



评级收费、评级监管与信用评级合理性

——基于实验研究的证据

朱松 黄海杰 许静静

摘要: 由于债券信用评级收费信息不公开, 因此评级收费如何影响信用评级合理性一直是个谜。基于对信用评级收费和评级合理性的实验研究, 本文发现: 评级收费与过度评级之间呈倒 U 型关系, 即较高的信用评级收费会诱导评级机构的过度评级, 但是更高的评级收费可能导致评级机构更加谨慎, 降低过度评级。监管能对过度评级现象起显著遏制作用: 随着监管强度的提高, 过度评级现象减少。而且, 监管的加强会提高评级机构的违约成本, 从而使得评级收费对评级合理性的影响拐点右移, 即更高的评级收费才能弥补因监管加强增加的违约成本。也就是说, 较严的监管不仅会直接遏制过度评级现象, 也会造成评级机构更加关注预期损失, 降低评级意见购买的可能。本文的研究为监管部门规范评级机构的信用评级收费及提高其信用评级合理性提供了参考和借鉴, 评级收费的公开化和透明化有利于信用评级的合理性。

关键词: 评级收费 信用评级合理性 过度评级 监管强度 实验研究

一、引言

2014 年的“11 超日债”结束了中国债券市场的“刚性兑付”, 自此债券违约事件时有发生。被认为安全的高等级 AAA 级债券也出现了违约^①, 国内信用评级的合理性及评级机构的客观公正不断受到各方的指责和质疑 (孔令强 等, 2021)。信用评

级的合理性与评级机构的专业能力及客观中立程度紧密相关。拥有评级资质的大型评级机构, 其专业能力是得到监管机构和市场认可的, 而且专业能力相差不大。因此客观中立程度决定了信用评级机构是否能够凭借其专业能力给出相对合理的信用评级水平。客观中立程度受到评级机构的

朱松 (通讯作者), 华东师范大学经济与管理学院会计系副教授; 黄海杰, 华东师范大学经济与管理学院会计系副教授; 许静静, 华东师范大学经济与管理学院会计系副教授。本研究得到国家自然科学基金“制度环境、代际传承与家族企业技术创新研究” (71802084)、教育部人文社科基金“国内信用评级机构的竞争策略对其市场地位的影响研究” (18YJA630156) 和上海市“科技创新行动计划”软科学研究项目“信用评级收费、评级监管与评级意见购买行为——基于实验研究的证据” (23692107500) 的资助。

① 违约主体上海某信初始主体评级即为 AAA。2016 年 12 月 29 日, 中诚信国际信用评级有限公司给予公司主体信用等级为 AAA, 中期票据信用等级也为 AAA; 2017 年 7 月 12 日联合资信评估有限公司给予公司主体信用等级为 AAA, 第二期中票信用等级也为 AAA。上海某信的债券评级从 2018 年开始被连续下调。3 个月时间, 联合资信和中诚信均将上海某信主体和相关债券的信用等级从 AAA 连续下调至 C。



职业道德标准及其对成本收益权衡的影响。其中，评级收费是评级机构的主要收益，很大程度上影响评级机构的独立客观性。但是，信用评级收费在国内外都是不公开的，评级收费如何影响评级质量及评级的合理性一直是个谜。在国内，激烈的市场竞争更是导致评级市场出现各种乱象（徐文鸣等，2022）——如东方某诚以5万元中标“2015年某省政府债券信用评级机构招标项目”，标价远低于项目预算55万元^①及行业收费均值25万元^②；某公国际则在为发行人提供评级服务的同时，直接向受评企业提供咨询服务并收取高额费用，且存在虚假表述和不实信息^③。评级公司的评级收费信息无法公开获取，导致研究者和监管机构不便对评级收费的合理性进行探究，监管机构也无法指导和监管评级机构的收费。

在债券违约事件发生后，国外投资者通常基于法律诉讼的途径寻求赔偿，监管机构则较少对评级机构进行处罚。2018年8月，美国证券交易委员会（SEC）首次发

起涉及信贷评级符号缺陷的执法行动，指控穆迪公司内部控制失责及其未能明确界定并始终如一地使用信贷评级符号，最终穆迪公司同意支付总计1.625亿美元的罚款。而在国内，由于法律诉讼机制尚不完善，金融领域对相关中介的管理和规范更多是直接依靠监管措施或者处罚来实现。国内债券违约自2014年开始，在违约事件发生不久，监管部门对少数评级机构的违规行为给予了在处罚性质和程度上均为较轻的警示处理^④（黄小琳等，2017）。随着国家多层次资本市场建设的逐步推进，债券市场在近年来进行了巨大的制度改革和监管改进。2019年12月26日，人民银行、发展改革委、财政部、证监会联合发布《信用评级业管理暂行办法》。2021年8月9日，人民银行会同发展改革委、财政部、原银保监会、证监会联合发布《关于促进债券市场信用评级行业健康发展的通知》，这些规定和政策将使我国债券市场和债券评级发展越发规范，评级收费乱象也将在一定程度上得到改变^⑤。而且，在实践中我们观

① 地方债评级5万元中标背后：价格战或将开打？[EB/OL].(2015-03-26)[2023-11-06].<https://www.yicai.com/news/4590644.html>.

