

Analysis of Differences and Demands of New and Old Income Standards Based on Text Mining

Xuejiao Huang, Jiantao Lai, Jian Liu, Huqin Yan

Xiamen National Accounting Institute, Xiamen Fujian

Email: 524088993@qq.com, 1435488610@qq.com, 1145777124@qq.com, yanhuqin@xnai.edu.cn

Received: May 1st, 2020; accepted: May 15th, 2020; published: May 22nd, 2020

Abstract

Since the new income standard was issued in 2017, it has been paid close attention to. The standard of income standard has an important impact on the financial statements and the financial data submitted to the outside world. Therefore, this paper uses python programming language, through the text mining of the old and new income standards and accounting vision forum posts, to make word clouds, and then compare and contact, sum up the differences between the old and new standards. At the same time, through the analysis, it is concluded that accountants are still in the stage of learning and exploring the practical application of the new standards, and hold a more positive attitude. Based on this, we also put forward some suggestions to make the income standard play a better role in economic activities.

Keywords

New Revenue Guidelines, High-Frequency Words, Hot Words, Accountants

基于文本挖掘的新旧收入准则差异及诉求分析

黄雪娇, 来建涛, 刘 剑, 阎虎勤

厦门国家会计学院, 福建 厦门

Email: 524088993@qq.com, 1435488610@qq.com, 1145777124@qq.com, yanhuqin@xnai.edu.cn

收稿日期: 2020年5月1日; 录用日期: 2020年5月15日; 发布日期: 2020年5月22日

摘 要

新收入准则自2017年发布至今一直备受关注, 收入准则的规范对财务报表以及对外报送的财务数据有着

重要影响。因此,本文利用Python程序语言,通过对新旧收入准则以及会计视野论坛帖子进行文本挖掘,制作词云,进而对比、联系,总结出新旧准则着重关注点的差异。同时,通过分析得出会计师对于新准则目前还处于学习与摸索实务运用的阶段,且持有较积极的态度结论。基于此,我们还提出了几点建议,使收入准则在经济活动中更好的发挥作用。

关键词

新收入准则, 高频词, 热点词, 会计师

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

会计准则是企业财务状况、经营成果及现金流量对外表达的规范标准。因此会计准则之于企业经营与经济社会有效运行具有一定的指导作用[1]。其中,为满足持续经营条件,规避破产清算,则必然要有收入产生以支持企业开展各项活动。所以,收入准则对企业的底层业务与活动进行了规范,保证了企业基础财务数据的规范性,也对经济社会的有效运行提供了基础规范作用。收入准则不仅关乎企业、经营者、监管层,对经济社会整体运行也有重要意义,因此,值得引起我们的关注。

2017年7月19日,财政部发布了新收入准则,并规定各类企业分期执行,业界对此的关注度至今仍持续不减。因此,挖掘新收入准则与旧收入准则的关注侧重点的差异,对于企业及时调整业务,适应准则变化具有重要意义。同时挖掘分析会计师对于新准则的态度反应,则有利于准则制定者的政策调整[2]。

因此,本文结合当前流行的文本挖掘方法,通过对会计准则委员会发布的新旧收入准则以及会计视野论坛有关新收入准则的252个帖子进行文本抓取,提取高频词,再通过对高频词的分析,发现新旧准则的特点以及会计师对新准则的态度。最后,通过新旧准则的对比分析,会计师态度与新旧准则的关系分析,发现准则变迁方向及提出我们的建议。

2. 数据描述与方法

2.1. 使用数据获取与数据预处理

本文一共采集了新旧两版收入准则以及与收入准则有关的252个帖子。新旧收入准则来自于财政部会计准则委员会官方网站<https://www.casc.org.cn/casc/zhence/quanwen/>,而252个帖子来源于对会计视野论坛新收入准则关键词相关内容的抓取。本文主要通过Python各个库函数的运用以及统计产品与服务解决方案(statistical product and service solutions, SPSS)的使用实现文本的抓取、处理与分析。首先通过BeautifulSoup4库函数对HTML的解析功能,实现对新旧会计准则及相关帖子的爬取。然后,因为现在已经有许多成熟的程序包可以调用,所以本文用现有的成熟数据统计软件及手工的方式对采集的文本预处理,再运用jieba库函数对文本进行分词处理[3]。将采集到的文本进行分词处理,即依据语义将一句话切分成一个个的词,例如,“识别与客户订立的合同”分词成:“识别”“与”“客户”“订立”“的”“合同”。句子里一些没有意义的停词,如例子中的“与”“的”将其过滤掉,该过程使用jieba自带的停词词库,实现停词过滤。分词这一步尤为重要,分词效果的好坏直接影响了后面分析的准确性。依据

