记录、商品照片、与企业聊天记录等,在此基础上才能成功发起投诉。此外,“黑猫投诉”平台还会邀请第三方平台进行核实。第三,投诉记录可追溯。“黑猫投诉”平台保留了全部的投诉记录,投诉一经成功发起,消费者就无法删除或撤回。

正是基于上述特点,“黑猫投诉”平台一经推出就获得了社会公众的广泛关注,迅速成为国内最有影响力的领先消费者在线投诉平台。截至2022年9月30日,该平台消费者在线投诉记录超过1000万条,产生了巨大的社会影响。本文利用百度指数考察了“黑猫投诉”平台的社会影响力。从图1可以看出,自“黑猫投诉”平台



推出以来，“黑猫投诉”百度指数呈现出总体上升趋势，并且在2019年6月稳步超过了“消费者权益”百度指数。此外，在每年的3月15日附近，“黑猫投诉”和消费者权益的百度指数均会有一个向上的波动，说明“黑猫投诉”和“消费者权益”一样，在消费者权益保护情绪高涨的节日中受到消费者更多关注。总的来说，“黑猫投诉”平台近年来影响力不断上升，逐渐成为消费者权益保护的重要机制，对监督经营者诚信经营、构造公平消费环境发挥了不可忽视的作用。

三、理论分析与研究假设

近年来，另类数据成为金融科技领域最为关注的热点和前沿问题，学界和业界都围绕另类数据进行了大量研究(笪等,2011;布道克等,2019;艾曼斯等,2022)。其中，与消费者行为有关的另类数据发展十分迅速，日益受到学术研究者、专业投资人、互联网企业等相关主体的重视(弗鲁特等,2017;黄,2018;廖理等,2021;张然等,2022)。消费者作为公司的重要利益相关者(姜、金,2020)，其购买行为和消费支出与公司销售活动密切相关，构成了营业收入和现金流的重要来源(夏皮罗,1983;黄,2018)。现代市场营销理论从消费者行为出发，将消费者的决策过程分为关注、兴趣、搜寻、行动、分享5个阶段(文人,2009)，从而为消费者行为相关的另类数据构建提供了指引。

与消费者行为相关的另类数据类型非常丰富，现有研究中使用较多的另类数据包括但不限于在线销售(朱,2019;廖理等,2021;张然等,2022)、在线搜索(笪等,2011;弗鲁特等,2017)、在线评论(黄,2018)、在线支付(阿加瓦尔,2021)。这些研究表明，与消费者消费行为相关的另类数据在一定程度上能够预测公司的未来业绩(笪等,2011;廖理等,2021)和股票收益(罗,2007;弗鲁特等,2017)。例如，黄(2018)通过研究亚马逊网站上消费者对于上市公司产品的评价发现，消费者群体产生的信息能够更加及时地反映上市公司的产品和现金流信息。朱(2019)通过分析上市公司零售商店停车场的卫星图像数据发现，停车场的汽车数量含有上市公司产品销售状况信息，能够预测上市公司未来的营业收入和股票价格。廖理等(2021)和张然等(2022)研究发现，电商平台的上市公司线上销售额同样包含了上市公司经营业绩信息，能够预测未来股票收益。与此同时，现有研究也指出消费者相关另类数据也存在着一些潜在的不足，例如在线搜索数据虽然在一定程度上可以刻画消费者的潜在需求，但是既难以区分消费者的正面情绪或负面情绪(黄,2018)，也无法反映实际发生的交易活动(廖理等,2021)。从另类数据的角度来看，投诉作为消费者行为的重要方面，一方面源自真实消费经历，另一方面具有明确的情感倾向，相对于其他消费者相关另类数据具有明显的优势。

消费者行为理论认为，投诉作为分享阶段的一种行为，是消费者遭遇不满意体验时寻找途径发泄的一种方式，反映了公司的产品或服务未能达到消费者的期望(黄敏学等,2010)，而当公司长期存在大量未能满足其期望的消费者时，其顾客忠诚度和满意度下降，进而导致经营业绩变差(福内尔等,1996;安德森等,2004)。具体来说，本文认为消费者在线投诉数据可能对上市公司经营业绩具有预测能力，主要源于两个方面。

第一，消费者在线投诉数据提供了有关上市公司当前经营状况的特质信息。“黑猫投诉”平台提供的在线投诉数据以消费者真实消费行为留下的直接购买记录为基础，不会受到管理层主观意识和任意操作信息披露行为的影响，在一定程度上包含了公司当前经营状况的信息。消费者投诉通过反馈公司在产品质量和服务过程等方面的缺陷，揭示出公司在研发创新、产品设计、生产控制、物流管理、客户服务等方面经营状况，特别是基本面有关的负面消息，从而使得市场参与者可以通过消费者在线投诉数据来透视上市公司经营信息。

第二，消费者在线投诉通过口碑效应对上市公司未来经营带来不可忽视的影响。消费者在市场中处于信息劣势方，他们在做出购买决策前需要搜集相关信息来降低感知风险(穆雷,1991;戴克等,2008;郑志刚等,2011;李焰、王琳,2013)。口碑是消费者购买决策时关注的主要因素之一(穆雷,1991;西尔弗曼,1997)，特别是随着数字时代的到来，网络社交平台极大地提升了口碑信息的人际影响力，消费者在购买产品和服务前也

越来越依赖于其他消费者的评价和意见(李研等,2018)。消费者投诉与负面口碑密切相关,消费者投诉越多,负面口碑传播越强,并且具有不可挽回的破坏性(伦纳德,1985)。正如俗话所说的“好事不出门,坏事传千里”,现有研究发现相对于正面口碑而言,消费者投诉形成的负面口碑具有更快的传播速度和更广的覆盖面,从而对企业经营活动造成更明显的影响(施陶斯,1997)。更多的消费者投诉导致更为负面的口碑,进而降低了消费者购买公司产品的可能性,使得企业未来营业收入增速出现明显下降。综合上述讨论,本文提出第一个研究假设。

研究假设1:在其他条件相同的情况下,消费者在线投诉具有良好的基本面预测能力,消费者在线投诉越多,上市公司未来营业收入增长率越低。

如前所述,消费者投诉在某种程度上反映了企业在提供产品和服务方面存在的缺陷(黄敏学等,2010)。在消费者行为理论看来,影响消费者决策的因素既包括了产品本身质量,也包括了购买过程中的体验。一般认为,产品质量是根本,因为消费者会优先考虑企业提供的产品能否满足自己的消费需求;与此同时,购买过程中的消费体验也是消费者购买决策的重要影响因素(约翰斯顿,1995;范秀成、刘建华,2004)。顺着这一思路出发,现有研究对消费者投诉的内容进行了细分,其中结果失败是指消费者最终所获得的产品或服务在质量上没有达到其期望,过程失败是指消费者在获取产品或服务的过程中没有获得良好的经历和体验(比特纳,1990;霍夫曼等,1995)。基于上述分析,本文提出第二个研究假设。

研究假设2:在其他条件相同的情况下,不同投诉类型的消费者在线投诉的基本面预测能力存在差异,其中结果失败的消费者在线投诉越多,上市公司未来营业收入增长率就越低。

企业对消费者投诉的回应是最为常见的补救措施之一(博肖夫,1999)。从本质上讲,企业对消费者投诉的回应是一种为消费者提供情绪补偿的主动行为,能够明显增加消费者的积极情绪(古、叶,2014)。这种回应的积极效应对现有消费者和潜在消费者都会产生显著影响:一方面,对于现有消费者而言,及时有效的回应有利于提高顾客满意度,重构消费者的购买意向(哈特等,1990;史密斯、博尔顿,1998);另一方面,对于潜在消费者而言,及时有效的回应有利于改善消费者认知,降低负面口碑传播的概率(袁兵、郑文哲,2019)。这就意味着,在提供产品或服务已经失败的情况下,如果企业仍采取不回应等消极方式来应对消费者,可能会进一步激发消费者负面情绪和恶化自身口碑,这会降低原有顾客的再购买概率并减少潜在顾客,最终使企业的未来经营业绩变差。基于上述分析本文提出第3个研究假设。

研究假设3:在其他条件相同的情况下,不同处理结果的消费者在线投诉的基本面预测能力存在差异,其中企业无回应的消费者在线投诉越多,上市公司未来营业收入增长率就越低。

四、研究设计

(一)数据来源与样本选择

本文使用的在线投诉数据来自“黑猫投诉”平台。“黑猫投诉”平台于2018年3月14日正式上线,是目前国内规模最大、影响力最广的消费者在线投诉平台之一,并依托新浪实现投诉信息的实时更新。我们使用网络爬虫工具爬取了“黑猫投诉”平台上线以来的全部投诉信息,并使用Python软件对初始非结构化数据进行了处理。本文的样本区间为2018年第二季度到2022年第三季度,一共获得了大约1020万条投诉信息,涉及到17593个市场主体,分布情况参见图2。