② 2007年发布的《银行间债券市场信用评级机构评级收费自律公约》显示：企业主体信用评级收费不应低于10万元，债项信用评级收费不应低于15万元，单笔债券评级收费不应低于25万元；跟踪评级收费一般是5万元，数据更新收费也是5万元。

③ 该评级公司被处罚的主要原因为提供了较多的咨询服务并高额收费。但据相关报道（如见https://www.sohu.com/a/250327363_467340），该评级公司并没有提供有价值的咨询服务，因此其被处罚的原因是通过变相的咨询收费（实际上还是购买评级的费用）来弥补其评级收费，本质上还是评级收费问题。本文在后续展开研究问题讨论和具体实验设计时，为保持研究问题聚焦和重点突出，没有单独区分咨询业务收费，评级费用中包含了此类变相的评级购买费用。

④ 如：深圳证监局，《关于对某元资信评估有限公司采取出具警示函措施的决定》，2013-08-02；中国证监会，《2015年证券评级机构现场检查情况通报》，2016-03-21。

⑤ 2021年5月中旬，中国信用评级行业重新签署行业自律公约，约定在2007年自律公约的基础上提高评级收费价格，以适应债券市场非强制评级的改革措施。据悉，修改后收费标准如下：企业主体信用评级收费不应低于25万元，债项信用评级收费不应低于20万元，单笔债券评级收费不应低于45万元（见<https://www.163.com/dy/article/GBIQLSD105198CJN.html>）；跟踪评级收费不应低于10万元，数据更新收费不应低于10万元。



察到对评级机构的监管和处罚也越来越严格和越来越频繁。2018年8月17日,某公国际被暂停债券、证券评级资格各一年。2020年和2021年东方某诚被责令整改、暂停评级业务,并处罚款594.5万元。强化对评级机构的监管和处罚能否带来信用评级质量的提升是现阶段国内监管机构迫切关注的问题。

因此,本文希望通过采用实验研究的方法设计不同信用评级收费下的信用评级情景,模拟实验对信用评级收费及其对评级合理性的影响进行探讨,弥补现实中无法获得评级收费信息而对评级合理性研究不足的问题;同时,实验加入了评级监管机构对评级失败的监管处罚,为评级机构监管从而抑制评级意见购买提供理论指导和借鉴。

二、理论分析与研究假设

(一) 文献回顾

经过相关文献回顾发现:在当下主要基于“发行人付费”的背景下(吴育辉等,2020),由于利益的驱动,评级机构会迎合发行人出具低质量的评级;而竞争机制、声誉机制、监管都无法很好抑制发行人的低质量评级。无论是低质量评级的出现,还是制约机制的失效,文献的讨论都是基于“发行人付费”模式下,评级机构对利益的追求。评级机构最主要的利益就是发行人支付的评级费用,正是出于对高评级费用的追求,才会导致低质量评级的出现。正如夏凡和姚志勇(2013)在解释

信用评级机构“顺周期”行为时认为,评级机构在评级收费与未来声誉成本之间做出权衡,如果评级收费足够高,评级机构将会高估评级从而收取评级费用。然而由于缺乏数据,目前并没有关于评级收费对评级质量影响的直接证据,本文将通过实验研究的设计对这一问题展开探讨,为“利益驱动”导致低质量评级提供直接证据。

具体文献综述详见附录1部分。^①

(二) 假设发展

信用评级机构会综合考虑信用评级业务的成本与收益,可能机会主义地配合被评级企业,给予其预期的信用评级水平,即实现被评级企业的“评级意见购买”。促使评级机构配合被评级企业的评级意见购买行为的因素包括更高的报酬(评级收费)、更高的市场占有率、维护和客户的关系等,而更高的报酬是最直接,也是最重要的一个因素。但是,信用评级机构的劳动付出、声誉补偿以及潜在损失等都会随着评级收费的提升而获得补偿,评级机构会对较高的评级收费更为重视。在各种成本得以补偿的情况下,评级机构不会为了再高一些的评级收费而冒更大的风险,其机会主义动机会下降,从而保证信用评级的合理性。

另外,声誉是中介机构最为重要的资产之一(Guembel & Rossetto, 2009),也是信用评级行业有效运行的重要保障。对声誉的重视则使得评级机构倾向于更加稳健的评级和出具更多的负面报告(Mariano,

① 附录1~7见增强出版,中国知网—《金融市场研究》。



2012 ; Atilgan et al., 2015)。如果评级机构的声誉成本较低，他们会倾向于给予被评级企业较高的信用评级，在赢得较高评级收费客户的情况下，评级机构能够提升其在行业内的声誉和影响力。一旦因违规而失去高质量且高收费的客户，会对评级机构产生更大的影响。因此，评级机构不会为了再高一些的评级收费而冒更大的风险，从而保证信用评级的合理性。

当严格监管和惩罚导致声誉成本提高时，评级机构则会提升信用分析的及时性和准确性 (Cheng & Neamtiu, 2009)。Becker & Milbourn (2011) 认为发行人付费模式下的评级制度存在一定缺陷，除非竞争不是那么激烈，否则会引发“评级购买” (Rating Shopping) 问题 (Bolton et al., 2012)，推高信用评级水平 (Alp, 2013 ; Bonsall IV, 2014)。监管弱化会降低声誉成本，进而导致信用评级机构迎合发债企业现象增多 (Kraft, 2015)。实际上，抑制评级机构的投机行为的因素有职业道德、监管强度 (即被查处的概率和被查处后的处罚方式)、声誉受损、客户流失等。由于职业道德难以衡量，而声誉受损和客户流失都在很大程度上与监管相关，或者说监管是造成声誉损失和客户流失的重要原因。因此，可以说监管是评级机构衡量成本收益时最重要的一个因素。

