

DOI: 10.16538/j.cnki.fem.20221128.201

前瞻性信息披露与分析师盈余预测

——基于文本分析和机器学习的证据

刘瑶瑶¹, 路军伟²

(1. 中国石油大学(华东)经济管理学院, 山东 青岛 266580; 2. 山东大学管理学院, 山东 济南 250100)

摘要: 前瞻性信息是管理层对公司未来发展前景进行展望的预测性信息。本文借助文本分析和机器学习量化年报的前瞻性信息, 研究其对分析师盈余预测行为的影响。研究发现, 年报前瞻性信息披露频率越高, 分析师盈余预测误差和分歧度越低, 即前瞻性信息披露能够提高分析师盈余预测准确性。分组检验显示, 在盈余质量和文本可读性较高公司, 前瞻性信息披露对分析师盈余预测准确性的积极效应更加显著。机制检验表明, 前瞻性信息披露主要通过缓解信息不对称、促进分析师实地调研, 来提高分析师盈余预测准确性。进一步研究发现, 前瞻性信息能够预测公司未来三年的基本面变化, 如盈余增长、创新产出和投资规模。研究结果表明, 前瞻性信息披露是有信息含量的, 能够为分析师等信息使用者提供决策有用的增量信息, 提高资本市场信息效率。研究结论在一定程度上肯定了年报前瞻性信息披露的积极效应, 为监管机构和上市公司改进信息披露提供了经验证据。

关键词: 前瞻性信息披露; 分析师盈余预测; 信息不对称; 文本分析; 机器学习

中图分类号: F270 **文献标识码:** A **文章编号:** 1001-4950(2023)10-0101-15

一、引言

信息披露是资本市场的核心问题, 文本信息是信息披露的重要形式(曾庆生等, 2018)。按照时态, 文本信息可分为过去时态、现在时态和将来时态信息。有效市场理论认为, 有效资本市场的股价至少已经包含所有历史信息, 这也意味着过去时态信息在投资者决策中的作用较为有限, 最有价值的当属现在时态和将来时态信息(王秀丽等, 2020)^①。将来时态信息即前瞻性信息, 预测了公司未来的发展战略、经营计划、收益风险、机遇挑战等, 对投资者估计未来现金流和风险具有重要的参考价值(马黎珺等, 2019; 潘越等, 2021; 田高良等, 2022)。因此, 我国证

收稿日期: 2021-10-19

基金项目: 国家自然科学基金项目(72072102); 山东大学人文社科青年团队项目(IFYT2001)

作者简介: 刘瑶瑶(1993—), 女, 中国石油大学(华东)经济管理学院讲师;

路军伟(1978—), 男, 山东大学管理学院教授, 博士生导师(通讯作者, sdljw@sdu.edu.cn)。

^①美国有线电视新闻网创始人特德·特纳曾经说道, “过去的信息是历史, 现在的信息是新闻, 未来的信息是情报”, 可见将来时态信息的重要性。

监会自2007年起多次修订信息披露准则,要求上市公司在年报对其未来发展前景进行展望,内容涉及行业格局趋势、发展战略、经营计划、面临风险等方面。在此背景下,前瞻性信息披露是否具有信息含量也逐渐引发学术界关注。有学者认为,前瞻性信息能够预测未来业绩(Schleicher等,2007;Li,2010),影响投资者决策(潘越等,2021;田高良等,2022),改善资本市场信息环境(Muslu等,2015),且与盈余质量形成正向补充关系(汪炜和袁东任,2014)。总体来看,这些研究较为一致地肯定了前瞻性信息披露的有效性,认为前瞻性信息能够为投资者等信息使用者提供增量信息。然而,也有学者认为,文本信息往往存在不同程度的“语言膨胀”,前瞻性信息可能也难逃例外(Feng等,2015),其策略性陈述将会加剧信息不对称(程新生等,2013)。可见,前瞻性信息披露的信息含量仍是现阶段一个亟待研究的开放性问题。

前瞻性信息以前瞻性和文本性为典型特征(王秀丽等,2020),相比历史信息和定量信息更为隐晦和复杂,解读该类信息也就需要更强的专业能力。Lehavy等(2011)、丘心颖等(2016)认为,分析师对于投资者解读前瞻性信息有着十分重要的作用。这给予本文一定启示,即从分析师角度研究前瞻性信息披露的信息含量,或许能为现有研究补充新的证据。分析师是专业的信息使用者,也是资本市场重要的信息中介,凭借着信息挖掘和信息解读优势,为资本市场提供盈余预测、投资建议等有价值的信息(方军雄,2007)。那么,分析师究竟能否有效挖掘并解读上市公司的前瞻性信息呢?目前,尽管已有国外学者对该问题进行初步探究(Bozzolan等,2009;Muslu等,2015),但研究结论是否适用于我国资本市场情境尚未可知。因此,本文对该问题进行研究,不仅能增强理论界对前瞻性信息披露信息含量的认知,还能为相关研究提供来自我国的经验证据。

本文以2007—2020年A股上市公司为研究对象,借助文本分析和机器学习量化年报的前瞻性信息,研究并检验其对分析师盈余预测行为的影响。研究发现,年报前瞻性信息披露频率越高,分析师盈余预测准确性也越高;盈余质量与文本可读性能够显著增强前瞻性信息披露对分析师盈余预测准确性的积极效应。机制检验发现,信息不对称的缓解和分析师实地调研的增加可能是前瞻性信息披露发挥积极效应的直接和间接路径。进一步地,前瞻性信息还能够预测上市公司未来三年的基本面变化,如盈余增长、创新产出及投资规模。研究表明,年报前瞻性信息披露是有信息含量的,分析师能够有效挖掘和解读前瞻性信息,提高盈余预测质量。

本研究可能的贡献包括:(1)基于分析师盈余预测视角,拓展了前瞻性信息披露经济后果的研究。目前,学术界就前瞻性信息披露是否具有信息含量尚未达成一致意见(Li,2010;程新生等,2013;汪炜和袁东任,2014;Muslu等,2015;Feng等,2015)。本文研究发现,年报前瞻性信息披露能够提高分析师盈余预测准确性,在一定程度上证明年报前瞻性信息披露具有信息含量,能够为分析师提供决策有用的增量信息。因此,本研究结论能够促进理论界对前瞻性信息披露信息含量的认知。(2)从前瞻性信息披露角度,丰富了分析师盈余预测影响因素的研究文献。近年来,相关研究从单纯关注定量信息对分析师盈余预测的影响,逐渐演变为同时关注定量信息和文本信息(林乐和谢德仁,2017;王雄元等,2017;李岩琼和姚颀,2020)。来自国外的研究表明,前瞻性信息作为面向未来的文本信息,也会影响分析师的盈余预测行为(Bozzolan等,2009;Muslu等,2015),但是在我国资本市场情境下,上市公司的前瞻性信息究竟是否也能被分析师有效挖掘和解读,尚未得到学术界的证实。因此,本文基于我国资本市场信息披露实践,研究前瞻性信息披露对分析师盈余预测的影响,能够为现有研究提供来自我国的经验证据。(3)从实践价值来看,本文对上市公司、监管部门、分析师具有一定的启示和借鉴。上市公司应规范前瞻性信息披露,以充分满足投资者、分析师等信息使用者的需求。监管部门应深化资本