根据本文研究需要,我们以上述样本区间存在在线投诉的上市公司作为基准样本,共包括431家A股上市公司。具体来说:首先,在数据处理上,黑猫投诉会明确公布投诉对象,我们以此作为依据确定被投诉企业名称。其次,在数据匹配上,本文根据上市公司及其子公司名称,将其与投诉数据库进行匹配,从而获得了样本期间存在着被投诉的431家A股上市公司,作为本文研究的基准样本。

在此基础上,参考既有文献的处理策略,我们对样本进行了如下筛选:(1)删除金融行业上市公司;(2)剔除上市不足一年的上市公司;(3)删除ST、PT上市公司,参见表1。经过上述处理,本文最终得到了包含359家

上市公司的5688个公司一季度样本。从具体的行业分布来说,排名前五的行业分别是:(1)零售业,一共涉及到31家上市公司,占该行业全部上市公司的比重为27%;(2)电气机械及器材制造业^①,一共涉及到29家上市公司,占该行业全部上市公司的比重为9%;(3)食品制造业,一共涉及到24家上市公司,占该行业全部上市公司的比重为32%;(4)纺织服装及服饰业,一共涉及到20家上市公司,占该行业全部上市公司的比重为46%;(5)酒、饮料和精制茶制造业,一共涉及到20家上市公司,占该行业全部上市公司的比重为40%。

本文涉及到的数据来自于以下几个方面:在线投诉数据由作者团队采用爬虫技术和Python工具自行构造;公司财务数据来自于中国数据研究平台(CNRDS);股票市场数据来自于万得数据库(WIND)。本文对所有连续变量在1%和99%的分位数上进行了缩尾处理。

(二)模型设定与变量定义

参考既有研究的做法(巴本科、森,2016),我们重点关注在线投诉数量是否包含上市公司未来基本面信息。需要说明的是,不少类似研究主要关注的是同期预测(弗鲁特等,2017;廖理等,2021),即在上市公司还未公布季度财务报告时,使用同季度的另类数据对该季度财务报告的公司业绩进行预测。本文在此基础上做出进一步努力,尝试进一步探索前瞻预测,即使用在线投诉数据对下一季度基本面信息进行预测。

基于上述考虑,本文借鉴现有文献的设定(巴本科、森,2016;张然等,2022),使用如下基准计量模型来检验在线投诉的信息含量:

$$SalesGrowth_{i,t} = \beta_0 + \beta_1 CPLT_{i,t-1} + \beta_e Controls_{i,t-1} + FE + \varepsilon_{i,t} \quad (1)$$

在(1)式中,被解释变量营收增长(*SalesGrowth*)是第*i*个上市公司在第*t*个季度的营业收入增长率,这是现有研究中最经常使用的反映上市公司基本面信息的指标之一(廖理等,2021);关键解释变量在线投诉(*CPLT*)是对数化的第*i*个上市公司在第*t-1*个季度的在线投诉次数合计。在这里,我们关注的是在线投诉变量的估计系数,如果 β_1 具有统计意义上的显著性,则意味着在线投诉能预测营收增长,即在线投诉具有信息含量。

在此基础上,本文进一步区分了不同类型的消费者在线投诉。首先,从投诉内容出发,本文进一步识别了针对质量的“结果失败”投诉和针对消费体验的“过程失败”投诉。为此,本文采用机器学习的方法对消费者在线投诉内容进行分析,将与产品和服务本身质量相关的投诉识别为结果失败(*CPLT_O*),将与消费体验相关的投诉识别为过程失败(*CPLT_P*)^②。其次,黑猫投诉平台对每一个消费者在线投诉进行了跟踪,本文根据投诉进度来判断企业回应情况,把处于“已回复”和“投诉已完成”进度的投诉识别成企业有回应的投诉(*CPLT_R*),将其他投诉识别为企业无回应的投诉(*CPLT_NR*)。

参考既有研究的做法(弗鲁特等,2017;廖理等,2021;张然等,2022),本文进一步控制了影响营收增长的潜在因素:包括总市值的对数(*LNSIZE*)、资产收益率(*ROA*)、账面市值比(*BM*)、股票收益率(*RET*)。与此同时,我们还进一步引入了多种固定效应(FE),包括公司固定效应(Firm FE)和季度固定效应(Time FE),以尽可能降低潜在的遗漏变量的影响。本文涉及到的主要变量及其定义参见表2。

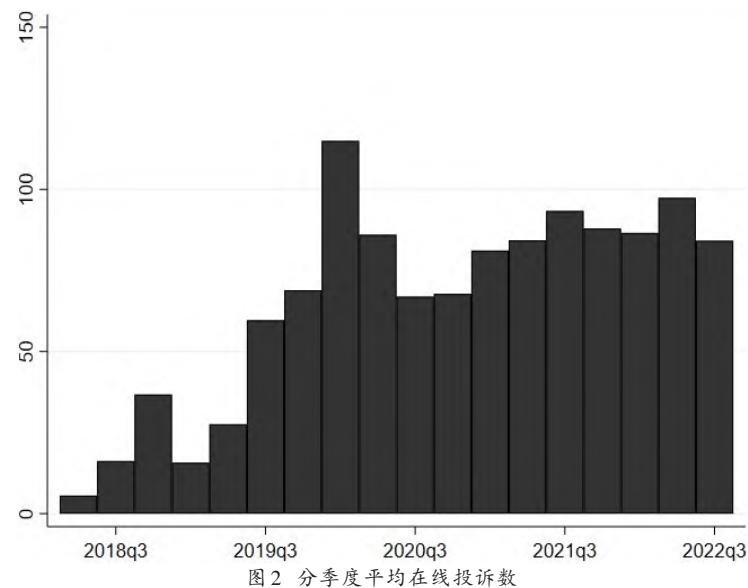


图2 分季度平均在线投诉数

表1 样本选择	
筛选过程	观测值
2018年第二季度~2022年第三季度 A股上市公司	7291
减去:	
金融行业上市公司	1110
上市不足一年的上市公司	322
ST、PT上市公司	171
最终样本	5688

(三)描述性统计

为了考察上市公司消费者在线投诉数量的变化趋势,本文绘制了上市公司的平均投诉总量及各类型平均投诉数量的季度分布图,参见图3。在图3中,针对上市公司的消费者在线投诉平均数量总体呈上升趋势,并且在2019年之后保持较为稳定的态势,各类型在线投诉的变化趋势与总体投诉情况保持一致。与此同时,“结果失败”投诉数量占全部投诉的84.55%,“企业有回应”投诉数量占全部投诉的72.52%。

表3是主要变量的描述性统计结果。从整体上看,上市公司季度营业收入同比增长率存在着较大的个体差异,表现为标准差较大。上市公司在线投诉量均值为17.5093,说明上市公司平均每个季度会收到17.51次在线投诉,相应的对数化指标

的均值为1.1531。与此同时,不同上市公司之间收到的投诉数量会有明显的差异,表现为较大的标准差。本文将其他变量的描述性统计结果和类似研究的描述性统计结果进行对比,并没有发现异常值,从而证实了本文数据是可靠的。

表2 主要变量定义与说明

变量类别	变量	符号	定义与说明
被解释变量 解释变量	营业收入同比增长率	<i>SalesGrowth</i>	$\ln(t\text{季度营业收入}) - \ln(t-4\text{季度营业收入})$
	在线投诉	<i>CPLT</i>	对数化的在线投诉,等于 $\ln(t\text{季度在线投诉数量}+1)$
	“结果失败”投诉	<i>CPLT_O</i>	对数化的“结果失败”在线投诉,等于 $\ln(t\text{季度“结果失败”在线投诉数量}+1)$
	“过程失败”投诉	<i>CPLT_P</i>	对数化的“过程失败”在线投诉,等于 $\ln(t\text{季度“过程失败”在线投诉数量}+1)$
	企业有回应投诉	<i>CPLT_R</i>	对数化的企业有回应在线投诉,等于 $\ln(t\text{季度企业有回应在线投诉数量}+1)$
	企业无回应投诉	<i>CPLT_NR</i>	对数化的企业无回应在线投诉,等于 $\ln(t\text{季度企业无回应在线投诉数量}+1)$
控制变量	公司规模	<i>LNSIZE</i>	$\ln(\text{第}t\text{季度期末总市值})$,总市值由第 <i>t</i> 季度末发行股票总数乘以第 <i>t</i> 季度末收盘价计算得到
	资产收益率	<i>ROA</i>	第 <i>t</i> 季度末净利润/第 <i>t</i> 季度末总资产
	账面市值比	<i>BM</i>	第 <i>t</i> 季度末总资产/第 <i>t</i> 季度末总市值
	股票收益率	<i>RET</i>	第 <i>t</i> 季度末考虑现金红利再投资的个股回报率

