信用评级机构:金字招牌从闪光到蒙尘

■夏春 诺亚控股香港有限公司研究总监

3月以来,穆迪和标普相继把中国和中国香港的主权信用评级展望由稳定下调为负面,国际三大评级机构之一的惠誉也以电子邮件评论称,如果中国政府中期经济战略不清晰,或对政府推进改革的能力和信心下降,他们也可能做出调整。

2008年金融危机爆发后,无数曾被信用评级机构冠以投资级(以标普为例,从BB+到AAA的评级)的房地产抵押债券和各种结构化金融产品,被重新评估为垃圾级。市场达成共识认为,评级机构的"评级膨胀"行为是造成金融危机最主要的原因之一。

我在《索罗斯们是金融危机的黑武士》一文里写道:复兴科技创始人西蒙斯认为评级机构以次充好,"把猪耳朵当丝质皮包来卖"是危机最大的责任人。当时,标普内部的一封电邮被媒体广泛报道:"母牛设计的结构化产品也能得到我们的评级"。电影《大空头》里有一出极具讽刺意味的场景:对冲基金经理鲍姆质问一家评级机构的主管为什么给予不合格的房屋抵押债券AAA的评级时,正在治疗眼疾,逮着硕大墨镜的主管回答说:如果我们不那么做,那些发行债券的公司就会去找穆迪,我们不合作的话,他们会从我们的竞争对手那里得到他们想要的评级。许多观众称,电影里最让人印象深刻的就是评级机构闭着眼睛工作的形象。

3月以来,穆迪和标普相继把中国和中国香港的主权信用评级展望由稳定下调为负面,国际三大评级机构之一的惠誉也以电子邮件评论称,如果中国政府中期经济战略不清晰,或对政府推进改革的能力和信心下降,他们也可能做出调整。中国和中国香港的政府财经官员纷纷出面评论,自然都是预料之内的不认同和不在乎的态度(巧合的是,就在3月31日标普宣布其决定时,香港证券及期货事务上诉审裁处公布复核裁决,认定穆迪在2011年7月发布的一份点评中国企业的报告未能满足《操守准则》相关要求,宣布判处穆迪1100万港元罚款并作公开谴责,穆迪同时需支付香港证监会60%的诉讼费用)。许多针对普通投资者的财经博客和公众号则以耸人听闻的评级机构搞垮中国的"阴谋论"来吸引眼球。至于市场,基本是波澜不惊,上证指数和恒生指数分别在3月份上涨了11.6%和8.9%。我在4月1日参加香港电台早间访谈节目《Money Talk》就听到另外一位嘉宾说,"现在除了一些因委托责任而必须听从评级机构意见的金融机构,主要是养老金和保险公司外,没多少人把他们当一回事。"

金字招牌的年代

看来,金融危机之后,评级机构的"坏孩子"形象已经深入人心。不过,最近以各种理由批评评级机构的文章都没有提到,2011年8月标普首次将美国长期债券评级从AAA下调到AA+,差不多同时,穆迪和惠誉都把美国主权信用评级展望下调为负面。相反,标普和穆迪在过去几年一直在上调中国的主权信用评级,从未下调过(惠誉曾经在2013年下调过中国主权信用评级)。

更少有人知道的是,在金融危机之前,存在历史长达100多年的三大评级公司称得上是金融业里受到批评指责最少的机构,由于对企业和政府发行债券时的偿债能力的评估与他们后来的表现相当吻合,评级机构的声誉曾经极佳。但评级机构的金字招牌究竟是如何坠落的,我敢说知情的人并不多。

就拿《大空头》电影那一幕来说,其解释实际上流于表面,离真相遥远。三大评级机构的竞争已经存在了100多年,为什么过去没有出现评级膨胀的事情?如果评级不合理,投资者为什么要大量购买被高估的债券和结构化产品?为什么普通的企业债券得到AAA评级的不到1%,而结构化产品获得AAA评级的比例接近60%?造成评级机构评级错误的真正原因在哪里?