市场信息披露改革,以确保前瞻性信息披露的可靠性、可读性、完整性和无偏性。分析师在盈余预测时应关注年报前瞻性信息,并实施调研活动来验证年报前瞻性信息。

二、文献综述

(一)文本信息披露与分析师盈余预测

随着计算机自然语言处理能力的发展和计算能力的提升,基于文本分析的实证研究逐渐成为会计研究的新方向。与标准化的定量信息相比,非标准化的文本信息篇幅更长、表达形式更丰富,通常具有定量信息难以体现的丰富内涵(曾庆生等,2018)。目前国内外学者相继从信息特征、信息内容等角度,对不同类型文本信息披露及其对分析师盈余预测的影响进行了研究。

一方面,从信息特征来看,由积极和消极两种感觉构成的语调,是文本信息的基本特征之一(曾庆生等,2018)。有学者认为,管理层语调能够传递未来业绩的增量信息,增加分析师荐股行为(林乐和谢德仁,2017),并提高其盈余预测准确性(钟凯等,2020)。也有学者发现,管理层语调可能包含刻意误导信息使用者的有偏信息,以配合内幕交易(曾庆生等,2018)、盈余管理(王华杰和王克敏,2018),这种行为无疑将误导分析师决策,降低其盈余预测准确性。除语调以外,可读性也是文本信息的基本特征之一,可能也会被管理者任意操控,以掩盖盈余管理(Lo等,2017;王克敏等,2018),缓解控制权转移风险(逯东等,2020)等。因此,较低的文本可读性也会损害分析师的盈余预测质量(刘会芹和施先旺,2020)。另一方面,从信息内容来看,相关文献聚焦特定类型文本信息对分析师盈余预测的影响。相关研究发现,或有事项信息披露越详细,明星分析师盈余预测准确性就越高(董小红等,2015)。此外,风险信息(王雄元等,2017)、研发信息(李岩琼和姚颐,2020)、核心竞争力信息(何雨晴,2021)等文本信息的披露频率,也与分析师盈余预测准确性显著正相关。这些研究给予我们一定启示,即整体而言,分析师对文本信息具有较好的处理和解读能力,但不可否认的是,文本信息的策略性陈述也会削弱其盈余预测质量。

(二)前瞻性信息披露的信息含量

若以时态为分类标准,文本信息可划分为过去时态、现在时态和将来时态信息。根据《公开发行证券的公司信息披露内容与格式准则第2号》,上市公司需主要在年报的管理层讨论与分析(MD&A)部分讨论和分析其未来发展战略、经营计划以及可能面对的风险等,这些内容都是涉及公司未来收入、成本、费用变化的前瞻性信息,势必会对投资者决策产生影响(马黎珺等,2019;潘越等,2021;田高良等,2022)。因此,前瞻性信息披露是否具有信息含量便逐渐引起学界关注。有学者发现,前瞻性信息披露能够预测公司未来业绩(Schleicher等,2007;Li,2010),改善资本市场信息环境(Muslu等,2015),且盈余质量高的公司倾向于披露更多前瞻性信息(汪炜和袁东任,2014)。但也有学者持相反观点,他们认为盈余操控公司会故意披露业绩指引等前瞻性信息来误导投资者决策,从而避免违规行为被发现(Feng等,2015)。程新生等(2013)也认为,现有制度难以对前瞻性信息披露进行约束,所以前瞻性信息披露并未有效缓解信息不对称。总体来看,前瞻性信息披露究竟是否具有信息含量,进而为投资者等信息使用者提供增量信息,仍是一个亟待研究的开放性问题。

分析师作为资本市场最重要的信息中介之一,扮演着信息使用者和信息提供者的双重角色(Schipper,1991),凭借信息优势对公司特质信息进行挖掘和解读,并将盈余预测和投资建议以研究报告形式提供给投资者(方军雄,2007)。其中,盈余预测是指分析师对所跟踪公司的未来财务状况进行预测,包括对未来会计期间利润总额、净利润、每股收益等关键财务指标进

行估计和测算。由此可见,前瞻性信息在分析师盈余预测过程中发挥至关重要的作用。国外有学者研究发现,前瞻性信息披露能够提高分析师盈余预测准确性(Bozzolan等,2009;Muslu等,2015),表明前瞻性信息披露能够为分析师决策提供增量信息。但是截至目前,国内学术界尚未对该问题给出具体答案。因此,在我国资本市场情境下分析师能否有效挖掘和解读前瞻性信息,便是值得研究的话题。

(三)文献述评

通过文献我们发现:(1)无论从信息特征,还是信息内容来看,分析师都对文本信息具有较好的解读能力,但文本信息的策略性陈述却会削弱分析师的盈余预测质量。(2)目前学术界对前瞻性信息披露的信息含量尚未达成共识,鉴于前瞻性信息相较于历史信息和定量信息的隐晦性和复杂性,从分析师这一专业信息使用者视角研究前瞻性信息披露的信息含量,或许能为现有研究补充新的证据。(3)尽管已有国外学者就前瞻性信息披露对分析师盈余预测的影响进行初步探究,但在我国资本市场实践中,分析师能否有效捕捉并解读前瞻性信息还未被学术界发掘。对此,本文基于我国资本市场情境,研究上市公司前瞻性信息披露对分析师盈余预测的影响,能够弥补前述研究不足,并为相关研究提供来自我国的经验证据。若前瞻性信息披露能够提高分析师预测准确性,则说明我国资本市场的前瞻性信息披露是有信息含量的;反之,倘若连分析师这类专业信息使用者尚且无法有效解读前瞻性信息,说明前瞻性信息披露可能并不能给我国资本市场信息使用者提供有效增量信息。研究结论对于监管部门有针对性地完善资本市场信息披露,以提高资本市场效率具有重要的理论意义和现实意义。

三、理论推导与研究假设

投资者对公司特质信息具有很大的需求,分析师能否为投资者提供有价值的信息,便成为评价其能力的重要指标(伊志宏等,2019)。出于对个人声誉和职业发展的考量,分析师往往有意愿和动机提高盈余预测质量,以向投资者传递有增量价值的特质信息。然而分析师盈余预测是否准确,归根结底还取决于其掌握的信息数量和信息质量(周开国等,2014)。那么,分析师究竟能否有效挖掘并解读前瞻性信息,进而提高盈余预测质量呢?本文从信息效应假说和机会主义假说两方面逻辑进行论证。

(一)信息效应假说

从信息供给角度,根据信号均衡理论,上市公司作为信息优势方能够通过信息披露,改善处于信息劣势方的投资者的信息不对称程度(Spence,1973)。前瞻性信息作为年报公开披露的信息,内含公司未来战略、行动、机遇和挑战等特质信息,已被不少研究证实能够预测未来业绩(Schleicher等,2007;Li,2010),影响投资者决策(潘越等,2021;田高良等,2022),改善资本市场信息环境(Muslu等,2015),并且与盈余质量正向补充(汪炜和袁东任,2014)。这些研究初步证实前瞻性信息具有信息含量,且从事后角度验证了前瞻性信息披露具有一定的可信度。由此可见,高质量的前瞻性信息披露既是上市公司向投资者预告未来发展前景的可行方式,也是投资者合理估计上市公司未来现金流和潜在风险的必要依据(马黎珺等,2019;潘越等,2021;田高良等,2022)。