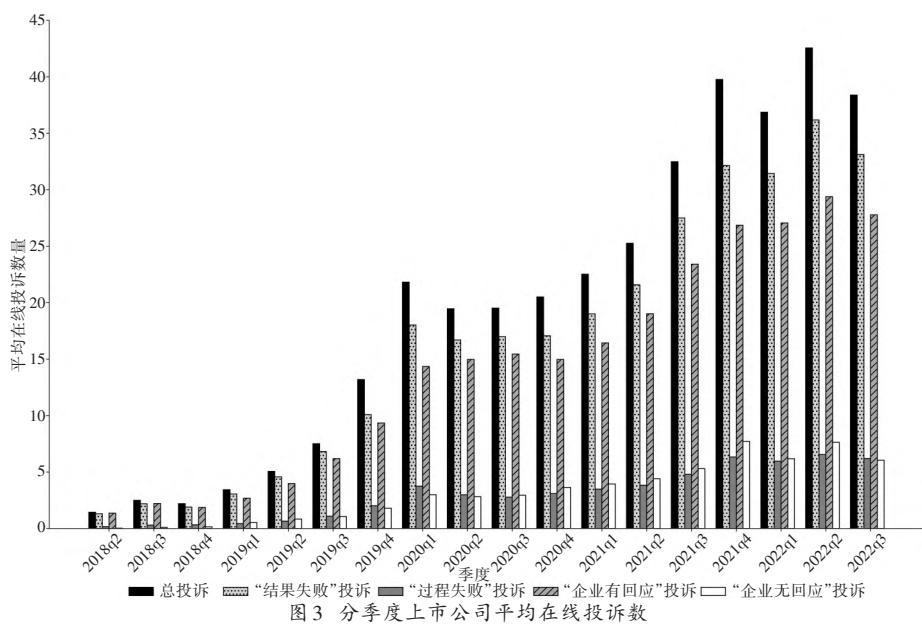


图3 分季度上市公司平均在线投诉数

Variables	Obs	Mean	S.D.	Min	P25	P50	P75	Max
<i>SalesGrowth</i>	5688	-0.0297	0.4443	-1.7965	-0.0892	0.0664	0.1875	0.8462
<i>CPLT</i>	5688	1.1531	1.5398	0.0000	0.0000	0.6931	1.9459	6.1092
<i>LNSIZE</i>	5688	16.3525	1.2727	14.2548	15.383	16.1387	17.1208	20.0745
<i>ROA</i>	5688	0.0314	0.0455	-0.1277	0.0067	0.0233	0.0513	0.1756
<i>BM</i>	5688	0.6539	0.2696	0.1033	0.4506	0.6526	0.8543	1.2103
<i>RET</i>	5688	0.0126	0.2121	-0.3667	-0.128	-0.0258	0.1188	0.8061

五、多元回归分析

(一)基准回归结果

我们使用(1)式检验在线投诉量(*CPLT*)与营业收入增长率(*SalesGrowth*)之间的关系,为在线投诉数据具有信息含量提供经验证据支持,估计结果见表4。在表4中,第(1)栏是没有控制任何其他因素下的单变量回归结果,第(2)~(4)栏是依次加入了公司层面的控制变量以及各种固定效应后的估计结果,第(5)栏报告了各变量的标准化贝塔系数。主要研究结果显示:第一,在没有加入任何控制变量的情况下,在线投诉量(*CPLT*)的估计系数为负,并且在1%的水平下显著,说明上市公司上期在线投诉量越多,则未来营业收入增长率越低,这与本文之前的理论预期是一致的。第二,在逐步加入各种控制变量以及固定效应之后,在线投诉量(*CPLT*)的估计系数和显著性尽管略有下降,但是依然在1%的水平下显著。最后,为了直观比较在线投诉量在基本面预测方面的能

工商管理

力,本文进一步报告了标准化贝塔系数,发现在线投诉量(*CPLT*)的标准化贝塔系数约等于-0.038,与*BM*、*RET*在一个量级,这在经济意义上是一个非常不容忽视的结果^③。综合起来,表4的估计结果表明,在线投诉量对未来营业收入增长率具有显著的预测效果,在线投诉量越高,未来营业收入增长率就越低,验证了研究假设1。

(二)基于投诉特征的进一步讨论

1. 结果失败与过程失败

表4的估计结果证实了消费者在线投诉具有良好的基本面预测能力。在此基础上,本文进一步检验不同投诉类型的消费者在线投诉的基本面预测能力是否存在差异,将前述两种不同类型的消费者在线投诉数量作为关键解释变量代入式(1),回归结果如表5所示。在表5中,*CPLT_O*是对数化处理后的结果失败投诉数量,*CPLT_P*是对数化处理后的过程失败投诉数量。我们发现结果失败投诉(*CPLT_O*)系数为负,且在1%水平下显著,而过程失败投诉(*CPLT_P*)系数虽然也为负,但并不具有统计意义上的显著性,这说明结果失败投诉对上市公司的未来营业收入增长率具有更强的预测作用。其内在逻辑在于,相较于服务态度等影响消费体验的因素,产品和服务质量是影响消费者购买决策的根本因素。

2. 企业有回应与企业无回应

与此同时,本文还进一步讨论了企业采取不同应对方式的消费者在线投诉的基本面预测能力差异,估计结果如表6所示。我们发现企业无回应的消费者在线投诉(*CPLT_NR*)的估计系数为负,且在1%水平下显著,而企业有回应的消费者在线投诉(*CPLT_R*)系数虽然也为负,但并不具有统计意义上的显著性,这说明企业无回应的消费者在线投诉对上市公司的未来营业收入增长率具有更强的预测作用。其内在逻辑可能源于企业积极回应是一种有效的挽救手段,从而降低了消费者在线投诉的负面效应,而对消费者投诉置之不理则进一步恶化了企业口碑。

(三)稳健性检验

为了验证估计结果的可靠性,本文还进一步展开了如下稳健性检验。

1. 替换被解释变量

借鉴已有研究的做法(杰格迪什、里福莱特,2006;廖理等,2021),本文使用标准化未预期营业收入(*SUR*)和标准化未预期盈余(*SUE*)来替代真实营业收入,检验在线投诉量能否预测标准化未预期营业收入和标准化未预期盈余。其中,*SUR*的计算方式如(2)式。其中,*Sales_{i,t}*是公司*i*在季度*t*的营业收入,*Sales_{i,t-4}*是公司*i*在去年同季度的营业收入, $\delta_{i,t}$ 是*t*季度之前4个季度营业收入的标准差;*SUE*的计算方式如(3)式,*Earnings_{i,t}*是公司*i*在季度*t*的营业利润,*Earnings_{i,t-4}*是公司*i*在去年同季度的营业利润, $\varphi_{i,t}$ 是*t*季度之前4个季度营业利润的标准差。接下来,本文以未预期营业收入(*SUR*)或标准化未预期盈余(*SUE*)作为被解释变量,来估计在线投诉量

表4 在线投诉与营业收入增长率

	<i>SalesGrowth</i>				
	(1)	(2)	(3)	(4)	(5)
<i>LCPLT</i>	-0.047*** (0.004)	-0.054*** (0.005)	-0.010*** (0.003)	-0.011*** (0.004)	-0.038
<i>LLNSIZE</i>		0.038*** (0.005)	0.016*** (0.003)	0.119*** (0.017)	0.340
<i>LROA</i>		1.420*** (0.203)	1.938*** (0.138)	1.603*** (0.164)	0.164
<i>LBM</i>		0.001 (0.025)	0.000 (0.018)	0.093* (0.055)	0.056
<i>LRET</i>		-0.087*** (0.027)	0.138*** (0.022)	0.061*** (0.020)	0.029
Constant	0.024*** (0.007)	-0.629*** (0.083)	-0.142*** (0.053)	-2.108*** (0.380)	
Time FE	No	No	Yes	Yes	
Firm FE	No	No	No	Yes	
N	5688	5688	5688	5688	
调整后 R ²	0.026	0.067	0.636	0.708	

注:(1)***、**、*分别表示1%、5%、10%的显著性水平;
(2)括号内为异方差稳健标准误;(3)L_t变量代表第t-1期的变量。下同。

表5 在线投诉的预测能力:

结果失败 VS 过程失败

	<i>SalesGrowth</i>	
	(1)	(2)
<i>LCPLT_O</i>	-0.012*** (0.004)	
<i>LCPLT_P</i>		-0.005 (0.006)
<i>LLNSIZE</i>	0.119*** (0.017)	0.118*** (0.017)
<i>LROA</i>	1.602*** (0.164)	1.605*** (0.164)
<i>LBM</i>	0.093* (0.055)	0.092* (0.055)
<i>LRET</i>	0.061*** (0.020)	0.061*** (0.020)
Constant	-2.110*** (0.380)	-2.111*** (0.380)
Time FE	Yes	Yes
Firm FE	Yes	Yes
N	5688	5688
调整后 R ²	0.708	0.707

