

【信息经济与管理】

基于流量、时间和费用的 信息消费水平测度方法研究

黄宝凤, 李娜

(南京邮电大学 经济学院, 江苏 南京 210023)

摘要:信息消费水平的准确测度是促进信息消费的基础。在移动互联网蓬勃发展的时代背景下,首先从信息消费的过程、信息消费的多样性、信息化水平的提高三个方面分析了综合利用流量、时间和费用三项指标进行信息消费水平测度的必要性;接着在信息消费具有惯性的假设下,利用固定权重和可变权重,构建了基于流量、时间和费用的信息消费水平测度模型,证明了该模型具有非负性、传递性、连续性和可加性;最后通过实证调查验证了所构建模型的正确性。实证结果表明:流量、时间和费用均是测度信息消费水平的重要指标;利用单一费用指标测度的信息消费水平与综合利用流量、时间和费用三项指标测度的信息消费水平有显著差异,而且会明显低估真实的信息消费水平。

关键词:移动互联网;信息消费;测度;效用

中图分类号:F49 **文章编号:**1673-5420(2018)01-0084-09

信息消费是一种直接或间接消费信息服务和信息产品的经济活动。进入21世纪,信息消费对经济发展的促进作用越来越强,据工业和信息化部门的相关数据显示:信息消费每增加100亿元,能带动国民经济增长338亿元^[1]。为促进信息消费,我国相继制定了一系列的相关政策,如2013年国家颁发的《国务院关于促进信息消费扩大内需的若干意见》等。各地方为促进本地区信息消费的发展,也先后出台了許多区域性的鼓励信息消费的政策,如上海市出台的《上海市贯彻〈国务院关于促进信息消费扩大内需的若干意见〉行动纲要(2014—2017年)》等。科学、准确的信息消费水平测度是执行这些政策的基础。

收稿日期:2017-10-21 本刊网址:<http://nysk.njupt.edu.cn>

作者简介:黄宝凤,教授,博士,研究方向:经济统计、数据挖掘等。

李娜,硕士研究生,研究方向:经济统计、数据挖掘等。

基金项目:国家统计局重点项目“居民信息消费的统计测度研究”(2014LZ50);国家社会科学基金项目“城市化背景下我国城乡居民信息消费研究”(14CTQ019)

一、信息消费水平测度的研究综述

目前国内外对于信息消费水平测度的研究大致可分为三类:第一,利用与信息化相关的指标来综合测度信息消费水平。如 Kellerman 提出信息消费水平可以通过软件买卖数量或国内、国际电话的数量与线路长度来测度,也可通过连接上有线电视系统的电视机数量来测度^[2];黄柳等通过每人每年通话次数和每年技术引进数量等指标综合测评企业的信息消费水平^[3]。以上研究是利用信息化设施的数量等指标进行综合测度,因此与其说是信息消费水平测度,不如说是信息化建设水平测度。第二,将居民的支出费用作为信息消费水平的测度。如张红历和丁志帆等人以城镇居民的医疗保健、交通通讯、娱乐文化三项消费支出的总和作为城镇居民的信息消费数额,研究了城镇居民信息消费的差异^[4-5]。李婵娟等人将我国农村居民人均消费性支出中医疗保健、交通通讯、教育文化娱乐服务三项支出的总和作为农村人均信息消费支出的替代,以评价我国农村居民信息消费水平^[6]。但是居民在交通和通讯、文化教育娱乐、医疗保健等方面的支出费用是信息消费的间接测度,与信息消费中的直接支出费用不能完全对等。第三,利用信息消费中直接消耗的字节(bytes)、单词(words)和时间(hours)三项信息消费的直接指标来测度信息消费水平。Roger 和 James 利用上述三项指标测度了 2008 年美国信息消费的总量^[7]。利用信息消费的直接指标来测度信息消费水平直观、明确,但由于数据收集较为困难,这类研究和应用在我国还不多见。

综上所述,目前信息消费水平测度这一课题已经受到了许多学者的关注,总体来看,这些学者的研究大多是利用单一指标(支出费用)对信息消费水平进行测度,不够全面和准确。随着移动互联网时代带来的信息消费多样性和免费 Wi-Fi 的普及,综合利用流量、时间和费用等多项指标,构建一个科学、合理和可操作的,适合我国居民信息消费特点的指标体系来测度信息消费水平是非常必要的,也是有意义的。

