

会计师事务所所有基于关键审计事项的审计风格吗

——基于中国上市公司披露新版审计报告的经验证据*

田高良 陈匡宇 齐保垒

(西安交通大学管理学院/管理教学实验中心 710049)

【摘要】 本研究手动收集首次披露关键审计事项的中国上市公司 2016-2018 年新版审计报告中关键审计事项段的内容及文本信息,并计算得到其文本特征指标,建立配对样本,运用多元回归方法分析关键审计事项在披露内容和文本特征的相似度和可比性,探究我国会计师事务所在披露新版审计报告时是否表现出基于关键审计事项的审计风格。结果表明,同一年份行业内,被同一家会计师事务所审计的公司在披露事项类别文本相似度、披露详细程度、语调指标和可读性指标上都具有显著更大的相似性,因此在关键审计事项角度有独立的审计风格。本研究有利于理解关键审计事项向市场传递信息的机制和意义,以促进审计工作者通过关键审计事项段披露更决策有用的信息,提升资本市场有效性。

【关键词】 审计风格;关键审计事项;新式审计报告;财经文本分析

一、引言

审计行业急需改变“三段式审计报告”的现状,这种格式化的审计报告和审计意见使报告使用者只能得到非常有限的同质化信息,无法识别出审计师了解公司时重点关注的对象和企业的重大风险隐患。提供信息过于有限的报告也会使审计师的责任有限,可能在审计工作中忽视一些必要的审计程序,甚至有意串谋、忽视风险,把本应缓解代理问题的审计监督机制转化成对客户公司贩卖标准无保留意见的渠道。

我国审计行业发展较晚、相对于国际审计行业较不成熟,但隐患和问题同样存在。我国部分会计师事务所的不良工作习惯很可能影响了财务报告的信息含量,也使审计服务作为代理问题的解决机制趋于无效化。

基于此背景,各个国家和地区都在推出加强审计报告信息含量的新规和政策,其中最重要的国际性文件是《国际审计准则第 701 号——在独立审计报告中沟通关键审计

事项》,其中指出审计师应当利用职业判断确定那些被视为关键审计事项的事项,并在审计报告中传达这些事项。ISA 701 也对关键审计事项做出了如下的定义:“在审计师的专业判断中,这些事项对本期财务报表的审计最为重要。关键审计事项选自与负责管理的公司人员沟通的事项。”这证明:关键审计事项是审计师设计审计工作时着重考虑的、最容易导致错报的重大风险点,这会使报告使用者理解审计师做出审计结论的思考过程,并且获得决策有用的额外信息。2016 年 12 月 23 日中国财政部印发《中国注册会计师审计准则第 1504 号——在审计报告中沟通关键审计事项》,该准则要求在上市公司的审计报告中增设关键审计事项部分,披露注册会计师根据职业判断为对本期财务报表审计最为重要的事项。关键审计事项准则对 A+H 股上市公司 2017 年 1 月 1 日生效,对沪深上市公司 2018 年 1 月 1 日生效。此次审计报告准则的改革对审计报告的信息含量提升具有重要意义,因而成为实务界和学术界研究的前

* 本文受国家自然科学基金项目(71672141)、国家社科基金后期资助项目(21FGLB006)、教育部人文社会科学研究项目(21YJA630082)资助。

沿热点。相对原来的审计报告,关键审计事项的披露将大大增加审计过程的信息含量。会计师事务所在审计过程中主要关注、识别、沟通哪些关键审计事项,将有效反应事务所的审计风格;因此可以通过研究不同事务所在审计过程中披露的关键审计事项,研究会计师事务所是否具有独立的审计风格。

审计风格的定义来源于 Francis 等 (2014) 的研究,是一种可比性的反映,体现了会计信息的基本质量特征。Kothari 等 (2010) 认为,每一家四大会计师事务所都有自己独特的内部工作规则,指导和规范审计师应用审计和会计准则。基于这种想法,我们认为会计师事务所在面对审计准则变化和修改时,会针对审计师展开培训及继续教育,使同一个会计师事务所的审计师在识别关键审计事项时会着重注意识别某类关键审计事项并偏爱披露这些事项,并且按照某种既定模板书写关键审计事项段;故同一个会计师事务所审计的公司披露的关键审计事项会有内容和文本特征的双重相似度和可比性。我们用两种方法衡量关键审计事项的审计风格:第一种方法是研究同一行业、同一年度中同一个会计师事务所审计的一对公司与不同的两个会计师事务所审计的一对公司在关键审计事项类别文本相似度和三种文本特征方面的差异。第二种方法为基于固定效应模型的联合检验,使用审计师事务所固定效应模型来检验审计报告关键审计事项段三种文本特征指标的共性。我们使用的样本来自 2016-2018 年披露新版审计报告的所有公司。与我们的主要假设一致,我们发现在同一行业、同一年度,由同一个会计师事务所审计的两家公司比由两个不同的会计师事务所审计的两家公司有更相似的关键审计事项段。这一结论具体表现为:在同一行业、同一年度内被同一个会计师事务所审计的一对公司相对于被不同会计师事务所审计的一对公司,有更大的关键审计事项类别文本相似度,更小的篇幅差异、语调指标差异和可读性指标差异。此外,会计师事务所固定效应是三种文本特征指标的一个具有统计学意义的决定因素。这些结果对于控制一系列可能产生冲击或以其他方式影响关键审计事项披露的其他因素都是稳健的,并且通过了多种稳健性检验。因此,我们认为,我国的会计师事务所表现了基于关键审计事项的审计风格。我们还研究了一对由同一家四大会计师事务所审计的公司是否比一对由同一家非四大会计师事务所审计的公司有更相似的关键审计事项段。这项检验基于这样一种观点:四大会计师事务所由于其规模、声誉及国际化水平都更好,将比非四大会计师事务所拥有更专业的技术团队和更多的准则解读专家,更早地了解国际审计准则的

变化的应用指南,导致会提供更高特异性的关键审计事项披露,即披露相似度和可比性更弱,更特定于企业的具体风险点,体现了高质量的审计服务与激进的披露风格。与上述预测一致,我们发现的实证证据表明,四大会计师事务所表现的关键审计事项审计风格弱于非四大会计师事务所。

本文对现有研究及实务领域有几点贡献。这是第一个通过除财务报表应计结构可比性外检验审计风格的研究。之前的审计风格研究,从 Francis 的研究起,往往使用审计师抑制盈余管理的程度(财务报告应计项目相似性)来衡量审计风格(Francis 等, 2014)。李留闯等 (2017) 将审计风格由会计师事务所层面延伸到个人审计师层面。本文尝试提出了使用非财务信息披露及文本信息的特征和可比性来研究审计风格,创新性地提出了一种新的审计风格度量方法,并在多种维度验证了其存在性。具体来说,我们首次尝试将关键审计事项及其文本特征与审计风格建立联系,并且用实证研究验证了这种风格的存在,为更多样的审计风格衡量提供了新思路,对于传统上使用应计项目衡量审计风格的方法做出了创新,相比之前 Francis 等提出的会计师事务所审计风格和李留闯等提出的个人审计师的审计风格具有一定的边际贡献。第二,本文拓展了对于关键审计事项的研究。之前对于关键审计事项的研究,主要是较为实践性地分析关键审计事项的披露偏好和披露事项,或对新版审计报告的改革使用事件研究法进行研究。本文通过实证研究方法、机器学习算法以及财经文本分析方法的应用,对文本信息、非财务信息为主的关键审计事项段进行了披露内容和文本特征的多种分析,提供了针对关键审计事项段进行信息挖掘和后续研究的一种可能性,也表明关键审计事项的信息含量值得进一步挖掘分析。第三,在试图解释本文的实证结果时,我们通过对关键审计事项披露背后所代表的会计师事务所工作流程、工作习惯和内部工作规则进行推演分析,在审计学术研究和审计实务之间建立了联系。审计工作具有一定的“黑箱效应”,传统的审计研究无法窥探到黑箱内部的工作流程和工作方法,只能通过审计费用、审计工时等间接研究。我们通过对关键审计事项这一具有信息沟通功能的非财务型文本信息进行财经文本分析,尝试去打破审计实务工作中的“黑箱效应”,在不接触审计底稿也不进行案例分析的情形下研究分析审计师的思维逻辑和工作惯例。我们推断审计师在这些关键审计事项相关会计科目的工作中付出了大部分的审计努力和工作时长,从而很大程度上影响了审计意见结论。本文通过对审计报告所隐含的信息含量、审计风险以及审

计质量等话题进行探讨,加之对实务工作者进行访谈,深度挖掘了关键审计事项信息沟通的机制和对资本市场传递信息的积极意义。

本文结构安排如下:第二部分为相关研究述评和理论分析;第三部分为假设建立;第四部分为研究设计;第五部分汇报实证结果;第六部分为稳健性检验;第七部分为与实务工作者及专业人士的探讨;最后为结论。

二、相关研究述评及理论分析

(一) 审计风格及其影响因素

审计风格最早由 Francis 等于 2014 年明确提出,将其定义为会计师事务所内部在执行审计准则、一般公认会计准则(GAAP)以及会计原则时的独特内部工作规则(Francis 等, 2014)。审计风格的研究大多都基于会计信息质量的重要属性:可比性。财务会计准则委员会(FASB)将可比性定义为一种信息质量,使用户能够识别两家公司财务业绩的异同。FASB 和国际会计准则理事会的联合概念框架强调:可比性是财务信息的一项基本属性,提高了其会计信息的实用性(FASB, 2010)。事实上, FASB 指出,财务报告的可比性是制定会计准则的主要原因(FASB, 1980)。

