

# Information Disclosure and Corporate Earnings Management: A Textual Analysis Based on the Annual Report

Meiqiong Zhou

Shaanxi Normal University, Xi'an

162066068@qq.com

**Keywords:** Text Mining, Risk Disclosure, Information Quality, Annual Report

**Abstract.** When enterprises disclose information, management, as the information supplier, usually uses surplus means to improve the self-interest to whitewash the contents of the annual report. The text analyzes and quantifies the risk information in the annual report by grabbing the annual report of all A-share listed companies from 2010 to 2017, and examines the impact of disclosure level on the quality of corporate information. The research results show that the annual report disclosure level is positively correlated with the information quality. The more the risk disclosure, the higher the information quality. This research is beneficial to the company's information users to make a reasonable assessment of the risks and values of the company, and to improve the company's internal management.

## 信息披露与公司盈余管理

### ——基于公司年报的文本分析

周美琼

陕西师范大学国际商学院, 西安, 陕西, 中国

162066068@qq.com

**关键词:** 文本分析; 风险披露; 信息质量; 年度报告

**中文摘要.**现代企业在披露信息时, 管理层作为信息供给方, 为提高自身利益通常采取盈余的手段粉饰年报内容。文本通过抓取 2010 年至 2017 年全部 A 股上市公司的年度报告进行文本分析, 对年报中的风险信息进行识别和量化, 检验披露水平对企业信息质量的影响。研究结果显示, 年报披露水平与信息质量呈正相关关系, 风险披露越多, 信息质量越高。本次研究有利于公司信息使用者对企业存在的风险和作合理的评估, 并提高公司内部管理水平。

## 1. 引言

年度报告是公司信息使用者判断其价值的重要媒介。上市公司在年报中将自身的经营现状及风险状态传递给外部信息用户。风险信息披露可以降低信息不对称(Lajili K&Zéghal D, 2005)<sup>[1]</sup>, 因此各国监管机构均颁布相应文件, 鼓励企业在年报中进行风险披露。1997 年, 美国证券交易委员会 (SEC) 要求企业在管理层讨论与分析 (MD&A) 部分披露潜在的市场风险信息。2005 年强制性要求公司在年报第 1A 节单独披露风险信息。2007 年, 中国证监会首次对风险披露提出具体规定, 要求上市公司及时披露相关事项的现状及其可能影响事件进展的风险因素, 并在 2012 年进一步要求在年度报告目录之后单独列示重大风险提示, 披露未来可能会对公司发展战略和经营目标的实施引发负面影响的重大风险。到目前为止, 我国证监会还

要求披露与金融工具有关的各类风险，不断强化风险信息披露规范。

尽管各国强制要求公司进行风险披露，但根据柠檬市场理论，管理层作为信息供给方，通常不愿披露坏消息(Kothari, 2009)<sup>[2]</sup>。出于公司竞争力及自身职业发展的考虑，管理层可能会通过降低年度报告的信息含量、模糊风险信息内容描述等手段向投资者、分析师及利益相关者粉饰公司的负面信息，诱导信息使用者忽略公司现有风险，高估公司价值。因此通过披露包含信息量的风险信息，可以让年报使用者更真实更全面的认识到公司现状，降低了信息不对称程度，信息更加透明，从而抑制管理层的盈余管理行为。随着政策的不断完善和监管力度不断增强，管理层实施盈余管理的潜在风险逐渐增强。在各国陆续颁布了相关文件，对风险披露的规定逐步补充和完善的环境下，我国如何提高年报中风险披露的内容、范围及形式等的规范性和有效性，有待深入探究。

不同于以往采用人工阅读的研究方法，本文利用 python 工具抓取上市公司年报中披露的风险信息，研究我国年报中非结构化文本风险披露对会计信息质量的影响。本文的研究具有一定理论和现实意义，补充了风险披露与会计信息质量两者之间存在联系的证据，为规范并完善我国年报中风险披露机制提供经验证据，对我国强化外部监管而言有一定的借鉴意义。

## 2. 文献综述与假设提出

年报中的非财务信息逐渐受到学者们的广泛关注，其中风险信息更是公司信息使用者研究的热点。会计信息反映了企业近期的财务状况和经营业绩，其披露的信息不仅影响投资者的投资效率，也是政府部门进行相关调控的依据。

