

年度报告的可读性水平、过往业绩与分析师预测

朱丹,李静柔,李世新

(重庆大学 经济与管理学院,重庆 400044)

[摘要]以我国2007—2018年的公司年报为样本,从普通中文文本复杂性和专业词汇密集度两个层面度量上市公司年报文本的可读性水平,研究发现:我国分析师在整体上对可读性水平低的年报采取了本能的“防御”措施,在盈利预测中对复杂年报的态度更加谨慎,且复杂年报确实加大了分析师预测的分歧程度;作为盈利预测的基准,过往业绩水平具有功能锁定效应,在年报文本可读性对分析师预测的影响关系中发挥了调节作用。进一步研究发现,分析师的专业能力在解读复杂年报和摆脱过往业绩的功能锁定效应中发挥了重要作用。

[关键词]年度报告;可读性水平;文本复杂性;专业词汇密集度;复杂年报;分析师预测;过往业绩;会计信息披露

[中图分类号]F234 **[文献标志码]**A **[文章编号]**1004-4833(2021)05-0077-09

一、引言

会计信息披露是现代企业所有权与经营权相分离背景下解决信息不对称问题的主要手段。高质量的信息披露有利于增强市场效率,引导社会经济资源的正确配置。纵观全球各大会计准则制定机构在概念框架中对会计信息的基本质量要求,都离不开“可理解性”或者“可读性”。我国《企业会计准则——基本准则》第二章第十四条提出,企业提供的会计信息应当清晰明了,便于财务会计报告使用者理解和使用。2012年9月19日,我国证监会发布的《公开发行证券的公司信息披露内容与格式准则第2号——年度报告的内容与格式》也明确要求,年报文本披露应该尽量力求简洁,避免重复累赘和误导性陈述,并应对理解难度较高的专业术语添加适当解释^[1]。

可读性是会计信息质量的基本保证^[2]。只有信息被投资者阅读并理解,信息披露才能真正解决企业管理层和各方利益主体之间的信息不对称问题,从而提高决策效率,优化社会资源配置。Loughran和McDonald研究发现,可读性强的会计信息披露更有助于维护资本市场的稳定^[3]。很长一段时间以来,尽管大家公认我国公司的年度报告越来越复杂,但关于年报可读性方面的研究却并不多见,这主要受限于汉语表达的特点,在早期研究中很难对中文文本的可读性进行计量,直到自然语言处理(NLP)技术让中文分词成为可能,文本分析在财务与金融领域才逐渐成为研究热点。当前,在我国公司会计信息披露的可读性方面,一些学者探索了公司年报可读性的计量^[4-5],还有不少学者关注了年报可读性的影响因素和经济后果^[6-8]。

本文从中文的文本复杂度和会计专业词汇的密集度两个层面来定义年度报告可读性计量指标,在年报编制要求越来越专业且复杂化的背景下,选择专业解读能力较强的证券分析师作为专业投资者代表,考察上市公司年度报告的文本可读性水平对分析师解读信息从而做出盈利预测的影响。本文的研究贡献主要体现在三个方面:首先,针对上市公司年度报告这种专业性极强的文本,本文整理了专业词汇库,将近年来伴随准则更新而诞生的“新”会计术语也纳入计量范围,并对专业术语的复杂程度进行了进一步区分,计量的准确性相对当前同类文献有了进一步提升。其次,从研究角度来看,卖方证券分析师作为连接公司与普通投资者之间的信息沟通桥梁,是公司年报信息披露的专业用户,考察年报文本信息的可读性水平对分析师行为的影响对于确保信息披露的有效性和信息披露监管工作具有重要意义。最后,从研究结果的角度来看,本文深入分析了年报文本的可读性水平对分析师预测行为的影响机制,以及分析师的专业胜任能力在理解复杂年报中的作用。

[收稿日期]2021-02-01

[基金项目]国家社会科学基金重点项目(18AGL009);教育部人文社会科学项目(19YJA790114);财政部国际化高端会计人才项目

[作者简介]朱丹(1974—),女,四川成都人,重庆大学经济与工商管理学院副教授,博士,从事财务会计理论与实践研究,E-mail:zhudan@cqu.edu.cn;李静柔(1996—),女,湖南常德人,重庆大学经济与工商管理学院硕士研究生,从事财务会计理论与实践研究;李世新(1974—),男,重庆人,重庆大学经济与工商管理学院副教授,博士,从事资本市场会计研究。

二、理论分析与研究假设

(一)年度报告的可读性与模糊假说

可读性本是应用语言学上的概念,是指文本易于阅读和理解的程度^[9]。国外学者从1950年左右就开始对英文文本的可读性问题展开了研究,但由于汉语不像英文那样在词汇之间存在天然空格,因此针对汉语言的文本分析一直受到分词技术发展的限制。近二十年来,随着NLP等技术的突破,汉语文本可读性方面的研究才逐渐丰富起来,“汉语文本可读性测评和分级的跨学科研究”还成为我国2017年国家社科基金重大项目招标选题^①。从20世纪末开始,陈阿林和张素使用公式法对中文文本的可读性计量进行了探索^[10],王进等基于自然语言处理技术建立了图书阅读难度分级模型与自动分级系统^[11],吴思远等通过机器学习模型考察了语言特征对可读性的预测能力^[12]。不过,上述研究主要局限在汉语教学领域。

企业年报是资本市场会计信息披露最重要的载体,与普通汉语教学领域文本相比,年报的财会专业性特别强,并且它所面向的用户通常已具备基础汉语阅读理解能力,因此,对这类专业文本的可读性计量应当与普通文本有所区分。Jones和Shoemaker指出,年报是一种大多数个人投资者都无法读懂的专业文本^[13];Heath等率先使用Fog指数计量了英文年报的可读性水平^[14];Li同时使用Fog指数和文本长度度量了公司年报的可读性^[15]。在针对中文年报所展开的研究中,孙蔓莉等使用完形填空法考察了我国公司的年报文本,得出了可读性不佳的研究结论,她还尝试针对年报中的“管理层讨论与分析”部分通过人工逐句阅读方法计量了其可读性^[16-17]。在中文分词难题逐步得到解决后,更多关于会计信息披露可读性方面的研究涌现出来,王克敏等学者研究了管理层披露策略和盈余管理对公司年报可读性水平的影响^[18];孙文章研究发现,声誉激励能促使董秘更好地履行信息披露职责,从而促进披露文本可读性水平的提高^[19]。