Becker 等(1998)和 Francis 等(2011)发现,四大审计师事务所客户的异常应计金额比非四大客户要小。Rahman 等(2012)发现被审计公司的自身属性如行业、公司财务杠杆、内部控制质量和规模等会影响审计师对于审计方法和工作风格的选择。Kothari 等(2010)认为会计师和审计师将制定“工作规则”,以规范会计实践。Francis 等(2014)发现同一年份、同一行业中被同一家事务所审计的公司会比不同事务所审计的公司拥有更加相似的应计结构和更可比的财务报告,即四大会计师事务所存在审计风格。

(二) 关键审计事项的相关研究

关于关键审计事项段的研究,已有文献主要由披露关键审计事项段后的影响,和关键审计事项段的披露特征、偏好两个部分展开研究。

Louis-Philippe 等(2014)分析审计报告中提到的事项影响了参与者的信息搜索过程,并增加了他们对审计报告中提到的财务报表披露的关注。Backof(2015)认为关键审计事项的确认和披露主要以和公司治理层的沟通为标准,同时也与审计师对合理保证的理解和定义有关。Reid 等(2015)认为强制披露关键审计事项这一事件会使报告使用者获取审计师对被审计公司的内部了解,从而提高审计报告的决策有用价值。Pinto 和 Morais(2019)发现审计费

用与披露的关键审计事项数量之间存在正相关关系。

很多学者讨论了关键审计事项的披露特征和披露偏好。柳木华和董秀琴(2018)整理了各行业平均披露事项条数和报告关键审计事项的识别依据,发现计价和分摊是最常见的两种认定来源。路军和张金丹(2018)从披露方式、数量分布、详细分类、行业特征和事务所特征等方面全面分析了披露特点和偏好。骆韦等(2019)整理了关键审计事项最常见的几种事项,其中在所有行业中都最常见的三类事项为:收入确认、应收账款类和商誉减值类。

很多研究将 A+H 股上市公司首次强制披露关键审计事项作为事件研究,以证明试行关键审计事项对市场的作用。叶妮(2018)发现披露关键审计事项这一事件提升了会计信息质量。李延喜等(2019)发现披露更多关键审计事项可以显著降低上市公司的应计盈余管理强度。王艳艳等(2018)发现披露关键审计事项段的公司的累计超额收益率在披露前后的变化显著高于未披露公司。

而部分学者持有相反的观点,即关键审计事项披露具有局限性:吴秋生和独正元(2018)发现虽然关键审计事项披露整体有效,但存在阐述不到位、信息不准确、未与企业状况充分联系等三种问题。岑敏儿和潘家恒(2019)认为关键审计事项的沟通整体有意义,然而少部分事项有格式化的缺陷,但绝大部分事项都缺失了审计程序的实施结果,也没有表述出超越一般审计处理的特殊重视。

(三) 财经文本特征的相关研究

财经文本特征,如语调指标、可读性指标等,含有超额信息,学术界对财经文本的意义展开了越来越多的重视。Li(2008)考察了年度报告可读性与公司业绩和盈余持续性之间的关系,发现收益较低的公司的年度报告难以阅读;年度报告更容易阅读的公司有更持久的正收益。Davis 等(2015)使用上市公司经理关于收益的公开电话会议文本为研究样本,发现经理使用积极词汇与消极词汇的语调信息可以影响投资者对电话会议披露信息的解释。谢德林和林乐(2015)发现管理层业绩说明会的语调可以预示公司下一会计年度的业绩。孟庆斌等(2017)发现管理层讨论与分析章节可读性越大,信息不对称程度就会越高,故对股价崩盘风险的降低作用越大。王雄元等(2018)基于管理层讨论与分析章节,发现以余弦度量的信息披露相似度与审计费用显著负相关。胡楠等(2021)发现管理者短视主义导致的研发支出减少和资本投资效率降低会损害企业的未来绩效。

根据上述文献可以看出:1)对于审计风格的研究基于会计信息基本质量特征之一的可比性,但大多以企业应

计结构的财务报告相似度作为衡量标准。2) 针对关键审计事项进行研究, 大多是使用其作为事件分析关键审计事项带来的积极意义或者不足之处; 抑或是总结关键审计事项在各个行业的披露特征、披露偏好和披露形式, 很少将关键审计事项与事务所的审计风格联系起来。3) 关于文本可读性及语调指标的研究多来自于管理层讨论与分析章节、业绩说明会等, 来研究文本背后隐含的信息和对市场反映的影响, 但鲜有针对审计报告中关键审计事项段文本分析审计师审计工作特点和审计风格的研究。因此, 本文试图联系关键审计事项、审计风格、财经文本分析三个领域, 用财经文本分析方法量化关键审计事项, 从而检验事务所是否有基于关键审计事项的审计风格。

三、假设建立

会计报告是财务报告系统特征复杂相互作用的结果, 这些特征包括会计准则、会计准则的解释、执行和诉讼等, 所有这些都会影响可比性。审计报告也应具有类似的特征, 对于审计师事务所这一盈利性机构, 由其内部理论研究者、准则解读专家深度理解会计准则、审计准则更新与变化, 并且形成自身独特的内部工作规则和范式, 从而指导审计师应用新版会计准则和审计准则是比较经济的, 有助于快速准确地执行新准则 (Skinner, 2010)。Francis 等 (2014) 认为审计师事务所具有差异化的追求, 会通过不同的工作方法、工作规则、工作范式 (甚至是专属的品牌色调) 来将自身与其他竞争者区分开, 这种主观选择也对风格产生具有贡献。在了解了国内外关于审计风格的相关讨论与概念形成过程后, 我们认识到审计风格来自于会计师事务所的独特内部工作规则。关键审计事项作为审计师在执行审计工作时与管理层沟通的、得到具有专业判断能力的审计师着重重视的重大风险点, 除了被审计公司的固有属性差异外, 它的披露应当会反映签字注册会计师在其供职的事务所受到的对会计准则的理解、解释和执行流程的专业判断指导与培训。换言之, 在面对新审计准则的要求时, 各大会计师事务所应当对注册会计师展开了审计准则解释和继续教育, 以保证注册会计师在负责年度财务报告审计时能良好应对这一新准则的变化, 而这种培训行为会使会计师识别和沟通关键审计事项时反映出会计师事务所的整体特征、思维方式和习惯, 故关键审计事项的相似度能成为继应计结构差异后有效反映会计师事务所审计风格的指标。

我们通过大量阅读审计师在审计报告中沟通的关键审计事项段, 发现关键审计事项段主要由如下三部分组成: 1. 审计师事务所识别的关键审计事项; 2. 审计师识别关

键审计事项的理由; 3. 审计师对于关键审计事项段展开了怎样的额外审计程序。由此可以看出, 关键审计事项段文本具有层层递进的逻辑关系。其中, 第二部分和第三部分是我们额外关注的, 因为审计师为什么识别关键审计事项隐含了审计师的职业怀疑, 这一过程不仅来源于签字注册会计师的个人判断, 也涵盖了会计师事务所的思维模式, 从而有可能出现同质化, 成为基于关键审计事项的审计风格的贡献来源; 第三部分: 关于关键审计事项段开展的额外审计程序更可以表明出会计师事务所的工作模式和工作方法, 显然也将带有审计师背后的会计师事务所的风格烙印。

2021 年 7 月, 证监会集中部署专项执法活动, 依法从严从快从重查办 16 起重大典型案件, 其中提到部分注册会计师和会计师事务所严重背离职业操守, 丧失“看门人”作用的案件, 包括上市公司相关审计机构单方面依赖公司提供信息, 搞“抄账式”审计, 未对舞弊风险事项执行充分的审计程序, 甚至还与上市公司提前商定审计意见类型。这类“抄账式”审计的客观存在可能暗示着在披露关键审计事项时, 识别对当期审计最重要的风险因素的机制也可能流于形式, 会计师事务所所有动机按照同行业可比公司或同一公司去年披露的关键审计事项搞“复制粘贴”, 使披露出的关键审计事项背离当期审计风险来源, 不符合公司实际情况, 无法向信息使用者传递足够的风险提示信息。因此, 关键审计事项段的逻辑结构和内容是我们假设会计师事务所具有基于关键审计事项段的审计风格的一个重要来源, 会计师事务所节省审计工时和减少审计投入所用的“抄账式”审计也可能使披露关键审计事项出现一定的同质化, 从而观测到审计风格。具体来说, 基于关键审计事项段的审计风格来源于如下因素:

第一, 审计报告中沟通的关键审计事项会呈现出按会计科目分类的集中特征。Kothari 等 (2010) 认为, 将笼统的、原则导向的会计准则下放给数千名个人审计师来自行解释, 将会是不经济的; 而 Francis 等 (2014) 也认为, 审计风格一旦存在, 将会使同一家会计师事务所在工作中系统地偏爱检验某些会计科目, 或忽视某些会计科目。通过和实务工作者及专家的沟通, 我们认识到: 关键审计事项是审计工作的要点、重点、难点, 是会计师事务所和注册会计师在整个审计流程和执行审计程序时的开展点。考虑到关键审计事项与审计师事务所设计、布局审计工作之间的紧密联系, 我们认为在审计师事务所的指导和范式影响下, 同一家会计师事务所的审计师会更偏爱于沟通某些关键审计事项, 并围绕这些事项展开审计工作。我们提取了