表6 在线投诉的预测能力:

企业有回应 VS 企业无回应

	<i>SalesGrowth</i>	
	(1)	(2)
<i>LCPLT_R</i>	-0.004 (0.004)	
<i>LCPLT_NR</i>		-0.016*** (0.006)
<i>LLNSIZE</i>	0.119*** (0.017)	0.117*** (0.017)
<i>LROA</i>	1.603*** (0.164)	1.613*** (0.164)
<i>LBM</i>	0.091* (0.055)	0.094* (0.055)
<i>LRET</i>	0.061*** (0.020)	0.062*** (0.020)
Constant	-2.125*** (0.381)	-2.049*** (0.375)
Time FE	Yes	Yes
Firm FE	Yes	Yes
N	5688	5688
调整后 R ²	0.707	0.708

与未预期信息之间的关系。同时,参考已有文献(弗鲁特等,2017;廖理等,2021),当被解释变量为未预期收入(SUR)时,本文加入滞后一期的标准化未预期营业收入(SUR)作为控制变量;当被解释变量为未预期盈余(SUE)时,本文加入滞后一期的标准化未预期盈余(SUE)和标准化未预期营业收入(SUR)作为控制变量,其他变量定义同(1)式。在这里,我们同样关注在线投诉量(CPLT)的估计系数,预期应该是显著为负。

$$SUR_{i,t} = \frac{Sales_{i,t} - Sales_{i,t-4}}{\delta_{i,t}} \quad (2)$$

$$SUE_{i,t} = \frac{Earnings_{i,t} - Earnings_{i,t-4}}{\varphi_{i,t}} \quad (3)$$

表7分别报告了在线投诉量与未预期营业收入(SUR)和未预期盈余(SUE)的估计结果。其中,第(1)栏和第(4)栏分别是未预期收入(SUR)作为被解释变量和未预期盈余(SUE)作为被解释变量的单变量估计结果,在此基础上,本文逐步加入了控制变量和固定效应,见第(2)~(3)栏和(5)~(6)栏。估计结果表明,不管在何种情形下,在线投诉量(CPLT)的估计系数始终为负,并且至少在5%的水平下显著,意味着作为另类数据的在线投诉量对未预期营业收入(SUR)和未预期盈余(SUE)均具有良好的预测能力。

2. 替换关键解释变量

基于在线投诉数据,本文构造了4个不同形式的解释变量进行稳健性检验,具体如下:首先,考虑到销售规模的影响,本文使用CPLT_Ratio(等于在线投诉数量除以销售收入)作为新的解释变量重新进行估计,结果见表8第(1)栏;其次,我们注意到一部分在线投诉诉求中提出了投诉金额,因此我们构建了在线投诉总金额(CPLT_Indemnity)用来替换(1)式中的在线投诉量,估计结果见表8第(2)栏;最后,考虑到被解释变量为上市公司营业收入的同比增长率,我们分别构建了过去3期的平均投诉量(CPLT_3)和相邻2期投诉数量变化率(CPLT_Change)作为新的核心被解释变量来探讨在线投诉的预测能力^④,估计结果分别见表8的第(3)栏和第(4)栏。我们发现,无论采用何种形式的在线投诉,估计系数始终显著为负,意味着使用不同形式的在线投诉数据并不改变本文研究结论的主要性质。

3. 内生性讨论

尽管在前面的讨论中,本文尽量排除了一系列如销售规模、用户增长等因素的影响,但在这其中仍有可能存在一些潜在的因素同时影响消费者在线投诉和企业的营业收入增长率。考虑到这一点,本文尝试通过工具变量法和双重差分法来缓解这一问题。

第一,本文构造除本公司之外城市内其他公司的平均投诉数量作为工具变量(IV)。一方面,一个地区其他公司的平均投诉数量越多,意味着这个地区的消费者权益保护更为突出,从而可能对该公司的消费者投诉带来影响,能够满足工具变量的相关性;另一方面,

表7 稳健性检验:替换被解释变量

	SUR			SUE		
	(1)	(2)	(3)	(4)	(5)	(6)
L.CPLT	-0.427*** (0.046)	-0.421*** (0.049)	-0.160** (0.065)	-0.275*** (0.039)	-0.276*** (0.040)	-0.204*** (0.075)
L.SUR		0.173*** (0.015)	0.172*** (0.014)		-0.036*** (0.012)	0.021 (0.015)
L.SUE					0.260*** (0.020)	0.153*** (0.020)
LLNSIZE		0.334*** (0.054)	0.033 (0.198)		0.245*** (0.045)	-1.426*** (0.233)
L.ROA		7.110*** (1.668)	3.674** (1.859)		4.003*** (1.367)	2.324 (2.074)
L.BM		-0.587** (0.241)	-1.353** (0.571)		-0.898*** (0.212)	-4.677*** (0.717)
L.RET		-1.340*** (0.286)	0.547* (0.297)		1.523*** (0.279)	2.345*** (0.321)
Constant	1.044*** (0.076)	-4.398*** (0.893)	3.397 (4.285)	0.411*** (0.074)	-3.181*** (0.739)	33.713*** (4.985)
Time FE	No	No	Yes	No	No	Yes
Firm FE	No	No	Yes	No	No	Yes
N	5596	5484	5484	5610	5498	5498
调整后R ²	0.019	0.080	0.468	0.009	0.113	0.205

表8 稳健性检验:替换解释变量

	SalesGrowth			
	(1)	(2)	(3)	(4)
L.CPLT_Ratio	-2.110*** (0.794)			
L.CPLT_Indemnity		-0.004*** (0.001)		
L.CPLT_3			-0.017*** (0.005)	
L.CPLT_Change				-0.007** (0.003)
LLNSIZE	0.110*** (0.017)	0.119*** (0.017)	0.114*** (0.019)	0.118*** (0.017)
L.ROA	1.602*** (0.163)	1.608*** (0.164)	1.571*** (0.170)	1.612*** (0.164)
L.BM	0.070 (0.054)	0.095* (0.054)	0.048 (0.056)	0.092* (0.055)
L.RET	0.063*** (0.020)	0.061*** (0.020)	0.040* (0.021)	0.061*** (0.020)
Constant	-1.938*** (0.376)	-2.088*** (0.378)	-2.028*** (0.400)	-2.111*** (0.380)
Time FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
N	5688	5688	5102	5688
调整后R ²	0.708	0.708	0.728	0.708

工商管理

这一工具变量显然不太可能直接影响到公司本身的营业收入增长率,从而满足工具变量的外生性要求。工具变量的估计结果见表9,其中,第(1)列是第一阶段的回归结果,第(2)列是第二阶段的估计结果。从第(1)列估计结果来看,工具变量的估计系数显著为正,这与我们的直觉是一致的,与此同时第一阶段F统计量大于10,不存在着弱工具变量的问题。从第(2)列的估计结果来看,在考虑潜在内生性问题之后,在线投诉数量(CPLT)的估计系数依然显著为负,验证了基准结果的稳健性。

第二,本文采用了多时点双重差分法重新识别在线投诉对上市公司未来营业收入的预测能力。具体地,本文采用样本期所有A股上市公司为研究样本,构造是否发生投诉的变量(*Treat*):第一次发生投诉之后($t+1$ 期及之后)的样本记为*Treat*=1,否则*Treat*=0。回归结果如表10所示。可以看出,无论是单变量回归结果还是逐步加入了控制变量和固定效应的回归结果,核心解释变量系数(*Treat*)始终显著为负,说明相较于没有发生投诉的上市公司来说,发生投诉的上市公司未来有着更低的营业收入增长率,进一步证明了本文研究结论主要性质的可靠性。

六、进一步讨论

(一)消费者在线投诉预测能力:多维度拓展

1. 在线投诉的预测能力:从公司到行业

前述研究已经表明,在公司层面的在线投诉具有良好的基本面预测能力。接下来,本文进一步探讨,基于上市公司汇总得到的行业在线投诉数据是否对同行业中的其他公司具有一定的基本面预测能力。在这里,我们认为存在着两种潜在的竞争性假说:一方面,其他公司可能遭受“池鱼之殃”,行业投诉量上升会使得整个行业都遭遇“灭顶之灾”(温弗里、麦克卢斯基,2005;李想、石磊,2014),此时行业在线投诉量对其他公司营业收入增长具有显著的负向影响。另一方面,其他公司可能“独善其身”,承接那些受到较多投诉公司的需求转移(佩蒂等,1983;郭国庆等,2007),此时行业在线投诉量对其他公司营业收入增长具有显著的正向影响。