2008年加入香港大学金融系的第一天,我就发现汤勇军教授在日以继夜地挖掘数据,他和其导师用两种方法找到了评级机构在结构化产品评级上的过失。

一是,他们对比了同一家评级机构的评级组和监察组。发现评级组在评级模型中使用了更利于产品获得更高评级假定(例如产品里不同资产的相关性被假设为很低,这有利于风险的分散),而监察组则使用了更真实的假定,与产品事后的表现比较更为接近。原因在于评级组为了拉到生意面对更多的利益冲突。

二是,他们发现,债券市场上的主要结构化产品债务担保凭证(CDO)获得被评为AAA比例,相比他们的基本面和评级机构的模型决定出来的比例高出许多,也就是说他们的评级被人为地调高了。例如他们发现1997年1月到2007年3月期间被评为AAA的CDO,只有1.3%符合评级机构制定的AAA评级需要满足违约概率标准,剩下的则达不到。92.4%获得AAA评级的份额实际只能满足AA的标准。

我对这些实证研究产生了浓厚的兴趣,不过,我对当时读到的关于评级膨胀的几种理论解释都不是太满意,各自都有一些明显的漏洞。后来我差不多花了两年时间才提出一个更加符合评级机构金字招牌从崛起到跌落的解释。

我首先发现,在金融危机爆发前,关于评级机构的学术研究异常地稀少,原因非常简单。第一,通过对历史数据的研究,大家发现评级机构给出依照债券违约概率表制定的评级和债券后来真正发生违约的概率对照,评级机构的准确度很高。第二,学者们早就发现评级和评级变动蕴含的信息,早就被市场其他的交易工具,如股票、期货、期权、信用违约互换(CDS)等提前揭示(缺点是他们往往会过度反应)。金融公司的信用分析师会对他们投资的债券信用风险进行内部评估,而这一债券的发行公司如果有股票公开交易,则股票分析师也会对公司的业绩进行全面评估,由于这两种分析师的工作是积极跟踪和动态调整的,他们掌握的信息往往通过股票和债券价格变化反应出来,相反,评级机构的人手有限,在动态跟踪和调整上往往落后市场。因此,过去评级机构受到的主要批评是他们在调整评级,尤其是向下调整时比较缓慢(相反,向上调整比较快,这当然是出于维持和发行企业良好关系的考虑)。比较明显的几次是美国一些大公司,比如雷曼兄弟,在宣布破产前,他们的债券评级依然很高。反过来说,这次穆迪和标普对中国和中国香港的主权信用展望评级调整,是早了还是迟了,考虑到中国和中国香港总体债务占GDP比重的不断上升,我个人倾向认为是迟了。

由于评级机构改变评级往往晚于市场的判断,一个直接的后果就是当评级机构宣布其决定时,市场往往反应不大。不出意外,评级机构最近的决定并未对中国和中国香港的股票和债券市场产生明显的负面影响,他们给出的理由市场早已经讨论了不知道多少遍了。

既然评级机构的作用可以被其他人或者市场本身所取代,为什么他们的声音仍然如此举足轻重?如果真的不在乎,财经官员们又何必愠怒之情溢于言表?实际上,许多人都听说过《世界是平的》作者托马斯·佛里德曼说的一句话,**"我认为整个世界现在有两个超级强权,一个是美国,一个是穆迪。美国可以用炸弹摧毁你,而穆迪可以下调你的债券评级而摧毁你。"**

评级发牌的超然权力

评级机构对市场的影响力来自于金融市场监管机构的权力外包。1934年制定的美国证券法要求公司发布标准化的财务报表,其中要求包含债券评级的内容,而这项服务主要由三大评级机构提供,但为评级付款的并非发行债券的公司,而是购买债券的投资者。1936年,银行业监管机构禁止银行投资"投机性证券"。换言之,银行只能投资于被评级机构认定的投资级债券。这样的话,即使银行内部的分析师认为债券的真实评级和评级机构的不一致,银行也必须服从评级机构的判断。随后,保险业和联邦养老金都采取了相同的要求。到了1975年,美国证监会进一步强化了评级机构的地位,他们要求投资银行和证券公司这些经纪商和自营商在满足最低资本要求时,必须参考评级机构对于债券的评级来计算和风险挂钩的资本额度。同时,证监会为了避免一些低质量的公司编制虚假的评级,设定了"全国认定的统计评级组织(NRSRO)"标准。三大评级机构成为元老。现在,全球有150多家各类评级机构,但真正能获得NRSRO认证,从而具备全球影响力的评级公司只有9家。不久,其他金融监管机构也都采取了NRSRO的标准,评级机构的话语权愈发强大,到了1990年初,连货币市场共同基金投资于短期商业票据也需要听从于评级机构的意见了。

