

世界银行扶贫协商小组 (CGAP)

特 刊 第 7 期

(2003 年 1 月)

等级评分：小额信贷领域的又一次突破？

MARK SCHREINER

本文面向那些希望初步了解有关等级评分系统的小额信贷经理们，因为他们想知道该系统如何工作，系统能做什么而不能做什么，以及该如何为系统的实施做准备。第一部分比较概括，接下来的部分就比较讲究技术性了。如果读者不想了解详情，可以直接浏览主要标题和方框里的内容。

本文不是一本使用手册。设计和实施一个等级评分系统需要高度专业化的技术，而且总体而言，它们也太复杂、太专业，无法在这有限的篇幅内解释清楚。

这里讨论的内容是基于在小额信贷等级评分过程中做过的一些最初试验。¹ 论文中的一些地方还吸取了高收入国家在等级评分方面的长期经验。² 文中使用的例子也反映了作者在拉丁美洲进行等级评分实践的经验。反过来，这也说明在拉丁美洲确实存在规模巨大、发展完善的小额信贷机构，它们拥有充足的电子数据库。所以，只要拥有足够强大的组织机构、足够规模的数据库，等级评分系统就将如在拉美一样，在非洲、亚洲和中欧占有一席之地。虽然它无法触及每一家小额信贷机构，但是，从长远来看，它将在全世界范围内普及。本文应该有助于经理们判断等级评分在自己机构中

¹ 见 Mark Schreiner, "对玻利维亚一个小额信贷机构客户退出风险的评分" (华盛顿大学社会发展中心, 手稿, St. Louis, Mo., 2001 年); "小额信贷的等级评分:它有效吗?" 《小额信贷杂志 2》, 第 2 期(2000 年):第 105-18 页; 和"预测玻利维亚一个小额信贷机构拖欠风险的评分模型"(华盛顿大学社会发展中心, 手稿, St. Louis, Mo., 1999 年)。另外见 Ulrike Vogelgesang, "危机时期的小额信贷:竞争, 增加的负债, 以及经济危机对还款行为的影响" Gk 工作论文系列 No. 2001-06 (曼海姆大学, 2001 年); 及 Laura Viganò, "开发银行的等级评分模型:一个非洲案例的研究" 《储蓄与发展 17》, 第 4 期(1993 年):第 441-82 页。

² 见 Elizabeth Mays, "等级评分手册"(芝加哥:Glenlake, 2000 年), 和"信贷风险模型:设计与申请(纽约:Glenlake, 1998 年); Lyn C. Thomas, "信贷和行为评分的考察:预测贷款给消费者的金融风险"《国际预测杂志》16 (2000 年):第 149-72 页; Peter McCorkell, "等级评分 101" (在联邦贸易委员会公共论坛上提交的论文, "消费者与等级评分" 1999 年); David J. Hand and W. E. Henley, "在消费者等级评分中的统计分类方法:一个回顾"《皇家统计协会期刊》, 系列 A 160, 第三部分 (1997 年):第 523-41 页; Loretta J. Mester, "等级评分的要点是什么?" 《商业周刊》(一九九七年, 九月/十月):第 3-16 页; Lyn C. Thomas, J. N. Crook, 和 D. B. Edelman, "等级评分和信贷控制论"(牛津:1992 年, Clarendon 出版社); 以及 Edward M. Lewis, "等级评分的引入"(加利福尼亚 San Rafael, : 一九九零年, Athena 出版社)。

的有效性。

摘要

小额信贷面临的挑战在于，穷人是否能够通过自我就业如约偿还贷款。等级评分，一种判断风险的新方法，是小额信贷领域的又一次突破吗？等级评分确实减少了拖欠行为，进而减少了花费在收款方面的时间；这种更高的效率提高了业务的扩展性和持久性。但是，等级评分不适用于大多数小额信贷机构。它最适合那些具有可靠的个人贷款技术和拥有大型贷款历史记录数据库的机构。即使等级评分系统已经运转，它也仅仅是一个逐步改进、提高的过程，并不能一蹴而就。特别是，等级评分系统不会取代小额信贷中的信贷员作用，因为许多自我就业型穷人的信贷风险与等级评分系统中可使用的信息无关。本文讨论的是，等级评分系统如何运作，小额信贷机构从中期待的是什么，以及机构如何使用它并且需要哪些数据。系统的成功不是来自技术的魔力，而是来自用户艰辛的训练：只有信贷员和分支机构的经理们懂得其工作的原理，并且确保系统在检测中有效，他们才会相信等级评分系统可以帮助他们做出决策。最重要的是，等级评分系统改变了小额信贷机构的思维方式，它营造了一种勤于分析的氛围，在这种氛围中，经理们定期发掘他们数据库中的信息以解决经营中出现的问题。

一、导言

小额信贷在两种判断自我就业型穷人还款风险的新方法中逐步发展起来：其一是联保小组，其二是依靠信贷员，他们对借款人和他们的家庭、经营情况以及担保品出具详细的个人和财务评估。等级评分系统（对于小额信贷来说）是判断还款风险的又一种新方法。它检测还款能力与贷款申请的量化特征之间的历史联系，并且它假定这种关系会随着时间的推移而持续下去，然后它再根据当前贷款申请的特征来预测未来的偿付风险。在高收入国家里，等级评分（以信用卡的形式）已经向成百上千万的人提供了小额、短期、无担保、低交易成本的贷款，堪称有史以来的最大成就。那么，它是否会成为小额信贷领域的又一次突破呢？

对于一些规模大、运营良好、拥有充足电子数据库的小额贷款机构来说，等级评分可以提高效率，扩展面向穷人的服务，并且能够提高机构的可持续性。它主要减少了向不良借款人催收逾期贷款的时间。（一个典型的信贷员，也许每周可以节约半天的时间。）信贷员可以利用这些时间，寻找更多良好的借款人，从而提高服务的深度和广度。

对于大型小额贷款机构而言，等级评分也可以带来盈利。举例来说，一项对玻利维亚历史数据的检验表明，2000年，如果12%最具风险的贷款被拒绝，那么逾期30天的贷款比例就能够减少28%。³即使保守地估计等级评分系统的成本、拒绝不良贷款的净收益、以及误拒良好贷款的净成本，等级评分也可以承担自身大约一年的费用。此外，它还会有大约一百万美元的净现值收入。

等级评分是一次显著的进步，但是，就联保小组和信贷员的个人评估而言，它不

³ Mark Schreiner, "玻利维亚FIE的等级评分系统" (给Fomento de Iniciativas Económicas的报告, 拉巴斯, 玻利维亚, 2001年)。

是一个突破。事实上，等级评分制对大多数以小组为特征的贷款机构或者村银行来说，效果不明显。另外，大多数向个人提供贷款的小额信贷机构尚没有为此做好准备。原因是，它们首先必须完善更为基础的程序，或者，它们的数据库不足以支持等级评分制。即使是对于已经做好准备的贷款机构而言，等级评分制也不会取代信贷员的作用，因为他们无法替代尚没有（或无法）量化入数据库的风险因素的个人评估。等级评分制并不是小额信贷领域的又一个突破，但是，它却是新理念（例如，面向需求量体裁衣、提供存款和支付服务、关注管理和激励、改善经营组织）中的一种，这些新理念，在未来很长一段时期内，必将为小额信贷带来看似微小却意义重大的进步。

等级评分制的核心挑战是组织机构的变化，毕竟在将它投入使用之前，它的预测能力可以通过历史数据来验证。信贷员以及分支机构的经理们很明显不相信这种魔力。在相信之前，他们需要了解等级评分工作的基础原理，然后通过在自己客户身上的实践来证实该系统的效用。理解和接受这一系统需要反复的培训，认真的模仿，以及对当前未偿还贷款预测的不断证实。从长远来看，一个优秀的等级评分项目应该改变一个组织的文化，使其经理在全职、内部分析师的帮助下明确分析尚未使用却在数据库中显示出经营问题的信息。

本文利用实际例子解释了等级评分项目的原理和操作。它详细描述了不同类型的计分卡，而且更为重要的是，它说明了如何在使用之前测试计分卡、如何在实地使用记分卡，以及如何记录操作等问题。同时，本文讨论了等级评分项目的优势和劣势，也破除了一些神话，特别是那些自称可以替代信贷员作用而且将加速贷款申请评估的神话。为了帮助小额贷款机构做好准备并充分利用等级评分项目，文章最后一段讨论了设计数据收集环节所需的基本细节。

二、主观评分与统计评分

小额贷款机构已经在使用主观评分，但没有使用统计评分。本节说明了主观评分和统计评分的基础概念，并解释了两者的互补性。等级评分是指，利用过去的风险与贷款特征之间的关系，由当前特征来预测未来风险的技术。将这种特征与风险联合起来研究的办法便是主观评分和统计评分。图一总体列出了二者之间的比较。

图一：主观评分和统计评分的比较		
比较项	主观评分	统计评分
信息来源	信贷员和机构的经验	数据库中贷款资金量化的历史记录
程序的连贯性	因信贷员而异、因日而异	相同贷款同等积分
程序的明确性	凭借生效的评估原则,区域信贷员的第六感或直观判断	以数学规则或公式确定量化特征与风险之间的关系

程序和产品	根据信贷员对客户本人的了解定性分类	根据计分卡确定的量化特征与风险之间的关系定量计算概率
认可程度	已经使用，被认为有效，管理信息系统和评估程序已经明确到位	改变了文化，可行性尚不知晓，改变了管理信息系统和评估程序
实施过程	信贷员需要长期培训和实地积累经验	对有关各方进行长期培训和跟踪
弊端	易受个人偏见、每天的情绪或者简单的人为错误影响	数据被伪造或被遗弃，或者没有充分利用数据以及过度使用数据等
灵活性	在有能力的经理调整下，应用范围很广	应用单一，在新环境中预测新一类风险需要重新计分
对平衡能力和“本该发生什么”的了解	根据经验或推测	通过检测计分卡记录的已偿贷款而得出

方框1：等级评分、小组贷款和村银行

由于数据方面的问题和小组贷款的性质，统计等级评分对于小组联保贷款或者村银行来说可能不太适用。一个基本的数据问题是大多数以小组形式发放贷款的机构不接受部分还贷：要么小组内的每一个人按时付款，要么谁也不付。这是一个很明智的策略，但是这意味着，数据库没有记录小组内的个体是否愿意并且有能力按时还贷。这就造成没有关于个人风险的数据。在这种情况下，等级评分可以预测小组的贷款风险，却不能预测小组中每一个人的风险。遗憾的是，小组风险与小组特征（例如，成员性别或者平均年龄等）的关联程度远没有个人风险与个人特征的密切。

即使贷款机构不接受个人支付，联保责任的性质也决定了小组成员的个人风

险被个体特征所掩盖。小组可以增加个人的支付意愿（通过小组压力和社会制裁），也可以增加个人的支付能力（通过小组指导和非正式保险）。另一方面，小组可能因为“多米诺骨牌般的拖欠效应”摧毁了个人的还贷意愿。于是，如果不考虑个体特征，贷款风险很大程度上取决于小组成员间的相互作用，这种相互作用的结果不可能用量化的特征很好地表示出来。

综上所述，量化特征对小组风险的表征作用不如对个人那样明显。这并不是坏事，它恰是小组贷款的目的。但是，它却使面向小组的贷款机构或村银行的等级评分工作变得更加困难，更加缺乏效果。

主观评分

小额贷款机构目前依靠主观评分来判断风险，它们根据自己对客户和贷款合同特征的量化指标（用数字计算的和记录在电子数据库内的）和定性了解（不是用数字计算，和/或也没有记录在电子数据库中）来预测贷款偿还风险。信贷员和信贷经理以及作为一个组织的小额贷款机构通过书面政策、培训和简单的口头形式，分享他们的经验。

虽然主观评分也使用定量的原则，举例来说，经营少于一年的客户将没有贷款的资格，但是，它更注重信贷员对客户个人特征的感性认识。以直觉为基础，主观评分划出了“不太危险，可以放贷”和“太危险，拒绝放贷”这样的定性判断。

正如小额信贷历史所证明的，主观等级评分是有效的。但是，它是否有改进的余地呢？举例来说，信贷员必须花费时间学习该组织积累的经验教训，必须培养对风险的第六感觉。此外，主观评分的准确性也会因官员的不同而不同，因信贷员某天情绪的变化而变化。而且主观判断可能会造成歧视，或者误将极少的（或者错误的）特征看作重点。

统计评分

统计评分是根据记录在数据库中的量化特征来预测风险。风险与特征的相关关系可以用一套能够明确测量概率的规定或公式来计算。例如，对于一个申请第一笔贷款的 25 岁男性木匠来说，预计其逾期 30 天的贷款风险是 20%，而对于一个在最近 3 笔贷款中都没有延付记录的 50 岁的女裁缝来说，预计其风险为 5%。金融乃风险管理，而统计评分通过一致的、明确的风险评估，为风险管理提供了便利。

等级评分的劣势是它是一种新鲜事物，只有少数小额贷款机构尝试过。利用数据库中的定量信息来帮助判断风险，与小额信贷的两个突破性创新（小组联保责任贷款与信贷员的一对一客户关系）形成了抵触，而这两个标志性创新恰好都利用了人对信用价值的主观评判。为了接受统计评分这样一种与众不同的事物，机构需要长期的培训和调整，还需要不断地证明它的预测能力。即使是在小额贷款机构采纳了等级评分之后，它们也要切记不能过分依赖它。因为，正如方框 1 中所说明的那样，统计评分与个人贷款的关系比对于小组贷款和村银行的关系更密切。

小额信贷的等级评分也有局限性，因为它需要一个电子数据库，用它记录下大量的过去贷款的还款历史和客户与贷款合同的特征。另外，数据必须相当准确。一些小额贷款机构，在它们的日常业务管理中积累了充足的数据，但是，其他许多机构没有电子数据库，没有对每笔贷款记录足够的信息，或者没有记录准确的数据。本文的一个目标就是，帮助经理重新设计信息系统以使他们的数据库在未来足以支持等级评分项目。

主观评分与统计评分的互补性

统计评分注重量化特征，而主观评分侧重于定性特征。哪种方法最好呢？在小额信贷领域二者都有一席之地，因为它们彼此互补。主观评分可以考虑统计评分忽略的问题，而统计评分可以分析对于主观评分来说太多、太复杂或者太细微的相关关系。两种评分办法均设定将来和过去相似，贷款特征与风险相关。当然，这些假定不一定都合理，但是，它们的紧密关联程度足以使等级评分变得有价值。

等级评分，无论是统计性的还是主观性的，都认为掌握过去的一些信息总比没有信息强。由于主观评分依靠的是那些具有丰富的现场经验和善于掌握各个方面信息的人，因此，当事态有别于过去时，它可以做出迅速、灵活的反映。而统计评分更具有连贯性，而且善于发现更多的（和更细微的）发展趋势，但是，它仅仅能够预测已经发生过多次的事情。

一些风险无疑与量化的特征相关，例如，欠债行为和过去的拖欠情况。可是，不是所有的特征都可以计量；即使是可以计量的特征也不总是可以量化的。在与小额信贷最相关的风险中，一定比例的（未知的）风险常常取决于个人特征，这些特征只有通过了解客户才能判断。量化特征与风险关联的比例究竟有多少呢？根据第三和第四节中的实验，本文认为，其相关程度大到足以证明统计评分的实用价值，而从另一方面这些实验也表明，主观评分可以忽略。

一些风险与更多利用统计评分计算的量化特征相关；而一些风险与更多利用主观评分方式判断的定性特征相关。在小额信贷领域，定性成分比例过大，以至于统计评分无法取代信贷员和他们的主观评分的作用。同样，统计评分不会解脱信贷经理对于信贷决策的责任。举例来说，它不可能发现借款人是否了解他们的经营或者是否浪费他们的贷款收益。统计评分在信贷委员会讨论中仅是第三种声音，它提醒信贷经理和信贷员要注意可能忽略的风险因素。

图 2：四叶树 1992-1999 年的数据（树形）

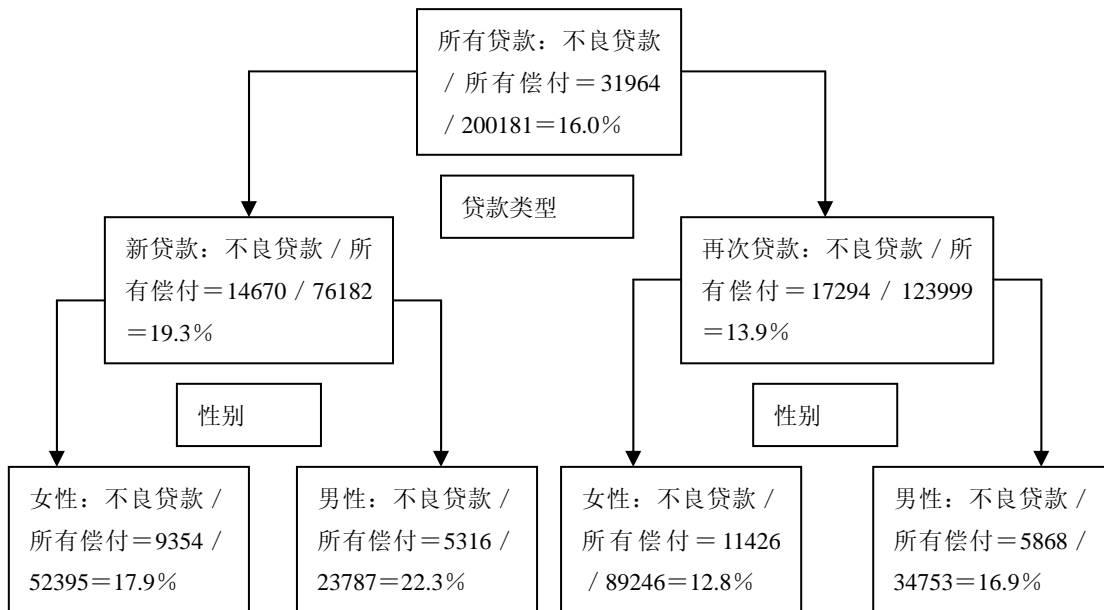


图 3：四叶树所示的历史风险，1992—99 数据（表格形式）

树枝			构造实例，1992-1999				
树叶	第一	第二	不良贷款	好贷款	叶上的实例总数	不良贷款比例%	占叶上所有实例总数比例%
1	新贷款	女性	9,354	43,041	52,395	17.9	26.2
2	新贷款	男性	5,316	18,471	23,787	22.3	11.9
3	再次贷款	女性	11,426	77,820	89,246	12.8	44.6
4	再次贷款	男性	5,868	28,885	34,753	16.9	17.4
所有贷款			31,964	168,217	200,181	16.0	100.0

来源：拉丁美洲小额信贷机构

三、计分卡工作原理以及如何测试计分卡

计分卡详细说明了未来风险与借款人、贷款和机构三者特征之间的可预见的关系。主观计分卡将明确的信贷评估原则与信贷员绝对的判断结合起来，而统计计分卡则是一套明确的规则或数学公式。本章提供了一个样例树，一种类型最简单的统计计分卡，

并且演示了如何在使用前测试它们的程序。

四叶树

图 2 和图 3 中的这个四叶树计分卡，是根据拉丁美洲一个大型小额信贷机构已还款的数据建立而成。贷款机构将"不良贷款"定义为所有被拖欠了至少 30 天的贷款或者每笔分期付款至少平均拖欠 7 天。⁴

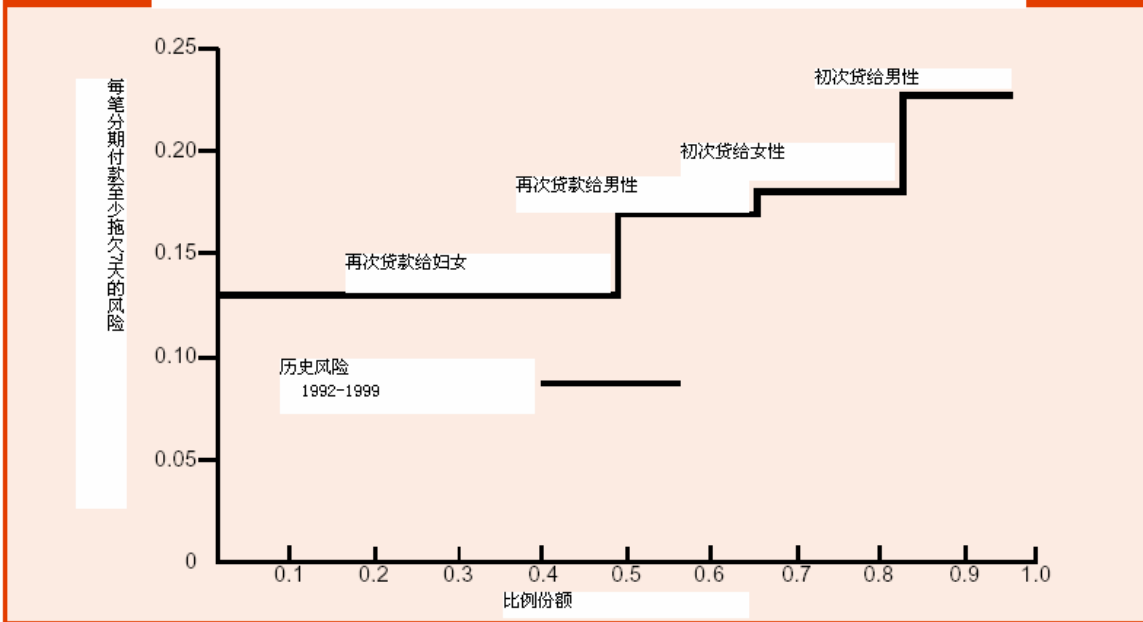
位于图 2 顶部的树根表明，在 1992 至 1999 年偿还的 200,181 笔贷款中有 31,964 笔是"不良贷款"。历史风险达到 16%，即不良贷款数占有所有贷款数的比例。图 2 中，树根以下的分枝，根据所贷款种类的不同（新贷款和再次贷款），然后再根据申请人的性别不同（男或者女），将"偿付贷款"（包括已偿还贷款和已注销的贷款）分成四片叶子。对于借新贷款的女性而言（下面左边的树叶），历史风险是 17.9%，即 9,354 笔不良贷款除以 52,395 笔总贷款。对于借新贷款的男性而言，历史风险是 22.3%，即 5,316 笔不良贷款除以 23,787 笔总贷款。对于再次贷款的女性而言，其历史风险是 12.8%。至于再次贷款的男性而言，其历史风险是 16.9%。图 4 描绘了与图 2 和图 3 相同的四叶树。四个分段代表四片树叶，按照从最小风险（左边）到最大风险（右边）的顺序排列。它们的高度代表各自的历史风险，每一个分段的长度则说明了在所有已还款中该叶对应笔数所占份额。举例来说，对于再次贷款的女性而言，对应笔数占已还款的比例是 44.6%（89,246 除以 200,181）。（见图 3，第三片叶子，最右边一栏）

这棵简单的四叶树给小额信贷机构以下几点启示：

- 当性别既定时，新的贷款比再次贷款风险大。
- 对于新贷款和再次贷款而言，贷给男性的风险比女性大。
- 风险最小段(再次贷给女性) 的风险大约是最大风险段（初次贷给男性）的一半。
- 最长段（再次给女性贷款，大约占有所有贷款的一半）风险最小。
- 最短段（初次贷给男性的贷款，大约占有所有贷款的 12%）风险最大。

⁴见 48 页有关的讨论，即为什么"坏"被定义为严重的不良行为，而不是违约。

图4 四叶树所显示的历史风险，1992—1999年数据（曲线图形式）



小额信贷机构如何根据这些启示来采取行动？因为新贷款（特别是贷给男性的）风险大，所以贷款机构也许要格外认真地筛选这一段的贷款申请。在申请低风险段，贷款机构也许要减少由信贷员完成的分析，或者降低对客户的要求。等级评分仅仅能预测风险，但是，它却无法告诉贷款机构如何去管理它。

这棵简单的四叶树所揭示的道理并不出人意料。也许大多数小额信贷机构知道，新贷款要比再次贷款风险大，贷给男性比贷给女性风险更高。但是，令一些人惊奇的是，初次贷给男性的风险几乎是再次贷给女性的两倍。这个简单的例子仅仅解释了等级评分的概念，并没有让我们由此深入了解特征与还贷风险之间先前未知的关系。

这种树是如何预测风险的呢？

等级评分假定风险及特征间过去的相关性在将来仍然成立。因此，一个分段内的历史风险就可以成为某段的预测风险。例如，假设一个有如图 4 所示四叶树结构的小额信贷机构收到一位女性再次贷款的申请，在对她进行传统的信贷评估后，机构初步批准了这个申请。而再次贷款给女性的历史风险是 12.8%，因此，从树型计分卡中得出的预测风险就是 12.8%。对于一个初次申请贷款的男性而言，如果根据贷款机构的传统标准获得了初步批准，那么他的预测风险就为相应分段的历史风险，即 22.3%。

无论是利用树型结构，还是以更加复杂的计分方式，等级评分都能够预测风险，但前提假定是，对于具有一定特征的贷款申请而言，它的未来风险将等同于具有相同特征申请的历史风险。主观等级评分也是一样，但是它是定性地、含蓄地测量历史关系，缺乏数量性和明确性。

