



中国证券投资基金业协会
Asset Management Association of China

ESG 研究专栏

(学术版)

【2020】第 4 期 总第 4 期

发布日期：2020 年 11 月 6 日

ESG 评级应该万流归宗吗？

上海交通大学上海高级金融学院教授 邱慈观
上海交通大学上海高级金融学院研究员 张旭华

随着 ESG（环境、社会及公司治理）投资在国内的崛起，ESG 评级随而出现，涉入的机构还真不少。目前纳入万得数据库的 ESG 评级或评分，就有富时罗素、嘉实、华证指数、商道融绿、社投盟、OWL 等 6 家，此外另有一些由民间组织和学者所做的评级，在圈内流传。

ESG 评级在国内才出现几年，但在欧美已崛起相当时间。首家评级机构 EIRIS（Ethical Investment Research Services）于 1983 年成立于英国伦敦，替教会及慈善组织提供企业的 ESG 信息，以引导负责任投资。另一家老牌的 ESG 数据和研究公司 KLD，以影响企业行为及推动世界更公平、更可持续为使命，于 1988 年在美国波士顿成立。

其后，ESG 投资的发展加快，驱动了数据需求，ESG 评级机构也急速涌现。依据统计，目前全球 ESG 评级机构有 600 家，其中包括彭博、MSCI、Vigeo-EIRIS、富时罗素、路孚特、

CDP 等，为市场提供多种产品与服务，从数据、研究、咨询、技术，到投资策略及股东议合服务。

然而，ESG 数据和评级机构，在法源、组织使命、法律身份、评级主旨、产品与服务等方面，都存在着巨大差异。以法源看，至少有大陆法系和英美法系两种，而前者的关切向利益相关者倾斜(如 Oekom)，后者的关切向投资者倾斜(如 MSCI)。以组织使命看，有些机构拟通过 ESG 数据来改变世界，另一些拟通过 ESG 数据来告知世界。以法律身份看，从非营利型到营利型都有，而这会对评级机构的独立性形成影响。从评级主旨看，有些机构只涉及单一维度(如 CDP)，另一些则涉及全科(如 Sustainalytics)。从产品与服务看，有的机构主营股东议合，有的机构专营数据和研究，另一些则提供更宽广的服务范围。

ESG 评级机构的数量繁多，背景迥异，表示这行业必定看法分歧，莫衷一是。

事实情况也的确如此。国际的 ESG 评级分歧巨大，数据来源不同，评级框架迥异、评级结果不一致，使投资者难以识别，甚至被误导，做出错误决定。

数据来源方面，有些 ESG 评级机构以企业公开披露的社会责任报告为主，有些评级机构在此基础上挖掘另类数据来源，如各政府及监管发布平台、新闻媒体、行业协会等。方法框架方面，有些评级机构以 ESG 风险评级来取代全面性的 ESG 评级，例如 Sustainalytics，而另一些则仍以全面性评级为主，例如 RobecoSAM。

评级结果方面，美国麻省理工学院的博格教授利用 2014

年数据，计算 KLD、MSCI、Vigeo-EIRIS 等六家机构的评级相关性，发现其平均相关性只有 0.54。日内瓦大学的吉卜森教授利用 2013 年到 2017 年间的数据库，计算另六家评级机构的评级相关性，发现其平均相关性为 0.46，而公司治理维度的相关性竟然低到只有 0.19。

分歧也存在于国内市场，ESG 评级机构对同一主体的评级并没有共识。以贵州茅台为例，华证指数评为“AA”，而商道融绿评为“C+”。依据今年平安数字经济研究院对国内 ESG 信息披露所发布的报告，穆迪与标普的信用评级相关性高达 0.99，而国内的 ESG 评级相关性只有 0.33。

