

上市公司信息质量综合优度指数构造

——基于奔福德定律

杨君岐(教授), 周雷雨

【摘要】 本文基于奔福德定律构建了针对上市公司单个指标信息进行质量评价的失真率指数,采用变异系数法确定了不同评价指标的权重,通过对各指标的失真率指数进行加权平均构造出了对上市公司信息质量评价的综合优度指数,并利用该方法对随机选择的2600多家沪深上市公司数据及49家ST公司数据进行了实证分析,取得了满意的效果。该方法不仅可用于上市公司信息的综合质量评价,而且可以为审计提供可能造假数据的痕迹。

【关键词】 上市公司; 奔福德定律; 信息质量; 优度指数

【中图分类号】 F239.1

【文献标识码】 A

【文章编号】 1004-0994(2016)06-0109-4

一、引言

上市公司是我国实体经济的主力军和核心力量,其产业发展水平对我国经济发展有着重要的影响。了解上市公司经营情况的重要方式之一就是获取上市公司披露的信息,但是近年来,信息披露造假、财务数据舞弊的现象屡禁不止。在评价上市公司信息质量方法方面,学者们进行了大量的研究。第一种方法是专业评分法,该方法对经验的要求比较高,具有较强的主观性;第二种方法是事件法,通过上市公司因违规被证券监管机构处罚的次数来衡量信息质量,客观性比较强,但是难以将一些隐蔽的违规行为纳入评价;第三种方法是理论模型法,例如收益不透明度等模型的构造,理论模型法具有一定的合理性,且操作简单。

近年来,有学者运用奔福德定律进行信息质量的评价,并取得了一定的成效。奔福德定律揭示了不同数字在自然数上出现的规律,通过大量的数据验证了数字分布的自然规律,在信息质量评价方面具有很强的客观性。以张苏彤教授为代表的学者主要是通过样本计算得到的1~9在首位分布的频率与奔福德定律的标准值的相关系数来进行信息质量的评判。该方法在识别信息是否造假方面具有较好的效果,但是没有对存在问题的信息进行定位捕捉,得到的评价结果(相关系数数值比较集中)不能较好地显示质量的差别,也未能解决基于多个指标的不同权重的信息质量综合评价问题。

本文首先基于奔福德定律构造不同数字在首位分布的失真率,由此来构造衡量单个指标信息质量的可信度指数,

帮助实现不真实信息的定位捕捉,然后利用变异系数法等来确定不同指标的重要性,最后通过加权计算各指标的可信度指数得到反映企业信息质量的综合优度指数。

二、基于奔福德模型的信息质量综合优度指数构造

要对上市公司信息质量进行客观评价必须考虑多个重要的指标,每个指标信息质量的好坏都影响着对上市公司信息质量的综合评价,但每个指标的重要性不一样,这里采用不同的权重来进行区别。权重的确定办法有很多,比如专家赋值法、层次分析法等,鉴于差别大的指标识别率高,本文采取变异系数法来确定权重。优度指数的构建是加权计算各个指标的信息质量指数的结果,为了方便大家使用该方法,文中给出了有关指数的EXCEL的实现途径。

为了方便叙述,不妨假设使用 n 个指标进行信息质量综合评价,指标集合为 $S=\{X_1, X_2, \dots, X_n\}$ 。

(一) 单个指标的信息质量评价

1. 数据预处理。为了方便数据处理,首先对原始数据进行预处理,使数据更加规范。对于众多首位为0的数据,可以根据小数位数,适当将数据整体乘以10的 n 次幂,消除首位为0的现象;对于出现负数的情况则取绝对值,在EXCEL中利用ABS函数可以取绝对值。

2. 统计首位数字出现的频率和奔福德标准值。

(1)首位数字出现1~9的频率。在EXCEL中利用“LEFT(数据源单元格,1)”可以提取各个数据的首位,利用函数“COUNTIF(数据源,i), $i=1,2,\dots,9$ ”统计数字 i 出现的频