审计报告中关键审计事项段的类别子标题, 这些子标题会概括披露的关键审计事项类别和科目, 如“应收账款减值准备”“收入确认”等。通过文本相似度分析方法, 我们对事项类别形成的文本信息进行配对文本相似度计算, 这一相似度指标可以衡量会计师事务所披露关键审计事项类别的可比性, 相似度越高, 意味着披露关键审计事项类别越接近。因此, 当公司配对样本中的两家配对公司是被一家会计师事务所审计时, 审计师将更倾向于披露类别一致的关键审计事项, 故两家公司应当会有类别文本相似度更高的关键审计事项段。

第二, 我们在阅读上市公司披露的审计报告时发现, 部分公司审计报告文件的文件名为“XX 会计师事务所新审计报告格式”, 这使我们推测: 针对新式审计报告, 部分会计师事务所制作了供审计师使用的报告格式模板, 来保证审计质量和审计结果的统一性, 这也符合会计师事务所的标准化工作流程。使用报告格式模板会使审计师书写审计报告更规范, 但同时也可能会使审计师在报告留白范围内描述与沟通关键审计事项, 故使字段篇幅接近。如前所述, 关键审计事项段包括审计师识别关键审计事项的理由和围绕关键审计事项执行的额外或特殊审计程序, 故这部分字段篇幅很可能体现了会计师事务所的工作规则、工作范式乃至工作量。篇幅可衡量披露关键审计事项的详细程度以及执行额外审计程序的复杂程度, 同一会计师事务所的披露篇幅将可能更接近, 体现了不同的信息沟通含量与不同的审计努力程度, 这在某种程度上也与审计质量有关。当公司配对样本的两家配对公司是被一家会计师事务所审计时, 关键审计事项段将会有更接近的篇幅规模。

第三, 审计师在识别关键审计事项到最后选择与报告使用者在审计报告中沟通某事项的过程中, 一直会存在与公司管理层的斡旋、妥协等。这使审计师可能会调用积极和消极词汇使语调存在差异。此外, 如前文所述, 审计师会倾向于用自身事务所偏好的专业词汇和句式来将自身与其他会计师事务所区分开。因此, 当公司配对样本的两家配对公司是被一家会计师事务所审计时, 关键审计事项段的语调指标差异会更小。

最后, 可读性衡量了文本信息易读与难读的水平, 会隐含不同的信息含量。审计师关键审计事项段可读性的差异同样会反映出会计师事务所的工作规则、遣词造句和文本书写习惯, 其中隐含的是审计师作为个人被会计师事务所这一整体潜移默化下形成的思维模式, 体现了审计风格的重要来源。因此, 当公司配对样本的两家配对公司是被一家会计师事务所审计时, 关键审计事项段的可读性指

标差异会更小。

审计风格同时来自于被审计公司和会计师事务所, 审计风格这一概念虽然增加了审计师事务所的工作效率, 但也可能带来思维惯性、思维惰性和错判风险, 如审计师没有准确识别被审计公司的核心风险点, 而是披露了其事务所更熟悉或其在培训中经常被提及的范式化风险点作为关键审计事项。Francis 等 (2014) 认为审计风格有可能以牺牲相关性、真实性和可比性为代价, 强加一致性。FASB (2010) 强调: 一致性可能使事物看起来相同, 但一致性不是可比性, 其目标是使“相似的事物”看起来相似, “不同的事物”看起来不同。本文不预测如果基于关键审计事项的审计风格存在, 这种风格的客观存在性是否会损害审计质量, 但我们认为高度一致性的关键审计事项披露有可能意味着审计师的职业怀疑被思维惰性和惯性思维影响, 使披露出的“关键审计事项”偏离被审计公司当期最重大的风险点, 从而增大审计风险。四大会计师事务所被认为拥有高水平的职业怀疑和审计质量, 并能更有效地执行会计准则和审计准则 (Becker 等, 1998; Francis 等, 2011), 因此我们假设被同一家四大会计师事务所审计的一对公司相对于被同一家非四大会计师事务所审计的一对公司, 会披露更多特异性事项和更多与客户公司相关的特异性文本, 故关键审计事项相似度会更弱。

应当注意的是: 审计师识别和沟通上市公司的关键审计事项是和被审计公司的固有属性分不开的。因此, 在研究审计风格时, 我们加入合适的控制变量来控制被审计公司固有属性, 来得到稳健的实证结果。

由上, 本文提出如下三个假设:

H1: 被同一家会计师事务所审计的一对公司会比被不同的会计师事务所审计的一对公司披露事项类别文本相似度更高的关键审计事项段。

H2: 被同一家会计师事务所审计的一对公司会比被不同的会计师事务所审计的一对公司披露文本特征更接近的关键审计事项段, 具体体现为:

H2a: 被同一家会计师事务所审计的一对公司会比被不同的会计师事务所审计的一对公司披露篇幅更接近的关键审计事项段;

H2b: 被同一家会计师事务所审计的一对公司会比被不同的会计师事务所审计的一对公司披露文本语调指标差异更小的关键审计事项段;

H2c: 被同一家会计师事务所审计的一对公司会比被不同的会计师事务所审计的一对公司披露文本可读性指标差异更小的关键审计事项段。

H3: 被同一家四大会计师事务所审计的一对公司的关键审计事项相似度低于被同一家非四大会计师事务所审计的一对公司。

四、研究设计

基于上述假设,我们通过披露关键审计事项在内容和文本特征上总计四个维度的指标,在其他文献的基础上检验了关键审计事项披露的横截面相似性。此外,我们通过会计师事务所固定效应模型来检验审计师客户公司关键审计事项文本特征的系统相似性,提高了稳健性。

(一) 变量界定

1. 被解释变量: 关键审计事项类别文本相似度 (TypeSimilarity)

我们认为,不同的会计师事务所会系统地检验某一类会计科目、披露某一类关键审计事项;同时系统地忽视某些会计科目,不沟通某些关键审计事项。这种披露事项类别的偏好,体现了会计师事务所总体的工作规则、思维范式、准则理解以及审计师职业培训结果。因此,我们阅读并记录了样本中公司披露的关键审计事项类别概述文本。事项类别概述文本来自于关键审计事项段的一级子标题,表明了披露事项的种类,审计师会在一级子标题下解释披露该事项的依据以及所做出的特殊审计程序。我们将这些事项类别标题形成一段短文本,在形成公司配对样本后,利用自然语言处理技术(NLP)计算文本相似度,以衡量两家配对公司之间的关键审计事项类别可比性。具体来说,我们主要利用N-Gram距离进行模糊匹配来计算中文文本相似度,可以综合考虑语意的相似性,对相似语意、同义词、缩写、倒序和简写等常见的相似度情况都能较好拟合。文本相似度得分在0和1之间,得分越高意味着两家公司关键审计事项的披露事项类别越接近。

2. 被解释变量: 篇幅(或披露详细程度)差异 (WordListDiff)

上文叙述了针对新式审计报告可能存在的会计师事务所报告格式模板,因此审计师会按照模板中留下的空白撰写文字,因此使关键审计事项展示出与会计师事务所风格特征有关的相应篇幅长度。同时,对于一条关键审计事项使用多大的篇幅去解释披露的理由和特殊的审计处理程序,象征着不同的信息含量,不同的审计努力程度,也在一定意义上体现了审计质量。因此,我们从文构财经文本数据库^①的词频数据库得到了审计报告中关键审计事项段的中文词数。该数据库通过中文财经文本分词的机器学习算法,

剔除了数字及英文字符,得到了关键审计事项段的中文词数。我们在经过配对形成公司对样本后,使用配对两公司每条关键审计事项段平均用词数的自然对数的绝对值差异来衡量配对公司间的篇幅差异(或披露详细程度差异),以检验是否被同一会计师事务所所审计的一对公司相对于被不同会计师事务所所审计的一对公司会披露篇幅规模更接近的关键审计事项段。具体计算方法为:

$$Lnwordlist_{Diff} = \left| \ln\left(\frac{word_a}{list_a}\right) - \ln\left(\frac{word_b}{list_b}\right) \right| \quad (1)$$

3. 被解释变量: 语调指标差异 (ToneDiff)

语调作为一种与内容无关的文本语言特征,主要表达书写者的情感倾向。在目前的研究中,具体体现为作者对于积极词汇和消极词汇的使用。

关键审计事项段象征着审计师在布局审计工作时认定的、对于本期审计工作最为重要的重大风险来源,但被审计师通过职业怀疑识别出的关键审计事项能否顺利得到披露,以及披露的程度如何,这取决于审计师与被审计公司之间的沟通、斡旋乃至妥协。然而,会计师事务所在出具关键审计事项段时,因为要给出披露关键审计事项的依据以及执行的特殊审计程序,即使隐瞒部分消极信息,仍可能暴露出隐含的情感倾向,因此不同的审计师事务所工作范式、工作规则、审计质量形成的不同的审计风格可能带来不同的关键审计事项段语调指标。

本文使用的中文语调特征数据来自文构财经文本数据库。具体构建过程如下:1)以Loughran和McDonald(2011)提供的金融情感英文词汇列表为基础,对词表中的英文词汇翻译后得到种子词集;2)利用同义词词林对词表进行扩充;3)对上述词表进行人工筛选之后,得到最终的情绪词汇列表。在有了财经情绪词汇列表之后,在中文文本分词技术的基础上,便可统计出积极词汇和消极词汇的相应词频,计算得到最终的语调指标。

