

付费问答社区用户围观行为研究

赵庆亮¹ 王培勇¹ 陈吉²

(1. 北京化工大学经济管理学院, 北京 100029; 2. 国家档案局档案科学技术研究所, 北京 100050)

摘要: 付费围观是知识付费领域重要的创新模式, 研究用户付费围观行为影响因素对该模式发展具有重要意义。本研究采集问答平台付费围观相关数据, 使用文本分析技术挖掘问题内容, 基于信号理论和信息处理双过程理论构建用户付费围观行为影响因素研究模型, 并分析各因素的影响效果。研究发现, 回答价格、回答被赞同次数、回答长度、相似问题被围观量和回答者声誉显著影响付费围观量, 回答者声誉对回答价格有调节作用。本研究将为社区运营者优化付费围观产品设计、合理开展社区运营提供指导。

关键词: 问答平台; 知识付费行为; 信号理论

中图分类号: C931.6

DOI: 10.3772/j.issn.1673-2286.2019.11.002

Web2.0以来, 用户内容生成 (User Generated Content) 模式不断发展, 借助在线问答社区社交化网络, 拥有认知盈余的知识生产群体能够将存在于脑海中排他性的隐性知识传播给知识需求方 (“知识”是指供需双方存在的信息差), 在线问答社区已成为用户信息搜寻的重要渠道。但由于审核机制的缺失, 网络信息同质化日趋严重, 信息质量参差不齐, 信息过载导致用户搜寻信息的时间成本增加。近年来, 随着消费结构升级、知识产权保护加强以及移动支付普及, 用户为优质信息付费以降低搜寻成本的意愿日益提升, 连接知识供给方和需求方的问答平台由此进入以付费制为核心的知识共享3.0阶段。该阶段对社会中分散的盈余知识技能和智力资源进行整合, 以付费的方式将信息共享、传递给社会大众或特定平台^[1]。

2016年以来, 知识付费产品和付费模式层出不穷, 付费围观是在行一点平台率先推出的知识付费创新模式, 微博问答上线的围观、知乎平台上线的旁听, 本质上都属于付费围观模式。与需要向单一回答者支付大额费用以获取回答的付费提问不同, 付费围观是指问答社区用户通过支付小额固定费用即可查看提问者以公开方式提问并被回答的问题, 其运营模式如图1所示。付费围观实现了对问答平台潜水用户“搭便车”行为的收费, 问答成为可被重复销售的产品, 围观者付出小额

费用即可获取自己感兴趣的内容, 提问者和回答者能够获取问答后续产生的经济收益。尤其是对提问者, 提出“好的问题”可以获得经济回报的预期会正向显著影响其付费提问意愿^[2], 问题是付费问答社区内容触发器和知识变现主要来源, 提升问题数量和质量对社区持续运营具有重要意义。在行一点在2016年5月上线后的42天内, 100万用户为内容产品买单, 重购率达43%^[3], 但随着对付费产品新奇感和体验感减弱, 用户由早期冲动消费向理性消费转变, 知识付费产品打开率和重购率出现缩水^[4]。付费问答社区的核心是用户, 因此洞悉用户行为态度, 了解用户消费特征, 对促进付费围观的发展至关重要^[5]。本文以在行一点平台为研究对象, 基于信号理论和信息处理双过程理论展开对用户付费围观行为影响因素的研究。本研究是对知识付费模式下用户行为研究的补充, 将对社区运营者优化付费围观产品设计、合理利用社区资源开展社区运营工作提供理论性指导。

