

电子科技大学

UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

《综合课程设计》报告

2019-2020-2

BACHELOR THESIS



论文 题目 基于在线旅游网站回复数据的民宿旅
游价值共创研究

学 院 公共管理学院

专 业 信息管理与信息系统

时 间 2020/7/9

小组成员及分工表

学号	姓名	主要工作
2017120101014	惠康欣	数据爬取、模型设计
2017120101015	韩思雨	模型设计、论文撰写
2017120101016	邓昕	数据爬取、结果分析
2017120101017	林云婷	结果分析、论文撰写

摘 要

在线旅行网站平台的兴起带动了民宿消费与行业发展，推动实现价值共创与服务质量的提升。本文基于携程旅行网的数据，使用普通最小二乘法模型探究了房东回复数据对民宿销量的影响。结果表明，房东回复数量的提升、回复时效的提升、回复内容的增加都会显著促进民宿销量的提升。交互效应分析表明，回复数量与回复时效的影响作用会随着民宿价格的提升而增强，而回复内容的影响作用会随民宿价格的提升而减弱。异质性分析表明，总体而言，回复数量、回复时效、回复内容对 5 种不同类型的房源销量影响均显著。最后，本文从在线旅游平台与房东角度为民宿发展提出了建议。

关键词 在线评论；价值共创；房东回复；消费者决策

ABSTRACT

The rise of online travel website platforms has driven the development of homestay consumption and industry, promoted the realization of consumer values co-creation and the improvement of service quality. Based on the data of Ctrip.com, this article uses the OLS model to explore the impact of landlord's response on the sales of homestay. The results show that the improvement of response quantity, quality and efficiency will significantly promote the sales of the homestay. However, the analysis of the interaction effect shows that the impact of response quantity and efficiency will weaken as the price of the homestay increases while the impact of response quality increases. Heterogeneity analysis shows that the number of responses, response efficiency, and response quality have significant effects on the sales of five different types of housing. Finally, this article put up some suggestions for the development of homestays from the perspective of online travel platforms and landlords.

Keywords: Online Comments, Value Co-creation, Response from Landlords, Consumers' Decision Making.

目录

摘 要.....	I
ABSTRACT.....	II
目录.....	III
第一章 引言.....	1
第二章 文献回顾.....	2
2.1 在线评论回复与销量.....	2
2.2 民宿平台研究现状.....	2
第三章 理论基础与研究假设的提出.....	4
3.1 评论回复率.....	4
3.2 评论回复时效.....	4
3.3 评论回复质量.....	5
第四章 实证研究设计.....	6
4.1 数据来源.....	6
4.2 数据收集.....	6
4.3 数据处理与计算.....	9
4.4 数据的描述统计.....	10
4.5 回归模型.....	11
5.1 基本回归结果.....	12
5.2 稳健性检验.....	13
5.4 异质性分析.....	15
第六章 研究结论与建议.....	18
6.1 研究结论.....	18
6.2 研究建议.....	18
参考文献.....	20

第一章 引言

“大众”旅游与乡村振兴背景下，旅游业蓬勃发展，住宿消费比例不断扩大。民宿，区别于传统的酒店与旅馆，为游客提供深入体验当地风情与享受民宿主人热情接待的机会，近年来越来越受到游客的喜爱，发展势头十分强劲。《2017 浙江民宿蓝皮书》显示，2017 年浙江省民宿总营业收入高达 53.5 亿元，其中客房直接收入 46.5 亿元，其他收入 7.1 亿元，直接就业人员 9.8 万人^[1]。

随着互联网的发展，越来越多人选择在线旅行预定。第 45 次中国互联网络发展状况统计报告^[2]显示，截至 2020 年 3 月，我国在线旅行预定用户规模达 3.73 亿，仅携程平台的上海民宿房源就有 32983 套^[3]，在线旅游平台的兴起为民宿主客双方提供了良好的沟通与交易平台，客观上促进了民宿的消费。在互联网时代，主客双方的沟通以在线评论的形式被公开展示，已成为网络口碑的延伸方式。而民宿作为一种体验型服务，在线评论与回复已成为潜在消费者购买决策的重要信息来源。

在服务价值创造方面，传统的服务价值是以商品为主导的，价值的产生主要是由卖方提供。而随着互联网的高速发展，顾客获取信息的渠道越来越多，顾客开始积极地参与到服务和产品的价值创造过程中^[4]，Vargo 和 Lusch^[5]将这样的价值逻辑归纳为服务主导逻辑（Service-dominant logic，简称 S-D 逻辑），并从服务系统和服务逻辑等视角研究了价值共创问题。所谓价值共创，是指价值是通过价值共创网络中成员（包括顾客与服务提供方）的共同参与及分享而产生的^[6]。以在线旅游平台的民宿商品为例，房东提供房源，顾客消费后评论提出建议，从而进一步促进民宿服务质量的提升。顾客与房东共同实现了服务质量的提升与服务价值的创造。

由此可见，在线旅游平台的兴起带动了民宿的消费与行业发展，也推动了价值共创从而实现服务质量的提升。顾客在线评论与房东回复则是记录了这一过程的重要信息载体。因此，本文试图研究顾客在线评论、房东回复与房源销量的关系，定量描述顾客评论与房东回复对房源销量的提升作用，并探究这一过程中的影响要素。

第二章 文献回顾

2.1 在线评论回复与销量

已有多位学者使用线上评论数据研究网络口碑对产品销量的影响。龚诗阳等^[7]使用线上图书评论数据研究网络口碑对图书销量的影响，发现考虑内生性后网络口碑中分数对销量的影响并不显著，影响主要来自于评论数量。杜学美等^[8]则通过实证研究验证了评论的数量、质量、效价都可以正向影响消费者的购买意愿，而潜在消费者的专业能力对评论的质量、数量与购买意愿之间的关系起着调节作用。亦有部分学者从文本内容出发，探究评论情感特征对消费者决策的影响。如焦梦蕾等^[9]利用美团酒店在线评论数据分析了文本 UGC 的情感特征对电子商务平台消费者决策行为的影响，研究发现情感特征对消费者决策有着正向影响，而随着样本量增加，正向影响逐渐减弱。