图5 构造样本与检测样本

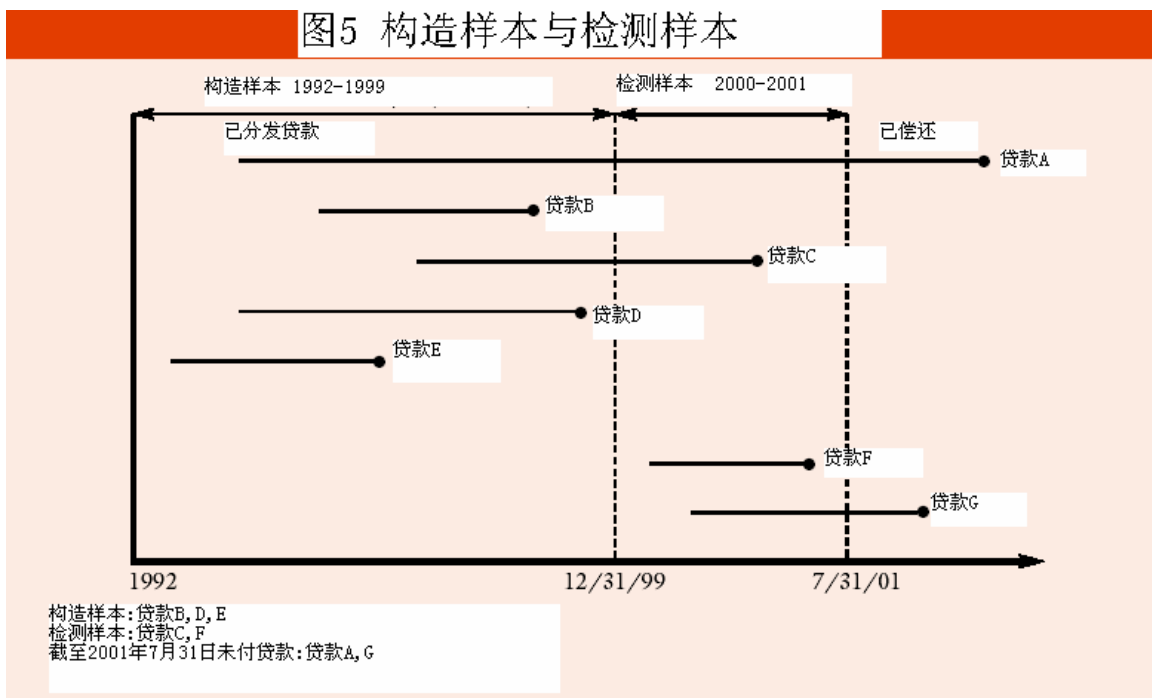


图 6：四叶树所示的实际风险，2000—2001

树枝			构造实例 2000-2001					
树叶	第一	第二	不良贷款	好贷款	叶上的实例总数	不良贷款预测比例%	不良贷款实际比例%	占叶上实例总数比例%
1	新贷款	女性	5,740	26,589	32,329	17.9	17.8	23.9
2	新贷款	男性	3,281	11,674	14,955	22.3	21.9	11.1
3	再次贷款	女性	7,752	56,575	63,327	12.8	12.1	47.6
4	再次贷款	男性	3,770	19,627	23,397	16.9	16.1	17.3
所有贷款			20,543	114,465	135,008	16.0	15.2	100.0

来源：拉丁美洲小额信贷机构

任何计分卡都可以预测风险，但是不是所有的预测都那么准。幸运的是，这种预测能力在使用前是可以测试的。历史测试表明，倘若过去你使用过计分卡，那么它表现出的效果的确很好。可以假设从现在起等级评分就会有相似的预测能力。

设想，有个想玩玩股票或者赌马的人，为了从股票市场获益或操纵马场，他搞出了一个新的系统。在投入自己的赌金之前，如果他没有用历史数据来测试这个新系统，看看它在过去能取得怎样的效果，那么他还不够聪明。同样，小额信贷机构应该在使用前也测试它们的计分卡，这样就能够避免灾难，并且会让那些持怀疑态度的人相信等级评分确实可行。

历史测试根据分发贷款已知的特征，用计分卡为已付清（或者注销的）的贷款预测风险。然后，测试工作将比较预测的风险与实际的风险，即：贷款（在分发后）最终变成好的、还是变成坏的。历史测试是等级评分的核心特点；贷款机构在没有首次测试预测能力前不应该进行评分。

历史测试

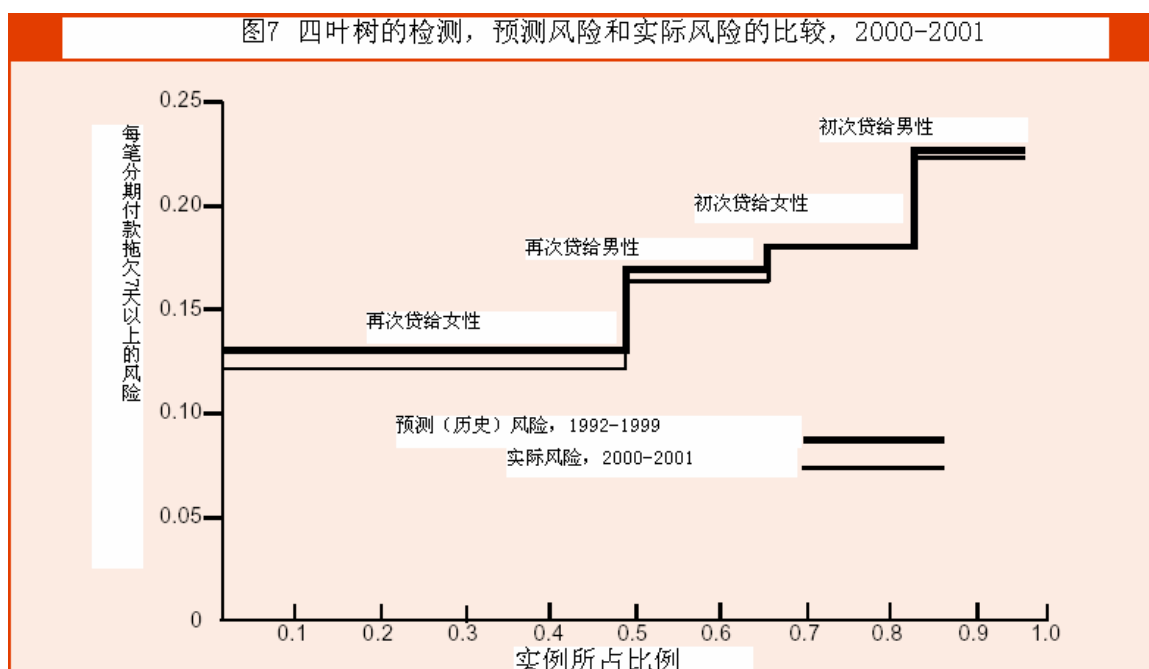
历史测试分为三个步骤：由构造样本中的贷款推导出计分卡；用计分卡来预测测试样本中的贷款风险，比较预测出的（历史的）风险与成为现实的实际风险。

历史测试将已偿还贷款（包括已注销的贷款）划分为两个样本。过去按指定最终期限偿还的贷款，组成用于建立计分卡的构造样本。在图 5 中，贷款 B、D 和 E 是在到期日之前偿清的，所以，它们进入构造样本。超过最终期限偿还，但是赶在数据库最后日期之前偿还的贷款，组成用于测试计分卡预测能力的测试样本。在图 5 中，测试样本是贷款 C 和 F，因为它们是在构造到期日以后、但在数据库记录截止期之前偿还的。到数据库截止期尚未收回的贷款（图 5 中的贷款 A 和 G）都没有纳入测试样本和

构造样本中，因为机构还不知道这些贷款的状况是好还是坏。

为了模拟实际生活中的等级评分，测试应该遵循三项原则。

第一，一笔给定的贷款要么用于构造样本，要么用于测试样本，但不能同时用于两者。在两个阶段使用同一笔贷款，这会夸大评分的预测能力。在构造阶段，必须对计分卡进行调整，使其反映构造样本中特征与风险之间明显的相关关系。但是，其中一些关系会随时间变化而变化，或者根本就不是真正的模型，它只是一个有限样本中机率发生的结果。这些关系在构造样本以外的贷款中是找不到的。因此，计分卡对构造样本中贷款风险的预测要比对其它贷款的准确。然而，在现实生活中，最重要的是预测不在构造样本中的贷款的风险。



第二，被测试的贷款必须是在构造样本贷款之后还清的。实际上，计分卡预测的是在构造样本中超过截止日期以后才偿还的贷款，这种测试应该模拟该情形。

第三，测试的根据必须是贷款分发时已知的特征。任何在放款之后获得的信息必须忽略，因为实际中的预测不能利用这种数据。

在图 3 的四叶树中，构造样本是在 1992 至 1999 年间已偿还的贷款，200,181 笔，而测试样本（如图 6 所示）是在 2000 年 1 月 1 日至 2001 年 7 月 31 日间支付的贷款，135,008 笔。给定贷款种类（新贷款或者再次贷款）和借款人性别（男性或者女性），计分卡就可以预测，测试样本的未来风险等同于具有相同特征的构造样本的历史风险。

举例来说，在图 6 中，预测再次贷款给女性的历史风险是 12.8%（叶子 3，“预测不良贷款比例”一栏），结果，在 2000 至 2001 年实现的实际风险是 12.1%（叶子 3，“实际不良贷款比例”一栏）。图 7 中，预测（历史）风险线与实际风险线之间的距离反映了计分卡的准确程度。⁵

⁵如图 7（也在图 10 中）所示的四个部分中的每一个部分，我们忽略已偿贷款随时间而变化的比例，以免造成混乱。

给男性的新贷款（最高风险段）的预测风险是 22.3%（叶子 2，“预测不良贷款比例”一栏）。这一数字同样接近 21.9%的实际风险（叶子 2，“实现不良贷款比例”一栏）。事实上，正如图 7 所示，这种树型预测的风险在所有四个分段里均接近实际发生的风险。

等级评分预测风险的假定是，风险与特征在过去的关联关系将来仍然成立。从历史角度检测它的预测能力就是将预测出的风险与已还清的贷款的实际风险进行对比。评分制的工作原理很像许多小额贷款机构已经使用的欠款等级制。但是，一旦评分制得到发展，它的使用将会更加便利，效果会更加强大（见方框 2）。

四、如何使用计分卡

信贷员和信贷经理在日常工作中将如何使用计分卡呢？本章用一棵十九叶树来说明一个贷款决策，它的根据是，将申请划入四个不同的风险等级。接下来，本章介绍了如何利用对预测能力的历史检测来制定策略界线以及估测风险、放款和利润之间的平衡点。

一棵十九叶树

如四叶树一样，图 8 中这棵十九叶树是根据一家大型小额信贷机构收集到的已还清贷款的数据构造而成。这个机构将“不良贷款”定义为逾期 30 天的贷款或者每笔分期付款平均拖欠 7 天的贷款。这棵十九叶树比四叶树有更多的树叶，但是它们的概念相同。更多的树叶能够使预测更加细化，能更清晰地区分高风险与低风险的差别。十九片叶子按七个变量被分为四个分支，这七个变量是小额信贷机构在传统评估过程中记录的一部分：

- 贷款种类（新贷款或者再次贷款）
- 电话的数量（没有，一个或者两个）
- 申请人的年龄（岁）
- 信贷员的经验（放贷的数量）
- 过去已还清的贷款中，每笔分期付款拖欠天数
- 负债率（负债除以资产）
- 担保率（动产担保的重置价值除以放贷额度）

第十一片叶子是最大的部分，占总贷款的 15%（“占叶中总笔数比例”栏），同时也是风险最小的部分，为 4.5%（“不良贷款比例”栏）。第十一部分包括再次贷款申请人，他们在过去已还清的贷款中每笔分期付款平均拖欠少于一天半，呈报拥有的电话数量为零或一，年龄超过四十岁。

方框2：等级评分与欠款等级

许多小额信贷机构根据申请人在过去贷款中的欠款情况给他们分级。评分与分级类似，只是评分更加准确。而且由于预测值之间的差别有明确的含义，所以评分的使用更加方便。如果

分级有用，评分则在三个方面更有用。

第一，评分以概率值量化风险，而分级只能给风险分类。举例来说，A级可能意味着“提供特别的激励，以保持忠诚”；B级，“接受并且允许增加额度和到期期限”；C级，“不更改条款地接受”；D级，“拒绝”。但是，放款机构没有对A级中将会变坏的比例做出预测，它也不知道A级比C级的风险高多少。相对比之下，评分不但按等级级排列风险，而且，一旦为求绝对准确（详见第五章），那么它可以确定各类风险的细致区别。举例来说，在预测风险为10%的贷款中，预计有10%的贷款出现不良贷款，这个概率只相当于预测风险为20%的贷款的一半。

第二，评分解释了风险与大量特征（包括拖欠行为）之间的关系，但是，分级却只关注欠款行为。当分级由于没有对新贷款拖欠行为进行记录而变得无用时，等级评分却对新贷款和再次贷款发挥同样的效力。

第三，等级评分利用历史数据库和统计技术，最大程度地确定风险和特征之间的相关关系。而相比之下，分级则根据策划系统的管理者的判断和经验来确定风险与拖欠行为的联系。当然，对数据库做一些简单的分析，可以使分级制度的设计更加合理，可是，管理者们很少这样做。同样，对预测能力的历史检测是等级评分的核心，但是，对于分级来说，事实上这是不存在的。

图8: 十九叶树所示的历史风险, 1992-1999年数据

树叶	树枝				构造样本, 1992-1999年				
	第一	第二	第三	第四	不良 贷款	好贷 款	贷款 笔数	不良贷款 比例%	占叶上总笔数 比例%
1	新贷款	没有电话	不适用	不适用	18 6	453	639	29.1	0.8
2		1部电话	年龄 ≤40	信贷员经验 ≤500笔	60 3	2, 459	3, 062	19.7 11.0	4.0 7.4
3				信贷员经验 >500笔	61 3	4, 980	5, 593		
4			年龄 >40	信贷员经验 ≤150笔	15 8	74 6	904 5,	17.5 8.2	1.2 7.1
5				信贷员经验 >150笔	44 6	4, 962	408		
6		2部电话	年龄 ≤40	信贷员经验 ≤700笔	99 3	3, 032	4, 025	24.7 14.6	5.3 5.5
7				信贷员经验 >700笔	61 4	3, 590	4, 204		
8			年龄 >40	信贷员经验 ≤700笔	49 0	2, 029	2, 519	19.5 11.8	3.3 3.6
9				信贷员经验 >700笔	31 9	2, 395	2, 714		
10	再次贷 款	每笔分期 付款拖欠 ≤1.5天	0或1部 电话	年龄 ≤40	67 0	9, 463	10, 133	6. 6	13.4
11				年龄 >40	51 3	10, 879	11, 392	4.5	15.0

12		2部电	年龄	98	7,	8,	11.0	11.7
13		话	≤40	0	895	875	8.2	11.4
			年龄	70	7,	8,		
			>40	6	945	651		
14	1.5天<每笔分期	0或1部	信贷员经验	47	1,	2,	22.3	2.8
		电话	≤2, 100	6	655	131		
15	付款拖欠≤7天		信贷员经验	10	96	1,	9.4	1.4
			>2, 100	0	0	060		
16		2部电	担保率	77	1,	2,	31.4	3.3
17		话	≤2.7	7	698	475	16.7	1.6
			担保率	20	1,	1,		
			>2.7	7	036	243		
18	每笔分期	负债率	不适用	10	293	40	26.9	0.5
	付款拖欠天数	≤0.03		8		1		
19	>7天		不适用		233		45.6	0.6
所有贷款		负债率		19	66,	42	12.1	100.0
		>0.03		5	703	8		
				9,		75,		
				154		857		

来源：拉丁美洲小额信贷机构

相比之下，第十九片叶子是最小的部分之一，占有贷款笔数的 0.6%（见“占叶上总笔数比例”栏），也是风险最大的部分，占 45.6%（见“不良贷款比例”栏）。它包含申请人的再次贷款，这些申请人在过去的贷款中每笔分期付款平均拖欠超过 7 天，负债率大于 0.03。

从图 8 所示的十九叶树的简单分析中，小额信贷机构可以得出一些启示。例如，尽管业务量集中在低风险段，但是有一些部分还是相当危险的。最糟糕的第十九片树叶，具有 45.6% 的风险，几乎是最好部分风险（具有 4.5% 的第十一片树叶）的十倍。小额信贷机构将会区别对待处于最高风险段的申请人与风险最低段的申请人。

与风险相关的特征如下：

- 年轻人比上年纪的人风险大。
- 在过去已偿还的贷款中，拖欠频繁的人比拖欠行为少的人风险大。
- 担保少的比担保多的风险大。
- 负债率高的比负债率低的风险大。
- 经验少的信贷员比经验丰富的风险大。
- 有一部电话的比没有或者有两部电话的风险大。也许因为在这个国家，小额信贷机构提供的这种服务，比起最贫困的人（没有电话），以及不太贫困的人（具有两部电话），更适合“中等”贫困者（有一部电话）的需要。

这些启示均与小额信贷机构的经历相吻合。这证实了等级评分的潜力以及放款机构的直觉。但是，评分不仅告诉机构已知的信息，而且它还量化与风险的关系。举例来说，放款机构已经知道在过去已偿还的贷款中，风险是随着拖欠的行为而增加的，但是，它不知道究竟增加了多少。如这棵树所示，在过去已收回的贷款中，每笔分期付款拖欠低于 1.5 天的再次贷款的风险为 7.3%（计算方法是，用叶子 10 到 13 中“不良贷款”的总数除以这些部分的贷款总数）。这个数字比相应拖欠 1.5 天至 7 天的再次贷款（第 14 至 17 部分）少 15.3 个百分点，比拖欠 7 天以上的再次贷款（第 18 至 19 部分）少 29.3 个百分点。

十九片树叶的历史测试

十九片树叶的历史测试，遵循了四叶树的相同程序。构造样本的时间段是从 1992 年至 1999 年，而测试样本的时间段从 2000 年至 2001 年。如前所述，从 1992 年至 1999 年这段时间的历史风险，被认为是对应段从 2000 年至 2001 年贷款的预测风险。于是，检测工作就可以对比预测风险与实际风险。

利用 1992 年至 1999 年数据所建立的十九叶树能够很好地预测 2000 至 2001 年的风险吗？图 8 列出了 1992 年至 1999 年十九个部分的历史风险；图 9 显示了 2000 至 2001 年实现的实际风险。图 10 将预测风险与实际风险做了比较。我们可以从三个方面来评价它的预测能力。

第一，绝对精度要看预测风险与实际风险之间的距离。图 10 中，一些距离很窄，而一些则很宽。举例来说，第十一部分（下边左角）的预测风险是 4.5%，而实际风险是 4.1%，误差大约是 9%（ $[4.5 - 4.1] / 4.5 = 0.09$ ）。但是，在第十三部分（从下边左角往上数两阶，预计风险是 8.2%，而实际风险则是 11.5%，结果造成 40% 的误差（ $(11.5 - 8.2) / 8.2 = 0.40$ ）。

第二，相对精度要看预测风险低的贷款是否比预测风险高的贷款实现的实际风险低。一个具有相对精度的计分卡可以正确地排列等级，哪怕它的绝对精度不高。这个十九叶树的相对精度就很高：

图9：十九叶树所示的现实风险，2000—2001年数据

树枝				构造样本，1992-1999年					
树叶 第一	第二	第三	第四	不良贷款	好贷款	贷款笔数	预测不良 贷款比 例%	实际不良 贷款比 例%	占叶上总笔数 比例%
1 新贷款	没有电话	不 适用	不适用	61	116	177	29.1	34.5	0.4
2	1部电话	年龄 ≤40	信贷员经验	460	1, 827	2, 287	19.7	20.1	4.6
3			≤500笔	508	3, 920	4, 428	11.0	11.5	9.0
4		年龄 >40	信贷员经验	126	436	562	17.5	22.4	1.1
5			≤150笔	387	4, 271	4, 658	8.2	8.3	9.4
6	2部电话	年龄 ≤40	信贷员经验	573	1, 293	1, 866	24.7	30.7	3.8
7			≤700笔	483	1, 603	2, 086	14.6	23.2	4.2
8		年龄 >40	信贷员经验	311	1, 005	1, 316	19.5	23.6	2.7
9			≤700笔	227	1, 164	1, 391	11.8	16.3	2.8
10 再次贷 款	每笔分期	0或1部 电话	年龄	477	6, 980	7, 457	6.6	6.4	15.1
			≤40	340	8, 027	8, 367	4.5	4.1	16.9
			年龄						

11			>40						
12		2部	年龄	612	3,465	4,077	11.0	15.0	8.3
13		电话	≤40	490	3,761	4,251	8.2	11.5	8.6
			年龄						
			>40						
14	1.5天<	0或1部	信贷员经验	447	1,526	1,973	22.3	22.7	4.0
15	每笔分期	电话	≤2, 100	144	1,079	1,223	9.4	11.8	2.5
	付款拖欠		信贷员经验						
	≤7天		>2, 100						
16		2部	担保率	527	1,015	1,542	31.4	34.2	3.1
17		电话	≤2.7	243	627	870	16.7	27.9	1.8
			担保率						
			>2.7						
18	每笔	负债率	不适用	68	106	174	26.9	39.1	0.4
19	分期	≤0.03	不适用	423	257	680	45.6	62.2	1.4
	付款拖欠		不适用						
	天数>7天	负债率	不适用						
		>0.03	不适用						
所有贷款				6,907	42,478	49,385	12.1	14.0	100.0

来源：拉丁美洲小额信贷机构

方框3：突变如何影响评分？

环境变化后，评分会失去绝对精度，^a但是，它通常保持着相对精度。在小额信贷中，变化是永恒的：竞争变得日益激烈，警察开始强制执行，或者经济开始衰弱。即使没有外部的变化，小额信贷机构随着成长，内部也在不断调整。

举例来说，小额信贷在玻利维亚的成功引起了1995至1996年智利的消费者金融公司之间的竞争^b。为了市场份额，这场争斗造成了三倍的欠款、双倍的客户退出率。评分制可以制止住这股客户退出流吗？一个客源流失计分卡（见第七章），根据1988至1996年的数据构建而成，并且接受了1997年数据的测试。它的构造样本和测试样本侧重于市场的突然变化。它的绝对精度不高，但是相对精度依然很高，具有使用价值。

a Edward M. Lewis, 《信贷评分入门》(San Rafael, 加利福尼亚: Athena出版社, 1990年).

b Elisabeth Rhyne, 《小额信贷的主流：在玻利维亚面向穷人的放款是如何起步、发展乃至成熟的》(Bloomfield, Ind.: Kumarian, 2001年); Jeffrey Poyo 和 Robin Young, “小额信贷的商业化：玻利维亚Banco Economico与Fondo Financiero Privado FA\$\$IL案例” (Bethesda, Md.: 微型企业的最佳实践, 1999年).

c Mark Schreiner, “对玻利维亚一家小额信贷机构的客户退出进行评分” (手稿, 社会发展中心, 华盛顿大学, St. Louis, Mo., 2001年)。

除了少数段之外，实际风险随着预测风险的增加而持续增加(见图 10)。总体来看，实际风险线从左边到右边向上倾斜。相对精度比绝对精度更有用，因为，正如第五段中讨论过的，经理可以使用“全球跟踪报告”将相对精度的分数转换成绝对精度的分数。另外，市场或者宏观经济的突变对相对精度的影响小于对绝对精度的影响（见方框 3）。