1 相关研究与理论基础

1.1 付费问答社区付费围观行为研究

现有付费问答社区用户行为研究主要集中在付费

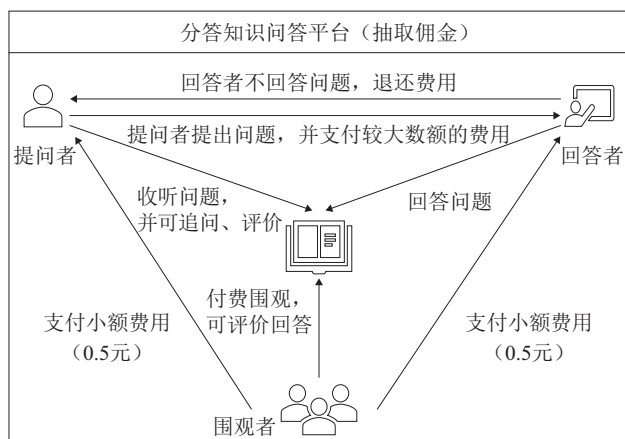


图1 付费问答社区中付费围观及付费提问运营模式——以在行一点平台为例

提问模式下用户行为影响因素的研究，对规模更为庞大的潜水用户付费围观行为影响因素探讨相对较少。付费围观和付费提问存在不完全替代关系，两者都可以满足用户知识获取需求，但付费围观是作为付费提问的价格补充机制存在，即用户付出更低的成本便可获得已有的问答，用户付费意愿影响因素也有所不同^[2]。张杨焱等^[6]基于感知价值理论研究了用户持续参与付费围观意愿的影响因素，其研究表明付费围观者对回答者的信任会对感知价值产生显著影响，而感知价值会进一步影响用户持续付费围观意愿。Fang等^[7]基于计划行为理论研究了问答社区用户持续参与行为的影响因素，其研究表明，心理层面上的知识需求对围观者持续参与行为有积极影响，而经济因素对其并没有起到正向作用；社交层面上，对平台忠诚度、共同愿景、共同交流方式和互惠规范对围观者持续使用问答社区有显著正向影响；功能层面上，感知互补性、感知相容性和感知外部威望对感知有用性有显著正向影响。李武等^[8]发现用户付费围观行为受到回答者粉丝量、回答价格和回答长度的影响，而回答价格在回答者粉丝量与付费围观量之间具有序列中介作用。上述研究多以问卷形式收集用户数据，仅李武采用网络数据对用户付费围观行为进行了研究，问卷调查数据收集成本大且存在不一致的风险，而网站用户行为数据能够客观地反映用户实际参与行为，这为洞察用户行为影响因素带来机遇和挑战。

1.2 信号理论

知识付费市场是典型的信息不对称市场，围观者除

非亲身体验否则不能对回答效用产生准确认知，其是在信息不确定性条件下对知识产品效用进行评估从而做出购买决策^[9]，因而交易过程中可能存在逆向选择，即由于交易双方存在信息不对称导致劣质信息内容驱逐优质信息内容，进而导致知识市场信息内容平均质量下降。信号理论由美国经济学家Spence^[10]提出，用以解释在信息不对称市场下交易双方是如何通过信号传递产品质量信息以减少信息不对称所产生的影响。信号理论指出，拥有高质量产品的卖主会发出标志产品质量的信号，且其发出信号的成本是低于低质量产品卖主的信号成本，而缺少信息的一方通过该信号推测产品的真实质量，从而达成最终交易^[11]。近年来，信号理论被广泛地应用于电子商务^[12]、政策评价^[13]、网络评价^[14]等领域。

1.3 信息处理双过程理论

James^[15]提出人类有逻辑思考与直觉（或推理系统和经验系统）两种不同的认知方式，由此提出人类信息处理双过程理论，该理论在信息处理研究领域得到广泛的应用和扩展，为解释信息处理、采纳行为提供了理论基础。以信息处理双过程理论为基础的信息处理模型中，详尽可能性模型被广泛地应用于个体信息处理研究，该模型认为个体在处理信息时有核心路线和边缘路线。核心路线是指用户具有高度动机，其关注的核心是信息内容本身。而边缘路线是指用户借助信息内容之外的线索来形成对信息有效性的认知，如信息来源可靠性、信息标的等。采用两种路线的相对权重是由个体动机和认知能力所决定的。

2 理论分析与研究假设

2.1 理论模型

本研究聚焦付费问答社区付费围观模式下用户围观行为影响因素的探索。由于知识产品交易中存在信息不对称问题，围观用户可通过对问答产品相关信号的观察形成对信息质量的判断进而做出付费决策。本文依据详尽可能性模型，从用户认知的核心路线和边缘路线出发，探索问答信息本身特征及信息来源特征对付费围观量的影响。本文研究因素包括回答价格、回答赞同数、回答时长、相似问题被围观量、提问者声誉、回答者声誉及其与价格的交互作用，研究模型如图2所示。

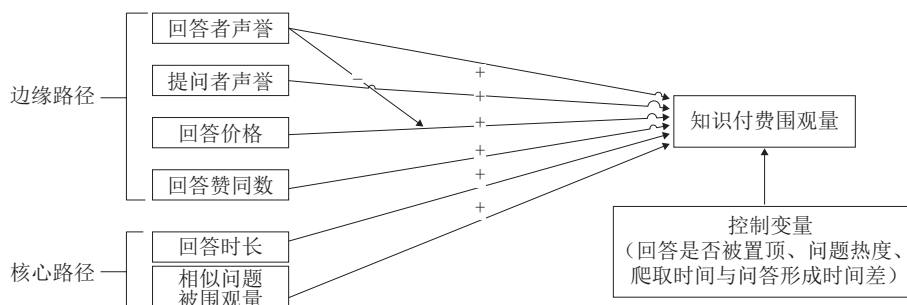


图2 研究模型

2.2 研究假设

价格是影响用户知识产品消费决策的重要因素^[16]，人们一般认为高价格代表高质量产品，价格与质量存在正相关性，而“便宜”意味着产品质量不佳^[17]。无论问答本身价格如何，付费围观者仅需支付小额费用即可查看问答内容，此时回答价格被围观者视为问答质量信号，问答设定的价格越高，用户越会产生“占便宜”心理。锚定效应理论认为，人们在面对不确定状态判断时，其最终决策会受到最初信息的影响^[18]，回答价格作为锚定值会对用户付费围观决策产生正向影响。基于上述分析，本文提出假设H1：回答的价格越高，该问题被付费围观的可能性越大。