从信息需求角度,分析师盈余预测是基于公司未来可能发生的事项或采取的行动,对公司未来价值及其价值驱动因子进行展望和预测(马黎珺等,2019)。如此来看,分析师进行盈余预测必然会对公司未来前景信息产生较高需求(Cheng等,2016)。从信息来源看,公开信息和私有信息是分析师盈余预测的重要依据(Schipper,1991;朱红军,2008)。年报所披露的前瞻性信

息,恰是以管理层对公司未来发展前景预测为基础的公开信息,因此高质量的前瞻性信息披露,理论上将能够缓解公司内外信息不对称,满足分析师对前瞻性公开信息的需求,提高其对未来盈余模型参数估计的精确度。此外,前瞻性公开信息的增加还为分析师提供了信息挖掘和信息解读的素材,可能会促使分析师进行更多的私有信息获取,以验证公开信息(朱红军,2008;杨青等,2019)。当年报前瞻性信息较多时,分析师可能会进行更多的实地调研活动,来获取经营现状和未来前景的私有信息(曹新伟等,2015;岳思诗和董大勇,2021)。公开信息与私有信息相互印证,有利于分析师及时准确地修正未来盈余模型,提高盈余预测准确性。这也就是说,前瞻性信息披露既可通过释放公开信息、缓解信息不对称,改善分析师盈余预测质量,还可通过促使分析师进行更多调研活动、获取私有信息,来提高盈余预测准确性。

综上,根据信息效应假说,高质量的前瞻性信息披露能够满足分析师对未来盈余预测的信息需求,进而提高其盈余预测质量。由此提出假设H1a:

H1a:控制其他因素,前瞻性信息披露能够提高分析师盈余预测准确性。

(二)机会主义假说

然而,基于机会主义假说,前瞻性信息披露可能无益于分析师盈余预测质量的提高。从信息供给角度,尽管现有证券监管规则已对上市公司做出强制披露前瞻性信息的要求,但并未强制规定前瞻性信息披露的格式、范围和数量。由此可见,前瞻性信息披露实则属于自愿披露范畴。如此一来,管理者便可能对前瞻性信息进行策略性陈述(程新生等,2013;Feng等,2015),以谋求某些短期利益如稳定股价等(王秀丽等,2020)。因此,若前瞻性信息存在策略性陈述,将无法将公司真实的特质信息传递给投资者及分析师等信息使用者,从而加剧其信息不对称程度。

从信息需求角度,分析师尽管具有信息挖掘和信息解读优势,但依然存在注意力有限问题(Driskill等,2020)。分析师往往需同时追踪多家公司,年报前瞻性信息的集中披露会分散其有限注意力,使其决策呈现有限理性。分析师身处有限注意力和有限理性的情境,未必能够及时识破管理层对前瞻性信息的策略性陈述,导致盈余预测准确性降低(刘会芹和施先旺,2020)。此外,前瞻性信息可能内含较高的不确定性,而信息不确定性往往也会降低分析师盈余预测准确性(Zhang,2006)。

综上,根据机会主义假说,若前瞻性信息披露的信息质量无法得到有效保障,可能将无法提高分析师盈余预测质量。由此提出假设H1b:

H1b:控制其他因素,前瞻性信息披露无法提高分析师盈余预测准确性。

四、研究设计

(一)研究样本

本文以2007—2020年全部A股上市公司为研究样本,剔除金融行业样本和财务数据缺失样本。研究模型中解释变量取滞后一期数据,因此前瞻性信息披露数据期间实则为2007—2019年,而分析师盈余预测数据期间实则为2008—2020年。为剔除异常值的影响,对所有连续变量进行1%至99%的缩尾处理。前瞻性信息披露数据来源于巨潮咨询网披露的上市公司年报,经文本分析获得;分析师盈余预测数据及其他财务数据来源于国泰安数据库。年报前瞻性词语抓取工具为Python2.7,数据处理软件为Stata16.0。

(二)研究变量

1.被解释变量

分析师盈余预测准确性为本文被解释变量。在王雄元等(2017)变量度量的基础上,分析师

盈余预测准确性包括盈余预测误差(*FERROR*)和盈余预测分歧度(*FDISP*),前者取值为“ $\frac{|每股盈余实际值 - 每股盈余预测值|}{(|每股盈余实际值| + 0.5)}$ ”,后者取值为“ $\frac{每股盈余预测值的标准差}{(|每股盈余实际值| + 0.5)}$ ”。

$$FERROR_{i,t} = \frac{|EPS_{i,t} - FEPS_{i,t}|}{|EPS_{i,t}| + 0.5} \quad (1a)$$

$$FDISP_{i,t} = \frac{sdFEPS_{i,t}}{|EPS_{i,t}| + 0.5} \quad (1b)$$

本文关注的是,第*t*-1年年报前瞻性信息披露与分析师对第*t*年年报盈余预测的关系。因此,每股盈余预测值 $FEPS_{i,t}$ 取值为第*t*-1年年报公布以后、第*t*年年报公布以前,分析师对第*t*年年报所作出的每股盈余预测值的均值。

2. 解释变量

前瞻性信息披露(*FTR*),参照胡楠等(2021)、张俊瑞等(2022)、田高良等(2022)等做法采用关键词法进行度量,遵循以下步骤:(1)确定前瞻性关键词的种子词集。主要基于Muslu等(2015)、张俊瑞等(2022)、田高良等(2022)设定的前瞻性词集,并结合中文语境进行修正和补充。种子词集包括“未来、将来、今后、以后、来年、明年”等将来时态词语,“前景、展望、愿景、发展、机遇、机会、时机、契机、目标、目的、挑战、风险”等前景词语,以及“希望、相信、期望、期待、预期、预计、预测、预估、计划、规划、预计、打算、寻求、可能、如果”等展望词语。该做法既能保证词集判断的权威性,又符合中文特殊的语法结构。(2)借鉴胡楠等(2021)做法,以Word2Vec机器学习技术对种子词集进行相似词扩充,扩充词集包括“新年度、下一年、下半年、下个月、下一步、下一阶段、短期内、近期、后期、之后、后续、接下来、新一轮”等将来时态词语,“使命、战略定位、战略方向、新形势、切入点、突破口、良机、有利条件、潜力、成长、崛起、扩张、趋势、面临、持续、不确定性、考验、严峻形势、冲击、威胁”等前景词语,“信心、有望、还将、终将、开拓、拓展、寻找、探寻、探索、谋求、发掘、开辟、争取、倘若、假如”等展望词语。(3)基于前瞻性关键词的种子词集和扩充词集,利用Python中文分词模块对年报文本^①进行自动分词,统计前瞻性关键词的词频。(4)前瞻性信息披露(*FTR*)取值为“ $100 \times \text{年报前瞻性关键词数} / \text{年报总词数}$ ”。

