

李雨洁^① 廖成林^② 李 忆^③ 伏红勇^④

摘 要 根据中国的商业环境,对淘宝网数据进行实证分析,文章作者发现在线商品评论呈近似反“L”形的非对称分布,且评论的均值不是商品质量的无偏估计量。这与现有研究认为在线商品评论呈正态分布且评价的均值是无偏估计量的结论不同,与国外的实证结果在线评论呈双峰分布亦不相同。为了深入研究在线商品评论存在偏差的原因,以及反“L”形分布的特征,文章构建了基于消费者效用的在线商品评论模型,该模型得出商品评论的均值作为无偏估计量的条件,并进一步揭示出中国电子商务网站“默认好评”机制、“退货评价”机制和消费者主动评价偏差是在线商品评论存在偏差的重要原因。

关键词 在线商品评论; 商品质量; 分布; 偏差; 矫正

在线商品评论可信吗

——在线商品评论的偏差分析及矫正策略^⑤

0 引言

在虚拟的网络购物环境中,在线商品评论是消费者获取商品质量信息、进行口碑传播的重要途径,是商家预测未来商品销量的重要依据,也是电子商务网站评估商家信誉的重要标准,因此受到学术界和实业界的共同关注(金立印等, 2008; Cui et al., 2012)。学者们从不同的角度对在线商品评论进行研究,然而在线商品评论是否存在偏见仍有争议。争议的焦点集中于在线商品评论能否反映商品的真实质量,评论均值的变化对商品未来销量是否有影响(Chevalier and Goolsbee, 2003; Duan et al., 2008; Liu,

2006)。这种争议的原因在于在线商品评论的分布尚不明确,无法确定商品评论的均值是否是商品质量的无偏估计量。研究在线商品评论的分布,以及该分布的均值和方差,可以帮助个体和组织迅速把握商品评论的特征和趋势,提高决策效率,因此具有理论和实践意义。

已有文献针对在线商品评论的分布有两种观点: Chevalier 和 Mayzlin(2006)以及 Clemons 等(2006)认为在线商品评论呈正态分布,评论的均值可以作为商品质量的无偏估计量,用来预测未来商品的销售量。Hu 等(2008, 2009)则认为在线商品评论呈“J”形双峰分布,评论不能直接反映商品的真实质量。这些研究的结论不完全一致,而且国外的商业和人文环境与中国有所差

① 李雨洁,重庆大学经济与工商管理学院博士研究生,E-mail: liyujie87@gmail.com。

② 廖成林,重庆大学经济与工商管理学院教授,博士生导师,E-mail: liaocl2005@126.com。

③ 李忆,重庆邮电大学经济与工商管理学院副教授,E-mail: cquliyi@163.com。

④ 伏红勇,重庆大学经济与工商管理学院博士研究生,E-mail: fuhongyong.cqu@gmail.com。

⑤ 非常感谢匿名评审专家提出宝贵和中肯的修改意见,特此致谢!

异,中国电子商务网站的交易机制、评价机制以及消费人群特征与国外有很大的不同,这会导致在线商品评论的分布产生哪些变化?这些变化对网络口碑特征及传播有效性有怎样的影响?为此,本文结合中国的现实背景,研究在线商品评论的分布,一方面对前人的研究结论进行检验,另一方面扩充了我国的相关实证结论,为中外跨文化的比较研究提供新的出发点。

关于在线商品评论的研究主要分为两类:一类是关于在线评论的有用性研究。在线商品评论的有用性,反映了消费者对评论质量的感知和评价。一般而言,有用性越高的商品评论,对消费者和商家的决策越有帮助。Mudambi 和 Schuff(2010)认为评论的等级和深度对评论有用性有重要影响。Pan 和 Zhang(2011)得出相同的结论,并认为商品类型对该影响有调节作用。Bobadilla 等(2010)利用协同过滤度量方法(collaborative filtering metric)同时考虑评论信息和有用性投票对在线商品评论有用性以及商品推荐系统的影响。随着在线口碑内容的不断丰富,一些学者研究口碑信息的特征对评论有用性的影响,如 Cao 等(2011)利用文本挖掘技术,从评论基本特征、文体特征、语义特征三个方面研究其对在线评论有用性投票的影响,得出语义特征对评论有用性投票的影响最大。Li 和 Zhan(2011)通过实证和实验的方法分析了语言风格、组织结构以及其他评论内容特征对感知评论有用性的影响。在线评论的有用性一方面影响消费者的采纳行为,同时也影响商品未来的销量(Cui et al., 2012; Zhu and Zhang, 2010)。Forman 等(2008)从评价者身份的视角,以亚马逊的数据实证分析得出评价者身份的披露对商品评论有用性和商品销售都有正向的影响。Ghose 和 Ipeirotis(2010)作了类似的研究,不同的是综合运用了计量经济学和文本挖掘的方法,得出商品评论的主观性、忠实度、可读性以及语言的正确性分析对评论有用性和商品销量有显著影响。总而言之,以上研究都肯定了在线评论对口碑传播及营销行为的重要性,然而这些研究并没有论证在线商品评论是否存在偏差。倘若

评价的数据本身是有偏误的,则会误导消费者和组织的决策者,为决策带来风险。

第二类研究主要涉及在线商品评论中不同类型的偏差。Rabin 和 Schrag(1999)较早开始关注首因效应对在线评论偏差的影响,指出第一印象会造成验证性偏差(confirmatory bias),使人们在获取新信息时受到先验的干扰。Kapoor 和 Piramuthu(2009)对这种偏见进一步定义为顺序偏差(sequential bias),研究顺序偏差随时间的动态变化。Sikora 和 Chauhan(2012)则更具体地研究了顺序偏差的特征,基于亚马逊网站19个不同商品类别的面板数据,采用卡尔曼滤波算法(Kalman Filtering Approach)估算在线商品评论中存在的顺序偏差,发现无论是体验类商品还是实物类商品,在线评论的评级存在显著的顺序偏差,且体验类商品的在线评论顺序偏差受评论者的个性特征的影响,而实物类商品在线评论顺序偏差则同时受评论本身和评论者个性特征的双重影响。由此可见,顺序偏差主要考虑时间因素对在线商品评论的影响。Staddon 和 Chow(2008)根据亚马逊网站的书籍评论,指出书籍的作者会委托熟人进行评价,造成在线书评存在评论者偏差(reviewer bias)。Hu 等(2010)的研究与之类似,并进一步证明了这种操纵性评价不仅存在于作者与评论者之间,还存在于卖家、出版商和评论者之间,指出企业的这种操纵性策略是商品真实质量的单调递减函数。由此可见,该类型的偏差主要考虑评价者的个体因素对在线商品评论的影响。

根据笔者所掌握的文献,国外已有一些在线商品评论偏差的成果,但主要集中于讨论时间因素和评论者个体因素对偏差的影响;国内学者彭惠和王永端(2012)的研究中提到,评价机制的弊端会影响商品评论的可信度;但以上研究都较少涉及以下问题:网站的评价机制对在线商品评论的偏差有何影响?如何反映评价机制对在线商品评论分布的影响?该机制下在线商品评论分布呈怎样的特征?如何纠正评价机制所导致的评论偏差?澄清该问题,将有助于我们更全面地理解在线商品评论的特点和网络口碑的