本文采用一个简单的模型来发展研究假设，本文模型基于 Bolton et al. (2012)、Opp et al. (2013)、周香芸等 (2019) 的博弈模型发展而来。具体模型假设和理论推演请见附录 2 部分。

类似于审计意见或法律意见，信用评级意见也存在很大程度的专业判断，合理水平的认定在不同机构之间存在一定差异。更重要的是，审计意见有参考的标准 (审计准则和财务报告准则)，而信用评级并没有规范或者信用水平没有相对统一的标准；虽然信用评级意见有市场公认的合理范围或者水平，但只要评级意见不是特别离谱，信用评级意见是很难被证明不合理的。另外，监管机构可以依靠其专业监管团队或者其他专家，对审计意见或者财务报告做复核性评判，以说明审计意见的合理性；但评级意见在复核性评判中更加困难，因为每家评级机构都有不同的信用评级模型，信用评级水平可以是在一定区间范围的，相对于审计意见的非黑即白 (标准意见和非标意见)，信用评级意见更具有操纵空间和灵活性。评级机构会综合考虑成本与收益，机会主义地配合被评级企业给予其预期的信用评级水平，即评级意见购买。激励评级机构配合被评级企业评级意见购买行为的因素包括更高的报酬 (评级收费)、更高的市场占有率、维护和客户的关系等，而更高的报酬是最直接，也是最重要的一个因素。在成本一定的情况下，评级收费越高，评级机构的机会主义动机就越强。

但是，信用评级机构劳务付出、声誉补偿、潜在损失等都会随着评级收费的提升而获得补偿，评级机构会对较高的评级收费更为重视。随着收费的增加，客户对评级机构的重要性也在增加。这种重要客户一旦评级出现问题，评级机构面临的成



本将更大, 包括负面新闻导致更多市场关注带来的隐性成本以及因高额收费导致的更严厉的处罚等。因此, 随着客户重要性的提升或者评级收费占评级机构收入较大比重, 评级机构反而会更加谨慎。而且, 不论是老客户还是新客户, 较高的评级收费, 尤其是超过评级付出的合理成本以及声誉溢价之上的费用, 往往意味着需要满足客户的评级需求, 即给予客户超过其合理评级的评级水平。这部分超额的收费越高, 意味着超过合理评级的程度就越大, 对于评级机构而言风险就越大。一旦该债券出现违约等严重事件, 会影响到评级机构的声誉, 以及由于监管处罚带来的其他损失。因此, 虽然较高的收费对于评级机构具有很大的诱惑力, 但是在更高的预期损失面前, 尤其是在监管处罚极为严厉的情况下, 越高的评级收费对评级机构而言反而更多的是预期损失超过收益, 评级机构会变得更加谨慎, 降低对于发债企业的评级迎合, 减少过度评级行为。因此, 评级收费与过度评级之间呈现先增后减的关系。因此, 本文提出假设 1。

假设 1: 评级收费与过度评级呈倒 U 型关系, 即随着评级收费的提高, 过度评级首先呈上升趋势; 当评级收费达到一定水平后, 过度评级呈现下降趋势。

对于评级机构而言, 如果配合被评级企业的评级意见购买行为, 给予被评级企业过于乐观的信用评级, 一旦被评级企业出现债券违约, 就会影响该评级机构的声誉, 更重要的是, 可能还会受到来自监管机构的关注。在国内监管较多的情况下,

会出现评级机构被监管机构约谈, 要求其进行整改, 可能也会被处以高额罚金; 更为严厉的, 还可能被监管机构暂停其债券评级资质, 严重影响其未来业务发展。约谈往往相对轻微, 会对评级机构的声誉产生一定的负面影响, 但在目前声誉机制并没有起到作用的市场环境下, 这种监管处罚方式对评级机构的影响有限。处以罚金意味着监管力度的增强, 对评级机构带来了实质性的处罚, 不仅造成更大的负面声誉影响, 而且也造成了更多的经济损失。而要求额外评级以及暂停评级资质就更为严厉, 是监管机构对评级机构不信任的表现, 也是监管最为严格的一种手段, 必然给评级机构带来声誉和实际经济利益的影响。随着监管强度的增强, 评级机构的机会主义行为所产生的违规成本也就越高, 评级机构进行机会主义行为的可能性就越低。因此, 本文提出假设 2。

假设 2: 随着监管强度的提高, 过度评级现象减少。

三、研究设计

(一) 实验设计

本实验目的在于探究监管对评级机构风险判断的影响。本实验在学校行为实验室进行, 被试同时进行实验。实验对象为国内某重点大学工商管理学院 MPAcc 专业硕士学生。实验为情景模拟实验, 被试在实验中扮演评级机构高层的角色, 有权决定对被评级单位出具不同水平的信用等级。实验为连续实验, 共分为 10 轮, 每轮结束后产生的后果将影响下一轮实验, 即



本轮获得报酬、被查处后的罚款金额不清零，累积到下一轮，直到 10 轮实验结束后计算总报酬。被试在实验中最终获得的报酬与被试实际得到的被试费挂钩，提高实验的真实感，增强了被试参与实验的体验感和积极性。具体的实验设计和实验流程请见附录 3。

(二) 模型与变量

为检验评级收费、评级监管对评级意见购买行为（过度评级或信用评级合理性）的影响，本文模型设定如下：

$$ResRating = \beta_0 + \beta_1 Fee + \beta_2 Feesq + \beta_3 Regulation + \beta_4 Wealth + \beta_5 StandRating + \beta_6 Learning + \text{个人固定效应} + \varepsilon \quad (1)$$