问答所积累的网络用户口碑是判断回答内容信息质量优劣的重要依据^[19]。网络用户口碑是指用户对产品形成的正面或负面评价，是用户通过互联网完成的关于产品质量的信息交换^[20]。电子商务领域的研究表明，已消费用户对产品的反馈会对潜在用户购买决策产生影响^[21]，方爱华等^[22]的研究发现知识产品的口碑能够调节感知价值对付费意愿的影响。因为知识问答所形成的答案较为简短、易于传播，为保护已付费用户对付费知识所享有的权益，平台不设置文本评论，已围观用户通过点赞的方式表达对回答内容的认可。回答被赞同次数是该回答受欢迎的正向信号，回答收到正向反馈的数量越高，意味着已围观用户越肯定该回答的质量。基于上述分析，本文提出假设H2：回答被赞同次数越多，该问题被付费围观的可能性越大。

用户会观察回答本身质量信号衡量付费决策，研究者多从文本特征（如长度、内容词密度）和非文本特征（评论数、回答数）两个方面衡量答案质量^[23-24]。Chen等^[25]发现提问者支付费用的高低与回答者提供回答的长度呈正向相关性，在行一点中回答者是以语音形式完

成回答，回答时长可作文本内容长度的近似替代。回答时长越长，意味着回答者投入的精力越多，答案所包含信息量越多，潜在围观者会以回答长度作为评判回答质量的信号，回答长度越长，该回答在同类问题中的质量就越高，用户付费围观的可能性越大。基于上述分析，本文提出假设H3：回答的时长越长，该问题被付费围观的可能性越大。

从内容角度对问题进行评估能够更加全面地了解问题对付费围观量的影响。目前，问答社区中以问题为中心的研究多集中于问题类型分类^[26]和相似问题发现^[27]。虽然在行一点平台尚未实现相似问题推荐，用户仍可通过主动搜寻模式寻找相似问题。本文认为相似问题被围观的数量能够在一定程度上代表该问题在社区中的流行程度和用户对该类问题的围观意愿，问题在社区中的流行程度越高、听众对该类型问题的围观量越高，该问题被用户付费围观的可能性越大。基于上述分析，本文提出假设H4：相似问题的围观量越高，该问题被付费围观的可能性越大。

基于详尽可能性模型的研究认为，信息来源是判断用户信息质量的重要依据^[28]，当问答社区用户缺少判断信息本身质量的途径时，信息来源可靠性会影响用户决策。声誉是问答社区用户分享知识、参与社区活动以及被社区成员认可的综合度量，能够反映用户在社区所处的地位，Till^[29]的研究证实知名度与消费者信任存在直接关系，名人会被认为是可靠的信息推荐者。张杨姝等^[6]的研究也表明，信息提供方的知名度、专业性与消费者的信任有直接的关系，而信任关系会直接影响消费者购买知识产品的决策。

问答社区中问答双方信息（如关注者人数、回答问题情况等）是公开的，潜在围观用户可通过观察这些信息形成对问答双方声誉的认知。沈宇飞等^[30]指出，用户知识分享答案获得的点赞数、回答问题数量可以用来

衡量社区用户声誉。Jin等^[31]的研究表明,用户所具有的社会资本对用户知识贡献有着显著的正向影响,而社会资本可以用关注者数量、被提问的数量来表征。在行一点中,用户被关注数、用户回答问题数量和用户已回答问题累计收听量反映了其在社区中所具备的影响力、认可度和其贡献的内容质量,可以用来衡量用户在社区中的声誉。社区用户具有的声誉越高,其所参与形成的问答被围观的可能性越大,一方面提问者声誉越高,用户认为提问者参与社区的程度更深、选择答主的能力更强,因此获得的信息更有保障;另一方面回答者声誉越高,用户认为回答者的专业性和知名度越高,其贡献内容的质量应当更高。此外,基于信息流的推送模式会使得声誉高的用户问答更容易被曝光,更多的曝光机会意味着更多的销售机会。基于上述分析,提出假设H5A:回答者的声誉越高,该问题被付费围观的可能性越大。H5B:提问者的声誉越高,该问题被付费围观的可能性越大。

在线问答社区中,用户可以很方便地了解到回答者在社区所积累的声誉,当用户对知识产品提供方有较为深入了解时,便不会对质量和价格不相符的知识产品付费^[9]。具有高声誉的回答者所提供的回答似乎更能得到用户的青睐,例如,平台“大V”回答的问题能够得到数以万计的收听量,由此所产生的围观收益已经远超回答价格所带来的收益。当回答者声誉较高时,一方