总的说来,现有研究对于中文年报文本可读性的计量还在探索之中,而且会计学科特有的技术和专业性使得年报使用者的理解能力在会计术语和日常汉语之间必然存在差异。由于年报使用者已普遍具有高中语文水平以上的阅读能力,因此普通字词的难易程度对于阅读者理解能力的影响有限,更大的障碍应该存在于专业的会计术语当中。所以,本文拟从普通文本的阅读负担和专业文本的理解难度两个层面计量年度报告的可读性。

Bloomfield的不完全反映假说(Incomplete Revelation Hypothesis, IRH)认为,从公开数据中获取某类信息的成本越高,市场价格对这类信息的反映就越不完全^[20]。在此基础上,他用模糊假说(Obfuscation)来解释年报可读性与公司盈余水平之间的关系:公司管理层可能会出于私利目的去操纵信息披露的可读性,从而通过影响使用者的信息解读效率来实施对已有利的印象管理。例如,通过降低年报文本信息的可读性水平(增加复杂度)来隐藏坏消息,或者通过提高年报文本信息的可读性水平使好消息得以有效传递,从而诱导信息使用者做出更加有利于管理层目的的价值判断^[20]。许多文献都证明公司业绩水平与年报文本可读性之间存在显著的正相关关系^[15,21-22],王克敏等认为这是公司管理层蓄意操纵造成的结果^[18];许文瀚和朱朝晖研究发现,上市公司“微盈利”现象与年报可读性水平显著负相关,并且“最严退市制度”显著提高了微盈利公司年报的可读性水平^[23],这说明企业可能会通过降低文本可读性水平来掩盖其盈余管理行为。综上,国内外研究成果均证实了模糊假说,说明即使按照相同的编报规则进行信息披露,通过增加使用者信息解读的难度也可以使得某些年报文本具有“烟幕”作用:减少不利信息对公司形象的影响,或者放大有利消息的作用。

(二)可读性水平、分析师预测与过往业绩的功能锁定影响

为了检验年报文本的可读性水平是否确实会对使用者的信息解读造成显著影响,本文选择从卖方证券分析师的角度来进行考察。一般来说,分析师预测是研究会计信息披露经济后果的重要视角,因为分析师具备相对专业的财会金融知识,对行业与市场也有着充分的了解。信息披露的质量越高,分析师预测的偏差就越小,整体的预测准确性也就越高。丘心颖等研究发现,分析师会更加积极地跟进文本信息复杂的年报,但年报可读性水平与分析师预测质量之间却不存在显著的相关关系^[7];邓可斌也认为,分析师对于复杂年报的跟进未能带来预测质量和预测信息含量的提升^[24]。然而,刘会芹等的研究结果表明,分析师不会特别关注可读性水平低的公司年报,并且当年报可读性水平较低时,分析师预测的准确度会下降,分歧度也会较高^[25]。综上,分析师盈利预测

①该项目目前由北京语言大学博士生导师、《世界汉语教学》期刊编辑江新教授主持。

的结果确实会受到公司年报信息披露可读性水平的影响,但影响方向并不一致,这提示我们:年报可读性水平对分析师预测的影响很可能还存在更复杂的机制或路径。

“功能锁定”(functional fixation)概念最早由心理学家 Dunker 和 Luchins 提出,指主体对客体的认识和利用存在某种功能性障碍^[26-27]。很多学者已研究证实资本市场存在“功能锁定”现象^[28],并用“功能锁定假说”特指投资者在决策过程中被锁定于某特定指标或表面信息,因而不能充分理解和利用其他相关信息来评估证券价值,从而做出不充分和有偏差估值的现象。本文认为,在年报可读性水平对分析师预测的影响机制中,过往盈利情况作为预测的基准信息,具有首要的影响力和重要的“锚定”作用,因此根据“功能锁定假说”,过往盈利情况会就年报可读性水平对分析师预测的影响作用进行调节。

从理论上来说,当年报文本的可读性水平较差时,分析师从复杂年报中获取有效信息会变得更加困难。在其他条件相同的情况下,由于对年度报告的信息解读效率下降,分析师就会更倚重过往盈利水平等定量信息和年报之外的其他信息来源。因此,本文预期分析师会更谨慎地看待文本复杂程度更高的公司年报,但前期业绩能对年报文本可读性水平与分析师预测乐观程度之间的关系进行“功能锁定”,具体表现为:当公司以往业绩较好时,复杂度高的年报能诱使分析师做出更乐观的估计;当公司以往业绩较差时,分析师面对复杂度高的年报则可能会表现得更加谨慎,控制预测的乐观程度。因此,本文提出假设 H1:

H1:企业过往业绩会约束年报可读性水平对分析师预测乐观程度的影响。具体而言,过往业绩越好,复杂年报越能提升分析师预测的乐观程度。

盈利预测的准确程度受乐观偏差和谨慎偏差以及两者偏差程度的共同影响。一般来说,复杂年报会加大分析师的信息解读难度,从而影响盈利预测的准确性。然而,公司年报并非预测的唯一信息来源,当年报解读难度增大时,必然会促使分析师通过其他渠道进行信息挖掘,如实地调研、电话访谈等。Stickel 和胡奕明等研究发现,这类信息挖掘和搜集活动能极大地提高分析师预测的准确性^[29-30]。因此,越是复杂的年报,传递给分析师的其他渠道信息挖掘压力就越大。