3. 控制变量

借鉴王雄元等(2017)、钟凯等(2020)等研究,选取公司规模(*SIZE*)、盈利能力(*ROA*)、成长能力(*GROWTH*)、财务杠杆(*LEV*)、盈余波动(*VOL*)、收益率波动(*SIGMA*)、管理层预测(*MF*)、两职合一(*DUAL*)、第一大股东持股(*TOP1*)、机构投资者持股(*INS*)、“四大”审计(*BIG4*)等作为控制变量。为控制分析师关注度对分析师盈余预测的影响,本文还将*FCOVER*作为控制变量。主要变量定义情况见表1。

(三) 研究模型

为检验年报前瞻性信息披露对分析师盈余预测准确性的影响,本文构建模型(2)进行实证检验。年报有关控制变量,如*SIZE*、*ROA*、*GROWTH*、*LEV*、*VOL*、*DUAL*、*TOP1*、*BIG4*,与解释变量类似取滞后一期数据;分析师盈余预测有关控制变量,如*SIGMA*、*MF*、*INS*、*FCOVER*,则与被解释变量类似取当期数据。此外,模型(2)不仅经过了公司层面的聚类调整,而且还控制了行业和年度固定效应。若 α_1 显著为负,则H1a得以验证;若 α_1 显著为正或不显著,则H1b得以验证。

^①尽管前瞻性信息披露主要集中在年报管理层讨论与分析(MD&A)部分,但不少公司(如万科)还会在年报其他章节(如致股东部分)披露前瞻性信息。因此,为更加全面地度量前瞻性信息披露,本文选择以年报为文本分析对象。

表1 主要变量定义表

符号	名称	定义
<i>FERROR</i>	盈余预测误差	取值为“ $ (每股盈余实际值-每股盈余预测值) /(每股盈余实际值 +0.5)$ ”
<i>FDISP</i>	盈余预测分歧度	取值为“每股盈余预测值的标准差/ $(每股盈余实际值 +0.5)$ ”
<i>FTR</i>	前瞻性信息披露	取值为“ $100 \times$ 年报前瞻性关键词数/年报总词数”
<i>SIZE</i>	公司规模	取值为“总资产”的自然对数
<i>ROA</i>	盈利能力	取值为“净利润/总资产”
<i>GROWTH</i>	成长能力	取值为“ $(本期营业收入-上期营业收入)/上期营业收入$ ”
<i>LEV</i>	财务杠杆	取值为“总负债/总资产”
<i>VOL</i>	盈余波动	取值为“前两期、前一期及当期总资产收益率”的标准差
<i>SIGMA</i>	收益率波动	取值为“考虑现金红利再投资的个股月收益率”的标准差
<i>MF</i>	管理层预测	若管理层发布业绩预告取值为1,否则取值为0
<i>DUAL</i>	两职合一	若董事长和总经理两职兼任则取值为1,否则取值为0
<i>TOP1</i>	第一大股东持股	取值为“第一大股东持股数/总股数”
<i>INS</i>	机构投资者持股	取值为“机构投资者持股数/总股数”
<i>BIG4</i>	“四大”审计	若为“四大”会计师事务所则取值为1,否则取值为0
<i>FCOVER</i>	分析师关注度	取值为“分析师跟踪人数+1”的自然对数

五、实证结果

(一)描述性统计

表2列示变量的描述性统计。*FERROR*和*FDISP*的均值分别为0.168和0.117,说明分析师盈余预测存在一定误差和分歧。最小值和最大值分别为0.001、1.090和0.000、0.690,说明不同公司的分析师盈余预测准确性存在差异。*FCOVER*的均值为2.150,每个公司平均有7个左右的分析师对其进行跟踪。*FTR*的均值为0.607,说明年报前瞻性关键词数占总词数的比例为0.60%左右,分布特征与张俊瑞等(2022)、田高良等(2022)类似。从控制变量来看,总资产收益率为5.00%,资产负债率为45.30%,营业收入增长率为25.10%,盈余波动为2.30%,股票收益率波动为13.20%,两职合一样本占22.60%,第一大股东持股为36.20%,机构投资者持股为49.60%，“四大”样本占8.20%。整体而言,变量分布特征符合预期。此外,从相关系数来看,*FTR*与*FERROR*、*FDISP*的相关系数均在1%的水平显著为负,说明前瞻性信息披露与分析师盈余预测准确性存在显著的正相关关系,可初步证明假设H1a。各变量之间相关系数的绝对值大都小于0.50,说明不存在严重的多重共线性,模型设定较为合理(篇幅所限,未列示相关系数表)。

表2 描述性统计表

变量	样本量	均值	标准差	最小值	中位数	最大值
<i>FERROR</i>	16173	0.168	0.208	0.001	0.091	1.090
<i>FDISP</i>	16173	0.117	0.123	0.000	0.075	0.690
<i>FTR</i>	16173	0.607	0.142	0.317	0.599	0.981
<i>SIZE</i>	16173	22.400	1.290	20.000	22.200	26.200
<i>ROA</i>	16173	0.050	0.048	-0.107	0.044	0.211
<i>GROWTH</i>	16173	0.251	0.544	-0.464	0.148	4.020
<i>LEV</i>	16173	0.453	0.200	0.060	0.455	0.876
<i>VOL</i>	16173	0.023	0.031	0.001	0.014	0.194
<i>SIGMA</i>	16173	0.132	0.060	0.043	0.119	0.366
<i>MF</i>	16173	0.629	0.483	0.000	1.000	1.000
<i>DUAL</i>	16173	0.226	0.418	0.000	0.000	1.000
<i>TOP1</i>	16173	0.362	0.152	0.091	0.347	0.752
<i>INS</i>	16173	0.496	0.236	0.013	0.531	0.915
<i>BIG4</i>	16173	0.082	0.274	0.000	0.000	1.000
<i>FCOVER</i>	16173	2.150	0.858	0.000	2.200	3.850

(三)回归结果

1.前瞻性信息披露与分析师盈余预测

表3的实证结果用以验证假设H1。*FTR*的系数分别为-0.042和-0.047,且分别在1%的水平显著为负,说明年报前瞻性信息披露频率越高,分析师盈余预测误差和盈余预测分歧度也就越低。这一结果支持了前瞻性信息披露的信息效应假说,表明年报所披露的前瞻性信息是具有信息含量的,分析师能够有效挖掘和解读这类信息并提高盈余预测质量,假设H1a得以验证。研究结论在一定程度上也肯定了我国资本市场前瞻性信息披露实践的可行性和必要性,呼应了已有的研究结论(Schleicher等,2007;Bozzolan等,2009;Li,2010;汪炜和袁东任,2014;Muslu等,2015)。

2.信息质量的调节效应

根据信息效应假说,高质量的年报前瞻性信息披露能够满足分析师对公司未来盈余预测的信息需求,提高盈余预测质量。由此可见,信息质量是前瞻性信息发挥积极效应的重要前提。对此,本文从定量信息质量和文本信息质量两个角度,研究了盈余质量和文本可读性两项年报信息质量特征,对前瞻性信息披露与分析师盈余预测关系的影响。