【基金项目】 国家软科学研究计划重大项目“西部地区科技促进经济发展方式转变机制研究——以‘关中—天水’经济区为例”(项目编号:2011GXS2D031)

□ 借鉴与参考

数,对频数用SUM求和,求得数据总数,接着通过定义单元格公式,用“频数÷总数”,得到各数在首位分布的频率 f_{ji} ($j=1,2,\dots,n;i=1,2,\dots,9$),其中 i 代表首位数字 $1\sim 9$, j 代表第 j 个评价指标, f_{ji} 代表第 j 个指标首位 i 出现的频率。

(2)各位数字奔福德标准值计算。 $1\sim 9$ 各位数字在首位出现的奔福德标准值为 T_i ,数字 i 在首位出现的概率为 T_i ,计算见公式(1)。

$$T_i = \log_{10}(1+1/i) \quad i=1,2,\dots,9 \quad (1)$$

3. 单个指标首位数字出现 $1\sim 9$ 的失真率计算。

(1)各个数字失真率的计算。失真率是指样本频率与相对应的标准奔福德值的相对误差,见公式(2)。

$$UR_{ji} = \frac{|f_{ji} - T_i|}{T_i} \quad j=1,2,\dots,n;i=1,2,\dots,9 \quad (2)$$

其中, UR_{ji} 表示第 j 个指标 i 在首位分布的失真率。

(2)单个指标的信息质量失真率计算。平均失真率是对 $1\sim 9$ 的失真率进行算术平均的结果,是对单个指标整体失真率的衡量。平均失真率计算见公式(3)。

$$UR_j = \frac{\sum UR_{ji}}{9} \quad j=1,2,\dots,n;i=1,2,\dots,9 \quad (3)$$

其中, UR_j 表示第 j 个指标的平均失真率。

在EXCEL中计算只需对单元格定义公式即可。奔福德定律揭示的是数字的自然规律, $1\sim 9$ 在首位分布的失真率越大,说明偏离标准值越远,数据可疑性越大,可以以此为线索,缩小范围追踪到造假数据的存在。一般来说,各个数字失真率在0.1以上说明数据比较可疑。平均失真率指数越大,说明信息质量越差,平均失真率超过0.05,则信息质量存在可疑。

4. 可信度指数。为了构造单个指标的信息质量评价的正向指标,本文用“1-指标的平均失真率”来构造可信度指数,衡量信息的真实可信度,计算见公式(4)。在EXCEL中计算只需对单元格定义公式即可。

$$FR_j = 1 - UR_j \quad j=1,2,\dots,n \quad (4)$$

其中, FR_j 表示第 j 个指标的可信度指数。

可信度指数越大,则说明信息质量越好;可信度指数低于0.95,则说明数据信息可能存在人为造假。

(二)上市公司多指标信息质量评价优度指数

依据上述办法对指标 X_1, X_2, \dots, X_n 进行评价,得到各指标的信息质量评价的可信度指数 FR_j ($j=1,2,\dots,n$)。

下面分别确定各个指标的权重,再利用加权平均的办法构建上市公司信息质量综合评价的优度指数。

1. 各指标权重的确定。权重确定的方法有专家评价法、层次分析法、变异系数法等方法。鉴于各指标的数据已采集,本文采用变异系数法来进行权重的确定。不妨设第 j 个评价要素 X_j 权重为 W_j ,均方差为 s_j ,平均值为 \bar{X}_j ,均方差系数