早期研究质疑年报中的风险信息实为模板式披露(Li, 2010)<sup>[3]</sup>，缺乏真实信息含量。而后的一些研究发现，风险披露包含信息量，可以产生市场反应，且年度报告中的风险情绪、风险词汇频率的变化值对股票回报波动性、异常交易量及分析师预测均有影响(Li, 2008)<sup>[4]</sup>。目前国内学者对于年报中的风险披露如何影响信息使用者的风险感知存在两种观点。一是风险披露揭示了潜在的风险因素，从而提高了信息使用者的风险感知(Bao&Datta, 2014)<sup>[5]</sup>。Kravet 和 Muslu (2013)<sup>[6]</sup>研究发现，风险披露信息量的增加与股票回报率和交易量的增加有关，这表明文本风险披露增加了信息使用者的风险感知。由于披露的风险信息具有主观性和自愿性，难以精确量化，信息使用者对于风险的判断存在异质性。二是风险披露增加了信息供给，解释了公司存在的风险及突发事件，缓解了信息不对称，从而降低了信息使用者的风险感知(Elmy et al., 2006)<sup>[7]</sup>。王雄元(2018)<sup>[8]</sup>研究发现，年报中风险披露文本的长度及增量越大，公司资本成本及其变化值越小。说明随着风险披露的增加，公司信息披露的透明度逐渐提高，信息使用者可以更多的了解公司存在的风险，从而降低了信息不对称性。

基于对国内外年报风险披露文献的回顾发现，风险披露可以解释公司在负债率、投资、研发等方面的不利变化(Elliot, 2014)<sup>[9]</sup>，满足了用户的信息需求，可使信息使用者通过年报中披露的风险信息分析公司未来的发展机遇和潜在的风险，减少对公司未来经营状况和现金流水平的质疑(Heinle & Smith, 2017)<sup>[10]</sup>。此外，公司信息透明度提高有助于分析师或投资者分析和理解市场，评价风险产生的原因及后果(Hodder, 2001)<sup>[11]</sup>。总体而言，风险披露可以降低信息不对称的程度，改善企业的信息环境，从而加大管理层隐藏负面信息的难度，公司采取盈余管理的行为更容易被发现，因此有助于抑制管理者对操控应计利润的盈余管理行为，提高应计质量。因此文本提出假设：

H1a: 风险披露水平与会计信息质量正相关。风险披露的比例越高，会计信息质量越高。

H1b: 风险披露水平与会计信息质量负相关。风险披露的比例越高，会计信息质量越差。

上市公司的会计信息主要通过年度报告对外部披露，年报的质量影响会计信息质量。在会计研究领域，衡量会计信息质量通常使用财务舞弊、盈余管理和注册会计师审计意见三种标准(蒋瑜峰, 2010)<sup>[12]</sup>。风险披露质量越高，公司信息的外部使用者可以通过披露内容了解公司的详细情况，缓解了双方的信息不对称，更有利于投资者和市场监管者的监督和制约。

公司管理层进行风险披露的动机，也受相关政策法规的影响，进而可能对会计信息质量产生一定影响。

产权性质作为上市公司的一个内部属性，会对企业的组织结构和经营行为产生重大影响。Bushman(2004)<sup>[13]</sup>通过研究各国上市公司的数据，发现产权性质为国有的公司会计信息质量较差，且国有控制程度越高，会计信息质量会越低。在公司监管方面，股权分散和高流动性是欧美公司的特点之一，股东们对信息质量的要求较高，公司披露信息的程度相应提高。我国公司的股权较为集中，特别是国有企业存在政企不分现象，公司的控制权掌握在少数人手中，股东们对信息披露的监管较弱。在市场调控方面，与欧美的资本主义市场经济相比，我国既有市场经济的调节，又有政府的宏观调控。在政府的干预下，国有产权性质为公司承担了隐性担保作用，他们所面临的融资、市场压力等限制要远远小于非国有上市公司，公司可能不愿意或较少的披露有关风险的信息，国有控股的企业更容易出现信息操纵的行为(朱宏泉，2010)<sup>[14]</sup>。综上，可以看出企业产权对公司会计信息质量存在一定影响，企业披露信息的质量和程度受到最终控股人及利益相关者的干预，国有控股的企业因此造成盈余管理监督失衡。因此，提出假设：