二、基于流量、时间和费用研究信息消费水平的必要性

流量、时间和费用均是反映消费者信息消费水平的指标^[8]。在移动互联网时代,由于流量、时间和费用三者之间的相关性减弱了,将流量、时间、费用三项指标结合起来,构建一个科学、合理和可操作的测度信息消费水平的综合评价体系是合理的,也是必要的。下面将从信息消费的过程、信息消费方式的多样性和信息化水平的提高三个方面进行理论分析。

(一)信息消费的过程

消费者行为学认为,消费是一个获取、使用和享受产品或服务的过程,所以信息消费也是一个获取信息、消耗信息、享受信息的过程。过去,由于网络覆盖率较低、带宽较窄,获取信息通常需要较高的费用,费用支出几乎是获取信息流量的唯一手段。但随着移动互联网的快速发展,尤其是免费 Wi-Fi 的普遍使用,消费者有更多的途径获取信息,流量和费用之间的高度相关性逐渐减弱。消耗信息流量是消费者信息消费的本质,费用支出只是消费者

获取信息流量的途径之一,因此流量和费用都是测度信息消费水平的重要指标。消费者在获取信息、消耗信息和享受信息的过程中都要花费一定的时间,而获取同样的信息,不同的消费者花费的时间不同,因此时间也是测度信息消费水平的重要指标。

(二) 信息消费方式的多样性

如果用流量、时间和费用三项指标来描述信息消费,每个指标按多、少进行二分,则有8种信息消费方式。随着智能手机的普及,移动互联网的发展,消费者获取信息更加便捷,信息消费方式也更加多样化,上述8种信息消费方式目前都普遍存在,其示例见表1。

表1 基于流量、时间和费用分类的8种信息消费方式

消费方式	举例
流量多、时间多、费用多	手机网络看付费视频,小说等
流量多、时间多、费用少	免费 Wi-Fi 下看视频,打游戏等
流量多、时间少、费用多	固定宽带条件下浏览付费视频等
流量多、时间少、费用少	免费 Wi-Fi 下浏览高清图片,逛淘宝等
流量少、时间多、费用多	阅读网上购买的创业企划书等
流量少、时间多、费用少	阅读下载的免费小说,玩手机小游戏等
流量少、时间少、费用多	购买信息设备,如手机,平板电脑等
流量少、时间少、费用少	浏览手机短信等

由表1可知,不同的信息消费方式在流量、时间和费用消耗上各有特点,任何单一的指标都不能全面、科学地表征当前多样化的信息消费方式。一般来说,不同的信息消费者具有不同的消费习惯,他们信息消费的主要方式一般是上述8种方式中的1~2种,所以不同的信息消费者在流量、时间和费用上的消费方式各不相同。

(三) 信息化水平的提高

过去,由于网络带宽较窄,网速较慢,通讯公司(电信、移动和联通)推出的上网套餐限定的流量额度较低而费用相对较高,使得费用和流量之间有着较强的相关性。随着信息化水平的提高,移动互联网的技术从2G升级到3G、4G,加上智能手机的普及,通讯公司推出了一系列复杂的套餐,使流量与费用之间的相关性趋于弱化。特别是免费 Wi-Fi 的普遍使用,使信息消费与费用之间的相关性进一步弱化,但与时间和流量之间的相关性进一步增强。消费者可以在费用较少甚至为零的情况下,花费很长时间进行流量消耗很多的信息消费活动,如利用免费 Wi-Fi 观看网络视频。所以费用这一单一指标不能全面测度信息消费的水平。

综上所述,信息消费的过程、信息消费的多样性和信息化水平的提高决定了流量、时间

和费用都是测度信息消费水平的重要指标。

三、基于流量、时间和费用的信息消费水平模型的构建

(一) 基本假设

根据消费者消费行为的习惯,即在短时间内,其信息消费的模式一般很少变化,我们可以做如下假设。

假设1:在一定时间内,一个消费者每次信息消费所花费的流量、时间和费用三项指标的比相同。

假设2:在一个地区,一定时间内,所有信息消费者的总信息消费的流量、时间和费用三项指标的比相同。

由假设1可知,一定时间内,某消费者每次信息消费在流量、时间和费用三项指标之间的比相同,所以某消费者在一定时间内所消耗的总流量、总时间和总费用的比值与任意单次信息消费中三项指标的比相同。