不过,在相同的年报复杂水平下,受业绩“功能锁定”作用的影响,过往业绩水平越好,业绩的“定心丸”作用就越强。因此,对于分析师预测来说,针对过往业绩好的企业去进行额外信息挖掘的压力相对要比过往业绩差的企业更小。也就是说,在同样的年报可读性水平下,随着分析师进行其他信息挖掘的努力程度的松懈,盈利预测的准确度反而可能会下降,即预测偏差增大。按照此逻辑,本文提出假设 H2:

H2:企业过往业绩会约束年报可读性水平对分析师预测偏差的影响。具体而言,过往业绩越好,复杂年报越会增大分析师预测的偏差程度。

就盈利预测的分歧程度来说,分歧主要产生于不同分析师对相同信息的理解差异和分析师之间对其他渠道信息挖掘的努力程度差别。此外,私有渠道的信息挖掘往往会使得不同分析师之间占有信息的质量差距显著拉大。

从理论上来说,复杂年报意味着信息披露的可读性水平差,这会促使分析师加大对其他渠道信息的挖掘力度。不过,在同样的年报复杂水平下,由于受到“功能锁定”作用的影响,企业的过往业绩越好,相应地,额外信息挖掘的压力就会减小很多,从而缩小了分析师之间所掌握信息的差异。因此,本文提出假设 H3:

H3:企业过往业绩会约束年报可读性水平对分析师预测分歧程度的影响。具体而言,过往业绩越好,复杂年报越会减弱分析师预测的分歧程度。

三、研究设计

(一)样本选择与数据来源

本文选择我国 2007—2018 年 A 股非金融类上市公司年报作为研究样本,其中年报文本来自 Wind 数据库,报表数据和分析师预测数据均来自 CSMAR 数据库。在研究过程中,我们对以下样本进行删除:(1)ST、*ST 以及所有者权益为负的公司样本;(2)因格式问题导致年报文本信息无法提取的样本;(3)数据缺失的样本。同时,对于分析师预测数据,仅保留预测年度前一年年报报出日到预测当年年报报出日之间的样本;若同一分析师在同一期间对同一企业进行了多次预测,则仅保留离预测期上一年度年报报出日最近的预测值。经筛选,本文最终获得 146280 个公司-年-分析师层面的分析师预测数据和 9861 个公司-年层面的上市公司数据。

接下来,本文首先使用 Total PDF Converter 对 PDF 格式的年报文件进行文字提取,然后利用 Python 中的 jieba 模块加入自定义的会计词典作为分词词典进行基本的分词和词频统计。本文使用 Python3.7 进行年报篇幅、平均句长以及专业词汇密度的计算,并对研究涉及的所有连续型变量进行上下 1% 水平的 Winsorize 处理。

(二) 变量定义

1. 因变量

对于分析师预测的乐观程度 (*Optimism*), 本文采用分析师盈利预测值高于公司实际盈利的幅度来衡量。参照 Jackson 的做法^[31], 我们将分析师预测乐观偏差定义为分析师预测的 EPS 减去当年实际 EPS, 再除以预测前一个交易日该股票收盘价格后予以标准化。本文使用不同分析师对于同一家上市公司同一预测期的预测乐观偏差均值来度量公司 - 年层面的分析师预测乐观偏差:

$$Optimism_{i,t} = \frac{1}{n} \sum (F_{i,t,n} - A_{i,t}) / P_{i,t} \quad (1)$$

在式(1)中, n 表示针对公司 i 第 t 年 EPS 发布预测的分析师数量, F 为分析师 n 对公司 i 在第 t 年的每股收益预测值, A 为公司 i 第 t 年的每股收益实际值, P 为分析师做出盈利预测前一个交易日的股票价格。

参考 Duru 的做法^[32], 本文把分析师预测的准确度偏差 (*Bias*) 定义为分析师预测的 EPS 减去当年实际 EPS 的绝对值, 再除以预测前一个交易日该股票收盘价格后予以标准化。本文以不同分析师对同一家上市公司同一预测期的不同预测偏差均值来度量公司 - 年层面的分析师预测偏差, 预测偏差越小, 意味着预测的准确度越高。式(2)中各变量的定义同式(1)。

$$Bias_{i,t} = \frac{1}{n} \sum |F_{i,t,n} - A_{i,t}| / P_{i,t} \quad (2)$$

此外, 参考薛雨佳的做法^[33], 我们把分析师预测的分歧程度 (*Dispersion*) 定义为所有分析师对同一家公司在相同年份的 EPS 预测结果的标准差。

$$Dispersion_{i,t} = STD(F_{i,t,n}) \quad (3)$$

2. 自变量

如前所述, 本文从普通文本的阅读负担和专业文本的理解难度两个层面度量年报文本的可读性水平, 其中, 前者使用年报总篇幅和平均句长进行刻画, 后者使用会计专业词汇密度来度量。目前仅有少数研究考虑了年报中专业术语的理解难度^[18], 但近二十年来, 会计准则越来越精细化和复杂化, 从而使得许多新的会计科目和专业词汇得以产生, 一些“新鲜”的专业词汇甚至让许多经验丰富的从业者都感到理解困难。因此, 会计词汇密度是影响财务报告这类专业文本可读性水平的重要因素, 专业理解难度越高的年报文本, 其可读性水平越低。

然而遗憾的是, 本文经过查证后发现, 随着电子辞典的普及, 目前并没有涵盖最新专业词汇的会计词典可以直接使用。因此, 在参考《英汉 - 汉英国际会计准则词汇》(2000 年) 的基础上, 本文根据 2000 年以来的我国会计准则变化情况, 对《企业会计准则》及其应用指南和 2019 年版 CPA 考试《会计》和《财务成本管理》教材中新出现的专业词汇进行手工整理, 并反复多次与小样本分词结果进行核对, 再结合数位会计学专家学者的意见与建议进行调整, 定制了在计量会计专业词汇密度时使用的会计专业词汇词典。