为 V_j ,均方差系数和权重的计算分别见公式(5)和公式(6)。均方差的计算可以在EXCEL中通过STDEV函数实现。

$$V_j = \frac{\sigma_j}{\bar{X}_j} \quad j=1,2,\dots,n \quad (5)$$

$$W_j = \frac{V_j}{\sum V_j} \quad j=1,2,\dots,n \quad (6)$$

2. 上市公司信息质量综合评价优度指数。利用上面计算的权重和可信度指数加权计算求得关于上市公司信息质量优度指数QEI,计算见公式(7)。

$$QEI = \sum W_j \times FR_j \quad (7)$$

三、案例分析

为了进一步检验模型的有效性,本研究采取两大类不同的上市公司进行分析。案例1是对随机抽取的沪深两市2600多家上市公司的部分数据进行分析,案例2是对随机抽取的49家ST公司数据进行分析。

(一)沪深两市2600家上市公司数据信息质量综合评价

本文收集了沪深两市2600多家上市公司2014年营业总收入、管理费用、营业总成本、营业税金及附加、财务费用、营业外支出、销售费用七项数据。

1. 描述性统计。收集的2600多家上市公司从其所处行业的角度来看,涉及制造业、软件和信息技术服务业、生态保护和环境治理业、银行金融业、建筑业等各行各业,覆盖面比较广。有关指标的计算结果见表1。

表1 上市公司描述性统计

	最大值(元)	最小值(元)	平均值(元)	标准差	变异系数
营业总收入	2115648000000	48168	7999417011	61350990633	7.669433
管理费用	63035000000	1499678	297636554	1811586505	6.086573
营业总成本	2048659000000	1845662	6434574975	56067133246	8.713417
营业税金及附加	172346000000	-31625697	267989234	4547882410	16.97039
财务费用	19693000000	-893079683	106248559	564051965	5.308796
营业外支出	5261000000	-50842391	11575279	121518232	10.49808
销售费用	121332000000	1001	579494628	5042215397	8.701056

如表1所示,营业总收入和营业总成本两项指标的最大值和最小值相差最大,其余指标之间的距离相对来说不大。营业税金及附加、营业外支出的变异系数最大。营业总收入最大的是中国石化,在2115亿元以上,最小的是ST宜纸,只有40万左右。营业总收入排名前十的主要为石油行业、建筑行业以及银行,分别是中国石化、中国石油、中国建筑、工商银行、上汽集团、中国中铁、建设银行、中国铁建、农业银行、中国人寿,这些企业营业总成本也都比较高,排名靠前。营业总收入排名垫底的主要以轻工业为主。

2. 信息质量综合评价。

(1)根据公式(1)、公式(2)、公式(3)计算得到各个财务指标的失真率、平均失真率,具体见表2。

表2 失真率和平均失真率

首位数	失真率						
	营业总收入	管理费用	营业总成本	营业税金及附加	财务费用	营业外支出	销售费用
1	0.024674	0.080399	0.049402	0.031694	0.032611	0.000184	0.009027
2	0.039825	0.074865	0.020078	0.023161	0.039537	0.036145	0.041687
3	0.057639	0.051837	0.051837	0.10548	0.042502	0.004781	0.004042
4	0.033457	0.147431	0.039107	0.036589	0.001013	0.054608	0.010856
5	0.091714	0.119545	0.011519	0.137586	0.119545	0.053669	0.08332
6	0.027316	0.086028	0.012702	0.037308	0.045374	0.001639	0.057216
7	0.004683	0.139752	0.066003	0.075276	0.041268	0.048744	0.006867
8	0.042236	0.042293	0.004289	0.192985	0.004289	0.009193	0.05145
9	0.084548	0.099315	0.05683	0.013519	0.036637	0.18206	0.086251
平均失真率	0.045121	0.093496	0.034641	0.072622	0.040308	0.043447	0.038969

从各个指标的失真率来看,营业总收入、营业总成本和销售费用首位数为1~9的失真率都较小,普遍质量比较好。在其他指标中,管理费用首位数为4的失真率最大,营业税金及附加首位数为8的失真率最大,财务费用首位数为5的失真率最大,营业外支出首位数为9的失真率最大,说明以这些数开头的数据存在较大的造假嫌疑,数据质量不高,可以以此为线索进行信息失真原因的追查。