ResRating 为被评级企业与评级机构之间的评级意见购买行为，采用实际评级水平与合理信用评级之间的差异表征，该指标为正意味着过度评级（存在评级意见购买），指标为零意味着合理评级（不存在评级意见购买），指标为负意味着评级低估。

评级收费类似于审计收费，收费在很大程度上取决于工作量及承担的风险。本文在实验中对工作量及基本风险做了相应控制和处理，因此评级机构给出的评级收费实际上相当于因承担额外风险而收取的费用。Fee 为评级机构给出信用评级水平后收到的评级费用，采用实验中的实际给出的评级水平对应的评级收费表征。为检验评级收费与过度评级之间的非线性关

系，本文加入了 Fee 的二次方项 Feesq。为降低变量的共线性问题，回归中采用评级收费取自然对数，并对其进行中心化。

Regulation 为评级监管变量。本文采用监管强度 (Reguforce) 表征监管部门对信用评级机构在债券出现违约后的监管措施，0 表示公司在债券存续期间出现违约时，信用评级机构不会被监管机构警告或者处罚；1 表示公司在债券存续期间出现违约时，信用评级机构会被监管约谈，要求提交评级资料，并进行整改；2 表示公司在债券存续期间出现违约时，信用评级机构会被监管约谈，要求提交评级资料，进行整改，没收评级收费并处 3 倍罚金。该指标越大，意味着监管强度越高，监管越严格。

关于评级意见购买、评级收费以及评级监管等关键变量的具体度量以及采取此度量方式的详细原因，请见附录 4 部分。

声誉是影响评级机构合理评级的重要机制，越是大型的评级机构往往越重视其声誉。因此，本文采用每轮实验开始前评级机构的累积财富 (Wealth) 表征其声誉。Wealth 为评级机构在上一轮结束时的累计收益，回归中采用累积收益的自然对数^①。

考虑到被测试者的风险厌恶程度以及对政策的预期会与实验的进度有一定的关系，即被测试对象存在学习曲线 (Learning)，因此，回归中采用实验的轮次表征学习曲线，若实验为第一轮，则

① 考虑到实验过程中有被测试者出现负资产的情况，回归中采用的是累计收益加 46（实验中的最低累计收费+1）然后取对数。



Learning 为 1, 若为第十轮, 则 Learning 为 10, 以此类推。

考虑到个人性别、年龄等造成的个体风险差异, 回归中本文也控制了被测对象的固定效应。

(三) 样本与数据

本文的样本均来自实验, 实验针对某重点大学工商管理学院 36 名 MPAcc 专业硕士学生。每组实验均为十轮, 每轮会随机出现一个被评级单位的背景信息并要求被测试者选择评级意见, 所以每个被测试者会做出十个选择。实验在某重点大学工商管理学院实验室上机测试。36 名实验对象, 每组实验会出现 10 个实验结果, 最终样本为 360 个实验样本。

四、实证分析

(一) 描述性统计

表 1 展示了不同信用评级收费水平下的过度评级情况。全部 360 个信用评级中, 有 4 个信用评级属于评级低估 (ResRating=-1), 占比 1%。100 个信用评级属于合理评级 (ResRating=0), 占比 28%。71% 的信用评级属于过度评级 (ResRating>0), 而且更激进的评级占比更高, ResRating=3 的信用评级占比 46%。总体而言, 过度评级的情况比较严重。

对比评级收费和过度信用评级的趋势可以看出, 评级收费越高, 过度信用评级越高; 评级收费越低, 反而容易引发评级低估。这意味着较高的信用评级收费很可能诱使评级机构给出激进的信用评级。

表 1 信用评级收费与过度评级

| Fee | ResRating | | | | | Total |
|-----|-----------|-----|-----|-----|-----|-------|
| | -1 | 0 | 1 | 2 | 3 | |
| 8 | 1 | 0 | 0 | 0 | 0 | 1 |
| 9 | 3 | 0 | 0 | 0 | 0 | 3 |
| 10 | 0 | 100 | 0 | 0 | 0 | 100 |
| 15 | 0 | 0 | 3 | 0 | 0 | 3 |
| 19 | 0 | 0 | 9 | 0 | 0 | 9 |
| 20 | 0 | 0 | 4 | 0 | 0 | 4 |
| 21 | 0 | 0 | 7 | 0 | 0 | 7 |
| 22 | 0 | 0 | 1 | 0 | 0 | 1 |
| 29 | 0 | 0 | 11 | 15 | 0 | 26 |
| 30 | 0 | 0 | 7 | 14 | 0 | 21 |
| 31 | 0 | 0 | 0 | 5 | 0 | 5 |
| 32 | 0 | 0 | 0 | 4 | 27 | 31 |
| 39 | 0 | 0 | 0 | 2 | 48 | 50 |
| 40 | 0 | 0 | 0 | 6 | 28 | 34 |
| 41 | 0 | 0 | 0 | 1 | 0 | 1 |
| 42 | 0 | 0 | 0 | 0 | 34 | 34 |
| 54 | 0 | 0 | 0 | 0 | 30 | 30 |
| 55 | 1 | 0 | 0 | 0 | 0 | 1 |
| | 4 | 100 | 42 | 47 | 167 | 360 |
| | 1% | 28% | 12% | 13% | 46% | |