jieba 包里自带的词典，文本被分为一个个的词后[4]。为了提取有用的信息，需要进一步过滤掉一个字的实词，比如文本中的“有”“向”等词。最后统计新旧准则及帖子的词频。截取排序前 25 的词，见表 1。

为了让统计结果更加直观，便于寻找主要属性特征，用 Python 的 Cloudword 库函数依据统计的词频制作词云图。鉴于一些高词频结果与研究目的无关(如关于新收入准则的词频统计里，“会计准则”一词出现的频率极高，但是，对于本研究问题却没有实际意义)。因此，筛选掉无意义的高频词，制作得到词云图，如图 1~3 所示。

Table 1. Word frequency

表 1. 词频

词序	新收入准则		旧收入准则		帖子	
	词	词频	词	词频	词	词频
1	企业	187	收入	44	收入	669
2	合同	156	企业	34	业务	262
3	商品	143	劳务	33	探讨	220
4	客户	119	商品	28	确认	215
5	履约	78	销售	27	企业	206
6	义务	69	金额	17	CPA	194
7	转让	53	确认	15	合同	175
8	确认	52	资产	12	公司	145
9	收入	50	合同	11	客户	130
10	对价	48	确定	11	请教	85
11	资产	41	成本	10	销售	80
12	履行	40	可靠	9	问题	78
13	确定	34	计量	8	商品	53
14	取得	32	交易	8	履约	44
15	成本	32	完工	8	是否	43
16	承诺	31	相关	7	服务	36
17	交易价格	30	价款	7	执行	30
18	单独	29	折扣	7	成本	25
19	金额	26	政策	6	对价	21
20	售价	24	应收	6	软件	21
21	变更	23	会计	5	义务	20
22	销售	23	让渡	5	案例	17
23	价值	21	使用权	5	折扣	15
24	分摊	20	售出	5	科目	12
25	可变	20	当期	5	产品	11
合计		1381		333		2807

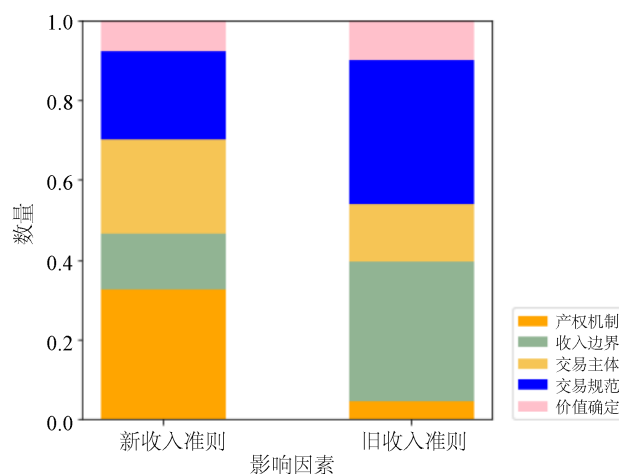


Figure 4. The proportion of characteristic in the previous principle and present principle

图 4. 新旧准则属性占比

Table 2. Characteristic and proportion of the Word frequency in the income principle

表 2. 准则词频属性与比重

属性	新准则		旧准则	
	词频	比重	词频	比重
产权机制	426	0.32569	11	0.04641
收入边界	184	0.14067	83	0.35021
交易主体	306	0.23394	34	0.14346
交易规范	291	0.22248	85	0.35865
价值判断	101	0.07722	24	0.10127
合计	1308	1.00000	237	1.00000

Table 3. Characteristic and proportion of word frequency of willingness

表 3. 诉求词频属性与比重

属性	词频	比重
学习	383	0.29037
考证	211	0.15997
准则	338	0.25625
实务	387	0.29340
合计	1319	1.00000