图10 19叶树的检测，预测风险与实际风险之比较，2000-2001年

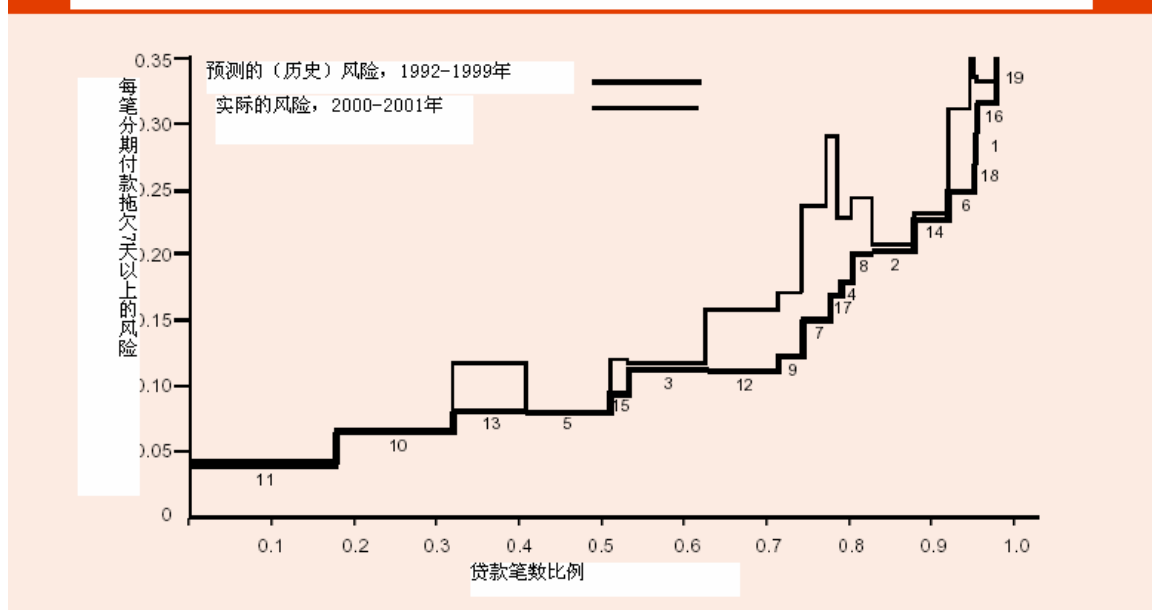
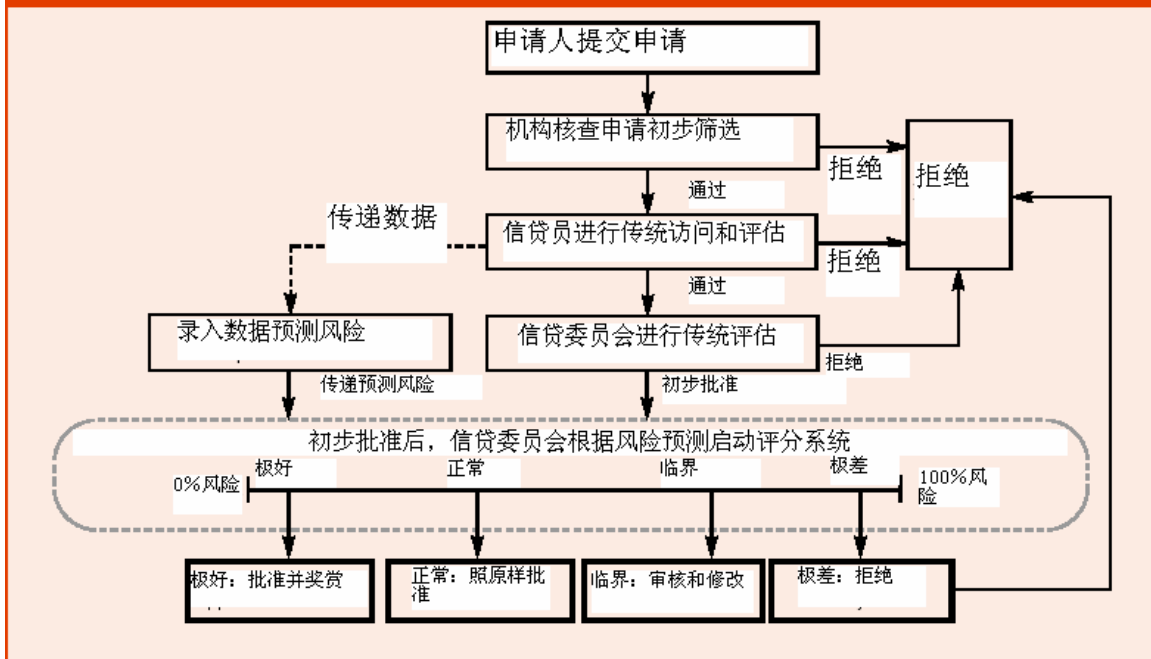


图11 一个四等级的评分政策和传统评估程序



第三，在最低和最高预测风险的贷款中，末端精度着眼于处于最关键部分的相对精度和绝对精度。毕竟，大多数贷款处于一般风险水平，而评分政策不影响中等风险的贷款。但是，它却影响风险最低的申请人（也许他们会受到特殊的奖励）和风险最高的申请人（他们的申请也许被修改甚至被拒绝）。十九叶树有极好的末端精度：具有最低预测风险的案例，同时也具有最低的实际风险，而具有最高预测风险的案例，同时也具有最高的实际风险。举例来说，具有最低预测风险的两个部分（图 10 中左下角的 10 和 11 行），同样具有最低的实际风险和很小的预测误差。具有最高预测风险的五个部分（图 10 中右上角的 6、18、1、16 和 19 行）的预测误差很大，但是它们的实际风险也最大。（树型通常具有系统预测误差和变量预测误差，特别是对于所占比例小的部分，这种情况更常见。）⁶

用四个风险等级进行评分

在对一个申请评分之前，小额信贷机构必须利用相同的信贷评估程序（即在没有评分系统时将会采用的程序）来初步批准贷款申请。在初步批准的贷款的特征给定的情况下，然后用评分预测风险。信贷委员会根据机构为贷款申请人制定的四个风险等级（极差、临界、正常和极好）的政策行事。这四个数值界线满足了机构的需要，考虑了服务覆盖面在广度、深度、和长度上的平衡。⁷因为评分忽略了定性特征而仅考虑量化特征，所以它不能取代传统评估（见方框 4）中的任何步骤。它仅仅是作为在传统程序结束后、赶在分发贷款之前附加的一个步骤。

图 11 用评分的方式为小额信贷机构描述了一个典型的评估程序。它从客户提交申请开始。在信贷员进行实地考察前，机构要根据基本的政策规定筛选贷款申请，例如，申请人最少要有一年的从业经验。如果申请通过这个关口，信贷员将进行实地考察（也许在办公室里

⁶ David J. Hand, Heikki Mannila, 及 Padhraic Smyth, "数据开发的原则"(剑桥, Mass.: 麻省理工学院出版社, 2001 年)。

⁷ Mark Schreiner, "覆盖面: 小额信贷社会收益的讨论框架"《国际开发期刊》第十四期 (2002 年): 第 591-603 页。

做一些简单分析), 然后决定是否将本申请提交给信贷委员会。通过这个阶段后, 申请被输入信息系统。该系统计算一个分数, 并且打印评分报告(评分模拟和特征报告影响在第五章讨论), 该报告将与其他例行报告一同交予信贷委员会。

方框4：为什么仅对按传统评估程序初步批准的实例评分？

被评分漏掉但是被主观评估捕捉到风险的比例很大，反之亦然。原则上讲，评分既可以在主观评估之前，也可以在其之后。如果评分在前，并且是预测低风险，那么放款机构就有可能在（成本昂贵的）主观评估上偷懒。这是很危险的，因为从定量因素看似低风险的贷款，在考虑主观因素后很可能具有高风险。因此，小额信贷机构应该仅仅针对已经通过主观评估程序、得到初步批准的申请评分。总体上，还款风险可以根据它与借款人、贷款和机构三者量化特征之间的关系分为三部分：

- 随机风险根本与任何特征（不管量化与否）不相关。
- 显化的风险与量化特征相关。
- 定性风险与非量化特征相关。随机风险（如雷电）是不可预见的。

评分方法测量而且仅仅测量显化风险。它揭示出相互关系，但不是原因；它只是指出它的存在，而不解释为什么特征属性同风险相关。最后，传统的小额信贷评估同时考察显化风险和定性风险。与评分方法相比，传统评估处理定性风险（评分是忽略定性风险的）更为出色，但是它对显化风险的处理不如前者。

打算利用评分方式跳过（或者敷衍）传统评估的小额信贷机构怀有侥幸心理，它们认为，已经通过的申请人的定性风险基本上等同于用传统评估方法初步批准的申请人的定性风险。

这就是说，定性风险不重要而且无法衡量，但是这种假设完全有悖于当前最流行的小额放贷技术。定性风险到底有多重要？我们能够知道的只是已分发贷款的表现，所以历史测试无法揭示那些在传统评估程序中由于定性风险原因而被拒绝的贷款申请倘若通过，它的业绩表现会如何。用评分方法替代主观筛选的小额信贷机构实在是作茧自缚。除非定性风险根本不重要，否则预测的风险将会过低。要想准确知道有多低，唯一的办法就是不经过主观筛选就批准一些贷款，然后看看结果如何。

信用管理局制度会随着时间发展得越来越好，不断普及，不断完善，而小额信贷机构将量化更多的特征。随着数据越来越多、越来越好，也许评分系统能够免去对主观风险评估的需要，但是，到目前为止这还不得而知。有一家玻利维亚成立的金融公司，仅仅通过评分来判断小额信贷借款人的风险，结果破产了^a。现在，评分系统是信贷员和传统评估程序的补充而不是替代。

^a Elisabeth Rhyne, 《小额信贷的主流：在玻利维亚面向穷人的放款是如何起步、成长乃至成熟的》(Bloomfield, Ind.: Kumarian, 2001).

图 12: 利用 19 叶树 (由 1992-99 年数据建造而成) 在 2000-01 年度使用四级评分策略得出的结果

叶	第一	第二	第三	第四	不良贷款	好贷款	总笔数	预测不良 贷款比例%	实际不良贷款 比例%	占叶 比例
所有贷款					6,907	42,478	49,385	12.1		14.0
极好:										
11	再次贷款	每笔分期支付拖欠≤1.5 天	0 或 1 部电话	年龄>40	340	8,027	8,367	4.5		4.1
极好贷款总计					340	8,027	8,367	4.5		4.1
正常:										
10	再次贷款	每笔分期支付拖欠≤1.5 天	0 或 1 部电话	年龄≤40	477	6,980	7,457	6.6		6.4
13	再次贷款	每笔分期支付拖欠≤1.5 天	2 部电话	年龄>40	490	3,761	4,251	8.2		11.5
5	新贷款	1 部电话	年龄>40	信贷员经验>150 笔	387	4,271	4,658	8.2		8.3
15	再次贷款	1.5 天<每笔分期支付拖欠≤7 天	0 或 1 部电话	信贷员经验>2100 笔	144	1,079	1,223	9.4		11.8
3	新贷款	1 部电话	年龄≤40	信贷员经验>500 笔	508	3,920	4,428	11.0		11.5
12	再次贷款	每笔分期支付拖欠≤1.5 天	2 部电话	年龄≤40	612	3,465	4,077	11.0		15.0
9	新贷款	2 部电话	年龄>40	信贷员经验>700	227	1,164	1,391	11.8		16.3
正常贷款总计					2,845	24,640	27,485	9.0		10.4
临界值										
7	新贷款	2 部电话	年龄≤40	信贷员经验>700	483	1,603	2,086	14.6		23.2
17	再次贷款	1.5<每笔分期支付拖欠≤7 天	2 部电话	担保率>2.7	243	627	870	16.7		27.9
4	新贷款	1 部电话	年龄>40	信贷员经验≤150	126	436	562	17.5		22.4
8	新贷款	2 部电话	年龄>40	信贷员经验≤700	311	1,005	1,316	19.5		23.6
2	新贷款	1 部电话	年龄≤40	信贷员经验≤500	460	1,827	2,287	19.7		20.1
14	再次贷款	1.5 天<每笔分期支付拖欠≤7 天	0 或 1 部电话	信贷员经验≤2100	447	1,526	1,973	22.3		22.7
临界贷款总计					2,070	7,024	9,094	18.1		22.8
极差										
6	新贷款	2 部电话	年龄≤40	信贷员经验≤700	573	1,293	1,866	24.7		30.7
18	再次贷款	每笔分期支付拖欠>7 天	资产负债率≤0.03	不适用	68	106	174	26.9		39.1
1	新贷款	没有电话	不适用	不适用	61	116	177	29.1		34.5

16	再次贷款	1.5 天<每笔分期支付拖欠≤7 天	2 部电话	担保率≤2.7	527	1,015	1,542	31.4	34.2
19	再次贷款	每笔分期支付拖欠>7 天	资产负债率>0.03	不适用	423	257	680	45.6	62.2
极差贷款总计					1,652	2,787	4,439	28.4	37.2

关于这一点，等级评分并没有改变传统评估程序的任何内容；它的使用仍然要等待申请的初步批准。那么在什么时候呢？如果信贷委员会对上报的所有申请都不经审核就批准，那么当信贷员决定提交申请时，初步批准就发生了。在这种情况下，该委员会靠评分来决定哪些申请需要深入调查，哪些不需要修改就可以通过。但是，如果由该委员会本身实施初步批准，那么评分结果必须在传统的筛选程序进行之后才能公布。（如果委员会先偷看到评分结果，他们也许会忍不住放弃定性风险判断就批准贷款。）分数在手，委员会就可以实行四级评分制度了（见图 11 最底行）。

“极好”风险等级

预测风险在最低限值以下的申请人，被归类为极好。为了保持这些低风险客户的忠诚度，放贷机构也许会采取一种策略，如通过提高信贷额度、减少收费、为及时还贷提供折扣、或者降低担保要求等办法，来为他们提高贷款价值。评分仅仅是识别“极好”客户的方法，但并不是保持这类申请人忠诚度的最好方法。它仅仅预测风险，后面的事情还需要经理去做决定。如果他们想根据风险来定价，那么在预测风险给定的情况下，他们还必须决定如何调整利率。

为了便于讨论，假设在图 12 中十九叶树上，极好的临界值是 5%。那么，所有预测风险低于 5%的贷款都被视为“极好”。所有“极好”的贷款都在第 11 片叶子上，其预测风险（根据 1992 年至 1999 年的数据）为 4.5%。极好贷款占有所有贷款笔数的 16.9%。

方框5：评分策略适用于再次贷款吗？

再次贷款申请人都有还贷记录，所以针对他们的评分工作要比对新申请人更有效。但是，一些小额信贷机构不愿意修改再次贷款的界线，更不用说拒绝“极差”的再次贷款了。其中一些原因是，他们怀疑评分的能力，另外的原因是，他们想通过提供额外贷款的方式来维护奖励守信用还款的名声。

如何去做？计分卡应该考虑贷款种类（新贷款或者再次贷款）以及他们的还款记录。实际上，在先前贷款中不经常欠款的再次贷款者风险很低，倘若出现这种情况，准确的计分卡会有所反映。但是，评分有可能接触到一些风险高却没有污点记录的借款人。如果历史测试没有拒绝再次贷款的申请，那么这些申请，事实上具有很高的风险。

而且，信贷机构不能拒绝这些申请人，因为它将向现在的借款人发出错误的信号，还因为信贷委员会宁愿拒绝评分，也不愿拒绝有极好记录的再次贷款申请。在这些情况下，处理极差贷款的策略应该是对评估进行仔细的审核，修改贷款合同，并且在放贷后进行有预防作用的“礼节性拜访”。

这“极好”的临界值 5%，表现如何呢？在 2000 至 2001 年，评分系统把经传统评估程序批准的 16.9%的贷款定为“极好”（见图 12 和 13）。在这些贷款中，有 4.1%变坏了，占

所有不良贷款的 4.9%。从另一个角度看，在“极好”的贷款中，每一笔不良贷款对应有 23.6 笔好贷款。评分既识别低风险又识别高风险贷款，它帮助积极的贷款机构处理两种极端风险。不想奖励低风险者的机构，可以选择将“极好”的界定值设置为零，因为预测的风险从来不可能那么低。

“正常”风险等级

预测风险超过“极好”的界定值、但是低于临界线界定值的申请人，称为“正常”。评分再次证实了这些申请的初步批准后，信贷委员会立即批复放款，并且对申请不做任何改动。大多数初步批准的申请符合正常贷款条件，所以，在大多数情况下，评分不影响评估，也不会给信贷委员会增加额外的成本。

再次假设，图 12 和 13 中十九叶树的“正常”界定值是 12%（叶子 10、13、5、15、3、12 和 9）。在 2000 至 2001 年，所有贷款中有一半以上（55.7%）是“正常”的，其预测风险大于“极好”的界定值 5%，但是小于“正常”的界定值 12%。在这些“正常”贷款中，有 10.4% 变坏了，占有不良贷款的 41.2%。每笔不良贷款对应 8.7 笔好贷款。

“临界线”风险等级

预测风险超过“临界”界定值、但是又低于“极坏”界定值的申请人，属于“临界线”等级。信贷委员会格外小心地审核这部分贷款，如果有担保，他们就修改放款额度、期限、担保要求、或者利率或费用（根据风险程度定价）等。该委员会还有可能拒绝一些“临界线”等级的贷款。

评分增加了信贷委员会评估临界线等级申请的时间。尽管大多数警惕的放款机构喜欢在登记这些临界等级申请之前就有机会处理它们，但是这增加了成本。

在图 12 和图 13 中，假设十九叶树的临界等级界定值是 23%（叶子 7、17、4、8、2 和 14）。在 2000 至 2001 年，临界等级贷款（预测风险大于正常界限，即 12%，但是又小于临界线值，即 23%）占有所有贷款的 18.4%。在这些临界线等级贷款中，有 22.8% 变坏了，占有不良贷款的 30%，每一笔不良贷款对应 3.4 笔好贷款。

“极坏”风险等级

预测风险超过最高界定值的申请被视为“极坏”。除极少实例之外（见方框 5），“极坏”贷款一概被拒绝。该委员会可能会审核“极坏”的贷款申请，看看是否有漏掉的，或者检查是否存在具有决定性、积极的定性因素可以推翻凭借评分而做出的决策。

回到图 12 和图 13，假定风险大于 24% 的贷款是“极坏”等级（叶子 6、18、1、16 和 19）。在 2000 至 2001 年，风险超过 24% 的贷款占有所有贷款的 9%，所以它们被定为“极坏”等级。在“极坏”贷款中，有 37.2% 变坏了，占有不良贷款的 23.9%。在它们当中，每一笔不良贷款对应 1.7 笔好贷款。

那些忽略历史测试的贷款机构，会深受那些高风险贷款之苦，因为倘若没有评分环节，它们早就被批准了（见方框 6）。机构可以将“极坏”设定为 100% 从而取消“极坏”等级，因

为风险从来不会那么高。

四级评分策略奖励低风险贷款，并且审核、修改或者拒绝高风险贷款。大多数贷款具有接近平均水平的风险，因而评分对于它们是没有效果的。它只能对信贷员或者信贷委员会给予的初步批准进行证实，所以，按传统标准拒绝的贷款，当然也会被评分所拒绝。

为评分策略制定界定值

界定值的选择取决于特定小额信贷机构对评分的预测能力以及它们对机构各种目标权衡的结果：⁸服务面的广度（贷款数量），深度（借款人的贫穷程度），以及长度（机构获得利润的持久性）。一个小额信贷机构必须为自己做出这些价值判断。然后，历史测试才可以指导它制定评分策略，以优化它的目标。具体方法是，该策略要显示假定不同的界定值如何影响批准的贷款的数量、错失的好贷款的数量和避免的不良贷款的数量。（这里假设，历史测试会揭示实际应用的未来结果。）

例如，图 14 表明了一棵具有极坏界定值范围的十九叶树的结果。如果该机构把 2000 至 2001 年的“极差”值设定为 24%，那么，为了避免一笔不良贷款，1.7%的好贷款就错失掉了。虽然避免了 23.9%的不良贷款，但是 9%的贷款也被评分策略拒绝掉了。

⁸ Schreiner, “覆盖面”。

方框6：评分将坏的贷款与好的贷款同时丢掉

一些被当作“极坏”拒绝的，本来是好的申请人，就如一些临界贷款如果不做修改本该是好的一样。一些人明白这一点后，他们几乎不可能接受统计评分方法。当然，传统的主观评估也不必要修改某些贷款以至于错误地拒绝某些申请人。也就是说，主观评分也将好的和坏的一起抛掉了。但是，历史测试可以通过统计评分量化预测误差，从而完善了严与松策略之间的合理选择。运用主观评分无法知道预测误差，所以，选择不容易优化。举例来说，在拉丁美洲，一些小额信贷机构在做个人贷款时，严格的程度就像吝啬鬼。在哥伦比亚，一个著名的小额信贷机构拒绝了一半的申请人、2/3的新申请人。玻利维亚还有一家更有名气的机构，几乎从来不准申请人请求的额度和期限。在这种严格的程度下，如果小额信贷机构更了解真实的风险或做好服务面的权衡，他们也许会更好地满足客户需求，并且保持甚至减少风险，这种情况是可能出现的。

图14：19叶树中在“极差”值范围内避免不良贷款而丧失好贷款的比例

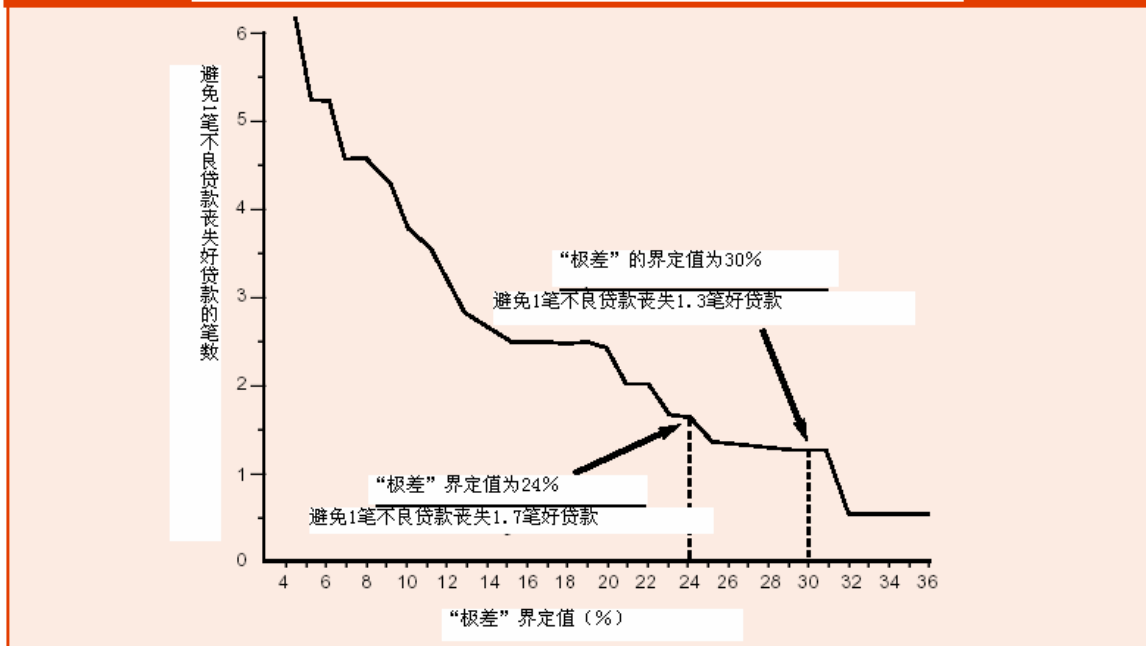
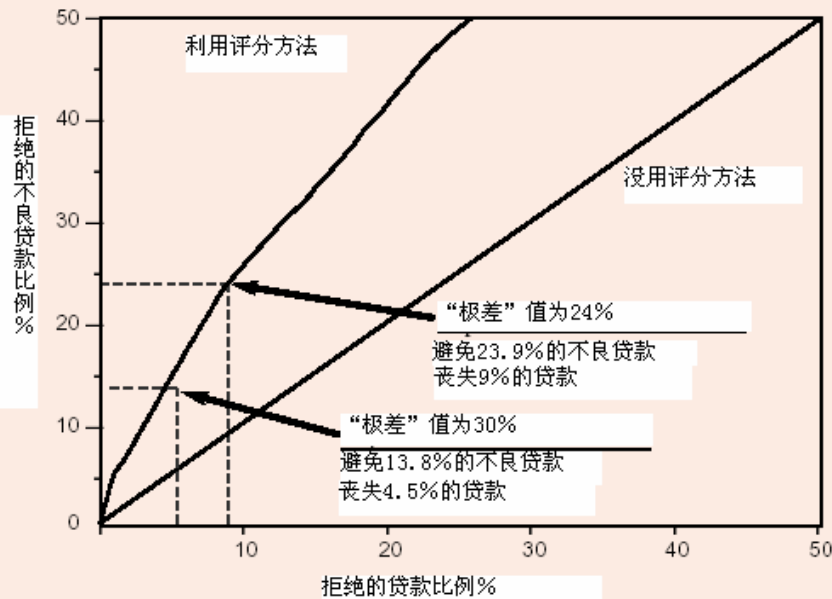


图15: 在19叶树的“极差”界定值范围内，拒绝贷款比例与避免不良贷款比例的对比



如果“极差”的界定值变化了，举例来说变为 30%，那么情况会发生何种变化呢？图 14 表明，每避免一笔不良贷款，就有 1.3 笔好贷款丢失了；图 15 表明，利用评分策略拒绝了 4.5%的贷款而避免的不良贷款比例是 13.8%。如果给出各种界定值可能产生的结果，那么，小额信贷机构就可以通过历史测试选择最适合自己目标的一种。

评分还表明了风险与显示扩展深度的特征（例如性别、收入或者年龄）之间的相关关系。这能够说明扩展深度和风险间的权衡结果。举例来说，评分可以表明，如果其他条件不变，仅能维持生计的农民出现拖欠逾期 30 天的可能性要比平常高 2%。这种信息便于小额信贷机构在深度（向处于生存线上的农民贷款）与广度（通过避免最大风险覆盖更多的借款人）和长度（通过避免最大风险赢得更多利润）之间做出明确的权衡。当然，在知道一笔贷款是有风险的，小额信贷机构不一定拒绝它。