面,反映了其所具有的独特经验受到用户的追捧,知识市场中粉丝经济会对知识产品销量产生影响^[8];另一方面,回答价格对潜在付费围观者决策所产生的影响将会减弱,回答价格作为回答质量信号时会受到其他信号的调节^[12],即回答者的声誉较高时,“价格-质量”模式可能会受到调节。基于上述分析,本文提出假设H6:回答者声誉会对回答价格起到调节作用,回答者声誉越高,回答价格对付费围观量的影响会减弱。

3 数据收集与处理

3.1 数据收集

本研究编写爬虫程序于2019年9月19日对付费问答社区在行一点问答模块下答主页和问答详情页面进行爬取。首先对网站各分类下认证的专家用户信息进行采集,然后对获得的专家用户已回答问题的提问者信息再次抓取,共计采集到81 639名用户信息。网站限制只能按照默认、最新和热门3种排序方式分别查看答主100条以内回答,因此只获取到单个用户每种排序下最多100条记录。数据预处理后,删除免费围观记录、存在缺失值的记录和重复记录后,共得到151 811条记录用于本研究,变量描述性统计结果如表1所示。

表1 变量统计性描述

变 量	变量符号	最小值	最大值	平均值	标准差
围观量	$Listen_count_i$	0	26 168	47.28	248.73
回答者回答问题量	$A_answercount_i$	0	27 456	299.69	805.68
回答者粉丝量	$A_follower_i$	0	237 276	4 122.33	18 412.45
回答者回答累计被围观量	$A_listened_i$	0	1 399 652	12 911.60	62 066.79
回答者声誉(标准化)	$A_reputation_i$	0	1	0.02	0.07
提问者回答问题量	$Q_answercount_i$	0	5 295	25.75	164.15
提问者粉丝量	$Q_follower_i$	0	237 276	228.73	3 621.75
提问者回答累计被围观量	$Q_listened_i$	0	954 944	702.56	8 204.50
提问者声誉(标准化)	$Q_reputation_i$	0	1	0.00	0.02
问题价格	$Price_i$	0	4 999	19.75	88.99
回答被赞同次数	$Vote_up_i$	0	644	1.14	6.39
回答时长	$Duration_i$	1	60	50.35	14.42
相似问题平均围观量	$Listen_mean_i$	0	4 612.40	31.97	82.95
问题热度	$Genre_i$	0.06	0.35	0.24	0.09
是否置顶	Is_sticky_i	0	1	0.15	0.12
回答形成到爬取时间点的间隔	$Timediff_i$	1	1 184	1 045.62	235.97

3.2 变量定义及测量

3.2.1 因变量

本研究因变量是回答被围观收听的数量 ($Listen_count_i$)。在行一点平台会记录问题被回答后用户付费围观收听的数量,同一用户多次收听仅被记录为一次收听,因此该数值反映回答被付费围观收听的人数。

3.2.2 自变量

本文自变量有回答者声誉 ($A_reputation_i$)、提问者声誉 ($Q_reputation_i$)、回答价格 ($Price_i$)、回答被赞同次数 ($Vote_up_i$)、回答时长 ($Duration_i$) 和相似问题被围观量 ($Listen_mean_i$)。其中,回答价格、回答被赞同数和回答时长可直接从网站获取,对声誉和相似问题被围观量需计算,本文做如下处理。

(1) 回答者和提问者声誉。在行一点社区未设置明确的声誉计算方式,本文综合用户被关注数、用户回答问题数量和用户已回答问题累计收听量来衡量回答者声誉和提问者声誉。具体而言,对用户被关注数、用户回答问题数量和已回答问题累计收听量分别进行标准化后求和再进行标准化。

(2) 相似问题平均围观量。在行一点平台未设置问题推荐系统,本研究采用自然语言处理方法寻找指定问题的相似问题。深度学习算法被广泛地应用于文本处理,Doc2vec模型是由Mikolov在词向量表述模型Word2vec的基础上扩展的语句级文本向量化模型^[32]。Doc2vec模型将句子表示为空间中向量,通过计算各句子间的余弦值即可以表示句子间相似度,该模型被广泛用于专利文件相似度检验^[33]、期刊论文热点选题识别^[34]等以文本相似度为基础的研究。本研究通过对所获得语句进行预处理(分词、去除停用词等),构建Doc2vec模型并计算各个问题最相似的5个问题,对这5个问题的付费围观量取平均值后作为本文自变量。

3.2.3 控制变量

本文关注的控制变量有回答是否置顶 (is_sticky_i)、爬取时间与问答建立时间差 ($Timediff_i$) 和问题热度 ($Genre_i$)。前两项可从网站直接获取。在行一点平台未从问题层面对问答类别细分,本文采用结合回答者