3. 热点话题分析

3.1. 收入准则分析

3.1.1. 新收入准则分析

新收入准则出现频率最高的是产权机制，累计词频达到 426 次。其次是交易主体和交易规范，分别达到 306 和 291 次。收入边界和价值判断较少，占比分别为 14.07%和 7.72%。因此可以看出，新收入准

则对企业及经济社会的引导也是偏向建立现代产权制度，厘清权责的经济环境。产权归属的前提是先确定主体，因而交易主体的确定是产权机制的建立基础。交易主体的确定为交易客体的归属权、控制权、收益权等的权属问题提供了研究对象。而交易规范则为产权机制的建立提供了程序保证。看似不相关的归属词，在会计准则的体现中，却表现出了相关性和整体性，都是围绕产权机制、权责分明建立的底层规范。

从高频词分析，产权机制表现在两个方面：一个是交易主体之间的产权关系，一个是合同内部不同商品的关系。“产权”“主体”、“客户”“履约”“义务”等词汇表现了交易主体之间的产权关系。新收入准则在广泛情景下规定了如何断定“商品”“服务”等的所有权权属及权属转移问题。微观来说，有利于企业合理判断交易完成程度，收入是否确认，保证了财务信息的真实可靠。宏观来说，有利于减少经济摩擦，提高经济运转效率。“分摊”“可变”“单独”等词汇则表现了合同内部不同商品的“产权”关系。新收入准则规范了如何划分“捆绑销售”，有利于对销售收入在不同商品间划分，同时用合同履行成本等科目对费用进行归集。有利于成本与收入的配比，实现商品之间的权责明确。从而有利于计算产品毛利，成本分析等活动的进行。有利于责任会计的实施，有效归因，有效评价及及时反馈与学习。

3.1.2. 旧收入准则分析

旧收入准则中出现频率最高的是收入边界和收入规范，其占比都超过 35%。交易主体、价值判断和产权机制则占比较低，其中以产权机制为最低，仅占比 4.64%。从各个热点话题来看，旧的收入准则则比较关注收入边界和交易规范二者。着重通过描述准则约束什么和如何约束来对企业收入确认进行规范和指导。旧收入准则对外通过明确规范的边界。对内“商品”“劳务”“使用权”等词，体现了旧准则对收入的分类思想。通过对收入进行分类，然后再通过收入规范对收入进行规定，使得准则对业务的规定更加的条理化。

收入边界的确定有利于厘清准则约束范围，明确准则与准则之间的衔接与各自的职责。但是以强调约束边界的准则则忽略本准则内部各部分之间的联系，更偏向于模块化的管理。同时，重视收入边界和收入规范的旧准则，更体现为一种流程思维。先通过把准则约束的范围框定，然后再对框定范围内的业务进行规定，这是典型的流程思维方式。与新准则体现了较为明显的差异性。这样的准则更体现为一种规定式的制度，而不是以核心理念为主的引导式制度。

3.1.3. 新旧准则对比分析

1. 相关性分析

新准则是在旧准则制定之后出现，都是为了规范同一类业务，因此新准则与旧准则之间是否具有传承性与相似性，就非常吸引人的思考。本文通过对统计产品与服务解决方案(statistical product and service solutions, SPSS)双变量相关性检验的运用，统计检验新旧准则之间是否具有相关性。具体情况见表 4。通常情况下，显著性大于 0.05 则为不显著。而新旧收入准则之间相关性为-0.262，显著性为 0.670。因此，由 Pearson 相关性可以看出二者具有不显著的负相关关系。从相关性检验的结果来看，不仅新准则不存在与旧准则的传承关系，而且存在一定的革新成分。这与我们的词频统计及热点词分析具有一致性，二者在关注侧重点上存在很大的差别。新准则侧重产权机制的关注，而旧准则则更关注收入边界。同时，二者的理念也存在较大的差异性。

2. 理念分析

从高频词和热点词的内在关系逻辑上来说，新收入准则更注重准则内部的统一性与一体化。而旧准则更偏重分类思想与流程管理。虽然都是为了收入业务的规范而制定，但是二者在制定理念上存在较大

的差异。具体如图 5、图 6 所示。

Table 4. Relevance and significant difference between the previous principle and the present principle