目前学术界针对商家回复的定量研究相对较少。程艳霞等^[10]基于感知价值理论，利用了结构方程模型，验证了商家回复对已经购买并发表评论的顾客具有重要影响，而影响程度取决于商家回复的有用性的三个维度——有效解决问题、满足情感需求、培育驾驭能力。李爱国等^[11]重点研究了商家回复在负面评论对消费者购买决策影响中的调节作用，发现商家回复对消费者购买决策呈正向调节，而回复策略与回复质量则呈负向调节。

2.2 民宿平台研究现状

目前已有部分学者对民宿预定量进行了定量研究。汪黄梅等^[12]基于信任理论建立了实证模型，爬取了小猪短租平台上的数据，探究了影响民宿预定量的影响因素，结果表明，订单平均确认时长、在线回复率、订单接受率、房东主页和房东性别对民宿预定量有着显著影响。而对于在线旅游网站的评价信息，已有的研究主要集中于对文本内容的分析，皮常玲等^[13]运用文本分析法分析了携程网上 1026 条民宿顾客抱怨的评论，分析得出顾客抱怨的主要问题与回复的主要内容。而李燕琴等^[14]则是从 Airbnb 平台收集评价信息，采用内容分析法分析了北京、台北两地民宿的差距，从位置选择、空间营造、房东特征、品牌管理等方面对民

宿发展提出了建议。

总体而言，目前学术界关于民宿的研究主要集中于个别案例的调研，定量研究，尤其是基于评论信息与回复信息的定量研究目前较少。基于此，本文收集了携程网上的民宿评论与回复信息，建立计量模型，探究在线评论与回复对民宿销量的影响，为定量研究评论回复与商品销量的关系提供了思路。

第三章 理论基础与研究假设的提出

不同于传统酒店，民宿是由私人房东提供闲置房源提供给有住宿需求的旅客，是个人对个人的交易，因此房东在在线旅游平台上展示的信息以及已消费顾客与房东的沟通互动成为了潜在消费者做出决策的重要依据。本文以在线旅游网站的评论与回复信息为研究对象，探究在线评论与回复对民宿消费的影响。

3.1 评论回复率

在传统的负面口碑的研究中，Laczniak 从归因理论的角度分析了潜在消费者的消费心理，他认为潜在消费者在进行负面评论归因时有两种重要的归因类型——品牌归因（顾客发布负面评论的原因是品牌的原因）和消费者归因（消费者发布负面评论的原因是消费者本人的原因）^[15]。而 Dens 等人的研究^[16]则表明商家针对消费者负面网络评论的有效回复可以降低消费者的品牌归因水平进而提升购买意愿。在在线旅游平台上，顾客进行消费后即可自主参与评论，评价的内容不一，往往是根据民宿的具体情况在各个方面表示赞赏或提出建议，房东的回复往往会对消费者提出的问题进行回应，若房东对顾客提出的问题视而不见，则会提高消费者的品牌归因水平，降低潜在消费者对商品的评价，降低潜在消费者的消费意愿。因而，本文提出如下假设

H1：房东的回复数量越高，房源销量就越好。

3.2 评论回复时效

消费心理学认为，当消费者对产品不满意，产生负面情绪时，心理上并无报复或对抗意识，反而对商家及时采取有效措施解除问题抱有强烈期待意识，而这种意识与回复及时程度呈负相关关系，即商家回复间隔时间越长，消费者对商家满意度越低，进而影响消费者购买行为^[17]。房东的回复越及时，顾客的心理诉求越容易得到满足，进而可以影响顾客撰写的评价内容与评分等级，提高潜在顾客的消费意愿。因而，本文提出假设：

H2：房东回复越及时，房源销量越好

3.3 评论回复质量

在网购行为中，我们可以发现商家回复主要有两种形式，一种是模板式的、单一的、结构化的回复，针对不同用户撰写的评论往往采用一模一样的回复，这会使得顾客反应的问题得不到针对性地解决；而另一种形式就是遵循具体问题具体分析的原则，对顾客提出的意见做出针对性的回应与改进。这两种回复的形式体现了商家对于服务的不同态度，针对性的回复不仅有利于消除负面评价的影响，还会增强潜在消费者对于服务质量提升的信心，进而可以提升潜在顾客的消费意愿。因而，本文提出假设：

H3：房东回复质量越高，房源销量越好

第四章 实证研究设计

4.1 数据来源

携程旅行网提供酒店、机票、车票等一站式专业预订服务，是中国最大的综合型在线旅游平台。携程网允许客户提供有关民宿住宿的评论，另外，民宿经营者可以对顾客的评论做出回应。下图为某民俗客栈下顾客评论和经营者回复的示例。本文将携程旅行网的“民宿客栈”板块作为研究数据来源，获取民宿房屋的基本属性特征、评论内容、回复内容等信息。



实际上，在网站评论内容中，部分评论可能属于虚假评论，但是在本文中忽略了这些虚假评论的影响，原因有二。首先，携程旅行网致力于确保在线客户评论的质量和准确性，仅允许真正的客户每次住宿后（每次在线交易一次评论）将评论发布到网站，并且利用专门的团队来识别和删除可疑的评论，这些质量控制流程一定程度上限制了虚假评论的发布。此外，鉴于携程旅行网站上的评论数量很多，虚假评论可能会的影响较为有限。

本文研究的重点是民宿经营者回复对后续绩效的影响，由于从在线旅游网站无法获取销量，考虑到民宿的销量和评论量成正比关系，因此使用民宿的评论量衡量绩效。具体来讲，按月份统计民宿的评论数据，分析上月经营者回复对本月民宿评论数量的影响。