就盈余质量而言,定量信息往往与文本信息相辅相成、相互补充,有研究证实,高盈余质量能够为文本信息提供可信度保障(汪炜和袁东任,2014),反过来,文本信

息操控往往也用来辅助盈余操控(王华杰和王克敏,2018)。这意味着,当年报盈余质量较高时,其文本信息的可信度可能也较高,而分析师将高质量的前瞻性信息纳入未来盈余预测模型,将能够提高盈余预测质量。就文本可读性而言,较低的可读性会提高分析师阅读和处理信息的难度和成本,降低其盈余预测质量(刘会芹和施先旺,2020)。由此可以合理预期,较低的可读性将阻碍分析师对前瞻性信息的挖掘和解读,即使前瞻性信息具有信息含量,分析师也未必能捕捉其所传递的价值信号,从而降低盈余预测质量。经由上述分析,在信息效应假说下,盈余质量和文本可读性两项信息质量特征,将会增强前瞻性信息披露对分析师盈余预测准确性的积极效应。

首先,借鉴修正Jones模型(Dechow等,1995),以操控性应计利润的绝对值的负数来度量盈余质量,*AQ*取值越大,盈余质量就越高。其次,借鉴胡楠等(2021)的神经概率语言模型,计算

表3 前瞻性信息披露与分析师盈余预测

变量	(1) <i>FERROR</i>	(2) <i>FDISP</i>
<i>FTR</i>	-0.042*** (-3.23)	-0.047*** (-6.20)
<i>SIZE</i>	0.009*** (4.24)	0.001 (0.46)
<i>ROA</i>	0.137*** (2.73)	0.122*** (4.09)
<i>GROWTH</i>	0.004 (1.04)	0.004** (2.09)
<i>LEV</i>	0.125*** (9.75)	0.071*** (9.18)
<i>VOL</i>	0.416*** (6.33)	0.218*** (5.15)
<i>SIGMA</i>	0.293*** (10.40)	0.306*** (17.36)
<i>MF</i>	0.030*** (8.16)	0.022*** (9.91)
<i>DUAL</i>	0.004 (0.84)	0.004 (1.61)
<i>TOPI</i>	0.002 (0.13)	0.023*** (2.59)
<i>INS</i>	-0.101*** (-10.26)	-0.040*** (-6.43)
<i>BIG4</i>	-0.016*** (-2.58)	-0.001 (-0.30)
<i>FCOVER</i>	-0.008*** (-3.22)	0.019*** (13.62)
常数项	-0.017 (-0.35)	0.023 (0.78)
行业/年度	控制	控制
样本量	16 173	16 173
调整R ²	4.69%	6.63%
F值	25.16	32.45

注:括号里为*t*值,*、**、***分别表示10%、5%、1%的显著性水平;下表同。

年报文本中各句子生成概率乘积的对数均值,作为年报文本可读性(*READ*)指标。模型(3)中,*N*为构成文本的句子数,*P_s*为句子生成的概率。*READ*取值越大,表示文本词对搭配顺序出现频率越高,文本可读性也就越高。最后,将盈余质量(*AQ*)和文本可读性(*READ*)作为调节变量,构建模型(4)进行实证检验。当调节变量为盈余质量(*AQ*)时,为缓解其与前瞻性信息披露之间的多重共线性问题,本文借鉴Balli和Sørensen(2013)的做法将交乘项(*FTR*×*AQ*)进行了去中心化处理。

$$READ_{i,t} = \frac{1}{N} \sum_{S=1}^N \log P_s \quad (3)$$

$$FERROR_{i,t}/FDISP_{i,t} = \beta_0 + \beta_1 FTR_{i,t-1} + \beta_2 AQ_{i,t-1}/READ_{i,t-1} + \beta_3 FTR_{i,t-1} \times AQ_{i,t-1}/READ_{i,t-1} + \sum \beta_i Controls + Industry + Year + \varepsilon \quad (4)$$

实证结果如表4所示,*FTR*的系数都显著为负,说明前瞻性信息披露与分析师盈余预测准确性的关系较为稳健。在前两列中,*FTR*×*AQ*的系数为-0.195和-0.078,分别在1%和5%的水平显著为负。结果表明,前瞻性信息披露对分析师盈余预测准确性的积极影响,在盈余质量较高时更为显著。同理,在后两列中,*FTR*×*READ*的系数也显著为负,表明文本可读性也会增强前瞻性信息披露对分析师预测准确性的积极效应。该结果完全符合预期,定量信息与文本信息相互佐证。因此,盈余质量高也可能是前瞻性文本信息质量高的表征。另外,可读性也是前瞻性信息在内的文本信息传递价值信号的必要前提。

表4 信息质量的调节效应

变量	(1) <i>FERROR</i>	(2) <i>FDISP</i>	(3) <i>FERROR</i>	(4) <i>FDISP</i>
<i>FTR</i>	-0.030** (-1.99)	-0.038*** (-3.61)	-0.120** (-2.26)	-0.097*** (-3.09)
<i>AQ</i>	-0.418*** (-10.38)	-0.123*** (-4.72)		
<i>FTR</i> × <i>AQ</i>	-0.195*** (-2.65)	-0.078** (-2.25)		
<i>READ</i>			0.001 (1.05)	0.001 (1.42)
<i>FTR</i> × <i>READ</i>			-0.003* (-1.74)	-0.002* (-1.66)
常数项	0.035 (0.60)	0.084** (1.96)	-0.015 (-0.24)	0.085** (2.39)
行业/年度	控制	控制	控制	控制
样本量	14 064	14 064	15 016	15 016
调整R ²	4.50%	2.89%	4.87%	3.88%
F值	24.10	17.72	24.18	19.04

(四)稳健性检验

1. 盈余预测的其他度量

为避免变量度量误差的潜在影响,以每位分析师最后一次盈余预测作为每股盈余预测值的计算基础,构造分析师盈余预测误差和分歧度,进行稳健性检验。另外,本文还借鉴宋乐和张然(2010)等研究的做法,将*FERROR*取值为“|每股盈余实际值-每股盈余预测值|/期初股票开盘价”,将*FDISP*取值为“每股盈余预测值的标准差/期初股票开盘价”,进行稳健性检验。整体而

言, *FTR*的系数依然在1%和5%的水平显著为负,说明前瞻性信息披露与分析师盈余预测之间的关系较为稳健。

2.前瞻性信息披露的细分

本文还借鉴李岩琼和姚颐(2020)的做法,将前瞻性信息披露(*FTR*)分解为正常前瞻性信息披露(*NFTR*)与超额前瞻性信息披露(*ABFTR*)进行稳健性检验。若分析师盈余预测质量的提高主要是由正常前瞻性信息披露带来的,而非超额前瞻性信息披露,则可进一步证实信息效应假说,排除机会主义假说。首先,选取公司规模(*SIZE*)、盈利能力(*ROA*)、成长能力(*GROWTH*)、财务杠杆(*LEV*)、盈余波动(*VOL*)、行业(*Industry*)、年度(*Year*)作为前瞻性信息披露(*FTR*)的影响因素变量。其次,将*FTR*作为被解释变量、上述影响因素变量为解释变量进行OLS回归,估算得到的拟合值和残差值即为正常前瞻性信息披露(*NFTR*)和超额前瞻性信息披露(*ABFTR*)。最后,将模型(2)中的解释变量替换为*NFTR*和*ABFTR*进行实证检验。实证结果显示,*NFTR*的系数在1%的水平显著为负,而*ABFTR*的系数则不显著。结果表明,分析师盈余预测准确性的提高主要是由正常前瞻性信息披露带来的,超额前瞻性信息披露对此无显著影响。该结果符合预期,且与信息效应假说的内在逻辑相呼应。