表 4. 新旧准则相关性与显著性

	新收入准则			旧收入准则		
	Pearson 相关性	显著性(双侧)	N	Pearson 相关性	显著性(双侧)	N
新收入准则	1		5	-0.262	0.670	5
旧收入准则	-0.262	0.670	5	1		5

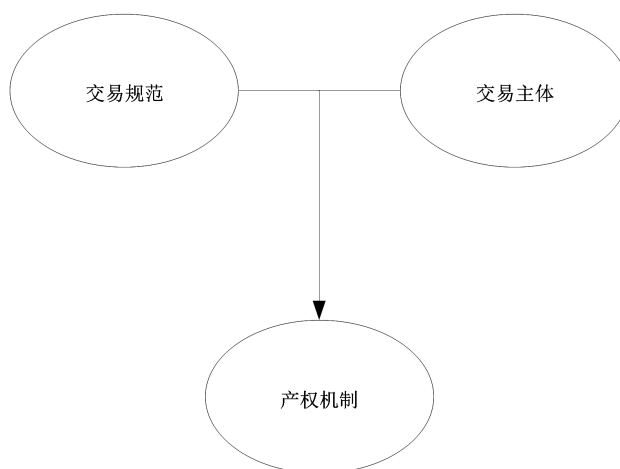


Figure 5. Logistic thinking in the present income principle

图 5. 新收入准则理念逻辑

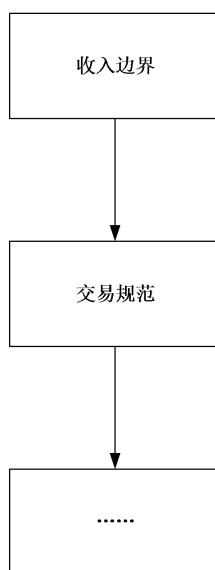


Figure 6. Logistic thinking in the previous income principle

图 6. 旧收入准则理念逻辑

3. 词频数分析

对于收集的排前 25 位的高频词进行统计, 新收入准则的高频词共计 1381 个, 而旧收入准则则只有

333 个。因此，从高频词的数量上，我们可以看出新收入准则的内容是旧收入准则的几倍。其规范的细致程度以及完善程度应该远超旧收入准则。准则的完备性能够降低准则运行出错的可能性，使得新准则能够更好地为企业财务计量提供基础。

4. 细节分析

因为收入边界、产权机制、交易主体新旧准则存在很大差异。同时，价值判断二者在热点词中占比均较低。而收入规范是二者共同的重点内容。因此，我们尝试通过对二者之间内部的高频词进行分析发现二者是否在细节上具有差异性。刨除共同存在的词汇，新收入准则更多的出现“对价”“单独”“变更”“分摊”等词汇，而旧收入准则更多的出现“可靠”“计量”“完工”“折扣”“政策”等词汇。从中可以看出，新收入准则在收入规范中更倾向于将合同的各个部分进行拆分与责任匹配，而旧收入准则更倾向于保证会计确认的谨慎性及对特殊情况的处理。因此可以看出，二者虽然都具有交易规范属性的词，却具有不同的偏向性，存在较大的理念差异。

3.2. 会计师态度分析

3.2.1. 热点词分析

通过对高频词进行归纳整理，发现会计师的态度主要集中在学习、考证、准则本身以及实务四个方面。其中实务和学习出现频率最高，分别出现了 387 和 383 次，分别占比 29.34%和 29.04%。考证出现频次最低，为 211 次，占比 16.00%。高频词展现了目前会计师们对于新准则的考虑方向与工作困扰。

因为学习和实务的出现比率极高，因此可以看出目前会计师们对于新准则的掌握和理解还不够深入，正处于互相交流与学习的阶段。同时，根据前面分析，新准则与旧准则存在理念上的根本不同，且内容含量是旧准则的数倍。因此，会计师们处于目前对新准则还处在学习与交流阶段也实属正常。对实务热点词进行分析，发现其下属高频词呈现多样化，说明各行各业的会计师均对新准则具有不同程度的疑问。对于考证的关注，也侧面印证了该观点。因为准则的变化才导致了考试内容的变化，因此而引起会计师对此关注度的提高。

通过分析，也可以看出会计师对于收入准则的变迁，关注点主要在实际方面，不管是实务的关注还是学习与考证的关注，都是关注准则的用法以及如何使自身适应收入准则的变化这类实际问题，而较少的关注准则变迁的理论问题及影响。