效用,因而值得深入探讨。

针对现有研究的不足,本文进行了以下研究工作:第一,对淘宝网的数据进行严格的采集与筛选,实证分析我国在线商品评论的分布,并与前人的相关研究结论进行比较;第二,建立基于消费者效用的在线商品评论模型,反映评价机制对在线商品评论分布的影响,给出了反映商品评论的分布函数,并计算该分布的均值和方差;第三,根据在线商品评论的分布特征,提出把在线商品评论的均值作为商品质量无偏估计量的条件以纠正评价机制对在线商品评论偏差的影响;第四,分别讨论“默认好评”机制、“退货评价”机制以及消费者主动评价的偏差对在线商品评论均值的影响,以及这些影响对管理实践的意义。

区别于之前的同类研究,本文针对中国的电子商务市场情境,研究在线商品评论的分布,揭示我国网络口碑传播的特征,为后续跨文化比较提供了新的思考。同时,本文从评价机制的角度,研究在线商品评论的偏差,利用分布建立模型反映我国在线商品评论的影响因素,并针对评价机制对评论偏差的影响机理进行了详细的讨论和解释。这丰富和完善了国内电子口碑传播的相关理论,也为中国的消费者和企业的营销实践提供参考。

1 背景

在中国的电子商务情境下,以淘宝网为例,其评价机制与国外电子商务网站有很大的不同。从计分方式看,淘宝网对商品评价的计分不是李克特量表式的5分制(1分表示最差,5分表示最好),而是定性地在“好评”“中评”“差评”中三选一。网站对评价计分进行累积,一方面反映为商品的好评率,另一方面反映为卖家的信用度。好评加1分,中评记0分,差评扣1分^①。只有消费者给予“好评”时,商家才能获得正向的激励(加1分)。从缺失评价看,存在“默认好评”机制,即消费者在交易成功后15天之内没有对商品进行评价,系统会代替消费者自动给出好评^②。这意

味着如果消费者没有及时对购买的商品进行评价,那么遗漏的评价只会变成好评,不会变成中差评。同时,发生退货的商品即交易关闭,这使一些因商品质量缺陷而退货的消费者不能再对商品进行评价。这种评价机制对我国的在线商品评论的分布有怎样的影响?在线商品评论能否真实地反映商品的质量?若不能,如何通过评论信息获知商品的真实质量?鉴于我国电子商务市场的特殊性,国内的相关研究不能直接照搬国外的研究结论;因此,亟须对我国在线商品评论进行深入探讨,而评价机制作为导致评论偏差的重要来源也为本文提供了研究视角。

2 我国在线商品评论分布的实证分析

为反映在线商品评论的真实分布,本文以淘宝网的数据为研究对象,对在线商品评论数据进行实证分析,初步反映分布的基本特征,并将我国的实证数据结果与国外的相关研究结果进行比较,深入分析国内外在线商品评论存在差异的原因。

2.1 数据来源

所选数据来源于淘宝网,数据收集遵从以下流程:(1)为保证样本的数量和多样性,筛选“淘宝指数”(淘宝官方数据分享平台)中热销类目排行前十的商品类别,分别是女装、美容护肤、数码配件、男装、手机、珠宝配饰、女鞋、笔记本电脑、零食、汽车用品。(2)从以上十种类别中随机选取女装、珠宝配饰、数码配件三类商品进行数据收集。(3)通过“合并相同卖家宝贝”,以卖家为单位,收集以女装、珠宝配饰、数码配件为主营业务的卖家的主营业务评价数据。

^① 淘宝网评价的规则与计分方法, <http://service.taobao.com/support/knowledge-882289.htm?dkey=catview>。

^② 淘宝网系统默认评价说明, <http://service.taobao.com/support/9910.htm>。

表1 样本的描述性统计数据

商品类别	评价等级			退货数
	差评	中评	好评	
女装	6 263	24 813	2 270 040	386 429
珠宝配饰	5 875	17 042	1 659 343	108 060
数码配件	1 573	3 081	256 081	35 872

为保证样本的有效性,基于以上数据收集,本文按照以下规则进行数据筛选:(1)天猫商城和淘宝网的评价体系存在差异,因此排除天猫商城的相关数据以保证数据来源的对等性;(2)按照“销量排名”,分别选取三类商品中排名前150名的卖家数据以保证数据的充足性;(3)排除主营业务占比低于90%的卖家数据,以保证退货数据主要产生于主营业务。(4)由于淘宝网只公布近30天的退货数据,为保证数据来源的对等性,商品评价的数据也只选取最近一个月(2012年9月)的样本;(5)为控制其他因素的干扰,排除退货数据中因“买家无理由退款”和“未收到货”而产生的数据,以保证“因商品质量问题而退货”数据的准确性。表1是样本的描述性统计数据。

2.2 数据分析

本文首先对女装、珠宝配饰、数码配件三类商品的在线商品评论数据进行分析。如图1所示,折线图显示出三类商品在线评价呈近似于反“L”形的分布。为了便于观察,主要纵坐标轴表示女装、珠宝配饰的刻度,次要纵坐标轴表示数码配件的刻度。三类商品的好评率分别高达98.6%,98.6%,98.2%,呈极度的正向偏移且非对称的分布。这与Lelis和Howes(2008)认为在线评论呈正态分布的结论不一致。

Jensen(1969)的研究中提到,任何检验只要拥有足够多的样本,最终都呈正态分布。为验证数据的稳定性,本文对样本再次筛选,选取卖家样本中评论数大于500条的商品进行同样的分析。如图2所示,折线图仍显示出近似于反“L”形的非对称分布,这和前人的研究结果不相同;

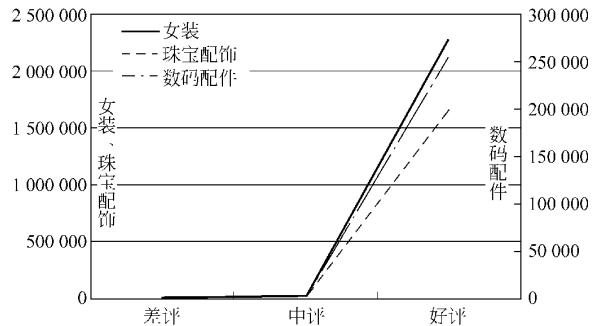
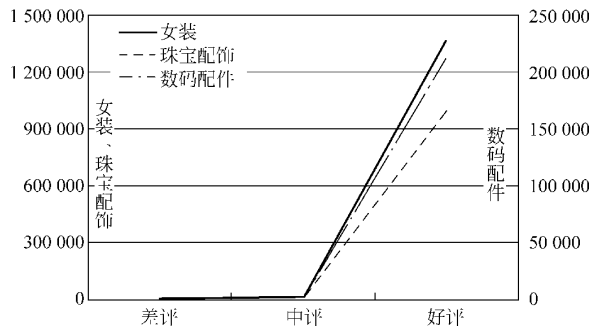


图1 女装、珠宝配饰、数码配件在线商品评论的分布

图2 女装、珠宝配饰、数码配件在线商品评论的分布
(每件商品评论数>500条)

因此,图2的结果进一步验证了在线商品评论极度正向偏移的非对称分布,并不是由于商品评论数太少的原因。

为检验在线商品评论分布的正态性,本文采用SPSS17.0对样本数据进行单样本的Kolmogorov-Smirnov检验(K-S检验)。结果如表2所示,“差评”“中评”“好评”三个评价等级的 p 值都等于0.000,在0.01、0.05和0.1水平上都不显著。表示接受原假设认为该分布是正态分布的概率是零,因此K-S检验的结果表明在线商品评论的分布是非正态分布。

