

招股说明书文本语气对新股市场表现的影响

——以创业板为研究对象

何智, 张飞, 周孝华(博士生导师)

【摘要】 选取2009~2012年346家创业板上市公司作为IPO样本,构建适用于中文语境的金融语气词库,应用Java算法模型作为文本分析工具,度量招股说明书文本中的不确定语气和负面语气,并实证检验不确定语气和负面语气对IPO抑价、新股波动性和长期收益的影响。研究表明:招股说明书中含有更多的不确定和负面语气会增加信息不对称程度,使投资者要求更多的风险补偿,也使承销商进一步压低IPO发行价格,进而导致更高的IPO抑价程度;不确定和负面语气会导致更多的投资者意见分歧,并使得新股产生更大的波动性和更加弱势的长期表现。该结果验证了招股说明书不确定和负面语气可以作为IPO事前不确定性的代理变量。

【关键词】 招股说明书; 文本语气; 新股市场表现; 不确定性

【中图分类号】 F832.5

【文献标识码】 A

【文章编号】 1004-0994(2019)08-0029-10

一、引言

招股说明书是经证监会批准、向公众介绍公司详细情况和发行事宜,并指导公众购买股票的重要文件。由于上市公司和投资者之间存在着信息不对称,上市公司通过招股说明书将其私人信息传递给缺乏信息的投资者^[1]。招股说明书第一次详细地向投资者披露了公司的经营情况、商业模式和财务信息等内容,其中包含了与公司价值相关的大部分信息。招股说明书披露的风险信息和价值信息会对IPO定价产生十分重要的影响^[2]。招股说明书中披露的信息不仅包括财务数据、经营数据和行业数据等量化信息,还包括描述公司前景、战略规划和竞争地位的文本信息。由于文本内容包含的软信息通常难以量化,导致学者们更偏重于对财务数据信息进行研究,却较少关注文本信息对IPO估值的影响。然而,许多在创业板上市的高科技公司通常没有相对长期的历史财务数据,这类公司的年龄很短,且处于

快速成长阶段,仅仅依靠有限的历史财务信息很难进行合理的IPO估值。与此同时,由于公司管理层比投资者更为深刻地理解公司的发展前景、盈利能力和投资机会等情况,投资者对公司的认知在很大程度上需要依赖管理层的描述。因此,充分挖掘招股说明书文本中隐含的价值信息对提升IPO估值的准确性非常必要。

文本分析法是一种通过挖掘文本中的特征词,并将其从定性表述转换成数据形式的量化分析方法,是学者们在企业环境信息披露领域和金融市场领域测度信息质量的重要方式^[3]。语气分析则是在文本分析的基础上,进一步深入分析文本的感情色彩,挖掘文本中潜在的倾向性和判断性信息。国外学者对金融文本分析的研究已经涉及投资者获取信息的各种渠道,包括新闻媒体、财务报告、招股说明书、分析师报告、管理层会议和社交网络等^[4,5],Loughran、McDonald^[6]对文本分析在财务和金融领域的应用进

【基金项目】 国家自然科学基金资助重点项目(项目编号:71232004);重庆市社会科学基金博士项目(项目编号:2017BS51)

进行了回顾总结和展望。

目前,国内有关金融领域文本分析的研究已经涉及资产定价方面,对IPO信息环境的文本分析主要集中在新闻媒体^[7]和网络信息^[8]等方面。但是,国内的金融文本研究与国外主流研究相比,无论在研究的广度还是深度方面都还有较大差距,对文本内容的挖掘和文本语气的分析也不够深入和细致。此外,将招股说明书文本分析应用到IPO定价领域的文献更是非常有限,仅有姚颐和赵梅^[2]、郝项超和苏之翔^[9]等从信息披露角度研究了招股说明书文本信息对IPO市场表现的影响。因此,本文试图为IPO定价研究提供新的思路,通过梳理招股说明书文本中包含的不确定词语、负面词语和正面词语的比例,对招股说明书的语气和措辞进行量化分析,并进一步研究招股说明书信息与IPO抑价、新股波动性和长期收益的关系。

本文通过文献综述和理论分析提出招股说明书语气与公司估值密切相关的研究假设,采用文本分析方法构建适用于中文语境的金融文本语气词库,测量了不确定语气和负面语气指数,并以2009~2012年在创业板上市的346家IPO公司作为样本进行实证研究,结果发现招股说明书含有更多的不确定和负面语言会导致更高的IPO抑价程度以及新股更大的波动性、更加弱勢的长期表现。本文可能存在的贡献体现在以下方面:第一,有别于现有文献以媒体和互联网等文本信息为研究对象,本文将中文招股说明书的文本分析应用到IPO定价领域,研究发现招股说明书语气能够较好地解释与预测IPO抑价和长期表现,为今后学者进行招股说明书信息的深入挖掘提供了借鉴和参考。第二,构建了适用于中文语境的金融文本语气词库,通过度量招股说明书不确定语气指数和负面语气指数来考察IPO估值中的不确定性,这种量化分析方法对传统的财务数据分析和价值的定性估计是一种很好的补充。第三,对招股说明书中隐含的不确定性信息进行深入挖掘,分析了不确定信息对发行人、承销商和投资者三方行为的影响,并结合实证研究,揭示出不确定语气和负面语气导致IPO高抑价、新股高波动性和长期弱勢表现的内在机理。

二、文献回顾与研究假设

国内外学者对IPO抑价的研究主要围绕信息不对称理论展开,在该理论框架下又派生出委托代理理论、信息生产假说和赢家诅咒假说等理论。Rock^[1]