基于上述讨论,本文构建了行业在线投诉量指标(*CPLT_Industry*),根据前述上市公司在线投诉量进行加总得到。在研究样本上,我们使用这些行业中没有被投诉的其他上市公司作为基准样本,来检验前述行业在线投诉数据的基本面预测能力,并尝试回答究竟是“池鱼之殃”还是“独善其身”。估计结果参见表11。在表11中,第(1)~(3)栏分别是单变量、加入控制变量、加入固定效应后的估计结果。估计结果表明:首先,不管是何种情形下,行业在线投诉量指标(*CPLT_Industry*)的估计系数始终为负,并且均在1%的水平下显著,意味着作为另类数据的行业在线投诉量确实对行业没有投诉的其他上市公司营业收入增长率具有良好的预测能力。再次,行业在线投诉量指标(*CPLT_Industry*)显著为负,意味着这些行业往往是“一荣俱荣、一损俱损”,行业负面新闻的增加会给全行业企业都带来系统性风险。最后,从经济意义来看,行业在线投诉量的估计系数

表9 稳健性检验:工具变量估计

	<i>L.CPLT</i>	<i>SalesGrowth</i>
	(1)	(2)
<i>Instrumented-L.CPLT</i>		-0.217** (0.098)
<i>IV:L.CPLT_City</i>	0.038*** (0.010)	
<i>LLNSIZE</i>	0.083* (0.043)	0.135*** (0.017)
<i>LROA</i>	-0.102 (0.391)	1.573*** (0.144)
<i>LBM</i>	0.221* (0.131)	0.133** (0.052)
<i>LRET</i>	-0.038 (0.063)	0.051** (0.024)
Time FE	Yes	Yes
Firm FE	Yes	Yes
F	14.02	
N	5688	5688

表10 稳健性检验:替换回归模型

	<i>SalesGrowth</i>		
	(1)	(2)	(3)
<i>Treat</i>	-0.090*** (0.008)	-0.117*** (0.008)	-0.032*** (0.009)
<i>LLNSIZE</i>		0.025*** (0.002)	0.111*** (0.007)
<i>LROA</i>		1.655*** (0.069)	1.941*** (0.057)
<i>LBM</i>		-0.009 (0.008)	0.081*** (0.020)
<i>LRET</i>		-0.070*** (0.010)	0.063*** (0.007)
Constant	0.004** (0.002)	-0.425*** (0.031)	-1.659*** (0.122)
Time FE	No	No	Yes
Firm FE	No	No	Yes
N	62781	62781	62781
调整后 R ²	0.002	0.027	0.637

表11 行业层面在线投诉与营业收入增长率

	<i>SalesGrowth</i>		
	(1)	(2)	(3)
<i>LCPLT_Industry</i>	-0.027*** (0.001)	-0.028*** (0.001)	-0.005*** (0.001)
<i>LLNSIZE</i>		0.026*** (0.002)	0.111*** (0.006)
<i>LROA</i>		1.643*** (0.073)	1.816*** (0.048)
<i>LBM</i>		-0.037*** (0.009)	0.088*** (0.016)
<i>LRET</i>		-0.063*** (0.010)	0.068*** (0.007)
Constant	0.084*** (0.003)	-0.336*** (0.035)	-1.823*** (0.095)
Time FE	NO	NO	YES
Firm FE	NO	NO	YES
N	52910	52910	52776
调整后 R ²	0.020	0.047	0.635

要明显小于公司层面在线投诉量,意味着行业层面在线投诉数据的预测能力比公司层面数据的预测能力弱。

2. 在线投诉的预测能力:从短期到长期

本文的基准结果表明在线投诉对于上市公司下一季度营业收入增长具有较好的预测能力。在此基础上,我们进一步关心这种预测能力随时间变化会出现何种趋势,是否具有一定的持续预测能力。为了检验预测有效期,本文进一步引入未来各期的营业收入增长率分别作为被解释变量,以检验在线投诉的预测能力是如何随着时间变化的,估计结果见表12。

在表12中,第(1)栏是基准结果(t),作为估计的参照系。第(2)~(5)栏分别是使用未来各期的营业收入增长率作为被解释变量的估计结果。从表12的估计结果可以发现:第一,从系数数值来看,一直到 $t+3$ 期,在线投诉量($CPLT$)的估计系数始终为负,并且在数值上基本没有明显变化,保持在相对稳定的水平;第二,从统计显著性来看,在线投诉量($CPLT$)估计系数的显著性水平随着预测期限的拓展逐渐呈现下降趋势,到 $t+4$ 期则完全不具备统计意义上的显著性。综合起来,表12的估计结果表明,在线投诉的预测能力至少在未来4个季度内仍然保持较好的水准,但预测效力会随着时间的推移而逐步减弱。

3. 在线投诉的预测能力:从营收增长到财务危机

现有研究表明(奥特曼,1968;姜付秀等,2009;余明桂等,2013),上市公司营业收入下降进而经营状况恶化很可能会导致净现金流的减少,从而导致发生财务危机的可能性变大。因此,我们将视角进一步延伸到财务危机,考察在线投诉对上市公司经营风险是否具有预测能力。参考已有文献做法(姜付秀等,2009),本文根据Z-Score^⑤构造了财务危机变量($RISK$),具体定义是,以1.8为临界值,将Z-Score低于1.8的样本视为经营风险较高,设定为 $RISK=1$,否则 $RISK=0$ 。在此基础上,本文用财务危机变量($RISK$)替代(1)式中的营业收入增长率,重新进行检验,估计结果如表13所示。

在表13中,第(1)~(3)栏分别是单变量、加入控制变量、加入固定效应后的估计结果。从显著性来看,在线投诉($CPLT$)系数为正,且均在1%水平下显著,说明上市公司在线投诉数量越多,未来发生财务危机的概率越大,表明在线投诉量($CPLT$)对上市公司发生财务危机概率具有较好的预测能力。

(二)消费者在线投诉预测能力:锦上添花还是雪中送炭

在这里,本文进一步尝试探索这种预测能力在不同上市公司之间的异质性。具体来说,考虑到公司信息透明度较低的上市公司的基本面预测更加困难(巴伦等,1999;方军雄,2007),从而在某种程度上也更有价值(方、佩雷斯,2009;罗进辉、杜兴强,2014)。本文参考既有研究的做法构造反映公司信息透明度的指标,在此基础上重点关注在线投诉量这一另类数据的预测能力在不同信息透明度的公司中的差异。对于这一问题的检验,有助于我们更好地理解在线投诉量在公司基本面预测中的价值。具体来说,我们着重关心在线投诉量这一另类数据的预测能力是“锦上添花”还是“雪中送炭”:如果在线投诉量的预测能力在高透明度的上市公司中更加显著,在某种程度上意味着这一另类数据更多的是“锦上添花”;与此相反,如果在线投诉量的预测能力

表12 在线投诉预测能力:从短期到长期

	$SalesGrowth_{t+1}$ (1)	$SalesGrowth_{t+2}$ (2)	$SalesGrowth_{t+3}$ (3)	$SalesGrowth_{t+4}$ (4)	$SalesGrowth_{t+5}$ (5)
$LCPLT$	-0.011*** (0.004)	-0.010** (0.005)	-0.011** (0.005)	-0.010* (0.005)	-0.002 (0.006)
$LLNSIZE$	0.119*** (0.017)	0.055*** (0.019)	-0.014 (0.020)	-0.085*** (0.022)	-0.112*** (0.023)
$LROA$	1.603*** (0.164)	0.709*** (0.199)	-0.147 (0.201)	-1.140*** (0.217)	-1.403*** (0.228)
LBM	0.093* (0.055)	-0.066 (0.057)	-0.131** (0.058)	-0.282*** (0.064)	-0.269*** (0.062)
$LRET$	0.061*** (0.020)	0.109*** (0.021)	0.168*** (0.022)	0.075*** (0.022)	0.031 (0.023)
Constant	-2.108*** (0.380)	-0.736* (0.410)	0.631 (0.437)	2.161*** (0.476)	2.586*** (0.487)
Time FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
N	5688	5342	5002	4662	4326
调整后 R ²	0.708	0.703	0.706	0.719	0.727

表13 在线投诉预测能力:从营收增长到财务危机

	$RISK$		
	(1)	(2)	(3)
$LCPLT$	0.052*** (0.004)	0.018*** (0.003)	0.014*** (0.004)
$LLNSIZE$		0.093*** (0.004)	-0.060*** (0.016)
$LROA$		-2.916*** (0.114)	-0.791*** (0.139)
LBM		0.828*** (0.018)	0.202*** (0.048)
$LRET$		0.045** (0.021)	-0.023 (0.017)
Constant	0.222*** (0.007)	-1.717*** (0.062)	1.919*** (0.335)
Time FE	No	No	Yes
Firm FE	No	No	Yes
N	5688	5688	5688
调整后 R ²	0.031	0.428	0.749

工商管理

在低透明度的上市公司中更加显著，在某种程度上意味着这一另类数据更多的是“雪中送炭”。比较而言，我们更希望另类数据扮演“雪中送炭”的作用，这意味着另类数据在基本面预测上能够提供更多的增量信息，具有更高的应用价值。

