

“绿色债券”的募集说明书文本如何影响其风险溢价？

展示人：方清源，付煜，&毛思文
2022年5月10日

研究背景

绿色金融

- 为支持环境改善、应对气候变化和资源节约高效利用的经济活动，即对环保、节能、清洁能源、绿色交通、绿色建筑等领域的项目投融资、项目运营、风险管理等所提供的金融服务

绿色债券

- 将所得资金专门用于资助符合条件条件的绿色项目，或为这些项目进行再融资的新型债券工具
- 《关于加快推进生态文明建设的意见》《关于在银行间债券市场发行绿色金融债券有关事宜公告》
- 截至2020年末，中国累计发行绿色债券约1.2万亿元，位居世界第二。

绿色债券核心特征

环保相关度高

- 多家机构开展绿色债券的评估认证业务，以第三方认证的形式加强债券“绿色”特征的可信度

政策关联度大

- 中央或地方的政策文件均会受到绿色债券发行与交易方的紧密关注

文献回顾

绿色债券定价

- 市场利率、国内通胀水平、发行主体注册资本、信用评级、媒体关注度 ...
- 绿色债券存在更多的发行成本优势。

债券募集说明书

- 非结构化数据：文本分析
- 环境保护相关度 + 政策响应力度
- 控制已发现的影响因子（文本复杂度、区别于其他债券的特异性信息等）

“绿色”词典构建

- 对绿色专利的研究

研究概要

研究概况

➤ 研究目的：

探究“绿色债券”募集说明书中包含的文本信息如何影响债券的风险溢价

➤ 研究流程：

1. 通过Wind数据库与巨潮资讯等网站收集所需数据
2. 基于文本信息构造解释变量，参考已有文献引入其他控制变量
3. 建立线性回归模型进行分析，通过检验解释变量前系数的显著性实现研究目的

数据选取与收集

数据来源与收集

- 文本语料数据来源：
2016年1月1日至2021年12月31日期间，在中国大陆上市发行的“绿色债券”募集说明书。
- 数据筛选：总计发现1091支满足要求的目标债券
 1. 得到Wind数据库认证并收录的“绿色债券”，共2616支。
 2. 排除所有中债登记分类为“国债”的债券
 3. 排除所有债务主体为金融机构的债券
 4. 排除所有债务主体不在中国大陆地区的债券
- 数据收集：通过巨潮资讯网与交易所信息披露门户，采用网络爬虫法下载募集说明书PDF文件

变量构造

解释变量构造

1. 环境保护相关度 Eco

- 意义：衡量绿色债券与环境保护、节能减排等绿色业务的相关性
- 构造流程：首先构建“环境保护”概念词库（结合文献中出现的词库与自行摘选），随后统计债券募集说明书中词库收录词汇出现的频数，并计算其与说明书总词汇数的比值。
- 计算方式：
$$Eco_i = \frac{i \text{ 募集说明书中环境保护相关词汇出现的频数}}{i \text{ 募集说明书的总词汇数}}$$
- 词库示例：“低碳、节能、环境、环保、生态、可持续、清洁、循环、排放、污染、能源”

变量构造

解释变量构造

2. 政策响应力度 *Policy*

- 意义：衡量绿色债券对国家政策、政府决议进行响应的程度或关注度。
- 构造流程：首先构建“政策响应”概念词库（结合文献中出现的词库与自行摘选），随后统计债券募集说明书中词库收录词汇出现的频数，并计算其与说明书总词汇数的比值。
- 计算方式：
$$\text{Policy}_i = \frac{i \text{ 募集说明书中政策响应相关词汇出现的频数}}{i \text{ 募集说明书的总词汇数}}$$
- 词库示例：“政府、国务院、中央、公共、基础设施、全会、决议、文件、号召、社会责任”