最早发现IPO抑价程度与公司内在价值的不确定性有关。由于不确定性难以直接衡量,随后的学者们通常选取与公司特征相关的指标来度量不确定性^[10]。Carter等^[11]认为有高声誉承销商和风险投资背景支持的公司可以在IPO过程中传递确定性更高的信息,降低信息不对称程度。Doran^[12]发现一些低质量发行人会发送较多的模糊信号来干扰信号接收者,使其难以区分发行人的优劣程度。Izhakian、Yermack^[13]将模糊信号与资产收益的不确定性联系起来,指出在特定机制下资产卖方可以通过发送模棱两可的信号来增加收益。因此,低质量的发行人有动机在信息披露中发送较多的模糊信号。姚颐、赵梅^[2]指出,在我国的IPO环境中,对风险信息的披露虽然是强制性要求,但本质上属于选择性披露。Guo等^[14]指出,过度风险披露会带来新股发行的“价格压力效应”,并降低市场流动性。但如果信息披露不充分,发行人又会面临监管压力和诉讼风险。在这种两难情形下,发行人极有可能采取折中策略,在招股说明书中使用更多不确定性的词语。在我国高度依赖语境沟通的文化背景下,招股说明书不确定语气的使用更是十分微妙,它可能源于未来战略缺乏清晰度、预期收益不够稳定或者潜在的不和谐股东关系。基本面不稳定或低质量的发行人在主观上试图通过传递不确定信号来增加收益,同时规避监管风险,但这在客观上增加了信息不对称程度。

那么,招股说明书中的不确定语气对机构投资者会有什么影响呢?Ellsberg悖论指出,由于对不确定性的厌恶,投资者在决策时会配置更大的权重在含有确定信息的事件上,而配置更小的权重在含有模糊信息的事件上,因此收益信息的模糊程度会影响投资组合的配置^[15]。Epstein、Schneider^[16]指出,投资者在处理不确定信息时,倾向于对信息质量进行最坏的评估,投资者厌恶信息质量差的资产,进而会导致不确定性溢价。Haemin、Pankaj^[17]研究了招股说明书披露信息的清晰程度与IPO抑价之间的关系,结果发现信息环境中的模糊信号会导致新股发行定价过低。显然招股说明书中更多的不确定语气会增加投资者对IPO估值的意见分歧程度,汪宜霞、张辉^[18]研究发现分析师对IPO估值的意见分歧越大,IPO首溢价的程度也越高。而负面语气对估值的影响则体现在两个方面:①更多的负面信息会对估值产生“压力效应”,机构投资者的负面判断将会反映在较低的报价中;②负面信息同不确定信息一

样,更多的负面语言也会增加IPO估值的难度和不确定性。Loughran、McDonald^[19]指出,对于历史业绩较差和负面信息较多的公司,投资者进行正的现金流预测的难度会更大,因此负面信息也会导致更多的意见分歧和不确定性。Arnold等^[20]研究表明,由于不同的投资者对招股说明书中的软信息有着不同的理解,招股说明书中的风险因素信息会导致更多的投资者意见分歧。姚颐、赵梅^[2]也认为,在我国证券市场中,披露风险信息对公司来说本身就是一种风险,对负面信息的披露会带来不确定的后果。因此,相对于正面信息,招股说明书中的负面信息会导致更多的估值不确定性。当难以判断IPO价值时,投资者会把不确定性视作资产的风险因素,并对持有该资产要求收益补偿^[16]。

承销商在IPO定价和招股说明书撰写中也发挥着非常重要的作用。Loughran、McDonald^[19]认为,发行人相对于承销商而言通常处于信息弱势地位。Baron^[21]提出的委托代理理论认为,新股发行的事前不确定性因素越多,承销商的承销风险也越高,IPO抑价程度就越大。Sherman、Titman^[22]研究发现,承销商承担着确保发行顺利完成的利益诉求,承销商会通过发行抑价来激励机构投资者致力于价值发现和息生产,并通过分配抑价股票而间接受益。创业板上市公司多为高科技和新兴产业企业,创业板公司IPO只需要通过初步询价便可由承销商直接确定发行价格,使得承销商对IPO定价拥有更多的主导权。创业板上市标准相对宽松,这些资质相对欠佳的公司往往会面临更大的不确定性,承销商为降低自身风险会更大程度地谨慎定价。由于招股说明书中含有更高程度不确定及负面语气的发行人通常处于更加弱势的地位,承销商因此可以更容易地压低发行价格,为自身谋利。

结合前人的研究成果,本文依次从发行人、投资者和承销商三方视角,分析了招股说明书语气影响IPO定价的内在机理,初步得出以下结论:①低质量发行人倾向于使用更多的不确定语言和负面语言,增加了信息不对称程度,进而可能导致更高的IPO抑价率;②更多的不确定语气和负面语气增加了机构投资者的估值难度和投资风险,投资者会对此要求更高的收益补偿,即导致更高的抑价率;③承销商为确保发行顺利完成并降低风险,面对招股说明书包含更多不确定和负面语气的公司,承销商会利用自身优势进一步压低IPO发行价格。为了验证以上

关于发行人、机构投资者和承销商行为分析结论的可靠性,本文提出如下假设:

假设1:招股说明书中包含的不确定和负面语气越多(少),则IPO抑价率越高(低)。

现有文献通常将股票收益波动率作为衡量事后不确定性的指标,并将上市公司的负面事件与股票收益率波动联系起来,如管理层变动和负面新闻报道等^[2,20]。由于招股说明书中的不确定和负面语气可能揭示了公司基本面中所隐含的不确定性,这种事前不确定性通常与股票上市后的事后不确定性是一致的。因此,我们可以将不确定和负面语气与IPO后新股的收益波动联系起来:①招股说明书中含有的不确定和负面语气往往与公司缺乏稳定的基本面有关,公司上市后的发展前景可能具有更多的不确定性,业绩可能有更大幅度的增长或下滑,即更多的事前不确定性可能意味着更大幅度的事后基本面波动。②在管理层讨论公司经营状况和未来战略时,招股说明书文本语言的确定性和精确度也会影响IPO的估值。较高比例的不确定和负面语气会使投资者准确理解估值信息变得更加困难,导致更多的投资策略分歧和预期差异。投资者对新股进行主观估值的不确定性越大,会导致股价波动的幅度也越大。因此本文提出第二个假设:

假设2:招股说明书中不确定和负面语气的含量越高(低),IPO后新股的收益波动率也越高(低)。

Arnold等^[20]研究表明,招股说明书中的模糊性软信息不仅会导致更多的投资者意见分歧,还会影响投资者的投资组合选择,进而影响股票上市后的价格变化。Gao等^[23]通过实证研究发现,投资者之间的意见分歧与新股长期表现存在显著的负相关关系。Miller^[24]提出的异质信念假说认为,在卖空限制下过度乐观的投资者使IPO价格严重偏离真实价值,但长期来看新的信息使投资者分歧逐渐减少,新股价格最终会逐渐回归真实价值。邹高峰等^[25]研究表明,IPO长期弱势表现的主要原因有:新股上市初期的高溢价、投资者的过度乐观情绪和投资者意见分歧。陈鹏程、周孝华^[7]研究了媒体情绪对IPO长期表现的影响,发现媒体情绪越高,IPO抑价程度越高,但新股长期表现反而更差。

通过梳理相关文献,本文认为招股说明书语气可能从以下三个方面对新股长期收益率产生影响:①招股说明书中使用较高比例的不确定和负面语气,会使投资者准确理解价值信息变得更加困难,导

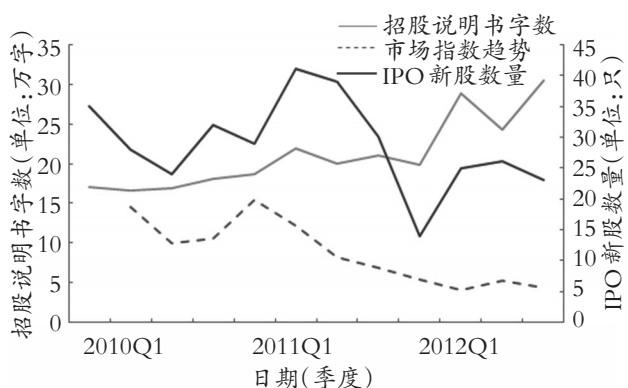
致更大程度的投资者意见分歧,进而导致新股更加弱势的长期表现。②更多的不确定和负面语气导致一级市场更高的抑价程度,吸引二级市场散户的过度追捧,并可能引导乐观投资者(盲目地或带有偏见地)把未来更高的不确定性看做更高的潜在收益,最终导致新股上市初期的过度溢价。③招股说明书含有更多不确定和负面信息可能意味着公司在上市后的发展前景具有更多的不确定性,其“业绩变脸”的可能性也更大,随着股价逐渐回归价值,新股的长期表现会更加弱势。综上所述,本文提出如下假设:

假设3:招股说明书中包含的不确定和负面语气越多(少),则新股长期收益表现越弱(强)势。

三、研究设计

1. 样本选择和数据来源。我国创业板于2009年10月正式开市。2012年10月,证监会暂停IPO,并在2014年IPO重启后设置了首日涨停板限制,此后大部分IPO的首日收盘价都达到了44%的涨停板,导致IPO首日涨幅不能作为衡量抑价程度的指标。因此,本文选取2009年10月~2012年10月在创业板上市的356家上市公司作为研究样本,并剔除10家数据缺失和金融行业上市公司,最终得到346个IPO样本。本文所使用的股票价格、收益和其他相关样本数据是从Wind数据库和国泰安数据库获取的,而招股说明书文件则从深圳证券交易所网站下载。

2. 招股说明书文件的处理。将下载的招股说明书文件统一转换为只包含汉字和数字信息的Word文件,删除所有招股说明书文件中的JPG和HTML等内容。然后将剩下的文本解析成单词,通过统计方法得到在随后的分析中需要使用的各种计数和比例。下图按季度统计了样本期间在创业板上市的新股数量和招股说明书的平均总字数。



招股说明书字数、IPO新股数量和创业板市场指数图

如图所示,招股说明书平均总字数在样本期间

处于明显的上升趋势。在2010年二季度以前,招股说明书平均总字数稳定在17万字左右,到2011年逐渐上升至20万字左右,然后从2011年四季度开始急剧上升,到2012年三季度停发新股时达到30万字。招股说明书总字数的不断攀升与我国IPO信息披露制度的逐渐完善有关。此外,熊市期间IPO数量的大幅减少也使得上市门槛变相提高,进而导致上市公司在招股说明书中需要提供更加完善的信息披露和更加详尽的商业计划。

从图中还可以看出,样本期间包含了2009年牛市末期、2010年的震荡市,以及2011~2012年的两年大熊市,创业板指数整体处于下降趋势。在此期间,创业板的新股发行数量波动较大,但同样呈下降趋势,且其波动特征与创业板指数基本同步。例如,在2009~2010年间有较多的公司在创业板上市,而在2011年熊市期间新股数量则大幅度减少。2011年四季度新股数量达到最低点,只有14只新股上市。

3. 招股说明书语气的度量。在金融文本分析领域,国内外学者主要通过统计正、负面词语的比例来度量文本语气^[4-6]。本文借鉴Loughran、McDonald^[19]以及汪昌云、武佳薇^[26]衡量文本语气的方法,将语气分析技术应用到对中文招股说明书的研究中,通过统计不确定、负面和正面词语在全文中的比例来度量招股说明书语气。首先,参考《最新汉英经济金融常用术语》,将Loughran、McDonald^[19]使用的情感词库列表翻译为中文。然后,根据《现代汉语词典》和中文语境对情感词库列表的中文翻译进行整理和补充,得到初步的不确定、负面和正面语气词库。最后,根据知网发布的中文情感词典和台湾大学发布的中文情感极性词典对正、负面词库进行匹配调整和补充,最终整理出适用于中文金融语境的不确定、负面和正面语气词库。本文构建的三类词库列表分别包括271个不确定词语、4370个负面词语和2256个正面词语。在金融语境中,通常有个别负面词语带有不确定的涵义,本文将这类词语同时归属于不确定词语列表和负面词语列表。