表 2 正态分布的 K-S 检验结果

	差 评	中 评	好 评
Kolmogorov-Smirnov Z	4.648	5.164	4.794
Asymp. Sig. (2-tailed)	0.000<0.01	0.000<0.01	0.000<0.01

2.3 讨论

Hu 等(2008)选取书籍、DVD 和录像机三类产品为研究对象,根据亚马逊网站的数据进行实证分析,发现在线商品评论的分布呈“J”形的双峰分布。本文借鉴该论文的研究方法,选取我国的实证数据进行分析,发现我国在线商品评论的分布是非正态的,呈现出极度正向偏移、近似于反“L”形的分布。导致国内外实证结果出现差异的原因,本文认为主要有以下三个方面:

第一,这和淘宝网的“默认好评”机制有关。并不是所有的消费者购买商品之后都会主动评价。根据消费者行为与口碑传播动机的文献,通常持有极端观点的消费者(非常满意或非常不满意的消费者)更倾向于主动评价(Richins, 1983; Sundaram et al., 1998)。Anderson(1998)的研究中也发现,相对于持一般评价的消费者来说,持非常满意和非常不满意评价的消费者更愿意进行口碑传播。这意味着由于消费者主动评价的偏差,在线商品评论会出现缺失,而我国电子商务网站的“默认好评”机制将缺失的评论直接默认到好评,因此推高了在线商品的好评率。本文收集了淘宝网一个月的女装、珠宝配饰、数码配件三类商品的系统“默认好评”数据进行统计分析,结果如图 3 所示,柱形图显示出三类商品的系统默认评分率分别占其评论总数的 47.7%、35.4% 和 33.2%。其中女装类商品的默认评分率高达 47.7%,接近评论总数的一半,另外两类商品的默认评分率大约是评论总数的三分之一,这进一步说明了“默认好评”机制对在线商品评论偏差的影响程度。国外的电子商务网站则不存在该机制的影响。

第二,这和评价等级的计分方式有关。淘宝网与国外的电子商务网站计分方式不同,淘宝网对商品的评价不是李克特量表式的 5 分制(1 分

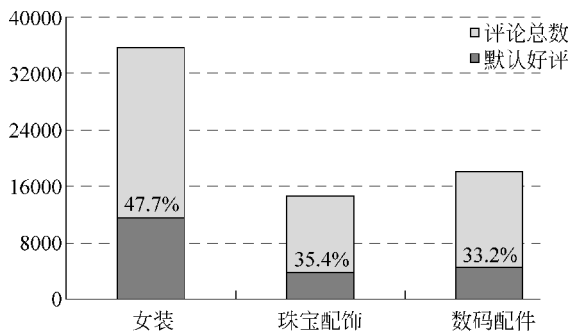


图 3 系统默认好评数与评论总数的比例关系

表示最差,5 分表示最好),而是定性地在“好评”“中评”“差评”中三选一。好评加 1 分,中评记 0 分,差评扣 1 分。只有消费者给予“好评”时,商家才能获得正向的激励(加 1 分)。这就使一些商家为提高店铺信誉盲目追求好评率而操纵买方评价,造成在线商品好评率的失真。此外,受消费者自身评价标准的不同,一些消费者认为给予“中评”就可以达到宣泄不满情绪的程度。这导致区分中、差评的界限不明显,中评与差评的数目相差不大(谈晓勇和任永梅,2008)。

第三,这和退货商品评价的缺失有关。消费者的异质性会导致其购后行为出现差异。一些消费者在对商品不满时,会以“差评”的方式抱怨,而另一些消费者则会以“退货”的方式结束交易。现行的评价机制使消费者在退货之后不能再对商品进行评价,这就阻碍了消费者对瑕疵商品的口碑传播,人为地降低了在线商品评论的差评率。本文收集了淘宝网 30 天内女装、珠宝配饰、数码配件三类商品的退货数据,并排除了物流及其他非商品质量因素导致退货的数据,进行统计分析,如图 4 所示。柱形图显示出三类商品的退货数分别占其总交易数量的 17%、6% 和 14%。在口碑传播中,“差评”往往比“好评”更具有说服力(Sen and Lerman, 2007),可信度更强,因此这些因退货而缺失的“潜在差评”对在线

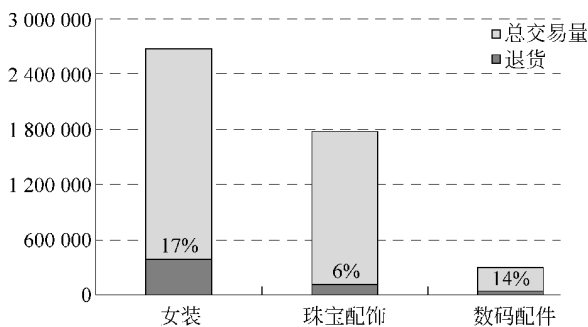


图4 退货商品与总交易量的比例关系

商品评论分布的偏差有重要影响。国外的电子商务网站则不会限制退货商品的评价,因此没有此方面的影响。

3 基于消费者效用的在线商品评论模型

为进一步反映在线商品评论分布的偏差,基于实证研究结果,本文将构建理论模型,讨论在线评论均值作为无偏估计量的条件以纠正偏差,旨在帮助消费者和商家不被偏差的商品评论所误导。

3.1 基本假设

本文提出三个基本假设:

第一,购买商品之后,若所有的消费者都主动撰写评论,即商品的在线评论不存在缺失,则在线商品评论服从正态分布。该假设已被前人通过实验法获得证明(Hu et al., 2008)。

第二,消费者在对商品非常满意和非常不满时,才会主动撰写商品评论,对商品持中庸态度的消费者不会主动撰写评论。现有营销学和消费者行为学的文献已证明了,消费者的极端情绪是他们进行口碑传播的动机,而在线口碑传播的方式就是对商品撰写评论。

第三,消费者是理性人,他们对商品的评价可以如实地反映消费者的效用。

3.2 符号定义

q : 商品的内在质量,表示商品本身固有的

属性。

v_i : 消费者的异质性,包括消费者的个人偏好、生活习惯、教育水平、社会地位等因素。

$u_i(q)$: 消费者 i 的效用。

$\underline{\delta}$: 消费者主动撰写差评的效用临界值。

$\bar{\delta}$: 消费者主动撰写好评的效用临界值。

$\bar{\delta}$: “默认好评”机制下,未主动评价部分被系统提高后的效用临界值。

$\underline{\delta}$: 消费者因不满意商品而退货的效用临界值。

\bar{r}_n : n 个消费者在线商品评论的均值。

r_i : 现实中人们实际观测到的消费者 i 的在线商品评论。

3.3 模型的建立

假设 q 代表商品的内在质量,表示商品本身固有的属性,不受特定消费群体的影响,因此设定为常量。 v_i 代表消费者的异质性,由于 v_i 受到消费者的个性特征、教育水平、社会地位、生活环境、收入水平等众多因素的影响,因此我们假设消费者的异质性 v_i 服从均值是 0,方差是 σ^2 的标准正态分布, $v_i: N(0, \sigma^2)$ 。令消费者 i 的效用 $U_i(q)$ 满足式(1):

$$u_i(q) = q + v_i \quad (1)$$

因此,消费者 i 的效用 $u_i(q)$ 服从均值是 q ,方差是 σ^2 的正态分布, $u_i(q): N(q, \sigma^2)$ 。根据之前的论述,主动撰写评论会产生时间和精力成本,并不是每一个消费者在购买商品之后都会主动撰写评论。只有在极端情绪的促使下,消费者才会通过口碑传播进行赞扬或者抱怨,以此来鼓励或劝诫其他潜在消费者。现有文献论证了消费者主动撰写商品评论的动机源于对商品极度满意和极度不满。当消费者对商品极为不满时,有可能会发生退货,鉴于淘宝现有的评价机制,退货即交易关闭,退货的商品不能再进行评价。