4.2 数据收集

本研究中使用 python 爬虫程序获取数据。我们使用了两种不同的技术爬取数据。搜索结果页面中包含了多个民宿的基本信息，属于静态页面。民宿详情页

面中包括了该民宿的所有评论和回复信息，属于动态页面。

对于搜索结果页面，通过分析网页 url，实现了网页的翻页和修改检索条件，为获得尽可能多的数据，本研究中选择了成都、重庆、西安、三亚、广州、上海、杭州、青岛等 56 个热门旅游城市，更改不同的价格区间，进行多次检索，然后使用 BeautifulSoup 库解析 HTML 文件，从中提取民宿的基本信息，包括 ID、民宿名称、民宿类型、所在城市、价格、评论数、评论星级。程序代码如下图所示。

```
# 民宿所在城市
# city='chengdu10'
# city='nanjing9'
city='shanghai23'

urllist=[]
# 构造爬取的urllist. 网页限制最多爬60页
for p in range(60):
    url='https://inn.ctrip.com/onlineinn/newlist/'+city+'/?d1=2020-09-24&d2=2020-09-25'+s1+p+str(p)
    urllist.append(url)

# 所有的房源信息列表
allitems=[]
for url in urllist:
    response=rq.get(url)
    texts=response.text
    if response.status_code==200:
        out_put(url,city+'_'+str(urllist[-1]),texts)
    else:
        print("访问出错")
        break
soup=BeautifulSoup(texts,'lxml')

# 找到带有unitid的所有房源信息
# houses=soup.find_all('script')[9].text
houses=soup.find_all('script',text=re.compile("unitId"))[0].text
houses=re.sub('___NEXT_DATA___',' ',houses).split(';')[2]
hjson=json.loads(houses)

items=hjson['props']['pageProps']['initialState']['items']
allitems.extend(items)
```

图 2 爬取 id 程序

```
1 allitems[0].keys()

dict_keys(['unitId', 'unitName', 'houseTypeName', 'pictureList', 'defaultPicture', 'pictureCount', 'logoUrl', 'geoCoordSysType', 'longitude', 'latitude', 'address', 'allowBooking', 'unitInfo', 'inIns', 'active', 'activityInfo', 'activityId', 'failType', 'productPrice', 'finalPrice', 'addPeopleAmount', 'cleaningAmount', 'priceMissingText', 'cityId', 'cityName', 'districtName', 'chineseLandlord', 'quickBooking', 'promotionPicUrl', 'smallPictureUrl', 'unitSummaries', 'adverUnit', 'adverUnitType', 'great', 'expressBooking', 'priceTags', 'rba', 'qualityCertification', 'checkInDate', 'checkOutDate', 'unitTags', 'distanceTip', 'showHouseVideo', 'showHouseVR', 'topListTag', 'urgencyTip', 'houseTags', 'finalPriceDisplayType', 'priceTipBadge', 'displayOriginPrice', 'houseListTagImage', 'productId', 'landlordLevelUrl', 'productPriceConstruct', 'promoText', 'rankScore', 'rankScore2', 'distance', 'extends4Ctrip', 'hotelId', 'recommendedGuests', 'bedCount', 'housingType', 'theHouseType', 'commentScore', 'checkInReceptionTime', 'checkOutLatestTime', 'preloadDetail', 'distanceDesc', 'rname'])
```

图 3 房源字段信息

```
# 需要获取的房源的属性
ikeys=['unitId', 'unitName', 'houseTypeName', 'longitude', 'latitude', 'address', 'unitInfo', 'productPrice', 'finalPrice', 'cityId', 'districtName']

ctrips=pd.DataFrame()
i=0

# 将数据写入dataframe
for item in allitems:
    for key in ikeys:
        ctriptions.at[i,key]=item[key]
# 房源的评分和评论数
for dict in item['unitSummaries']:
    summary="none"
    if dict['textType']=='COMMENT':
        summary=dict['text']
    break
if summary=="none":
    ctriptions.at[i,'score']="none"
    ctriptions.at[i,'numcomments']="none"
else:
    summary=summary.split('/')
    ctriptions.at[i,'score']=summary[0].replace("分","")
    ctriptions.at[i,'numcomments']=summary[1].replace("点评","")
i+=1
```

图 4 保存数据

第四章 实证研究设计

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	unitId	unitName	houseType	longitude	latitude	address	unitInfor	productPr	finalPrice	cityId	districtNo	score	numcomments
2	16404270	「百分」 烦恼太古里	104.08862	30.653815	镜铠街	近春熙路/		880	880	10	锦江区	5	27
3	1197681	北站舒适大床房/双地	104.09507	30.697085	柚米国际社	近春熙路/		498	449	10	青羊区	4.9	96
4	261732	「格调」近春熙路,宽	104.07535	30.678562	壹号公馆	(近春熙路/		238	238	10	青羊区	5	1787
5	1170373	已消毒*放心住*近熊	104.10704	30.695985	永立龙寓	(近春熙路/		288	245	10	成华区	5	1516
6	3961015	近宽窄巷子~文殊院	104.07535	30.678562	成都市青羊	近春熙路/		160	160	10	青羊区	4.9	2235
7	11355	双地铁春熙路柠檬宿	104.09259	30.667205	红星国际	(近春熙路/		428	386	10	锦江区	4.9	508
8	10612463	春熙路/太古里/熊猫	104.10589	30.694866	永立龙寓	近春熙路/		213	213	10	成华区	5	14
9	1170221	春熙路宽窄巷子文殊	104.0751	30.678841	壹号公馆	(近春熙路/		208	208	10	青羊区	5	1409
10	15108262	歌今·海若 复工放心	104.09836	30.647663	歌今Gerzi	近春熙路/		368	332	10	锦江区	5	68
11	78457	市中心/春熙路/奢三	104.07558	30.6701	富力·公作	近春熙路/		1380	1380	10	青羊区	4.9	634
12	10829211	设计师Loft民宿/近	104.06407	30.683299	成都途家	近春熙路/		268	268	10	青羊区	5	52
13	6568683	太古里/春熙路/锦里	104.10697	30.696013	驷马桥四路	近春熙路/		188	170	10	成华区	5	136
14	4670941	【已消毒】春熙路/近	104.10683	30.695964	永立龙寓	近春熙路/		200	200	10	成华区	4.9	551
15	12092690	春熙路宽窄巷子 套	104.07353	30.676666	证大·正	近春熙路/		599	540	10	青羊区	5	106
16	11971821	◆『安心消毒』春熙	104.08078	30.667412	恒大广场	(近春熙路/		488	488	10	青羊区	5	445
17	16404270	「百分」 烦恼太古里	104.08862	30.653815	镜铠街	近春熙路/		880	880	10	锦江区	5	27
18	1197681	北站舒适大床房/双地	104.09507	30.697085	柚米国际社	近春熙路/		498	449	10	青羊区	4.9	96
19	261732	「格调」近春熙路,宽	104.07535	30.678562	壹号公馆	(近春熙路/		238	238	10	青羊区	5	1787
20	1170373	已消毒*放心住*近熊	104.10704	30.695985	永立龙寓	(近春熙路/		288	245	10	成华区	5	1516
21	3961015	近宽窄巷子~文殊院	104.07535	30.678562	成都市青羊	近春熙路/		160	160	10	青羊区	4.9	2235
22	11355	双地铁春熙路柠檬宿	104.09259	30.667205	红星国际	(近春熙路/		428	386	10	锦江区	4.9	508
23	10612463	春熙路/太古里/熊猫	104.10589	30.694866	永立龙寓	近春熙路/		213	213	10	成华区	5	14