参考既有文献，本文使用证券分析师关注度来衡量上市公司信息透明度，一般而言关注上市公司的证券分析师越多，信息透明度就越好（霍普，2003；朱红军等，2007；潘越等，2011）。具体地，本文定义了低分析师关注（*LowAnalyst*）变量，当某一年上市公司的跟踪证券分析师人数低于其所在行业的中位数时，*LowAnalyst* 取值为 1，否则取值为 0。进一步地，本文引入在线投诉量与低分析师关注的交互项（*CPLT×LowAnalyst*），以检验在线投诉量的基本面预测能力在不同证券分析师关注度样本中的差异，估计结果见表 14。

在表 14 中，第(1)栏是没有添加控制变量的估计结果，第(2)栏是添加控制变量之后的估计结果。本文发现交互项（*CPLT×LowAnalyst*）的估计系数均在 1% 的水平下显著为负，说明在线投诉量对低证券分析师关注度样本的基本面预测能力更强。证券分析师作为资本市场的重要信息中介，在发掘和传递上市公司信息方面发挥了重要作用，而本文的估计结果表明在线投诉在某种程度上与证券分析师扮演了“互补”作用。在那些证券分析师关注比较少的上市公司中，作为另类数据的在线投诉量更有助于基本面预测，这一结论验证了在线投诉量“雪中送炭”的价值。

（三）消费者在线投诉预测能力：消费者业务重要性的影响

现有研究认为，公司对消费者的依赖程度是影响口碑效应大小的重要因素，公司对消费者的依赖程度越高，就越容易受到口碑效应的影响（肖静华等，2015）。因此，我们预计消费者在线投诉的基本面预测能力在那些消费者业务更重要的样本中更显著。考虑到现有数据中缺乏对于消费者业务占比的分析，本文基于公司文本数据创新性地构造了反映消费者业务重要性的指标。具体来说，本文计算了公司年报管理层讨论中出现的与消费者相关词数占总词数的比例，来近似地刻画上市公司对消费者的依赖程度^⑥。在此基础上，本文构造了高消费者依赖度（*High toC*）变量，当某一年消费者相关词数占总词数比例高于其所在行业的中位数时，*High toC* 取值为 1，否则取值为 0。进一步地，本文引入在线投诉量与高消费者依赖度的交互项（*CPLT×High toC*），以检验在线投诉量的基本面预测能力在不同消费者依赖度样本中的差异，估计结果见表 15。在表 15 中，第(1)栏是没有添加控制变量的估计结果，第(2)栏是添加控制变量之后的估计结果。本文发现交互项（*CPLT×High toC*）的估计系数均在 5% 的水平下显著为负，这说明在线投诉量预测基本面的能力在不同消费者依赖度的上市公司中存在差异。总的来说，其在高消费者业务依赖度的上市公司中有着更强的预测基本面的能力。

七、主要结论

顾客至上，市场经济条件下的消费者口碑对企业生存至关重要。本文首次收集了超过 1000 万条消费者在线投诉数据，系统研究了消费者在线投诉是否包含了公司基本面信息。研究结果表明：首先，消费者在线投诉数据含有与公司基本面相关的信息，对公司未来营业收入增长率具有良好的预测效果。其次，从不同类型的消费者在线投诉的预测能力来看，针对质量的“结果失败”投诉和企业无回应的投诉具有更强的基本面预测效果。再次，我们针对消费者在线投诉的预测能力开展了进一步讨论，发现消费者在线投诉具有行业溢出效应，

表 14 在线投诉预测能力：
锦上添花还是雪中送炭

	<i>SalesGrowth</i>	
	(1)	(2)
<i>L.CPLT</i>	-0.007 (0.004)	-0.008* (0.004)
<i>L.CPLT×LowAnalyst</i>	-0.038*** (0.012)	-0.031*** (0.012)
<i>L.LowAnalyst</i>	0.043** (0.021)	0.052*** (0.020)
<i>LLNSIZE</i>		0.118*** (0.017)
<i>LROA</i>		1.598*** (0.164)
<i>LBM</i>		0.082 (0.055)
<i>LRET</i>		0.059*** (0.020)
Constant	0.283*** (0.046)	-2.077*** (0.381)
Time FE	Yes	Yes
Firm FE	Yes	Yes
N	5688	5688
调整后 R ²	0.687	0.708

表 15 在线投诉的预测能力：
消费者业务重要性的影响

	<i>SalesGrowth</i>	
	(1)	(2)
<i>L.CPLT</i>	-0.001 (0.006)	-0.003 (0.006)
<i>L.CPLT×High toC</i>	-0.012** (0.006)	-0.011** (0.005)
<i>L.High toC</i>	0.020 (0.012)	0.017 (0.012)
<i>LLNSIZE</i>		0.119*** (0.017)
<i>LROA</i>		1.601*** (0.164)
<i>LBM</i>		0.091* (0.055)
<i>LRET</i>		0.060*** (0.020)
Constant	0.275*** (0.046)	-2.113*** (0.380)
Time FE	Yes	Yes
Firm FE	Yes	Yes
N	5688	5688
调整后 R ²	0.687	0.708

对同行业其他企业同样具有一定的基本面预测能力；从预测窗口来看，具有一定的持续预测能力；从预测类型来看，对企业财务危机也具有较好的预测能力。此外，消费者在线投诉的基本面预测能力具有显著的异质性，在低信息透明度和消费者业务重要性更高的样本中更加显著。本文的研究结论对于监管机构、上市公司、专业投资者等相关主体具有重要的启发价值。

首先，对于市场监管部门而言，可以借鉴“黑猫投诉”等非官方在线投诉平台的差异化优势，实现与官方投诉渠道的有机联动，将在线投诉平台信息作为市场监管和消费者权益保护工作的有力补充，不断提升监管效能。具体来说，一方面，聚焦“求同”，将官方投诉渠道和非官方在线投诉平台数据进行分析比对，将二者重合的行业、领域、产品和服务作为市场监管和消费者权益保护工作的重点领域；另一方面，重视“求异”，将官方投诉渠道少而在线投诉平台多且增速较快的投诉线索及时纳入有关部门的监管视野。

其次，对于金融监管部门而言，在全面落实中央金融工作会议关于“对风险早识别、早预警、早暴露、早处置”要求背景下，鉴于消费者在线投诉数据具有良好的公司基本面前瞻性预测能力，金融监管部门可以考虑将消费者在线投诉数据纳入上市公司监测体系，并据此针对性地关注上市公司，以更好地防范化解资产价格剧烈波动风险，从而维护金融市场稳定。

再次，对于上市公司而言，消费者在线投诉数据反映出公司经营状况方面存在的潜在问题，需要高度重视对消费者诉求的反馈和解决。一方面，消费者投诉实际上指明了公司产品和服务的薄弱点，通过加强对消费者投诉数据的监测有利于公司制定产品和服务改进方案，不断提高公司产品和服务的消费者满意度。另一方面，研究结果显示消费者投诉后是否回复具有不同效果，因此高度重视售后服务、建立专业售后服务队伍、提高消费者投诉的及时回复率等就成为提高公司管理水平的重要内容。

最后，对于专业投资者而言，消费者在线投诉这一另类数据对于公司基本面具有良好的预测能力，从而为投资者提供了前瞻性信息，有助于及时了解上市公司发展情况，为投资决策提供参考借鉴。具体来说，专业投资者可以将消费者在线投诉信息作为公司财务等信息的有力补充，并在此基础上构建更有效的投资组合，从而获得更高的收益^⑦。

（作者单位：蔡卫星，广东财经大学国家金融学研究中心、广东财经大学金融学院；蒲雨琦，广东财经大学金融学院；李浩民，武汉大学经济与管理学院）

注释

①该行业被投诉企业主要从事家电制造，如果以家电制造行业作为基准，则被投诉企业占家电制造行业比重约为30%。

②识别结果失败投诉和过程失败投诉的具体过程如下：(1)通过人工阅读的方法给5000个消费者投诉打上过程失败或结果失败的标签，并将其中4000条投诉作为训练集，剩下1000条投诉作为预测集；(2)将得到的训练集和预测集运用到机器学习中的Bert模型进行训练，得到一个专门用来判断消费者在线投诉属于结果失败还是过程失败的标签分类模型；(3)运用训练模型对消费者投诉进行识别。根据黄等(2023)的研究，Bert模型的优势在训练样本规模较小且包含一般文本中不常用的金融词汇的文本中尤为突出。

③我们计算了各个变量Partial R²，发现CPLT(0.15%)高于BM(0.11%)，与LNSIZE(0.16%)相当接近，低于ROA(3.33%)和RET(0.22%)，表明具有较好的解释能力。感谢审稿专家的建设性意见。