表 2 展示了不同监管强度下的过度评级情况。在债券存续期间出现违约时, 信用评级机构不会被监管机构警告或者处罚的情况下 (Reguforce=0), 没有出现企业的信用评级被低估的情况 (占比为 0%), 4% 的信用评级属于合理评级, 而 96% 的信用评级则属于过度评级, 而且更激进的评级 (ResRating=3) 的占比更是高达 72%。

在债券存续期间出现违约时, 信用评



级机构会被监管约谈，要求提交评级资料 and 进行整改的情况下 (Reguforce=1)，依旧没有出现企业的信用评级被低估的情况 (占比为 0%)，6% 的信用评级属于合理评级，而信用评级则属于过度评级的占比下降为 94%，而且更激进的评级 (ResRating=3) 的占比也下降到 61%。

在债券存续期间出现违约时，信用评级机构会被监管约谈，要求提交评级资料，进行整改，没收评级收费并处以 3 倍罚金的情况下 (Reguforce=2)，合理评级的占比立刻提高到 63%，而过度评级占比进一步下降到 34%，更激进的评级 (ResRating=3) 的占比也进一步下降到 13%。在这种严监管的情况下，甚至出现了信用评级被低估的情况 (评级低估占比 3%)。

以上趋势表明，随着监管强度的增强，过度评级的情况明显下降，尤其是更为激进的评级会减少。

表 2 监管强度与过度评级

| Reguforce | ResRating | | | | | Total |
|-----------|-----------|-----|-----|-----|-----|-------|
| | -1 | 0 | 1 | 2 | 3 | |
| 0 | 0 | 6 | 12 | 22 | 104 | 144 |
| | 0% | 4% | 8% | 15% | 72% | |
| 1 | 0 | 4 | 8 | 16 | 44 | 72 |
| | 0% | 6% | 11% | 22% | 61% | |
| 2 | 4 | 90 | 22 | 9 | 19 | 144 |
| | 3% | 63% | 15% | 6% | 13% | |
| 3 | 4 | 100 | 42 | 47 | 167 | 360 |
| | 1% | 28% | 12% | 13% | 46% | |

表 3 展示了变量的描述性统计情况。过度评级 (ResRating) 平均为 1.758，意味着以上实验中的企业的信用评级总体上是被高估的，不过信用评级存在被低估的情况 (ResRating 最小值为 -1)。信用评级收费 Fee 平均为 32.677 万元，最低为 8 万元，最高为 55 万元。预期的处罚金额均值为 20.711 万元，最大值为 220 万元 (评级收费的 4 倍^①)。被测试者每轮测试时的初始财富平均为 142.341 万元，最低为 -45 万元 (上一轮存在债券违约而评级机构被处以高额罚金的情况)，最大为 419 万元。

测试中存在 26.9% 的样本出现了债券违约的情况，即 360 个测试样本中有 97 例债券违约情况。

(二) 回归分析

表 4 给出了评级收费、监管强度对信用评级合理性 (过度评级) 的影响。第 (1) 列和第 (2) 列检验评级收费对信用评级合理性的影响，审计收费变量 Fee 的回归系数在第 (1) 列和第 (2) 列中都在 0.01 水平上显著为正，第 (2) 列加入评级收费的平方项后，评级收费二次项 Feesq 的系数显著为负，评级收费 Fee 的系数显著为正，即评级收费 Fee 与过度评级之间呈倒 U 型关系；而且，拐点在第一象限，即信用评级收费为正的某个水平。具体而言，随着信用评级收费的提高，过度评级情况会愈发严重；但是当信用评级收费达到一定水平后，过度评级则会呈现下降趋势。本文的假设 1 得到验证。以上结果意味着较高

① 若受到处罚，则罚金为该次评级的收费+3倍罚金，即4倍评级收费。



表 3 变量描述性统计

| 变量名 | 变量 | 样本量 | 均值 | 标准差 | 最小值 | 中位数 | 最大值 |
|----------|-------------|-----|--------|--------|-----|------|-----|
| 评级意见购买 | ResRating | 360 | 1.758 | 1.318 | -1 | 2 | 3 |
| 评级费用 | Fee | 360 | 32.677 | 14.336 | 8 | 35.5 | 55 |
| 监管措施 | Reguforce | 360 | 1 | 0.895 | 0 | 1 | 2 |
| 监管约谈 | Talk | 360 | 0.2 | 0.401 | 0 | 0 | 1 |
| 监管处罚 | TalkPanelty | 360 | 0.4 | 0.490 | 0 | 0 | 1 |
| 预期处罚金额 | ExpPanelty | 360 | 20.711 | 53.647 | 0 | 0 | 220 |
| 合理信用评级水平 | StandRating | 360 | 12 | 6.008 | 6 | 12 | 18 |
| 债券违约 | Default | 360 | 0.269 | 0.444 | 0 | 0 | 1 |
| 实验轮次 | Learning | 360 | 5.5 | 2.876 | 1 | 5.5 | 10 |