H2：在其他条件不变的情况下，与国有上市公司相比，非国有上市公司的风险披露水平的提高更有利于会计信息质量的提高。

### 3. 研究设计

#### 3.1 样本选择与数据来源

本文选择 2010-2017 年全部 A 股上市公司作为样本。剔除了金融类、ST、\*ST 公司及缺失主要变量数据的样本。利用公司盈余质量来对会计信息质量加以衡量，计算所需的财务数据和公司特征数据均来自 CSMAR 数据库；风险披露数据是在上市公司年报中通过 python 技术抓取关键词构建风险披露水平，本文使用的上市公司年报均来自于巨潮资讯网；高管特征数据来自于 CSMAR，并从 wind 数据库、RESSET 数据库及百度百科手工搜集了部分缺失的数据。考虑到样本中的异常值会对研究结果的准确度造成影响，本文对所有连续变量进行了 1% 缩尾处理。最终得到 14309 个观测样本。

#### 3.2 模型设定与变量定义

为研究风险披露水平与会计信息质量之间的关系，我们构建了以下回归模型：

$$IQ = \alpha + \beta_1 Riskdsic + \beta_2 Control + \sum year + \sum ind + \varepsilon$$

(1)

*Riskdsic* 为风险披露水平，本文参考了 Kravet&Mulsu(2013)<sup>[4]</sup>的做法对风险披露进行度量，数据处理过程如下：下载 A 股上市公司 2010 至 2017 年的全部年报；剔除文件名称中含有“英文版”、“修订版”的年报；剔除金融行业、ST 和 \*ST 类企业，以及上市时间不足一年的企业；使用 python 计算机编程将 PDF 格式文件处理为纯文本格式文件 (.txt)，剔除所有数字及各类符号；建立停用词表，剔除停用词；对提取出的文本清理页眉页脚，删除年报中出现“年度报告”的行数据；利用 python 编程计算年报中出现的“风险”、“不确定性”、“潜在”、“可能”、“波动”词汇的频数。本文用所有风险关键词出现的总次数除以报告总字数来衡量风险披露水平。即  $Riskdsic = (\text{全文风险频数} / \text{年报总字数}) * 100$ 。

*IQ* 为会计信息质量。本文采用 Dechow, Sloan 和 Sweeney (1995)<sup>[15]</sup>修正的 Jones 模型，通过分年度分行业回归 (2) 式，提取各组回归系数分别代入 (3) 式，计算每个上市公司的每年的可操控性应计利润 (*ACC*)，再对 *ACC* 取绝对值的负值，以衡量会计信息质量 (*IQ*)，*IQ* 值越大，会计信息质量越高。其中： $TA_t$  为上市公司的总应计利润，等于企业经营利润-经营活动现金净流量； $Asset_{t-1}$  为公司上年度资产总额； $\Delta REV_t$  为公司销售收入的增加额； $\Delta REC_t$



为公司应收账款变动额； $PPE_t$ 为固定资产原值。

$$\frac{TA_{i,t}}{Asset_{i,t-1}} = \alpha_1 \frac{1}{Asset_{i,t-1}} + \alpha_2 \frac{\Delta REV_{i,t}}{Asset_{i,t-1}} + \alpha_3 \frac{PPE_{i,t}}{Asset_{i,t-1}} + \varepsilon_{i,t} \quad (2)$$

$$ACC_{i,t} = \frac{TA_{i,t}}{Asset_{i,t-1}} - \left( \alpha_1 \frac{1}{Asset_{i,t-1}} + \alpha_2 \frac{\Delta REV_{i,t} - \Delta REC_{i,t}}{Asset_{i,t-1}} + \alpha_3 \frac{PPE_{i,t}}{Asset_{i,t-1}} \right) \quad (3)$$

*Control* 包括以下控制变量。产权性质 (*SOE*)：如果公司最终控制人是国有控股，则为 1，否则为 0；资产负债率 (*Lev*)：公司总负债/总资产；资产收益率 (*ROA*)：公司年净利润/年末总资产；第一大股东持股比 (*First*)；上市年限 (*Age*)：上市至今的年数取自然对数；高管性别 (*Gender*)：女性高管占高管团队人数的百分比；换手率 (*Turnover*)：上市首日的股票换手率；公司规模 (*Size*)：对公司总资产取自然对数；账面市值比 (*BM*)：股东权益与公司市值之比。

## 4. 实证结果与分析

### 4.1 描述性统计

表 1 列出了各变量的描述性统计结果。*Riskdisc* 表示公司年报的风险披露水平，最大值和最小值分别为 1.1101 和 0.0067，可以保证充分的差异性，为本文的计量分析提供了基础。从总体平均值来看，我国 A 股上市公司在年报披露中的风险信息含量相对较低。*IQ* 是衡量会计信息质量的指标，最大值和最小值分别为-0.0008 和-0.3934，与其他文章里计算结果基本一致。