贵州茅台 600519.SH



ESG评分						
	富时罗素	社会价值投资联盟	商道融绿	嘉实	华证指数	OWL
最新评级日期	2020-07-14	2020-06-30	2020-06-30	2020-05-31	2020-04-30	2020-01-31
总分	1.10	57.11	65	78.05	65	42.56
评级	--	BBB+	C+	--	AA	--
全部均分	1.29	56.31	65	49.43	65	50.00
公司排名	278/551	156/300	779/800	296/3843	778/3828	1792/2405
行业均分/排名 (日常消费)	0.99 (12/36)	56.09 (9/20)	47.85 (51/53)	47.72 (15/209)	84.41 (45/209)	48.45 (117/144)
行业均分/排名 (食品、饮料与...)	1.01 (9/27)	56.15 (8/18)	47.71 (39/41)	47.59 (14/173)	83.84 (36/173)	48.73 (105/123)
行业均分/排名 (饮料)	0.98 (5/13)	54.86 (4/9)	48.06 (15/15)	51.91 (5/40)	86.53 (14/40)	47.91 (23/31)
行业均分/排名 (白酒与葡萄酒)	0.92 (3/9)	55.71 (4/8)	48.32 (11/11)	52.99 (4/29)	87.28 (12/29)	47.84 (17/22)
	申请权限	更多细分数据	更多细分数据			更多细分数据

图 1：贵州茅台的 ESG 评级、评分一览（来源：Wind）

针对这种混乱，业者、学者及跨界组织都先后涉入，做出相当的努力。例如，有业者以一统天下为目标，开发了 ESG 生态系统图谱；有业者以提高透明度为宗旨，启动了“对评级者评级”（Rate the Raters）报告。麻省理工学院以“层层混淆”（Aggregate Confusion）为名，成立了一个研究 ESG 评级分歧的专项，以探讨产生分歧的原因。另外，世界企业可持续发展委员会更形成了一个 ESG 评级工作小组，以拨云见日为旨，帮助投资者了解各评级之间的差异，以振兴 ESG 评级的价值。

一、ESG 评级为何分歧？

ESG 评级相比于信用评级，主要区别在于前者关注评级主体在非财务维度的情况，后者关注其财务相关情况。对于企业财务维度的界定及度量，普遍接受的准则已经建立，如今争议不多。但对于企业非财务维度的界定及度量，由于建立在一些模糊概念上，其具体化及框架化都涉及诠释者的社会背景与价值观系统，故迄今非但没有普遍接受的准则，未来是否应该标准化更是必须严肃讨论的问题。

然而，ESG 数据和评级是产品，ESG 数据商和评级机构是组织。组织的社会文化背景、历史渊源、使命、结构、法律身份等因素都带有价值观成分，对组织的客观特质和主观理念框架形成影响，随而对其产品产生作用。特别是，当产品涉及社会判断及价值观时，这些影响因素就更为重要。所以，在探讨产品时，显然不应该把产品背后的组织因素强行抽离掉。

因此，针对造成 ESG 评级分歧的原因，目前至少有两种看法。一种是由产品看，纯粹从技术视角来看造成分歧的原因。另一种是跳脱出技术层次，从机构视角来看造成分歧的原因。

从技术角度看 ESG 评级分歧，已经存在相当时间，其中涉及对 ESG 指标、方法、度量及权重等进行深度解析，以找出产生评级分歧的技术性原因。从机构背景角度看 ESG 评级分歧，则才开始，由可持续会计准则委员会创始人、牛津大学的伊克雷教授领衔，其中涉及对评级机构之社会脉络及历史渊源的解析，以找出认知模式和价值观体系对 ESG 评级的

影响。

二、从技术角度看 ESG 评级分歧

从技术角度看 ESG 评级的分歧，学者、业者及政府监管者都有涉入。技术本身不涉及价值观，它造成的分歧比较可能通过标准化要求和监管流程来统一，故这个视角有政策意涵，例如对 ESG 评级行业设置准入条件并进行监管、对评级制订标准框架等。

针对造成评级分歧的技术面原因，学者做了很多研究，而以麻省理工学院的“层层混淆”专项为代表。它由博格教授领衔，通过对 KLD、Sustainalytics、Vigeo-EIRIS、Refinitiv、MSCI 和 RobecoSAM 等六家欧美的 ESG 评级机构，研究产生分歧的原因。对分歧来源进行解析后，博格教授归纳出主题覆盖差异、指标度量差异和权重设置差异三个来源。

首先是 ESG 主题覆盖存在差异。ESG 评级虽然主要考量 E、S 和 G 三个维度，但每个维度下的议题却各有不同。例如，MSCI 的 ESG 评级分为 3 个维度，10 大主题，37 个关键议题。富时罗素的 ESG 评级在 3 个维度下面，有 14 大主题，300 多个指标。路孚特的 ESG 评级则包含 10 大主题，450 多个指标，而 CDP 则只关注环境维度，下设 3 大主题。