表 4 评级收费、监管强度与过度评级

| 变量 | 预期符号 | Resrating | | | | |
|-----------------------|------|----------------------|-----------------------|-----------------------|---------------------|----------------------|
| | | (1) | (2) | (3) | (4) | (5) |
| Fee | + | 52.005*** (9.01) | 45.535*** (9.30) | | 64.130*** (3.81) | 7.041*** (5.30) |
| Feesq | - | | -35.831*** (-7.63) | | -42.634* (-1.95) | -0.071*** (-5.34) |
| Reguforce | - | | | -2.041*** (-12.26) | -1.746** (-2.56) | -2.021*** (-3.13) |
| Wealth | + | -4.490*** (-3.48) | -4.352*** (-3.40) | -1.293* (-1.87) | -3.897** (-2.03) | -3.820** (-1.98) |
| StandRating | ? | 0.111** (2.40) | 0.070 (1.57) | 0.082*** (4.11) | 0.143* (1.86) | 0.181*** (2.67) |
| Learning | - | 0.669*** (2.84) | 0.631*** (2.74) | 0.128 (1.00) | 0.512 (1.57) | 0.490 (1.48) |
| 个人固定效应 | | 控制 | 控制 | 控制 | 控制 | 控制 |
| N | | 360 | 360 | 360 | 360 | 360 |
| Pseudo R ² | | 0.8950 | 0.9045 | 0.3130 | 0.9183 | 0.9168 |

注：被解释变量为过度评级 ResRating；模型为 Ologit 模型；***p<0.01, **p<0.05, *p<0.1；括号内为 Z 值。

的信用评级收费会诱导评级机构的过度评级，但是评级机构也会注重自身的声誉以及潜在的违规风险，更高的评级收费可能

导致评级机构更加谨慎，降低过度评级。因此，对于提高评级收费不能完全否定，合理的评级收费才能保证合理的信用评级



水平。

第(3)列单独检验了监管强度对信用评级合理性(过度评级)的影响。可以看到,监管强度变量 *Reguforce* 的系数在 0.01 水平上显著为负,意味着监管强度越大,过度评级的水平越低,符合预期,支持了本文的假设 2。

第(4)列综合了评级收费和监管强度,综合考察其对信用评级合理性(过度评级)的影响。评级收费变量 *Fee* 的系数依旧显著为正,二次方项 *Feesq* 的系数依旧显著为负,与前两列的结果一致,即评级收费与过度评级之间呈倒 U 型关系,进一步支持了本文的假设 1。监管强度变量 *Reguforce* 的系数也依旧在 0.01 水平上高度显著为负,意味着监管强度的提高会降低过度评级,进一步支持了假设 2。

为了更好体现评级收费与过度评级之间倒 U 型关系的经济意义,第(5)列采用未对数化的评级收费(实验中的实际给出的评级水平对应的评级收费)进行回归。可以看到,评级收费 *Fee* 和 *Feesq* 的回归系数的显著性与符号与前几列结果也都一致,即评级收费 *Fee* 与过度评级依旧呈现倒 U 型关系。此时,倒 U 型的拐点为 50.604 (即 $5.364 / (2 \times 0.053)$),即当评级收费低于 50.604 万元时,评级收费越高,过度评级水平越高;但当评级收费高于 50.604 万元时,过度评级水平反而下降,评级机构在评级收费与潜在的成本之间做出权衡,不会因更高的收费而给予更高的过度评级。假设 1 进一步得到验证。

评级机构的初始财富 *Wealth* 的回归系

数基本不显著,意味着初始财富对其风险判断的影响不大。可能的原因在于实验中初始财富以及每轮获得的收益并不高。被评级企业的合理信用评级水平 *StandRating* 的回归系数显著为正,意味着过度评级也会基于合理评级水平发生变化,合理评级越高的情况下,评级机构可能更加倾向于给出更高的评级水平,因为合理评级水平意味着被评级企业违约的可能性,合理评级水平越高,违约的可能性越低,过度评级被察觉的可能性就越低,从而导致评级机构机会主义投机行为更加明显。学习曲线 *Learning* 的系数不显著,意味着被测试者并没有在后面的实验中更谨慎,可能的原因在于影响其风险判断的更多是评级收费和监管处罚问题,而每轮实验都是随机的,难以体现出一定的趋势和规律。

总之,表 4 的结果表明:评级收费 *Fee* 与过度评级之间呈倒 U 型关系,即较高的信用评级收费会诱导评级机构的过度评级,但是更高的评级收费可能导致评级机构更加谨慎,降低过度评级。而监管强度的提高能够有效降低和抑制过度评级问题,也会造成评级机构更加关注预期损失,降低评级意见购买的可能。

监管的加强是否也影响了评级机构在评级意见购买中的风险判断,即影响到评级收费对评级合理性的作用上?对此,本文进一步进行了检验,结果见表 5。首先,本文对全样本、无监管样本和有监管样本分别进行回归,比较评级收费的拐点,以此来评价监管和评级收费给评级合理性带来的交互影响。



表 5 评级收费与监管强度的交叉影响

| 变量 | Resrating | | | | |
|-----------------------|------------|-----------|------------|-----------|----------|
| | 全部样本 | 无监管样本 | 有监管样本 | 全部样本 | 全部样本 |
| | (1) | (2) | (3) | (4) | (5) |
| Fee | 45.535*** | 76.065*** | 52.496*** | 78.910** | 95.053 |
| | (9.30) | (2.71) | (4.23) | (2.20) | (1.27) |
| Feesq | -35.831*** | -64.277** | -33.249*** | -69.004** | -87.963 |
| | (-7.63) | (-2.43) | (-2.63) | (-2.09) | (-1.13) |
| Reguforce | | | | -3.424** | -3.284** |
| | | | | (-2.17) | (-2.10) |
| Reguforce × Fee | | | | 8.403* | -0.373 |
| | | | | (1.72) | (-0.04) |
| Reguforce × Feesq | | | | | 24.746 |
| | | | | | (0.69) |
| Wealth | -4.352*** | -13.340** | -5.658 | -3.975* | -4.113* |
| | (-3.40) | (-2.17) | (-1.58) | (-1.94) | (-1.88) |
| StandRating | 0.070 | 0.214** | 0.105 | 0.165* | 0.179 |
| | (1.57) | (2.22) | (1.47) | (1.96) | (1.53) |
| Learning | 0.631*** | 1.896* | 0.877 | 0.479 | 0.498 |
| | (2.74) | (1.87) | (1.36) | (1.44) | (1.45) |
| 个人固定效应 | 控制 | 控制 | 控制 | 控制 | 控制 |
| N | 360 | 360 | 360 | 360 | 360 |
| Pseudo R ² | 0.9045 | 0.9214 | 0.9268 | 0.9257 | 0.9262 |