表 1 主要变量的描述性统计

Variable	Obs	Mean	Std. Dev.	Min	Max
<i>IQ</i>	14,309	-0.0652	0.1428	-0.3934	-0.0008
<i>Riskdisc</i>	14,309	0.1529	0.0455	0.0067	1.1101
<i>BM</i>	14,309	0.9733	1.0187	0.0109	12.1002
<i>Lev</i>	14,309	0.4504	0.2249	0.0071	8.6118
<i>Size</i>	14,309	22.2373	1.3159	17.7569	28.5087
<i>First</i>	14,309	0.5691	0.1544	0.0132	1.0116
<i>Turnover</i>	14,309	2.1970	1.6926	0.0090	18.5607
<i>Age</i>	14,309	2.1165	0.7834	0	3.2958
<i>ROA</i>	14,309	0.0390	0.1292	-6.7760	10.4009
<i>Gender</i>	14,309	0.1467	0.1527	0	1

### 4.2 相关性分析

Pearson 相关系数 *IQ* 与 *Riskdisc* 在 1%水平上显著为正，说明在不考虑其他因素影响时，风险披露水平越高，会计信息质量越好，与假设 1a 相符。由于相关性分析没有控制其他因素的影响，并且已有研究表明企业规模 (*Size*)、负债 (*Lev*)、高管性别 (*Gender*) 及产权性质 (*SOE*) 等控制变量均对会计信息质量有显著影响，且这些变量从 pearson 相关系数表可以看出与风险信息披露也有高度相关性。因此需要通过多元回归分析，控制其他因素对会计信息质量的影响，得到更可靠的结论。为了排除共线性问题对回归结果的影响，笔者通过计算各变量的 VIF 值，结果均小于 5，表明没有严重的共线性问题。

### 4.3 多元回归分析

表 2 分别进行了普通 OLS 回归和固定效应模型回归，可以看出年报风险披露水平 (*Riskdisc*) 均与信息质量 (*IQ*) 显著正相关。说明公司年报中提供的风险披露文本能够传递公司层面的信息，有助于提高公共信息的供给量和信息透明度，风险披露的信息量越多，越有助

于抑制公司的盈余管理空间，公司的会计信息质量越高，支持了假设 1a。同时可以看到，再加入控制变量后，风险披露水平与信息质量直接依然是显著正相关的，说明风险披露提高了会计信息质量。在控制变量中，上市年限、账面市值比与信息质量显著正相关，负债率、第一大股东持股比例、换手率、规模和资产报酬率与信息质量显著负相关。

表 2 风险披露与信息质量回归分析

变量	IQ					
	OLS		固定效应模型		非国有企业组	国有企业组
	(1)	(2)	(3)	(4)	(5)	(6)
riskdsic	0.0863*** (0.0129)	0.0722*** (0.0128)	0.155*** (0.0164)	0.0938*** (0.0183)	0.0885*** (0.0253)	0.0422* (0.0236)
bm		0.00349*** (0.000892)		0.0135*** (0.00139)	0.0195*** (0.00244)	0.00570*** (0.00152)
lev		-0.0731*** (0.00366)		-0.0861*** (0.00662)	-0.0917*** (0.00853)	-0.0614*** (0.00971)
size		0.00698*** (0.000646)		-0.00369** (0.00164)	-0.00452** (0.00209)	-0.00245 (0.00247)
large		-0.0403*** (0.00420)		-0.0736*** (0.00894)	-0.0804*** (0.0117)	-0.0675*** (0.0126)
Turnover		-0.00267*** (0.000374)		-0.000151 (0.000443)	6.27e-08 (2.38e-06)	1.88e-06 (2.64e-06)
female		-0.0172*** (0.00366)		-0.0166** (0.00741)	0.00458 (0.00328)	0.0157*** (0.00409)
age		-0.00851*** (0.000918)		0.0105*** (0.00253)	-0.0239 (0.0227)	-0.0570** (0.0238)
soe		0.0107*** (0.00129)		-0.00254 (0.00570)		
roa		-0.0955*** (0.0133)		-0.0980*** (0.0173)	-0.0337*** (0.00934)	0.00785 (0.0108)
Constant	-0.0770*** (0.00205)	-0.152*** (0.0130)	-0.0875*** (0.00256)	0.0571* (0.0312)	0.0879** (0.0394)	0.0180 (0.0474)
Industry & Year	NO	NO	YES	YES	YES	YES
Observations	14,309	14,309	14,309	14,309	8,007	6,302
R-squared	0.003	0.051	0.008	0.039	0.044	0.042