进一步地, Dupree 将会计专业词汇分为描述性和技术性两大类^[34], 后者通常仅在专业文献或会计准则中出现, 对于普通信息使用者来说, 理解难度更大。有鉴于此, 本文在度量年报中会计专业词汇的密度时, 也反复参考多位会计学专家学者和证券公司分析师的反馈意见, 将词典中的专业词汇按照理解难度分为两级: 将“收入”“费用”“资产”等普通文本中也常出现、非专业人士一般也能理解的专业词汇归为描述性专业词汇, 将“递延所得税负债”“其他权益工具”等需要经过必要的会计专业知识学习才能真正理解的词汇归为技术性专业词汇。

本文使用的总会计专业词汇密度 (*ATDensity*) 是指包括描述性和技术性专业词汇在内的所有会计专业词汇数占年报总词汇数的比例; 技术性会计专业词汇密度 (*TTDensity*) 则单指技术性专业词汇数占年报总词汇数的比例; 年报总篇幅 (*Length*) 采用年报全文的中文总字符数进行度量; 平均句长 (*Sentence*) 由总的中文字符数除以总句数计算得到。一般而言, 篇幅越长或逻辑关系复杂的长句越多, 阅读者的理解负担越重, 文本的可读性水平相对越低。本文将上述四个指标按照公式(4)进行离差标准化处理, 合成年报文本复杂度计量指标 (*Complexity*), 该指标值越大, 年报文本的可读性水平就越低。

$$Complexity_{i,t} = \sum \frac{x_{i,t} - \min(x)}{\max(x) - \min(x)} \quad (4)$$

在式(4)中, x 分别取总会计专业词汇密度 ($ATDensity$)、技术性会计词汇密度 ($TTDensity$)、年报总篇幅 ($Length$) 和平均句长 ($Sentence$)。

3. 调节变量与控制变量

过往业绩水平是分析师预测的基准。由于盈利属于定量信息且具有“功能锁定”效应,在探究年报文本信息可读性水平对分析师预测的影响时不应忽视,因此本文将公司的上年每股收益水平 (EPS) 作为调节变量。

此外,本文从被预测客体公司的个体特征、预测主体分析师层面的特征以及分析师预测所可能用到的其他信息来源等三个维度对后续实证检验过程进行了控制。公司个体特征方面,本文使用总资产的自然对数 ($Size$) 和资产负债率 (Lev) 分别控制公司规模与杠杆水平的影响。由于盈利波动与公司经营风险程度相关,波动越大,分析师越有动机跟进,其预测越有优势^[35],故本文使用每股收益变化率 (ΔEPS) 控制经营风险程度。鉴于 Schipper 提出的分析师对亏损企业的预测乐观偏差程度更高这一观点^[36],本文控制了公司上年度盈亏性质 ($Loss$) 的影响。此外,由于公司的业务复杂程度直接关系到盈利预测的难度大小,因此本文将参控股公司数量作为衡量业务复杂度 ($Subc$) 的指标,并加以控制。最后,本文还对公司信息披露质量 ($Iquality$) 进行了控制。

对于分析师层面的因素,本文考虑并控制了分析师对于股票的推荐程度 ($Stdrank$)、预测区间长短 ($FHorizon$)、分析师关注的程度 ($Analyst$) 以及是否为明星分析师的重点关注对象 ($Star$) 等指标。对于公司年报之外的其他信息披露渠道,本文使用其他交流平台变量 (OCP) 控制了公司启用官网问答平台、官网微博和微信平台的数量,使用投资者互动变量 ($Question$) 控制了证交所等投资者互动平台的提问数量。

变量的具体定义如表 1 所示。

表 1 变量定义

变量类型	变量名称	变量符号	变量定义
因变量	乐观偏差	<i>Optimism</i>	第 t 年跟踪公司 i 的分析师预测的乐观偏差平均值
	预测偏差	<i>Bias</i>	第 t 年跟踪公司 i 的分析师预测的准确度偏差平均值
	预测分歧度	<i>Dispersion</i>	第 t 年跟踪公司 i 的不同分析师预测结果的标准差
自变量	文本复杂度	<i>Complexity</i>	公司 i 第 t 年年报文本的复杂(或不可读)性水平计量指标
	总会计词汇密度	<i>ATDensity</i>	公司 i 第 t 年年报中所有会计专业词汇词频数占总词频数的比例
	技术性专业词汇密度	<i>TTDensity</i>	公司 i 第 t 年年报中技术性会计专业词汇词频数占总词频数的比例
	文本篇幅	<i>Length</i>	公司 i 第 t 年年报的中文字符总数,取自然对数
	平均句长	<i>Sentence</i>	用公司 i 第 t 年年报的中文字符总数除以句子总数得到平均每句字符数,取自然对数
调节变量	盈利水平	<i>EPS</i>	公司 i 第 t 年的每股收益
控制变量	公司规模	<i>Size</i>	公司 i 第 t 年的总资产,取自然对数
	杠杆水平	<i>Lev</i>	公司 i 第 t 年的资产负债率
	盈利变动	ΔEPS	公司 i 第 t 年的每股收益与第 $t-1$ 年每股收益之差
	业务复杂度	<i>Subc</i>	公司 i 第 t 年参控股公司数加 1,取自然对数
	其他交流平台	<i>OCP</i>	公司 i 第 t 年其他交流平台的数量
	投资者互动	<i>Question</i>	公司 i 第 t 年收到的投资者网络平台提问数加 1,取自然对数
	信息披露质量	<i>Iquality</i>	公司 i 第 t 年被交易所出具的信息披露工作评价等级
	盈亏性质	<i>Loss</i>	虚拟变量,公司 i 第 t 年亏损时取 1,否则取 0
	推荐程度	<i>Stdrank</i>	第 t 年跟踪公司 i 的分析师所报告的 <i>GTA</i> 标准化评级之均值
	预测区间	<i>FHorizon</i>	第 t 年跟踪公司 i 的分析师预测时间和年报公布日距离天数之差的平均值
	分析师跟踪	<i>Analyst</i>	第 t 年跟踪公司 i 的分析师数量加 1,取自然对数
	明星分析师	<i>Star</i>	虚拟变量,公司 i 在第 t 年被明星分析师跟踪程度高于平均值时取 1,否则取 0
	年度	<i>Year</i>	年度虚拟变量
	行业	<i>Industry</i>	行业虚拟变量