图 5 部分数据展示

对于民宿详情页面，分析网站请求过程可知，页面中展示的数据是从数据接口中获取的，因此，对某一个特定的民宿，直接调用 API，模拟发送 post 请求，获得 json 格式的数据，解析获得每条评论的文字，评论时间，回复内容，回复时间等。

```

workbook = xlwt.Workbook() #创建一个Excel文件,用于存储输出数据
worksheet = workbook.add_sheet('text') #创建一个sheet
row=0
rbook = xlrd.open_workbook(r'C:\Users\dengx\Desktop\ctrips.xlsx')
# sheets方法返回对象列表, [xlrd.sheet.sheet object at 0x103f147f0]
rbook.sheets()
# xls默认有3个工作簿, Sheet1, Sheet2, Sheet3
rsheet = rbook.sheet_by_index(0) # 取第一个工作簿
# 循环工作簿的所有行
curl="https://m.ctrip.com/restapi/soa2/16593/json/getHouseCommentList"
for i in rsheet.get_rows():
    print(i)
    unitID = int(i[0].string)
    pageSize=int(i[12].string)
    payloadData={
        "unitID":unitID,"pageIndex":0,"pageSize":pageSize,"tagParameter":"1,0","head":{"cid":"09031029212120301011","ctok":"","cver":"1.0","la
    }
    payloadHeader={
        "content-type": "application/json",
        "origin": "https://inn.ctrip.com"
    }
    print(unitID, pageSize)
    cresponse=rq.post(curl, data=json.dumps(payloadData), headers=payloadHeader)
    # 评论长度, 对应pageSize
    data=json.loads(cresponse.text) #json就是字典数据
    #pprint.pprint(data)
    comments=data['data']['comments']
    for comment in comments:
        commentDetail=comment['commentDetail'] #评论内容
        createTime=comment['createTime'] #评论创建时间
        hotelName=comment['hotelName'] #名称
        overall=comment['overall'] #总分
        replyDetail=comment['replyDetail'] #回复内容
        replySubmitTime=comment['replySubmitTime'] #回复时间
        scoreTitle=comment['scoreTitle'] #评分名称
        unitID=comment['unitID'] #房源ID

```

图 6 评论数据爬取

1	房源ID	评论内容	评论时间	房源名称	评分	回复内容	回复时间	评分名称
2	4665607	这家公寓真	2020-04-2	爵府度假公	5		2020-04-3	超棒
3	4665607	房东在线沟	2020-04-1	爵府度假公	5	感谢您的入	2020-04-1	超棒
4	4665607	第一次住过	2020-04-2	爵府度假公	5		2020-04-3	超棒
5	4665607	过来出差顺	2020-04-2	爵府度假公	5		2020-04-3	超棒
6	4665607	在网上对比	2020-04-1	爵府度假公	5	感谢您的入	2020-04-1	超棒
7	4665607	非常漂亮的	2020-04-0	爵府度假公	5	感谢您的入	2020-04-1	超棒
8	4665607	管家人特别	2020-03-1	爵府度假公	5	感谢您的入	2020-03-1	超棒
9	4665607	很喜欢这里	2020-04-2	爵府度假公	5		2020-04-3	超棒
10	4665607	疫情导致入	2020-02-2	爵府度假公	5	感谢您的入	2020-03-0	超棒
11	4665607	总体来说，	2020-02-0	爵府度假公	5	感谢您的入	2020-02-0	超棒
12	4665607	民宿地理位	2020-02-0	爵府度假公	5	感谢您的入	2020-02-1	超棒
13	4665607	管家人特别	2020-01-1	爵府度假公	5	感谢您的入	2020-02-1	超棒
14	4665607	能在这个非	2020-02-1	爵府度假公	5	感谢您的入	2020-02-1	超棒
15	4665607	房间干净整	2020-04-2	爵府度假公	5		2020-04-3	超棒
16	4665607	房间给人非	2020-04-2	爵府度假公	5		2020-04-3	超棒
17	4665607	很满意的一	2020-01-0	爵府度假公	5	感谢您的入	2020-01-0	超棒
18	4665607	第一次利用	2020-01-0	爵府度假公	5	感谢您的入	2020-01-0	超棒
19	4665607	第一次入住	2020-02-0	爵府度假公	5	感谢您的入	2020-02-0	超棒
20	4665607	房间收拾非	2020-02-0	爵府度假公	5	感谢您的入	2020-02-1	超棒
21	4665607	非常棒的一	2020-01-2	爵府度假公	5	感谢您的入	2020-01-2	超棒
22	4665607	房间环境非	2019-11-3	爵府度假公	5	感谢您的入	2019-12-0	超棒
23	4665607	第一次近非	2019-12-2	爵府度假公	5	感谢您的入	2019-12-2	超棒
24	4665607	下次来成者	2020-01-1	爵府度假公	5	感谢您的入	2020-02-1	超棒
25	4665607	非常棒的一	2020-01-1	爵府度假公	5	感谢您的入	2020-01-2	超棒
26	4665607	公寓位置非	2020-01-1	爵府度假公	5	感谢您的入	2020-02-1	超棒