没有任何理由强迫面向穷人客户的小额信贷机构拒绝高风险的贷款，但是，它们也不想忽视风险预测。没有人放款全然不顾风险，即使是最有献身精神的小额信贷机构，它们也会为实现一定的扩展深度而限制所能接受的成本。评分只是为权衡提供线索；机构仍然必须自己决定如何去做。另外，拒绝并非总是伤害申请人。小额信贷不仅有利，而且还有害，特别是对那些最穷的人。⁹一些高风险的贷款，即使没有变坏，借款人也会殚精竭虑为按时归还分期付款而努力，而在这种情况下，如果申请在第一次被拒绝，他们的日子就会更好过些。

如果已估计出错失好贷款的净财务成本和避免不良贷款的净财务收益，那么，评分会有助于在广度和长度（利润）之间直接估算首轮的结果。这种影响可能会大得惊人。如果做出合理的假设，那么，在 2000 至 2001 年的十九叶树中，24%的“极差”界定值也许

⁹ Paul Mosley, “玻利维亚的小额信贷与贫困”《开发研究期刊》第三十七期，第四号(2001年)：第 101-32 页。

会为小额信贷机构节约 20 万美元以上的费用（见方框 7）。

在实践中（如图 12 和 13 中的十九叶树所示），大多数小额信贷机构倾向界定值的结果是，10%为极好、60%为正常、20%为临界和 10%为极差。这种宽度模式具有四种优点。首先，它保持很低的“极好”比例，便于贷款机构向他们最好的客户提供特殊激励，并且还能控制激励成本。第二，大多数贷款为“正常”，因此，评分不会改变针对多数贷款的标准评估程序，这可以进一步使一线工作人员安心，并且鼓励他们接受评分系统。第三，大多风险贷款处于“临界线”等级，信贷员和信贷经理不情愿仅靠评分方法而拒绝申请人。将多数风险借款人归类到“临界线”，可以促使信贷委员会不是拒绝，而是审核风险贷款，并且考虑调整贷款合同的条件。第四，“极差”部分比例很低。其中一些“极差”贷款是极端危险的。因为一组占很大比例的贷款要变坏，所以，信贷员会倾向于注意还贷表现（和他们奖金）上的差异。随着时间的推移，这会增强员工对评分系统的信心。

通过界定值实现贷款比例在宽度范围内的重新分配，评分可以同时提高服务面的广度、深度和长度。扩展的广度有可能增加，这是因为信贷员完全可以将因拒绝一些极端危险贷款而节约出来的催款时间用于提高放款，从而补偿拒绝贷款带来的损失。扩展的长度（持续性获利）可以增加，这是因为增加的放贷量产生的收入有可能超过评分产生的成本。扩展的深度可能增加，这是因为在增加的贷款量中，可能有一部分流向更穷的借款人。总之，评分是一种创新，它提高效率，因而有潜力维系各方面正常水平的平衡。¹⁰如果评分帮助小额信贷机构用较少的费用做成更多的事，那么，它将行百善而无一弊。

评分的成本

评分具有五种成本：数据收集成本、构造成本、操作成本、策略导向成本和过程成本。首先，收集和输入数据以建立一个计分卡，这形成数据收集成本。对于最初级的小额信贷机构，这不仅仅包括原原本本地输入收集到的申请数据，还包括提高信息系统的性能以处理额外的信息。对于这些贷款机构来说，评分应该不是首选；除了他们使用的评分以外，改进他们的信息系统是更值得的。

方框7：估算评分对利润的影响

贷款机构可以估算评分对利润影响的效果，甚至可以在实行评分系统之前。这种盈利性评估可以使风险投资者确信评分是值得的。^a

如果给出“极差”的界定值，历史测试就会显示出每避免一笔不良贷款而丧失的好贷款的数目。那么，假设贷款机构知道签下一笔好贷款的平均净财务收益，也知道做一笔不良贷款的平均财务成本（这种成本大多是信贷员花在收款上，而不是市场营销、评估和放款上的时间的机会成本）。事实上，即使小额信贷机构以利润为导向进而（仅仅盲目地）推行放贷策略（不管使用还是不使用评分系统），也很少有机机构测算过这些利润和成本。但是，贷款机构知道，一笔不良贷款的成本远远超过一笔好贷款的收益。举例来说，在富裕国家，信用卡发放机构一般认定，用十笔以上的好贷款才能弥补一笔不良贷款的损失。如果贷款机构采用计分卡，那么不良贷款的数量减少了（减少成本），好贷款的数量（至少是头一轮的效果）也减少

¹⁰ Elisabeth Rhyne, "小额信贷的阴阳：接近穷人并保持可持续性"《小额信贷公告》第二期(一九九八年)：第 6-8 页；以及 Claudio Gonzalez-Vega, "小额信贷：更广泛的成就与新的挑战"《经济与社会不定期论文》第 2518 号 (俄亥俄州，哥伦布：俄亥俄州立大学，1998 年)。

了（减少收益）。评分对利润的实际效果可以按如下测算：
 （不良贷款的单位成本 x 避免的不良贷款数目） - （好贷款的单位收益 x 失去的好贷款数目）。

对于19叶树，不良贷款的假定成本是300美元，好贷款的假定收益是100美元。使用24%的“极差”界定值，历史测试（图12的最底行，“贷款总笔数”一栏）表明，4439笔贷款会被当作“极差”贷款。在这些之中，1652笔结果变成不良贷款（见“不良贷款”栏），而2787笔结果是好贷款（见“好贷款”栏）。那么，每一笔不良贷款对应1.7笔好贷款。如果所有“极坏”贷款按照2000至2001年的策略全部被拒绝，那么所估计的利润变化将为：

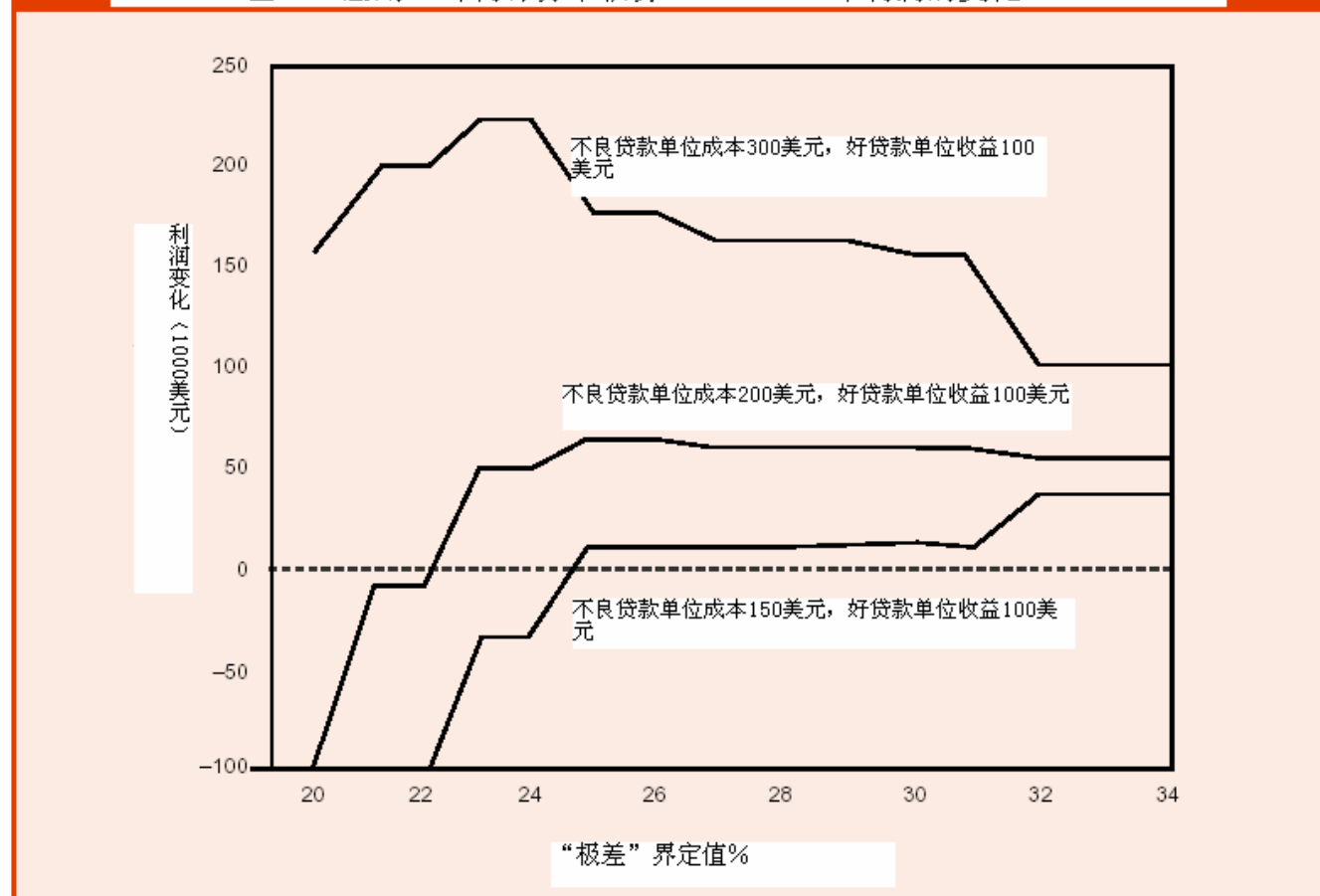
$$(\$300 \times 1,652) - (\$100 \times 2,787) = \$216,900.$$

即使仅仅拒绝在第19片树叶的（最危险的部分，见图12，“占叶上所有贷款笔数比例”一栏）1.4%的申请人，也会增加101,200美元的利润： $(\$300 \times 423) - (\$100 \times 257) = \$101,200$ 。

图16按照不良贷款的成本和好贷款的收益之间三组可能的假定值，列出了十九叶树不同的利润变化。在此有两点应该注意。第一，盲目遵循一个粗心设定的“极坏”值有可能促使信贷机构很快破产。第二，不良贷款的成本相对于好贷款的收益的比率越大，评分的潜在收益就越大。

a John Coffman, “针对微型、小企业放贷的评分入门”（该论文提交于世界银行关于“来自微型企业贷款的利润”主题的会议，华盛顿特区，2001年4月2日至3日）

图16：应用19叶树计分卡估算2000—2001年利润的变化



大多数发展水平较高的小额信贷机构数据积累的成本已经降低；所有申请均是原原本本地按标准输入。对于这些贷款机构来说，一旦数据库有足够的案例支持计分卡的构建，评分立即可以实行。还有第三种信贷机构，他们拥有足够的信息系统，但是还没有输入申请信息。与其雇佣一支录入人员的队伍输入存档的书面申请，他们还不如立即着手捕获电子形式的数据。

第二，评分项目本身（计分卡的建立以及和信息系统、培训和跟踪的整合）产生一次性创立成本。特别是调整信息系统以便自动计算和报告风险预测，可能是一个无法预计的漫长而又困难的过程，这会耗费项目预算的很大一部分。事实上，许多评分项目在这个阶段上失败了。

第三，日常的评分工作需要数据录入人员、信贷员和信贷经理投入时间，这造成操作成本。这些成本是低的。举例来说，信贷员已经收集到了计分卡所需的大多数特征。信息系统负责计算分数，所以，主要的操作成本是信贷委员会用于审核临界等级贷款上的额外时间以及正在进行的人员培训的成本。

第四，奖励“极好”者或者拒绝“极差”者，都产生策略成本。奖励不是总有效果；而一些“极差”的贷款，如果当初批准通过，它们也许会有好的结果。

第五，也是最重要的，评分的出现使机构处于不断变动中，这产生了过程成本。一些权力从信贷部门转移到信息部门。一些雇员可能会公开地反对评分所带来的变化，其他人可能用编造的数据或者忽略政策规则巧妙地避开评分系统。还有一些人在传统评估中敷衍了事，非有意地破坏评分效果。培训和跟踪（见第五章）是管理这些过程成本的最好方法。

评分的收益

评分的收益包括贷款损失的减少，客户忠诚度的提高以及根据风险（基于风险的定价）调整利率和费用的能力。最重要的是评分可以减少催款的时间，有利于小额信贷机构进行明确的定量分析便于管理者做出决策。

减少贷款损失也许是评分带来的最小收益，倘若如此，那是因为大多数使用评分的小额信贷机构很少遭受拖欠的困扰。“极好”贷款带来的更大忠诚度，可能比减少贷款损失这种收益更大。

给定一个分数，小额信贷机构可以拒绝贷款申请或者修改贷款合同从而避免风险。有的修改试图用增加利率或费用的办法来补偿风险。但是在实际操作中，搞清楚如何调整价格是很复杂的，特别是在没有对各种成本和收入组成进行精确计算的时候。

评分的最大收益在于，信贷员花费在收款上的时间减少了，他们可以有更多的时间开发新的业务。不良贷款成本高，大多因为催收贷款花费了很多时间。评分对利润产生影响（见方框 7），是因为拒绝“极差”贷款和修改“临界线”贷款意味着信贷员必须找出更少数量的不良贷款。然后，他们可以腾出手来，把节省下来的时间用在市场推广、评估和放贷上。许多小额信贷机构期望评分能使他们更多地节约评估时间而不是收款时间。但是，大

多数信贷员花费在收款上的时间和评估上的时间相同，这里必须再次重申，评分不能代替定性评估（见方框 4）。

假设信贷员每周收款可能花费两到三天的时间。假设，在十九叶树中“极差”的界定值是 24%，那么，评分减少大约 10%的放贷，同时减少约 25%的不良贷款（见图 17）。再（保守地）假设，在评分前，信贷员每周花费两天收款，那么，评分每周为他们节省的时间为半天（两天的 25%）。

继续假设，信贷员每周一般用两天的时间做营销、评估和放贷。如果他们利用额外的半天时间照以前那样有成效地招揽新客户，那么放贷量将增加 25%。在扣除由评分所拒绝的 10%的“极差”贷款以后，评分的最终结果是不良贷款减少 25%，放贷量增加了大约 12.5%。方框 7 讨论了可能出现的最基本的影响。

即使评分使一些本该可以得到批准的贷款被拒绝，它还是可以提高服务面扩展的广度和长度。深度会怎么样呢？在高收入国家，评分已经增加了深度。¹¹大多数家庭已经可以获得有史以来最灵活的小额信贷产品——信用卡，因为利用评分系统可以以较低成本评估数量巨大的小型、短期、无担保贷款的风险。

在小额信贷中，评分也应该增加深度。首先，用每周多出的半天时间去找新客户，这有可能使信贷员在他们的总贷款量中增加给穷人的数量。（即使大多数新借款人相对富裕，至少还有一些比较贫穷。）第二，评分还保护一些贫穷的借款人以免他们自身做出最糟糕的判断。

拒绝或者修改高风险的贷款，不仅减少了贷款机构的成本，而且还帮助了有些借款人，因为如果真的得到贷款，他们反而忧心忡忡、担心催款的到访，并且为了拼命还清债务他们还要变卖家产。评分可以减少小额信贷对这些人的伤害。第三，也是最根本的，小额信贷的前提是穷人是值得信赖的，而不是贷款机构缺乏合适的工具来判断他们的风险。评分完善了风险评估工具，因此有助于消除评估过程（见方框 8）中的偏见和错误。如果穷人真的是值得信赖的，那么评分将有助于展示前所未有的、更加深入的客户覆盖面。

也许从长远来看，评分最为重要的收益是，直接从数据库分析得来的量化权衡信息有助于刺激管理层的决策欲望。举例来说，一旦经理运用权衡结果建立了一种评分策略，如图 12 所示的那样，他们肯定不愿回到原来的状态，即仅凭感觉就模糊地判断其他信贷策略的结果。

金融离不开信息，而许多小额信贷机构数据库中的信息是一个尚未开启的宝库。评分经验可能会促使小额信贷机构指定一到两个雇员专门负责“数据开采”（利用历史信息预测未来的行为），从而为商业决策提供信息。预测还贷风险（信用等级评分）就是一个例子，但是，“数据开采”还能够预测客户退出的风险，¹²或者预测那些最有可能回应市场的潜在客户类型。¹³在室内进行的数据挖掘不必过于复杂。例如，简单的记录表（就如文

¹¹ W. Scott Frame, Michael Padhi, 及 Lynn Woosley, "等级评分对低收入和中等收入地区的小额商业贷款的影响"《亚特兰大联邦储蓄银行工作论文》，第 2001-6 号(2001 年)； Peter McCorkell, "等级评分 101"；与 Edward M. Lewis, "等级评分的引入"

¹² Mark Schreiner, "对玻利维亚一个小额信贷机构客户退出风险的评分"。

¹³ Rhonda Delmater 和 Monte Hancock, "数据开发的解释：一位经理的以消费者为中心的商业意识"(波士顿, Mass.: 数

中的实例树) 花费不多, 然而很有信息价值。简单、实用的分析加上灵活的变通促使经理跳出现在信息系统提供的范围, 促使他们开始思考哪类信息更有助于他们的决策。

图17: 评分的收益—信贷员在收款上节省了时间

活动	评分前		评分后	
	时间比例%	每周的天数	时间的比例%	每周的天数
会议和日常管理	20	1	20	1
营销、评估和分发贷款	40	2	50	2.5
收款	40	2	30	1.5
变化:	由于信贷员时间增加而增加的贷款申请			25%
	由于使用评分系统而减少的批准数量			-10%
结果:	已批准贷款申请的净增加			12.50%

来源: 假设的例子

五、培训员工, 建立评分政策, 以及监督操作

纯粹从技术角度讲, 小额信贷评分的工作原理如前所述。但是, 从人的角度讲, 评分并不如想象中的简单。利用评分完善决策, 不仅仅需要对工作原理有理智的理解, 而且还需要发自内心地对评分的信任, 此外再加上尝试创新的恒心。信念来自于理解和感知, 尝试创新的意愿来自可见的利益。最后, 评分系统成功的关键更多的来自培训和仿效, 而很少依靠技术手段。

培训是评分的核心, 因为风险投资者—资助者、高层管理者、信贷经理以及信贷员—在正常情况下心里会有些疑问。为了理解和接受评分系统固有的体例, 机构需要经过数月的反复培训与消化。

首先, 一个可能操着陌生口音(如果他或她确实会讲这种语言)的顾问来到这里, 他没有会见小额信贷机构的员工或客户, 就声称有一个秘密的计算公式, 这个公式可以帮助一线工作者完成他们最困难的工作: 推测出到底谁值得信贷。第二, 评分有别于依靠小组联保贷款或者信贷员亲自访问的传统小额信贷评估。新的方法不是靠个人对借款人品行的评估, 而是靠机构对量化特征的计算。

字出版社, 2001年); 以及 Michael J. A. Berry 和 Gordon Linoff, "掌握数据开发: 消费者关系管理的艺术和科学"(纽约: John Wiley 和Sons, 2000年)。

方框8：统计评分是一种歧视吗？

统计评分确实存在歧视：它假定每一个申请人只是同一个旧情况的翻版，而不可能是有别于数据库中其它明显相似案例的独特新版。但是，主观评分存在的歧视如果不比统计评分多，至少也和它一样。信贷员根据他们和他们的顾问从其他借款人身上获得的信息评估风险，而不是靠开发经验或者偏见以外一些神奇的智慧来评估。在确实独特的案例中（或者如果小额信贷机构刚刚创立或信贷员是个新手），因为没有经验，所以决策只能从随意的猜测或者偏见中得出。

根据相似的其他人的经历来评估一个人是不公平的，但是，又没有其他方法可选择。仅有的不存在歧视的贷款机构是那些批准所有贷款申请的机构。因此，问题不是是否存在歧视，而是如何尽可能公平合理地对待这种歧视。这种合理的“歧视”就是将两种类似的情况做对比。举例来说，统计评分将申请人和先前在同一个贷款机构贷款并具有相似量化特征的借款人做对比。如果女性比男性有更好的还款历史，那么，计分卡就会如实反映。相反，主观评分吸取了小额信贷的普遍经验、机构的经验和特殊信贷员和信贷经理的经验。如果仅因为创造历史要花费时间，那么，不可避免地，部分经验要来自小额信贷机构本身历史之外。

合理的歧视有意识地选择所要求的特征。在主观评分中所采用的特征（和它们与风险的联系）是明确的；而在统计评分中，它们至少是部分含蓄的。意识到所有评估都存在歧视，这有助于人们理解，评估程序无法让小额信贷试图废止的这种郁闷永久存在。（见方框10）。

合理的歧视仅仅使用与风险真实相关的特征。而且，合理的歧视设法发现新的与风险相关的特征，以便提高预测的准确性，以及更好地将经历转化为评估风险的经验。历史测试是合理歧视的关键，因为它指出推测关联的真实性。与主观评分相比，统计评分更容易测试。

总的来说，评分促进了合理的“歧视”，因为它增加了小额信贷机构对自身经验的了解。这只会减少偏见，更正错误的推理。

第三，信贷员和信贷经理为他们的营生判断风险。这并不奇怪，因为他们不愿意将他们的饭碗托付给一个神奇的魔盒。信任的产生不仅仅需要见到计分卡，人们还需要懂得评分预测的原理，亲眼目睹预测的命中率，而且拥有跟踪预测的手段。

正如所有的项目一样，评分需要“购得”管理层技巧，需要来自内部的拥护。同样，某人必须证明评分系统运作的原理、可以解决的问题。这个道理并不新鲜，把它做出来就是了。

引入评分机制

大多数高层管理者和资助者都听说过评分。但是，一些人认为它能包治百病，另一些人认为它是骗人的把戏，不管怎样，大家对它都报有一种神秘感。正如本文所述，这篇介绍性的文稿，其实际意义是纠正错误的概念、从头设定实际的期望目标。尽管评分的工具是数学计算，但是，评分比黑板上的数学推导意义更大，它带来的是一个组织文化地震般的变化。建立一个评分项目所投入的工作量和时间可能会比大多数管理者想象中的要大、

长久，并且困难重重。当管理者掌握了评分技术，一些人会变得激动起来，而另外一些人就会摆出一副防御性的姿态，因为他们感觉到了生存的威胁，而所有人都会持有怀疑态度。为了营造适应这种变化的氛围，评分项目必须不断地向管理者提出问题，输入新内容，然后照着下列问题匹配答案：

- 您的任务是什么？
- 评分如何促使您完成任务？
- 您发现什么特征最关乎风险？
- 定性特征有多重要？
- 对于您来说什么是“坏贷款”？
- 您想预测何种风险？
- 您愿意牺牲多少好贷款以避免一笔坏贷款？
- 您认为追溯多长的历史才能准确预测未来？
- 放贷政策何时改变的？
- 信贷评估程序何时改变的？
- 目标环境何时变化的？
- 竞争何时开始的？
- 最近的宏观经济繁荣和衰退何时发生的？
- 您不相信数据库中哪部分信息？
- 为适应评分，信息系统（和信息部门的人员）该如何调整？
- 对于实施评分项目，您预见将会出现什么样的障碍？

构建和测试计分卡

下一步是建立计分卡，并且进行历史测试。得出结果后，评分项目经理再次会见高级管理层，以便回顾基本概念，并提交具体的贷款机构所要求的结果，包括历史测试的结果和已探明的风险与特征之间的关系。然后，项目组分小组巡访机构，将评分介绍给所有信贷员和信贷经理。这种介绍很少侧重抽象的概念，而是更多的介绍来自历史测试和所建计分卡的具体实例。

这些会议是很耗费成本的，但是，忽略它们将是一个错误。甚至当信贷员和信贷经理看到评分在历史测试中的效果，他们也许仍然很难接受它。一线工作人员不可避免地要经历排斥和怀疑的过程。最好在实施计分卡之前，给予他们足够的时间使其度过这个过程。

在这里，同样关键的一步是询问问题并且寻求反馈：

- 风险和特征间的关系与您的经验相符吗？
- 您认为用现实生活中的那些原因可以解释这种关系？
- 当您进行实地访问时，您希望寻找什么？
- 在您从实地得到的数据中哪些拿到数据库中就变得不适用了？
- 您会推荐记录什么样的特征为将来计分卡所用？
- 您何时初步批准一个申请？
- 您将如何修改贷款合同的条款和条件以控制风险？
- 您每周用多少时间收款？
- 您每周用多少时间进行市场推广、评估和放贷？

计算分数和形成报告

下一步是要将评分和评分报告的产生过程自动化。总的来说，管理者不愿意改变他们的信息系统，但是，为了在分支机构中使用评分，除了设计自动化过程没有其他选择。有两种主要的办法。第一种，小额信贷机构从咨询公司购买一个做好的评分系统（软件还可能硬件）。这样做，既快捷又简单，但是昂贵。还可以要求输入两次数据，一次是为常规的信息系统，一次是为评分系统。另外，数据录入后，平行的评分系统不会自动执行，必须靠人工操作启动。最后，在平行系统中，预测的风险不容易和贷款机构已经使用的定期报告融为一体。如果用户必须花些功夫、费些力气才能使用评分系统，那么他们很可能放弃它。

实现自动化的第二种方法就是，小额信贷机构将计分卡和相关报告直接集成到现有的信息系统中。这不是一项简单的任务。小额信贷机构（或者它的软件提供商）必须能够修改系统，而且还要指派一名程序员专门负责评分。依靠该系统，集成工作需要三至六个人一个月的工作量，贷款机构的系统管理员不能在晚上和周末做这项工作。集成面临的技术挑战依贷款机构的不同而不同，所以，所有的问题都不可能事先预见。但是，集成具有很重要的优点：数据只需要一次性输入，评分分数可以自动产生，风险预测可以轻易地结合到贷款机构使用的标准报告中。权衡利弊，集成是更可取的办法。

图18：修改贷款条款后风险预测评分模拟报告的样例

客户：某简	分支：中央	申请号：12345		
信贷员：约翰·史密斯	委员会：2001年1月3日	申请日期：2001年1月1日		
	分发额	到期期限	担保（占分发额比例）	预测风险（%）
要求的条款	1,000	10	100	40
分发额	900	10	100	38
	800			33
	700			29
到期期限	1,000	9	100	37
		8		32
		7		27
担保（占分发额比例）	1,000	10	125	39
			150	37
			200	36

来源：作者的样例

一旦计分卡实现自动化后，项目就进入一个非强迫式的“逐渐了解评分”阶段。几个月内，系统会对所有案例进行评分。但是，项目明确要求信贷员和信贷经理不许对预测的风险采取行动，只有在信贷委员会作出决策后才能查看分数。

这使得贷款工作人员要慢慢适应评分，而且鼓励他们在没有要求立即改变的压力下，思考使用风险预测的方法。这一阶段必须要为这种反馈过程创造足够的时间和空间。人们可能会抓住任何明显的弱点或者错误来贬低评分，所以，机构必须倾听意见并着手解决问题。这也许需要再次寻访分支机构，以便检查它们对内容与概念的理解，同时通过新的测试展示评分自执行以来对所放款项的预测能力，并提出更多的问题：

■ 您认为这些预测合理吗？

■ 评分预测准确的频率有多少？

■ 有没有这种情况，被预测为高风险的案例，您根据自己的经验判断实际上它是 低风

险？

■对于高风险的案例您将如何对付？

■您将如何奖励低风险案例？

■何种报告将有助于您对评分的利用？

■您对评分程序的改变有何建议？

■如果当初根据评分的预测值来决定对待“极差”和临界申请人的策略，那么，您的业绩奖金将如何变化？

Filename: OP07Part1_ch.doc
Directory: N:\ForNeeranToPost\Final Chinese
Template: C:\Documents and Settings\wb294243\Application
Data\Microsoft\Templates\Normal.dot
Title: 等级评分：小额信贷领域的又一次突破？
Subject:
Author: xingda
Keywords:
Comments:
Creation Date: 11/25/2005 4:02:00 PM
Change Number: 4
Last Saved On: 12/18/2005 12:35:00 PM
Last Saved By: Staff
Total Editing Time: 2 Minutes
Last Printed On: 12/21/2005 11:19:00 AM
As of Last Complete Printing
Number of Pages: 41
Number of Words: 14,943 (approx.)
Number of Characters: 16,289 (approx.)