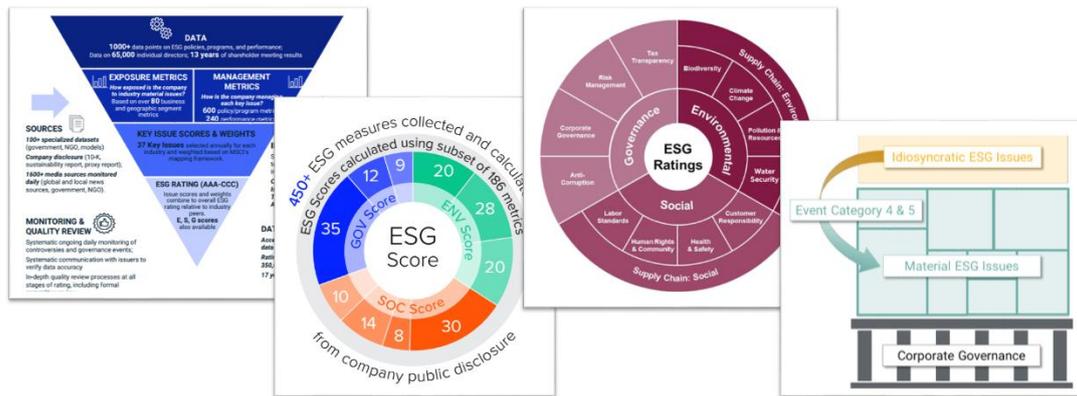


图 2：ESG 评级框架，由左至右分别为 MSCI、路孚特、富时罗素和 Sustainalytics

以环境维度为例，MSCI 的 ESG 评级关注气候变化、自然资源、污染及废弃物、环境机会四个主题，而富时罗素的环境维度则包括气候变化、污染及自然资源、生物多样性、水资源安全、供应链。CDP 则关注气候变化、森林资源、水资源安全三大主题。依此可见，富时罗素单列的生物多样性主题，MSCI 却并未涉及。主题覆盖差异会导致最终 ESG 评级的差异。

其次是指标度量存在不同。针对同一议题，指标选取也出现差异。譬如，员工管理可以从员工流失率指标来看，也可以从员工满意度或劳动纠纷指标来看。又譬如，商业道德是从企业政策来看，还是事件发生频率？各家 ESG 评级机构在指标选取上的差异，同样会产生影响。

第三是权重设置的差异。不同 ESG 评级在权重设置上也各有不同。MSCI 会考虑各议题对公司及行业的影响程度和影响时间长短。如果影响程度大，实质性发生快，MSCI 会给予该议题更高的权重。国内的商道融绿则根据行业，设置了通用指标和行业指标，并给予不同权重。

博格教授的研究发现，在三个来源里，ESG 评级分歧主

要由主题覆盖差异和指标度量差异所形成，而权重设置差异则比较次要。这个研究目标恢宏，但主要围绕着表象做分析，而未能再深入一层，追究达成评级一致性是否需要满足一些先决条件。

针对于此，杜克大学伽特奇教授的研究就提供了更深一层的洞见。他先提出评级一致性的两个先决条件：共同的理论架构（common theorization）和共同的度量（commensurability），再以六家评级机构为对象，检视这两个条件是否满足。

简单的说，共同的理论架构关乎评级机构是否有共同的想法，譬如什么构成“有担当的企业”，而在此基础上形成所关切的概念和维度。共同的度量关乎评级机构是否能以相同的方式来度量同一现象，譬如对于失责行为的度量是以定性为主，抑或以定量为主。当这两个先决条件都成立时，各评级就可能达成一致性结果，反之则不然。当伽特奇教授把六家机构的 ESG 评级放在这个框架下检视时，发现两个条件都不成立，因而导致了现实世界里的评级分歧。

得到这样的结论并不奇怪，因为光是从表面看，对于什么构成“有担当的企业”，大陆法系下评级机构的看法就和英美法系下评级机构的看法不同。更具体地，针对企业承担责任的对象，大陆法系倾向于多方利益相关者，包括供应商、顾客、员工、股东等，而英美法系则倾向于独尊股东。法源不同会影响评级机构的看法，造成大陆法系下之 Sustainalytics 和英美法系下之 MSCI，其评级项目重点和最终结果都产生差异。

三、从机构角度看 ESG 评级分歧

伽特奇教授所言之共同的理论架构，应该与评级机构的社会脉络有关：当评级机构源于相同的机构背景，有共享的法律法规、社会常规及文化认知模式时，则会通过共同的理论框架来理解 ESG，也才能塑造出相近的产品——ESG 数据和评级。