类别和文本主题分类的方法归并问题类别,从而计算问题热度。首先形成各个问题的细分类别标签Tag,对每个问题分别采用百度文本分类模型形成细分类别标签Tag_1和按平台标注的回答者所属类别形成细分类别标签Tag_2,最终问题细分类别标签Tag的确定如下:①若问题回答者是非认证用户,则Tag=Tag_1;②若问题回答者是认证用户,依据经验规则进行匹配,如Tag_1=“科普”且Tag_2=“外科手术”,则Tag=“健康”;③对于匹配规则之外的未标注问答,采用人工标注的方式形成类别标签Tag。然后合并细分类别形成主类别,如“音乐”“娱乐”“时尚”归类于“兴趣”主类别下,最终得到“财经”“健康”“教育”“社会”和“兴趣”五大问题类别;各个类别与总数相比即可得到不同类别的问题热度,如表2所示。

表2 各个问题类别的热度

类 别	频 次	占比/%
财经	31 586	20.80
健康	23 491	15.50
教育	9 675	6.40
社会	53 332	35.10
兴趣	33 727	22.20

3.3 变量间相关性分析

在进行OLS线性回归前,对部分偏态变量进行对数变换后标准化,使之符合回归模型要求,变换后得到的各个变量间的相关性见表3。

4 研究模型与结果分析

本文采用分层回归方法对主效应和调节效应进行研究^[35],首先建立仅包含控制变量的模型一。然后,建立包含自变量直接作用和控制变量的模型二。有研究指出统计上显著性与真实显著性有所区别,在使用大样本数据回归分析时,变量间较小的相关性也可能导致统计上显著^[36]。由表3可知,提问者声誉与围观量的相关性为0,且不显著,本文尝试建立剔除提问者声誉的模型三。最后,在模型三的基础上加入价格和回答者声誉交叉项构建模型四。4个模型的参数估计结果如表4所示。

为进一步验证回答者声誉对价格的调节作用,本

表3 变量间相关性分析

		1	2	3	4	5	6	7	8	9	10
1	ln(围观量)	1.00									
2	ln(回答者声誉)	0.34**	1.00								
3	ln(提问者声誉)	0.00	-0.02**	1.00							
4	ln(回答被赞同次数)	0.47**	0.29**	0.03**	1.00						
5	ln(回答时长)	0.17**	0.07**	0.00	0.14**	1.00					
6	ln(回答价格)	0.41**	0.31**	-0.08**	0.14**	0.13**	1.00				
7	ln(相似问题围观量)	0.07**	0.05**	-0.02**	0.03**	0.08**	0.09**	1.00			
8	是否置顶	0.01**	-0.03**	0.02**	0.05**	0.02**	-0.04**	-0.01**	1.00		
9	爬取时间间隔	0.15**	-0.26**	0.04**	-0.02**	-0.08**	-0.18**	-0.04**	0.03**	1.00	
10	问题流行度	-0.01**	0.01**	0.02**	0.02**	-0.11**	-0.05**	-0.03**	0.01**	0.10**	1.00

注: ** $p<0.01$, * $p<0.05$

表4 模型参数估计结果

变量名	模型一			模型二			模型三			模型四		
	β	T	P	β	T	P	β	T	P	β	T	P
ln(回答者声誉)	-	-	-	5.86**	64.34	0.00	5.86**	64.35	0.00	13.87**	77.11	0.00
ln(提问者声誉)	-	-	-	0.82**	4.77	0.00	-	-	-	-	-	-
ln(回答价格)	-	-	-	0.44**	155.20	0.00	0.44**	155.67	0.00	0.49**	172.96	0.00
ln(回答被赞同次数)	-	-	-	0.93**	168.90	0.00	0.93**	169.38	0.00	0.94**	175.28	0.00
ln(回答时长)	-	-	-	0.27**	40.69	0.00	0.27**	40.73	0.00	0.25**	37.99	0.00
ln(相似问题围观量)	-	-	-	0.03**	11.07	0.00	0.03**	10.99	0.00	0.03**	10.35	0.00
问题热度	-4.27**	-9.22	0.00	-0.39**	-11.09	0.00	-0.39**	-11.05	0.00	-0.37**	-10.67	0.00
是否置顶	0.10**	2.66	0.01	0.05	1.53	0.13	0.05	1.58	0.11	0.08**	2.75	0.01
爬取时间间隔	0.00**	56.56	0.00	0.02**	129.90	0.00	0.00**	130.01	0.00	0.08**	136.07	0.00
ln(回答价格)* ln(回答者声誉)	-	-	-	-	-	-	-	-	-	-2.03**	-56.65	0.00
常数	-	-	-	-1.946**	-60.99	0.00	-1.95**	-60.94	0.00	-2.07**	-65.50	0.00
Adjust R_squared	0.02			0.43			0.43			0.45		