(五)内生性检验

1.工具变量法

前瞻性信息披露与分析师盈余预测之间可能存在互为因果的内生性问题。对此,本文借助工具变量法解决该问题。李岩琼和姚颐(2020)在研究研发文本信息披露与分析师盈余预测时,选择以“同行业同年度公司研发文本信息披露的均值”作为工具变量。然而,分析师通常会同时跟踪来自多个行业的多家公司,因此工具变量的选择还需考虑同行业关联公司的影响,以更好地满足相关性和外生性要求。本文在李岩琼和姚颐(2020)研究基础上,以“同行业同年度公司(除本公司以及与本公司被同一分析师跟踪公司以外)的前瞻性信息披露的均值(*FTRIND*)”作为工具变量。

从相关性来看,弱工具变量检验中的Kleibergen-Paap rk Wald F统计量为314.55,表明工具变量(*FTRIND*)并非弱工具变量。从外生性来看,借鉴方颖和赵扬(2011)、潘越等(2021)的研究思路,同时以前瞻性信息披露(*FTR*)和工具变量(*FTRIND*)来解释分析师盈余预测准确性;结果发现,工具变量(*FTRIND*)对分析师盈余预测的影响主要是通过前瞻性信息披露(*FTR*)来实现的,表明工具变量也符合外生性要求。工具变量法的结果如表5第(1)(2)列所示,结果表明,在控制潜在的互为因果问题后,前瞻性信息披露依然能够提高分析师盈余预测准确性。

表5 工具变量和Heckman模型的结果

变量	(1) <i>FERROR</i>	(2) <i>FDISP</i>	(3) <i>FERROR</i>	(4) <i>FDISP</i>
<i>FTR</i>	-0.089*** (-3.47)	-0.137*** (-9.03)	-0.029** (-2.13)	-0.054*** (-6.74)
<i>IMR</i>			0.118*** (3.21)	0.053** (2.41)
常数项	-0.100* (-1.89)	0.003 (0.10)	-0.097* (-1.82)	-0.022 (-0.68)
控制变量	控制	控制	控制	控制
行业/年度	控制	控制	控制	控制
样本量	14 584	14 584	14 584	14 584
调整R ²	4.61%	6.38%	4.79%	7.15%
F值	22.66	32.37	22.26	31.80

2. Heckman模型

前瞻性信息披露与分析师盈余预测之间还可能存在着样本选择偏误导致的内生性问题,为解决这一问题,本文以Heckman模型重复主检验。首先,根据行业年度中位数,构造前瞻性信息披露虚拟变量(*FTRDUM*)。其次,借鉴“工具变量法”的检验思路,将*FTRIND*作为排他性变量。然后,以*FTRDUM*为被解释变量,*FTRIND*以及所有控制变量为解释变量进行Probit回归,计算逆米尔斯比率(*IMR*)。最后,将*IMR*作为控制变量,重复检验模型(2),结果如表5第(3)(4)列所示,*FTR*依然与*FERROR*和*FDISP*显著负相关,表明在控制潜在的样本选择偏误问题后,前瞻性信息披露对分析师盈余预测的影响依然稳健。

3. 行业均值调整

前瞻性信息披露还可能存在着明显的行业特征。对此,本文还对前瞻性信息披露进行行业均值调整(*FTRADJ*)。结果显示,在进行行业均值调整后,*FTRADJ*与*FERROR*与*FDISP*依旧显著负相关,表明本文研究结论受行业因素干扰较小。

4. 固定效应模型

为控制不随时间变化因素对实证结果的影响,本文还以固定效应模型进行内生性检验。实证结果显示,*FTR*的系数依然在1%的水平显著为负,研究结论依然稳健。

六、进一步研究

(一) 机制检验:直接路径和间接路径

1. 直接路径:缓解信息不对称

根据信息效应假说,前瞻性信息披露能充分释放未来相关的公司特质信息,有利于分析师提高未来盈余模型参数估计的精确度,提高盈余预测质量。由此推测,信息不对称的缓解可能是前瞻性信息披露发挥积极效应的潜在路径。对此,本文实证检验了信息不对称在前瞻性信息披露与分析师盈余预测之间的中介效应。

参照Kajüter等(2019)的研究,以股票非流动性度量信息不对称。借鉴Amihud(2002)、苏冬蔚和熊家财(2013)的研究,股票非流动性的度量如模型(5)。其中, r_{iid} 和 v_{iid} 分别是第*t*年第*d*天股票*i*考虑现金红利再投资的收益率和交易金额(单位:百万), D_{it} 是第*t*年股票*i*的交易天数。*ILLIQ*取值越大,股票非流动性越高,即信息不对称程度也越高。

$$ILLIQ = \frac{1}{D_{it}} \sum_{d=1}^{D_{it}} \left(\frac{|r_{iid}|}{v_{iid}} \right) \times 100 \quad (5)$$

表6是中介效应检验的结果。前三列是盈余预测误差(*FERROR*)的结果,后三列是盈余预测分歧度(*FDISP*)的结果。以前三列为例,第(1)列中*FTR*的系数在1%的水平显著为负,系数及显著性与表3第(1)列相比未发生实质改变,说明前瞻性信息披露对盈余预测准确性的改善效应依然存在。第(2)列中*FTR*系数也显著为负,说明前瞻性信息披露能够降低股票非流动性,缓解信息不对称。第(3)列中*ILLIQ*系数显著为正,说明股票非流动性越高,即信息不对称越严重,盈余预测准确性也就越低;此外,*FTR*的系数依然显著为负,说明信息不对称在前瞻性信息披露对分析师盈余预测准确性的影响中发挥部分中介效应,Sobel检验结果也对该中介效应进行了验证。同理,第(4)(5)列的实证结果也符合预期。总体而言,前瞻性信息披露通过释放公司未来特质信息,能够缓解信息不对称,提高分析师盈余预测准确性。

2. 间接路径:促进分析师调研

根据信息效应假说,前瞻性信息披露还可能促使分析师进行实地调研以获取更多私有信

息,来验证前瞻性公开信息的可信度。因此,分析师实地调研也可能是前瞻性信息披露发挥积极效应的间接路径。对此,本文还实证检验了分析师实地调研在前瞻性信息披露与分析师盈余预测之间的中介效应。

表 6 信息不对称的中介效应

变量	(1) <i>FERROR</i>	(2) <i>ILLIQ</i>	(3) <i>FERROR</i>	(4) <i>FDISP</i>	(5) <i>FDISP</i>
<i>FTR</i>	-0.042*** (-3.23)	-0.030*** (-10.77)	-0.036*** (-2.79)	-0.047*** (-6.20)	-0.044*** (-5.86)
<i>ILLIQ</i>			0.181*** (4.15)		0.077*** (3.00)
常数项	-0.017 (-0.35)	0.396*** (38.15)	-0.088* (-1.74)	0.023 (0.78)	-0.008 (-0.25)
控制变量	控制	控制	控制	控制	控制
行业/年度	控制	控制	控制	控制	控制
Sobel Z		-4.390***		-3.292***	
样本量	16 173	16 173	16 173	16 173	16 173
调整R ²	4.69%	25.00%	4.82%	6.63%	6.69%
F值	25.16	84.04	24.77	32.45	31.56