同理,由假设2,可将上述分析延伸到一个消费群体(如一个家庭或一个小区),得到与之相同的结论,即某一消费群体在一定时间内所消耗的总流量、总时间和总费用的比与任意单次信息消费中三项指标的比相同。

(二) 信息消费水平的综合评价模型

在移动互联网时代,流量、时间和费用均是表征信息消费水平的重要指标,为将三项指标转化为一个综合反映我国信息消费水平的指标,本文利用多指标综合评价法建立信息消费水平的综合评价模型。

1. 无量纲化处理

无量纲化处理方法的选择关系到模型结果的准确性和真实性。常见的无量纲化处理方法有“极差化”处理、“中心化”处理、“极小化”处理、“极大化”处理和“均值化”处理5种。

“中心化”处理是目前进行无量纲化处理最普遍使用的方法,虽然该方法可以消除数量级和量纲影响,但是各原始指标数据之间的差异性也被消除,所以中心化后的数据不能准确地反映原始信息。“极差化”处理、“极小化”处理以及“极大化”处理的分母只与数据的最大值和最小值有关,与其他数据无关。进行“极差化”处理时,如果极差很大,最后得到的无量纲指标值很小,相当于降低了该指标的权重;相反如果极差很小,最后得到的无量纲指标值很大,相当于增大了该指标的权重^[9]。用“极大化”和“极小化”进行无量纲化处理时,也会有同样的不足。“均值化”处理不仅可以消除各指标的数量级和量纲的影响,而且还能保留各指标变异程度的信息,是一种比较好的处理方法^[10]。综合考虑,本文采用“均值化”方法进行无量纲化处理。

为方便计算某地区任意信息消费者的信息消费水平和比较该地区内任意两个群体的信息消费水平,本文在用“均值化”进行无量纲化处理时,选取某地区内信息消费的流量、时间

和费用的均值。^①

设某地区信息消费的流量、时间和费用的均值分别为 $\bar{F}, \bar{T}, \bar{M}$, 共有 n 个信息消费者, 每个消费者在一定时间内共信息消费 m 次, $y_{ik1}, y_{ik2}, y_{ik3}, i = 1, 2, \dots, n, k = 1, 2, \dots, m$ 分别为某地区第 i 个消费者一定时间内进行的第 k 次信息消费在流量、时间和费用三项指标上的观察值, $x_{ik1}, x_{ik2}, x_{ik3}$ 为对应指标无量纲化处理后的值, 通过“均值化”无量纲化处理得到:

$$x_{ik1} = \frac{y_{ik1}}{\bar{F}}; x_{ik2} = \frac{y_{ik2}}{\bar{T}}; x_{ik3} = \frac{y_{ik3}}{\bar{M}} \quad (1)$$

显然, 由假设 1 可得出结论 1。

结论 1: 在一定时间内, 一个消费者信息消费的流量、时间和费用三项指标总量无量纲化处理后的比与一次信息消费三项指标无量纲化处理的比相同。

证明: 设 $y_{ik1} : y_{ik2} : y_{ik3} = a_1 : a_2 : a_3$, 其中 a_1, a_2, a_3 为常数, y_{i1}, y_{i2}, y_{i3} 分别为某地区第 i 个消费者一定时间内进行的信息消费在流量、时间和费用三个指标上的总消费值, x_{i1}, x_{i2}, x_{i3} 为对应指标无量纲化后的值, 由假设 1 可得:

$$\begin{aligned} x_{ik1} : x_{ik2} : x_{ik3} &= \frac{y_{ik1}}{\bar{F}} : \frac{y_{ik2}}{\bar{T}} : \frac{y_{ik3}}{\bar{M}} = \frac{a_1}{\bar{F}} : \frac{a_2}{\bar{T}} : \frac{a_3}{\bar{M}} = \frac{ma_1}{\bar{F}} : \frac{ma_2}{\bar{T}} : \frac{ma_3}{\bar{M}} \\ &= \frac{\sum_{k=1}^m y_{ik1}}{\bar{F}} : \frac{\sum_{k=1}^m y_{ik2}}{\bar{T}} : \frac{\sum_{k=1}^m y_{ik3}}{\bar{M}} = \frac{y_{i1}}{\bar{F}} : \frac{y_{i2}}{\bar{T}} : \frac{y_{i3}}{\bar{M}} = x_{i1} : x_{i2} : x_{i3} \end{aligned}$$