本文采用不确定、负面和正面词语字数占文本总字数的比例来量化文本语气。由此我们编写出计算字频的Java程序,字频计算程序在Java Runtime Environment(JRE)环境中运行,通过输入处理过的招股说明书文本和语气词库列表,便可输出字数频率。

4. 计量模型和变量定义。构建普通最小二乘估

计(OLS)回归模型,将IPO招股说明书中的不确定、负面和正面语气(在模型中统一用Prop表示)作为解释变量,IPO抑价率(IPO_underpricing)、新股收益波动率(IPO_volatility)和新股长期收益率(Long_return)作为被解释变量,检验了招股说明书语气对IPO抑价、新股收益波动率和长期收益率的影响。本文构建的三个OLS回归模型分别如下:

$$IPO_underpricing = \alpha + \beta_p Prop + \sum \beta_i Control + \varepsilon_i \quad (1)$$

$$IPO_volatility = \alpha + \beta_p Prop + \sum \beta_i Control + \varepsilon_i \quad (2)$$

$$Long_return = \alpha + \beta_p Prop + \sum \beta_i Control + \varepsilon_i \quad (3)$$

在上述模型中,IPO抑价率为上市首日收盘价相对于发行价的变动比例,新股长期收益表现采用新股上市1月后买入并持有36个月的收益率衡量。在回归模型(2)中,对新股收益波动率的度量则借鉴Andersen等^[27]提出的实际波动率计算方法,根据个股IPO之后第6~65日期间的每日收益数据(r_i)计算个股收益波动的RMSE值,并将其作为对实际波动率的估计值。

用 r_i 表示第*i*日股价变动的对数收益率, r_i 的计算公式如下:

$$r_i = \ln(P_i) - \ln(P_{i-1}) \quad (4)$$

个股在上市*t*日后的一段期间内,实际波动率的计算公式如下:

$$\sigma_t^n = \sqrt{\frac{1}{n-1} \sum_{i=t+1}^{n+t} (r_i - r_m)^2} \quad (5)$$

式中, σ_t^n 表示第*t*+1~*n*+*t*日共*n*日的真实波动率。为了减少首发初期非常规交易活动对波动率计算的影响,剔除前5个交易日的数据,也即*t*取5日,*n*取60日。 r_m 表示*n*日内新股收益率的平均值。

通过梳理以前文献中常用的IPO控制变量^[7,8,26],根据其对于IPO定价的解释程度选取了11个控制变量:公司年龄、发行规模、市场收益、风险投资背景、承销商声誉、网下认购倍数、每股收益、净资产收益率、资产负债率、资产规模和招股上市间隔期(从招股说明书发布日到上市日的间隔天数)。在研究新股长期表现时,为了消除市场因素和投机因素的影响,还选择了同期的市场收益率(长期市场收益)和同期个股换手率(长期换手率)作为控制变量。具体变量定义见表1。

四、实证分析

1. 描述性统计。主要变量的描述性统计见表2。由表2可见,我国创业板IPO有很高的抑价程度,平

表1 变量定义

变量类型	变量符号	变量名称	变量说明
被解释变量	IPO_underpricing	IPO抑价率	首日收盘价相对于发行价变动的百分比
	IPO_volatility	新股收益波动率	新股上市后第6~65日期间收益的真实波动率
	Long_return	新股长期收益率	新股上市1月后买入并持有36个月的收益率
解释变量	P_Uncertainty	不确定语气指数	不确定词语字数占招股说明书总字数的比例
	P_Negative	负面语气指数	负面词语字数占招股说明书总字数的比例
	P_Positive	正面语气指数	正面词语字数占招股说明书总字数的比例
控制变量	Age	公司年龄	公司成立年数
	Size	发行规模	初始发行股票数量
	M_return	市场收益	发行日的市场收益率
	VC	风险投资背景	新股有风险资本背景时取值为1,否则为0
	Underwriters	承销商声誉	排名前十的承销商取值为1,否则为0
	Multiple	网下认购倍数	网下机构认购总量除以配售数量
	EPS	每股收益	上市前一年净利润除以发行总股本
	ROE	净资产收益率	净利润与净资产的比值
	Debt	资产负债率	负债占总资产的比例
	Asset	资产规模	上市前一年总资产的自然对数
	Days	招股上市间隔期	从招股说明书发布到正式上市的间隔天数
变量	LM_return	长期市场收益	买入并持有股票36个月期间的同期市场指数收益率
	L_turnover	长期换手率	买入并持有股票36个月期间的股票换手率

均抑价率为34.362%,最大值高达209.736%。另外,我们还统计了不同时期的IPO抑价情况:2009年牛市期间,35家公司的IPO平均抑价率为92.484%;2010~2011年一季度震荡期间,154家公司的平均抑价率为30.556%;2011年二季度~2012年熊市期间,157家公司的平均抑价率为25.151%。由此可见,市场环境对IPO抑价率的影响也十分明显。新股收益波动率的平均值和中值分别为3.481和2.292,这说明在新股上市后的6~65日期间每日收益有着较

大的波动性。新股长期收益率的平均值为-16.906%，中值为-15.835%，远低于长期市场收益率的平均值-7.203%和中值-8.784%，这说明新股上市后一年的平均表现往往比市场指数更加弱势。该结果与邹高峰等^[25]的观点一致，他们指出新股在上市后1~3年内存在长期弱势表现。