变量构造

解释变量构造

3. 文本特质性信息含量 *Heter*

- 意义：衡量募集说明书文本信息中包含的无法被其附近时间区间（前后各90自然日）内，其他绿色债券募集说明书信息所解释的比重。
- 构造流程：首先采用文本向量化方法计算各支债券募集说明书文本的词频向量。随后，划定每一支债券发行前后各90自然日内的其他绿色债券为“可比债券”。最后，将债券 i 的词频向量与其全部“可比债券”的词频向量的算术平均值做回归，得到拟合优度 R_i^2
- 计算方式：
$$\text{Heter}_i = \ln \left(\frac{1-R_i^2}{R_i^2} \right)$$

变量构造

被解释变量

- 意义：债券发行时的风险溢价 *Spread*
- 构造流程：计算债券的票面利率超过发行时无风险利率的差值，选用上海银行间同业拆放利率（Shibor）作为无风险利率标准。其中，Shibor的具体数值为发行前五日利率的平均值。
- 计算方式： $\text{Spread}_i = \text{Coupon Rate} - \text{Shibor}$

变量构造

控制变量

变量类型	变量名称	变量简称	变量定义
财务信息 控制变量	公司规模	Size	公司总资产的自然对数
	资产负债率	Lev	企业总资产/企业总负债
	净资产收益率	ROE	企业净利润/企业总资产
	流动比率	Current	流动资产/流动负债
债券特征 控制变量	债券期限	Maturity	债券期限
	信用评级	Rating	Wind数据库标注的债券信用评级，当债券评级为AA-时取值为1，评级为AA时取值为2，评级为AA+时取值为3，评级为AAA时取值为4。
	期权属性	Convert	债券属于含期权债券则取值为1，否则取值为0
	发行规模	Volume	债券发行总规模的自然对数
固定效应 环境参数	企业存续年限	Age	债券发行时债务主体已经存续的年限
	行业	Industry	债务主体企业所属于的Wind行业分类
	年份	Year	债券发行年份，2016年至2016年分别设置为0-5
	地区	Province	发债企业所属省份

模型与假设

研究假设

H1：绿色债券募集说明书与环境保护的相关度（Eco）越高，债券的风险利差（Spreads）越低。

H2：绿色债券募集说明书对政策的响应力度（Policy）越高，债券的风险利差（Spreads）越低。

H3：绿色债券募集说明书中含有的特质性信息含量（Heter）越高，风险利差（Spreads）越低。

回归模型

Spreads_i

$$\begin{aligned} = & \alpha + \beta_1 \text{Eco}_i + \beta_2 \text{Policy}_i + \beta_3 \text{Heter}_i + \beta_4 \text{Size}_i + \beta_5 \text{Lev}_i + \beta_6 \text{ROE}_i + \beta_7 \text{Current}_i + \beta_8 \text{Maturity}_i \\ & + \beta_9 \text{Rating}_i + \beta_{10} \text{Convert}_i + \beta_{11} \text{Volume}_i + \beta_{13} \text{Age}_i + \sum \text{Industry} + \sum \text{Year} + \sum \text{Province} + \varepsilon_{i,t} \end{aligned}$$

➤ 对回归系数 β_1 , β_2 , β_3 进行统计检验即可验证本研究假设是否成立

模型拓展

回归模型

Spreads_i

$$= \alpha + \beta_1 \text{Eco}_i + \beta_2 \text{Policy}_i + \beta_3 \text{Heter}_i + \beta_4 \text{Size}_i + \beta_5 \text{Lev}_i + \beta_6 \text{ROE}_i + \beta_7 \text{Current}_i + \beta_8 \text{Maturity}_i \\ + \beta_9 \text{Rating}_i + \beta_{10} \text{Convert}_i + \beta_{11} \text{Volume}_i + \beta_{13} \text{Age}_i + \sum \text{Industry} + \sum \text{Year} + \sum \text{Province} + \varepsilon_{i,t}$$