机构因素之所以会影响 ESG 产品，与其非财务性特质有关，其中涉及认知模式和价值判断，必须通过评级者的主观框架来进行选择和解读。特别是，ESG 数据背后有可持续性 & 实质性两个关键理念，其概念化与框架化都取决于评级机构的社会背景、组织使命与法律身份，而这些因素又决定了评级机构所选择的市场定位，最终形成其产品与服务组合。换言之，有别于财务数据之价值中立性，ESG 数据受到评级机构思维模式和价值观体系的影响，而通过社会脉络视角来理解 ESG 数据的建构流程，就更能凸显各评级机构的个别特性和相互差异性。

事实上，依据组织战略管理文献，一家企业的外部环境和内部组织流程，共同形塑了它的企业文化、战略定位和产品组合。因此，从社会脉络角度看 ESG 评级，要求欧系的 Vigeo-EIRIS 和美系的 MSCI 产生相同的评级结果，是欠缺理论依据的看法。

特别是，由创始原因、组织使命、文化背景、法源等所反映的社会脉络看，Vigeo-EIRIS 和 MSCI 的概念化框架不同，造成它们在实质性的界定、度量方式的选择上都分歧，最终反映于 ESG 评级。实质性的界定方面，Vigeo-EIRIS 强

调其慈善组织及工会的历史传承，旨在替利益相关者服务，而实质性包含了对各种利益相关者造成影响的外部效应。相比之下，MSCI 界定的实质性聚焦于投资者，而只纳入了对企业长期盈利会造成影响的 ESG 风险。度量方式上，Vigeo-EIRIS 的关注围绕着程序正义和民众权益等软议题，度量方式以定性为主。相比之下，MSCI 的关注围绕着实质效益等硬议题，度量方式以定量为主，譬如绩效指标。

因此，从机构视角看评级分歧，看到的是一些基础条件对评级的影响，而这方面的异质性形成了一个隐性因素，阻碍了评级结果达成一致性的可能。

四、应该万流归宗吗？

ESG 评级分歧的事实，已经广为人知，市场上有很多报导，指出它们如何分歧，譬如平均相关系数多低、哪些维度的评级分歧最大等。但是，迄今为止，讨论一直围绕着表象，而超越这个层次的讨论并不多见。譬如，评级分歧对股票收益有何影响？评级分歧的产生原因为何？

本文聚焦于 ESG 评级分歧的原因，把学者对于分歧原因的研究归纳为两派，一派由技术面因素来解释评级分歧，另一派则由超越技术层次的社会面因素来解释评级分歧。两派的视角不同，而目前研究结论似乎否定了技术派路线，理由是当追溯评级机构何以使用不同技术时，最后的归因又回到社会面因素。

当 ESG 评级分歧由机构背景和社会脉络形成时，对投资者、基金经理及监管者有何意涵？

首先，对投资者和基金经理等使用者而言，各种 ESG 评

级都有一定程度的使用价值，但使用者必须理解各评级之间的差异，而后依自身偏好来选择符合需求的评级。譬如，CDP 的环境评分可能更具深度，RobecoSAM 的 ESG 评级可能更为欧洲投资者信赖，而 MSCI 的评级可能更强调会影响股东财富的实质性 ESG 风险。另外，如路博迈等资源丰富的大型资管公司，可以在各机构的 ESG 数据基础上，增加内部大数据团队的发现，以开发合适的 ESG 评级。

其次，针对某些监管者要求 ESG 评级标准化的主张，从评级本身是基于社会脉络而建构的视角看，通过统一数据来源和评级框架等方法来解决分歧问题，欠缺本质上的意义。固然，针对企业的 ESG 信息披露、针对行业的 ESG 评级“漂绿”等问题，监管者仍应制订对策，但这不同于对 ESG 评级框架、指标选取和度量方式的监管，它们都涉及评级组织的社会价值观系统，而任何相关措施都将难以落实。

最后，如果评级分歧的原因超乎表面，而关乎评级机构的社会脉络，则万流归宗可能是不实际的想法。相比于“求同”，对于 ESG 评级“存异”，洞察评级分歧背后的真正原因，并明智地选取运用，方为可行之道。

声明：本文仅代表作者个人观点，中国证券投资基金业协会（AMAC）出于交流目的经作者授权刊发，不意味着 AMAC 支持、认可其全部或部分观点，不代表 AMAC 立场。