语调有两种常见的衡量方法,第一种的计算方法为积极词汇和消极词汇的差除以积极词汇与消极词汇之和(谢德仁和林乐,2015);第二种则为积极词汇和消极词汇的差除以文本段总词汇数(Davis等,2012)。我们使用第一种语调指标进行主检验,使用第二种语调指标作为替代变量进行稳健性检验,最终两种语调衡量方法的结果都是稳健并显著的。对于本文形成的公司配对样本,我们对每个公司对计算了其语调指标的绝对值差异。本文所用语调指标和语调差异指标的具体计算方法为:

^① 文构财经文本数据库网站: www.wingodata.com。

$$Tone1 = \frac{POSWords - NEGAWords}{POSWords + NEGAWords} \quad (2)$$

$$Tone2 = \frac{POSWords - NEGAWords}{TOTALWords} \quad (3)$$

$$Tone_{Diff} = |Tone1_a - Tone1_b| \quad (4)$$

4. 被解释变量：可读性指标差异（Readability_Diff）

可读性衡量了文本易读与难读的水平，我们针对整个关键审计事项段文本进行可读性指标分析，来寻找审计报告书写者写作风格的一致性证据；即同年同行业中被同一个事务所审计的一对公司相对于被不同的会计师事务所审计的一对公司会有更接近的可读性水平，体现了审计师事务所的整体思维模式、遣词造句体系对于其员工（个人注册会计师）的影响。

传统的文本可读性度量方式，往往忽略了词汇的前后搭配顺序，或更适用于英文文本。本文使用的可读性指标来自文构财经文本数据库。在朴素贝叶斯的假设下，即假设文本中句子相互独立，并考虑句中词汇的前后搭配顺序，将文本中各个句子生成概率乘积的对数似然的均值，作为此文本的可读性度量。数据库的具体指标采用深度学习算法构建而成，主要构建过程如下：1）利用 Word Embedding 将每个词表示成一个密集的固定长度的实值向量，然后语义相近的词在向量空间上具有相同的向量表示；2）借鉴 Hierarchical Softmax 和 Negative Sampling 的优化思想，可计算得到句子的生成概率；3）再将各个句子生成概率乘积的对数均值作为该文档的可读性度量。可读性用公式表示为：

$$Readability = \frac{1}{N} \sum_{s=1}^N \log P_s \quad (5)$$

其中， P_s 表示句子 s 生成的概率， N 表示构成文本的句子数。其值越高，表示文本中词对搭配顺序在语料中出现的频率越高，文本越容易被理解，文本可读性越高；反之，则表示词对在语料中出现的频率越低，文本越不容易被理解，文本可读性越差。由于可读性指标为负数，我们在配对形成公司对样本后，对可读性指标取绝对值后再取自然对数，再计算两家公司间的绝对值差异，以衡量两公司关键审计事项段的可读性差异水平。具体计算公式为：

$$Readability_{Diff} = |\ln(|Readability_a|) - \ln(|Readability_b|)| \quad (6)$$

（二）样本选择与数据来源

本文选择 2016–2018 年中国所有披露关键审计事项的上市公司作为初始样本，其中 A+H 股上市公司于 2016 年开始披露关键审计事项，A 股上市公司于 2017 年开始披露

关键审计事项。我们剔除了 ST 或 *ST 等经营状况异常的公司，因为这些公司会披露更多与持续经营假设不确定性有关的关键审计事项，语调指标也会显著消极；剔除了审计意见非标准的公司；剔除了数据缺失的公司。由于本文的初始样本最终会形成基于同一年度、同一行业配对的配对样本，因此本文的初始面板数据虽为非平衡面板，但并不会影响结果。我们最后保留了 3435 家独立公司披露的总计 7135 份审计报告，其中共披露了 14953 条关键审计事项。本文所用的审计报告为手动收集，来源于巨潮资讯网，系证监会指定的上市公司信息披露网站；变量数据来自于国泰安数据库（CSMAR）；财经文本分析技术、机器学习方法以及文本特征数据均来自文构财经文本数据库。本文所用的所有连续变量均进行了上下 1% 的缩尾（Winsorize）处理，以排除异常值带来的影响。

本文的主检验样本为差异及相似度样本，在该样本中的每个观测都是同一行业、同一年份的一对公司（Francis 等，2014）。每个公司都将和与其属于同一行业一年度的其他公司均两两配对一次。例如，如果某一行业一年度有 4 家公司 A、B、C、D，那么最后将会形成 6 个公司对，分别为 A-B、A-C、A-D、B-C、B-D、C-D。本文配对依据的行业为证监会 2012 年新版行业分类，其中对于制造业 C 保留两位细分行业编码，对于其他行业均保留一位行业编码。本文最后形成了 1767801 个公司配对样本观测。

（三）研究模型和其他变量

借鉴 Francis 等（2014）使用的研究方法，参照本文提出的假设，共建立如下的四个模型：

$$TypeSimilarity_{i,j,t} = \alpha_0 + \alpha_1 SameAuditor_{i,j,t} + \alpha_2 Controls_{i,j,t} + YearFixedEffect + IndustryFixedEffect + \varepsilon_{i,j,t} \quad (7)$$

$$WordListDiff_{i,j,t} = \beta_0 + \beta_1 SameAuditor_{i,j,t} + \beta_2 Controls_{i,j,t} + YearFixedEffect + IndustryFixedEffect + \varepsilon_{i,j,t} \quad (8)$$

$$ToneDiff_{i,j,t} = \delta_0 + \delta_1 SameAuditor_{i,j,t} + \delta_2 Controls_{i,j,t} + YearFixedEffect + IndustryFixedEffect + \varepsilon_{i,j,t} \quad (9)$$

$$ReadabilityDiff_{i,j,t} = \theta_0 + \theta_1 SameAuditor_{i,j,t} + \theta_2 Controls_{i,j,t} + YearFixedEffect + IndustryFixedEffect + \varepsilon_{i,j,t} \quad (10)$$

其中，SameAuditor 是一个虚拟变量，在这对公司被同一家会计师事务所审计时取 1，否则取 0。模型（7）用来检验假设 H1，我们预测系数 α_1 显著为正，则假设 H1 得到验证；在同一年度同一行业内，同一家会计师事务所审

计的一对公司相对于被不同的会计师事务所审计的一对公司会披露事项类别更一致的关键审计事项。模型 (8)、(9)、(10) 用来检验假设 H2, 我们预测系数 β_1 、 δ_1 、 θ_1 均显著为负, 则假设 H2 得到验证: 在同一年度同一行业内, 同一家会计师事务所审计的一对公司相对于被不同的会计师事务所审计的一对公司会披露文本特征更接近的关键审计事项段, 具体体现为披露详细程度更接近、语调指标差异更小、可读性指标差异更小。

为了验证假设 H3, 我们生成了子样本, 其中仍为配对样本, 但只包括被同一家四大会计师事务所审计的公司对和被同一家非四大会计师事务所审计的公司对。我们将自变量换为 SameBig4 进行检验, 是衡量是否被同一家四大会计师事务所审计的虚拟变量; 之后重新检验上述四个模型, 来验证四大会计师事务所相对于其他会计师事务所

的风格强弱。

我们主要控制了一对公司的审计费用、公司规模、应计项目规模、操纵性应计项目、资产负债率、资产收益率和一系列内部控制水平指标, 来试图在缺乏相关文献参考情况下捕捉可能影响两公司披露关键审计事项的其他差异或冲击。所用控制变量缩写为 Controls。和 Francis 等 (2014) 的控制模式一致, 对于上述所用控制变量, 我们均引入了两公司之间变量的绝对值差异和最小值 (Diff/Min), 来针对公司配对样本的性质同时控制变量的绝对水平和相对差异。我们还控制了行业和年份的固定效应, 以捕捉潜在的经济系统性冲击, 并在公司水平上进行了聚类 (Clustering) 处理 (Petersen, 2009)。我们也尝试将聚类水平延伸到独立的公司对上, 结果依然维持稳健。本文所有变量如表 (1) 所示:

表 1 变量定义

变量符号	变量标记	变量说明
TypeSimilarity	关键审计事项类别相似度	使用 N-Gram 方法计算的两公司关键审计事项披露类别文本模糊匹配相似度
WordListDiff	关键审计事项披露详细程度差异	两公司关键审计事项段每条事项使用的中文词数的自然对数的绝对值差异
ToneDiff	关键审计事项语调差异	两公司净语调指标的绝对值差异
ReadabilityDiff	关键审计事项可读性差异	两公司可读性指标绝对值的自然对数的绝对值差异
SameAuditor	是否被同一家会计师事务所所审计	虚拟变量, =1 时两公司被同一家会计师事务所所审计, =0 时则反之
SameBig4	是否被同一家四大会计师事务所所审计	虚拟变量, =1 时两公司被同一家四大会计师事务所所审计, =0 时两公司被同一家非四大会计师事务所所审计
Af_min	两公司审计费用自然对数最小值	$= \min \{ \ln(AF_a), \ln(AF_b) \}$
Af_Diff	两公司审计费用自然对数绝对值差异	$= \ln(AF_a) - \ln(AF_b) $
Age_min	两公司上市年限自然对数最小值	$= \min \{ \ln(Age_a), \ln(Age_b) \}$
Age_Diff	两公司上市年限自然对数绝对值差异	$= \ln(Age_a) - \ln(Age_b) $
Accurals_min	两公司应计项目规模最小值	$= \min \left\{ \left(\frac{Acc_a}{Asset_{a,t-1}} \right), \left(\frac{Acc_b}{Asset_{b,t-1}} \right) \right\}$
Accurals_Diff	两公司应计项目规模绝对值差异	$= \left \left(\frac{Acc_a}{Asset_{a,t-1}} \right) - \left(\frac{Acc_b}{Asset_{b,t-1}} \right) \right $
AbnAcc_min	两公司操纵性应计规模最小值	使用 Kothari 等 (2005) 的业绩修正琼斯模型计算操纵性应计项目 $= \min (AbnAcc_a, AbnAcc_b)$