(三) 模型构建

为了探究年报文本可读性水平对分析师盈利预测的影响,针对三个研究假设,本文建立检验模型如下:

$$Forecast_{i,t} = \beta_0 + \beta_1 (Complexity_{i,t-1} \times EPS_{i,t-1}) + \beta_2 Complexity_{i,t-1} + \beta_3 EPS_{i,t-1} + \sum ControlVariables_{i,t-1} + \sum ControlVariables_{i,t} + \sum Year + \sum Industry + \varepsilon_{i,t}$$

其中,因变量 $Forecast$ 在回归检验中分别由预测乐观程度 $Optimism$ 、预测偏差程度 $Bias$ 和预测分歧程度 $Dispersion$ 替代;滞后一期的控制变量包括分析师层面控制变量和其他沟通渠道控制变量。

四、实证结果及分析

(一)描述性统计

表2报告了核心变量的描述性统计结果。分析师预测乐观偏差的均值大于零,且乐观偏差和预测偏差的均值分别为其中位数的近两倍,表明研究样本中的部分分析师在盈余预测方面过于乐观,偏差程度较大;在年报文本的可读性水平计量方面,总篇幅和平均句长的分布比较均匀;特别值得一提的是,上市公司年度报告中出现会计专业词汇的频次很高,总专业词汇占比高达35%,其中理解难度更大的技术性专业词汇密度约为12%,可见年度报告对于普通投资者的理解难度是客观存在的;文本复杂度的均值为2.289,分布均匀;企业盈利水平EPS指标的取值区间为(-0.74385,2.43508),说明样本公司的盈利水平差异较大,在这种情况下研究不同的过往业绩水平对分析师盈利预测的“功能锁定”影响是很有意义的。

(二)回归分析

在回归分析前,本文对主要变量进行了相关性检验,结果显示变量之间不存在高度相关,且VIF值域区间为(1.02,3.36),说明变量之间不存在严重的多重共线性问题。

本文采用OLS方法进行回归分析,结果如表3所示。在表3中,主检验的列(1)结果显示,企业过往盈利水平和年报文本的复杂度均对分析师盈利预测的乐观程度具有显著的负向影响,这说明:一方面,在相同的过往盈利水平和其他条件控制下,年报文本的复杂程度越高,分析师预测的乐观程度被抑制得越厉害,或者说对于复杂程度更高的年度报告,分析师具有本能的防御心理,会更谨慎地发表其盈利预测。另一方面,在其他条件相同时,企业的过往盈利越好,分析师的预测越不容易盲目乐观,即“好”是有一定的天花板限制的;而对于过往盈利越差的公司,分析师普遍更容易产生美好预期,做出偏乐观的估计。

从文本复杂度与盈利水平的交乘项在1%水平上显著为正可以看出,过往盈利水平对于分析师预测的乐观程度具有功能锁定影响,年报文本复杂度对于分析师预测乐观程度的边际影响为 $-0.00470 + 0.00789 \times EPS$,这意味随着公司上年EPS值的增大,年报文本复杂度对分析师预测乐观程度的负向影响越不明显,甚至当EPS明显大于某个临界值(以本文研究样本为例,即0.6左右)时,这种负向影响反而能转为正向。也就是说,在控制了其他变量的影响之后,如果公司过去的业绩特别好,那么复杂年报能让分析师做出更乐观的未来盈利预测,或者说复杂年报具有放大好消息的影响力;而如果公司的过去业绩很糟糕,那么当分析师面对同样水平的复杂年报时,其防御心理会更明显,因此盈利预测也会更谨慎。研究结论支持了假设H1,说明在面对可读性差

表2 核心变量的描述性统计结果

变量	均值	标准差	最小值	最大值	中位数
乐观偏差	0.00898	0.01950	-0.03124	0.12030	0.00461
预测偏差	0.01304	0.01823	0.00020	0.12180	0.00716
预测分歧度	0.09900	0.12800	0.00000	0.71400	0.07011
文本复杂度	2.28900	0.14410	1.92202	2.62440	2.29120
总专业词汇	0.35512	0.02894	0.27887	0.42077	0.35580
技术性专业词汇	0.12457	0.01614	0.08636	0.16435	0.12399
总篇幅	11.25564	0.24950	10.66303	11.88402	11.25214
平均句长	4.63912	0.12878	4.32694	4.97705	4.63762
盈利水平	0.46119	0.48558	-0.74385	2.43508	0.36099

表3 主检验回归结果

变量	(1)	(2)	(3)
	Optimism	Bias	Dispersion
Complexity × EPS	0.00789 *** (2.92439)	0.00595 ** (2.46871)	-0.13407 *** (-5.32117)
Complexity	-0.00470 ** (-2.40758)	-0.00163 (-0.78650)	0.05433 *** (3.87605)
EPS	-0.02586 *** (-4.10986)	-0.01904 *** (-3.36969)	0.40375 *** (6.81605)
Size	0.00175 *** (7.56838)	0.00267 *** (11.20121)	-0.00333 ** (-2.08323)
Lev	0.00560 *** (5.52090)	0.00913 *** (8.25109)	0.07302 *** (9.18648)
Stdrank	0.01556 *** (6.50829)	0.00690 *** (2.58560)	0.04532 *** (3.33772)
Fhorizon	0.00689 *** (9.33548)	0.00667 *** (8.93111)	0.02982 *** (6.34517)
Analyst	-0.00029 (-1.29089)	-0.00103 *** (-4.58887)	0.04070 *** (23.58076)
OCP	-0.00017 (-1.02086)	-0.00023 (-1.27200)	-0.00415 *** (-3.01903)
Question	0.00061 *** (4.27660)	0.00075 *** (5.05066)	0.00620 *** (5.75814)
Iquality	-0.00248 *** (-8.96395)	-0.00283 *** (-9.68487)	-0.00929 *** (-4.77214)
ΔEPS	-0.03314 *** (-36.27125)	-0.02122 *** (-19.63396)	-0.06647 *** (-10.44748)
Loss	0.01013 *** (8.94832)	0.01107 *** (9.56808)	0.10021 *** (14.15816)
Subc	0.00001 (0.07174)	-0.00038 * (-1.81688)	-0.01016 *** (-6.33680)
Star	0.00062 ** (2.07431)	0.00066 ** (2.15534)	0.00411 * (1.69004)
Industry		控制	控制
Year		控制	控制
Constant	-0.07263 *** (-8.55731)	-0.07667 *** (-8.81150)	-0.24430 *** (-4.49290)
Observations	9861	9861	9861
R ²	0.445	0.318	0.329
F	42.42 ***	32.57 ***	72.44 ***