两个有用的报告

信贷委员会通常想要求了解看修改边缘临界线的案例是如何的办法会怎样如何影响风险预测的。评积分模拟报告满足回答了这一点。举例来说，图 18 十八表明了，由于贷款合同的条款逐一变化，因此内部成分组成的各不相同，预测的风险有可能出现改变的如何可能变化。这一些风险预测是合同条款修改后，再次在使用计为申请使用了积分卡衡量申请进行申请所得的结果以后，又修改了贷款合同的一个条款后。”评积分模拟报告”积分模拟报告有分两良种形式。第一种是在信息系统内的一种选项，为供信贷委员会不开会时飞行匆忙地测试修改情况使用的一种选择。第二种是一个书面报告，它装包括在每天为信贷委员会准备的包裹中的一个论文报告。

另一个第二个报告，“特征的影响效果”，是针对就风险预测背后的原因做出了，特性的影响，回答了要知道风险预测后面的原因的要求。对于给定所给的申请，这个报告指出了如下两种特征：那些背离哪些与，它表明了从平均历史值导致风险极大提高的价值评估观念存在形成差异的特征以及最大地提高了风险，以及，大多增加了风险的特性，哪些特征最大程度地及大多减少降低了风险的特征的特性。（图 19 十九就是一个例子）。一个论文这个报告的打印稿书面版，同样也将会包括在信贷委员会的每日日常报告中。

设立制定评积分策略

一旦贷款工作人员在花了几个月的时间来使他们自己熟悉了评积分系统之后，小额信贷机构就要制定一个评积分策略、分发成文的书面的评积分策略手册、并且开始在实际操作案例中使用评积分。为什么要求是一个成文的书面的策略呢？因为假如没有书面策略和精确明确的规定，机构很难防止工作人员转而使用传统的信贷评估的现象。一个精确明确的政策策略还可以减少错误和不连贯一致的使用评积分。正如传统的信贷评估，评积分需要一个书面的策略。

图19：特征报告影响效果的范例

客户：某简	案例：A12345	风险：一行内拖欠30天	
信贷员：约翰·史密斯	批准日期：2002.6.2	历史：1995.1.1至2002.5.1	
特征	实际值	历史均值	影响（百分点）
1. 拖欠天数 / 分期付款（在过去已偿贷款中）	8.7	1.7	5.8
2. 迟付分期付款的数目（在过去已偿贷款中）	6	4	4.2
3. 信贷员的经验（以分发贷款的笔数衡量）	77	535	3.4
4. 商业活动的类型	木工	不适用	1.5
5. 家庭电话	无	有	1.1
6. 到期期限（在过去已偿还贷款中；以月计）	8	10.5	0.6
7. 资本循环（%）	缺失	326	0.3
8. 还款负担（%）	20	18	0.1
.....	
36. 担保范围（%）	350	300	-0.4
37. 客户性别	妇女	妇女	-0.7
38. 雇员的数目	0	0.25	-1.9
39. 客户经历（以月计）	36	14	-2.3
40. 客户年龄	55	43	-4.4
风险预测	23.2	9.3	13.9

资料来源：作者举例

一个书面的评分策略应该专门规定风险界定值和相应于每个界定值所采取的行动。举例来说，策略为界定“极好”和“极差”设立了不同的风险水平，在相应风险水平之下的为“极好”，而在相应风险水平之上的为“极差”。它也设立了与“正常”和“临界线”相对应的风险水平。而且，书面评分策略说明了如何奖励“极好”的贷款。对于“临界线”而言，它详细说明了信贷委员会为降低风险所做各项努力的优先级别——是通过降低贷款规模、缩短到期期限、还是增加担保范围。它还指导他们如何使用“评分模拟报告”（见图 18）来观察对贷款合同所做的修改可能产生的效果。最后，书面评分策略强调，“极差”的贷款申请必须被拒绝。

方框9：为什么评分对这样的借款人不公平？

像好的天气预报一样，好的评分预测也是取平均值，而不是指向哪一天，或者哪一笔个体贷款。事实上，风险预测从来做不到每个案例都百发百中；预测的风险永远大于零而小于100%，但是，现实中实际风险永远是零（没有变坏）或者100%（确实变坏了）。对于某个具体的贷款而言，评价评分系统的正确与错误，毫无意义。

评分做出的预测是概率值，而不是确定值。准确程度的测量是通过比较一组平均预测风险和平均不良贷款率（实现了的风险）而得出的。如果评分系统运行正常，那么，一些具有高预测风险的贷款会维持好的状态，而一些具有低预测风险的贷款会变坏。举例来说，如果评分系统有效，那么，风险为50%的借款人中有一半会维持好的状态，而对于风险为5%的借款人而言，20个人中就有1个会变坏。

当然，评分政策（不同于评分预测）对于具体的个案来说，最终结果有可能正确，也有可能是错误。例如天气预报说下雨的机率为60%，你选择了带雨伞，这既有可能正确（如果下雨了）也有可能错误（如果没有下雨）。同理，在评分的帮助下做出的批准或者拒绝贷款的决策，在事后看来可能正确也可能错误（尽管我们永远不会看到拒绝贷款的决策带来的正确结果）。

专断策略

评分最有价值的一面是，这种方法可以识别出信贷委员会认为是安全的高风险案例。但是，信贷员和信贷经理都具有人类的弱点，当评分结果与他们的判断相抵触时，他们可能会藐视评分，并且寻找任何微小的瑕疵来贬低评分（例如，一笔低风险的贷款变坏了，或者一个高风险的贷款保持良好——见方框 9）。怀着相同的心态，他们可能会要求知道为什么风险如此之高。

反对评分策略的选择是专断。（对于小额信贷而言，专断体现在经核准的“极差”贷款和未受审核的“临界线”贷款上。）专断策略通过三种方式发生作用。第一，它通过全面跟踪报告不断地测试预测能力。第二，它指出风险是如何通过“特征影响报告”与特征相关联的。第三，它不仅仅力劝用户重视积分，而且还详细指出后果。

举例来说，小额信贷机构可以在业绩奖金上惩罚过度的专断¹。如果专断超过“极坏”贷款的某个百分比，奖金就要被扣减。时间长了，直接惩罚就没有多大必要了，因为信贷员认识到，滥用专断会导致欠款增多和业绩奖降低。

谨慎的专断的确需要。信贷委员会也许知道，个别的案例是例外，只有凭人的判断才可以评估被计分卡忽略的定性特征。关键要适度。正如不是所有人都能在平均水平以上一样，专断不可不可能评估所有高风险的贷款。在高收入国家里，贷款机构试图将专断范围限制在“极坏”的百分之十以内。对于小额信贷而言，一个合适的范围也许该是百分之二十五。

小额信贷机构必须跟踪专断行为以便向信贷员提供反馈意见。总的来说，最终专断判断的风险比预测的风险少（原因是信贷委员会知道一些计分卡未显示的事情，而且信贷员为使他们的预言成真工作异常努力），但是，比其它贷款的风险大（因为，计分卡知道一些信贷委员会不知道的事情）。

寄生策略

应对专断策略是防止对评分的过少依靠，而应对寄生策略是设法防止对评分的过度依赖。特别是，书面策略必须强调（正如本文所强调的）评分只处理经过传统评估程序初步批准的贷款申请。所以，有必要不断地提醒那些还有疑虑的人以帮助他们找到平衡点。在人们看到评分可行后，他们可能会忽略传统的评估。如果贷款尚未通过机构传统主观评估的初步批准，就被实施评分，那么风险将会被低估——低估的程度也许很大。在这里，再次重申本文的核心观点，小额信贷机构不能用评分代替它的主观评估。它应该在完成主观评估后实施评分，否则拖欠贷款会急速增加。全面跟踪报告此报告跟踪评分的进展情况。如历史测试一样，它比较预测风险和实际风险，但不同的地方是，它还应用于未偿贷款。全面跟踪报告是评分报告的核心，甚至比历史测试更有用。它检查评分系统对运转中贷款的有效性。像其它评分报告一样，它通过系统自动产生。在评分阶段的头几个月里，贷款机构每周查阅它以便核实评分的预测能力，并且指导策略的调整。在这之后，机构每月检查一次。

第一份全面追踪报告包括分发前没有评过分的未偿贷款，并且和历史测试一样，它显示了假定的预测能力。几个月以后，报告展示它对于决策完全依赖于评分的贷款的预测能力。图 20 是一个根据一家拉丁美洲小额信贷机构回归计分卡（在第六章讨论）形成的全面跟踪报告。“不良贷款”被定义为每笔到期分期付款平均拖欠四天，或者拖欠期限为三十天。

图 20 最左边一栏（“预测风险%”）规定了每一行的预测风险范围。贷款机构限定了范围数量和临界线。从左边数第二栏（“未偿贷款数量”）是未收回贷款所占的比例，它的预测风险落入某一行的范围内。它表明在未偿贷款总量中预测风险的分布情况。举例来说，截止到 2001 年 7 月 31 日，有 0.5% 的未偿贷款的预测风险落入 0 到 2% 的范围。同时，有 9.5% 的贷款的预测风险超过 40%（将以下这些栏相加），而有 19.5% 的贷款的预测风险在 30% 以上。（未偿贷款栏内的数字相加为 100）。

¹ Martin Holtmann, “设计金融激励制度以增加信贷员的产出率：谨慎从事”，《小额信贷公告》第六号（2001 年）：第 5-10 页。

四个中央栏（“自放款之日起以天数计的实际风险比例%”）表明了给定预测风险和账龄的未偿贷款的实际风险。逐行比较实际风险和预测风险可以看出计分卡的能力。预测的风险越接近实际风险，预测能力就越强。（这些栏内的数字相加，不到100）。

举例来说，预测风险为8%到10%，并且已放出0至90天（见图20）的贷款，其实际风险是5.3%。也就是说，在符合这两个标准的1394笔未偿贷款中，截止到报告日，有74笔（占5.3%）变成不良贷款。另外一个例子，预期风险超过70%，并且已放出271天以上的贷款，其实际风险为77.9%。

图20说明了一个基本观点：实际风险随着放款时间的增加而增加。这里有两个原因。第一，一些近期的贷款还没有面临到期的分期付款，所以，他们还没有机会变“坏”。第二，贷款将近结束时，拖欠贷款会增加。²因此，检验预测能力的最好方法是着眼于最近已偿还的贷款和（或）已经放出很长时间且尚未偿还的贷款。

全面跟踪报告最右边一栏表明了最近已还贷款的实际风险。（贷款机构决定报告将追溯多长的历史：例子中是12个月。这一栏很关键，不仅因为它包括了所有到期期限的贷款，而且因为最近偿还的贷款已经有足够的时间可以变坏。

核查预测能力

全面跟踪报告检查计分卡是否有效。绝对的准确性意味着实际风险贴近预测的风险。在图20中，预测风险为0到2%的近期已偿还贷款，其实际风险为3.2%（右边第一行）。这是在预测范围以外，但是它接近预测值。对于预测风险为2-4%、4-6%和6-8%而言，实际风险落在其内，然而对于其他范围的预测风险，其实际风险超过了相应范围的最高值。绝对准确是好的，但是它不完美。因为，对于预测风险高的贷款而言，预测的风险比实际的风险略低。

相对准确性意味着预测风险较低的贷款和预测风险较高的贷款相比，前者的实际风险比后者的低。图20中的计分卡具有很好的相对准确性。除了最低的两个范围，实际风险随着图中的数字自上而下不断增加。

末端准确性意味着，在风险分布的末端（尾部），绝对和相对准确性都很好。末端准确性很重要，因为评分策略并不影响近乎平均风险水平（正常）的贷款。评分仅仅影响风险极低（极好）和风险极高的（临界和极坏）贷款。

图20中的计分卡具有很好的末端准确性。举例来说，预测风险为0到2%的近期已还贷款的实际风险是3.2%。预测风险在2-4%、4-6%和6-8%范围内的实际风险恰好落在该范围内。在高端，预测风险超过70%的近期已还贷款中，有75.4%变坏（右下角）。在预测风险超过40%的已还贷款中，超过一半的贷款变坏了。

图20：全面跟踪报告范例

风险：同一行内每笔分期支付风险数量：贷款笔数
拖欠4天或欠款30天

分支机构：所有

² Ulrike Vogelsang, "危机时期的小额信贷"

检测日期：2002.6.2

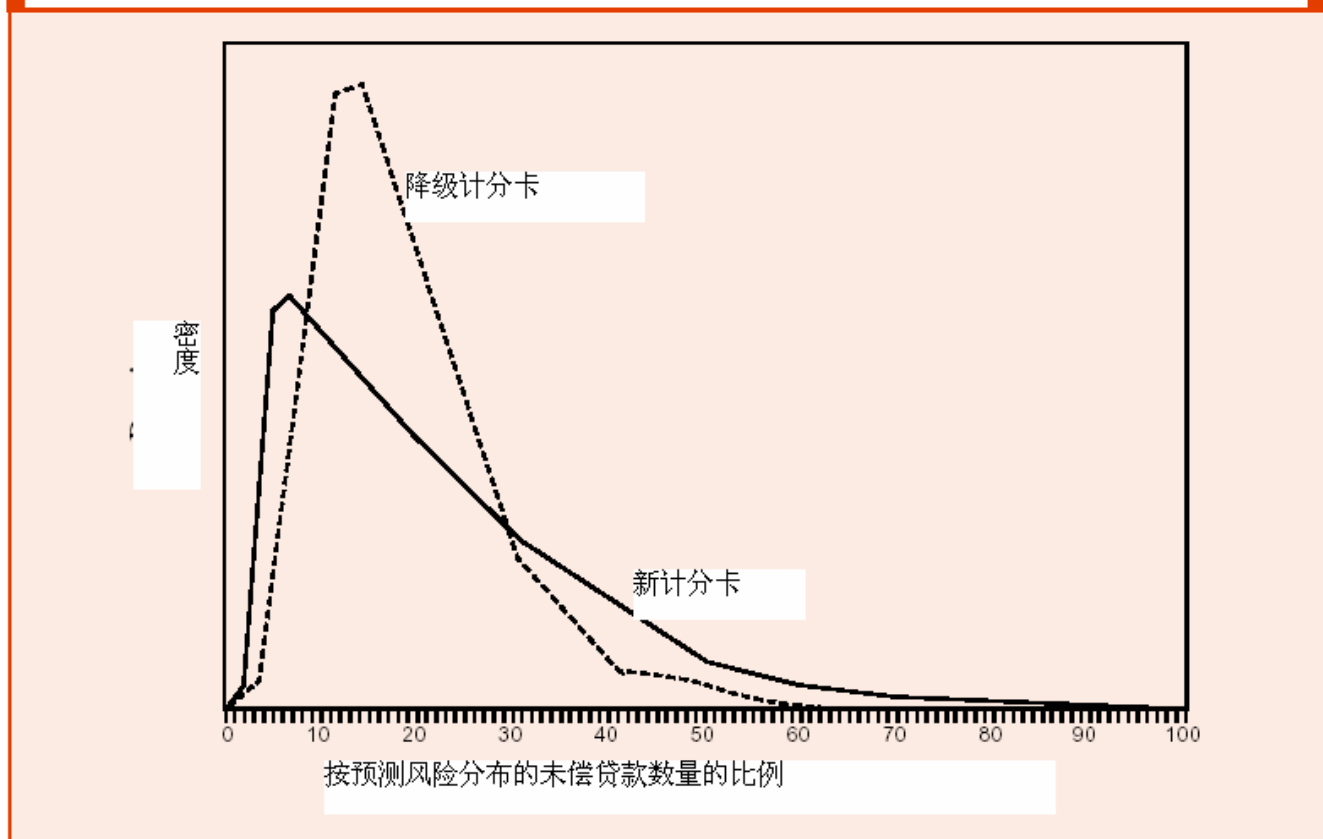
计分卡构建日期：2001.7.31

预测风险未偿贷款数 自贷款分发以来以天数计的实际风险(%) 在最近12个月已偿贷款的实际风险(%)

	91-18 181-27					
	0-90	0	0	271+		
0~2	0.5	1.4	2	0	4	3.2
2~4	5.1	2.8	2.8	2.1	3.5	3.1
4~6	7.8	3	4	4	5.1	4.7
6~8	8.1	3.9	4.8	5.5	8.1	7.8
8~10	7.7	5.3	6.7	6.4	11.5	10.6
10~15	17	5.5	8.1	11.6	18.1	16.3
15~20	14.5	6.8	12.1	17.9	27.6	24.7
20~25	11.4	9	16.9	23.8	33.1	27.2
25~30	8.4	11.4	19.4	30.4	37.8	36.3
30~40	10	14.6	25	37.3	45.8	43.1
40~50	5.1	18.4	30.4	50.9	53.6	52.6
50~60	2.7	23	42.3	57.2	60.4	60.1
60~70	1.2	32.4	42.6	65.2	70.5	70.3
70~100	0.5	34.3	62.9	65.5	77.9	75.4

资料：对一家拉丁美洲小额信贷机构业务实施评分

图2 1：在新的和降级的计分卡中预测风险分布变化的例证



固定绝对误差

具有绝对精度的计分卡更易于使用。相对精确仅能将贷款按预测风险排序。举例来说，具有10%预测风险的贷款比具有20%预测风险的贷款的实际风险小。但是，实际上，这两组的实际风险有可能是7%和25%。如果使用绝对精确，那么，具有10%预测风险的贷款不仅会有10%的实际风险，而且其实际风险还恰好是具有20%预测风险的一半。

遗憾的是，计分卡不存在完美的绝对准确性。但是，全面跟踪报告可以显示相应于给定预测风险水平的实际风险水平。有了这个信息，用户可以调整预测风险的级别，使调整后的预测更加准确。

假定全面跟踪报告显示，预测风险水平总是高出五个百分点。举例来说，贷款机构就把具有25%预测风险的贷款仅仅当作20%的预测风险来处理。在现实生活中，误差的模式更复杂，但是这种转换原则仍然有效，并且信息系统可以自动实现这种转换。

跟踪专断

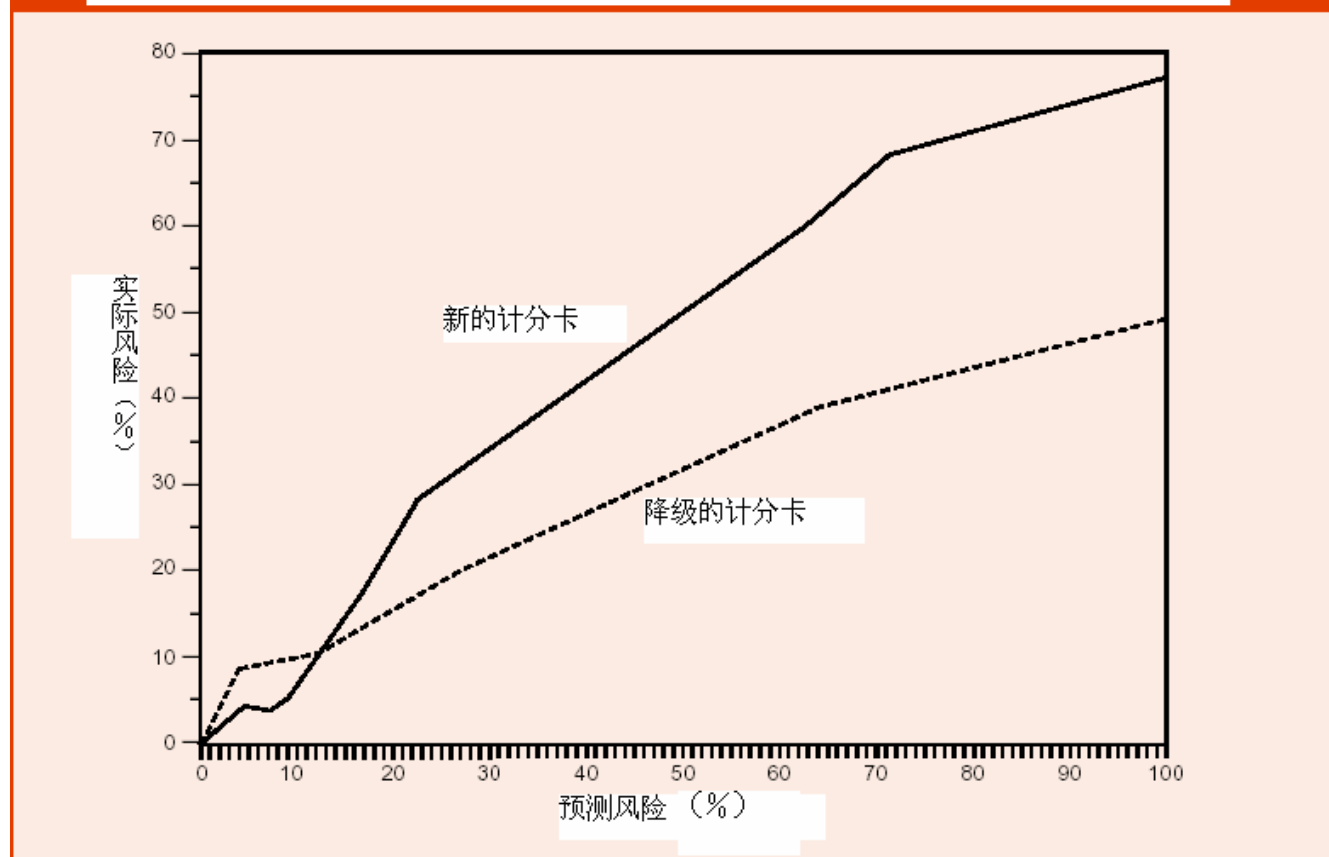
根据定义，预测风险超过“极差”界定值的已放出贷款即为“专断”贷款。专断可能被滥用，所以，管理者必须跟踪它们的结果。采取的方式是，在已放出的极差贷款中检验实际风险随时间而变的情况。对比的基点是评分开始前的实际风险。正如评分系统下发放的贷款一样，如果“极差”一类中的实际风险比预测风险低得多，那么，一般来说，对预测风险被高估的贷款的专断行为已经成功地被制止住了。如果实际风险降低很大，以至于小额信贷机构打算批准已知具有这种风险水平的贷款，那末，目前对专断的这种限制应该保持。否则，应该加强这个限制，直至专断中的实际风险降低到可以接受的水平。举例来说，如