因为监管机构把监管权外包,具有垄断地位的NRSRO评级机构实际上成了"评级执照"的发牌人,哪怕他们的评级不准确,其他人也几乎无法挑战他们,这基于两个原因。一是,虽然市场上的公司和机构都要按评级机构的评级来满足各种监管要求,但评级机构却总是在评级报告后附有免责声明:"任何使用报告信息的用户,都不应依赖评级和意见去做投资决策"。二是,评级机构把他们的评级定义为"观点/opinion",享受美国宪法的言论自由保护,因此在金融危机前针对评级机构的诉讼非常罕见。而评级机构拥有了这种超然的地位后,对企业基本面发生变化的反应迟缓则毫不让人意外,动画电影《疯狂动物城》拿行动缓慢的树懒来嘲讽规管驾驶执照的机动车辆管理局,其意图一目了然。

不过,也必须承认,要了解快速变化的全球经济中隐藏的风险,需要系统性、数据驱动的分析。虽然其他交易工具可以揭示市场已经知道的信息,但毕竟不够完善,评级机构的"事后诸葛亮"也有一定的客观原因,比如他们可能缺少跟踪重要因素的工具,这些因素包括一国的创新能力和私营部门的财务健康状况。特别是在国家主权信用分析,金融危机的预警上,评级机构的系统性和一致性仍然有其价值的。

多年来评级机构对于企业债券的评级相当准确,赢得了市场的尊重。特别是在1970年,面临大量需求的三大评级机构相继从"向投资者收费"商业模式转变成"向发行者收费"之后,仍然保持了评级的基本准确(目前的9家NRSROs里仍然有只向投资者收费和不收费的公司)。大家通常认为,发行者付费就会有很强的动力要求评级机构给予更高的评级,这会带来严重的利益冲突,尤其当评级机构展开收费竞争的话,评级膨胀应该会自然而然地产生。那为什么《大空头》给出的"竞争说"并未在1970–2000年产生明显的问题?

原因有这样几点。首先,评级机构认为,"声誉"是他们可以持续经营的最重要因素,因此尽可能地维持评级的公正。尤其是当企业和政府债券作为发行主体时,由于结构比较简单和透明,任何不准确的评级都很容易被市场上的分析师发觉,从而破坏评级机构的声誉。其次,发行债券的企业和政府数量之大,使得任何一家想和评级机构做讨价还价的企业或政府最后都发现自己实际处在弱势地位,别忘了前面提到的"穆迪可以下调你的债券评级而摧毁你"。最后,学者们研究发现,评级机构内部的薪酬和升职制度都和评级结果关系不大,因此潜在的利益冲突受到了很好的控制,声誉机制压倒了利润动机。经济学原理也清楚地表明,只要有制衡和约束,竞争的结果通常是好的,而不是坏的。

评级膨胀的原因

金融危机后,评级机构备受谴责,研究评级机构的论文开始如雨后春笋一般涌现。关于评级膨胀的研究很多,理论上可以概括为下面几种。第一种可称为"产品复杂论",结构化产品要远比过去的企业和政府债券复杂和不透明(里面包含了房地产贷款,个人贷款,信用卡和汽车贷款,以及这些结构化产品的衍生品),但同时结构化产品的多样化特征,又使得其具有降低违约率风险的"魔法"(即使垃圾债券放到一起,由于多元化分散风险,其收入流也可以用来构造出具备投资级的债券),只不过一旦发生系统系风险,则会魔法失灵。而由于衡量结构化产品的数学工具的缺陷,尤其是对于历史数据的依赖,导致对房地产价格下跌和系统系风险爆发的概率低估,以及不同资产之间相关性的低估。因此评级机构的错误是无心之失,例如穆迪的大股东巴菲特在参加金融危机调查委员会举行的听证会上以此理由来为评级机构辩护。但这点被前面提到的实证研究否定,评级机构存在明知故犯,人为调高评级的欺骗性错误。