同理, 根据假设 2 也可以进一步得到结论 2。

结论 2: 在一段时间内, 一个消费群体消费的流量、时间和费用三项指标总值无量纲化处理后的比与一次消费三项指标无量纲化处理的比相同。

2. 权重的确定

权重是多指标综合评价中的重要因素, 本文从固定权重和可变权重两个视角分别探讨信息消费三项指标的权重设定问题。

(1) 固定权重

流量、时间和费用均是反映信息消费水平的客观性指标, 所以三项指标的权重可以设为固定值。假设通过专家评估或类似的方法确定的三项指标对信息消费水平的贡献率分别为 a, b, c , 且 $a + b + c = 1$, 则第 i 个消费者在流量、时间和费用三指标的权重可设为: $\omega_{i1} = a, \omega_{i2} = b, \omega_{i3} = c$ 。

(2) 可变权重

消费者进行信息消费的目的是为了享受信息消费带来的满足感, 也可称为信息消费的效用, 这种效用由信息消费中消耗的流量、时间和费用共同“贡献”。由假设 1 和结论 1 可以看出, 在一定时间内, 一个消费者每次信息消费所花费的流量、时间和费用三项指标的比相同, 三项指标无量纲化处理后的比也相同。即对一个信息消费者来说, 流量、时间和费用三项指标对每次信息消费效用的“贡献”率是不变的, 某个指标的比值越大, 其对信息消费

^① 为方便计算, 本文用样本的平均值代替某地区流量、时间和费用的均值。

效用的“贡献”率就越大,对应的权重也越大,因此可根据消费者在流量、时间和费用三项指标无量纲化处理后的比来确定权重,其具体权重的大小可由消费者花费的流量、时间和费用对其享受信息消费的满足感的影响,即效用函数来确定。为便于测度总体的信息消费水平,设消费者有相同的效用函数为 $u(x)$,则消费者在流量、时间和费用三项指标上的权重可

$$\text{设为: } \omega_{ij} = \frac{u(x_{ij})}{\sum_j u(x_{ij})}, j = 1, 2, 3。$$

由效用理论可知,对同一个信息消费者来说,由于每次信息消费时三项指标的比相同,因此三项指标对应的效用函数值是不变的,其权重也是不变的。对不同信息消费者来说,由于三项指标的比是不同的,因此三项指标对应的效用函数值是变化的,其权重也是变化的。对一个消费者来说,由于每次信息消费所花费的流量、时间和费用三项指标的比相同,我们有理由认为比值越大,其所对应指标的花费给消费者带来的信息消费满足感越大;比值越小,其所对应指标的花费给消费者带来的信息消费满足感越小。因此该效用函数可假设为中立类型或冒风险类型。

3. 信息消费水平的计算

目前,进行多指标综合评价的模型有多种,包括线性和非线性综合评价模型。本文为便于研究,无论是在固定权重还是在可变权重的情况下,均采用应用较为普遍的线性综合评价模型:

$$s_{ik}(x_{ik1}, x_{ik2}, x_{ik3}) = \omega_{i1}x_{ik1} + \omega_{i2}x_{ik2} + \omega_{i3}x_{ik3} \quad (2)$$

s_{ik} 表示第 i 个消费者进行第 k 次信息消费所产生的信息消费水平。由式(1)可以明显看出信息消费水平 s_{ik} 具有非负性、连续性和传递性,这三个特点是信息消费水平必须满足的,否则计算出的信息消费水平将失去现实意义。

事实上,可加性也是信息消费水平模型必须具有的。因为一个消费者或消费群体总的信息消费水平应该是消费者或消费群体每一次信息消费水平的加总。下面对式(2)的可加性进行证明。