图 7 评论数据示例

最终，删除了过少（小于 20）的数据后，共得到 19620 个房源信息，1172869 条评论，其中 813225 条有民宿经营者回复。

4.3 数据处理与计算

本文关注了经营者回复对民宿销量的影响，因此将获取到的评论数据汇总到民宿级别，分别计算每个民宿的评论数、回复数等特征。考虑经营者发布的回复内容的影响作用发生在潜在消费者浏览回复内容后，即当前的回复内容对未来的民宿销量产生影响，因此本文以月为基本统计单位，研究上月经经营者回复内容对本月民宿销量的影响，将数据按月分别计算当月评论数、上月回复数、上月平均评论数、平均回复时效、平均回复字数、平均回复间隔，变量测定如表 1 所示。删除了部分错误数据后，共得到 254691 条有效数据。

表 1 重要变量描述

变量	测量	计量指标	标记
当月评论数	当月民宿的评论总数	非负整数	$Comments(m)$
上月回复数	上月经经营者对评论的回复总数	非负整数	$RespCount(m-1)$
回复间隔	上月经经营者回复时间与对应评论时间间隔的平均值	非负整数	$RespEfficiency(m-1)$
回复长度	上月经经营者回复内容字数的平均值	非负整数	$RespQuality(m-1)$
价格	民宿的价格	非负整数	$Price$
上月平均评论数	上月所有民宿评论数的均值	非负整数	$Avg_Comments(m-1)$
上月相对评论数	上月评论数与平均评论数的差	实数	$\Delta Comments(m-1)$

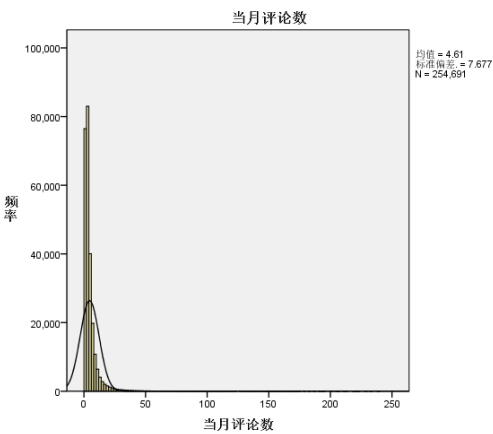
4.4 数据的描述统计

对获取的数据进行处理后，通过对各指标属性的量化统计，各指标要素字段的基本描述如表 2 所示。

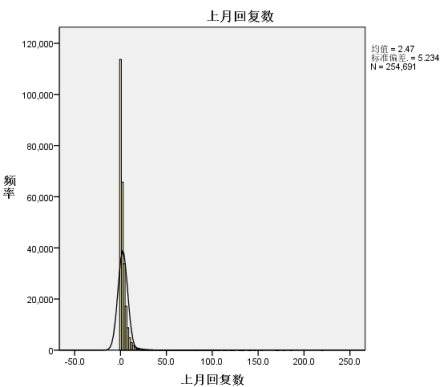
表 2 变量描述性统计

	count	mean	std	min	25%	50%	75%	max
当月评论数	254691	4.605067	7.676962	1	1	3	5	238
回复间隔	254691	32.48816	95.75159	0	0	1	11	2462
回复长度	254691	42.93972	55.68636	0	13	26	56	1035
价格	254691	443.2285	525.3075	32	228	299	399	10000
当月相对评论数	254691	-1.38E-06	7.575758	-5.4155	-3.1449	-1.6177	0.5112	233.833
上月平均评论数	254691	3.08291	2.379365	0	0	3.6244	5.4224	8.5
上月回复数	254691	2.467209	5.23359	0	0	1	3	219
上月相对评论数	254691	0.345464	6.21953	-5.4155	-1.9895	0	0	233.833
上月评论数	254691	3.428374	6.74664	0	0	2	4	238

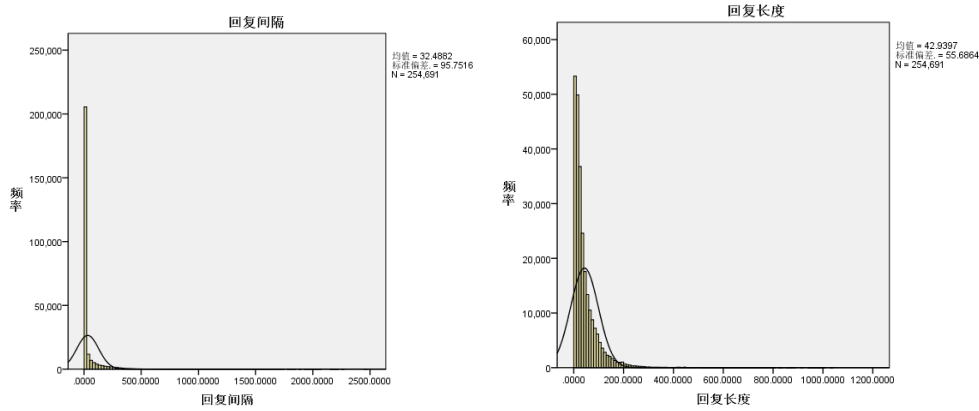
下面对各个属性的分布情况进行分析。



当月评论数是主要关注的因变量，用来表示民宿当月的订单数，针对每个民宿按月份统计，根据爬取的数据，绘制的当月评论数分布情况如图所示，可以看出多数民宿的评论数小于 50，最多的当月评论数超过了 250，呈长尾分布。



经营者对上月消费者评论的回复数量是主要关注的自变量，其分布情况如上图所示，可以看出其分布情况和当月评论数类似。



我们认为，回复间隔和回复长度能够反映民宿经营者的回应的积极性和回复态度，使用回复间隔和回复长度分别作为度量回复及时性与回复质量的指标。

我们按天数统计评论的回复间隔，用评论的字数统计回复长度，其分布情况如上图所示，可以看出大多数民宿评论的回复间隔较短，也有少数间隔天数较长，回复长度大多在 200 字以内。