④感谢审稿专家的建设性意见。

⑤根据奥特曼(1968)，Z-score计算公式为： $Z\text{-score} = 1.2 \times \text{营运资金}/\text{总资产} + 1.4 \times \text{留存收益}/\text{总资产} + 3.3 \times \text{息税前利润}/\text{总资产} + 0.6 \times \text{股票总市值}/\text{负债账面价值} + 0.999 \times \text{销售收入}/\text{总资产}$ 。

⑥具体来说，我们采用如下过程来获得消费者业务相关词库：第一，参考李(2010)的做法，我们阅读了500份上市公司管理层讨论文本，确定了与消费者相关的种子词；第二，参考胡楠等(2021)和姜富伟等(2021)的做法，我们基于Word2Vec算法的CBOW模型对种子词进行相似词扩充，并邀请相关领域专家在此基础上对词语进行筛选，进而得到与消费者相关的词集。

⑦中外文人名(机构名)对照：潘诺利文(Purnawirawan)；笪(Da)；弗鲁特(Froot)；伯格(Berg)；布道克(Boudoukh)；法玛(Fama)；弗伦奇(French)；侯(Hou)；罗(Luo)；洪堡(Homburg)；艾拉德(Allard)；艾曼斯(Edmans)；黄(Huang)；姜(Jiang)；金(Kim)；夏皮罗(Shapiro)；文人(Fumito)；朱(Zhu)；阿加瓦尔(Agarwal)；福内尔(Fornell)；安德森(Anderson)；穆雷(Murray)；戴克(Dyck)；西尔弗曼(Silverman)；伦纳德(Leonard)；施陶斯(Stauss)；约瑟斯顿(Johnston)；比特纳(Bitner)；霍夫曼(Hoffman)；博肖夫(Boshoff)；古(Gu)；叶(Ye)；哈特(Hart)；史密斯(Smith)；博尔顿(Bolton)；巴本科(Babenko)；森(Sen)；杰格迪什(Jegadeesh)；里福莱特(Livnat)；温弗里(Winfree)；麦克卢斯基(Mccluskey)；佩蒂(Petty)；奥特曼(Altman)；巴伦(Barron)；方(Fang)；佩雷斯(Peress)；霍普(Hope)；李(Li)。

参考文献

(1)范秀成、刘建华：《顾客关系、信任与顾客对服务失败的反应》，《南开管理评论》，2004年第6期。

(2)范秀成、郑秋莹、姚唐、穆林：《顾客满意带来什么忠诚？》，《管理世界》，2009年第2期。

(3)方军雄：《我国上市公司信息披露透明度与证券分析师预测》，《金融研究》，2007年第6期。

- (4)郭国庆、杨学成、张杨:《口碑传播对消费者态度的影响:一个理论模型》,《管理评论》,2007年第3期。
- (5)胡楠、薛付婧、王昊楠:《管理者短视主义影响企业长期投资吗?——基于文本分析和机器学习》,《管理世界》,2021年第5期。
- (6)黄敏学、王峰、谢亭亭:《口碑传播研究综述及其在网络环境下的研究初探》,《管理学报》,2010年第1期。
- (7)姜富伟、孟令超、唐国豪:《媒体文本情绪与股票回报预测》,《经济学(季刊)》,2021年第4期。
- (8)姜付秀、张敏、陆正飞、陈才东:《管理者过度自信、企业扩张与财务困境》,《经济研究》,2009年第1期。
- (9)李想、石磊:《行业信任危机的一个经济学解释:以食品安全为例》,《经济研究》,2014年第1期。
- (10)李研、金慧贞、李东进:《社交网络情境下消费者口碑生成的影响因素模型:基于真实口碑文本的扎根研究》,《南开管理评论》,2018年第6期。
- (11)李焰、王琳:《媒体监督、声誉共同体与投资者保护》,《管理世界》,2013年第11期。
- (12)廖理、崔向博、孙琼:《另类数据的信息含量研究——来自电商销售的证据》,《管理世界》,2021年第9期。
- (13)罗进辉、杜兴强:《媒体报道、制度环境与股价崩盘风险》,《会计研究》,2014年第9期。
- (14)潘越、戴亦一、林超群:《信息不透明、分析师关注与个股暴跌风险》,《金融研究》,2011年第9期。
- (15)施文、赵雪:《在线消费者投诉对产品销量的影响机制:召回的中介作用》,《南开管理评论》,2023年网络首发。
- (16)孙瑾、苗盼:《口碑效价因人而异:调节定向与社会距离对消费者口碑效价的影响》,《南开管理评论》,2021年第5期。
- (17)肖静华、谢康、吴瑶、廖雪华:《从面向合作伙伴到面向消费者的供应链转型——电商企业供应链双案例研究》,《管理世界》,2015年第4期。
- (18)余明桂、李文桂、潘红波:《管理者过度自信与企业风险承担》,《金融研究》,2013年第1期。
- (19)袁兵、郑文哲:《互联网环境下企业家口碑对品牌评价的影响——顾客语言认知视角》,《南开管理评论》,2019年第2期。
- (20)张然、平帆、汪荣飞:《线上销售与未来股票收益》,《金融研究》,2022年第6期。
- (21)郑志刚、丁冬、汪昌云:《媒体的负面报道、经理人声誉与企业业绩改善:来自我国上市公司的证据》,《金融研究》,2011年第12期。
- (22)朱红军、何贤杰、陶林:《中国的证券分析师能够提高资本市场的效率吗——基于股价同步性和股价信息含量的经验证据》,《金融研究》,2007年第2期。
- (23)Agarwal, S., Qian, W. L. and Zou, X., 2021, "Disaggregated Sales and Stock Returns", *Management Science*, Vol.67(11), pp.7167~7183.
- (24)Allard, T., Dunn, L. H. and White, K., 2020, "Negative Reviews, Positive Impact: Consumer Empathetic Responding to Unfair Word of Mouth", *Journal of Marketing*, Vol.84(4), pp.86~108.
- (25)Altman, E. I., 1968, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *Journal of Finance*, Vol.23(4), pp.589~609.
- (26)Anderson, E. W., Fornell, C. and Mazvancheryl, S. K., 2004, "Customer Satisfaction and Shareholder Value", *Journal of Marketing*, Vol.68(4), pp.172~185.
- (27)Babenko, I. and Sen, R., 2016, "Do Nonexecutive Employees Have Valuable Information? Evidence from Employee Stock Purchase Plans", *Management Science*, Vol.62(7), pp.1878~1898.
- (28)Barron, O. E., Kile, C. O. and O' Keefe, T. B., 1999, "MD&A Quality as Measured by the SEC and Analysts' Earnings Forecasts", *Contemporary Accounting Research*, Vol.16(1), pp.75~109.
- (29)Berg, T., Burg, V., Gomovic, A. and Puri, M., 2020, "On the Rise of FinTechs: Credit Scoring Using Digital Footprints", *The Review of Financial Studies*, Vol.33(7), pp.3401~3450.
- (30)Bitner, M. J., 1990, "Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Response", *Journal of Marketing*, Vol.54(2), pp.69~82.
- (31)Boshoff, C., 1999, "RECOVSAT: An Instrument to Measure Satisfaction with Transaction-specific Service Recovery", *Journal of Service Research*, Vol.1(3), pp.236~249.
- (32)Boudoukh, J., Feldman, R., Kogan, S. and Richardson, M., 2019, "Information, Trading, and Volatility: Evidence from Firm-Specific News", *The Review of Financial Studies*, Vol.32(3), pp.992~1033.
- (33)Da, Z., Engelberg, J. and Gao, P., 2011, "In Search of Attention", *Journal of Finance*, Vol.66(5), pp.1461~1499.
- (34)Dyck, A., Volchkova, N. and Zingales, L., 2008, "The Corporate Governance Role of the Media: Evidence from Russia", *Journal of Finance*, Vol.63(3), pp.1093~1135.
- (35)Edmans, A., Fernandez-Perez, A., Garel, A. and Indriawan, I., 2022, "Music Sentiment and Stock Returns Around the World", *Journal of Financial Economics*, Vol.145(02), pp.234~254.
- (36)Fama, E. F. and French, K. R., 2015, "A Five-Factor Asset Pricing Model", *Journal of Financial Economics*, Vol.116(1), pp.1~22.
- (37)Fang, L. and Peress, J., 2009, "Media Coverage and the Cross-Section of Stock Returns", *Journal of Finance*, Vol.59(5), pp.2023~2052.
- (38)Fornell, C., Johnson, M. D., Cha, J. and Bryant, B. E., 1996, "The American Customer Satisfaction Index: Nature, Purpose and Findings", *Journal of Marketing*, Vol.60(4), pp.7~18.
- (39)Froot, K., Kang, N., Ozik, G. and Sadka, R., 2017, "What do Measures of Real-time Corporate Sales Say about Earnings Surprises and Post-announcement Returns?", *Journal of Financial Economics*, Vol.125(1), pp.143~162.
- (40)Fumito, K., 2009, "The Modeling of AISAS Marketing Process", *Japanese Journal of System Dynamics*, Vol.(8), pp.95~102.
- (41)Gu, B. and Ye, Q., 2014, "First Atep in Social Media: Measuring the Influence of Online Management Responses on Customer Satisfaction", *Production and Operations Management*, Vol.23(4), pp.570~582.
- (42)Hart, C. W. L., Heskett, J. L. and Sasser, W. E., 1990, "The Profitable Art of Service Recovery", *Harvard Business Review*, Vol.68