果“极差”的界定值是 70%，并且假设在开始评分后的第一天，全面跟踪报告显示在本应定为“极差”类别的过去贷款中，实际风险为 78%。评分一年后，假如全面跟踪报告显示，在专断的贷款（所放出的贷款的预测风险超过 70%）中实际风险变成 35%。这意味着，信贷委员会总体上限制了对风险高估的案例的专断。这个 35%的比例可能比贷款机构愿意承受的风险还要大，如果是这样，贷款机构将加紧对专断的限制。如果贷款机构愿意冒此风险，那么，目前应对专断的策略应当保持。

设定或者调整策略界定值

全面跟踪报告显示了每一个风险范围内贷款所占的比例，以及相应于给定预测风险的实际风险水平。因此，小额信贷机构可以使用全面跟踪报告去确定或者调整策略的界定值。

图22：在新的和降级的计分卡中预测风险和实际风险之间关系变化的样例



对于图 20 中的计分卡，2%的“极好”界定值将会影响 5%尚未收回的贷款（左边第二栏，第一行），而一个 4%的“极好”界定值将可能影响 5.6%的未偿贷款。70%的“极差”界定值会拒绝 0.5%现尚未收回的贷款。进一步讲，对于这样一种“极坏”贷款策略而言，每放弃一笔好贷款就能够避免三笔坏贷款的发生（因为在这个范围内，实际风险大约是 75%—见右下角）。如果“极差”界定值减少至 30%，那么，19.5%的贷款会被拒绝，大约有一半是不良贷款。

检测计分卡降级

因为将来更类似于近期，而不是很远的过去，所以，计分卡的预测能力随时间而退化。

全面跟踪报告通过两种方式表明了这一特征。第一，预测风险分布更加尖耸（很少展开）。降级的形式使典型的预测更加靠近平均水平。图 21 是一个假设的例子，在这个例子中，新计分卡预测风险的分布图是根据图 20 中全面跟踪报告的前两栏绘制而成的。

表明预测能力退化的第二个指标是预测风险和实际风险之间较为平缓的坡度。随着降级，实际风险超过低预测风险水平的预测风险。而且，降级意味着在高预测风险水平那里，实际风险比预测风险要小。图 22 是一个假设的例子，在这个例子中，新计分卡的预测风险和实际风险之间的关系是根据图 20 中全面跟踪报告的倒数第二栏绘制的。

如果给定的计分卡是依据最新的报告，那么，为了观察降级的程度，管理者就要对比全面跟踪报告中预测风险的分布情况（和/或者预测风险与实际风险之间的关系）。象图 21 和 22 这样的图形就形象地突出了报告中数据的变化。

下降的速度取决于放款策略变化的频率、目标环境、竞争程度、资金总量增长、宏观经济以及其他随时间而变化且影响风险的各种不确定因素。在评分发生严重退化情况之前（也许在两至三年后），小额信贷机构应该更新计分卡。更新要比初创评分项目简单而且迅速。将自首张计分卡以来积累的数据包括进去，建立一个新的计分卡，然后将它嵌入现存系统中。

信贷员跟踪报告

全面跟踪报告是评分工作的核心。但是，对于信贷员和信贷经理而言，它也许太抽象（因为它比较不同组别贷款的预测风险与实际风险），太宽泛了（因为它包含了所有未偿贷款和所有近期已偿还的贷款）。从技术上讲，全面跟踪报告是对评分预测能力的最好测试。但是，一线工作人员似乎更喜欢简单的报告，因为这种报告便于他们对比预测的风险和自己熟悉的借款人的还贷表现。

像这样一种信贷员跟踪报告，它把对预测风险和还贷表现（实际风险）的测量，加入到信贷员和信贷经理每日或者每周收到的总业务报告中。图 23 和 24 是两份有关一家拉丁美洲小额信贷机构递减计分卡（见第六节）的简单报告。这家小额信贷机构将“不良”定义为在整个贷款期内至少发生一段 30 天的拖欠行为。这种信贷员跟踪报告与历史测试不同，它涉及未偿贷款；它也与全面跟踪报告不同，因为它列出了具体借款人的名字。

图 23: 信贷员跟踪报告显示的风险最高的 30 笔贷款 (发放的贷款已超过 270 天)

报告日期: 2001.7.31

分支机构: 全部

风险: 30 天以上

贷款: 已发放超过 270 天

列表: 30 笔最高风险贷款

实际风险

贷款号	客户名	发 放 天 数	发 放 额 (美 元)	月 支 付 额	下 笔 到 期 日	目 前 欠 款	欠 款 次 数	每 期 拖 欠 天 数	最 长 拖 欠 期	贷 款	是 否 不 良	预 测 风 险 %
79922	Javela, María	308	2,106	83	8 月 3 日	23	2	42.5	77	是		90
50973	Posada, María	334	1,860	71	8 月 29 日	0	3	21.1	36	是		81
71596	Arboleda, Nivelly	336	1,323	132	8 月 29 日	2	3	14.8	25	否		80
80816	Beltrán, Dioselina	304	1,032	48	8 月 29 日	0	3	14.8	42	是		80
62037	Nuñez, Dolly	337	5,683	316	8 月 2 日	0	1	22.7	28	否		72
45638	Cruz, Leonor	304	377	22	8 月 29 日	0	3	45.5	101	是		71
64823	Rivera, Antonia	304	603	39	8 月 29 日	23	2	22.2	39	是		68
61653	Marín, Graciela	337	5,763	283	8 月 2 日	0	4	14.5	25	否		62
78800	Muñoz, Marco	304	2,003	111	8 月 29 日	0	3	25.7	67	是		60
24893	Silva, Oswaldo	304	388	29	8 月 29 日	86	2	36	86	是		59
65323	Ruíz, Asia	308	56	12	8 月 3 日	58	4	24.7	58	是		59
59506	Cardona, Graciela	334	188	51	8 月 29 日	0	2	11.9	18	否		59
54093	Tejada, María	285	14,638	790	8 月 11 日	0	1	0.3	2	否		58
71243	Castillo, Rosa	293	630	70	8 月 18 日	0	2	6.1	15	否		58
22692	Tavárez, María	348	143	39	8 月 13 日	0	1	0.4	2	否		58
99155	Marroquín, Libia	334	77	41	8 月 29 日	0	1	11.1	22	否		58
18634	Rivera, Melida	334	470	50	8 月 29 日	191	2	82.7	191	是		57
74810	Marulanda, Pablo	304	331	27	8 月 29 日	23	3	25.8	54	是		56

20410	Valencia, Claudia	356	323	53	8月21日	0	4	5.5	14	否	55
60737	Suárez, Yolanda	335	275	40	8月3日	0	1	0.5	2	否	55
85854	Marín, Jorge	308	1,275	106	8月3日	0	4	7.7	20	否	55
42074	Lozano, Nevalia	292	251	19	8月18日	86	2	52	93	是	54
30986	Berrios, Fanny	318	2,449	136	8月13日	0	2	4.4	15	否	54
31208	Gomez, Diafanor	306	6,049	291	8月1日	0	3	4.5	12	否	54
89020	Calderón, Editha	319	259	38	8月14日	0	1	7	14	否	54
8408	Marulanda, María	306	332	42	8月1日	0	2	61.6	131	是	53
36244	Castillo, Brunilda	279	383	46	8月5日	0	1	0.9	3	否	52
5699	Ortiz, Nubia	334	570	46	8月29日	0	2	15.5	39	是	52
7719	Montoya, Javier	281	100	17	8月7日	36	3	12.6	36	是	52
40373	Moreno, Peregrino	304	381	50	8月29日	177	4	68.9	177	是	51
来源：拉丁美洲小额信贷机构递减计分卡和数据库						平均风险：			50		

61

图 24：信贷员跟踪报告显示的风险最低的 30 笔贷款（发放的贷款已超过 270 天）

报告日期：2001.1.12

分支机构：全部

风险：30 天以上

贷款：已发放超过 270 天

列出：30 笔最低风险贷款

实际风险

贷款号	客户名	发放天数	发放额（美元）	月支付额	下笔到期日	目前欠款	欠款次数	每期拖欠天数	最长拖欠期	是否不良贷款	预测风险%
62225	Valencia, Lucero	292	59	60	8月18日	0	0	0	0	否	0.5
38388	Betancourt, Jose	305	73	26	8月1日	0	1	0.1	1	否	0.5
88687	Valencia, Juan	279	35	36	8月5日	0	0	0	0	否	0.5

94799	Fernandez,Zorrilla	281	289	38	8月7日	0	0	0	0	否	0.5
8154	Sanchez,Hernan	290	102	36	8月16日	0	0	0	0	否	0.5
38563	Escobar,Patricia	316	117	32	8月11日	0	1	7	13	否	0.5
27819	Echandia,Henry	322	102	36	8月17日	0	0	0	0	否	0.6
21502	Jaramillo,Ema	285	289	103	8月11日	0	1	0.1	1	否	0.6
71907	Guervara,Cesar	295	87	31	8月20日	0	0	0	0	否	0.6
49562	Paz,Maria	336	768	167	8月1日	0	1	0.8	5	否	0.6
93142	Escobar,Monica	284	35	36	8月10日	0	0	0	0	否	0.6
11221	Palomino,Fe	287	73	26	8月13日	0	0	0	0	否	0.7
88301	Garcia,Alberto	308	289	38	8月3日	0	0	0	0	否	0.7
77258	Arce,Eduardo	305	116	41	8月2日	0	1	1	5	否	0.7
1582	Contreras,Elena	318	147	77	8月13日	0	1	0.1	1	否	0.7
79476	Sanchez,Gonzalo	323	293	65	8月18日	0	1	1.4	5	否	0.7
985	Lopez,Flor	295	35	36	8月20日	0	0	0	0	否	0.7
85657	Torres,Maria	280	347	46	8月6日	0	0	0	0	否	0.7
16697	Chacon,Emilsa	293	73	26	8月18日	0	1	4	20	否	0.7
53165	Gutierrez,Lucila	356	153	55	8月21日	0	0	0	0	否	0.7
80399	Lopez,Alejandro	291	460	86	8月17日	0	1	0.1	1	否	0.7
32949	Castano,Alvaro	323	68	36	8月18日	0	0	0	0	否	0.7
94131	Duque,Lucia	287	219	78	8月13日	0	0	0	0	否	0.7
28050	Polanco,Gerardo	294	76	79	8月19日	0	1	0.1	1	否	0.7
30709	Fajardo,Carmen	349	101	103	8月14日	0	0	0	0	否	0.7
54730	Aristiza,Morena	287	73	26	8月13日	0	0	0	0	否	0.7
18377	Ceballos,Luis	314	168	45	8月9日	0	0	0	0	否	0.7
28881	Escobar,Jose	323	78	41	8月18日	0	0	0	0	否	0.8
34129	Munoz,Edison	283	461	86	8月9日	0	0	0	0	否	0.8
74078	Tabarez,Jesus	341	50	51	8月6日	0	1	0.2	1	否	0.8

来源：拉丁美洲小额信贷机构递减计分卡和数据库

平均风险：

0

0.6

对于“极差”的贷款而言，图 23 列出了 30 笔最高风险的未偿贷款，截至报告日它们至少已贷出 270 天。在这组未收回贷款中，平均预测风险是 61%（右下角），平均实际风险为 50%。即使是 15 笔“好的”贷款，结果也不是那么理想：这 15 笔贷款都有过拖欠记录，其中差不多有四笔贷款的拖欠时间超过十天。当信贷员看到自己的借款人出现在这份名单上时，当他们回忆起向这些借款人收款遇到的麻烦时，他们可能会意识到评分的价值。

在“极好”这一方面，图 24 指出了 30 笔最低风险的贷款。平均预测风险在 1%以下（见右下角），并且没有一笔贷款真的变坏了。事实上，在 30 笔贷款中有 19 笔根本没有拖欠行为，在其余 11 笔有拖欠行为的贷款中，有 6 笔拖欠期仅为一天，只有 2 笔超过十天以上。³

就信贷员和分支机构经理而言，从他们看见自己的借款人列在报告中（如图 23 和 24 所示），到打消他们的疑虑，使他们认识到评分能够从信贷委员会已批准的那些贷款中识别出高风险和低风险，确实经历了很长一段时间。实施评分的小额信贷机构应该在发给信贷员和信贷经理的标准日报和周报告中，加上信贷员跟踪报告。

如果雇员给评分一个机会，他们会看见它的成效，但是，他们必须理解它，相信它能够成功。于是，这成为培训和测试的任务。一旦接受评分，那么，恰当使用它的方法便取决于书面政策，严格的对专断控制，以及持续不断的监控。将未偿贷款预测风险与实际风险进行比较的跟踪报告，既为全面贷款业务，也专为每一个信贷员，提供了必要的、不变的稳固基础。

六、递减计分卡和专家系统

此章介绍递减计分卡，一种比树型更加复杂、更加强大的计分卡。同时，还介绍专家系统（计分卡的第三种类型），然后将递减计分卡、树型结构和专家系统加以对比。

递减计分卡

递减计分卡是一个数学公式，通过将借款人、贷款和贷款机构的特征加权并相加来计算预测值（概率）。公式所选择的特征和使用的权重来自复杂的统计学知识，在这里不再详述。但是，递减预测有点像树型预测方法，而且所有的运算都是由信息系统完成。与树型和专家系统相比，递减预测效果最好，并且最清楚地表明了风险和特征之间的联系。

假设统计数据发现，随着借款人年龄的增长，风险每年按 0.1 个百分点的速度减少。而且进一步假设，“基础风险”是 10 个百分点。那么，预测一笔贷款变坏的概率的递减公式就是：

$$\text{风险} = 10 - 0.1 \times \text{年龄}$$

利用这个公式为一个三十岁的借款人预测风险，结果将是 $10 - 0.1 \times 30 = 7$ 个百分点。对于一个五十五岁的人，预测风险便是 $10 - 0.1 \times 55 = 4.5$ 个百分点。（这些权重只是举例。真正的权重取决于具体的贷款机构。）

³是评分正好这么幸运吗？如果贷款机构的历史坏帐率为 9.6%，那么，随便挑选的 30 笔贷款并且不变坏的机率（如在低风险信贷员跟踪报告中）则小于 5%。30 笔中有 15 笔变坏的比率，如在高风险信贷员报告中，则小于十亿分之一。

第二个例子，假设统计数据发现，风险以每月 0.25 个百分点的速度随到期期限而增加。如果基础风险是十个百分点，那么递减预测风险公式是：

$$\text{风险} = 10 + 0.25 * \text{到期期限}$$

因此，对于一笔三个月期限的贷款而言，它的预测风险是： $10 + 0.25 \times 3 = 10.75$ 个百分点。对于一笔十二个月期限的贷款而言，其预测风险是： $10 + 0.25 \times 12 = 13$ 个百分点。

图25：在递减计分卡中风险与自第一次放款以来月数之间的关系

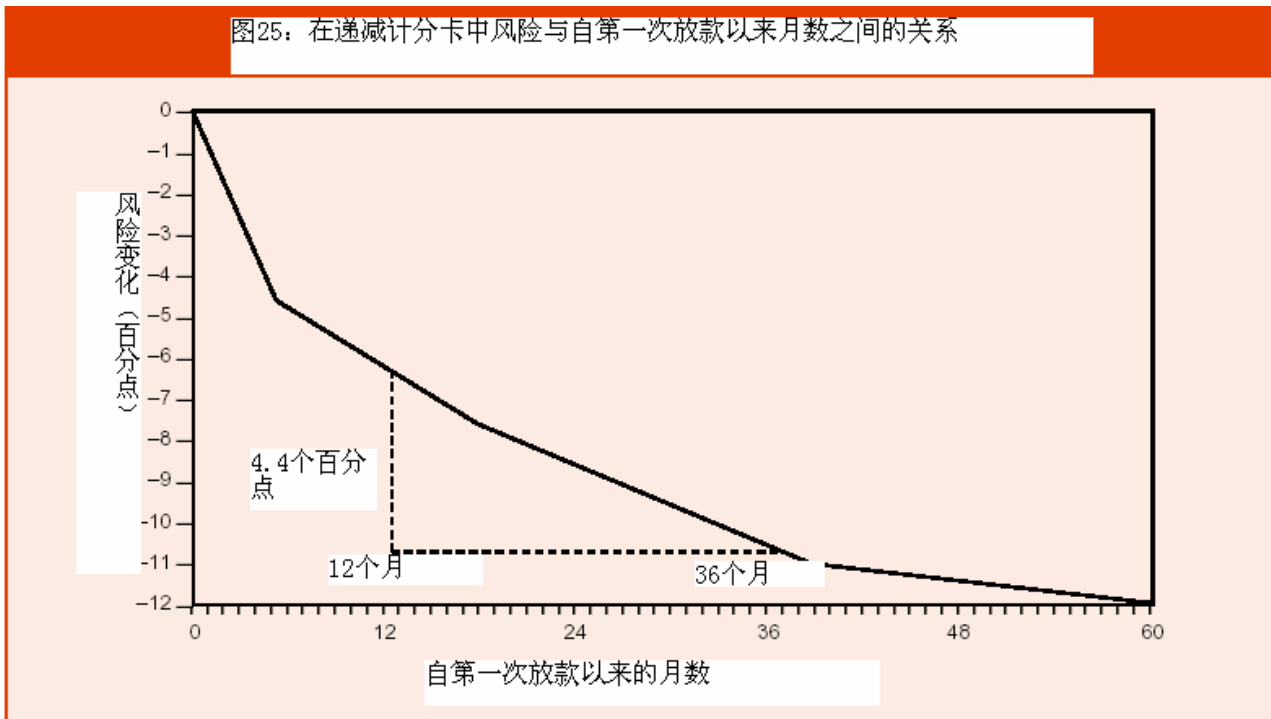
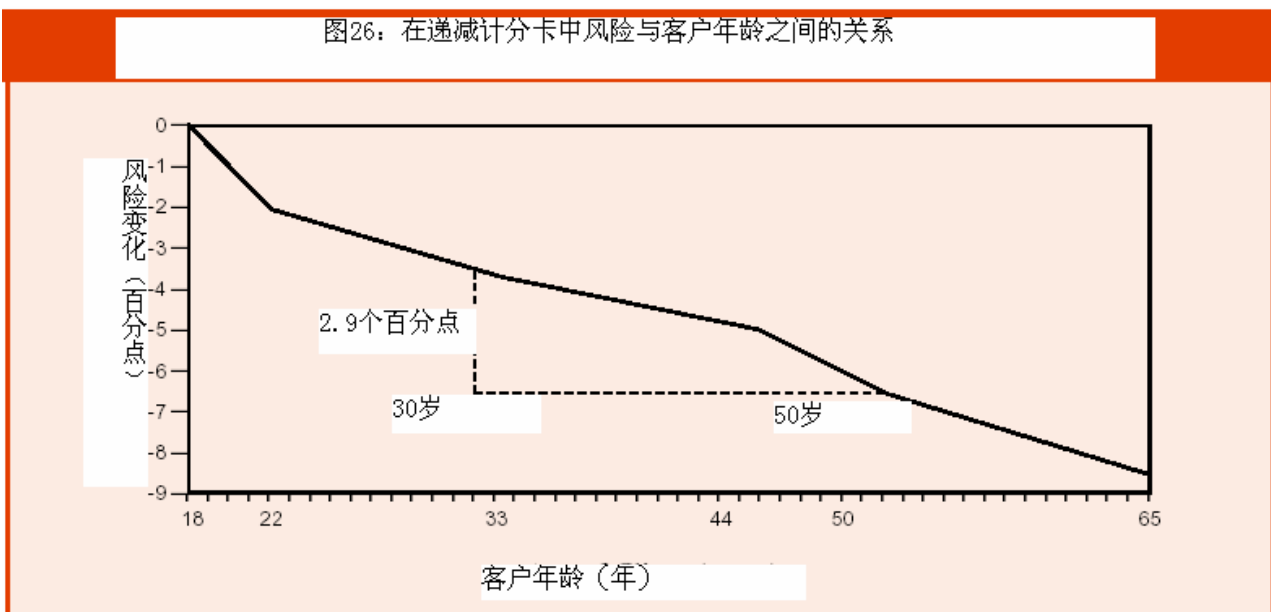


图26：在递减计分卡中风险与客户年龄之间的关系



实践中，递减计分卡包括的特征范围十分广泛。举例来说，一个结合上述两个单一特征公式的计分卡，可以更细致地区别高风险与低风险：

$$\text{风险} = 10 - 0.1 \times \text{年龄} + 0.25 \times \text{到期期限。}$$

举例来说，一个三十岁的借款人，有一笔三十六个月期限的贷款，那么，预测风险为 $10 - 0.1 \times 30 + 0.25 \times 36 = 16$ 个百分点。相比之下，一个五十五岁、有一笔三个月期限贷款的借款人，她的预测风险是 $10 - 0.1 \times 55 + 0.25 \times 3 = 5.25$ 个百分点。

事实上，递减计分卡可能包括三十至五十个特征，并且从具体的小额信贷机构数据库中，它能推导出所有的权重。在信息系计算算出所有预测值后，贷款机构就通过前面章节描述的程序来使用它。

递减计分卡中的风险与特征之间的关系

虽然递减方式的预测能力在所有计分卡种类中是最好的，但是，它的最大优势，也许还在于它清楚地表明了风险与特征间的关系。对某一特征配置权重，这不仅表明了计分卡中其他特征保持不变的情况下该特征是增加、还是减小了风险，而且还指出了变化的程度。这些关系只有在申请通过了传统评估的初步批准以后，才成立。此处所举出的例子，是来自一家拉丁美洲小额信贷机构真实的递减计分卡。

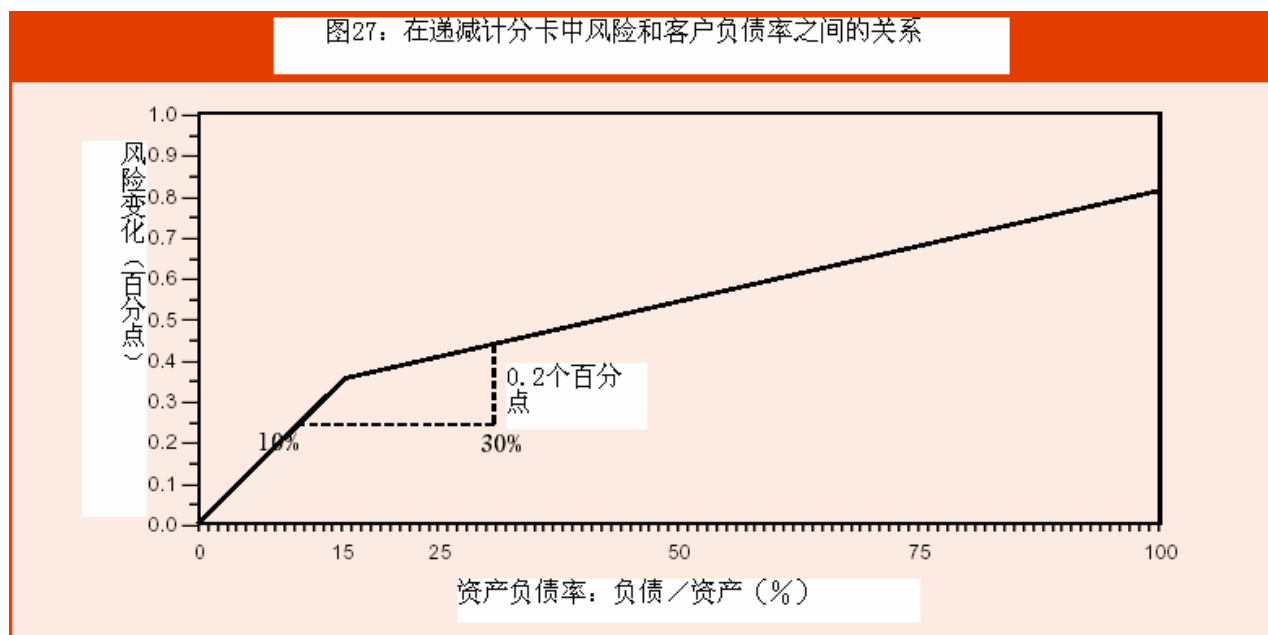


图28：在递减计分卡中，过去三笔贷款的风险和拖欠之间的关系

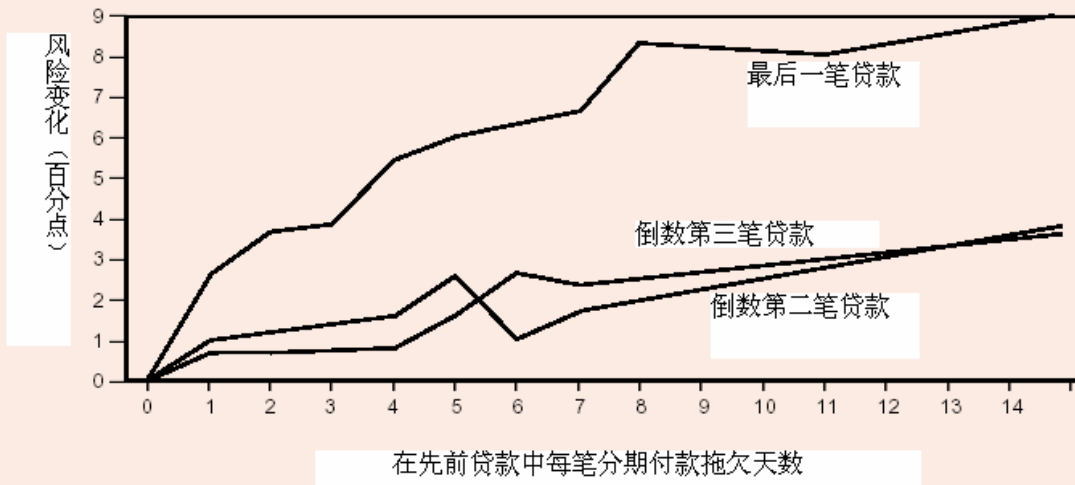


图29：在递减计分卡中风险与商业类型之间的关系

商业类型	对风险影响程度 (%)	贷款额所占比例 (%)
货车运输和驾驶出租车	-3.6	0.4
销售水果和蔬菜	-3.5	2.3
住宅区杂货店	-2.6	4.3
小型家庭用品店	-2.1	6.4
流动售货	-2	4.4
美容店	-2	2.7
面包店	-1.9	2.3
化妆品销售	-1.9	1.6
杂货店	-1.7	2.3
女裁缝和服装制作	-1.3	11.1
精制食品销售	-1	1
学校	-1	0.6
食品加工	-1	0.6
汽车备件商店	-0.6	0.7
街头快餐食品	-0.6	0.5
肉食品市场	-0.5	1.4
家庭用具销售	-0.5	1
服装店	-0.2	1.6
其它或者未不知的	0	39.5
鞋店	0.1	2.5
药店	0.3	1.9
餐馆	0.7	1.7
五金商店	0.8	1.1
百货商店	0.9	4.1
专业服务	1	0.6
艺术品	1.2	0.8
锁匠和金属加工	1.6	0.7
汽车机修	1.7	0.5
制鞋	2.1	1
木工	2.6	0.5