4.5 回归模型

我们决定使用回归模型验证前文提出的假设。将房源的当月评论数作为被解释变量，上月回复数、回复间隔和回复长度作为解释变量进行回归分析。此外，需要消除其他无关因素的影响，我们加入一系列控制变量（房源类型、价格、上月平均评论数、时间）进入回归模型。回归公式如下：

$$\begin{aligned} Comments(m) = & \beta_0 + \beta_1 \cdot RespCount(m-1) + \beta_2 RespEfficiency(m-1) \\ & + \beta_3 RespQuality(m-1) + \beta_4 \cdot AvgComments(m-1) + \beta_5 \cdot Price(m) \\ & + \beta_6 \cdot Type + \beta_7 \cdot City + \varepsilon \end{aligned}$$

第五章 实证结果与分析

5.1 基本回归结果

根据前文公式所示的计量模型，本文基于在线旅游网站（携程网）的在线评论数据通过普通最小二乘法(OLS)研究在线评论数据的民宿旅游价值共创，基本回归结果如表 3 所示。表 3 通过逐步加入控制变量的方式估计了在线评论数据对民宿旅游价值共创的影响。

第（1）列未加入任何控制变量，上月回复数与回复长度的系数均在 1%的水平上显著为正，表明上月回复数与回复长度均与当月评论数之间存在显著的正相关关系，而回复间隔的系数在 1%的水平上显著为负，说明回复间隔与当月评论数存在显著的负相关关系，由于遗漏重要解释变量，还不能得出明确的因果关系。第(2)列在第(1)列的基础上加入了城市固定效应和房源类型固定效应，上月回复数以及回复长度的估计系数仍在 1%的水平上显著为正，且系数小于第（1）列的估计结果，可见遗漏重要解释变量将会高估上月回复数及回复长度对当月评论数的促进效应。同时回复间隔仍显著为负，表明回复间隔越长，当月评论数量越少。第（3）列控制了上月平均评论数与价格等控制变量，三个主要自变量的估计系数仍在 1%的水平上显著。

表 3 基本回归结果

解释变量	(1)	(2)	(3)
上月回复数	0.4128*** (0.0096)	0.3977*** (0.0096)	0.3933*** (0.0105)
回复间隔	-0.0043*** (0.0001)	-0.0043*** (0.0001)	-0.0042*** (0.0001)
回复长度	0.0065*** (0.0004)	0.0051*** (0.0004)	0.0051*** (0.0004)
上月平均评论数			0.0270*** (0.0094)
价格			0.0001** (0.0000)
城市固定效应	否	控制	控制
类型固定效应	否	控制	控制
样本量	254691	254691	254691
拟合优度	0.0876	0.1111	0.1112

注：*、**、***分别代表 10%、5%与 1%显著性水平；括号内为估计系数的稳健标准差。

表 3 的基本回归结果说明，以各个房源为基本单位，上月回复数、回复间隔

以及回复长度均能显著影响当月评论数, 具体来说, 上月回复数每增加 1 条, 当月评论数增加 0.4128 条; 回复长度每增加 1 条, 当月评论数增加 0.0065 条; 回复间隔每增加 1 天, 当月评论数减少 0.0043 条。进而说明, 在线旅游网站的在线评论数据具有显著的价值共创效应, 民宿经营者可以通过在线评论数据的回复进而提高民宿的销售量。这样, 前文提到的三个假设都已经初步得到验证。

根据第 (3) 列的回归结果, 我们还可以得到一些关于控制变量的有意义的结论。上月平均评论数显著为正, 说明上月平均评论数越多, 当月评论数就越多, 这也表明了上个月顾客的评论信息为下个月顾客的旅游民俗选择提供了一定的参考价值, 从而顾客通过在线评论在在线旅游网站中创造了价值; 而价格在 5% 的水平上显著, 说明价格越高的民宿, 其销售量越高, 这可能是顾客认为价格越高相应的服务会越好, 为了有一个愉快的旅途, 因此更倾向于选择价格高的民宿, 但价格的估计系数仅为 0.0001, 也表明了虽然价格对旅游民宿的销售量有促进效果, 但是该影响微乎其微。

5.2 稳健性检验

本文的基本回归模型的因变量与自变量都采用了绝对的评论与回复数据, 但绝对的数据可能会由于季节、天气等多种因素有较大的浮动, 导致结果的不稳定, 例如: 也有可能是恶劣天气导致了当月评论数减少, 民宿销售量下降, 为了解决这一问题, 下文我们将绝对的评论与回复数据, 替换成相对数据进一步评估在线旅游网站在线评论数据的民宿旅游价值共创。稳健性检验使用的模型如下:

$$\begin{aligned} Comments'(m) = & \beta_0 + \beta_1 \cdot RespRate(m-1) + \beta_2 RespEfficiency(m-1) \\ & + \beta_3 RespQuality(m-1) + \beta_4 \cdot AvgComments'(m-1) + \beta_5 \cdot Price(m) \\ & + \beta_6 \cdot Type + \beta_7 \cdot City + \varepsilon \end{aligned}$$

我们将因变量当月评论数替换成当月相对评论数, 该变量以当月所有民宿的销售量平均值为基准, 在该基准上上下下浮动, 如某一民宿当月相对评论数小于 0 表示该民宿当月销售量小于平均值, 该变量显示了某一民宿销售量在行业整体中的相对位置; 自变量上月回复数替换为上月回复率, 上月回复率通过总回复数除以总评论数计算得到; 上月平均评论数替换为上月相对评论数, 该变量与当月相对评论数计算方式与意义相同。回归结果如表 4 所示。

表 4 稳健性检验结果

解释变量	(1)	(2)
上月回复率	0.2249*** (0.0309)	0.2740*** (0.0293)
回复间隔	-0.0014*** (0.0001)	-0.0011*** (0.0001)
回复长度	0.0095*** (0.0004)	0.0075*** (0.0004)
上月相对评论数		0.4100*** (0.0100)
价格		0.0001* (0.0000)
城市固定效应	控制	控制
类型固定效应	控制	控制
样本量	254691	254691
拟合优度	0.0378	0.1485