设消费者 i 在一定时间内共进行了 m 次信息消费,由式(1)消费者在这段时间内花费的总流量、总时间和总费用可分别表示为: $\sum_{k=1}^m \bar{F}x_{ik1}$, $\sum_{k=1}^m \bar{T}x_{ik2}$, $\sum_{k=1}^m \bar{M}x_{ik3}$ 。根据结论1、结论2和固定权重与可变权重的设定可知,同一消费者在信息消费总流量、总时间和总费用上的权重与其每次信息消费在流量、时间、费用上的权重相同,结合式(2),消费者的信息消费水平 s_i 为:

$$\begin{aligned} s_i &= \omega_{i1}x_{i1} + \omega_{i2}x_{i2} + \omega_{i3}x_{i3} = \omega_{i1} \frac{\sum_{k=1}^m \bar{F}x_{ik1}}{\bar{F}} + \omega_{i2} \frac{\sum_{k=1}^m \bar{T}x_{ik2}}{\bar{T}} + \omega_{i3} \frac{\sum_{k=1}^m \bar{M}x_{ik3}}{\bar{M}} \\ &= \sum_{k=1}^m (\omega_{i1}x_{ik1} + \omega_{i2}x_{ik2} + \omega_{i3}x_{ik3}) = \sum_{k=1}^m s_{ik} \end{aligned}$$

所以该信息消费水平模型的可加性成立。

四、实证分析

2016年6月至7月,我们采用方便抽样的方法对南京邮电大学的350名大学生进行了信息消费问卷调查,并成功回收320份问卷,其中有效问卷305份。利用上述调查数据,验证本文提出的信息消费水平测度模型的正确性。

(一)信息消费中流量、时间和费用之间的关系

经Shapiro-Wilk检验发现调查数据中流量、时间和费用三项指标均不服从正态分布,所以本文通过流量、时间和费用三项指标间的Spearman相关关系来讨论三项指标之间的关系。表2给出了流量、时间和费用三项指标间的Spearman相关系数矩阵。

表2 流量、时间和费用三项指标间的Spearman相关系数矩阵

相关系数	流量	时间	费用
流量	-	0.23	0.15
时间	0.23	-	0.02
费用	0.15	0.02	-

由表2可以看出,费用与流量之间的相关系数为0.15,费用与时间之间的相关系数仅为0.02,这说明在移动互联网时代,费用与信息消费中的流量和时间的相关性减弱,费用的多少不再是决定信息消费水平高低的必要条件。由表2还可以看出,流量、时间和费用三项指标之间的相关系数的最大值只有0.23,这说明大学生信息消费中流量、时间和费用三项指标间几乎没有相关性,^①任一指标都不能被其他两项指标全面表征,因此利用流量、时间和费用三项指标共同测度信息消费水平是必要的。

(二)不同信息消费水平测度模型的比较

我们分别利用四种模型测度了大学生的信息消费水平的中位数。^②在模型1中,只利用单一费用指标来测度信息消费水平;模型2综合利用流量、时间和费用三项指标,通过固定权重综合测度信息消费水平,固定权重通过层次分析法确定;^③模型3综合利用流量、时间和费用三项指标,通过可变权重测度信息消费水平,可变权重基于中立的效用函数确定;模型4综合利用流量、时间和费用三项指标,通过可变权重测度信息消费水平,可变权重基于冒风险的效用函数确定。表3给出了四种模型测度信息消费水平和模型1与其他三种模型测度的信息消费水平之间的Spearman相关系数。

① 相关系数的 p 值小于0.3,可以视为变量之间几乎没有相关性。

② 由于模型采用的是“均值化”方法进行的无量纲化处理,处理后数据的均值为1,故这里用中位数代替均值对信息消费水平进行度量。

③ 在模型2中利用层次分析法确定的流量、费用和时间的固定权重分别为: $\omega_{11} = 0.65, \omega_{12} = 0.23, \omega_{13} = 0.12$ 。

表3 四种模型测度的信息消费水平和信息消费水平之间的相关系数

参数	模型1	模型2	模型3	模型4
信息消费水平	0.391	0.731	1.000	1.117
Spearman 相关系数	-	0.472	0.467	0.456
Wilcoxon 秩检验 p 值	-	$7e^{-10*}$	$2e^{-16*}$	$2e^{-16*}$