来源：拉丁美洲小额信贷机构

风险与特征之间的关系

在图 25 中，递减计分卡表明随着放款以来月份的增加，风险大大减少。例如，一个借款人在放款 36 个月后，在其它条件不变的情况下，风险要比一个借款人发放 12 个月后的风险少 4.4 个百分点。另外，风险还随着年龄的增加而大幅减少。举例来说，在图 26 中，一个五十岁的人（其余条件不变）相对于一个三十岁的人，风险大约要少 2.9 个百分点。

如图 27 所示，风险随着家庭或企业中负债率（负债除以资产）的增加而增加。一个负债率为 10% 的人（其它条件不变）比负债率为 30% 的人，风险要少 0.2 个百分点。在过去三笔贷款（见图 28）中，风险还随贷款每笔分期支付的平均欠款天数而增加。举例来

说，最后一笔贷款十天的欠款期使当前风险增加了 8%；倒数第二笔贷款七天的欠款期使当前风险增加了 2%。倒数第三笔贷款的拖欠对目前风险的影响，与倒数第二笔贷款的影响非常相似。

因此，与具有良好记录的借款人相比，在过去的三笔贷款中，平均拖欠 10 天、7 天和 7 天的借款人，使其在目前贷款中的风险大约增加 $8 + 2 + 2 = 12$ 个百分点（假设它已经按照传统的评估程序被初步批准了）。

图 28 对于既定借款人将来拖欠和过去拖欠两者之间的关系提出了四点主要结论。第一，过去已实现的风险越大，将来预测的风险越高。第二，过去很久的欠款比近期欠款的预示作用要弱。第三，相对于良好的记录而言，过去哪怕只有很短的欠款记录，它也预示着未来更大的风险。举例来说，在先前的贷款中，平均拖欠一天就会使目前风险增加两个百分点。假定这个小额信贷机构总的坏贷款率在 15% 以下，那么，这两个百分点的变化也是很大的。第四，风险随着过去的欠款记录而增加，但增加比率递减（这种关系只是对初步批准的贷款申请成立）。

其它关联

经营的种类与风险紧密相关。对于图 29 中的小额信贷机构而言（“对风险的影响”一栏，按降序排列），最低的风险是：

- 出租汽车和卡车司机
- 存货周转快的店铺（水果和蔬菜、杂货、小型家用物品）
- 街头食品小贩（快餐食品, 面包）
- 美容店和化妆品柜台
- 女裁缝们

对于这个贷款机构，具有最高风险的经营类型是：

- 制造商（木工、鞋匠、汽车机修和锁匠）
- 专业人员和艺术家们
- 存货周转慢的商店（五金、医药、鞋、衣服、家用器具和汽车零件）
- 餐馆

图30：在递减计分卡中风险和信贷员个人之间的关系

信贷员	对风险的影响 (%)
Carmen Ochoa	-10.1
Catalina Gonzalez	-9
David Soto de los Santos	-5.7
Rosario Sosa Almanecer	-3.9
Mariangeli Cintron Ruiz	-2
Rosa Justiniano Ornes	-0.2
其他人	0
Ma. Eugenia Mariscal	1.1
Marcos Orta	2.3
Eldo Parra Barriga	3
Oscar Navajas	3.3
Teresa Guzman	4.9
Enrique Flores Santos	7
Maria Padilla Ruiz	13.6

图 29 显示了每种经营类型的历史贷款业务量在总量中所占的比例。这个贷款机构明显关注低风险的商业活动。

信贷员

递减计分卡还可以揭示风险和一个具体的信贷员之间的关系。在图 30 中，这种关系在不同信贷员之间呈现出广泛的多样性。在这个例子中，顶层和底层的信贷员大约相差二十四个百分点。负责贷款的信贷员对风险的影响很大，但是，只有递减计分卡（而不是树型或者专家系统）利用这种特征去提高预测风险的准确性。它可以帮助信贷机构确定培训、激励和奖金分发的目标。

对于图 30 有必要说明一点。信贷员管理风险的方法是，在放款前筛选申请人，在放款后监控贷款。递减计分卡揭示了监控的效果，但没有显示筛选的效果。这是因为在其他特征保持不变的情况下用递减方法衡量信贷员对风险的影响，就如同所有信贷员管理具有相同量化特征的贷款业务。

事实上，信贷员管理着不同的贷款业务，其组成（定量的和定性的）取决于信贷员筛选申请人的效果。有一些信贷员通过筛选那些本不需要过多监管的申请者来达到既定的贷款风险水平；另外一些信贷员使用较少的筛选程序但是利用较多的监控来达到同等贷款风险的水平。此外，一些信贷员被安排到棘手的地区，在那里，一定水平的技术和努力发挥出来的效果不如在其他地方大。因此，贷款机构不应该立即解雇那些在递减计分卡中排名低的信贷员，而是应该调查排名低的原因并且设法去排除它们。

专家系统

从经理的经验和判断中取得，而不是来自数据的统计分析的计分卡，被称为专家系统。专家系统与传统的评分不同，评分使用模糊的判断，而专家系统使用明确的规则

或者数学公式。专家系统的强大之处在于它们不需要一个数据库，并且，由于它们是小额信贷机构经理和信贷员共同构建的成果，因此，在本组织中销售它们并不困难。专家系统的弱点在于，它们比树型或者递减计分卡的预测能力差。另外，因为专家系统假定了风险与特征之间的关系，因此它们不能解释这种关系。如今大多数声称要使用评分的小额信贷机构都知道专家系统意味着什么。

专家系统树好象统计树，唯一不同的是，前者的分支不是来自一个顾问所做的数据库统计分析，而是来自小额信贷机构经理和信贷员的经验、判断以及推测。这样，一棵树的叶子表明的是定性的等级，而不是量化的概率。举例来说，图 2 中的统计树，预测再次贷给一个妇女的风险是 12.8%，但是，在图 31 中专家系统树将这种同样贷给妇女的再次贷款列为“非常安全”。如今，小额信贷中使用最普遍的专家系统树是根据拖欠等级分数建立的（见方框二）。

专家系统递归是一组数学公式（就象统计递归一样），但是，经理们自主选择哪些特征和它们的权重，而不是从数据分析中得出。专家系统递归也产生一个数字，这个数字标明一个级别，而不是一个概率。所以，分数可能会大于 100 或者成为负数。因此，尽管专家系统的递归也许能达到某一水平的相对准确性，但是，它缺乏绝对的准确性。

所有专家系统的改进，无论是树型结构还是递归方式，都可以通过测试预测能力将等级转化为概率来实现。历史测试以及跟踪报告，适用于统计计分卡，同样也适用于专家系统。但是，不同于预测风险与实际风险之间概率的比较，专家系统测试对比的是预测出的等级与实际风险。贷款机构可以使用测试将非概率化的等级转化为数字型的概率，然后在工作中仅仅使用概率。

图 32：五种类别风险的政策范例

待预测的风险类型	所采取的政策
1. 放款之前：如果放出，此贷款在它的放款期内是否会出现某个水平的欠款？	极坏：拒绝 临界线：修改条款 正常：不加修改地放款 极好：提供奖励和增加额度
2. 放款之后：这位借款人的下一次分期付款是否会迟付？	假定他有罪：进行“礼节性拜访”，打电话，或者写信 假定他清白：拭目以待
3. 收款：这笔目前拖欠 X 天的贷款，会出现 X+Y 天的拖欠吗？	高风险与高风险价值：立即访问，放弃缓和的策略

	<p>高风险或者高风险价值：立即访问但是使用缓和的策略</p> <p>低风险和低风险价值：稍后再访问，并且缓和地催债</p>
4. 抛弃：一旦目前这笔贷款还清，此借款人还会申请另外一笔贷款吗？	<p>剔除：因为糟糕的还贷表现，不能续贷</p> <p>不安全的犹豫中的人：拭目以待，没有激励</p> <p>安全的犹豫中的人：提供激励以重来</p> <p>忠诚的人：拭目以待，没有激励</p>
5. 访问：在信贷员进行实地访问后，贷款机构会拒绝申请吗？	<p>没有希望的：不进行实地访问就拒绝</p> <p>有希望的：继续进行访问</p>

更为重要的是，历史测试和跟踪报告指出了预测能力的程度。如果经理选择的分支和权重偏偏不太理想，专家系统仍然可以发挥预测作用。⁴另外，专家系统可能会以他们对数据的低需求和便于使用的优势来弥补它们预测能力低的缺陷。

小额信贷机构应该非常乐于试用简单的自制的计分卡。⁵但是，他们应该在使用之前和过程中测试这些计分卡。令人难以置信的是，大多数使用专家系统的小额信贷机构并没有测试它们。它们的错误不在于偏爱专家系统而冷落统计计分卡，而在于他们忽视了对预测能力的测试。那些使用评分的贷款机构应该遵循这样一个原则：眼见为实，而不能搞盲目崇拜。

递归法的预测能力最强，而且它比树型和专家系统，更能显示出风险与特征间的关系。但是，递归复杂，并且对数据库的要求最高。只有最大规模和发展最成熟的小额信贷机构才具备使用递减计分卡的条件。

树型结构（甚至是自己动手建造的树）的预测能力也奇好，并且它要求的数据比递归少。和专家系统一样，树型结构易于解释、易于向员工推广，但是，它并不是总能明确地揭示风险与特征间的联系。

专家系统很容易构建，因为它不需要数据。这一点使它们成为与当今小额信贷机构关系最密切的一类计分卡，但是，它们的不足在于预测能力不如树型和递归。缺少统计计分卡所需数据的小额信贷机构可以先从专家系统开始，但是，它们还应该着手收集必要的的数据以便支持预测效果更好的计分卡。

七、准备评分：预测何种风险？

第一个评分项目应该简单地构造一张单一的计分卡。贷款机构必须在放款前评分、放款后评分、收款评分、客源流失评分和访问评分中（见图 32）做出选择。大多数机构会选择

⁴ A. D. Lovie and P. Lovie, "平面的最大效果与线性评分预测模型"《预测期刊》第五期：第 159-68 页； Peter Kolesar and Janet L. Showers, "一个强大的使用绝对数据的信贷筛选模型"《管理科学》第三十一期，第二号(1985 年)：第 123-33 页； William G. Stillwell, F. Hutton Barron, 及 Ward Edwards, "评估信贷申请：对多属性效用加权计算技术的确认"《组织行为及人工操作规程》第 32 期(1983 年)：第 87-108 页；和 Howard Wainer, "线性模型中的评估系数：它值得关注"《心理学公告》第 83 号(1976 年)：第 213-17 页。

⁵ Mark Schreiner, "自己动手构建小额信贷评分树"(2001 年 11 月 11 日至 12 日，在多米尼加共和国首都圣多明各举行的 Tercer 研讨会上所做的有关拉丁美洲及加勒比地区小额信贷的演讲)

放款前评分（这是本文到目前为止一直讨论的类型），这既是因为这种四级策略简单而实用，又是因为对放款前的风险预测可以为放款后和收款的评分打下基础。

放款前评分

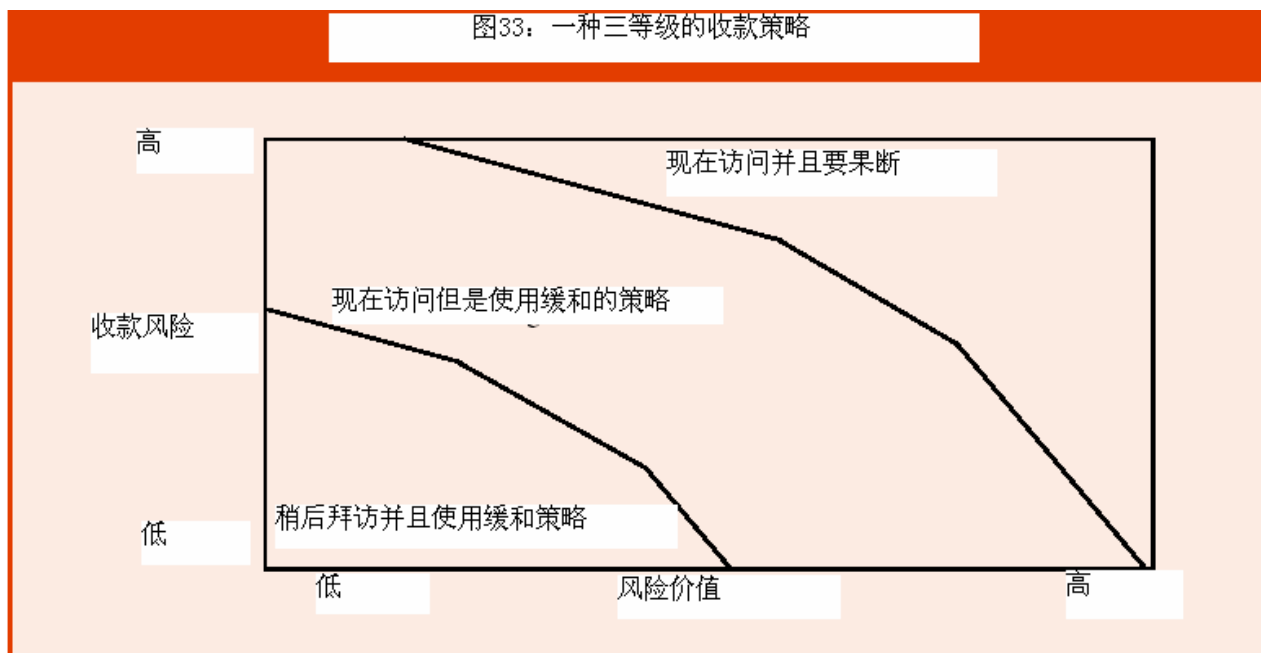
放款前评分，预测的是如果一笔初步批准的贷款被放出去，它在贷款期内变坏的概率。贷款机构必须决定如何定义“坏的”概念，通常包括下列几个方面：

- 一次欠款期超过 X 天
- 欠款超过 Y 次，不考虑时间长度
- 每笔分期付款的平均拖欠时间超过 Z 天

给评分系统中“坏的”概念下定义也许会成为一种有益的实践。它迫使小额信贷机构仔细思考有关拖欠与成本的问题。例如，欠款的次数和它的长度，哪一个更重要？多次的短期欠款是否可以忍受？贷款机构还应反思，目前应该使用什么样的标准去决定是否向在先前的贷款中有拖欠记录的客户再次续贷。

放款前的评分不应该把“坏的”概念定义为“违约”。在一定的技术水平上，大多数小额信贷机构遇到的违约情况极少，以至于无法揭示风险与特征之间的关系。更重要的是，大多数小额信贷机构早在贷款陷入“违约”境地之前，就已经把它视为一笔坏贷款了。信贷员不会反问自己：“如果我批准了这笔贷款，最终我能收回它吗？”，而是这样问自己，“如果我批准了这笔贷款，我是否需要费很大力气才能把它收回来？”有鉴于此，大多数小额信贷机构都设有自己的对策，对于那些在某一时点上有多次欠款记录的借款人，即使他在此前的贷款中没有违约，机构也会拒绝向他们再次贷款。

图33：一种三等级的收款策略



放贷后评分

放款后评分，预测的是一笔未收回贷款的下一笔分期付款可能被拖后的概率。放款后

的风险与放款前的风险高度相关。两种计分卡都从同样一组特征中预测，不同的是，放款后的计分卡还包括当前贷款的还款记录、当前贷款已偿还的分期付款次数、以及其余尚未收回的贷款余额。放款前的评分分数可以有效地替代放款后的分数，因为放款前具有很高风险的贷款在放款后同样具有很高的风险。只有在放款后的风险对于贷款机构而言已经非常明显的情况下（例如，在放款后，未偿贷款出现严重拖欠情况），放款前的分数才失去替代的作用。

如果忽略用于预测放款后风险的计分卡，那么，机构就可以选择一种简单的两级策略（见图 32）。具有最高风险的贷款（或者，也许具有最高的风险价值）被推定为“有罪”，这一等级的贷款可能占有所有贷款的 5%。甚至在麻烦开始之前，机构就要对这些借款人进行一次预防性的礼节拜访、电话通话或者信件联络。如果其余贷款被推定为“清白”，那么，直到借款人真的陷入拖欠危机时，小额信贷机构才采取特殊的措施。

信贷员跟踪报告（见图 23 及 24）帮助信贷员决定拜访对象。举例来说，从图 23 得出的候选人，将包括三种还没有变坏的高风险、高价值贷款：

- 预测风险为 54%的分发额度为 6,049 美元的贷款
- 预测风险为 58%的分发额度为 14,638 美元的贷款
- 预测风险为 72%的分发额度为 5,683 美元的贷款

图34：一种四等级客源流失评分策略

		传统信贷评估标准		
		不合格	合格	
			贷前风险高	贷前风险低
流失风险	高	开除：不提供激励	不安全的犹豫中的人：不提供激励	安全的犹豫中的人：提供激励
	低		忠诚者：不提供激励	

在礼节性拜访中，信贷员仅仅是拜访借款人，并不涉及任何目前的收款事宜，讨论的问题没有任何威胁性。信贷员决不可向客户透露他们已被评为高风险的情况，以免它真的由预言变成现实。有一定身份的借款人，如果他们发觉自己成为被怀疑的对象，他们很可能会表现出敌意。仅仅是信贷员的访问性出现就足以增强借款人按时还贷的观念。信贷员可以利用拜访的机会进行信息反馈，如询问客户这笔贷款是如何运转的，了解他们对于贷款机构提供的服务有什么意见，以及打听一下他们对贷款合同是否存在疑虑等等。

在贷款机构刚开始使用评分时，礼节性拜访显得尤其有价值。在这一点上，许多“极坏”贷款已经记录在案，即使贷款机构无法收回这些贷款，它仍然可以为风险管理做点什么。

收款评分

收款评分，预测的是一笔目前已迟付 X 天的贷款将出现 $X + Y$ 天拖欠的概率。最普遍的情况是，它预测一笔昨天开始欠付、今天迟付一天的贷款最后成为一笔拖欠 30 天的贷款的风险。实际操作中，收款评分的分数将被列入关于拖欠贷款的日常报告中。然后，信贷员根据收款的风险和风险价值，通过一种三等级策略来决定首次拜访的对象，以及如何温和地催促他们还债（见图 32 和 33）。对存在高风险和高风险价值的案例要立即访问，并且要直截了当地表明态度。对于高风险或者低风险价值（但两项不是同时具有）的案例，也要进行立即拜访，但是态度的表达要较为缓和。最后，对于低风险和低风险价值的案例，首先要保持沉默，几天后以温和的方式进行第一次接触。低风险的客户，对信贷员因他们欠付而进行的到访也许会感到懊恼。他们会很好地纠正自己，如果情况没有什么改变，一个善意的警示就足以使他们步入正轨。

与放款后计分卡一样，收款计分卡使用的特征几乎与放款前计分卡使用的相同，所以，放款前评分的结果可以代替收款评分。如此说来，放款前计分卡的分数是一次结果三处使用。

流失客源评分

流失客源评分，预测的是借款人在还清目前贷款以后再申请另一笔贷款的概率。⁶ 小额信贷机构希望避免客户流失，因为利润通常会随着一笔重复贷款的循环而增加。⁷ 如果贷款机构知道哪一位客户存在退出的风险，它可以提高吸引力以实现重复放贷，例如，降低利率或者免除 51 种贷款手续费，这当然是临时性的，而且是在借款人保持令人满意的还款表现的前提下。这种激励对于贷款机构来说成本很大，而流失客源评分可以帮助机构将这种激励锁定在最有可能中途退出的客户身上，这样一来，成本就控制住了。

在一笔未偿贷款最后一次分期付款前的一个月里，贷款机构计算客户流失评分的分数和放款前评分的分数，并且假设下次贷款的合同将采用与目前合同相近的条款。随后，机构开始实施一种四级策略（见图 32 和 34）：

开除。 开除是一种强制性退出。根据传统的评估标准，他们在目前贷款中的拖欠行为已使他们丧失取得再次贷款的资格。

不安全的犹豫中的人。 即使他们目前贷款的状态没有变坏，“不安全的犹豫中的人”还是面临退出和变坏两方面的危险。他们可以申请续贷，但是，贷款机构不会向他们提供特殊的激励。

安全的犹豫中的人。 “安全的犹豫中的人”存在退出的风险，但是他们没有还贷问题。这些好的客户有可能流失，因此，贷款机构应向他们提供激励以便他们续贷。

忠诚者。 忠诚者既不存在退出风险，也没有变坏的风险。贷款机构不必向他们提供特殊的激励，因为他们无论怎样都会再次贷款。

⁶ Mark Schreiner, "对玻利维亚一个小额信贷机构客户退出风险的评分"。

⁷ Craig F. Churchill and Sahra S. Halpern, "构筑用户的忠诚度：小额信贷制度的实践指导"《小额信贷网络技术指导》，第 2 号（2001 年）；Richard Rosenberg, "对客户保留度的测量"，《小额信贷公告》，第 6 期（2001 年）：第 25-26 页。

拜访评分

在实地拜访之前，拜访评分预测在实地拜访之后拒绝放贷的概率。这种被拒的案例会花费信贷员许多时间，却毫无收益。拜访评分是根据书面申请上的特征预测拒绝放贷的风险，从而削减毫无结果的拜访数量。图 34 中的两级策略就在没有访问的情况下拒绝了没有希望的客户（也许占拜访评分中最糟糕的 5%），但是对于有希望的客户还是按照传统的评估程序进行了实地拜访。

拜访评分仅用于未经访问就拒绝的情况，它不适用于未经拜访就接受的情况。如在方框 4 中讨论的那样，代理风险低并不意味着风险的定性水平低，但是，表现值高的风险也许会使定性风险悬而未决。

不仅可以预测拜访后拒绝放贷的风险，拜访计分卡还可以预测放款后贷款变坏的概率。这是一种不包含拜访中收集到的特征的放款前的评分。假定预见到的还贷问题会导致访问后的拒绝决定，那么，即使无从知晓被拒申请人的还贷表现，量化特征（它与初获批准的申请人还贷的高风险相关）仍然还有可能与所有申请人被拒绝的高风险相关。因此，针对偿还表现的拜访评分可以代替针对拒绝决策的拜访评分，反之亦然。

仅用一个实际的测试即可揭示拜访评分对还贷的预测能力。相比之下，针对拒绝决策的拜访评分要通过历史数据先行测试。但是，与建立一个用于还贷的拜访计分卡不同，用于拒绝决策的拜访计分卡需要来自被拒申请人的特征，然而，只有少数小额信贷机构将这些数据输进它们的信息系统。

大多数小额信贷机构将从放款前的评分开始，也许他们还会把它当作放款后评分和收款评分的替代品。一旦他们掌握了放款前评分的使用方法，他们就可能加入客源流失评分这一程序，而对于那些拥有充分数据的贷款机构，加入拜访评分这一程序也是指日可待。

八、准备评分：收集何种数据？

评分只能预测发生多次的事情，而且前提是这一事件已被记录在数据库中。最有效的风险评估来自于对平常数据的收集。

大多数小额信贷机构没有足够合格的数据去建立计分卡，所以，一旦他们确定了为某个风险评分，那么，紧接着的一步工作就是积累更多更好的数据。⁸这个任务有三个方面。第一，简单地积累更多的坏案例。这个工作需要时间，而且对于贷款业务量小的机构来说，还需要等待机构的发展壮大。第二，收集关于借款人、贷款和贷款机构的附加特征；第三，改进数据的质量。