注：*、**、***分别代表 10%、5%与 1%显著性水平；括号内为估计系数的稳健标准差。

上月回复率与回复长度均在 1%的水平上显著为正，回复间隔显著为负，更换变量后的回归结果与基本回归结果保持一致，从而证明了基本回归模型的稳健性，进一步验证了假设的正确性。

5.3 交互效应检验

前文通过研究发现回复数量、回复间隔、回复长度与价格均在 1%的水平上显著促进了民宿销售量，为了进一步研究这几个变量的作用机制，本文接下来将研究价格与几个自变量的交互效应。

本研究通过逐步加入交互效应的方式进行研究，首先加入了价格与上月回复数的交互效应，表中第（1）列展示了该交互效应的影响，该交互效应在 1%的水平上显著为正，表明回复率对销量的影响随着价格的升高而加强，因此价格越高的民宿，房东更应注重回复顾客的评论；第（2）列展示了价格与回复间隔的交互效应，回归结果均在 1%的水平上显著为负，表明回复间隔对销量的影响随着价格的升高而增强，但交互效应的估计系数为-1.49e-06，表明该交互效应影响较弱。第（3）列展示了价格与回复长度的交互效应，根据表中结果可以看出，在 1%的显著性水平上，回复长度对销量的影响随着价格的升高而减弱，这一点可能是因为顾客回复没有提出实质性问题或者没有引发争议时，商家较多的采用了复制粘贴的方式给与统一的回复。

表 5 交互效应检验结果

解释变量	(1)	(2)	(3)	(4)
上月回复数	0.2840*** (0.0124)	0.3933*** (0.0105)	0.3932*** (0.0105)	0.2820*** (0.0126)
回复间隔	-0.0040*** (0.0001)	-0.0036*** (0.0001)	-0.0042*** (0.0001)	-0.0037*** (0.0002)
回复长度	0.0017*** (0.0005)	0.0051*** (0.0004)	0.0058*** (0.0004)	0.0054*** (0.0005)
上月平均评论数	0.0581*** (0.0087)	0.0270*** (0.0094)	0.0269*** (0.0094)	0.0577*** (0.0087)
价格	-0.0021*** (0.0002)	0.0001*** (0.0001)	0.0002*** (0.0001)	-0.0017*** (0.0001)
价格*上月回复数	0.0006*** (0.0000)			0.0006*** (0.0000)
价格*回复间隔		-1.49e-06*** (3.14e-07)		-9.30e-07** (4.61e-07)
价格*回复长度			-1.49e-06*** (3.58e-07)	-7.32e-06 *** (1.06e-06)
城市固定效应	控制	控制	控制	控制
类型固定效应	控制	控制	控制	控制
样本量	254,691	254,691	254,691	254,691
拟合优度	0.2294	0.1112	0.1113	0.2318

注：*、**、***分别代表 10%、5%与 1%显著性水平；括号内为估计系数的稳健标准差。

5.4 异质性分析

房源类型的不同为顾客提供了不同的服务与居住体验，甚至不同的顾客对于不同房源类型的偏好不同，这都导致了不同的房源类型在市场需求中存在着差异，因此不同的房源可能对当月评论数产生的影响可能也存在差异性，表 6 便探讨了这种差异性。表 6 的（1）-（5）列展示了使用该评论数据中房源数最多的前 5 名房源类型进行标准化回归的结果，这 5 种房源分别是普通公寓、酒店式公寓、客栈、Loft 复式以及别墅单间。

5 种类型房源的上月回复数标准化回归系数均在 1%的水平上显著为正，但值得关注的是它们的系数存在着差异，其中 Loft 复式最大为 0.3712，别墅单间的估计系数为 0.3227，普通公寓及酒店式公寓的估计系数接近，均在 0.29 左右，而客栈的上月回复数的估计系数最小为 0.2217。总体来说，5 种房源类型均在较大程度上受到房东回复数的影响，这可能是由于回复数在一定程度上代表了房东

对于顾客的服务态度,选择这五种房源类型的顾客均会通过回复数这一信息来查看房东的服务态度,进而决定是否订购。同时 Loft 复式的估计系数最大,表明 Loft 复式受到上月回复数的影响最大,因此 Loft 复式的房东更应该注重对于客户评论的回复,这将有效提高其销售量。

表 6 不同房源类型对民宿旅游价值共创影响的回归结果

解释变量	(1)	(2)	(3)	(4)	(5)
	普通公寓	酒店式公寓	客栈	Loft 复式	别墅单间
上月回复数	0.2978*** (0.0135)	0.2886*** (0.0256)	0.2217*** (0.0206)	0.3712*** (0.0260)	0.3227*** (0.0498)
回复间隔	-0.0552*** (0.0001)	-0.0526*** (0.0002)	-0.0559*** (0.0003)	-0.0426*** (0.0003)	-0.0655*** (0.0005)
回复长度	0.0417*** (0.0005)	0.0410*** (0.0009)	0.0403*** (0.0013)	0.0282*** (0.0007)	0.0004 (0.0017)
上月平均评论数	0.0036 (0.0110)	0.0220*** (0.0239)	-0.0005 (0.0283)	0.0216** (0.0177)	0.0022 (0.0459)
价格	0.0176*** (0.0001)	0.0162** (0.0001)	-0.0076* (0.0001)	0.0031 (0.0001)	0.0146 (0.0009)
城市固定效应	控制	控制	控制	控制	控制
样本量	119629	48733	42929	25093	6928
拟合优度	0.1196	0.1160	0.0908	0.1629	0.1735

注: *、**、***分别代表 10%、5%与 1%显著性水平;括号内为估计系数的稳健标准差。

此外 5 种类型的房源的回复间隔的标准化估计系数也均在 1%的显著性水平上为正,其中别墅单间的回复间隔标准化估计系数最小为-0.065,表明客栈的回复间隔相比于其他几种类型的房源来说,对价值共创影响更大,客栈房东应该注重回复效率以提高销售量,Loft 复式的回复间隔标准化估计系数最大为-0.0426,表明其对当月评论数的负影响作用最小。