借鉴周冬华和张启浩(2021)的做法,*SURVEY*取值为“分析师调研次数+1”的自然对数。与*FERROR*和*FDISP*类似,*SURVEY*界定为第t-1年年报公布以后、第t年年报公布以前的期间。表7第(1)(4)列结果显示,前瞻性信息披露对分析师盈余预测准确性的积极影响依然存在;第(2)列中,*FTR*系数在5%的水平显著为正,说明前瞻性信息披露能够显著增加分析师实地调研频率;第(3)(5)结果表明,分析师实地调研在前瞻性信息披露与分析师盈余预测准确性之间起到部分中介效应。Sobel检验结果,也支持了分析师实地调研的中介效应。

表 7 分析师调研的中介效应

变量	(1) <i>FERROR</i>	(2) <i>SURVEY</i>	(3) <i>FERROR</i>	(4) <i>FDISP</i>	(5) <i>FDISP</i>
<i>FTR</i>	-0.042*** (-3.23)	0.245** (2.09)	-0.039*** (-3.05)	-0.047*** (-6.20)	-0.046*** (-6.08)
<i>SURVEY</i>			-0.010*** (-8.15)		-0.004*** (-5.06)
常数项	-0.017 (-0.35)	-0.601 (-1.44)	-0.022 (-0.47)	0.023 (0.78)	0.020 (0.71)
控制变量	控制	控制	控制	控制	控制
行业/年度	控制	控制	控制	控制	控制
Sobel Z		-3.029***		-2.772***	
样本量	16 173	16 173	16 173	16 173	16 173
调整R ²	4.69%	17.10%	5.03%	6.63%	6.78%
F值	25.16	27.23	26.48	32.45	32.44

(二)前瞻性信息披露的预测价值

根据信息效应假说,倘若前瞻性信息是管理层针对公司未来经营发展有理有据的合理预测,将得到公司未来基本面特征的验证(马黎珺等,2019)。进而,若将信息细分为积极信息和消极信息,前瞻性信息披露又会侧重于哪类信息?若侧重于前者,前瞻性信息披露可能预示未来基本面向好发展,若侧重于后者,则可能预示未来基本面较差。对此,本文选取盈余增长、创新产出、投资规模三者作为基本面特征,检验前瞻性信息披露对公司未来基本面信息的反映程

度,在进一步验证信息效应假说的基础上,可探究前瞻性信息披露究竟偏向于积极信息披露还是消极信息披露。

首先,定义未来基本面特征。未来盈余增长 $EG_{i,t+n}$ 取值为“(第 $t+n$ 年净利润—第 t 年净利润)/第 t 年总资产,未来创新产出 $RD_{i,t+n}$ 取值为“第 $t+n$ 年的专利授予数量+1”的自然对数,未来投资规模 $INV_{i,t+n}$ 取值为“第 $t+n$ 年购建固定资产、无形资产和其他长期资产支付的现金/第 t 年总资产”。其次,将 $EG_{i,t+n}$ 、 $RD_{i,t+n}$ 、 $INV_{i,t+n}$ 三者作为被解释变量, FTR 为解释变量进行实证检验。表8的结果显示,前瞻性信息披露频率越高,公司未来两年的盈余增长、创新产出,以及未来三年的投资规模也越高。结果表明,前瞻性信息披露能够预测公司未来基本面变化,再次证明前瞻性信息是有信息含量的,而分析师依据高质量的前瞻性信息,也就能够改善盈余决策质量。此外该结果还表明,管理层可能更倾向于披露公司未来发展前景相关的积极消息,而非消极消息。这可能是因为,消极消息相对积极消息往往具有更高的不确定性,披露过多无益于信息使用者的信息挖掘和解读,反而可能引发股价大幅波动。基于上述考虑,管理层的前瞻性信息披露便可能侧重于积极信息。该结果也为前瞻性信息披露监管提供了新思路,即无论积极信息还是消极信息皆应均衡披露,以保证前瞻性信息的完整性和无偏性。

表 8 前瞻性信息披露的预测价值

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	EG_{t+1}	EG_{t+2}	EG_{t+3}	RD_{t+1}	RD_{t+2}	RD_{t+3}	INV_{t+1}	INV_{t+2}	INV_{t+3}
FTR	0.013*** (2.63)	0.017* (1.96)	0.020 (1.58)	0.244*** (2.63)	0.269*** (2.70)	0.084 (0.77)	0.010* (1.71)	0.023*** (2.65)	0.029** (2.36)
常数项	0.151*** (11.45)	0.370*** (13.39)	0.573*** (13.20)	-1.986*** (-7.19)	-1.915*** (-6.38)	-1.406*** (-4.68)	0.042*** (2.60)	0.158*** (6.73)	0.329*** (9.08)
控制变量	控制	控制	控制	控制	控制	控制	控制	控制	控制
行业/年度	控制	控制	控制	控制	控制	控制	控制	控制	控制
样本量	21 002	21 002	21 002	21 002	21 002	21 002	21 002	21 002	21 002
调整 R^2	3.64%	5.99%	7.01%	1.67%	1.56%	3.06%	7.77%	5.45%	5.08%
F 值	18.05	17.51	15.28	20.10	17.53	37.23	23.60	16.51	12.84

七、研究结论与启示

(一)研究结论

信息披露制度是资本市场健康发展的制度基石。我国资本市场的前瞻性信息披露制度自2007年正式确立以来,得到了实务界和学术界的广泛关注。然而,现有研究对前瞻性信息披露的信息含量尚未形成一致结论。本文认为该分歧的原因可能在于,前瞻性信息以前瞻性与文本性为典型特征,相比历史信息 and 定量信息更为隐晦和复杂,解读该类信息也就需要更强的专业能力。分析师作为资本市场重要的信息中介,已被证实具有信息挖掘和信息解读优势。那么,分析师究竟能否有效挖掘并解读前瞻性信息呢?目前国内学术界尚未就该问题给出具体答案。

本文借助文本分析和机器学习解析年报前瞻性文本信息,对该问题进行了研究。研究发现,前瞻性信息披露频率与分析师盈余预测误差、分歧度显著负相关,说明前瞻性信息披露能够提高分析师盈余预测准确性。盈余质量和文本可读性会显著增强前瞻性信息披露对分析师盈余预测的积极效应,说明信息质量是前瞻性信息传递价值信号的必要前提。机制研究发现,前瞻性信息披露主要通过缓解信息不对称、促进分析师实地调研,来提高分析师盈余预测质量。进一步研究发现,前瞻性信息成功预测了公司未来三年的基本面变化。这些证据表明,年报前瞻性信息是有信息含量的,能够为分析师决策提供增量信息。分析师依据前瞻性信息进行盈余预测,既可以提高预测质量,又可以改善资本市场信息环境,进而提高资本市场信息效率。