续表

变量符号	变量标记	变量说明
AbnAcc_Diff	两公司操纵性应计规模绝对值差异	$= AbnAcc_a - AbnAcc_b $
Size_min	两公司规模最小值	规模水平以总资产的自然对数衡量 $= \min \{ \ln(Asset_a), \ln(Asset_b) \}$
Size_Diff	两公司规模绝对值差异	$= \ln(Asset_a) - \ln(Asset_b) $
PPE_min	两公司固定资产占比最小值	以固定资产净值占总资产比例计量 $= \min \left\{ \left(\frac{PPE_a}{Asset_a} \right), \left(\frac{PPE_b}{Asset_b} \right) \right\}$
PPE_Diff	两公司固定资产占比绝对值差异	$= \left \frac{PPE_a}{Asset_a} - \frac{PPE_b}{Asset_b} \right $
Lev_min	两公司资产负债率最小值	资产负债率 = 总负债 / 总资产 $= \min \left\{ \left(\frac{Lib_a}{Asset_a} \right), \left(\frac{Lib_b}{Asset_b} \right) \right\}$
Lev_Diff	两公司资产负债率绝对值差异	$= \left \frac{Lib_a}{Asset_a} - \frac{Lib_b}{Asset_b} \right $
Mtb_min	两公司市场价值比率最小值	$= \min (Mtb_a, Mtb_b)$
Mtb_Diff	两公司市场价值比率绝对值差异	$= Mtb_a - Mtb_b $
Top1_min	两公司第一大股东持股比例最小值	$= \min (Top1_a, Top1_b)$
Top1_Diff	两公司第一大股东持股比例绝对值差异	$= Top1_a - Top1_b $
Boardsize_min	两公司董事会人数的自然对数最小值	$= \min \{ \ln(Board_a), \ln(Board_b) \}$
Boardsize_Diff	两公司董事会人数的自然对数绝对值差异	$= \ln(Board_a) - \ln(Board_b) $
Inde_min	两公司独立董事占比最小值	$= \min (Inde\%_a, Inde\%_b)$
Inde_Diff	两公司独立董事占比绝对值差异	$= Inde\%_a - Inde\%_b $
Roa_min	两公司资产收益率最小值	$= \min (Roa_a, Roa_b)$
Roa_Diff	两公司资产收益率绝对值差异	$= Roa_a - Roa_b $
Ana_min	两公司分析师跟踪水平最小值	$= \min \{ \ln(Ana_a + 1), \ln(Ana_b + 1) \}$
Ana_Diff	两公司分析师跟踪水平绝对值差异	$= \ln(Ana_a + 1) - \ln(Ana_b + 1) $
Return_min	两公司考虑红利再投资的年个股回报率最小值	$= \min (Return_a, Return_b)$

续表

变量符号	变量标记	变量说明
Return_Diff	两公司考虑红利再投资的年个股回报率 的绝对值差异	$= Return_a - Return_b $
Same_soe	两公司股权性质是否为国有的虚拟 变量	= 1 表示两公司均为国企或均不为国企；= 0 表示一家为国企 而另一家不为国企
Same_Div	两公司股利政策是否一致的虚拟 变量	= 1 表示两公司都分配股利或都未分配股利；= 0 表示一家分 配股利而另一家未分配股利
Same_Dua	两公司两职分离情况是否一致的虚 拟变量	= 1 表示两公司均两职兼任或均不是两职兼任；= 0 表示一家 为两职兼任而另一家不是两职兼任

五、实证结果和分析

(一) 描述性统计结果与分析

表 2 描述性统计

变量名	N	Mean	SD	Median	P25	P75	Min	Max
TypeSimilarity	1767801	0.388	0.211	0.387	0.226	0.529	0.000	1.000
Wordlistdiff	1767801	0.270	0.218	0.219	0.103	0.381	0.004	1.093
ToneDiff	1767801	0.190	0.150	0.155	0.072	0.271	0.002	0.681
ReadabilityDiff	1767801	0.305	0.243	0.244	0.109	0.453	0.004	1.032
SameAuditor	1767801	0.079	0.270	0.000	0.000	0.000	0.000	1.000
Af_Diff	1767801	0.676	0.561	0.539	0.251	0.950	0.000	2.705
Af_min	1767801	13.522	0.413	13.487	13.218	13.768	12.612	14.724
Age_Diff	1767801	1.084	0.859	0.916	0.357	1.705	0.000	3.178
Age_min	1767801	1.368	0.952	1.609	0.693	2.079	0.000	3.045
Accurals_Diff	1767801	0.088	0.087	0.062	0.029	0.117	0.001	0.466
Accurals_min	1767801	-0.046	0.077	-0.033	-0.069	-0.001	-0.391	0.089
AbnAcc_Diff	1767801	0.115	0.110	0.080	0.035	0.159	0.001	0.539
AbnAcc_min	1767801	-0.057	0.098	-0.039	-0.091	0.001	-0.401	0.132
Size_Diff	1767801	1.359	1.071	1.114	0.521	1.937	0.020	4.857
Size_min	1767801	21.501	0.884	21.438	20.840	22.063	19.614	24.098
PPE_Diff	1767801	0.137	0.115	0.107	0.048	0.197	0.002	0.514
PPE_min	1767801	0.136	0.087	0.121	0.072	0.184	0.002	0.414
Lev_Diff	1767801	0.216	0.161	0.184	0.087	0.314	0.003	0.715
Lev_min	1767801	0.306	0.150	0.289	0.186	0.410	0.056	0.698

续表

变量名	N	Mean	SD	Median	P25	P75	Min	Max
Mtb_Diff	1767801	0.896	1.200	0.499	0.212	1.053	0.008	7.231
Mtb_min	1767801	1.368	0.409	1.270	1.080	1.541	0.827	3.008
Top1_Diff	1767801	0.156	0.119	0.130	0.061	0.225	0.002	0.517
Top1_min	1767801	0.248	0.097	0.236	0.174	0.303	0.074	0.523
Samesoe	1767801	0.622	0.485	1.000	0.000	1.000	0.000	1.000
Samediv	1767801	0.613	0.487	1.000	0.000	1.000	0.000	1.000
Samedua	1767801	0.573	0.495	1.000	0.000	1.000	0.000	1.000
Boardsize_Diff	1767801	0.203	0.189	0.201	0.000	0.251	0.000	0.788
Boardsize_min	1767801	2.005	0.184	1.946	1.946	2.197	1.609	2.303
Inde_Diff	1767801	0.056	0.055	0.042	0.000	0.095	0.000	0.267
Inde_min	1767801	0.349	0.030	0.333	0.333	0.353	0.333	0.429
Roa_Diff	1767801	0.074	0.103	0.042	0.018	0.084	0.001	0.672
Roa_min	1767801	-0.006	0.105	0.017	0.003	0.036	-0.627	0.105
Ana_Diff	1767801	1.357	1.054	1.099	0.511	2.110	0.000	3.784
Ana_min	1767801	0.664	0.889	0.000	0.000	1.099	0.000	3.178
Return_Diff	1767801	0.266	0.262	0.189	0.087	0.350	0.003	1.404
Return_min	1767801	-0.381	0.160	-0.397	-0.487	-0.296	-0.706	0.160

表2中,针对公司配对样本的描述性统计显示,自变量 SameAuditor 的均值为 0.079,方差为 0.270,这与样本内 30 多家会计师事务所的随机配对结果比较符合。因变量部分:关键审计事项类别相似度 TypeSimilarity 均值为 0.388,方差为 0.211,最小值为 0,最大值为 1,展现了较

大的变异性,较适合后续回归分析;三种文本特征变量 WordListDiff, ToneDiff, ReadabilityDiff 也都具有合适的均值水平、方差水平和足够的极差。其他所有控制变量都显示了合理和符合经济实质的描述性统计指标。

(二) 多元回归结果与分析

表 3

假设 H1、H2 的多元回归结果^②

	(1)	(2)	(3)	(4)
	TypeSimilarity	WordListDiff	Tone_Diff	Readability_Diff
SameAuditor	0.023 *** (9.08)	-0.028 *** (-10.58)	-0.015 *** (-9.50)	-0.099 *** (-30.59)
Constant	0.940 *** (11.40)	0.019 (0.18)	-0.092 (-1.41)	0.339 *** (3.86)

② 限于篇幅,本文未列示控制变量。完整回归结果请联系作者邮箱: tian-gl@xjtu.edu.cn。

续表

	(1)	(2)	(3)	(4)
Control Variables	Control	Control	Control	Control
Ind Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
N	1767801	1767801	1767801	1767801
Adjusted R ²	0.07	0.01	0.01	0.02