表2 各变量的描述性统计

变量	平均值	标准差	最小值	中值	最大值
IPO_underpricing	34.362	37.101	16.683	25.261	209.736
IPO_volatility	3.481	2.061	0.414	2.292	11.409
Long_return	-16.906	20.872	-67.714	-15.835	32.206
P_Uncertainty	0.202	0.033	0.110	0.200	0.320
P_Negative	2.772	0.338	1.820	2.755	4.120
P_Positive	4.094	0.391	3.020	4.080	5.460
Age	8.681	4.374	0.496	8.781	24.093
Size	18.992	9.553	6.997	16.012	80.005
M_return	0.204	4.303	-9.977	0.245	10.603
VC	0.583	0.498	0	1	1
Underwriters	0.344	0.474	0	0	1
Multiple	39.821	34.537	1.506	27.178	226.405
EPS	0.587	0.347	0.157	0.517	3.717
ROE	26.331	9.155	10.881	24.811	102.203
Asset	19.508	0.578	18.098	19.478	21.779
Debt	38.848	14.683	4.654	39.091	79.170
Days	13.901	5.306	9.003	13.001	51.005
LM_return	-7.203	14.340	-22.447	-8.784	51.654
L_turnover	1595.455	676.734	337.254	1533.286	4484.668

表2还显示出，不确定语气指数的平均值(0.202%)和中值(0.200%)几乎相同，负面语气指数和正面语气指数的平均值和中值也非常接近。正面

语气的平均值是负面语气的1.48倍，这意味着招股说明书中的正面词语明显多于负面词语，招股说明书对正面信息和乐观预期进行描述时往往不吝言辞，而对负面词语的使用则非常谨慎。有风险投资背景的IPO样本占了一半以上，而由顶级承销商承销的样本只有三分之一左右。上市前净资产收益率的平均值为26.331%，说明大部分公司在上市前有着较好的盈利能力。网下认购倍数的数据悬殊，最大值高达226.405倍，最小值却只有1.506倍。新股平均的长期换手率高达1595.455，说明二级市场投资者对新股和次新股有着明显的投机行为。

表3展示了主要变量的相关系数。不确定语气和负面语气与IPO抑价率有着较高的相关性(相关系数分别为0.241和0.148)。不确定语气与负面语气之间的相关系数只有0.044，意味着个别词语同时出现在两类词库列表中对统计结果影响不大。此外，招股说明书语气与部分公司特征变量存在相关性，但相关系数都在0.25以下，因此回归模型不存在严重的多重共线性问题。

2. 回归结果。表4将IPO抑价率作为被解释变量，报告了对假设1的检验结果。第(1)列为只含有11个IPO控制变量的回归结果，后三列回归分别增加了不确定语气、负面语气和正面语气作为解释变量。从调整R²的变化可以看出，在加入不确定语气和负面语气之后，回归模型的拟合程度从40.15%分别提升到41.80%和42.15%。不确定语气和负面语气与IPO抑价率均显著正相关。其中，不确定语气对IPO抑价率的回归系数为0.9864，在5%的水平上显著(t统计量为2.04)，意味着含有更高比例不确定语气的IPO有着更高的抑价率；平均来说，不确定语气

表3 主要变量的相关系数

变量	IPO_underpricing	P_Uncertainty	P_Negative	P_Positive	Age	VC	Underwriters	EPS	ROE	Asset	Debt
IPO_underpricing	1.000										
P_Uncertainty	0.241	1.000									
P_Negative	0.148	0.044	1.000								
P_Positive	-0.056	-0.036	-0.122	1.000							
Age	0.021	-0.091	-0.110	0.130	1.000						
VC	0.048	0.194	0.239	-0.024	0.013	1.000					
Underwriters	-0.061	0.021	-0.027	0.020	-0.007	0.046	1.000				
EPS	-0.168	-0.024	-0.149	0.044	0.058	0.027	0.038	1.000			
ROE	-0.086	0.023	-0.097	0.065	-0.112	-0.088	0.081	0.461	1.000		
Asset	-0.204	-0.112	0.018	0.161	0.071	0.117	0.027	0.155	-0.110	1.000	
Debt	-0.058	0.081	0.102	-0.072	0.068	-0.049	-0.112	-0.087	-0.091	0.451	1.000

含量每上升1%, IPO抑价率约显著提高0.99%。负面语气对抑价率的回归系数为0.1174,在5%的水平上显著(t统计量为2.41),因此更多的负面语气对IPO抑价率也存在正向影响;平均而言,招股说明书负面语气含量每上升1%,抑价率约显著提高0.12%。

表4 招股说明书文本语气与IPO抑价率

变量	(1)	(2)	(3)	(4)
P_Uncertainty		0.9864** (2.04)		
P_Negative			0.1174** (2.41)	
P_Positive				-0.0570 (-0.67)
Age	0.0096** (2.54)	0.0093** (2.46)	0.0097** (2.56)	0.0098** (2.57)
Size	0.1119 (0.40)	0.0683 (0.24)	0.0825 (0.30)	0.1120 (0.34)
M_return	2.6982*** (6.82)	2.5428*** (6.33)	2.6635*** (6.78)	2.6920*** (6.80)
VC	0.0270 (0.83)	0.0143 (0.44)	0.0073 (0.2)	0.0270 (0.83)
Underwriters	0.0089 (0.27)	0.0103 (0.31)	0.0006 (0.02)	0.0086 (0.25)
Multiple	0.0050*** (10.17)	0.0049*** (9.77)	0.0049*** (9.96)	0.0050*** (10.06)
EPS	0.0129 (0.18)	0.0073 (0.10)	-0.0094 (-0.13)	0.0140 (0.20)
ROE	-0.0613 (-0.27)	-0.0751 (-0.33)	0.0054(0.02)	-0.0594 (-0.26)
Asset	-0.1485** (-2.59)	-0.1352** (-2.35)	-0.1355** (-2.37)	-0.1517** (-2.64)
Debt	0.2958* (1.94)	0.2800* (1.84)	0.2463* (1.61)	0.2945* (1.93)
Days	0.0066* (1.95)	0.0069** (2.05)	0.0073** (2.16)	0.0066* (1.95)
截距	2.7198*** (2.65)	2.2959** (2.19)	2.1647** (2.07)	2.9942*** (2.70)
调整的R ²	0.4015	0.4180	0.4215	0.4022