注:括号内为t值;***表示 $p < 0.01$, **表示 $p < 0.05$, *表示 $p < 0.1$ 。下同。

的年度报告时,分析师受到了公司历史业绩“功能锁定”的影响,容易得出好公司更好、差公司谨慎说好的预测结论。

表3中列(2)的检验结果显示,历史盈利水平 *EPS* 与分析师预测的偏差程度显著负相关,即公司过去的业绩越好,分析师盈利预测偏差就越小,或者说预测准确度越高。但是,年报文本的可读性水平本身对分析师预测的偏差并没有显著影响,这可能是因为复杂度高的年报一方面会导致预测难度增大,另一方面也会激励分析师通过其他渠道进行信息挖掘,努力纠正复杂年报对预测偏差的不利影响。不过,由于 *Complexity* 与 *EPS* 的交乘项对分析师预测偏差的影响在1%水平上显著为正,说明在同样的历史业绩水平下,可读性差的年度报告会加大分析师预测的偏差,预测难度加大,结果更不准确,这正是复杂年报的“模糊”作用。因此,假设 H2 并没有得到完全支持,实证研究结果表明,复杂年报对分析师预测的偏差并没有显著的直接影响,但它能减缓较差的过往业绩对预测偏差的不利影响,为“模糊假说”提供了证据。

表3中列(3)的检验结果显示,年报文本复杂度和公司历史业绩水平与分析师预测分歧程度均呈显著的正相关关系,且交乘项系数在1%水平上显著为负,这意味着在其他条件相同的情况下,年报文本复杂度对分析师预测分歧程度的边际影响为 $0.05433 - 0.13407 \times EPS$,即公司过去的业绩越糟糕,年报复杂度与分析师预测分歧程度的正相关性越强,复杂年报让分析师们达成一致预测的难度就越大。这是因为在年报复杂程度一定的情况下,分析师对过往业绩更好的公司信任程度更高,而对过往业绩不佳的公司信任程度更低,会更有动机通过各种私有渠道去进一步挖掘信息,从而加大了预测结果的分歧程度。

(三)进一步检验

总体而言,对于复杂或可读性差的年度报告,分析师的专业胜任能力将显得更加重要。在客观信息披露不够充分或可读性较差时,分析师的主观职业判断或理解能力对盈利预测结果的影响更大,能否被复杂披露“模糊”印象或被过往盈利水平“功能锁定”均取决于分析师的专业胜任能力。因此,为了进一步检验年报可读性水平对分析师盈利预测的作用机制,本文选择明星分析师作为专业能力更强的分析师代表,将样本分为明星分析师参与度高于平均水平 ($Star = 1$) 和低于平均水平 ($Star = 0$) 两组分别进行回归,并对分组差异进行费舍尔检验,检验结果如表4和表5所示。

表4 分组回归结果

变量	(1)		(2)		(3)	
	<i>Optimism</i>		<i>Bias</i>		<i>Dispersion</i>	
	Star = 1	Star = 0	Star = 1	Star = 0	Star = 1	Star = 0
<i>Complexity</i> × <i>EPS</i>	0.00195 (0.58058)	0.01365 *** (3.34178)	-0.00005 (-0.01562)	0.01126 *** (2.97092)	-0.10851 *** (-2.96400)	-0.16207 *** (-4.63500)
<i>Complexity</i>	-0.00054 (-0.20257)	-0.00776 *** (-2.81351)	0.00412 (1.56048)	-0.00557 * (-1.82825)	0.06000 *** (2.69565)	0.05365 *** (2.96975)
<i>EPS</i>	-0.01110 (-1.40919)	-0.03981 *** (-4.17365)	-0.00467 (-0.68545)	-0.03146 *** (-3.53350)	0.33520 *** (3.92402)	0.47710 *** (5.77173)
<i>Controls</i>	yes	yes	yes	yes	yes	yes
<i>industry</i>	yes	yes	yes	yes	yes	yes
<i>year</i>	yes	yes	yes	yes	yes	yes
<i>Constant</i>	-0.07563 *** (-6.14571)	-0.06618 *** (-5.69318)	-0.09350 *** (-7.83812)	-0.05917 *** (-4.82459)	-0.37100 *** (-4.02256)	-0.15315 ** (-2.25407)
Observations	4423	5438	4423	5438	4423	5438
R ²	0.377	0.487	0.263	0.361	0.310	0.349
F	15.77 ***	34.16 ***	14.15 ***	27.79 ***	31.53 ***	44.38 ***

表5 组间差异检验结果

变量	(1) <i>Optimism</i>			(2) <i>Bias</i>			(3) <i>Dispersion</i>		
	$b_0 - b_1$	Freq	P 值	$b_0 - b_1$	Freq	P 值	$b_0 - b_1$	Freq	P 值
<i>Complexity</i> × <i>EPS</i>	0.014	1	0.020	0.012	0	0.000	-0.013	23	0.460
<i>Complexity</i>	-0.006	44	0.120	-0.009	50	0.000	-0.017	36	0.280
<i>EPS</i>	-0.034	49	0.020	-0.026	50	0.000	0.044	23	0.460
<i>Controls</i>		yes			yes			yes	
<i>_Cons</i>	0.032	1	0.020	0.050	0	0.000	0.017	15	0.300