***, **, * 分别表示在 0.01, 0.5, 0.1 的水平上双尾显著;

上述模型均使用 OLS 模型回归, 结果在公司层面聚类, 括号内为稳健的 t 统计量。

本文所用变量具有最小值和绝对值差异的特殊属性, 因此不存在严重的多重共线性问题。事实上, 未列出的 VIF 表显示, 绝大多数变量的 VIF 都小于 5, 所有变量的 VIF 都小于 10。

表 3 列示了对于假设 H1、H2 的实证检验结果。(1) 列为对假设 H1 的检验结果, SameAuditor 的系数为 0.023, 在 $p<0.01$ 的水平上正向显著 (双尾)。(1) 列的实证结果验证了假设 H1, 即在同一行业一年度内, 被同一家会计师事务所审计的一对公司, 相对于被不同的会计师事务所审计的一对公司, 会有更大的关键审计事项类别相似度和可比性。(2)、(3)、(4) 列的实证结果均为对 H2 的检验结果, 分别检验了子假设 H2a、H2b、H2c。在三个回归结果中, SameAuditor 的系数分别为 -0.028, -0.015,

-0.099, 均在 $p<0.01$ 的水平上负向显著 (双尾)。(2) (3) (4) 列的实证结果共同验证了假设 H2 的准确性, 即在同一行业一年度内, 被同一家会计师事务所审计的一对公司, 相对于被不同的会计师事务所审计的一对公司, 会有更小的关键审计事项段文本特征差异, 具体体现为: 有更小的每条事项篇幅或披露详细程度差异; 有更小的语调指标差异; 有更小的可读性指标差异。上述所用所有 OLS 回归模型均控制了行业和年份的多元固定效应, 同时进行了公司层面的聚类处理, 结果依然维持稳健。

表 4 对于假设 H3 的实证检验结果

	(1)	(2)	(3)	(4)
	TypeSimilarity	WordListDiff	ToneDiff	ReadabilityDiff
SameBig4	-0.046 *** (-3.75)	0.031 * (1.87)	0.036 *** (4.68)	-0.010 (-0.81)
Constant	0.798 *** (6.41)	0.268 ** (2.04)	0.105 (1.27)	0.372 *** (3.15)
Control Variables	Control	Control	Control	Control
Ind Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
N	139833	139833	139833	139833
Adjusted R ²	0.09	0.01	0.02	0.01

***, **, * 分别表示在 0.01, 0.5, 0.1 的水平上双尾显著;

上述模型均使用 OLS 模型回归, 结果在公司层面聚类, 括号内为稳健的 t 统计量。

表 4 列示了对于假设 H3 的实证结果, 该回归结果来自于一个子样本, 其中仅包含被同一家会计师事务所审计的公司对; 检验变量为 SameBig4, =1 表示两公司被同一

家四大会计师事务所所审计, =0 表示两公司被同一家非四大会计师事务所所审计。该假设的提出基于一种观点: 四大会计师事务所所有更高的国际化水平与更高的审计质量,

并且可能会更早的了解国际审计准则的有关变化,因此可能会做出更激进的披露,表现为更弱的公司对之间的关键审计事项相似度。具体来说,四大会计师事务所可能更加根据被审计公司当期实际的最重大风险点识别出关键审计事项,进行范式化披露的可能性更小,因此我们观测到基于关键审计事项的审计风格弱于其他会计师事务所。表4的列(4)中,SameBig4的系数为负但不显著。(1)列中,SameBig4的系数为-0.046,在 $p<0.01$ 的水平上负向显著(双尾)。(2)列中SameBig4的系数为0.031,仅在 $p<0.1$ 的水平上正向显著(双尾)。(3)列中,SameBig4的系数为0.036,在 $p<0.01$ 的水平上正向显著(双尾)。因此,我们认为,在同一年份行业中,被同一家四大会计师事务所审计的一对公司,相对于被同一家非四大会计师事务所审计的一对公司,会有更弱的关键审计事项类别文本相似度、更大的披露详细程度差异、更大的语调指标差异(但没有可读性差异的证据)。总体上,这表明四大会计师事务所表现出的审计风格或者说关键审计事项段的可比性弱于非四大会计师事务所。考虑到关键审计事项段的过高可比性或相似度可能是会计师事务所集体思维、思维惰性和思维惯性的产物,我们认为这种弱化的基于关键审计事项的审计风格可能在一定程度上是高质量审计和高水平职业怀疑的象征,意味着四大会计师事务所更激进地披露了关键审计事项,沟通了更多与报告使用者决策有用相关的信息。

(三) 会计师事务所固定效应模型

另外一种检验风格效应的方法是使用固定效应模型。在会计师事务所固定效应模型检验中,我们不使用配对样本,而是使用传统的公司水平观测,使用如下的OLS回归模型检验会计师事务所固定效应是否对文本特征指标产生系统性的统计差异:

$$\text{TextVariables}_{i,t} = \alpha_0 + \gamma + \alpha_1 \text{Controls}_{i,t} + \text{YearFixedEffect} + \text{IndFixedEffect} + \varepsilon_{i,t} \quad (11)$$

其中, $\text{TextVariables}_{i,t}$ 是上文所用的三种文本特征指标,即每条关键审计事项披露用词数的自然对数、语调指标和可读性指标绝对值的自然对数。 γ 是会计师事务所的固定效应回归系数集。我们依然控制了一系列公司水平的控制变量,以及行业和年份固定效应,以控制潜在的系统性冲击。此时的固定效应系数可以捕捉每个会计师事务所在披露文本特征时控制公司属性和其他固定效应后的增量效应。我们将在固定效应模型中使用F检验,并预测我们将会拒

绝所有会计师事务所披露文本特征相似的零假设。(Francis等,2014)

在未列出的回归结果中,以WordList、Tone、Readability为因变量的回归,检验会计师事务所一系列系数相等的F统计量分别为22.04,3.75,85.75。这三个F统计量均在 $p<0.01$ 的水平上显著,否定了每个会计师事务所披露关键审计事项文本特征完全一致的零假设,证明会计师事务所是控制其他因素和冲击后,使关键审计事项段文本特征出现统计上系统性差异的一个来源。

固定效应模型的局限为F检验只能否定所有的会计师事务所关键审计事项段文本特征是完全一致的,换言之,只能证明对于每种文本特征变量,至少有一家会计师事务所不同于其他会计师事务所。因此,我们认为会计师事务所固定效应的检验仅作为参考和提升稳健性,主要结果还应参照基于配对样本的OLS回归。

六、稳健性检验

(一) 更换关键审计事项类别相似度的计量方法

本文对于假设H1的检验使用了关键审计事项类别描述文本相似度来代表两公司披露关键审计事项在科目和类别上的可比性和相似度。然而,本文所用的基于N-gram的文本相似度可能并不完美,因为这种文本相似度算法不考虑其背后的经济和会计意义。因此,我们换用如下的方法衡量关键审计事项类别相似度,并重新检验假设H1。

我们基于骆韦等(2019)、柳木华和董秀琴(2018)对于关键审计事项类别的描述性总结文献、现行企业会计准则和会计科目,将所有行业披露的关键审计事项分类为31类。^③

之后,我们通过手动阅读并理解关键审计事项段,人工将样本中公司披露的关键审计事项分类为上述31类,之后对每个公司一年观测值形成基于关键审计事项披露类别的31元向量,由1和0组成,表示该公司在该年披露的关键审计事项和未披露的关键审计事项。在相同的基于同一年、同一行业的配对方法后,我们使用余弦相似度算法计算一对公司关键审计事项类别向量的相似度,以衡量两公司在关键审计事项类别上的可比性。余弦相似度具体的计算方法为:

$$\text{Similarity} = \frac{\sum_{i=1}^{31} A_i * B_i}{\sqrt{\sum_{i=1}^{31} A_i^2} * \sqrt{\sum_{i=1}^{31} B_i^2}} \quad (12)$$

此处所用的余弦相似度算法是基于会计学科目和经济

^③ 具体分类的会计科目限于篇幅未列示,请联系作者邮箱。

实质形成的类别特征向量余弦相似度，可以在不考虑文本特征本身差异，仅仅考虑事项经济及会计实质的情况下衡量关键审计事项段的类别可比性，符合本文假设 H1 对于关键审计事项类别的检验需求。

我们使用新的余弦相似度作为因变量重新检验（7），依然控制了相同的控制变量、固定效应，并且在公司层面聚类处理。

表 5 更换关键审计事项类别相似度算法的回归结果

	(1)
	Similarity
SameAuditor	0.019 *** (5.56)
Constant	1.373 *** (10.78)
Control Variables	Control
Ind Fixed Effect	Yes
Year Fixed Effect	Yes
N	1767801
Adjusted R ²	0.08

表 5 可看出，更换关键审计事项类别相似度算法后，依然维持了在 $p < 0.01$ 水平的双尾正向显著性，为会计师事务所具有基于关键审计事项类别的审计风格提供了补充证据。

（二）更换语调指标计量方法

上文对于语调指标变量的解释中，我们提到了与本文主检验中所用语调指标（谢德仁和林乐，2015）不同的另一种语调计量方法，区别在于分母由积极词汇和消极词汇之和换为总词汇数。我们使用第二种语调指标重新进行配对处理，并重新检验模型（9），使用同样的控制变量、固定效应和聚类处理方法。