注:括号中为t统计量,***、**和*分别表示1%、5%和10%的显著性水平。下同。

第(4)列回归结果显示正面语气的回归系数并不显著,该结果与其他研究金融文本语气的文献是一致的^[19, 26]。有文献指出,正面信息和负面信息对证券市场的影响是不对称的,同等程度的负面信息对股市的影响更加强烈^[26]。前景理论也指出,投资者对负面信息和正面信息的敏感程度不同,对负面信息的反应更加敏感而深刻,但持续时间相对短暂。

此外,发行人和承销商倾向于对IPO做出偏乐观的预期,对负面信息的使用会十分谨慎。同时,机构投资者对招股说明书披露的业绩波动和经营困境等负面信息也会重点关注。因此,相对于正面信息,负面信息对IPO定价有更强的解释作用是合理的。市场收益和网下认购倍数都与IPO抑价率显著正相关。年龄更长的公司有着更高的抑价程度,由于这些公司的基本面通常更加稳健,知名度也相对更高,因此更易获得投资者的信赖和关注。招股上市间隔期与IPO抑价率显著正相关,较长的间隔日期使投资者获得更充分的资金准备,也可能使发行信息的传播更广泛,引发更多的投资者关注。综上所述,该实证结果验证了假设1,证明了招股说明书不确定语气和负面语气对IPO抑价有显著的解释作用。

表5详细报告了招股说明书文本语气与IPO后6~65日期间的收益波动率的回归结果。第(1)列为只含有控制变量的回归结果,在第(2)、(3)列中,不确定和负面语气与新股收益波动率显著正相关,相关系数分别为4.7441(t统计量为4.79,在1%的水平上显著)和0.1456(t统计量为4.17,在1%的水平上显著)。平均来说,招股说明书中的不确定语气和负面语气含量每上升1%,新股收益波动率分别显著增加约4.74%和0.15%。此外,第(4)列结果显示,正面语气的回归系数不显著。在控制变量方面,公司基本面变量与新股收益波动率显著负相关,也即公司规模越大、每股收益越高,则新股收益波动率越小。同时,风险投资背景也与IPO后新股波动性密切相关,有风险投资背景的IPO往往含有更多的估值不确定性,也可能导致二级市场的交易价格产生更多波动。此外,IPO发行规模与新股收益波动率显著负相关,发行规模越小的新股在二级市场的波动性越大。综上所述,招股说明书中含有的不确定和负面语气越多,意味着IPO内在价值的事前不确定性也越高,从而导致IPO之后的新股收益率产生更大的波动性。因此,假设2得到验证。

表6报告了招股说明书文本语气与新股长期表现的回归结果。后三列回归分别增加了不确定、负面和正面语气作为解释变量,模型的调整R²值从32.44%分别上升到35.36%、34.62%和35.04%,表明回归模型的拟合程度有所提升。不确定语气、负面语气与新股长期收益率都呈显著负相关关系,相关系数分别为-0.9348(t统计量为-3.14,在1%的水平上显著)和-0.0706(t统计量为-2.44,在5%的水平

表5 招股说明书文本语气与新股收益波动率

变量	(1)	(2)	(3)	(4)
P_Uncertainty		4.7441*** (4.79)		
P_Negative			0.1456*** (4.17)	
P_Positive				0.2886 (0.65)
Age	-0.0278 (-0.15)	-0.0510 (-0.29)	-0.0319 (-0.18)	-0.0340 (-0.19)
Size	-1.3616** (-2.48)	-1.3205** (-2.26)	-1.3631** (-2.48)	-1.3577** (-2.46)
M_return	0.0099 (0.49)	0.0173 (0.50)	0.0159 (0.47)	0.0108 (0.48)
VC	0.3873** (2.23)	0.3580** (2.06)	0.3832** (2.20)	0.3901** (2.24)
Underwriters	0.1181 (0.68)	0.0865 (0.50)	0.1154 (0.66)	0.1096 (0.63)
Multiple	0.0163 (0.35)	0.0154 (0.11)	0.0163 (0.35)	0.0162 (0.32)
EPS	-0.3666*** (-3.45)	-0.2879*** (-3.15)	-0.3560*** (-3.41)	-0.3541*** (-3.41)
ROE	2.3548 (0.72)	2.0100 (0.47)	2.3083 (0.68)	2.3167 (0.69)
Asset	-0.6913** (-2.04)	-0.6703** (-2.03)	-0.6943** (-2.06)	-0.6893** (-2.01)
Debt	-1.3879 (-0.43)	-1.3767 (-0.29)	-1.3976 (-0.44)	-1.3685 (-0.41)
截距	3.6303* (1.68)	2.9122 (1.32)	4.2248* (1.87)	2.7016 (0.98)
调整的R ²	0.2679	0.2723	0.2674	0.2675

上显著);平均来说,招股说明书文本的不确定语气或负面语气含量每上升1%,则新股长期收益率分别显著降低约0.93%或0.07%。正面语气指数与新股长期收益率显著正相关(系数为0.1433,t统计量为2.86,在1%的水平上显著),招股说明书中的正面语气含量越高,新股长期收益率也越高;平均地,正面语气含量每提升1%,新股长期收益率显著提高0.14%。在IPO环境下,投资者无法获取公司的历史交易数据和其他公开信息,只能依靠招股说明书信息对公司进行估值,招股说明书中的不确定语气和负面语气会增加投资者的异质信念程度。邹高峰等^[25]指出,投资者意见分歧越大,IPO的长期弱势程度也越大,本文的研究结论与此一致,且进一步支持了异质信念假说。随着时间的推移,投资者可获取的公开信息不断增多,投资者的估值趋于一致,股价趋于真实价值,股票的长期表现将由公司的基本面决定。公司