针对模型(1)和模型(2)的进一步检验结果显示,在明星分析师参与水平高的样本中,分析师预测的乐观程度和预测偏差均不受年报文本复杂水平以及公司历史业绩的显著影响,这说明相较于非明星分析师,明星分析师在进行盈利预测时确实具有更高的业务水平和能够保持更加客观的态度,分析师专业能力的提升可以很好地削弱年报文本可读性水平的的影响以及过往业绩水平的“功能锁定”效应。

针对模型(3)的进一步检验结果显示,尽管年报文本的复杂水平和公司历史业绩水平对分析师预测分歧度均具有显著影响,但不同明星分析师参与水平的组间差异并不显著,这是因为明星分析师更高的专业水平主要影响自身判断结果,而对整体的盈利预测分歧度并没有实质性影响。

(四) 稳健性检验

为进一步检验结论的稳健性,本文对解释变量和被解释变量进行替换:(1)借鉴王玉涛等衡量分析师盈余预测分歧度的方法^[37],本文将预测分歧度指标定义为各分析师预测的标准差除以当年实际盈余,即 $Dispersion_{i,t} = STD(F_{i,t,n})/EPS_{i,t}$; (2)使用年报总会计术语密度(*ATDensity*)作为年报复杂度的替代指标。检验结果如表6所示,显示本文研究结论是稳健的。

五、结论性评述

本文从普通文本的阅读难度和专业文本的理解难度两个层面计量了公司年报的可读性水平,然后以分析师作为专业的信息使用者代表,检验了年报文本的可读性水平对分析师盈利预测的影响。研究发现:第一,分析师对于复杂的年度报告进行了本能的防御,复杂年报降低了分析师预测的乐观度,但年报的复杂化意味着上市公司公开披露的信息质量的降低,促使分析师对其他渠道来源信息进行挖掘,进而增大了分析师预测的分歧度。第二,作为盈利预测的基准,公司的过往业绩水平具有明显的“功能锁定”效应,会对年报文本可读性与分析师预测之间的关系进行调节。当公司过往业绩水平较高时,复杂年度报告的“模糊”作用明显,容易诱使分析师放松警惕并做出更乐观且不准确的盈利预测,随着过往业绩水平的下降,这种影响也随之减弱。在公司年报文本的可读性差且过往业绩也不佳的情况下,分析师会表现得更加谨慎,且具有通过其他私有渠道进行信息挖掘的更强动力,这反而能缩小盈利预测的准确度偏差。第三,分析师的专业胜任能力在年报可读性对分析师盈利预测的影响中发挥了重要作用,明星分析师更能看透复杂年报对信息解读的不利影响,并摆脱过往业绩的“功能锁定”束缚,在盈利预测中保持更理性和更客观的工作态度。

当前,全球会计准则建设和信息披露监管日趋严格,公司年度报告的篇幅越来越长,新兴的专业词汇出现得越来越多,公司年报的可读性水平整体上呈现下降趋势。作为全球经济整体的一部分,我国的会计准则已与国际接轨并持续趋同,信息披露规则的复杂化将成为未来发展方向。本文的研究结果表明,年报可读性的降低增加了信息使用者的阅读负担和信息解读成本,影响了信息传递的效率,那么如何在保障信息披露效率提升的同时,进一步提高信息使用者的专业解读能力呢?本文提出以下建议:第一,进一步增强对上市公司信息披露的监管力度,尤其是要逐步加强对文本信息披露的监管,在规范“应披露”内容和格式的同时,逐步引导上市公司重视信息披露的可读性要求,保障披露信息的质量。第二,信息不对称的存在使得中小投资者的劣势地位很难扭转,从某种程度上看,分析师在资本市场中发挥着“信息解说员”作用。因此,可通过严格分析师从业门槛、完善激励机制等手段,培养拥有足够专业胜任能力的证券分析师,促进其业务能力和水平的持续提高,从而更好地发挥分析师群体的信息解读与传递作用。第三,重视普通投资者的财经知识学习问题。当今社会,商业模式不断创新,经济交易越来越多样化和复杂化,信息披露的专业性越来越强,因此对投资者进行适当的专业教育、提高信息使用者的专业素质,从而帮助他们能够更好地理解交易环境和解读专业信息是非常重要的。

参考文献:

- [1] 中国证券监督管理委员会. 公开发行证券的公司信息披露内容与格式准则第2号——年度报告的内容与格式(2012年修订)[S]. 中华人民共和国国务院公报,2012.
- [2] Courtis J K. Readability of annual reports: Western versus Asian evidence[J]. Accounting, Auditing & Accountability Journal, 1995, 8(2): 4-17.
- [3] Loughran T, McDonald B. Measuring readability in financial disclosures[J]. The Journal of Finance, 2014, 69(4): 1643-1671.
- [4] 葛伟琪. 我国上市公司中英文年报可读性对比研究[J]. 宁波大学学报(人文科学版), 2007(1): 108-113.
- [5] 张星星. 金融危机下上市公司年度报告可读性特征研究[J]. 湖北工业大学学报, 2010(3): 77-79+82.
- [6] 任宏达, 王琨. 社会关系与企业信息披露质量——基于中国上市公司年报的文本分析[J]. 南开管理评论, 2018(5): 128-138.