表 6 更换语调指标计量方法的回归结果

	(1)
	Tone2_Diff
SameAuditor	-0.001 *** (-8.59)
Constant	0.004 (0.76)

续表

	(1)
	Tone2_Diff
Control Variables	Control
Ind Fixed Effect	Yes
Year Fixed Effect	Yes
N	1767801
Adjusted R ²	0.01

表 6 可看出，在更换第二种净语调衡量方式进行回归后，SameAuditor 变量依然维持了 $p < 0.01$ 水平上的双尾负向显著性。

（三）更换披露详细程度指标计量方法

本文对于假设 H2a 的检验使用了每条关键审计事项平均使用词数的自然对数来衡量关键审计事项披露的详细程度。我们换用另外一种计量方法，即关键审计事项字段总中文词数的自然对数，试图在不考虑事项条数的情况下衡量关键审计事项段的整体篇幅和披露详细程度。在更换因变量后，我们重新估计模型（8），使用同样的控制变量、固定效应和聚类方法。

表 7 更换篇幅指标的回归结果

	(1)
	WordDiff
SameAuditor	-0.040 *** (-10.27)
Constant	0.385 *** (2.62)
Control Variables	Control
Ind Fixed Effect	Yes
Year Fixed Effect	Yes
N	1767801
Adjusted R ²	0.01

表 7 可看出，在更换篇幅指标计量方法进行回归后，SameAuditor 变量依然维持了 $p < 0.01$ 水平上的双尾负向显著性。

（四）子样本检验

我们生成了两个子样本：四大会计师事务所配对子样

本和八大会计师事务所配对子样本。在这两个子样本中，配对公司来源仅为被四大（或国内八大）会计师事务所所审计的公司。因此，配对样本中，一对公司要么被相同的四大（或八大）会计师事务所所审计，要么被不同的四大（八大）会计师事务所所审计。我们利用子样本重新检验模型（7）（8）（9）（10），并且预测在两个子样本中，依然可以维持相同的符号和显著程度，故假设 H1 和 H2 可以

被子样本的实证结果继续支持。这部分也能检验四大会计师事务所的中国分支机构和本土八大会计师事务所是否也表现出与本文主要结果一样的审计风格。

在未列示的子样本描述性统计结果中，SameAuditor 在四大会计师事务所配对子样本和本土八大会计师事务所配对子样本中的均值分别扩大到了 0.275 和 0.164，均值水平明显变大。

表 8 四大会计师事务所配对子样本的实证结果

	(1)	(2)	(3)	(4)
	TypeSimilarity	WordListDiff	ToneDiff	ReadabilityDiff
SameAuditor	0.038 *** (6.11)	-0.086 *** (-6.52)	-0.038 *** (-4.04)	-0.105 *** (-9.63)
Constant	0.765 *** (3.44)	0.443 (0.98)	0.813 *** (3.48)	0.076 (0.32)
Control Variables	Control	Control	Control	Control
Ind Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
N	5219	5219	5219	5219
Adjust R ²	0.32	0.09	0.06	0.10

表 9 本土八大会计师事务所配对子样本的实证结果

	(1)	(2)	(3)	(4)
	TypeSimilarity	WordListDiff	ToneDiff	ReadabilityDiff
SameAuditor	0.017 *** (7.29)	-0.022 *** (-8.89)	-0.005 *** (-3.50)	-0.126 *** (-32.34)
Constant	0.827 *** (7.94)	0.288 ** (2.46)	0.179 ** (2.42)	0.436 *** (3.70)
Control Variables	Control	Control	Control	Control
Ind Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
N	793955	793955	793955	793955
Adjusted R ²	0.08	0.01	0.02	0.04

由上文的表 8、表 9 可以看出，在使用四大会计师事务所配对子样本和八大会计师事务所配对子样本重新检验 4 个 OLS 回归模型时，SameAuditor 变量在两个样本中对于四个因变量分别在 $p < 0.01$ 的水平上双尾正、负、负、负显著，

与本文对假设 H1、H2 的主检验结果维持一致。这证明会计师事务所基于关键审计事项段的类别和文本特征审计风格在更小的配对样本维度上依然存在，这也意味着基于关键审计事项的审计风格客观、普遍存在于各级别的会计师事务所中。

七、与实务工作者及专业人士的探讨

目前为止,我们通过实证方法将审计风格拓展到关键审计事项段这一非财务型文本化信息层面,发现中国会计师事务所具有基于关键审计事项段的审计风格。我们对此的解释是:基于关键审计事项段的审计风格是会计师事务所的集体思维模式对于个人审计师识别关键审计事项以及设计特殊审计程序的影响。具体来说,我们推测在新审计报告执行前,审计师事务所中会有专家建立关键审计事项段模板,并对注册会计师进行继续教育与集中培训,使个人注册会计师在执行审计工作、识别关键审计事项时具有对某些其在继续教育和集中培训中较为熟悉的关键审计事项的偏好,从而可能会系统地识别某些关键审计事项,并存在没有识别当期审计中最为重大的风险点作为关键审计事项的可能性;“四大”会计师事务所被认为具有更高水平的职业怀疑和审计质量,且它们接触国际审计准则变更较早,故受到会计师事务所的思维模式影响较小,这是我们观测到“四大”会计师事务所基于关键审计事项的审计风格较弱的原因。

为了证实我们推演的逻辑,我们与十余位审计一线工作者和实务专业人士进行了访谈调研,其中包括国际“四大”会计师事务所、国内“八大”会计师事务所以及地方大事务所的资深注册会计师、合伙人等。不同层级的实务工作者对于关键审计事项段审计风格的看法类似,而他们从业的会计师事务所对关键审计事项段及新式审计报告发布和实施时做出的回应不尽相同。所有受访的实务工作者均标识其会计师事务所新审计报告准则发布前对其注册会计师进行了继续教育和集中培训,但“四大”会计师事务所的会计师接触这一概念普遍较早,进行的培训次数也较多,这可能跟“四大”的国际化水平较高有关系。在其进行的培训中,对于出现概率较高的关键审计事项进行的培训时长最多,也针对这些出现概率较高的关键审计事项设计了较多案例指引,受访者表示其对这些关键审计事项最为熟悉,在被审计公司存在该等重大风险时偏好去识别其为关键审计事项,包括“收入确认”“应收账款减值准备”“存货跌价准备”“商誉”等。这些受访者表示最为熟悉并建立为案例的关键审计事项也是我们在整理数据中发现出现概率最高的关键审计事项,均在所有关键审计事项出现频率中排名前五。受访的实务工作者表示其所在的会计师事务所都对新式审计报告和其中的关键审计事项段建立了模板性文件。多位来自于“四大”会计师事务所、“八大”会计师事务所的合伙人表示,其任职的事务所对于每个细分行业建立了关键审计事项数据库,在他们执行

审计程序和识别关键审计事项时会优先在被审计公司所在行业数据库中寻找相符的事项作为关键审计事项,并按照数据库相应地执行额外的审计程序,除非其中的某些关键审计事项对于被审计公司并不适用(如没有商誉)。由于行业数据库的存在性,受访者普遍认同基于关键审计事项的审计风格的客观存在。受访者还指出,在不存在重大变动和较高审计风险时,他们不会优先选择在数据库之外识别出新的、独特的关键审计事项;在连续数年审计一家上市公司时,如不存在经营风险重大变动,他们会优先识别去年披露过的关键审计事项并继续使用。

关键审计事项行业数据库的存在性很好地符合了本文得出的结论和推演的逻辑,即个人注册会计师会从其所任职的会计师事务所形成的集体智慧(自行建立的行业数据库)中寻求指引和帮助,并且会执行数据库中出现额外审计程序。各个事务所分别建立的关键审计事项段行业数据库,包括该行业中较为常见的关键审计事项和相应的特殊审计程序,已经形成了一种知识资产,因此会计师事务所的审计师在识别关键审计事项、执行审计程序时必定带有其任职的会计师事务所的思维模式、工作规则和工作规范,故基于关键审计事项段的审计风格普遍地存在于各级别的会计师事务所中。然而,我们仍不知道由各个事务所建立的关键审计事项行业数据库是否足够涵盖被审计公司的所有风险点,是否会具有“管中窥豹”的效应,“集体智慧”和“羊群效应”究竟谁占主导。因此,在注册会计师不会优先选择由自己的职业怀疑来识别特异的关键审计事项时,我们不知道目前存在的基于关键审计事项的审计风格是积极的还是消极的,以及是否对于审计质量有损害作用。我们可以确认的是,关键审计事项段的信息沟通作用可以很好地向资本市场传递额外的信息,并且使信息使用者可以一瞥注册会计师和其背后的会计师事务所的思考过程和审计结论的得出过程,从而更好理解审计工作的流程。关键审计事项本身对于研究者也具有继续挖掘的价值。

八、结论

审计风格体现了作为会计信息基本质量特征之一的可比性。之前的研究发现会计师事务所的风格会为审计后的财务报告带来更大的应计结构可比性。我们继续拓展了这一系列的研究,采用在国外和国内都尚属新鲜事物的关键审计事项段的披露事项类别和披露文本特征的可比性和相似度来衡量审计风格,并验证了其存在,具有一定的边际贡献及创新性。我们的结论也通过与实务工作者进行访谈得到了额外的证据。具体来说,我们基于2016-2018年所有披露新版审计报告的中国上市公司披露的关键审计事项,