回复长度的影响效果中,别墅单间的标准化估计系数不显著,表明回复长度对选择别墅单间的顾客无显著影响,其他四种房源类型的回复长度均在 1%的水平上显著为正,其中普通公寓的回复长度标准化估计系数最大为 0.0417,酒店式公寓与客栈的回复长度标准化估计系数次之,Loft 复式最小为 0.0282,表明在这四种房源类型种回复长度的影响效果逐渐减小。表 6 的回归结果表明,不同的房源类型对于在线旅游网站的在线评论数据的价值共创存在差异,其中 Loft 复式应着重增加上月回复数来增加旅游民宿销售量,别墅可以通过缩小回复间隔来增加客栈销售量,而普通公寓更应该注重回复质量。总体来说,上月回复数、回复间隔、回复长度均对 5 种房源类型有显著影响,故 5 种房源类型的房东都应注重

在回复评论的过程中为顾客创造价值。

第六章 研究结论与建议

6.1 研究结论

本文爬取了携程旅行网上 56 个热门旅游城市民宿的房东回复数据，以房源本月销量为被解释变量，房东上月回复率（数量）、回复时效、回复质量作为解释变量，房源类型、价格等作为控制变量，构建了面板数据进行回归分析。最终我们可以得到如下结论：房东上月回复数量的提高、回复时效的提升、回复质量的提升都可以显著促进房源本月销量的提升。这与李爱国等学者^[17]发现的商家的回复时效和回复质量可对负面评论内容评分与体验性产品销量关系存在调节作用的结论一致。此外，通过交互效应检验，我们可以发现，回复数量与回复时效的影响作用会随着民宿价格的提升而增强，而回复内容的影响作用会随民宿价格的提升而减弱；通过异质性分析，我们发现，对于别墅单间，回复质量对销量的提升作用不显著。类型为 Loft 复式的房源，上月回复数与回复时效的提高对于本月民宿销量的提升作用较大。

6.2 研究建议

本文基于携程在线旅行平台的回复数据，从价值共创的角度探讨了民宿经营者回复对民宿销量的影响。基于本文的发现，我们从在线旅行平台和房东两个方面为民宿发展提出建议：

1. 对在线旅行网站的建议

在线旅行网站不仅仅是民宿商品的展示与选购的平台，更是促进民宿的经营者与消费者沟通的平台。被公开展示的已消费人群与经营者的沟通与互动对潜在消费者的决策具有重要作用。在线旅行网站应鼓励民宿的经营者积极与消费者互动，建立多种沟通机制促进消费者与服务提供者的良性互动，这有助于民宿商品线上口碑的传播与商品销量的提升。

2. 对民宿经营者的建议

民宿经营者对已完成消费的顾客的回复会显著影响潜在消费者的决策，因此，民宿经营者应耐心阅读与回复顾客的评论，并尽可能地缩短回复的时间、回复较多的内容。这不仅可以从顾客的评论中吸取建议从而提升服务的质量，更可以展

第六章 研究结论与建议

示房东的诚意与友好，建立起房东在已消费顾客和潜在消费者心中的友善与好客的形象，促进民宿销量的提升。

参考文献

- [1]浙江发布国内首部民宿蓝皮书 助推民宿产业发展.[EB/OL][2020-06-24].<https://www.meadin.com/155247.html>.
- [2]第 45 次《中国互联网络发展状况统计报告》(全文).[EB/OL][2020-06-24].http://www.cac.gov.cn/2020-04/27/c_1589535470378587.htm.
- [3]上海民宿客栈-上海住宿公寓-携程民宿.[EB/OL][2020-06-24].<https://inn.ctrip.com/onlineinn/newlist/shanghai23/?d1=2020-06-10&d2=2020-06-11>.
- [4]杨学成. 出行共享中的用户价值共创机理——基于优步的案例研究[J]. 管理世界, No.287(08):154-169.
- [5]VARGO,S. L. Evolving New Dominant Logic for Marketing[J]. journal of marketing, 2004.
- [6]简兆权, 肖霄. 网络环境下的服务创新与价值共创:携程案例研究[J]. 管理工程学报, 2015, 029(001):20-29.
- [7]龚诗阳, 刘霞, 刘洋,等. 网络口碑决定产品命运吗——对线上图书评论的实证分析[J]. 南开管理评论, 2012(04):120-130.
- [8]杜学美, 丁璟好, 谢志鸿,等. 在线评论对消费者购买意愿的影响研究[J]. 管理评论, 2016, 028(003):173-183.
- [9]焦梦蕾,赵涛,徐勇,武雅利.一种考虑文本 UGC 情感特征的消费者行为预测模型[J].情报理论与实践,2019,42(12):139-143+85.
- [10]程艳霞,王紫穗.感知价值视角下商家回复有用性的结构维度与测量[J].商业经济研究,2020(11):93-96.
- [11]李爱国, 邓召惠, 赖胜强. 商家回复在负面评论对消费者购买决策的影响中的调节效应研究[J]. 现代经济探讨, 2017(06):83-90.
- [12]汪黄梅,丁文辉.基于信任理论的民宿预定量影响因素研究——以长三角民宿区为例[J].福建商学院学报,2019(01):37-43.
- [13]皮常玲, 郑向敏. 基于在线评论的民宿顾客抱怨研究——以厦门鼓浪屿民宿为例[J]. 旅游论坛, 2017(3).
- [14]李燕琴, 于文浩, 柏雨帆. 基于 Airbnb 网站评价信息的京台民宿对比研究[J]. 管理学报, 2017, 014(001):122-128,138.
- [15]Laczniak R N , Ramaswami D C N . Consumers' Responses to Negative Word-of-Mouth Communication: An Attribution Theory Perspective[J]. Journal of Consumer Psychology, 2001,

参考文献

11(1):57-73.

[16]Pumawirawan N , Eisend M , De Pelsmacker P , et al. A Meta-analytic Investigation of the Role of Valence in Online Reviews[J]. Journal of interactive marketing, 2015, 31(Aug.):17-27.

[17]李爱国,邓召惠,毛冰洁. 在线负面评论对体验型产品销量的影响——基于商家回复视角[J]. 商业研究(7 期):138-144.