注: ** $p<0.01$, * $p<0.05$

文对回答者声誉按高低分组($M\pm 0.5SD$ 为标准)后绘制调节作用示意图如图3所示。当声誉较低时,价格对围观量有正向影响($\beta=0.41, p=0.00$),当声誉较高时,价格对围观量仍有正向影响($\beta=0.37, p=0.00$),但斜率降低,这表明价格对围观量的正向效应随回答者声誉提高而降低。

由表4可知,模型二调整后 R^2 为0.43,较模型一有较大提升,说明自变量的确对因变量起到了重要的解释作用。模型二中自变量均显著,控制变量除问答是否置顶不显著外,其他控制变量均显著。模型三剔除与付费围观量相关性不显著的自变量(提问者声誉),对模型重新回归后,新模型解释力未发生明显变化,这表明提

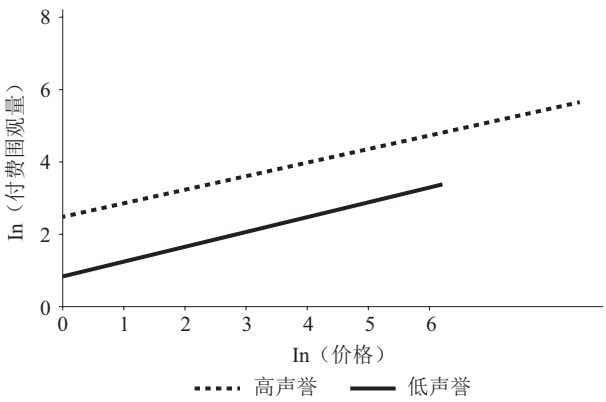


图3 回答者声誉对价格的调节作用示意图

问者声誉对围观量所产生的影响微弱,可不予考虑,即H5B在统计学意义上成立,但在实际中参考意义不大。

模型三中,价格($Price_i$)系数为正,且影响显著,H1成立,即回答者付出的精力和成本越高,回答质量相应的会更高,回答者会设置更高的价格来传递回答内容的信息质量,因此回答价格作为质量信号的作用得到验证。回答被赞同次数($Vote_up_i$)系数为正,且影响显著,H2成立,回答被赞同的次数越多,已购围观者对该回答质量的整体认可度高,潜在围观者出于从众心理选择对更多人产生高效用的问答进行付费围观。回答时长($Duration_i$)的系数为正,且影响显著,H3成立,这与Zhang等^[37]对信息长度和信息质量之间相关性研究一致,在无法获取回答内容的情形下,回答的长度是回答质量的可靠度量。相似问题被围观量($Listen_mean_i$)的系数为正,且影响显著,H4成立,相似问题被更多地围观,说明该类型的问题在社区中得到了更多用户的关注,这表明用户有为与已购问题类型相似问题付费的

意愿,潜在围观用户数量越多,问题被付费的可能性越大。回答者声誉($A_reputation_i$)的系数为正,且影响显著,H5A成立,回答者通过贡献自己的知识和经验得到社区中其他用户的信任,用户在缺少对回答内容本身特征了解时,信息提供方的可靠程度会降低其付费决策过程中的不确定性,用户更容易信任声誉高的回答者。

模型四中加入价格和回答者声誉的交叉项,调整后 R^2 为0.45,模型解释力更强。该交叉项变量显著,系数为负,H6成立。由于在行一点采用回答者先设定回答价格,提问者再进行提问的模式,因此在一段时期内,某一回答者所有回答问题的价格是不变的,当围观者进行付费决策时,回答者的声誉信号比价格信号的可信程度更高。围观者很容易获取到回答者在社区中的历史贡献和成就,当围观者对回答者声誉的认可度高时,其对价格和质量之间关系的认知会受到影响。对本文假设及结论如表5所示。

表5 假设及结论

假设编号	假设内容	结 论
H1	回答的价格越高,问题被付费围观的可能性越大	成立
H2	回答被赞同次数越多,问题被付费围观的可能性越大	成立
H3	回答的时长越长,该问题被付费围观的可能性越大	成立
H4	相似问题的围观量越高,问题被付费围观的可能性越大	成立
H5A	回答者的声誉越高,问题被付费围观的可能性越大	成立
H5B	提问者的声誉越高,问题被付费围观的可能性越大	不成立
H6	回答者声誉会对回答价格起到调节作用,回答者声誉越高,回答价格对付费围观量的影响会减弱	成立

5 结论与讨论

本研究基于信号理论和信息处理双过程理论构建了付费问答社区用户付费围观行为影响因素研究模型,探讨了这些因素如何对付费围观量产生影响。研究发现回答的价格、回答被赞同的次数、回答时长、相似问题被围观数量和回答者声誉会对围观量产生正向影响,而回答者声誉会对回答价格起到调节作用,回答者的声誉越高,回答价格对围观量的正向影响随之减弱。本研究在理论层面上丰富了付费围观领域的学术研究,且对问答社区运营优化也有一定指导性意义。本研究现实意义和结论如下。