第二种"评级购买"理论和《大空头》的解释接近:由于改成了"向发行者收费"的商业模式,如果标普不给我一个更高比例的AAA,发行企业可以去找穆迪。前面已经解释,在过去债券结构比较简单透明时,这种竞争和购买行为对评级机构的压力不大。因此,这个理论只有在结构性产品非常复杂的前提下才能生效,也有一定的实证支持(比如和始终向投资者收费的机构相比,标普的评级要好一些),但并不足够有说服力。除了前面讲的原因,主要是超过80%的结构性产品发行企业获得两家到三家的评级(一项研究发现99%的企业获得两家评级)。而且这个理论有一个很大的不足,就是必须假定投资者是愚蠢的,他们或者不知道评级被高估了(如果他们知道,就可以用更低的价钱买到那些产品),或者说他们知道评级购买行为的存在,但还是低估了评级机构犯错的可能性。但事实是,债券投资者通常是非常精明的机构投资者。

第三个则是前面提到的"评级执照"说,这个理论认为评级膨胀和收费模式改变没有关系,由于受监管的公司和机构只能投资于被评为投资级的产品,投资者哪怕知道评级膨胀(AAA的比例明显过高),也乐于购买。这种动力在近些年变得更强,因为他们可以由此而更容易满足新法规(比如巴塞尔协议)下的最低资本要求,把多出来的钱去投资风险更高但回报也可能更高的项目,或者可以省下自己聘请信用分析师的费用,当结构化产品越来越复杂的情况下,投资者越热衷于服从评级机构的判断。而评级机构则迎合了投资者这种监管套利的需求,把猪耳朵贴上了丝质皮包的标签。这个理论认为,只要改变监管对投资级评级的优惠待遇,则评级膨胀现象将会消失。

我提出了一个全新的理论,一言以蔽之,"评级咨询"导致了"评级膨胀"。应该说这个解释早已被金融界从业人员反复提及,但它就好比数学上的著名猜想,学术界一直没有把数学模型做出来。当我做出来的那个时候的激动,是无法用言语来形容的。我大致以非技术性方式解释一下。

假设企业发行的债券质量不同,他们自己清楚,但是评级机构和投资者只知道质量分布而非具体每一家企业的质量。由于评级机构的评级技术是不完美的,有可能把高质量的企业错误地先内部评为低质量(但犯错的可能性不大),然后再选择告诉外部投资者该企业质量是高还是低(诚实还是欺骗策略)。只有被外部评级为高质量的企业才能成功发行债券,并向评级机构付费。数学可以证明,在这种三方博弈情况下评级机构一定会采取诚实策略:外部评级和内部评级一致。虽然是向发行者收费,但与大家猜测的这会导致利益冲突恰恰相反,评级机构不会把内部评定为低质量的债券对外评定为高质量。为什么?因为不需要事前缴费,低质量的企业也会去碰碰运气,希望被评级机构内部评定为高质量。但这种可能会被精明投资者提前预期到,所以会要求企业提供更高的补偿,相应地,评级机构可以从发行人和投资者取得的回报会相应减少。因此评级机构的最优策略是诚实。在这种情况下,低质量企业的最优策略是不要求评级和不发行债券,高质量企业的最优策略是要求评级,虽然有可能被误评为低质量。

假如企业一开始并不知道自己的债券质量,或者需要评级机构的咨询来重新包装债券以提升质量,并缴付一笔咨询费。接下来评级机构再进行评级,其余过程同上。在这样的三方博弈下,评级机构的最优策略可能变成欺骗,即把内部评定为低质量的企业对外宣布为高质量。原因是采取如前的诚实策略,则对质量缺乏信心的企业没有动力去寻求咨询的帮助,而评级机构也会因此减少咨询收入。虽然投资者会明白其中的奥妙,但是他们要求的额外赔偿可能会低于评级机构从咨询获取的额外收入。而只要咨询能够从平均水平上帮助企业提高债券质量,则企业也愿意缴付咨询费,三方皆大欢喜。决定评级机构选择诚实还是欺骗的关键因素就变成收费的高低,评级技术犯错的概率,企业质量的高低等。另外,当投资者基于评级机构过去的准确性表现,而低估了评级技术犯错的概率时(尽管投资者知道结构化产品更复杂),评级机构就有更强的动力去欺骗。而由于债券的投资期很长,投资者要发现评级机构的错误就需要很长的时间。