注：被解释变量为过度评级 ResRating；模型为 Ologit 模型；*** p<0.01, **p<0.05, *p<0.1；括号内为 Z 值。

在前三列的回归中，评级收费 Fee 和 Feesq 的回归系数都显著，且与之前结果一致，此时评级收费 Fee 与过度评级倒 U 型关系的拐点同样在第一象限，进一步支持了假设 1。而第 (1) 列全样本中的拐点为 0.64 [即 $45.535/(2 \times 35.831)$]，即随着评级收费的自然对数 lFee 达到 0.64 后，评级收费越高，过度评级会下降。第 (2) 列

为对没有监管措施的样本 (Reguforce=0) 进行的回归，也就是说即使出现债券违约（不论是否存在评级虚高），信用评级机构也不会受到任何监管措施。第 (2) 列无监管措施样本的拐点为 0.59 [即 $76.065/(2 \times 64.277)$]。第 (3) 列为有监管措施的样本 (Reguforce>0) 进行的回归，只要出现债券违约且存在评级虚高，信用评级机



构就会受到监管措施，这种情况下的拐点为 0.79 [即 $52.496 / (2 \times 33.249)$]。也就是说，全样本回归的拐点较无监管措施样本的拐点更大。这意味着随着监管强度的提高，评级收费 Fee 与过度评级的倒 U 型关系进一步右移，即拐点向右平移。本文认为可能的原因在于：相对于没有监管的情况，由于存在较强的监管，评级机构会要求更高的评级收费从而补偿过度评级可能带来的损失。换言之，在存在较强的监管情况下，与没有监管相比，同样的评级收费可能无法达到评级机构对预期损失的补偿预期，因此就不太可能出现预期中的过度评级。以上结果表明，信用评级中较高的监管不仅会直接遏制过度评级现象，也会造成评级机构更加关注预期损失，降低评级意见购买的可能。

（三）稳健性检验

本部分进行稳健性检验。①区分了不同的监管方式衡量评级监管；②利用预期处罚金额表征评级监管；③直接采用评级机构实际给出的信用评级为被解释变量；④采用信用评级过度收费替换评级收费变量；⑤剔除中途破产样本；⑥区分高违约率和低违约率样本；⑦考虑参与者预期的影响等，得到的结果与主要回归分析的结论一致。具体结果和讨论详见附录 5 稳健性检验部分。

（四）进一步检验：债券违约后的评级收费与过度评级

本部分进行进一步检验，主要讨论了

债券违约后，评级机构的评级收费与过度评级之间的关系是否受到影响。基于实验数据的组间差异比较及回归分析发现：上一轮违约以及违约后的监管方式对之后的信用评级都没有产生明显的影响；对评级风险判断影响最大的依旧是评级收费和预期的处罚金额。这一结果与黄小琳等 (2017) 一致，可能的原因在于现行制度环境尤其是监管强度较轻，不足以从实质上促使评级机构改变其过度评级的动机。只有更加严厉的评级监管措施，如被监管机构禁止从业，或许能够较好地约束评级机构的激进行为。如 2014 年“11 超日债”打破中国债券市场的刚性兑付后，监管机构并没有对评级机构表现出严格的监管态度和措施。而 2018 年 8 月某公国际被暂停债券、证券评级业务各一年^①，监管态度和措施才严格起来。具体结果和讨论详见附录 6 进一步检验部分。

五、结论

本文基于实验研究发现：评级收费与过度评级之间呈倒 U 型关系，即较高的信用评级收费会诱导评级机构的过度评级，但是更高的评级收费可能导致评级机构更加谨慎，降低过度评级。不同的监管方式对抑制过度评级的效果不同，简单的监管约谈不仅不会起到预期效果，反而会刺激过度评级；存在实质性处罚（罚金或禁止从业等）对抑制过度评级起到了显著作用，

① 北京证监局依据《暂行办法》拟责令大公国际限期整改，期限一年，整改期间不得承接新的证券评级业务，更换不符合条件的高级管理人员。



预期处罚金额越高,越会加大评级机构机会主义行为的违约成本,从而降低其机会主义动机,提高信用评级的合理性。而之前是否违约及采用何种监管方式对之后的信用评级合理性都没有产生明显的影响,评级机构并没有表现出更为谨慎的做法。对于评级机构风险判断影响最大的,还是当期的评级收费和预期处罚方式(处罚金额)。本文的研究结果为评级机构的监管和引导提供了一定的借鉴和参考,也为评级收费的公开化和透明化需求提供了理论基础和实证支持。

由于本文采用的是实验研究方法,被测试者并不是评级机构从业人员,因此不可避免地存在不完全符合现实从业人员判断的可能。不过被测试者为会计专业硕士(MPAcc),有较强的会计财务学知识基础,了解债券评级机构的内容和职责;其中一些MPAcc专业硕士学生也有一定的相关行业的从业或实习经验,因此其认知偏差对实证结果的干扰较少。基于实验研究被试