表6 招股说明书文本语气与新股长期表现收益率

变量	(1)	(2)	(3)	(4)
P_Uncertainty		-0.9348*** (-3.14)		
P_Negative			-0.0706** (-2.44)	
P_Positive				0.1433*** (2.86)
Age	0.0043* (1.95)	0.0043* (1.98)	0.0042* (1.92)	0.0040* (1.81)
Size	-0.0005 (-0.28)	-0.0001 (-0.07)	-0.0003 (-0.18)	-0.0005 (-0.28)
LM_return	0.0772*** (8.11)	0.0642*** (7.48)	0.0772*** (8.16)	0.0754*** (8.11)
VC	-0.024 (-1.27)	-0.0125 (-0.64)	-0.0128 (-0.65)	-0.0240 (-1.26)
Underwriters	0.0229 (1.14)	0.0239 (1.21)	0.0171 (0.86)	0.0217 (1.10)
L_turnover	-0.0001*** (-5.87)	-0.0001*** (-5.62)	-0.0001*** (-5.80)	-0.0001*** (-6.16)
EPS	0.0440*** (3.04)	0.0465*** (3.12)	0.0565*** (3.25)	0.0403*** (3.07)
ROE	0.0534 (0.40)	0.0647 (0.49)	0.0113 (0.38)	0.0448 (0.34)
Asset	0.0849** (2.44)	0.0745** (2.16)	0.0776** (2.24)	0.0915*** (2.65)
Debt	-0.0410 (-0.45)	-0.0349 (-0.39)	-0.0151 (-0.17)	-0.0379 (-0.42)
截距	-1.7317** (-2.75)	-1.3690** (-2.16)	-1.4108** (-2.21)	-2.3864*** (-3.59)
调整的R ²	0.3244	0.3536	0.3462	0.3504

成立时间越长,股票的长期表现越好;基本面越好(每股收益越高、资产规模越大),长期表现也会越好。此外,长期市场收益也与新股长期收益率显著正相关,但长期换手率与新股长期收益率则显著负相关,换手率越大,新股的长期表现越差。鉴于较高的换手率通常伴随着较高的股价波动率,这是由投资者较高程度的意见分歧造成的,因此该结果也符合前文的异质信念假说。综上所述,实证结果也验证了假设3。

3. 稳健性检验。鉴于样本期间的市场行情表现较弱,而且国家对IPO审批的指导政策和上市企业的性质分类也可能影响IPO后新股的市场表现,因此作为对实证结果的稳健性检验,本文在回归分析中添加新的控制变量,考察了招股说明书文本语气对IPO抑价率和新股长期表现的影响。添加的相关控制变量包括:①上市前3个月的IPO过会率代理

作为政策变量；②上市前15个交易日的市场指数收益率作为市场整体行情变量；③招股说明书文本的总字数；④IPO前公司的账面总资产；⑤国有企业虚拟变量，公司属于国有控股背景取值为1，否则为0；⑥高科技企业虚拟变量，公司属于高科技企业（判定依据为国务院批准发布的《高科技企业认定管理办法》）取值为1，否则为0。添加这些额外的控制变量后重新检验，发现对不确定、负面语气的系数值和显著性水平影响很小，这说明本文的实证研究结论具有较好的稳健性（限于篇幅，未列出回归结果）。

五、结论

本文针对招股说明书中隐含的不确定信息和负面信息，结合信息不对称理论、信息生产假说和异质信念假说等理论，提出了招股说明书语气影响IPO抑价率、新股收益波动率以及新股长期表现的研究假设。接下来，采用文本分析方法构建了适用于金融语境的文本语气词库，对招股说明书中的不确定语气、负面语气和正面语气进行了度量。然后，以2009~2012年在创业板上市的346家公司作为样本进行了实证研究，结果发现招股说明书中的不确定和负面语气对IPO抑价率、新股收益波动率以及新股长期表现都存在显著影响。本文得到的主要结论包括：

第一，招股说明书不确定和负面语气揭示出公司在发展历程和未来前景中存在的确定性因素。①不确定语气源于未来战略缺乏清晰度、预期收益不够稳定或者潜在的不和谐股东关系。基本面不稳定或低质量的发行人为增加收益和规避监管风险，更有可能使用较高比例的不确定语气，这在一定程度上增加了IPO的信息不对称程度。②更多的负面语气会对估值产生“压力效应”。机构投资者对IPO的负面判断将会反映在较低的报价中，且更多的负面语气也会增加IPO估值的难度和不确定性。

第二，招股说明书中含有更多不确定和负面语气会导致更高的IPO抑价率。更多的不确定语气和负面语气增加了机构投资者的估值难度和投资风险，投资者会对此要求更高的收益补偿，即导致更高的发行抑价程度。承销商在IPO定价过程中占据强势地位，对发行定价拥有更大的主导权。由于招股说明书中含有更高比例不确定及负面语气的发行人通常处于更加弱势的地位，为了降低发行风险，承销商可以更容易地利用自身优势压低发行价格，进而导致更高的发行抑价程度。

第三，招股说明书中的不确定和负面语气越多，

导致IPO之后新股的收益波动率也越高。更多的事前不确定性可能意味着更大幅度的事后基本面波动，招股说明书含有更多不确定和负面语气的公司往往缺乏稳定的基本面，上市后股票的业绩和估值也可能产生更大幅度的波动；招股说明书中较高比例的不确定和负面语气会使投资者准确估值变得困难，投资者对新股进行估值的不确定性也更大，会导致股价波动的幅度也更大。

第四，招股说明书中更多的不确定和负面语气会导致新股更加弱势的长期表现。不同投资者对招股说明书中的软信息有着不同的理解，更多的不确定语气和负面语气会导致更大程度的投资者意见分歧。在投资者意见分歧严重的情况下，高抑价发行会引致二级市场的过度追捧，乐观投资者（盲目地或带有偏见地）会把未来更高的不确定性看做更高的潜在收益，最终导致新股上市初期的过度溢价。招股说明书中含有更多不确定和负面信息可能意味着公司在上市后的发展前景具有更大的不确定性，出现“业绩变脸”的可能性也更高，随着股价逐渐回归真实价值，新股的长期表现会更加弱势。