(二)研究启示

本研究具有一定的启示:(1)前瞻性文本信息是定量信息基础上的增量信息源,能够改善分析师决策质量,提高资本市场信息效率。因此,上市公司应严格遵守信息披露准则,尤其文本信息披露更要保证专业性、规范性、可靠性和可读性,以充分满足投资者、分析师等信息使用者的需求。(2)对监管部门而言,应加大对资本市场文本信息披露的规范和引导力度,不仅需规范文本信息披露的广度和深度,还要针对文本信息虚假陈述建立事后追溯惩罚机制。对前瞻性信息披露而言,监管部门尤其应引导上市公司对积极前瞻性信息和消极前瞻性信息披露进行均衡披露,确保前瞻性信息的完整性和无偏性,以保障前瞻性信息披露制度的有效性。(3)分析师作为资本市场重要的信息中介,承担着信息使用者和信息提供者双重角色,在挖掘和解读公司特质信息时不仅要关注定量信息,还要兼顾文本信息尤其是前瞻性信息。此外,对年报公开披露的前瞻性信息进行挖掘和解读时,分析师还可以进行实地调研活动,以获取更多非公开信息来验证前瞻性公开信息的可信度,从而提高盈余预测质量、改善资本市场信息环境。

主要参考文献

- [1]程新生,谭有超,程昱. 前瞻性信息缓解了信息不对称吗?[J]. *财经研究*, 2013, 39(3): 42-52,63.
- [2]胡楠,薛付婧,王昊楠. 管理者短视主义影响企业长期投资吗?——基于文本分析和机器学习[J]. *管理世界*, 2021, 37(5): 139-156.
- [3]马黎珺,伊志宏,张澈. 廉价交谈还是言之有据?——分析师报告文本的信息含量研究[J]. *管理世界*, 2019, 35(7): 182-200.
- [4]潘越,林淑萍,张鹏东,等. 语言将来时态标记特征与公司股利政策——基于投资者语言认知效应的跨国研究[J]. *经济研究*, 2021, 56(7): 127-143.
- [5]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, 15(4): 1483-1506.
- [6]田高良,薛宇婷,李星,等. 投资者重视管理者的前瞻性吗?——基于年报文本分析的经验证据[J]. *管理工程学报*, 2023, 37(1): 225-236.
- [7]王秀丽,齐荻,吕文栋. 控股股东股权质押与年报前瞻性信息披露[J]. *会计研究*, 2020, (12): 43-58.
- [8]汪炜,袁东任. 盈余质量与前瞻性信息披露:正向补充还是负向替代?[J]. *审计与经济研究*, 2014, 29(1): 48-57.
- [9]曾庆生,周波,张程,等. 年报语调与内部人交易:“表里如一”还是“口是心非”?[J]. *管理世界*, 2018, 34(9): 143-160.
- [10]张俊瑞,仇萌,张志超,等. “深港通”与前瞻性信息披露——基于上市公司年报语言将来时态特征的研究[J]. *证券市场导报*, 2022, (4): 15-26.
- [11]Amihud Y. Illiquidity and stock returns: Cross-section and time-series effects[J]. *Journal of Financial Markets*, 2002, 5(1): 31-56.
- [12]Balli H O, Sørensen B E. Interaction effects in econometrics[J]. *Empirical Economics*, 2013, 45(1): 583-603.
- [13]Bozzolan S, Trombetta M, Beretta S. Forward-looking disclosures, financial verifiability and analysts' forecasts: A study of cross-listed European firms[J]. *European Accounting Review*, 2009, 18(3): 435-473.
- [14]Cheng Q, Du F, Wang X, et al. Seeing is believing: Analysts' corporate site visits[J]. *Review of Accounting Studies*, 2016, 21(4): 1245-1286.
- [15]Driskill M, Kirk M P, Tucker J W. Concurrent earnings announcements and analysts' information production[J]. *The Accounting Review*, 2020, 95(1): 165-189.
- [16]Kajüter P, Klassmann F, Nienhaus M. The effect of mandatory quarterly reporting on firm value[J]. *The Accounting Review*, 2019, 94(3): 251-277.
- [17]Muslu V, Radhakrishnan S, Subramanyam K R, et al. Forward-looking MD&A disclosures and the information environment[J]. *Management Science*, 2015, 61(5): 931-948.
- [18]Schipper K. Analysts' forecasts[J]. *Accounting Horizons*, 1991, 5(4): 105-121.
- [19]Zhang X F. Information uncertainty and analyst forecast behavior[J]. *Contemporary Accounting Research*, 2006, 23(2): 565-590.

Forward-looking Information Disclosure and Analysts' Earnings Forecasts: Evidence from Text Analysis and Machine Learning

Liu Yaoyao¹, Lu Junwei²

(1. School of Economics and Management, China University of Petroleum, Qingdao 266580, China;

2. School of Management, Shandong University, Jinan 250100, China)

Summary: Information disclosure is the institutional basis of the healthy development of the capital market. With the increasing improvement of investors' ability, information disclosure based solely on historical operations cannot fully meet their information needs. Consequently, CSRC revised the information disclosure principles in 2007, which required listed companies to provide outlooks and forecasts on their future development in annual reports. Forward-looking information can forecast a company's future development strategies, business plans, earnings risks, and opportunities and challenges, which contains a significant value for investors to estimate future cash flows and risks. Compared with historical information and quantitative information, it is more obscure and complex, so it requires more professional ability to interpret such information. As professional information users and important information intermediaries of the capital market, analysts have been proven to have advantages in information mining and interpretation. Therefore, can analysts effectively mine and interpret forward-looking information?

Taking A-share listed companies from 2007 to 2020 as the sample and using text analysis and machine learning to measure forward-looking information disclosure, this paper investigates the above question. The results show that the higher the frequency of forward-looking information disclosure, the higher the analysts' earnings forecast accuracy. In firms with high earnings quality and readability, the above positive relationship is more significant. The mechanism test indicates that forward-looking information disclosure can improve the earnings forecast accuracy by alleviating information asymmetry and increasing analysts' site visits. Further research also reveals that forward-looking information can successfully predict firms' fundamental changes in the future, which indicates that such disclosure has its information value and can provide incremental information for analysts and other users of the capital market, which in turn can improve the information efficiency of the capital market.

The possible contributions of this paper are as follows: First, based on the perspective of analysts' earnings forecast, it expands the research on the economic consequences of forward-looking information disclosure. Currently, there is no academic consensus on whether such disclosures are informative, and the findings can promote the understanding of this issue. Second, from the forward-looking disclosure perspective, it enriches the literature on the influencing factors of analysts' earnings forecasts. Research from abroad suggests that forward-looking disclosure can improve the quality of analysts' earnings forecasts, but whether the conclusion applies to China's capital market has not been verified. Therefore, this paper supplements the empirical evidence from China. Third, it has certain enlightenment for listed companies to regulate forward-looking information disclosure, for regulators to deepen the reform of capital market information disclosure, and for analysts to focus on forward-looking information in annual reports when making earnings forecasts.

Key words: forward-looking information disclosure; analysts' earnings forecast; information asymmetry; text analysis; machine learning

(责任编辑:王 孜)