对象本身参与实验积极性不够导致参与度较低的问题,本文采取了一定的激励措施,增强被试对象参与实验的体验感和积极性。虽然有上述实验对象的选择和实验过程的控制在一定程度上削弱了被试认知偏差对实证结果的干扰,但不可否认,实验模拟距离具体实务操作依然有一定距离,这也是本文的局限性之一。另外,现实中每家评级机构的评级模型和标准会有一些的差异,评级机构的专业程度和评估师的胜任能力也会有所差别,被评级企业的财务和非财务信息存在不好识别的问题,导致比实验更为复杂的决策情景,评级机构也可能会因为获取市场而出现虽然收费低但仍然给予企业高评级的现象,这也会导致结果存在一定的偏误。最后,由于实验设定的监管处罚中并没有极其严厉的禁业措施,因此对于之前的监管措施是否能够更好地发挥事前威慑作用没有能够较好地验证,后续实验有待改善^①。N

学术编辑:卢超群

参考文献

- [1] 黄小琳,朱松,陈关亭.债券违约对涉事信用评级机构的影响——基于中国信用债市场违约事件的分析[J].金融研究,2017(03):130-144.
- [2] 孔令强,黄雨昕,俞春江.完善中国评级机构公司治理,促进行业高质量发展[J].金融市场研究,2021(05):62-75.
- [3] 徐文鸣,杜宇轩,王磊.全面注册制时代信用评级监管的改革经验[J].金融市场研究,2022(10):15-23.
- [4] 吴育辉,翟玲玲,张润楠,等.“投资人付费”vs.“发行人付费”:谁的信用评级质量更高?[J].金融研究,2020(01):130-149.
- [5] 夏凡,姚志勇.评级高估与低估:论国际信用评级机构“顺周期”行为[J].金融研究,2013(02):184-193.

① 关于更加严格的监管措施是否有效的问题,本文也通过调查问卷的方式对36名被试者在出现债券违约后的风险倾向进行了统计。如果受到两次处罚则会被监管机构禁止从业,有94.44%的被试者会变得谨慎,但依旧有2.78%的被试者评级态度不会发生变化,而且有2.78%的被试者会更加激进地评级。这一结果显示,相较于高额的处罚金额,禁业的处罚措施会使评级机构变得谨慎起来,但是依旧会有机会主义行为的出现。



- [6] 周香芸,田益祥,李立,等.新评级监管对中国企业债券信用评级的声誉效应——基于新评级选购博弈模型[J].金融学季刊,2019,13(01):151-175.
- [7] Alp A.Structural Shifts in Credit Rating Standards[J].Journal of Finance,2013,68(6):2435-2470.
- [8] Atilgan Y,Ghosh A,Yan M,et al.Cross-Listed Bonds, Information Asymmetry, and Conservatism in Credit Ratings[J].Journal of Money, Credit and Banking,2015,47(5):897-929.
- [9] Becker B,Milbourn T.How Did Increased Competition Affect Credit Ratings?[J].Journal of Financial Economics,2011,101(3):493-514.
- [10] Bolton P,Freixas X,Shapiro J.The Credit Rating Game[J].Journal of Finance,2012,67(1):85-111.
- [11] Bonsall IV S B.The Impact of Issuer-Pay on Corporate Bond Rating Properties: Evidence from Moody's and S&P's Initial Adoptions[J].Journal of Accounting and Economics,2014,57(2-3):89-109.
- [12] Cheng M,Neamtiu M.An Empirical Analysis of Changes in Credit Rating Properties: Timeliness, Accuracy and Volatility[J].Journal of Accounting and Economics,2009,47(1-2):108-130.
- [13] Guembel A,Rossetto S.Reputational Cheap Talk with Misunderstanding[J].Games and Economic Behavior,2009,67(2):736-744.
- [14] Kraft P.Do Rating Agencies Cater? Evidence from Rating-Based Contracts[J].Journal of Accounting and Economics,2015,59(2-3):264-283.
- [15] Mariano B.Market Power and Reputational Concerns in the Rating Industry[J].Journal of Banking & Finance,2012,36(6):1616-1626.
- [16] Opp C C,Opp M M,Harris M.Rating Agencies in the Face of Regulation[J].Journal of Financial Economics,2013,108(1):46-61.

Rating Fees, Regulatory Intensity and Credit Rating Rationality

—Evidence Based on Experimental Research

ZHU Song HUANG Haijie XU Jingjing

(Department of Accounting, School of Economics and Management, East China Normal University)

Abstract Since credit rating fees are not publicly disclosed, the effect of rating fees on credit rating decisions has long been something of a mystery. Based on an experimental study of credit ratings and rating rationality, this paper finds that there is an inverted U-type relationship between rating fees and over-rating. Higher credit rating fees can induce over-rating by rating agencies, though they can also lead to more cautious behavior by heightening rating agency concerns over potential loss of business. The existence of substantial penalties, such as curbs on conducting business, for example, has played a significant role in suppressing excessive ratings. The higher the expected penalty amount, the greater the potential cost of opportunistic behavior by a rating agency. The study also could serve as a reference for regulatory authorities in supervising credit rating charges and improving the reasonableness of credit ratings. It also calls for greater transparency in setting rating charges.

Keywords Rating Fees, Regulatory Intensity, Over-rating, Rationality of Credit Rating, Experimental Research

JEL Classification D53 G24 M41