因此,给定某个消费者 i 主动撰写评论的概率 p (如图 5 中的阴影部分)为:

$$p = \begin{cases} 1 & \underline{\delta} < u_i(q) < \bar{\delta} \quad \text{或} \quad \bar{\delta} < u_i(q) \\ 0 & \underline{\delta} < u_i(q) < \bar{\delta} \end{cases} \quad (2)$$

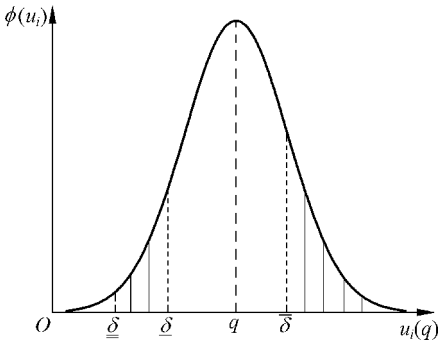


图5 消费者*i*主动撰写评论的概率分布

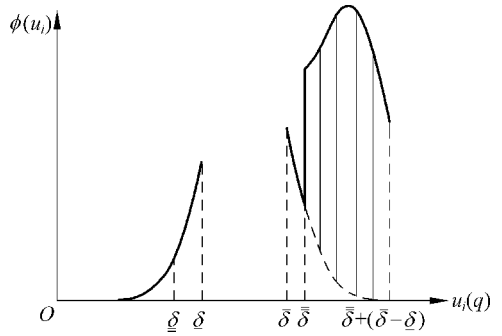


图6 未主动评价的消费者*i*的效用被“默认好评”提高之后的概率分布

其中, $\underline{\delta}$ 和 $\bar{\delta}$ 是消费者主动撰写评价的两个临界值, δ 是消费者退货的效用临界值。本文假设 $\underline{\delta}$ 和 $\bar{\delta}$ 以及 δ 是常数, 并且 $\underline{\delta} < \delta$, 表示消费者对商品极度不满, 其效用低于主动差评的效用。当 $u_i(q) < \underline{\delta}$ 时, 即消费者 i 的效用低于退货的临界值, 表示消费者对商品无法忍受, 需要退货; 当 $\underline{\delta} < u_i(q) < \delta$ 时, 即消费者 i 的效用介于退货和差评临界值之间, 表示消费者对商品非常不满, 通过差评来抱怨商品的不足, 但不需要退货; 当 $\delta < u_i(q) < \bar{\delta}$ 时, 即消费者 i 的效用介于差评与好评临界值的中间, 表示消费者认为商品一般, 没有撰写评论的意愿, 不主动进行评价; 当 $\bar{\delta} < u_i(q)$ 时, 即消费者 i 的效用高于好评的临界值, 表示消费者对商品非常满意, 通过好评来赞扬商品的优点。

由于“默认好评”机制的存在, 未主动评价部分消费者 i 的效用 $\delta < u_i(q) < \bar{\delta}$ 被系统人为地提高了。为了反映这种现象, 本文提出 $\bar{\delta}$, 用于表示未主动评价的消费者效用被提高后的临界值。因为系统将未主动评价部分消费者的效用人为地提高到极好的水平, 因此在连续化的函数中, 本文认为这种被动好评的临界值不低于消费者主动好评的临界值, 即 $\bar{\delta} \geq \delta$; 并且在 $\underline{\delta} < u_i(q) < \bar{\delta}$ 之间的消费者 i 的效用被提高到 $\bar{\delta} < u_i(q) < \bar{\delta} + (\bar{\delta} - \delta)$ 的范围内 (如图 6 阴影部分所示)。为了便于处理, 将 $\underline{\delta} < u_i(q) < \bar{\delta}$ 区间内

的密度量向 $\bar{\delta} < u_i(q) < \bar{\delta} + (\bar{\delta} - \delta)$ 区间内进行平移叠加。

本文用 \bar{r}_n 表示 n 个消费者在线评论的均值。当 $\underline{\delta} = \bar{\delta}$ 且 $\underline{\delta}$ 足够小时, 所有消费者都主动评价, 在线商品评论近似地呈正态分布, 此时在线商品评论的均值如实地反映商品的质量。利用中心极限定理:

$$\lim_{n \rightarrow \infty} \Pr \left[\frac{\bar{r}_n - q}{\sigma} \leq x \right] = \Phi(x) \quad (3)$$

由此得到 \bar{r}_n 近似地服从均值是 q 、方差是 σ^2/n 的正态分布。此时, 人们可以从在线商品评论中获取商品质量的真实信息。

然而, 实际情况并非如此。当 $\underline{\delta} < \delta < \bar{\delta}$ 时, 用 r_i 表示现实中人们实际观测到的消费者 i 的在线商品评论, 它包括一部分消费者的主动评论和另一部分被“默认好评”的商品评论, 以及由于退货而缺失的评论。由于在线商品评论反映了消费者的效用, 因此 $r_i = u_i(q)$, 且 $i \in \{i \mid \underline{\delta} < u_i(q) < \delta \text{ 或 } \bar{\delta} < u_i(q)\}$ 。为讨论在线商品评论的均值是否是商品质量的无偏估计量, 本文得出定理 1, 计算出在线商品评论分布的均值和方差。

定理 1 在线商品评论的均值 $\bar{r}_n = q + \left((\bar{\delta} - \delta) \left(\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\delta - q}{\sigma}\right) \right) + \Phi\left(\frac{\delta - q}{\sigma}\right) \right) \Delta$, 方差 $\sigma_n^2 = ((\bar{\delta} - q - 2(\bar{\delta} - \delta)\Psi)\sigma\Omega + (1 - \Omega^2))\sigma^2 +$

$$(\bar{\delta} - \underline{\delta})^2 \Psi - (\bar{\delta} - \underline{\delta})^2 \Psi^2 + 2\sigma(\bar{\delta} - \underline{\delta})\Theta)/n, \text{其中 } \Lambda = 1/\left(1 - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\right), \Psi = \left(\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right)\right)\Lambda, \Omega = \left(\frac{\bar{\delta} - q}{\sigma}\right)\Lambda, \Theta = \left(\phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \phi\left(\frac{\underline{\delta} - q}{\sigma}\right)\right)\Lambda.$$

证明：见附录 1。

从定理 1 中不难看出，现实中商品评论均值 \bar{r}_n 是一个有偏差的估计量，同时受 $(\bar{\delta} - \underline{\delta})\left(\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right)\right)$ 、 $\phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\sigma$ 和 $1/\left(1 - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\right)$ 的影响。

3.4 模型的讨论

为了纠正偏差，本文根据定理 1 讨论在线商品评论的均值作为商品质量无偏估计量的条件，并分别讨论“默认好评”机制和“退货评价”机制以及消费者主动评价的偏差对在线商品评论均值的影响。