第一,回答者声誉对付费围观量有正向影响,而提问者声誉影响并不显著。无论是对理性围观者还是非

理性围观者,回答者声誉是付费决策的重要依据,而对提问者声誉关注度不高。第二,相似问题围观量会正向影响目标问题围观量。本研究对问题从内容角度进行量化研究,采用文本相似度算法寻找目标问题最相似的5个问题,并将其量化进入模型,研究结果表明用户对感兴趣问题的相似问题具有更强的付费围观意愿,这与用户自身的信息需求特征有关。在社区运营层面,设置问答相似问题推荐系统引导用户更加便捷地发现其感兴趣的问题,能够优化用户信息搜寻路径,改善用户体验。第三,回答者声誉对价格产生负向调节作用,回答者声誉越高,价格作为质量信号对付费围观量的正向影响减弱。这表明回答者声誉足够高时,围观者对其回答价格的关注程度降低,现阶段围观行为有明显的粉丝经济特点。

本研究仍有缺陷和不足之处。首先,本文研究对象限定在在行一点,研究结论对其他有付费围观功能的平台是否适应还需进一步验证。其次,本研究使用截面数据,一方面变量存在内生性,如付费围观量可能会对回答被赞同次数产生影响;另一方面在用户付费决策过程中,该回答历史围观量会对其付费决策产生影响,而截面数据无法反映该影响过程。因此,本研究结果应被视为探索性研究结果,此方面的研究有必要继续收集面板数据进行更加深入的探讨。最后,本文采用爬虫技术获取到大规模问答数据集和用户数据集,在线数据能够真实客观地反映用户行为,但是无法从认知层面对用户付费围观行为进行深入研究。

参考文献

- [1] 王传珍. 知识付费奇点与未来[J]. 互联网经济, 2017(1): 68-73.
- [2] 李武, 艾鹏亚, 许耀心. 在线付费问答平台的用户付费模式及付费意愿研究[J]. 图书情报工作, 2018, 62(13): 24-29.
- [3] 周昶帆. 狂飙42天, 朋友圈现象级产品分答是怎样做出的?[EB/OL]. (2016-07-07) [2019-09-01]. <https://36kr.com/p/5049218.html>.
- [4] 艾媒报告. 2018-2019中国知识付费行业研究与商业决策分析报告[R/OL]. [2019-09-09]. <https://www.iimedia.cn/c400/63439.html>.
- [5] 李武. 在线知识付费平台: 何去何从?[J]. 图书情报知识, 2018(3): 2.
- [6] 张杨隼, 彭子健, 刘齐凯. 问答平台用户付费围观持续参与意愿的影响因素[J]. 图书馆论坛, 2018, 38(6): 86-94.
- [7] FANG C, ZHANG J. Users' continued participation behavior in social Q&A communities: a motivation perspective[J]. Computers in Human Behavior, 2019, 92: 87-109.
- [8] 李武, 艾鹏亚, 宾锋. 粉丝力量与知识付费: 在线问答平台用户付费围观行为研究[J]. 图书馆杂志, 2019, 38(4): 74-81.
- [9] 蔡舜, 石海荣, 傅馨, 等. 知识付费产品销量影响因素研究: 以知乎Live为例[J]. 管理工程学报, 2019, 33(3): 71-83.
- [10] SPENCE M. Signaling in retrospect and the informational structure of markets[J]. American Economic Review, 2002, 92(3): 431-459.
- [11] 张维迎. 博弈论与信息经济学[M]. 上海: 上海人民出版社, 1999.
- [12] XU Y, CAI S, KIM H W. Cue consistency and page value perception: Implications for web-based catalog design[J]. Information & Management, 2013, 50(1): 33-42.
- [13] 夏清华, 何丹. 政府研发补贴促进企业创新了吗——信号理论视角的解释[J/OL]. 科技进步与对策: 1-10 [2019-11-19]. <http://kns.cnki.net/kcms/detail/42.1224.G3.20191119.1321.022.html>.
- [14] NAM S, MANCHANDA P, CHINTAGUNTA P K. The effect of signal quality and contiguous word of mouth on customer acquisition for a video-on-demand service[J]. Marketing Science, 2010, 29(4): 690-700.
- [15] JAMES W, BURKHARDT F, SKRUPSKELIS I K. The principles of psychology[M]. Beijing: China Social Sciences Publishing House, 1999.
- [16] VERMA D P S, GUPTA S. Does higher price signal better quality?[J]. Vikalpa, 2004, 29(2): 67-78.
- [17] LICHTENSTEIN D R, RIDGWAY N M, NETEMEYER R G. Price perceptions and consumer shopping behavior: a field study[J]. Journal of Marketing Research, 1993, 30(2): 234-245.
- [18] 李斌, 徐富明, 王伟, 等. 锚定效应的研究范式: 理论模型及应用启示[J]. 应用心理学, 2008, 14(3): 269-275, 281.
- [19] GODES D, MAYZLIN D. Using online conversations to study word-of-mouth communication[J]. Marketing science, 2004, 23(4): 545-560.
- [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(1): 38-52.
- [21] DELLAROCAS C. The digitization of word of mouth: Promise and challenges of online feedback mechanisms[J]. Management Science, 2003, 49(10): 1407-1424.
- [22] 方爱华, 陆滕滕, 刘坤锋. 虚拟社区用户知识付费意愿实证研究[J]. 图书情报工作, 2018, 62(6): 105-115.
- [23] FU H, WU S, OH S. Evaluating answer quality across knowledge domains: using textual and non-textual features in social Q&A[C]// Proceedings of the 78th ASIS & T Annual Meeting: Information Science with Impact: Research in and for the Community. American Society for Information Science, 2015: 88.
- [24] 李晨, 巢文涵, 陈小明, 等. 中文社区问答中问题答案质量评价和预测[J]. 计算机科学, 2011, 38(6): 230-236.
- [25] CHEN Y, HO T H U A, KIM Y M I. Knowledge market