从平均失真率来看,管理费用和营业税金及附加的平均失真率比较大,并且大于0.05,说明这两项指标信息质量不高,应该加大对这两项数据的审查力度。

(2)根据公式(4)、公式(5)、公式(6)、公式(7)计算得到各个指标的可信度指数和权重,并利用CORREL函数求得相关系数,具体结果见表3。从表中可以发现,可信度比较低的有管理费用和营业税金及附加,都在0.95以下,依据张苏彤教授相关系数判断法来看,这两项的相关系数小于0.998,也属于比较可疑的信息造假项。

表3 相关指数计算结果

	营业总收入	管理费用	营业总成本	营业税金及附加	财务费用	营业外支出	销售费用
相关系数	0.998237	0.993878	0.998066	0.997391	0.998464	0.998628	0.999022
可信度	0.954879	0.906504	0.965359	0.927378	0.959692	0.956553	0.961031
权重	0.119933	0.09518	0.136258	0.265379	0.083018	0.164167	0.136065

(3)绘制相关系数和可信度指数的图表,见图1。从中可以看出,可信度指数与相关系数具有一致性,但是比相关系数更能突出数据之间的差异,因为可信度指数之间的差距比较大。可信度指数大于0.95,则信息质量比较可靠;小于0.95,则信息质量可疑。

根据表3数据求得质量综合评价指数为0.9459,具体计算如下:

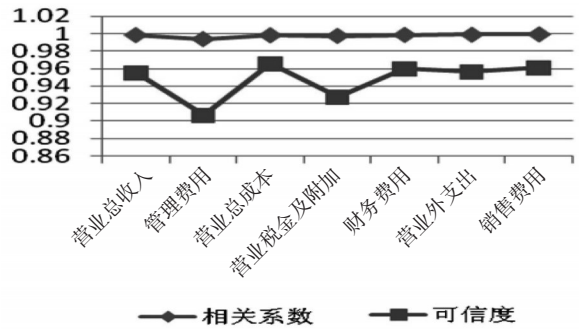


图1 相关系数和可信度

$$QEI=0.954879 \times 0.119933 + 0.906504 \times 0.09518 + 0.965359 \times 0.136258 + 0.927378 \times 0.265379 + 0.959692 \times 0.083018 + 0.956553 \times 0.164167 + 0.961031 \times 0.136065 = 0.9459$$

综合指数接近0.95,说明整体信息质量较好,但存在着造假的可能性。这说明我国大多数上市公司的信息披露质量比较真实可靠,但小部分的公司数据存在人为修改、造假的可能。为了进行对比,下面对ST公司信息质量进行评价。

(二)运用EXCEL实现ST公司单指标信息质量评价

本文收集了49家ST公司2014年的部分数据,涉及营业总收入、管理费用、营业总成本、营业税金及附加、财务费用、营业外支出、销售费用。ST公司是被特殊处理的公司,说明这些公司在财务状况上出现异常,比如连续亏损,或者是其他异常情况。

1. 数据预处理。本案例中只有营业外支出的数据存在负数情况,需要对其取绝对值,利用ABS函数进行处理;其他数据都为较大整数,不需要进行小数位处理。

2. 首位统计。由于ST公司进行分项计算的数据样本量偏小,因此对它们首先进行整体的信息质量评价。首先提取各项数据的首位,以第一列营业总收入为例,*ST钢构的首位提取需在D2单元格中定义公式“LEFT(C2,1)”,取得首位为6,然后进行下拉,提出所有首位。

3. 统计频数。进行1~9在首位出现的频数统计,以营业总收入为例,对于*ST钢构,在E2单元格中定义公式“COUNTIF(D2:D50,1)”,统计得到1出现的频数,在E3单元格中定义公式“COUNTIF(D2:D50,2)”,依次类推,得出1~9在首位出现的频数。由于有些公司数据不全,所以每项财务指标的数量是不一样的。在E12单元格中定义SUM求和公式,求得频数总数,即每列数据的数量。管理费用等按照此方法定义公式,或者直接复制粘贴营业总收入的统计单元格,可以得到其他财务指标的统计数据。对各项财务指标1~9在首位出现的频数求和,在Y2单元格中定义公式“SUM(E2, H2, K2, N2, Q2, W2, T2)”,下拉公式求得1~9各个数的出现频数,以及Y12单元格中的数据总量。