命题 1：在线商品评论的均值 \bar{r}_n 收敛于商品的真实质量 q 的条件是：所有交易成功的消费者购后都主动评价商品并纠正“退货评价”机制，使发生退货的商品也可以被评价。

证明：由定理 1 知，当 $\bar{\delta} - \underline{\delta} = 0$ 或者 $\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) = 0$ ，且 $\phi\left(\frac{\bar{\delta} - q}{\sigma}\right) = 0$ 时，

$\bar{r}_n = q$ ，表示在线商品评论的均值不存在偏差。

结合实际背景， $(\bar{\delta} - \underline{\delta})$ 表示“默认好评”机制的影响， $\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right)$ 表示未主动评价部分的人数所占的比例， $\phi\left(\frac{\bar{\delta} - q}{\sigma}\right)$ 表示商品退货人数的影响。

若取消“默认好评”机制， $\bar{\delta} - \underline{\delta} = 0$ ；当所有消费者都主动评价时， $\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) = 0$ ；若退货评价机制允许消费者在退货之后，仍然可

以对商品进行评价，则 $\phi\left(\frac{\bar{\delta} - q}{\sigma}\right) = 0$ 。

但“默认好评”机制取消之后，未主动评价的消费者依然存在，因此只有当上述条件满足 $\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) = 0$ 且 $\phi\left(\frac{\bar{\delta} - q}{\sigma}\right) = 0$ 时，在线商品评论的均值才收敛于商品的质量。否则，相对于商品真实质量 q ，在线商品评论均值将存在非零偏差：

$$\left((\bar{\delta} - \underline{\delta})\left(\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right)\right) + \phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\sigma\right)\Lambda, \Lambda = 1/\left(1 - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\right) \quad (4)$$

命题 1 证毕。

由命题 1 可知，当取消“默认好评”机制且纠正“退货评价”机制时，由于消费者主动评价的偏差依然存在，所以在线商品评论仍有偏差。因此取消“默认好评”机制，纠正“退货评价”机制的同时，还要使所有消费者主动评价才可以纠正在线商品评论的偏差。基于此，本文认为不仅网站的评价机制会对商品评论的偏差产生影响，消费者是否主动撰写评论也是导致在线商品评论出现偏差的关键原因。

由于“默认好评”机制和“退货评价”机制以及消费者主动评价的偏差都对在线商品评论的分布都有重要的影响，本文将分别讨论这些因素的影响机理。当只考虑“默认好评”机制作用时，得出命题 2；当只考虑“退货评价”机制的作用时，得出命题 3；当只考虑消费者主动评价的偏差时，得出命题 4。

命题 2：当只考虑“默认好评”机制的作用时，在线商品评论的均值 \bar{r}_n 大于商品的真实质量 q ，并且 \bar{r}_n 受“默认好评”机制作用的效用临界值 $\bar{\delta}$ 、消费者主动评价的效用临界值 $\underline{\delta}$ 与 $\underline{\delta}$ 的影响。其中 $\bar{\delta}$ 对 \bar{r}_n 有正向影响且该影响随未主动评价人数所占比例的增大而增强， $\underline{\delta}$ 对 \bar{r}_n 有正向影响， $\underline{\delta}$ 对 \bar{r}_n 有负向影响。

证明：当只考虑“默认好评”机制，不考虑“退

货评价”机制的影响时,可以理解为 $\underline{\delta} \rightarrow -\infty$, 则

$$\lim_{\underline{\delta} \rightarrow -\infty} \phi\left(\frac{\bar{\delta}-q}{\sigma}\right) = 0, \lim_{\underline{\delta} \rightarrow -\infty} \Lambda = 1 / \left(1 - \Phi\left(\frac{\bar{\delta}-q}{\sigma}\right)\right) = 1,$$

因此,当考虑“默认好评”机制作用时,在线商品评论的均值为:

$$\bar{r}_n = q + (\bar{\delta} - \underline{\delta}) \left(\Phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta}-q}{\sigma}\right) \right) \quad (5)$$

因为 $(\bar{\delta} - \underline{\delta}) \left(\Phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta}-q}{\sigma}\right) \right) > 0$, 所以 $\bar{r}_n > q$;

由式(4)得: $\frac{\partial \bar{r}_n}{\partial \bar{\delta}} = \Phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta}-q}{\sigma}\right) >$

0, 表示“默认好评”机制作用下,未主动评价部分被系统提高后的效用临界值 $\bar{\delta}$ 对 \bar{r}_n 有正向影响;

$\Phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta}-q}{\sigma}\right)$ 表示未主动评价人数所占比例,因此, $\bar{\delta}$ 对 \bar{r}_n 的正向影响随其增大而增强。

$$\frac{\partial \bar{r}_n}{\partial \underline{\delta}} = - \left(\Phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta}-q}{\sigma}\right) \right) - \frac{(\bar{\delta} - \underline{\delta})}{\sigma}$$

$\left(-\Phi\left(\frac{\underline{\delta}-q}{\sigma}\right) \right) < 0$, 表示消费者主动差评的效用临界值 $\underline{\delta}$ 对 \bar{r}_n 有负向影响;

$$\frac{\partial \bar{r}_n}{\partial \bar{\delta}} = \Phi\left(\frac{\bar{\delta}-q}{\sigma}\right) \cdot \frac{1}{\sigma} > 0, \text{表示消费者主动好评的效用临界值 } \bar{\delta} \text{ 对 } \bar{r}_n \text{ 有正向影响。}$$

命题 2 证毕。

命题 2 证毕。

命题 2 论证了“默认好评”机制对在线商品评论均值的影响,并强调了未主动评价的消费者所占比例越大,该机制造成均值的偏差越大。同时,消费者主动好评的效用临界值越大,反映了消费者对商品的正向期望越强,对商品越不容易满意,消费者主动撰写好评的意愿越低,造成在线商品评论的缺失增大,导致“默认好评”机制将缺失评价被拔高到“好评”,因此推高了商品评论的均值。同理,主动差评的效用临界值越小,意味着消费者对商品的负面容忍度越强,消费者主动撰写差评的意愿越低,也造成在线商品评论的

缺失增大。在“默认好评”机制的作用下,缺失的“差评”被默认到“好评”,这也推高了商品评论的均值。

命题 3: 当只考虑“退货评价”机制作用时,在线商品评论的均值 \bar{r}_n 大于商品的真实质量 q , 且 \bar{r}_n 随退货人数的增多而增大。

证明: 当只考虑“退货评价”机制的作用时, $\Phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta}-q}{\sigma}\right) = 0$, 由定理 1 得:

$$\bar{r}_n = q + \phi\left(\frac{\bar{\delta}-q}{\sigma}\right) \sigma / \left(1 - \Phi\left(\frac{\bar{\delta}-q}{\sigma}\right)\right) \quad (6)$$

因为 $\phi\left(\frac{\bar{\delta}-q}{\sigma}\right) \sigma / \left(1 - \Phi\left(\frac{\bar{\delta}-q}{\sigma}\right)\right) > 0$, 所以 $\bar{r}_n > q$;