注: * 表示在 $p=0.01$ 水平上与模型1的结果有显著差异

由表3可知,模型1测度的信息消费水平最低,只有0.391,分别是模型2、模型3和模型4测度的信息消费水平的53.5%,39.1%和35.0%。对模型2、模型3和模型4进行Wilcoxon秩检验,在 $p=0.01$ 水平上都与模型1有显著差异。比较模型2、模型3和模型4测度的信息消费水平可知,基于固定权重的模型2测度的信息消费水平最低,而基于冒风险效用函数的可变权重模型4测度的信息消费水平最高。一般来说消费者都会努力追求自己信息消费效用的最大化,基于效用函数的可变权重测度的信息消费水平更加准确地表征了消费者的信息消费效用的最大化,因此模型3和模型4测度出的信息消费水平更高是合理的。

由表3还可知,模型1与模型2、模型3和模型4测度的信息消费水平之间的Spearman相关系数都小于0.5,这说明模型1测度的信息消费水平与其他三个模型测度的结果是低度相关。

上述结果说明利用单一费用指标测度的信息消费水平不但不能全面反映真实的信息消费水平,而且会显著低估真实的信息消费水平。

五、结论

在移动互联网时代,信息消费的方式更加多样,任何单一指标都不能准确、全面、科学地测度信息消费的真实水平,流量、费用和时间都是测度信息消费水平的重要指标,利用流量、费用和时间三项指标综合测度信息消费水平是可行的、必要的和有意义的。通过实证分析发现:流量、时间和费用均是衡量信息消费水平的重要指标,且三项指标之间几乎没有相关性;利用单一费用指标测度的信息消费水平与综合利用流量、时间和费用三项指标测度的信息消费水平有显著差异,而且会显著低估真实的信息消费水平。限于数据的可得性,基于流量、时间和费用三项指标构建的信息消费水平测度模型还有待于在更大范围内验证,这也是我们今后进一步的研究方向。

参考文献:

- [1] 张坤晶,易明.我国信息消费的现状与潜在问题[J].理论导刊,2014(1):94-96.
- [2] KELLERMAN A. Phases in the rise of the information society[J]. Info,2000(6):537-541.

- [3] 黄柳,况能富. 企业技术创新的信息消费评价[J]. 中国软科学,2002(6):2-5.
- [4] 张红历,梁银鹤. 中国省域城镇居民信息消费差异分析[J]. 情报科学,2016(2):9-14.
- [5] 丁志帆. 城镇居民信息消费的差异化福利效应研究——基于1993—2011年经验数据的数值模拟分析[J]. 财经科学,2014(2):51-61.
- [6] 李婵娟,王子敏,马万里. 农村居民信息消费省际差异的影响因素与贡献测度——基于shapley值过程的实证分析[J]. 情报科学,2017(8):146-152.
- [7] ROGER B, JAMES S. Measuring consumer information [J]. International Journal of Communication, 2012(1):980-1000.
- [8] 张卫华,赵铭军. 指标无量纲化方法对综合评价结果可靠性的影响及其实证分析[J]. 统计与信息论坛,2005(3):33-36.
- [9] 叶宗裕. 关于多指标综合评价中指标正向化和无量纲化方法的选择[J]. 统计科学与实践,2003(4):24-25.
- [10] 马哲明,靖继鹏. 国内信息消费研究综述[J]. 情报科学,2007(3):471-475.

(责任编辑:楼启炜)

On the measurement of information consumption in terms of flow, time and cost

HUANG Baofeng, LI Na

(School of Economics, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: The accurate measurement of information consumption is the basis for promoting information consumption. Based on the rapid development of mobile Internet, this paper analyzes the necessity of measuring information consumption in terms of the indexes of flow, time and cost from three aspects: the process of information consumption, the diversity of information consumption and informatization level; Under the assumption that information consumption is inertial in a certain period, and based on fixed weights and variable weights, a measurement model is constructed by using the indicators of flow, time and cost, and the measurement is proved to be nonnegative, transitive, continuous and increasable; Besides, a survey based on this model is analyzed. The results show: flow, time and cost are important indicators for measuring information consumption; an information consumption measured with the single cost indicator is significantly underestimated, and has a significant difference from the measurement in terms of three indexes of flow, time and cost.

Key words: mobile Internet; information consumption; measurement; utility