将样本按产权性质 (SOE) 分为国有企业和非国有企业两组。表 2 第 (5) 和 (6) 组的分组回归结果显示：两组系数均为正，说明风险披露可以提高信息质量。非国有企业组在 1% 水平上显著，而国有企业组仅在 10% 水平上显著，说明在其他条件不变的情况下，相比于国有企业，非国有上市公司的风险披露水平和会计信息质量的显著性更高，即非国有企业的风险披露的提高更有利于会计信息质量的提高，与假设 2 相符。

### 5. 研究结论

本文以上市公司的年报文本为研究对象，采用文本挖掘方法衡量风险披露水平，将盈余管理作为上市公司的信息质量的代理变量，讨论了我国年报风险披露能否提高信息质量，并考虑产权性质作为调节变量对风险披露和信息质量的影响。研究发现：第一，年报中风险信息的披露可以降低公司管理层盈余管理程度，提高了信息质量。第二，国有上市公司的政府背景有一定的隐性担保的作用，其在披露风险时的信息质量较低，非国有上市公司的风险披露的提高更有利于信息质量的提高。

总体而言，年报中的文本部分能够披露风险，且披露的风险信息能够降低内外界信息不

对称性，因此，对管理层的盈余管理起到一定抑制作用，提高公司的信息质量。上市公司在年度报告中自愿披露的风险信息越多，越是能向外传达一种积极信号并加强信息使用者的信任，有利于利益相关者对公司的价值和风险做出正确的估计，这对我国信息使用者是一种保护。完善风险披露机制对我国强化外部监管而言有一定的借鉴意义。

## References

- [1] Lajili K, Zéghal D. Labor cost voluntary disclosures and firm equity values: Is human capital information value-relevant? [J]. *Journal of International Accounting Auditing & Taxation*, 2005, 14(2):121-138.
- [2] Kothari S P, Li X, Short J E. The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis[J]. *The Accounting Review*, 2009, 84(5):1639-1670.
- [3] Li F. Textual Analysis of Corporate Disclosures: A Survey of the Literature[J]. *Social Science Electronic Publishing*, 2010a, 29: 143–165.
- [4] Li F. Annual report readability, current earnings, and earnings persistence[J]. *Journal of Accounting & Economics*, 2008, 45(2–3):221-247.
- [5] Bao Y, Datta A. Simultaneously Discovering and Quantifying Risk Types from Textual Risk Disclosures[J]. *Management Science*, 2014, 60(6):1371-1391.
- [6] Kravet T, Muslu V. Textual risk disclosures and investors' risk perceptions[J]. *Review of Accounting Studies*, 2013, 18(4):1088-1122.
- [7] Elmy F J, Leguyader L P, Linsmeier T J. A review of initial filings under the SEC's new market risk disclosure rules[J]. *Journal of Corporate Accounting & Finance*, 2006, 9(4):33-45.
- [8] Wang, Gao. Annual Report Risk Disclosure and Equity Capital Costs[J]. *Financial Research*, 2018(01):174-190.
- [9] Elliot W B, Rennekamp K M, White B J. The Paradoxical Behavioral Effects of a Directional Goal on Investors' Risk Perceptions and Valuation Judgments[J]. *Social Science Electronic Publishing*, 2014:1-20.
- [10] Heinle M S, Smith K C. A theory of risk disclosure[J]. *Review of Accounting Studies*, 2017, 22(4):1459-1491.
- [11] Hodder L, Koonce L, Mcanally M L. SEC Market Risk Disclosures: Implications for Judgment and Decision Making[J]. *Accounting Horizons*, 2001, 15(1):49-70.
- [12] Jiang Yufeng, Yuan Jianguo. Accounting Information Quality Evaluation: Summary and Prospects[J]. *Finance and Accounting News*, 2010(6):3-5.
- [13] Bushman R, Qi Chen, et al. Financial accounting information, organizational complexity and corporate governance systems[J]. *Journal of Accounting & Economics*, 2004, 37(2):167-201.
- [14] LI Yajing, ZHU Hongquan. Board Control and Earnings Quality: An Empirical Analysis of China's Securities Market[J]. *Mathematics Practice and Cognition*, 2010, 40(21):000012-25.
- [15] Dechow P M, Sloan R G, Sweeney A P. Detecting Earnings Management[J]. *Accounting Review*, 1995, 70(2):193-225.