在2000年以前,评级咨询还没有普及,一来传统的企业和政府债券结构简单透明,对咨询服务没有什么需求。二来,最初美国发行房地产抵押债券的政府机构对于贷款人的信用要求很高。后来,利润驱使许多商业机构和投资银行也加入到游戏中,对贷款人信用要求降低,如何对结构化产品进行包装以获得更高比例的高评级变得关键,评级咨询也由此而蓬勃发展,但问题也随之而来。我的这个解释和之前最大的区别在于,博弈的三方都是精明的,没有任何一方愚蠢。特别是,即使改变监管对投资级评级的优惠待遇,但继续允许评级咨询收费的存在,那么评级膨胀现象仍然有可能出现。

因此,监管机构对评级机构的咨询和评级双重身份的放任,才是问题的关键。可惜的是,这个问题早在2001–2002年安然和安达信的财务造假丑闻里早就曝光(安达信具有咨询和审计双重身份),但监管机构却对类似行为听之任之,对于金融风险的积累和蔓延失察和失职。

评级机构因为金融危机前的评级错误,付出的不仅仅是声誉的损失。标普在2015年同意向美国证交会和另外两家监管机构支付总计7700万美元以达成和解,同意向美国司法部支付6.875亿美元和解金、向美国19个州和哥伦比亚特区支付6.875亿美元和解金。近年来,在学术界和业界的共同推动下,监管机构要求评级机构采取更加公平公正的方式对待被评级的企业和政府,并与评级咨询服务脱钩,同时建议市场降低对评级的依赖。应该说这些都是正确的方向,只不过监管机构自身的改革需要时间,围绕评级机构的是非争议仍然在相当长的时间内继续,评级机构重获市场的信任更加需要时间。

热视角

美联储对再次加息时间意见分歧

美联储(Fed)多名政策制定者在最近会议上警告全球动荡给美国经济带来"可观"风险,分歧严重的官员们在会上就多快进一步上调短期利率展开辩论。

3月15日至16日的利率会议深深笼罩着这种担忧:如果美国意外触礁,这家央行有多大能力放松政策?在这次会议上,除一位委员以外,联邦公开市场委员会(FOMC)其他所有委员都投票赞成维持利率不变。

3月会议的会议纪要显示,关于在去年12月加息后下一次何时收紧货币政策,政策制定者存在分歧,一些人提倡在4月就采取行动,另一些人则警告称这将表现出不必要的紧迫感。

在这次会议上,美联储将其对2016年利率的预测下调至两次加息,每次加息25个基点,接近低迷的市场预期,同时美联储提出市场和海外较疲弱增长对美国前景构成的风险。

此后,美联储主席珍妮特·耶伦(Janet Yellen)已强调央行在考虑下次何时加息的问题上应"谨慎前行",这暗示,她无意匆忙进行第二次加息。然而,公开声明表明美联储官员之间存在分歧,很多地区联储银行行长提出乐观观点。

会议纪要进一步显示出了这些分歧,例如,在美国是否近乎处于完全就业以及在多大程度上重视最近通胀上升以及通胀预期低迷等问题上,政策制定者有不同意见。

几名美联储官员警告称,本月加息将"表明一种他们认为不适当的紧迫感",另一些官员支持加息,强调劳动力市场和通胀前景日益走强。

会议纪要证实,对于美联储多名政策制定者而言,今年年初以来的金融市场动荡仍是一个值得积极关切的问题,似乎表明他们不急于加息。

会议纪要称:"几名与会者认为,导致今年早些时候全球金融状况严重恶化(尽管是暂时的)的基本因素尚未完全化解,因此构成下行风险。"

(FT中文网)