
答审稿专家的宝贵意见

第一份意见

审稿意见 (1.1):

一是，论文标题中有机器学习方法一词，但是综述对于机器学习的综述比较少，是否可以介绍一下债券机器学习研究，与股票市场的区别与难点。

修改说明:

修改稿在“利用机器学习算法推动预警模型创新”部分介绍金融机器学习相关文献后，进一步介绍了机器学习在债券市场和股票市场中应用研究的区别，指出债券市场的相关研究要滞后于股票市场，导致这一区别的可能原因是股票市场异象因子的研究较为成熟，大量因子构成的高维指标天然地可以作为机器学习算法的输入变量。然而，债券市场因其 OTC 交易特征，在基础研究数据采集、数据库构建，尤其是针对市场连续交易数据的获取方面，较股票市场有显著的难度。参见修改稿第 12 页。

进一步，修改稿还指出将机器学习应用于中国债券市场的难点在于构建具有中国特色的债券市场预测指标库，研究者需要花费大量精力考虑研究数据的可得性、可追溯性以及数据质量等诸多问题。参见修改稿第 12-13 页。

审稿意见 (1.2):

二是，前面债券违约风险影响因素的综述，是否相关影响因素是否应该考虑在机器学习指标里。就是说，前后两块内容是否要有一定的联系。

修改说明:

修改稿在“利用机器学习算法推动预警模型创新”部分指出，“为开展基于大数据和机器学习的债券违约风险预警模型研究，本文对债券违约风险影响因素进行了全面梳理，这些因素可以自然地考虑在机器学习算法的输入指标里”，加强了前后两块内容的联系。参见修改稿第 13 页。

第二份意见

审稿意见 (2.1):

在债券违约风险研究进展方面有综无述。对于风险影响因素，为什么从宏观经济、金融市场、企业特征和债券特征四个方面去体现？这四个方面体现出现有研究的哪些特点？文中仅仅简单列表总结了现有成果。对于这些研究进展体现出了现有研究的哪些优势、哪些劣势以及具有哪些意义，都没有论述。

在预警模型部分同理。作为综述类文章来讲，文章中所总结的模型来源是哪里，模型经历过什么样的改进，针对了债券违约风险的什么问题，存在哪些优势与不足，这些问题在文中都没有回答。让人不清楚这些总结的贡献。

修改说明:

修改稿在风险影响因素部分的第一段指出，文章按照“从宏观到微观”、“从外部到内部”的梳理思路，从宏观经济、金融市场、企业特征、债券特征四方面切入，对代表性研究进行系统的归纳和总结，这样能够显得更有层次感。参见修改稿第3页。

修改稿进一步在影响因素部分的最后一段对风险影响因素研究进展的优劣势以及意义进行了论述，表述为：债券违约风险影响因素的相关研究较为全面，包括宏观经济、金融市场、企业特征、债券特征等方面。一方面，现有研究具有积极意义，表现在丰富的研究结论为监管部门提供了充分的理论依据，拓展了债券违约风险识别需要注意的方面。另一方面，现有研究也存在一定局限性，体现在部分国外学者的研究结论是否适用于中国市场有待检验。由于市场环境存在一定差异，同一因素对国内外债券市场违约风险的作用效果可能是不同的，不能盲目将国外学者的研究结论迁移到中国，在未来需要对中国市场开展针对性研究，为相关领域的探讨提供来自中国资本市场的独特证据。此外，影响因素的多元性在一定程度上给监管部门带来了无形的挑战，如何把众多影响因素纳入整体监管框架、构建更为有效的风险预警模型是值得思考的问题。参见修改稿第6页。

在预警模型部分，修改稿补充了所总结模型的文献来源，介绍了代表性模型（包括线性模型中的 Altman 模型、非线性模型中的 Merton 模型）的改进历程，指出其分别针对的是企业财务预警问题、债券风险定价问题。修改稿还在不同模型的介绍中分析了模型的优缺点和不足。参见修改稿第7页。

在预警模型的最后一段，本文进一步指出经典模型的两大不足之处，其一是学术界和业界尚未对最优模型的选择达成共识，不同模型的表达式差异较大，使用不同模型得到的违约风险测算结果可能相差甚远。其二是不同模型使用的指标各异、假设各不相同，考虑到债券违约风险影响因素的多样性，上述经典模型很可能存在遗漏重要变量、假设条件难以满足的问题，构建一个变量更为全面、形式更为灵活的预警模型是未来研究的重要方向。参见修改稿第 7-8 页。

审稿意见 (2.2):

未来研究方向部分，为什么选择从机器学习的视角切入？纵观全篇，未发现针对债券违约风险而应用机器学习的优势，缺少针对性理论分析。

修改说明:

修改稿在“利用机器学习算法推动预警模型创新”部分从理论上分析了使用机器学习模型预测债券违约风险相比于经典模型具有的三点优势，分别是充分拓展模型使用的信息集、有效规避现有模型的局限性、自然引入变量系数动态时变性。参见修改稿第 11-12 页。