若退货人数增多,从图 4 中可以看出, $\underline{\delta}$ 将增大,由于 $\underline{\delta} < q$, 则 $\phi\left(\frac{\bar{\delta}-q}{\sigma}\right)$ 增大, $\left(1 - \Phi\left(\frac{\bar{\delta}-q}{\sigma}\right)\right)$ 随之减小,因此 $\phi\left(\frac{\bar{\delta}-q}{\sigma}\right) \sigma / \left(1 - \Phi\left(\frac{\bar{\delta}-q}{\sigma}\right)\right)$ 增大。

由于商品质量 q 不变,则退货人数越多,在线商品评论的均值 \bar{r}_n 被提高得越多。

命题 3 证毕。

由于现行的机制是退货即交易关闭,消费者不能再对商品进行评价,因此使一些由于商品质量问题而不满的消费者无法以在线评论的方式进行口碑传播。这会造成部分差评缺失,使评论的均值升高,出现偏误。在现实中退货现象不可避免,因此,命题 3 进一步说明,通过改进“退货评价”机制,使发生退货的商品能够被评价,可以纠正在线商品评论的偏差。

命题 4: 当只考虑消费者主动评价的偏差作用时,消费者对商品的正面期望较高会导致在线商品评论的均值 \bar{r}_n 小于商品的真实质量 q ; 反之,则使在线商品评论的均值 \bar{r}_n 大于商品的真实质量 q 。

证明: 当只考虑消费者主动评价的偏差作用时,在线商品评论的均值为:

$$\bar{r}_n = q + \frac{\phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \phi\left(\frac{\delta-q}{\sigma}\right)}{1 + \Phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \Phi\left(\frac{\delta-q}{\sigma}\right)} \sigma \quad (7)$$

式(7)的推导见附录2。

$$\textcircled{1} \text{ 若 } \underline{\delta} < \bar{\delta} < q, \text{ 则 } \phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \phi\left(\frac{\delta-q}{\sigma}\right) > 0,$$

从而 $\bar{r}_n > q$ 。表示消费者对商品的正面期望较低,此时消费者容易被满足,主动撰写好评的意愿较强,主动好评的人数增多,导致在线商品评论的均值大于商品的真实质量。

$$\textcircled{2} \text{ 若 } q < \underline{\delta} < \bar{\delta}, \text{ 则 } \phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \phi\left(\frac{\delta-q}{\sigma}\right) < 0,$$

从而 $\bar{r}_n < q$ 。表示消费者对商品的正面期望较高,此时消费者不易被满足,主动撰写好评的意愿较低,主动好评的人数减少,导致在线商品评论的均值小于商品的真实质量。

$$\textcircled{3} \text{ 若 } \underline{\delta} < q < \bar{\delta}, \text{ 则当 } |\bar{\delta}-q| > |\underline{\delta}-q| \text{ 时,}$$

$\phi\left(\frac{\bar{\delta}-q}{\sigma}\right) - \phi\left(\frac{\delta-q}{\sigma}\right) < 0, \bar{r}_n < q$,表示相对于消费者对商品的负面容忍度,其正面期望更大时,主动撰写好评的意愿更低,导致在线商品评论的均值小于商品的真实质量。反之,则 $\bar{r}_n > q$ 。

综合①②③可得,消费者对商品的正面期望较高会导致在线商品评论的均值 \bar{r}_n 小于商品的真实质量 q ,反之则使在线商品评论的均值 \bar{r}_n 大于商品的真实质量 q 。

命题4证毕。

根据命题4,消费者对商品的正向期望越高,越不容易满足,这会降低其主动好评的意愿,最终导致在线商品评论的均值低于商品的真实质量。反之,当消费者对商品的负面容忍度越高,越不容易抱怨,这会降低其主动差评的意愿,最终导致在线商品评论的均值高于商品的真实质量。因此,消费者对商品过高的正向期望和负面容忍度都会影响其主动评价的意愿,不利于口碑信息的正确传播。这提示网站和商家应客观

对商品进行描述,使消费者对商品建立正确的期望。

4 结论

4.1 研究结论与理论贡献

本文从评价机制的角度研究了在线商品评论的偏差的影响因素及纠正偏差的措施。基于淘宝网数据,本文实证分析了在线商品评论的分布特征,根据该分布建立数学模型,反映评价机制对在线商品评论分布的影响,讨论商品评论的均值作为无偏估计量的条件,并深入分析评价机制对在线商品评论偏差的影响机理,以此来纠正偏差。通过前面的研究有如下发现:①在线商品评论的分布是非正态的,且呈现极度正向偏移、近似于反“L”形的分布。②在线商品评论的均值不是商品质量的无偏估计量,导致其偏差的原因是“默认好评”机制和“退货评价”机制以及消费者主动评价的偏差。③为了纠正偏差,本文得出纠正在线商品评论的均值偏差的条件是,使所有购买商品的消费者都主动撰写商品评论,并修改现行的“退货评价”机制,使发生退货的交易仍然可以被评价。

本文为网络口碑及在线商品评论的相关理论补充了中国情境下的经验证据,发现了一些与国外研究不同的结果,进一步丰富了该领域的文献。

既有的国外文献认为在线商品评论呈正态分布(Chevalier and Mayzlin, 2006; Clemons et al., 2006)或者“J”形双峰分布(Hu et al., 2008, 2009);但是,与此不同的是,本文从实证数据和数学模型中都发现了我国电子商务情境下在线商品评论的分布是非正态的,呈现出极度正向偏移、近似于反“L”形的分布。这一结果的不同反映了在线商品评论偏差的来源不同造成了评论偏差的机制差异,也进一步证明了电子商务网站的评价机制对在线商品评论特征有重要影响(彭惠和王永瑞,2012)。同时,从侧面反映

了我国电子商务环境中卖家的不良行为可能会使消费者迫于压力给予虚假好评(李维安等, 2007)。本文借鉴前人的研究方法,结合我国的现实背景进行理论和实证研究,得出了不同于国外研究的结论,为后续研究提供了参考,这种中外研究结果的差异,也为后续研究提供了出发点。

既有文献中,关于在线商品评论偏差的研究结论是较为含混的(Hu et al., 2010; Mudambi and Schuff, 2010; Staddon and Chow, 2008)。本文从评价机制的视角进行研究,结果揭示了偏差的原因是“默认好评”机制和“退货评价”机制,以及消费者主动评价的偏差。根据在线商品评论的分布,本文提出商品评论均值作为商品质量无偏估计量的条件,以此来纠正偏差。这为网络营销及在线口碑研究提供了新的视角,且丰富和完善了网络环境下消费者行为的相关理论,并为企业实践提供有益的参考。

4.2 管理启示

对于管理实践而言,本文提示了网络交易的各参与方应克服现行评价机制导致的商品评论的偏差,维护良好的交易环境。(1)对于商家来说,应根据商品评论的分布,准确地识别真实的好评与差评,适时评估消费者偏好,调整供货与销售策略,传递商品价值。同时采取激励措施鼓励消费者积极参与购后评论,主动进行口碑传播。(2)对于消费者来说,一方面积极主动评价,有利于降低在线商品评论的缺失率,从而降低“默认好评”机制的负面影响,还可以帮助其他消费者获取真实的评论信息。另一方面针对可能存在的评论偏差,可以通过仔细阅读差评,或者阅读有评价内容的好评,以及从其他渠道获取信息为购买决策提供帮助。(3)对于电子商务网站来说,应尽快纠正评价机制,使评价数据尽量客观,同时配合商家一起采取措施尽量鼓励消费者积极参与购后评价,最大限度地调动持“中庸”态度的消费者主动评价,营造公平交易的商业环境。