➤ 对回归系数 β_1 , β_2 , β_3 进行统计检验即可验证本研究假设是否成立

基于概念词库的词频



“环境保护” 概念词



“政策响应” 概念词

来源：《福建福清核电有限公司 2016 年公开发行绿色公司债券（第一期）（专项用于碳中和）募集说明书》

关联词发现

Word2vec权重模型

- 所有募集说明书文档作为语料训练模型
- 计算词与词相似性，匹配关联词：概念词库 “环境保护” “政策响应”
- 用关联词扩展概念词库，构造新解释变量 Eco & Policy
 - $$\text{Eco}_i = \frac{i \text{ 募集说明书中环境保护相关词汇出现的频数}}{i \text{ 募集说明书的总词汇数}}$$
 - $$\text{Policy}_i = \frac{i \text{ 募集说明书中政策响应相关词汇出现的频数}}{i \text{ 募集说明书的总词汇数}}$$
- 以相似度（词向量距离）为权重

gensim官网：<http://radimrehurek.com/gensim/index.html>

特征性信息

TF-IDF权重调整

- 构造特质性信息解释变量(Hanley & Hoberg, 2010)时,
 - 词频向量回归，一个维度（词）为一个样本
 - $Heter_i = \ln \left(\frac{1-R_i^2}{R_i^2} \right)$
- 不同词重要程度不同
- 调整词频向量，使用TF-IDF向量，而非词频（TF）向量
- 以IDF值为权重，加权回归WLS

其他文本特征

主题词的相似度

- 调用模型获取文档主题词：AI平台api或预训练LSA模型
- 计算文档主题词的相似度，作为额外解释变量
 - 基于Word2vec的词向量空间距离
 - 或基于词频向量的文档特质性(Hanley & Hoberg, 2010)

情感/语调

- 调用模型评估语段情感倾向：SnowNLP或AI平台api
 - 筛选核心语段：高频出现概念词&TF-IDF均值高
- 合成文档情绪得分，作为额外解释变量

感谢！

- [1] 朱家明, 祝静远. 基于线性回归的绿色债券票面利率影响因素的实证分析[J/OL]. 辽宁石油化工大学学报, 2019, 39(04): 87–92[2022–05–07].
- [2] 许欣怡. 我国绿色债券发行利率影响因素实证分析——基于可贷资金模型[J/OL]. 经营与管理, 2019(06): 21–23[2022–05–07].
- [3] 龚玉霞, 滕秀仪, 赛尔沃, 等. 绿色债券发展及其定价研究——基于二叉树模型分析[J/OL]. 价格理论与实践, 2018(07): 79–82[2022–05–07].
- [4] 倪弋菁. 社会关注度对绿色债券定价影响研究[D/OL]. 上海交通大学, 2019[2022–05–07].
- [5] 王玲. 绿色因素对我国绿色公司债券发行定价影响效果研究[D/OL]. 四川大学, 2021[2022–05–07].
- [6] 吴武清, 甄伟浩, 杨洁, 等. 企业风险信息披露与债券风险溢价——基于债券募集说明书的文本分析[J/OL]. 系统工程理论与实践, 2021, 41(7): 1650–1671.
- [7] 林晚发, 方梅, 沈宇航. 债券募集说明书文本信息与债券发行定价[J/OL]. 管理科学, 2021, 34(4): 19–34.
- [8] 吴育辉, 唐浩博. 风险信息披露能预测债券违约风险吗? ——来自债券募集说明书的经验证据[J/OL]. 财务研究, 2021(5): 15–28.
- [9] Xiang X, Liu C, Yang M, et al. Confession or justification: The effects of environmental disclosure on corporate green innovation in China[J/OL]. Corporate Social Responsibility and Environmental Management, 2020, 27(6): 2735–2750[2022–05–07].
- [10] Messeni Petruzzelli A, Maria Dangelico R, Rotolo D, 等. Organizational factors and technological features in the development of green innovations: Evidence from patent analysis[J/OL]. Innovation, 2011, 13(3): 291–310[2022–05–07].
- [11] Li D, Huang M, Ren S, et al. Environmental Legitimacy, Green Innovation, and Corporate Carbon Disclosure: Evidence from CDP China 100[J/OL]. Journal of Business Ethics, 2018, 150(4): 1089–1104[2022–05–07].