上述结论对监管部门和投资者的启示主要体现在以下方面：①重视和加强IPO文本信息披露的规范程度。监管部门不仅应要求公司披露的信息真实、准确和完整，还应该对招股说明书文本信息的披露进行标准化、规范化约束，使投资者能够准确地了解公司基本面。对信息造假和欺诈行为采取强制退市措施和法律制裁。②通过推进IPO注册制改革来弱化承销商和发行人之间信息和地位的不对称。通过降低直接融资门槛和加强企业上市辅导，提升发行人在IPO过程中的地位和话语权，形成促进发行人披露真实、准确信息的激励机制。③加强对投资者的教育，倡导长期投资和价值投资。营造以内在价值作为投资标准的市场氛围，鼓励投资者不仅要重视数据信息使用，还应重视对招股说明书文本价值信息的挖掘。

主要参考文献：

- [1] Rock K.. Why New Issues Are Underpriced[J]. Journal of Financial Economics, 1986(1): 187~212.
- [2] 姚颀, 赵梅. 中国式风险披露、披露水平与市场反应[J]. 经济研究, 2016(7): 158~172.
- [3] 张秀敏, 汪瑾, 薛宁等. 语义分析方法在企业环境信息披露研究中的应用[J]. 会计研究, 2016(1):

- 87~96.
- [4] Ahmed Y., Elshandidy T.. The Effect of Bidder Conservatism on M&A Decisions: Text-based Evidence from US 10-K Filings [J]. International Review of Financial Analysis, 2016(7):176~190.
- [5] Milian J. A., Smith A. L.. An Investigation of Analysts' Praise of Management During Earnings Conference Calls [J]. Journal of Behavioral Finance, 2017(1):65~77.
- [6] Loughran T., McDonald B.. Textual Analysis in Accounting and Finance: A Survey [J]. Journal of Accounting Research, 2016(9):1187~1230.
- [7] 陈鹏程,周孝华. 媒体情绪与IPO市场表现——询价制下的理论与实证 [J]. 管理评论, 2016(7): 22~30.
- [8] 张维,翟晓鹏,邹高峰等. 市场情绪、投资者关注与IPO破发 [J]. 管理评论, 2015(6):160~167.
- [9] 郝项超,苏之翔. 重大风险提示可以降低IPO抑价吗?——基于文本分析法的经验证据 [J]. 财经研究, 2014(5):42~53.
- [10] Jenkinson T., Ljungqvist A.. Going Public: The Theory and Evidence on How Companies Raise Equity Finance [M]. Oxford: Oxford University Press, 1996:1~256.
- [11] Carter R., Dark F., Singh A.. Underwriter Reputation, Initial Returns, and the Longrun Performance of IPO Stocks [J]. Journal of Finance, 1998(1):285~311.
- [12] Doran D. T.. Stock Splits and False Signaling Cost Within a Management Reputation Framework [J]. Journal of Applied Business Research, 2011 (3):115~26.
- [13] Izhakian Y., Yermack D.. Risk, Ambiguity, and the Exercise of Employee Stock Options [J]. Journal of Financial Economics, 2017(1):65~85.
- [14] Guo R. J., B. Lev, N. Zhou. Competitive Cost of Disclosure by Biotech IPOs [J]. Journal of Accounting Research, 2004(2):319~355.
- [15] Garlappi L., R. Uppal, T. Wang. Portfolio Selection with Parameter and Model Uncertainty: A Multiple-prior Approach [J]. Review of Financial Studies, 2007(1):41~81.
- [16] Epstein L., Schneider M.. Ambiguity, Information Quality, and Asset Pricing [J]. Journal of Finance, 2008(1):197~228.
- [17] Haemin Dennis Park, Pankaj C. Patel. How Does Ambiguity Influence IPO Underpricing? The Role of the Signalling Environment [J]. Journal of Management Studies, 2015(6):796~818.
- [18] 汪宜霞,张辉. 卖空限制、意见分歧与IPO溢价 [J]. 管理学报, 2009(9):1204~1208.
- [19] Loughran T., McDonald B.. IPO First-day Returns, Offer Price Revisions, Volatility, and Form S-1 Language [J]. Journal of Financial Economics, 2013(5):307~326.
- [20] Arnold T., Fische R. P. H., North D.. The Effects of Ambiguous Information on Initial and Subsequent IPO Returns [J]. Financial Management, 2010(4):1497~1519.
- [21] Baron D. P.. A Model of the Demand for Investment Banking Advising and Distribution Services for New Issues [J]. Journal of Finance, 1982(4): 955~976.
- [22] Sherman A., Titman S.. Building the IPO Order Book: Underpricing and Participation Limits with Costly Information [J]. Journal of Financial Economics, 2002(6):3~29.
- [23] Gao Y., Mao C., Zhong R.. Divergence of Opinion and Long-term Performance of Initial Public Offerings [J]. The Journal of Financial Research, 2006(1):113~129.
- [24] Miller E. M.. Risk, Uncertainty and Divergence of Opinion [J]. Journal of Finance, 1977(4):1151~1168.
- [25] 邹高峰,张维,常中阳. 询价制度下中国IPO长期表现 [J]. 管理科学学报, 2012(11):66~75.
- [26] 汪昌云,武佳薇. 媒体语气、投资者情绪与IPO定价 [J]. 金融研究, 2015(9):174~189.
- [27] Andersen T. G., T. Bollerslev, F. X. Diebold, H. Ebens. The Distribution of Realized Stock Return Volatility [J]. Journal of Financial Economics, 2001(1):43~67.

作者单位:重庆大学经济与工商管理学院,重庆400044