4.3 研究局限与展望

本文的研究局限在于实证研究只选取了三种商品的数据进行统计,未来研究可以扩大样本,进行类似研究来充实研究结论。为了使模型更容易处理,本文的模型假设持中评的消费者不主动评价,未来研究可以采用更严格的假设,获取更精确的条件。除此之外,从本文的研究问题出发,仍有许多可以发展的空间。首先,可以考虑将消费者的购买偏好引入模型,考虑消费者对商品的偏好对在线商品评论的影响。其次,消费者的异质性可能会影响其主动评价的意愿,进而影响在线商品评论的偏差。因此可以考虑通过实证的方法,研究消费者的特征因素对评论偏差的影响机理。再者,考虑引入实验设计的方法,分析现行的计分方式对在线商品评论分布的影响。

参考文献

- 金立印,王如意,邹德强. 2008. 消费者对网络评论信息的可信性评价[J]. 营销科学学报, 4(4): 45-55.
- 李维安,吴德胜,徐皓. 2007. 网上交易中的声誉机制——来自淘宝网的证据[J]. 南开管理评论, 10(5): 36-46.
- 彭惠,王永瑞. 2012. C2C电子商务市场整体卖家信用分布研究[J]. 情报杂志, 31(5): 122-127.
- 谈晓勇,任永梅. 2008. C2C电子商务网站信用评价中的主要问题及其对策研究[J]. 经济理论研究, 11(19): 39.
- Anderson E W. 1998. Customer satisfaction and word of mouth [J]. Journal of Service, 1(1): 5-17.
- Bobadilla J, Serradilla F, Bernal J. 2010. A new collaborative filtering metric that improves the behavior of recommender systems [J]. Knowledge based System, 3(6): 520-528.
- Cao Q, Duan W J, Gan Q W. 2011. Exploring determinants of voting for the “helpfulness” of online user reviews: A text mining approach [J]. Decision Support Systems, 50(2): 511-521.

- Chevalier J A, Mayzlin D. 2006. The effect of word of mouth on sales: Online book reviews [J]. *Journal Marketing Research*, 43(3): 345-354.
- Chevalier J, Goolsbee A. 2003. Measuring prices and price competition online: Amazon and Barnes and Noble [J]. *Quantitative Marketing and Economics*, 1(2): 203-222.
- Clemons E K, Gao G D, Hitt L M. 2006. When online reviews meet hyper differentiation: A study of craft beer industry [J]. *Journal of Marketing Research*, 23(2): 149-171.
- Cui G, Lui H K, Guo X N. 2012. The effect of online consumer reviews on new product sales [J]. *International Journal of Electronic Commerce*, 17: 39-58.
- Duan W, Gu B, Whinston A B. 2008. Do online reviews matter? An empirical investigation of panel data [J]. *Decision Support Systems*, 45(4): 1007-1016.
- Forman C, Ghose A, Wiesenfeld B. 2008. Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets [J]. *Information Systems Research*, 19(3): 291-313.
- Ghose A, Ipeirotis P. 2010. Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics [C]. *IEEE Transactions on Knowledge and Data Engineering*, 23 (10): 1498-1512.
- Hu N, Liu L, Sambamurthy V. 2010. Fraud detection in online consumer reviews [J]. *Decision Support Systems*, 50(3): 614-626.
- Hu N, Pavlou P, Zhang J. 2009. Overcoming the J-shaped distribution of product reviews [J]. *Communications of the ACM*, 52(10): 144-147.
- Hu N, Pavlou, P Zhang J. 2008. Why do online product reviews have a J-shaped distribution? Overcoming biases in online word-of-mouth communication. Working paper. Singapore Management University.
- Jensen A. 1969. How much can we boost IQ and scholastic achievement? [J]. *Harvard Educational Review*, 39(1): 1-123.
- Kapoor G, Piramuthu S. 2009. Sequential bias in online product reviews [J]. *Journal of Organizational Computing and Electronic Commerce*, 19: 85-95.
- Lelis S, Howes A. 2008. A bayesian model of how people search online consumer reviews [J]. *Cognitive Science Society*, 553-558.
- Li J, Zhan L J. 2011. Online persuasion: How the written word drives WOM [J]. *Journal of Advertising Research*, 51(1): 239-257.
- Liu Y. 2006. Word-of-Mouth for movies: Its dynamics and impact on box office revenue [J]. *Journal of Marketing*, 70(3): 74-89.
- Mudambi S M, Schuff D. 2010. What makes a helpful online review? A study of customer reviews on Amazon.com [J]. *MIS Quarterly*, 34(1): 185-200.
- Pan Y, Zhang J Q. 2011. Born unequal: A study of the helpfulness of user-generated product reviews [J]. *Journal of Retailing*, 4(87): 598-612.
- Rabin M, Schrag J L. 1999. First impressions matter: A model of confirmatory bias [J]. *The Quarterly Journal of Economics*, 114(1): 37-82.
- Richins M L. 1983. Negative word-of mouth by dissatisfied consumers: A pilot study [J]. *Journal of Marketing*, 47(1): 68-78.
- Sen S, Lerman D. 2007. Why are you telling me this? An examination into negative consumer reviews on the web [J]. *Journal of Interactive Marketing*, 21(4): 76-94.
- Sikora R T, Chauhan K. 2012. Estimating sequential bias in online reviews: A Kalman filtering approach [J]. *Knowledge-Based Systems*, 27: 314-321.
- Staddon J, Chow R. 2008. Detecting reviewer bias through web-based association mining [C]. *Proceedings of the Second ACM Workshop on Information Credibility on the Web*, 5-10.
- Sundaram D S, Mitra K, Webster C. 1998. Word of mouth communications: A motivational analysis [J]. *Advances in Consumer Research*, 25: 527-531.
- Zhu F, Zhang X Q. 2010. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics [J]. *Journal of Marketing*, 74(2): 133-148.

Are Online Product Reviews Credible: The Deviation and Correction of Online Product Reviews' Distribution

Li Yujie^①, Liao Chenglin^①, Li Yi^②, Fu Hongyong^①,

① School of Economics and Business Administration, Chongqing University;

② School of Economics and Business Administration, Chongqing University of Posts and Telecommunications)

Abstract Based on the Chinese business environment, the paper firstly conducts an empirical analysis using Taobao's data. Results are that the online product review follows a reversed "L"-type distribution, of which the mean is a biased estimator of the product's true quality. These findings are not only different from the existed results which indicate that online product reviews follow a normal distribution, but also distinct from the empirical results of foreign e-business data. In order to study the reasons for the biased online product reviews and the features of the reversed "L"-type distribution, this investigation constructs a theoretical model to obtain the conditions for the mean of the online product reviews to be an unbiased estimator. And the authors further demonstrate that the "default good reviews" mechanism, "return evaluation" mechanism and the deviation of consumer's initiative evaluation are the main reasons for the biased online product reviews.