- (4), pp.148~156.
- (43) Hoffman, K. D., Kelly, S. W. and Rotalsky, H. M., 1995, "Tracking Service Failures and Employee Recovery Efforts", *Journal of Services Marketing*, Vol.9(2), pp.49~61.
- (44) Hope, O., 2003, "Analyst Following and the Influence of Disclosure Components, IPOs and Ownership Concentration", *Asia-Pacific Journal of Accounting & Economics*, Vol.10(2), pp.117~141.
- (45) Hou, K., Xue, C. and Zhang, L., 2015, "Digesting Anomalies: An Investment Approach", *Review of Financial Studies*, Vol.28(3), pp.650~705.
- (46) Huang, A. H., Wang, H. and Yang, Y., 2023, "FinBERT: A Large Language Model for Extracting Information from Financial Text", *Contemporary Accounting Research*, Vol.40(2), pp.806~841.
- (47) Huang, J., 2018, "The Customer Knows Best: The Investment Value of Consumer Opinions", *Journal of Financial Economics*, Vol.128(1), pp.164~182.
- (48) Jegadeesh, N. and Livnat, J., 2006, "Revenue Surprises and Stock Returns", *Journal of Accounting and Economics*, Vol.41(1~2), pp.147~171.
- (49) Jiang, F. and Kim, K., 2020, "Corporate Governance in China: A Survey", *Review of Finance*, Vol.24(4), pp.733~772.
- (50) Johnston, R., 1995, "The Determinants of Service Quality: Satisfiers and Dissatisfiers", *International Journal of Service Industry Management*, Vol.6(5), pp.53~71.
- (51) Leonard, D. A., 1985, "Experts as Negative Opinion Leaders in the Diffusion of a Technological Innovation", *Journal of Consumer Research*, Vol.11(4), pp.914~926.
- (52) Li, F., 2010, "Textual Analysis of Corporate Disclosures: A Survey of the Literature", *Journal of Accounting Literature*, Vol.29(1), pp.143~165.
- (53) Luo, X. and Homburg, C., 2008, "Satisfaction, Complaint, and the Stock Value Gap", *Journal of Marketing*, Vol.72(4), pp.29~43.
- (54) Luo, X., 2007, "Consumer Negative Voice and Firm-Idiosyncratic Stock Returns", *Journal of Marketing*, Vol.71(3), pp.75~88.
- (55) Murray, K. B., 1991, "A Test of Service Marketing Theory: Consumer Information Acquisition Activities", *Journal of Marketing*, Vol.55(1), pp.10~25.
- (56) Petty, R. E., Cacioppo, J. T. and Schumann, D., 1983, "Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement", *Journal of Consumer Research*, Vol.10(2), pp.135~146.
- (57) Purnawirawan, N., De Pelsmacker, P. and Dens, N., 2012, "Balance and Sequence in Online Reviews: How Perceived Usefulness Affects Attitudes and Intentions", *Journal of Interactive Marketing*, Vol.26(4), pp.244~255.
- (58) Shapiro, C., 1983, "Premiums for High Quality Products as Returns to Reputation", *Quarterly Journal of Economics*, Vol.98(4), pp.659~680.
- (59) Silverman, George., 1997, "How to Harness the Awesome Power of Word of Mouth", *Direct Marketing-Internet Marketing*, Vol.60(7), pp.32~37.
- (60) Smith, A. K. and Bolton, R. N., 1998, "An Experimental Investigation of Customer Reactions to Service Failure and Recovery Encounter: Paradox or Peril", *Journal of Service Research*, Vol.1(1), pp.5~17.
- (61) Stauss, B., 1997, "Global Word of Mouth: Service Bashing on the Internet Is a Thorny Issue", *Marketing Management*, Vol.6(3), pp.28~30.
- (62) Winfree, J. A. and Mccluskey, J. J., 2005, "Collective Reputation and Quality", *American Journal of Agricultural Economics*, Vol.87(1), pp.206~213.
- (63) Zhu, C., 2019, "Big Data as a Governance Mechanism", *The Review of Financial Studies*, Vol.32(5), pp.2021~2061.

Customer First: The Fundamental Prediction Ability of Consumer Online Complaints

Cai Weixing^{a,b}, Pu Yuqi^b and Li Haomin^c

(a. Research Center for National Finance, Guangdong University of Finance and Economics; b. School of Finance, Guangdong University of Finance and Economics; c. School of Economics and Management, Wuhan University)

Abstract: In the era of the digital economy, data elements have become the core engine of this new economic form, giving rise to more alternative data. As a typical alternative data, consumer online complaint information has a crucial impact on the development of enterprises. This study delves into the information content of consumer online complaints based on a dataset comprising over 10 million complaints. The results show that: first, the online complaint data of consumers contains information related to the company's fundamentals, which exhibits an excellent predictive effect on the company's future revenue growth rate. Second, among various consumer online complaints, those concerning "outcome failure" in quality and complaints that companies do not respond to have a more robust predictive effect on the companies' fundamentals. Third, we further discuss the predictive ability of consumer online complaints, pointing out that consumer online complaints have an industry spillover effect and have a specific predictive ability on the fundamentals of other companies in the same industry. From the perspective of the prediction window, these complaints demonstrate continuous predictive abilities over time. From the perspective of the prediction type, it also has great predictive ability for enterprise financial distress. In addition, the predictive ability of consumer online complaints on the company's fundamentals varies significantly, especially in samples with lower information transparency and higher importance to consumer business. These findings hold significant implications for regulators, enterprises, investors, and other stakeholders.

Keywords: consumer complaints; information content; alternative data; fundamental prediction

Customer First: The Fundamental Prediction Ability of Consumer Online Complaints

Cai Weixing^{a,b}, Pu Yuqi^b and Li Haomin^c

(a. Research Center for National Finance, Guangdong University of Finance and Economics; b. School of Finance, Guangdong University of Finance and Economics; c. School of Economics and Management, Wuhan University)

Summary: The emergence of the digital economy has positioned data elements as its core engine, with alternative data gaining significant attention as a novel data type. This type of data offers distinct advantages over traditional data, including larger volumes, improved timeliness, enhanced authenticity, and finer granularity, making it pivotal in the fields such as economics, finance, and accounting. The digital era has introduced new means for consumer rights protection, with online complaints serving as a new avenue for safeguarding consumer interests. This development has made consumer online complaint data a typical member of alternative data. Over the years, the research on consumer complaints was constrained by limited access to the large-scale, real-world data, often relying on small-scale surveys or experimental data. The utilization of consumer online complaint information helps to overcome these limitations by offering insights that were previously inaccessible. Set against the backdrop of policy, theoretical frameworks, and practical applications, this study adopts a multidisciplinary perspective to investigate the predictive ability of consumer online complaint data on the company's fundamentals. It aims at providing guidance on harnessing new data elements, exploring application scenarios, and facilitating value transformation in the digital economy.

Utilizing big data methods, this study collected over 10 million real consumer online complaints from "the Black Cat Complaints," which is a leading online complaint platform in China. These complaints were matched with listed company data to construct a comprehensive database for analysis. The research empirically tested the predictive ability of consumer online complaints on the company's fundamentals and examined the variance in predictive ability across different kinds of complaints. The findings indicate that consumer online complaint data, containing information relevant to the company's fundamentals, effectively can predict future revenue growth. Complaints regarding quality failures and lack of corporate response were particularly potent in forecasting the company's performance. Moreover, the study discussed the broader industry-wide implications, sustained predictiveness, and crisis predictive ability of consumer complaints, noting significant heterogeneity in predictive ability based on company transparency and consumer business reliance.

This study makes several contributions. first, it innovatively employs large-scale consumer online complaint data for analysis, enriching alternative data research from multiple perspectives. Second, it broadens the base of fundamental prediction research using Chinese data and a wide range of consumer company data. Third, it leverages millions of real complaints from "Black Cat Complaints" to scrutinize consumer complaint behaviors, thereby offering substantial data support for studying corporate reputation. Future research could explore online complaints' predictive effects on financial assets. Additionally, given that the targets of consumer online complaints encompass both listed and unlisted companies and considering the challenges associated with data accessibility for non-listed companies, future research may focus on analyzing their impact on non-listed companies.

Keywords: consumer complaints; information content; alternative data; fundamental prediction

JEL Classification: G11, G12, M14