图 2 EXCEL 中 ST 公司信息质量评价指数计算

4. 计算频率和失真率。在 Z2 单元格中定义公式“Y2/\$Y”,得到 1 在首位出现的频率,公式下拉,得到全部频率。标准值的公式为“AA2=(LOG10(X2))”,各个数的失真率“AB2=ABS(Z2-AA2)/AA2”,平均失真率“AC2=AVERAGE(AC2:AC10)”,计算可信用指数通过对 AD2 单元格定义公式,图 2 为计算结果。

5. 结果分析。通过数据可以发现,平均失真率为 0.16,大于 0.05,说明存在造假嫌疑,同时首位为 2、3、4、7、8、9 的失真率都超过 0.1,首位为 9 的失真率达到 0.28 以上,失真严重,可信用指数为 0.84,小于 0.95。综合来看,该数据样本造假征兆明显,应该按照这些失真率由大到小进行造假证据追查。

根据实际调查和资料收集发现,2014 年*ST 国创、*ST 新都、*ST 国恒都出现了审计无法给出意见的情况,存在财务信息造假,再次验证了失真率对信息质量评价的有效性。

四、结束语

本文基于奔福德定律构造出一种上市公司信息质量评价的新方法,利用该方法对沪深两市普通上市公司和 ST 公司计算分析发现,ST 公司信息质量普遍不如一般上市公司,信息调整和造假的可疑性比较大。相对于 ST 公司,就选择的有关指标来看,我国上市公司数据质量基本可信,管理费用和营业税金及附加的可信用比较低,这也与上市公司普遍采用“税收筹划”有关,证明该方法具有非常好的实用性。

本文在对企业进行信息质量评价时,提出了利用失真率来进行信息失真的定位,利用可信用指数进行单个指标的信息质量判断,同时构建基于指标重要性差异的信息综合评价优度指数来进行整体的质量判断,完善了现有的奔福德定律在信息质量评价方面的应用理论。1~9 在首位分布的失

真率大于 0.1,说明存在信息失真可能;质量综合评价指数和可信用在 0.95 以下,说明信息存在造假可能。

信息质量综合评价优度指数可以用来进行单个公司各年财务数据的评价,可以观察公司信息质量变化的轨迹;也可以对各个公司之间的信息质量进行评价,以便横向比较;此外,可以对某项防止、减少信息造假的政策或者措施的实施效果进行结果评估。

根据已有的研究成果,总量指标比较符合奔福德定律,而资产负债率等比率型指标则不太符合,不能利用奔福德定律进行质量评价。但是,由于比率型指标也是由总量指标计算得到的,可以利用总量指标的失真率和可信用对比率型指标进行评价,具体评价方法还需进一步研究。

主要参考文献:

张苏彤,康智慧.信息时代舞弊审计新工具——奔福德定律及其来自中国上市公司的实证测试[J].审计研究,2007(3).
 王福胜,李勋,孙逊.奔福德定律在审计领域的应用[J].财会月刊,2007(8).
 刘福清,陈义文.浅谈奔福德定律在审计中的应用[J].会计之友,2010(12).
 王大江.基于奔福德定律的现金流量表审计质量研究[J].会计之友,2014(25).
 许存兴,张芙蓉,吕国钧.基于奔福德定律的上市公司审计意见效应分析[J].财会通讯,2010(24).
 邱兆东,彭桃英.审计舞弊侦查方法研究[J].财会月刊,2014(5).
 作者单位:陕西科技大学管理学院,西安 710021