表6 稳健性检验结果

变量	(1)	(2)	(3)
	<i>Optimism</i>	<i>Bias</i>	<i>Dispersion</i>
<i>ATDensity</i> × <i>EPS</i>	0.03140 ** (2.36965)	0.02376 ** (1.97946)	-3.45310 ** (-2.11847)
<i>ATDensity</i>	-0.00954 (-1.12118)	0.00472 (0.53241)	3.41847 ** (2.18194)
<i>EPS</i>	-0.01867 *** (-3.89928)	-0.01384 *** (-3.15762)	-0.41152 (-0.70496)
<i>Controls</i>	yes	yes	yes
<i>Industry</i>	yes	yes	yes
<i>Year</i>	yes	yes	yes
<i>Constant</i>	-0.07652 *** (-10.27782)	-0.08198 *** (-10.75606)	1.56224 (1.36507)
Observations	9861	9861	9861
R ²	0.440	0.315	0.178
F	44.66 ***	34.59 ***	65.53 ***

- [7]丘心颖,郑小翠,邓可斌. 分析师能有效发挥专业解读信息的作用吗?——基于汉字年报复杂性指标的研究[J]. 经济学(季刊),2016(4):1483-1506.
- [8]崔文娟,郝佳赫,盖亚洁. 财务报表可读性对股票收益率的影响[J]. 会计之友,2019(14):52-57.
- [9]李绍山. 易读性研究概述[J]. 解放军外国语学院学报,2000(7):1-5.
- [10]陈阿林,张素. 中文阅读难度模型及易读性公式探索[J]. 计算机科学,1999(11):42-44+27.
- [11]王进,周慧,罗国峰,等. 基于自然语言处理的图书阅读难度自动分级研究[J]. 计算机时代,2017(8):1-5.
- [12]吴思远,于东,江新. 汉语文本可读性特征体系构建和效度验证[J]. 世界汉语教学,2020(1):81-97.
- [13]Jones M J, Shoemaker P A. Accounting narratives: A review of empirical studies of content and readability[J]. Journal of Accounting Literature, 1994, 13(1):142-.
- [14]Heath R L, Phelps G. Annual reports II: Readability of reports vs. business press[J]. Public Relations Review, 1984, 10(2):56-62.
- [15]Li F. Annual report readability, current earnings and earnings persistence[J]. Journal of Accounting and Economics, 2008(45):221-247.
- [16]孙蔓莉. 上市公司年报的可理解性研究[J]. 会计研究, 2004(12):23-28+97.
- [17]孙蔓莉,蒋璐,孙健. 业绩归因的自利性披露及市场反应研究——汇率单边升值情境下的纺织业表现[J]. 会计研究, 2013(4):46-51+95.
- [18]王克敏,王华杰,李栋栋,等. 年报文本信息复杂性与管理者自利——来自中国上市公司的证据[J]. 管理世界, 2018(12):120-132+194.
- [19]孙文章. 董事会秘书声誉与信息披露可读性——基于沪深A股公司年报文本挖掘的证据[J]. 经济管理, 2019(7):136-153.
- [20]Bloomfield R J. The 'incomplete revelation hypothesis' and financial reporting[J]. Accounting Horizons, 2002(16):233-243.
- [21]张思捷. 企业年报可读性与盈余持续性关联性之研究[D]. 台湾:国立台湾大学, 2010.
- [22]陈银娥,江媛. 管理层权力、制度环境与董事会报告可读性——来自我国上市公司的经验证据[J]. 珞珈管理评论, 2017(3):84-99.
- [23]许文瀚,朱朝晖. 上市公司“微盈利”现象、盈余管理与年报可理解性[J]. 首都经济贸易大学学报, 2019(2):93-103.
- [24]邓可斌. 过度自信的分析师——分析师跟进复杂性年报动机研究[J]. 清华金融评论, 2018(6):105-106.
- [25]刘会芹,施先旺. 年报可读性对分析师盈余预测的影响[J]. 证券市场导报, 2020(3):30-39.
- [26]Dunker K. On problem-solving[J]. Psychological Monographs, 1945, 58(5):1-11.
- [27]Luchins A. Mechanisation in problem solving[J]. Psychological Monographs, 1942, 6:1-95.
- [28]赵宇龙,王志台. 我国证券市场“功能锁定”现象的实证研究[J]. 经济研究, 1999(9):56-63.
- [29]Stickel S. Reputation and performance among security analysts[J]. The Journal of Finance, 1992, 5(47):1811-1836.
- [30]胡奕明,林文雄. 信息关注深度、分析能力与分析质量——对我国证券分析师的调查分析[J]. 金融研究, 2005(2):46-58.
- [31]Jackson A R. Trade generation, reputation and sell-side analysts[J]. The Journal of Finance, 2005, 60(2):673-717.
- [32]Duru A, Reeb D M. International diversification and analysts' forecast accuracy and bias[J]. The Accounting Review, 2002, 77(2):415-433.
- [33]薛雨佳. 董事海外背景与分析师盈余预测——基于企业绩效与信息质量的中介效应[J]. 北京工商大学学报(社会科学版), 2020(5):83-94.
- [34]Dupree J M. Users preferences for descriptive v. technical accounting terms[J]. Accounting and Business Research, 1985(Autumn):281-289.
- [35]岳衡,林小驰. 证券分析师 VS 统计模型:证券分析师盈余预测的相对准确性及其决定因素[J]. 会计研究, 2008(8):40-49+95.
- [36]Schipper K. Analysts' forecasts[J]. Accounting Horizons, 1991(4):105-121.
- [37]王玉涛,王彦超. 业绩预告信息对分析师预测行为有影响吗[J]. 金融研究, 2012(6):193-206.

[责任编辑:王丽爱]

The Readability of Annual Report, Past performance and Earnings Forecast by Analysts

ZHU Dan, LI Jingrou, LI Shixin

(School of Economics & Administration, Chongqing University, Chongqing 400044, China)

Abstract: We measure annual reports' readability by the standard text's complexity and the intensity of professional concepts. Based on the annual report data from 2007 to 2018, we found that analysts are more cautious with complex reports, and complex reports always increase the divergence of earnings forecasts. However, as a benchmark for earnings forecasts, past performance has a functional fixation effect on the relationship between the annual reports' readability and analysts' forecasts. Further research proved that professional skills play an essential role in interpreting complex reports and avoiding past performance's functional fixation effects.

Key Words: annual reports; readability level; text complexity; the intensity of professional terms; complex annual reports; analyst's forecast; past performance; accounting information disclosure