3.2.2. 情感分析

仅通过对热点词的分析不足以看出会计师对新收入准则持有的情感态度。因此，通过观察取出的高频词汇，对高频词汇进行分类，分别分为消极词汇、中性词汇和积极词汇[5]。并对各类词汇进行赋分-1、0、1。通过各类词汇的词频进行加权平均计算，可以看出会计师对新收入准则的态度，具体如表 5。总体来看，会计师对于新收入准则呈现中性偏积极的心态。同时，从获取数据的 252 篇帖子中随机取样 10%，26 篇进行阅读分析，发现会计师态度与分析结果基本一致。会计师们发帖主要是询问问题、表达疑惑与分享经验。其对准则变化发生的原因及影响缺少探究，同时对准则自身的消极与积极影响也缺少表达。重点放在了如何学习准则与如何适应准则变化的影响上。

4. 对策与建议

4.1. 对准则制定者的建议

从会计师对新会计准则的态度来说，会计师们对于准则大多处于学习阶段，对于实务中的问题存在较多的疑惑。因此，准则制定者应该尽到自己的责任。努力实现准则制定与执行之间的对接。对会计准

则的实施与执行情况进行跟踪调查,从而针对性解决会计准则在执行中遇到的各种问题。撰写与发布准则执行指引与指南,为会计师理解会计准则提供参考。充分发挥财政部下属的三家国家会计学院的培训能力,开展新准则的培训工作。编写参考书目以及推动制作网络课程,从而让大量的会计能够熟悉并理解新收入准则,有利于新准则的实施。

Table 5. The proportion of word attitude
表 5. 情感词汇比例

	消极词汇	中性词汇	积极词汇
词频	78	2429	300
比例	0.10688	0.86534	0.02779
分数	-78	0	300
总分		222	

4.2. 对企业的建议

因为新收入准则相对旧收入准则有很大程度的改变,所以企业首先要了解本企业实施新准则的时刻表,做好新旧转换的衔接工作,一方面要做到会计合规,另一方面要避免新旧准则造成的会计收益的波动。同时对内部会计师进行培训,保证新收入准则能够正确实施。并且,基于文本挖掘我们发现新收入准则着重关注现代产权制度的建立。因此,为了适应新收入准则的变化,企业应该对外重视与交易企业的权责关系的划分与执行,重视合同在交易中的地位。有必要的要形成自己的律师团队或者强化法律咨询在业务中的重要性,规避收入准则实施的法律问题,也是适应产权保护制度日渐清晰明确的趋势。另一方面,要注重责任会计的实施,使得成本与收益在不同责任主体之间划分,从而有利于新收入准则的实施,并顺应准则所暗含的关注方向。

4.3. 对会计师的建议

从会计师的态度中,我们可以看出会计师对于新准则一直处于学习与交流的氛围中,可以看出大部分会计师对于新收入准则并不熟悉。建议会计师们能够持续这种交流与学习,同时能够扩大交流与学习的对象。走出专业圈子,能够与业务人员进行接触,这样才能够更好的理解公司业务与收入情况的机理与实现过程,从而更好的把准则应用在自己公司的业务上,使财务数据更真实有效的反映业务。同时,强化会计准则的学习,实现对新收入准则的掌握。

基金项目

本论文得到了厦门国家会计学院 2019 年“云顶课题: Python 财务数据分析”项目的支持。

参考文献

- [1] 周大琳. 新旧收入会计准则对比分析[J]. 纳税, 2019, 13(34): 160.
- [2] 吴霏雨. 新旧收入准则对收入的确认对比分析[J]. 商业会计, 2019(21): 59-62.
- [3] 宋英慧, 黄麒. 基于文本挖掘技术的财务报表附注披露研究[J]. 会计之友, 2019(1): 142-147.
- [4] 范珈瑜. 基于文本挖掘的游客对古镇旅游态度的分析[J]. 大数据, 2017, 3(6): 93-101.
- [5] 黄方亮, 崔红燕, 任晓云, 李字庆, 张晓波. 年报管理层讨论与分析的语调倾向——基于 A 股市场的检验[J]. 投资研究, 2019, 38(5): 19-42.