Key words Online product reviews, Product quality, Distribution, Deviation, Correction

专业主编：陈煜波

附录 1

密度函数为：

$$\lambda(r_i) = \begin{cases} 0 & r_i < \underline{\delta} \text{ 或 } \underline{\delta} < r_i < \bar{\delta} \\ \frac{\phi\left(\frac{r_i - q}{\sigma}\right)}{\sigma\left(1 - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\right)} & \bar{\delta} < r_i < \underline{\delta} \text{ 或 } \bar{\delta} < r_i < \bar{\delta} \text{ 或 } r_i > \bar{\delta} + (\bar{\delta} - \underline{\delta}) \\ \frac{\phi\left(\frac{r_i - q}{\sigma}\right) + \phi\left(\frac{(r_i - \bar{\delta} + \underline{\delta}) - q}{\sigma}\right)}{\sigma\left(1 - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\right)} & \bar{\delta} < r_i < \bar{\delta} + (\bar{\delta} - \underline{\delta}) \end{cases} \quad (A1)$$

均值 μ_r 为：

$$\begin{aligned} \mu_r = \int r_i \lambda(r_i) dr_i = & \int_{\underline{\delta}}^{\bar{\delta}} \frac{r_i \phi\left(\frac{r_i - q}{\sigma}\right)}{\sigma\left(1 - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\right)} dr_i + \int_{\bar{\delta}}^{\bar{\delta} + (\bar{\delta} - \underline{\delta})} \frac{r_i \phi\left(\frac{r_i - q}{\sigma}\right)}{\sigma\left(1 - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\right)} dr_i + \\ & \int_{\bar{\delta}}^{\bar{\delta} + (\bar{\delta} - \underline{\delta})} \frac{r_i \left(\phi\left(\frac{r_i - q}{\sigma}\right) + \phi\left(\frac{(r_i - \bar{\delta} + \underline{\delta}) - q}{\sigma}\right) \right)}{\sigma\left(1 - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\right)} dr_i + \end{aligned}$$

$$\int_{\bar{\delta} + (\bar{\delta} - \underline{\delta})}^{+\infty} \frac{r_i \phi\left(\frac{r_i - q}{\sigma}\right)}{\sigma \left(1 - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)\right)} dr_i =$$

$$q + \left((\bar{\delta} - \underline{\delta}) \left(\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \right) + \phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \sigma \right) \Lambda \quad (\text{A2})$$

$$\Lambda = 1 / \left(1 - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \right) \quad (\text{A3})$$

方差 σ_r^2 为:

$$\sigma_r^2 = \int r_i^2 \lambda(r_i) dr_i - \mu_r^2 =$$

$$\int_{\underline{\delta}}^{\bar{\delta}} \frac{r_i^2 \phi\left(\frac{r_i - q}{\sigma}\right)}{\sigma \left(1 - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right)\right)} dr_i + \int_{\bar{\delta}}^{\bar{\delta}} \frac{r_i^2 \phi\left(\frac{r_i - q}{\sigma}\right)}{\sigma \left(1 - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right)\right)} dr_i +$$

$$\int_{\bar{\delta}}^{\bar{\delta} + (\bar{\delta} - \underline{\delta})} \frac{r_i^2 \left(\phi\left(\frac{r_i - q}{\sigma}\right) + \phi\left(\frac{(r_i - \bar{\delta} + \underline{\delta}) - q}{\sigma}\right) \right)}{\sigma \left(1 - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right)\right)} dr_i +$$

$$\int_{\bar{\delta} + (\bar{\delta} - \underline{\delta})}^{+\infty} \frac{r_i^2 \phi\left(\frac{r_i - q}{\sigma}\right)}{\sigma \left(1 - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right)\right)} dr_i - \mu_r^2 =$$

$$(\underline{\delta} - q - 2(\bar{\delta} - \underline{\delta})\Psi)\sigma\Omega + (1 - \Omega^2)\sigma^2 + (\bar{\delta} - \underline{\delta})^2\Psi -$$

$$(\bar{\delta} - \underline{\delta})^2\Psi^2 + 2\sigma(\bar{\delta} - \underline{\delta})\Theta \quad (\text{A4})$$

其中,

$$\Psi = \left(\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \right) / \left(1 - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \right), \quad \Omega = \phi\left(\frac{\underline{\delta} - q}{\sigma}\right) / \left(1 - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \right),$$

$$\Theta = \left(\phi\left(\frac{\underline{\delta} - q}{\sigma}\right) - \phi\left(\frac{\bar{\delta} - q}{\sigma}\right) \right) / \left(1 - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \right) \quad (\text{A5})$$

利用中心极限定理:

$$\lim_{n \rightarrow \infty} \Pr \left[\frac{n^{1/2} (\bar{r}_n - \mu_r)}{\sigma_r} \leq x \right] = \Phi(x) \quad (\text{A6})$$

当在线商品评论的样本 n 足够大时,在线商品评论的均值 \bar{r}_n 和方差 σ_n^2 分别为:

$$\bar{r}_n = q + \left((\bar{\delta} - \underline{\delta}) \left(\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \right) + \phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \sigma \right) \Lambda \quad (\text{A7})$$

$$\sigma_n^2 = ((\underline{\delta} - q - 2(\bar{\delta} - \underline{\delta})\Psi)\sigma\Omega + (1 - \Omega^2)\sigma^2 + (\bar{\delta} - \underline{\delta})^2\Psi - (\bar{\delta} - \underline{\delta})^2\Psi^2 + 2\sigma(\bar{\delta} - \underline{\delta})\Theta) / n \quad (\text{A8})$$

$$\text{其中 } \Lambda = 1 / \left(1 - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \right), \quad \Psi = \left(\Phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) \right) \Lambda, \quad \Omega = \left(\frac{\phi\left(\frac{\underline{\delta} - q}{\sigma}\right)}{\sigma} \right) \Lambda,$$

$$\Theta = \left(\phi\left(\frac{\underline{\delta} - q}{\sigma}\right) - \phi\left(\frac{\bar{\delta} - q}{\sigma}\right) \right) \Lambda. \quad (\text{A9})$$

附录 2

密度函数为：

$$\lambda(r_i) = \begin{cases} 0 & \underline{\delta} < r_i < \bar{\delta} \\ \frac{\phi\left(\frac{r_i - q}{\sigma}\right)/\sigma}{1 + \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)} & -\infty < r_i < \underline{\delta} \text{ 或 } \bar{\delta} < r_i < +\infty \end{cases} \quad (\text{A10})$$

均值 μ_r 为：

$$\begin{aligned} \mu_r &= \int r_i \lambda(r_i) dr_i = \\ &= \int_{-\infty}^{\underline{\delta}} \frac{r_i \phi\left(\frac{r_i - q}{\sigma}\right)/\sigma}{1 + \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)} dr_i + \int_{\bar{\delta}}^{+\infty} \frac{r_i \phi\left(\frac{r_i - q}{\sigma}\right)/\sigma}{1 + \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)} dr_i = \\ &= q + \frac{\phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \phi\left(\frac{\underline{\delta} - q}{\sigma}\right)}{1 + \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)} \sigma \end{aligned} \quad (\text{A11})$$

利用中心极限定理,式(A6)：

$$\lim_{n \rightarrow \infty} \Pr \left[\frac{n^{1/2}(\bar{r}_n - \mu_r)}{\sigma_r} \leq x \right] = \Phi(x)$$

当在线商品评论的样本 n 足够大时,在线商品评论的均值 \bar{r}_n 为：

$$\bar{r}_n = q + \frac{\phi\left(\frac{\bar{\delta} - q}{\sigma}\right) - \phi\left(\frac{\underline{\delta} - q}{\sigma}\right)}{1 + \Phi\left(\frac{\underline{\delta} - q}{\sigma}\right) - \Phi\left(\frac{\bar{\delta} - q}{\sigma}\right)} \sigma \quad (\text{A12})$$