- design: a field experiment at Google Answers [J]. Journal of Public Economic Theory, 2010, 12 (4): 641-664.
- [26] HARPER F M, MOYD, KONSTAN J A. Facts or friends? distinguishing informational and conversational questions in social Q&A sites [C] //Proceedings of the Sigchi Conference on Human Factors in Computing Systems. ACM, 2009: 759-768.
- [27] WANG K, MING Z, CHUA T S. A syntactic tree matching approach to finding similar questions in community-based qa services [C] //Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2009: 187-194.
- [28] SUSSMAN S W, SIEGAL W S. Informational influence in organizations: an integrated approach to knowledge adoption [J]. Information Systems Research, 2003, 14 (1): 47-65.
- [29] TILL B D. Using celebrity endorsers effectively: lessons from associative learning [J]. Journal of Product & Brand Management, 1998, 7 (5): 400-409.
- [30] 沈宇飞, 廖博, 徐扬. 社会化问答社区中声誉系统对知识分享活动的影响研究 [J]. 情报学报, 2018, 37 (11): 1154-1163.
- [31] JIN J, LI Y, ZHONG X, et al. Why users contribute knowledge to online communities: an empirical study of an online social Q&A community [J]. Information & Management, 2015, 52 (7): 840-849.
- [32] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space [EB/OL]. [2019-09-09]. <https://arxiv.org/abs/1301.3781>.
- [33] 曹祺, 赵伟, 张英杰, 等. 基于 Doc2Vec 的专利文件相似度检测方法的对比研究 [J]. 图书情报工作, 2018, 62 (13): 74-81.
- [34] 阮光册, 夏磊. 基于Doc2Vec的期刊论文热点选题识别 [J]. 情报理论与实践, 2019, 42 (4): 107.
- [35] BILLINGS R S, WROTE S P. Use of path analysis in industrial/organizational psychology: Criticisms and suggestions [J]. Journal of Applied Psychology, 1978, 63 (6): 677.
- [36] 艾尔·巴比. 社会研究方法: 第11版 [M]. 邱泽奇, 译. 北京: 华夏出版社, 2009: 465-466.
- [37] ZHANG L, PENG T Q, ZHANG Y P, et al. Content or context: Which matters more in information processing on microblogging sites [J]. Computers in Human Behavior, 2014, 31: 242-249.

作者简介

赵庆亮, 男, 1975年生, 副教授, 通信作者, 研究方向: 用户行为分析、知识管理, E-mail: zhaoql@mail.buct.edu.cn。
王培勇, 男, 1995年生, 硕士研究生, 研究方向: 用户行为分析、知识管理。
陈吉, 男, 1994年生, 助理馆员, 研究方向: 档案信息化、知识管理。

Research on the Behavior of Users' Watching on Paying Q&A Platform

ZHAO QingLiang¹ WANG PeiYong¹ CHEN Ji²

(1. The college of Economics and Management, Beijing University of Chemical Technology, Beijing 100029, China; 2. Institute for Archival Scientific Technology, the State Archives Administration of China, Beijing 100050, China)

Abstract: A new model of user paying to see the answers of questions asked by some others is an important innovation form in the field of knowledge payment, understanding the influencing factors of users behavior is significant to promote the development of this model. We crawled the relevant data in question-and-answer platform. Using text analysis technology to dig into the content of problems, based on signal theory and elaboration likelihood model, we constructed a research model on the influencing factors of users' behavior, and then the effects of factors are also analyzed. This paper found that the price of answer, the approved number of answer, the length of answer, the number of related questions' listening volume and the reputation of respondents significantly affected listening volume, while the reputation of respondents had a moderating effect on the price of answer. This study provides guidance for community operators to optimize the paid product design and use community resources to carry out community operation.

Keywords: Question-and-Answer Platform; Knowledge Paying Behavior; Signal Theory

(收稿日期: 2019-10-22)