审稿意见 (2.3):

“文本分析应用于债券违约风险领域的研究可以从三方面进行拓展”这一部分，提到要拓展至新闻媒体报道、社交媒体评论、政府报告等文本，而表 6 中已经清晰地呈现了已有大量此类研究。此处分析缺乏创新性，研究深度不足。

修改说明:

修改稿对这部分的论述进行了适当修改，指出大量的此类研究目前集中在股票市场和宏观经济领域，并没有在债券市场研究中得以应用，修改后的表述为：

目前，金融文本分析已经被广泛应用于股票市场、宏观经济研究，涉及的文本类型包括公司年报、新闻媒体报道、社交媒体评论、政府报告等。很多学者发现文本情绪等指标对股票市场、宏观经济具有显著影响（见表 6）。由于债券市场与股票市场具有联动性，也会受到宏观经济状况的影响，因此文本情绪等指标也很可能会影响债券市场，在未来有必要将上述研究拓展至债券市场，从文本分析这一机器学习视角更深入地挖掘债券违约风险的影响因素。

参见修改稿第 9 页。

第三份意见

审稿意见 (3.1):

是否有必要厘清债务违约、债券违约、财务预警等概念？例如，债券是公司债务的类型之一；z-model 实际是财务预警模型而不仅仅是衡量债务违约风险。

修改说明:

修改稿在影响因素部分指出“债券违约风险是企业债务违约风险的重要表现形式”，厘清了债务违约和债券违约的概念。参见修改稿第 2 页。

修改稿还在预警模型部分指出“Altman 模型侧重于分析企业财务预警问题，由该模型计算得到的企业财务预警得分可以预测债券违约风险”，厘清了财务预警和债券违约的概念。参见修改稿第 7 页。

审稿意见 (3.2):

关于债券违约风险影响因素部分，文章对各方面因素主要梳理了文献的具体变量和结果，可否在每部分对该因素影响所涉及的核心理论和关键逻辑作总结提炼。此外，列示文献中涉及了诸多不同的被解释变量，是否需要就违约风险的代理变量稍作说明（信用利差？融资成本？）。

修改说明:

修改稿在各方面影响因素梳理的第一段最后提炼了该方面因素影响债券违约风险的核心观点，有助于把握每部分内容的核心理论和关键逻辑。参见修改稿第 3-5 页。

修改稿在影响因素部分的第一段指出，“本文对债券违约风险影响因素的梳理主要参考债券信用利差相关文献，因为多数研究将信用利差作为债券违约风险的代理变量（史永东等，2021；杨国超和蒋安琪，2022）。此外，债券违约风险是企业债务违约风险的重要表现形式，企业融资成本等特征与违约风险相关性较大，因此在梳理影响因素过程中也借鉴了部分主题为企业债务风险、融资成本等文献的结论”，以此对违约风险的代理变量（信用利差、债务风险、融资成本等）进行了补充说明。参见修改稿第 2-3 页。

审稿意见 (3.3):

未来方向部分，一是可否考虑将文本分析提取影响因素和机器学习算法构建模型结合起来；二是使用机器学习方法的动机之一可补充传统评级膨胀的原因。

修改说明：

修改稿在未来方向部分指出，“可充分借鉴债券违约风险影响因素相关文献，构建包含宏观经济、金融市场、企业特征、债券特征等在内的原始预测指标库，还可以在指标库中加入文本分析提取的影响因素”。参见修改稿第 13 页。

修改稿在“机器学习算法推动预警模型创新”部分，在论述机器学习优势的第二点中补充了传统评级膨胀的原因，表述为：“国内债券市场的信用评级因受到‘发行人付费、发行人选择’选聘机制的影响，一定程度上持续存在数据粉饰、信息失称程度高的问题，在评级购买、评级迎合的作用下易导致评级膨胀现象。机器学习模型中能够输入充分全面的客观数据，尽可能避免数据选择性披露和人为因素带来的负面影响”。参见修改稿第 12 页。

审稿意见（3.4）：

机器学习方法应用展望部分，作异质性检验是否真的能拆解“黑箱”？该部分用语可以再斟酌。可考虑人为将经济学先验信息变量和机器学习方法相结合。

修改说明：

修改稿在机器学习方法应用展望部分删去了“黑箱”的表达，相关表述修改为“对机器学习预测背后的经济机制开展讨论”，并且指出“在运用机器学习模型时，还可以尝试人为将经济学先验信息变量和机器学习方法相结合，使模型更具经济学意义”。参见修改稿第 13 页。

提示：外审评议和作者回复（若有）是文章的过程性材料，本刊选择其中有价值的部分，隐名后予以公开。增强出版内容与纸刊正文的版权同属于《金融市场研究》，如研究中使用该内容，请遵守本刊版权声明。