选取关键审计事项披露类别和关键审计事项段文本特征两个维度进行配对可比性检验。研究发现,在同一年份同一行业中,被同一家会计师事务所审计的一对公司相对于被不同的会计师事务所审计的一对公司会披露事项类别更接近、事项披露详细程度差异更小、事项语调指标差异更小、可读性指标差异更小的关键审计事项段,因此我国会计师事务所具有基于关键审计事项部分的审计风格。此外,在同一年份同一行业中,被同一家四大会计师事务所审计的一对公司相对于被同一家非四大会计师事务所审计的一对公司会表现出相似度和可比性更差的关键审计事项信息,因此四大会计师事务所基于关键审计事项的审计风格比非四大会计师事务所更弱。会计师事务所固定效应模型表明事务所的固定效应是产生关键审计事项文本特征差异的一个具有统计学意义的因素。研究对于一系列更换变量和样本范围的稳健性检验均可以维持显著性。

会计师事务所具有专业的准则研究人员,会对准则变更进行提前的深入研究,并形成工作范式、工作规则和整体的思维模式提供给个人注册会计师。此外,会计师事务所具有将自身与其他竞争者区分的差异化动机,因此针对关键审计事项执行的特殊审计程序也具有差异化特征。因此,在审计师识别和披露关键审计事项的过程中,会反映出其供职的会计师事务所对准则的解读、培训与继续教育的结果,在披露关键审计事项类别时表现为会按照事务所的思维模式和工作范式系统地偏爱识别某些关键审计事项,并系统地忽视其他类的关键审计事项;在书写关键审计事项段文本时可能会套用事务所下发的指引性模板,并且会偏爱使用事务所熟悉的某些专业用词和遣词造句风格。因此,我们认为这些因素都是我们观测到会计师事务所具有基于关键审计事项段类别和文本特征的双重风格的原因。通过与实务工作者的探讨,我们认识到一些会计师事务所已经形成了关键审计事项行业数据库这一知识资产,进一步强化了事务所披露关键审计事项的同质化,基于关键审计事项的审计风格得到了普遍验证。

四大会计师事务所的审计风格存在但相对于非四大会计师事务所更弱,我们认为这是一种激进披露的象征,可能意味着更高的职业怀疑水平和审计质量,因此披露了更多特定于被审计公司审计风险的非同质化信息。

我们无法预测基于关键审计事项的审计风格是好或坏,但可比性不等同于一致性,应当使相似的事物表现出相似,而使不相似的事物表现出鲜明的差异。然而,披露关键审计事项的高度同质化可能意味着会计师事务所的个人注册会计师没有充分表现出足够的职业怀疑水平,而是可能使

事务所工作范式和工作规则产生的思维惯性和思维惰性凌驾于职业怀疑之上,使关键审计事项不能充分反映企业当期审计工作最大的审计风险来源和最重要的审计工作侧重,最终无法起到向报告使用者披露决策有用信息的目的,而是流于形式。

结合《新证券法》改革要求提升审计师责任和目前对于上市公司信息披露的强化监管,我们认为监管者应当考虑和各地注册会计师协会合作,持续追踪新版审计报告的信息含量,积极研究国际审计准则的变化和调整,促使审计师破除格式的桎梏,积极、激进地披露关键审计事项,以提升新版审计报告的决策有用性,最终增加我国资本市场的有效性。

主要参考文献

- 岑敏儿,潘家恒. 2019. 关键审计事项准则执行效果分析——以深交所主板 A 股公司 2017 年年报为例. 北方经贸, 3: 96~97
- 胡楠,薛付婧,王昊楠. 2021. 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. 管理世界, 5: 139~156
- 李延喜,赛骞,孙文章. 2019. 在审计报告中沟通关键审计事项是否提高了盈余质量? 中国软科学, 3: 120~135
- 柳木华,董秀琴. 2018. 关键审计事项的披露特点——基于 2017 年上市公司新式审计报告的统计分析. 证券市场导报, 11: 12~19
- 路军,张金丹. 2018. 审计报告中关键审计事项披露的初步研究——来自 A+H 股上市公司的证据. 会计研究, 2: 83~89
- 骆韦,吴喆,韩维芳. 2019. 关键审计事项的披露、确定依据及审计应对——以 2016 年 A+H 股公司为例. 财会月刊, 10: 95~103
- 孟庆斌,杨俊华,鲁冰. 2017. 管理层讨论与分析披露的信息含量与股价崩盘风险——基于文本向量化方法的研究. 中国工业经济, 12: 132~150
- 王雄元,高曦,何捷. 2018. 年报风险信息披露与审计费用——基于文本余弦相似度视角. 审计研究, 5: 98~104
- 王艳艳,许锐,王成龙,于李胜. 2018. 关键审计事项段能够提高审计报告的沟通价值吗? 会计研究, 6: 86~93
- 吴秋生,独正元. 2018. A+H 股公司关键审计事项准则执行效果分析. 会计之友, 12: 86~90
- 谢德仁,林乐. 2015. 管理层语调能预示公司未来业绩吗? ——基于我国上市公司年度业绩说明会的文本分析. 会计研究, 2: 20~27
- 叶妮. 2018. 关于关键审计事项披露与会计信息质量的

分析——基于最新审计报告的经验证据. 审计与理财, 12: 12~16

Backof, A. G. 2015. The Impact of Audit Evidence Documentation on Jurors' Negligence Verdicts and Damage Awards. *The Accounting Review*, 90 (6): 2217~2204

Becker, C. L., M. L. Defond, J. Jambalvo, K. R. Subramanyam. 1998. The Effect of Audit Quality on Earnings Management. *Contemporary Accounting Research*, 15 (1): 1~24

Bertrand, M., A. Schoar. 2003. Managing with Style: The Effect of Managers on Firm Policies. Working Paper, Massachusetts Institute of Technology (MIT) & Sloan School of Management, 118 (4): 1169~1208

Davis, A. K., W. Ge, D. Matsumoto, J. L. Zhang. 2015. The Effect of Manager-specific Optimism on the Tone of Earnings Conference Calls. *Review of Accounting Studies*, 20 (2): 639~673

Francis, J. R., E. L. Maydew, H. C. Sparks. 2011. The Role of Big 6 Auditors in the Credible Reporting of Accruals. *Auditing A Journal of Practice & Theory*, 18 (2): 17~34

Francis, J. R., M. L. Pinnuck, O. Watanabe. 2014. Auditor Style and Financial Statement Comparability. *Accounting Review*, 89 (2): 605~633

Ge, W., M. Dawn, J. L. Zhang. 2011. Do CFOs Have Style? An Empirical Investigation of the Effect of Individual CFOs on Accounting Practices. *Contemporary Accounting Research*, 28 (4): 1141~1179

Kothari, S. P., A. J. Leone, C. E. Wasley. 2005. Performance Matched Discretionary Accrual Measures. *Journal*

of Accounting and Economics, 39 (1): 163~197

Kothari, S. P., K. Ramanna, D. J. Skinner. 2010. Implications for GAAP From an Analysis of Positive Research in Accounting. *Journal of Accounting & Economics*, 50 (2-3): 246~286

Li, F. 2008. Annual Report Readability, Current Earnings, and Earnings Persistence. *Journal of Accounting and Economics*, 45 (2-3): 221~247

Li, L., B. Qi, J. Z. 2021. The Effect of Engagement Auditors on Financial Statement Comparability. *Auditing: A Journal of Practice and Theory* (Forthcoming)

Loughran, T., B. McDonald. 2011. When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance*, 66 (1): 35~65

Petersen, M. A. 2009. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *The Review of Financial Studies*, 22 (1): 435~480

Pinto, I., A. I. Morais. 2019. What Matters in Disclosures of Key Audit Matters: Evidence from Europe. *Journal of International Financial Management & Accounting*, 30 (2): 145~162

Reid, L. C., J. V. Carcello, C. Li, T. L. Neal. 2015. Are Auditor and Audit Committee Report Changes Useful to Investors? Evidence from the United Kingdom. Working Paper, University of Tennessee

Sirois, L. P., B. Jean, P. Bera. 2018. The Informational Value of Key Audit Matters in the Auditor's Report: Evidence from an Eye-Tracking Study. *Accounting Horizons*, 32 (2): 141~162

Do Audit Firms Have Independent Auditor Style Based on Key Audit Matters? Empirical Evidence from the New Version of Audit Report Disclosed by Chinese Listed Companies

Tian Gaoliang et al.

Abstract: This article manually collected the content and text information of the key audit matter part in the 2016-2018 new version audit reports disclosure by the Chinese listed companies, and calculated their multiple text feature indicators. In the case of establishing a firm-pair sample, multiple regression methods were used to analyze the similarity and comparability of key audit matters in four dimensions, investigating whether China's auditors have an independent auditor style based on key audit matters when they are mandatory to disclose key audit matters in the new version audit reports. The results show that under the premise of controlling other variables that may affect the company's disclosure of key audit matters, companies audited by the same auditor have greater similarities in disclosure KAM subjects, disclosure KAM part length, KAM part tone indicators and KAM part readability indicators, so there is an independent auditor style in the perspective of key audit matters.

This study is helpful for scholars and practitioners to further understand the mechanism and significance of key audit matters transmitting information to the capital market, in order to facilitate auditors to disclose more useful information for decision-making in the KAM part and improve the effectiveness of capital markets.

Key Words: Auditor Style; Key Audit Matters (KAMs); New Audit Report; Text Analysis