所要求的坏贷款的数量

没有任何方法能够准确地计算出建立一个计分卡所需要的坏贷款的数量。统计理论仅对最简单的统计方法（例如平均值）提供准确的抽样数量。即使如此，所需样本规模还要

⁸此节的部分取自Mark Schreiner所著的"在Prizma评分：如何做好准备"（提交给Prizma 和扶贫协商小组，Mostar、波斯尼亚和黑塞哥维那的报告，sean@prizma.ba）。

取决于取得样例之后才知晓的参数。这里，不存在用于计算递归或者树型结构所需抽样数量的公式。⁹

在高收入国家中，广为认可的意见是计分卡至少需要五百笔坏贷款的案例。¹⁰但是，它同时假定客户不仅有获取薪金的工作，而且在信用局中留有信贷记录。¹¹在这种特定情况下，一个具有十至十五个特征（它们大多来自信贷报告）的计分卡足以构建一张功能强大的计分卡。但是，在目前小额信贷中，大多数借款人为自营型就业者，而且，即使信用局存在，大多数借款人还没有被记录在内。

因此，小额信贷计分卡中典型特征表现出的预测能力不如它在一个高收入国家中那样有效。¹²小额信贷要获得足够的预测能力就需要更多的特征，并且，它要得出风险与更多特征间的关系就需要更大规模的坏贷款的数量。

建立一个有效的小额信贷计分卡也许至少需要一千笔坏贷款的案例。这个猜测很可能过低了，而不是过高了。越多越好，但是，满足评分总体需要（以及小额信贷领域的评分）的准确数目还不知道。这种数目的权衡还取决于贷款机构和具体的环境。这种不确定性便是创新的代价。

小额信贷机构能够象在美国的小企业贷款机构那样共享他们的数据吗？¹³小额信贷的不幸之处是，机构各异，一种规模不能适用所有的人。只有小额信贷机构在同一国家运作、具有同样的目标市场、并且使用相同的传统评估程序，这样才可以说，有一个共享的数据计分卡总比什么都没有强。跨国界借鉴计分卡是一种愚蠢之举。

收集适合的特征

一个小额信贷机构现在应该开始收集什么样的特征才能够在若干年后构建一个积分卡（或者一个功能更强大的计分卡）？在下面的列表中，标有星号的是一些核心的必需特征。大多数经营个体贷款的小额信贷机构已经收集了大部分这样的核心数据作为传统评估的一部分。能够增加预测能力的附加特征也在下面列出，尽管没有它们，同样可以构建出有效的计分卡来（例如，我们在图 20、23 和 24 中看到的结果）。其中大部分特征可以让申请人在初次申请时提供出来。

借款人的特征

人口统计数据。申请人的人口特征是最有预测性的一类：

- 性别*
- 出生年份*
- 婚姻状态*（已婚/同居，一直未婚/从未同居，离婚/分居，丧偶）结婚/同居年份
离婚/分居/丧偶年份
- 最后的学业等级*

结婚或者分居日期是考察家庭稳定性最便利的表征量。一些小额信贷机构可以选择忽略与

⁹ William G. Cochran, 《抽样技术》第三版(纽约: John Wiley 和 Sons, 1977 年)。

¹⁰ Edward M. Lewis, "等级评分的引入"

¹¹ 信用局是与众多贷款机构共同收集、存储、发布有关单个借款人还款信息的数据库。

¹² Mark Schreiner, "小额信贷的等级评分"。

¹³ Latimer Asch, "利用评分提高社会中介法贷款的效率"(手稿, San Rafael, 加利福尼亚: Fair, 艾萨克, 2000 年)。

人口统计特征相关的风险，而这些特征是申请人无法为自身选择的（见方框 10）。

联络信息。在数据库中记录下电话号码和联络信息，对预测风险有作用：

- 可以找到家中的电话号码*（可以是邻居的电话）
- 可以找到单位的电话号码*（可以是邻居的电话）
- 从家（或从单位）到最近的分支机构的距离、花费的时间以及所支付的公共交通费用（如果需要的话）

到最近的分支银行的距离（以及有电话可用）可以间接衡量交易成本。¹⁴更大的交易成本增加了借款人拖欠的风险，而且也增加了信贷员监控的难度。

家庭人口统计数据。家庭的组成影响现金流和风险：

- 十八岁（含）以上的人口数（包括申请人）
- 十七岁（含）以下的人的数。

家庭财产。家庭财产的存在（以及随时间推移的变化）可预示风险：

- 家庭房屋的使用和占有（所有者，租赁人，其它）

搬至目前居所的年份

搬至以前居所的年份

目前居所的房间数目（不包括洗澡间及厨房）

- 土地所有权

冠名的田产土地（存在还是不存在）

未冠名的田产土地（存在还是不存在）

其它冠名的土地（公顷数）

其它未冠名的土地（公顷数）

- 住房建筑

马口铁房顶（存在还是不存在）

水泥地面（存在还是不存在）

供水管线（存在还是不存在）

排污管线（存在还是不存在）

供电线（存在还是不存在）

- 使用的车辆

汽车，拖拉机，卡车，或者公共汽车（存在还是不存在）

摩托车（存在还是不存在）

自行车（存在还是不存在）

- 电器

冰箱（存在还是不存在）

天然气炉或者电炉（存在还是不存在）

工作用彩色电视（存在还是不存在）

发电机（存在还是不存在）

- 正规的储蓄帐户（存在还是不存在）

相关家庭财产视当地的具体情况而定。

假如在没有贷款的情况下，财产不会发生变化，那么，这些数据就表示“产生了影响”。另外，其中一些财产的测量也出现在评估贫困的指标中。所以，贷款机构收集这些数据，除了在评分系统中使用外，可能还有其它方面的原因。

¹⁴ Mariano Rojas 和 Luis Alejandro Rojas, "墨西哥优惠贷款的交易成本", 《发展政策评论》第 15 期(1997 年): 第 23-46 页; 和 Carlos E. Cuevas, "发展中国家金融中介的交易成本", 《经济与社会不定期论文》, 第 1469 号(俄亥俄州, 哥伦布: 俄亥俄州州立大学, 1998 年), ccuevas@worldbank.org

评分可能会显示出更穷困的客户（举例来说，那些拥有更少财产的借款人）具有更大的风险。以扶贫为目标的小额信贷机构，可能将一些与贫困相关的特征排除在计分卡外，或者为了更穷的客户而接受更大的风险。评分无法改变借款人的风险，它仅仅提高对已经存在的风险的认识。

商业经营的统计。最基本的商业活动特征有助于预测还贷风险：

- 行业*（制造、服务、贸易、农业）
- 经营的具体类型*
- 在本行业起家时的年份
- 自己开办企业的年份*
- 正式注册（存在还是不存在）
- 现金流的书面记录（存在还是不存在）
- 场所类型（商店的沿街铺面，流动性，锁盒，以家庭为基础，其它）
- 场所的使用和占有（所有，租用，以及其它）
- 搬至目前场所的年份
- 以全职工作量计算的人员工作月
- ? 每年所雇佣的工人
- ? 付工钱的家庭成员
- ? 不付工钱的家庭成员
- ? 付工钱的非家庭成员

许多小额信贷机构已经记录了“雇佣人员的数目”。但是，对于评分来说，这经常是无用的数据。因为它把季节性雇佣与永久性雇佣，兼职雇员与全职雇员，家庭成员与非家庭成员，以及付工钱和不付工钱的雇员都混在了一起。雇员的工作量应该针对不同的工种按每年员工的工作月计量。

家庭/企业的财务流。

每月的现金流是预测信贷风险强有力的指标：

- 经营收入*
- 家庭收入中来自工资的部分*
- 家庭收入中来自其它渠道的部分*
- 购买货物发生的经营消费*
- 企业工资支出*
- 其它的经营花费*
- 租金的支出
- 其它的家庭支出*
- 每月其它债务到期的分期付款（包括家庭抵押）*

由于现金流的波动不定，小额信贷机构还应该询问销售份额中现金与信贷之间的比例。在进行实地拜访时，信贷员必须收集财务数据。大多数小额信贷机构目前记录销售状况、其它收入、经营支出、以及家庭消费。对于评分来说，不集中的现金流是有好处的，因为风险有一部分取决于现金流是否有规律地流动，或者取决于现金流的强制性与自愿性。

企业的股票。大多数小额信贷机构已经记录了资产与负债的价值：

- 总资产*
- ? 固定资产*

- ? 存货*
- ? 现金及银行帐户*
- 总负债*
- ? 非正式债务*
- ? 正式债务*

还贷记录。预测未来还贷情况的最好指标就是观察过去的还贷表现。对于每一笔贷款到期的分期付款而言，贷款机构应该记录到期日，以及还贷日期。由这可以得出测量拖欠行为的以下几个方面：

- 最长的拖欠期*
- 每笔分期付款拖欠的天数*
- 迟付的分期付款的次数*

在还清每一笔贷款后，贷款机构还应该要求信贷员主观地对总体还款表现划定等级，范围从1（最好）到5（最坏）。

信用局。信用局数据的预测能力十分强大。¹⁵如果贷款机构收到信用局关于一些借款人的报告，那么，他们应该将如下的数据输入他们的信息系统：

- 目前以及过去债权人的身份
- 目前和过去贷款的放贷日期(和清偿日期)
- 目前和过去贷款的放贷额度
- 目前和过去贷款的每月分期付款
- 目前和过去债权人的最大信贷限度
- 目前和过去贷款的拖欠情况
- 欠目前债权人的额度
- 质询的次数

方框10：评分应该使用受保护的特征吗？

任何人都无法选择他们的性别、种族、母语或者年龄，而且许多人，特别是妇女和少数民族，在选择婚姻以及居住地上都受到了限制。然而，所有这些特征都是一目了然的，因此，它们可能（或者已经或者正在）被用来压制一个集团而满足另一个集团的利益。传统的贷款机构，曾经不合理地将带有这些标记（“受保护的特征”）的人拒之门外，一方面是因为贷款机构加入到压制这些人的行列中，另一方面是因为他们的压制行为加剧了这些申请人的风险。小额信贷一个核心的目的就是要帮助转变这种状态。

在一些高收入的国家，法律禁止在计分卡中使用受保护的特征。法律的目的旨在排除来自非统计评分中的明确的压制，并且防止统计评分产生这样一种观念：社会上其他地方的压制行为使得风险与受保护的特征挂起钩来。但是，在大多数低收入的国家，像这样的法律根本不存在。受保护的特征对还款风险有预示作用，小额信贷

¹⁵ Michael E. Staten, "综合信贷报告的价值：来自美国的经验教训" (在世界银行年会上提交的关于"小额商业贷款收益"的论文，华盛顿特区，2001年，4月2日至3日)；和 Elinor Haidor, "信用局：对微型企业收益产生杠杆作用的信息"，《小型企业发展评论2》，第2号（2000年）：第1、5-8页。

计分卡应该使用它们吗？

这个问题不是很好回答。一种做法就是收集更多和更好的数据。毕竟，基因不会直接引起风险。受保护的特征与风险间接相关，因为它们与社会所产生的特征相联系，而后者反过来又直接影响风险。举例来说，Y染色体缺乏并不影响一个妇女的还款风险，但是，社会对妇女成为女裁缝而不是铁匠的认可却会影响还款风险。随着有更多更好关于特征的数据直接与风险相关，作为间接变量的受保护特征的重要性将会减少。

当然，这还不能解决问题。即使妇女有更大（或者更小）的风险（不是因为她们是妇女，而是因为社会限制妇女），她们也无法选择自己的特征。在某种程度上，那些不受保护的特征也许是非自愿的。举例来说，穷人没有选择去受穷。即使是明显选择的特征也是一些我们没有察觉到的选择和强制之间冲突的产物。一些人仍然相信，这里根本不存在选择，只有听命于自然法则无情的摆布。

归根结底，风险依然存在，而且大部分风险还与未选中的特征相关。小额信贷机构必须决定如何评估风险，前提假定是任何一种评估都必然建立在经验和偏好上。在追求更好的预测效果和加剧不公平的歧视之间总会存在一种平衡。最终，小额信贷机构必须对收集何种数据以及如何使用它们作出有价值的判断。评分有助于提高这种判断的效果，通过量化方法，在使用某种特征和追求预测精度之间找到平衡。

个人特征的代表量。谨慎对待评分的小额信贷机构应该设法记录能够代表个人品质、并且与还款纪律高度相关的特征。举例来说，在拉丁美洲，没有喝酒习惯的人更有可能认真对待还债问题。同样，每周（或者每天）参加宗教活动的人更有可能忠诚地遵守还款制度。宗教或者恶习在一些地方可能会很敏感（或者是不恰当或者是不合法），所以，贷款机构应该结合地方实际情况来采纳这些原则：

- 在过去年份中的饮酒量
- 在过去年份中的吸烟量
- 在过去年份中购入的彩票量
- 在过去年份中参加宗教仪式的次数
- 目前居民委员会或教堂小组的成员人数
- 最后一份领薪职位的时间
- 参加循环储蓄与信贷协会（ROSCA）
- ? 最近参加的日期
- ? 定期捐助额
- ? 捐助频率

参加循环储蓄与信贷协会标志着客户具有储蓄人和债务人的经历。循环储蓄和信贷协会还可成为一种可靠的资金来源，以资助客户偿还小额信贷机构的分期付款。

量化的主观判断。筛选定性风险的唯一方法就是派信贷员到现场，让他们面对面地了解申请人的情况（见方框 4）。然而，一个信贷员的主观判断可以被量化。这样一来，评分就可以揭示，比方说，变坏的概率与主观判断（是“平均”，还是“平均以上”）之间的关系了。打算在将来实施评分的小额信贷机构应该就以下三点（“平均以下”、“平均”以及“平均以上”）开始着手对主观判断进行量化：

- 总体信贷风险
- 回复的可信度与透明度
- 参照物的质量
- 企业家的雄心和创造力
- 商业前景
- 现金流的可变性
- 最近在家庭和企业上投资的力度
- 对贷款合同条款的掌握
- 家庭关系与非正式的支持
- 家庭和企业的纯洁性和构成

很明显，如果所有被批准的申请人都定级在平均以上，那么，它就发挥不了什么作用了。
贷款特征

小额信贷机构已经记录了大部分有预测作用的贷款特征：

- 申请提交日期*
- 放贷日期*
- 全部清偿日期*
- 所要求的额度*
- 所放款的额度*
- 分期付款额度*
- 分期付款次数*
- 分期付款周期*
- 利息率*
- 费用和佣金*
- 宽限期*
- 重新安排情况*
- 贷款用途*
- 担保类型*
- 担保的评估值*
- 联保人的身份

申请日期是用来衡量从贷款申请到贷款发放经历的天数。了解联保人可以让评分把他们的信贷记录（如果他们有的话）合并到申请人的评分结果中。如果联保人后来又为自己申请贷款，那么，他们所担保的贷款的还款记录，同样可以作为一个预测器供贷款机构使用。

放款机构的特征

放款机构的特征，更确切的说，具体的分支机构和指定的信贷员，对风险有很强的影响力。放款机构应该记录一些信贷员的简单特征。这样一来，评分不仅会描述出理想中的信贷员的样子，而且还可以更好地预测计分卡构建后所雇信贷员放出的贷款的风险：

- 性别
- 出生年份
- 婚姻状态(已婚或者未婚)
- 家庭人口数目
- 大学所学科目
- 最后完成的学业等级

附加数据的价值

如果有足够的坏贷款，根据以上标有星号的主要特征，机构就可以构建一个功能强大并且

有用的计分卡，因为小额信贷机构已经收集了其中的大部分特征。由上述所有特征编制而成的计分卡，其预测能力也许要比一个仅有主要特征的计分卡好 20%至 40%。

方框11：评分对“嘈杂”数据或者“肮脏”数据起作用吗？

小额信贷数据，如所有数据一样，总是有一些“肮脏”（误差）和“嘈杂”（围绕真实值的随机变动）的因素。举例来说，固定资产的价值是受“嘈杂”因素影响的，因为它很难评估。它也可能是“肮脏的”，因为信贷员有可能操纵评估，以至于他可以让他认可的申请符合小额信贷机构评估政策所要求的财务比率。

计分卡构造过程中的统计工作尽一切可能把“肮脏”和“嘈杂”的因素从预测（风险与特征间的联系）中过滤掉。如果没有任何征兆（或者，如果一个特征确实与风险不相关），那么，统计过程就显示出这种情况，并且从计分卡中删除这一特征。在许多情况下，明知它们是“肮脏”和“嘈杂”的数据，但它们仍然包含有用的信息征兆。

积累附加的数据会提高预测能力，但是，也将消耗更多的成本。这些成本主要产生于重新设计申请表、修改信息系统以接受附加的数据，以及输入新数据等方面。尽管信贷员不得不多做一些工作，但是，一个有文化的客户可以在首次申请时轻松地提供大部分的附加内容。

关于储存优质数据的指导原则

在获得人力资源之后，一个小额信贷机构最大的资产便是信息。但是，在通常情况下，正规的信息系统很薄弱，除了用于跟踪贷款，别无他用。评分的出现以及电子数据库频繁的使用使得数据质量成为焦点。

多年来，大多数小额信贷机构已经收集了一部分主要特征，但是，他们从来没有使用过数据。结果导致信贷员和数据输入员产生这样的认识：关注数据质量耗费时间，却得不到任何回报。随着评分的发展，数据质量变得重要起来。为了做好付出必要努力的准备，一线员工必须知道这种旧习惯已经不再适用，必须了解不适用的原因，以及他们自身将从这种变化中得到的收益。

为经营类型确定前后一致的定义

商业经营类型是三个最重要的有预测作用的特征之一，另外两个是过去的欠款情况，以及信贷员的身份。遗憾的是，关于经营类型的数据，质量通常不高，因为给定的代码可能包含的商业类型太广泛，因而不能明确区分高风险和低风险。但是，不管怎样，具有“肮脏”和“嘈杂”的数据总比没有数据强（见方框 11）。

商业经营类型编码不当的原因通常有三。第一，信贷员们没有遵循共同的定义。一个信贷员认为是酒吧，可在另一个信贷员眼中它却是餐厅。第二，信贷员着眼于产品，而

不是经营活动，举例来说，他们把制鞋者、修鞋人和卖鞋者统统归类为“鞋类”，但是他们的经营活动分别属于制造业、服务业和商业，并且具有大不相同的风险。第三，数据录入人员趋向于将信息归在总的标题下，例如，“食品销售”或者“商店”，而不是在一张很长的代码列表中寻找匹配者。

使一线员工意识到这个问题是改进工作的第一步。第二步，就是制作一张包含五十种或者更多的最常用的经营类型的列表，仔细定义每种类型，并教信贷员和数据输入员坚持这样做。这样一来，大约有 90% 的经营活动会在这五十个代码中找到，其余的 10% 可以编码到“其它”一栏中。第三步，是精确地确定经营活动（行业）的种类：

- 贸易（不可转变项目的交易）
- 制造业（可转变项目的交易；例如，贸易商、制造商的买和卖，但是他们所买的与他们所卖的不同）
- 服务业（专业劳动的交易，或者体力劳动的使用）
- 农业（直接来自土地的植物、动物或者矿产品的生产）

第四步是出台一项正式的书面政策，根据五十种经营类型和四项活动的定义为每个企业编码。第五步，在信贷员填写的表格中加入所有行业（附有定义）和所有经营类型的清单。第六步也就是最后一步，监控新系统的执行情况。

这是一项很大的工作量，但是，如果经营类型能恰当地被记录，那么它将有很强的预见性。如果借款人没有工薪收入，而且他们没有信用局的记录，那么，上面提到的这三个最主要特征对小额信贷计分卡而言缺一不可。

不要扔掉数据

与花费数年来建立计分卡相比（因为旧的数据已被丢弃），电子存储并不昂贵。长期未使用的数据是今天评分系统获得活力的源泉，并且它是未来市场调研和监控客户的关键。¹⁶掌握一条原则：数据一旦键入，一律保存。

从被拒的申请中收集数据

许多小额信贷机构喜欢使用拜访评分，以此缩短（或者跳过）一些实地访问调查。这意味着机构或者预测还款问题，或者预测拜访后的拒绝问题。在拜访之前预测还款问题可能凑效，但是，只有现场的测试，才可以确定预测的能力。（全面跟踪报告不能帮忙。）因为拜访计分卡的建立仅仅来自于通过定性筛选且获得批准的借款人的数据，所以，预测未经筛选的借款人就充满太多的不确定性（见方框 4）。信贷员仍然要拜访那些通过了拜访评分的申请人，因为，如果没有经过定性的筛选，评分将不予批准，只有拒绝。

现场访问后进行决策是一个更好的选择。要做到这一点，小额信贷机构必须首先将几千笔被拒申请的数据输入他们的信息系统。一旦他们拥有关于拜访后拒绝和拜访后批准两方面的数据，他们就可以构建计分卡，然后，根据拜访前所知道的特征来预测决定拒绝与否。（即使使用来自被拒申请的数据，预测还贷风险的拜访计分卡仍然不能批准未经访问的申请人，因为未经筛选的借款人的还贷行为仍然充满着变数。）

¹⁶关于市场研究，见 Inez Murray, “小额信贷机构的客户重视什么？来自三大洲的对比分析”（手稿，世界妇女银行业务，哥伦比亚，2001 年）。关于客户监控，见 Gary Woller, “在设计 and 实施一个效果监控系统时应注意的重要问题”（手稿，Brigham Young 大学，Provo，犹他州，2001 年）。

记录筛选员和监控员的特征

三个最重要的有预测作用的特征之一是信贷员的身份。负责贷款的人员有时会由于内部重组、工作量重新分配、或者工作调动的原因而改变。当此种情况发生的时候，大多数系统取消原来筛选员的特征，并且只记录目前监控员的情况。这在两个方面减少了评分的预测能力。第一，在一定程度上由原来筛选员构成的风险，被评分卡错误地归咎于监控员。第二，归咎于筛选员的风险忽视了贷款转给他人带来的后果。

解决的办法是在数据库中增加一个记录筛选员的字段。原来的“信贷员”的字段继续记录目前的监控员。如果一个人员至始至终地跟着一笔贷款，那么，筛选员与监控员即为同一个人。

记录这两种信贷员身份也许看似微不足道，因为大多数贷款通常仅有一个信贷员。但是，事实上，被评分定为高风险的信贷员经常说，他们继承了许多坏贷款，或者他们不得不放弃所有的好贷款。因此，信贷员的身份特征对于风险预测有很强的影响力。为了说服信贷员及信贷经理接受这种要求，机构就需要在计分卡的构建过程中对转手的贷款进行记账。反过来，这也要求数据库要有跟踪筛选员和监控员的记录。

空白的记录是未知的，而不是零

有时，申请人在申请表上留下一个空白记录，或者信贷员在拜访后，忘记了记录实地调查中的某项内容，或者数据录入员偶尔会漏掉一个字段。其结果是那里存在一个缺失的（未知的）数值。举例来说，如果申请人漏填“出生年份”，那么，他的年龄不是零，而是未知。

缺失的数值通常情况下也很有预测作用。举例来说，档案中没有经营收入记录的贷款可能会比有经营收入记录的贷款有风险。通常，缺少数据和发生还贷风险有共同的原因（例如，忽略了拜访，或者申请人有意隐藏某些情况）。遗憾的是，大多数小额信贷信息系统没有正确地记录漏掉的项。他们或者在空白处填上零、或者强制每个字段不能为空，这致使信息录入员将所有空白都改为零，或者虚构数据，或者为缺失的项编造前后矛盾的代码。（例如，一个大型小额信贷机构为上百个九十多岁的人发放了贷款，他们都出生于 1911 年 11 月 11 日，这太明显了。）

没有正确录入漏掉的项会对评分系统产生两种危害。第一，它排除了把缺失项用作预测特征的可能性。第二，它混淆了缺失项和真正“零”值项与风险之间的关系。举例来说，孩子的数目，通常可以记录为“非零”、“零”、或者漏记。没有提供孩子数目信息的借款人的风险，也许不同于填写没有孩子的借款人的风险。然而，将未知项替换成零，导致的结果就是，评分系统将同样的风险分配给这两组。解决方法是，为漏掉的项建立明确的代码，然后培训信贷员和数据录入员使用它。一些数据库语言已经为漏掉的项保留了编码。对于其它的程序语言，小额信贷机构可以使用“-99”。

无论要预测哪种类型的风险，统计评分都需要大量高质量的数据。即使是少数拥有充足数据库的小额信贷机构也应该着手在他们的信息系统中，输入信贷员的判断、信用局报

告以及被绝申请等内容。至于其他大多数的小额信贷机构，如果他们打算在几年后使用评分系统，那么，他们必须现在就修改信息系统。改善数据库的质量是一项十分艰苦的工作，但是，它不会比在永远没有评分的帮助下判断风险更艰难。

九、结论

评分将自我就业型穷人不如约还贷的风险量化。它还将还款风险与借款人、贷款和放款机构特征之间的关系明确化。最重要的是，评分提供了一种以量化风险和明确的权衡策略为依据的决策方式。这有可能加速组织文化的转变，因为经理们懂得他们寻求的是更完善、更准确的推断与决策。虽然简单的数据分析可以为决策提供信息，但是，大多数小额信贷机构还没有投资建立（更不要说利用）一个被视为财富的全面、准确的数据库。

小额信贷的评分，总体上说，在接近这个目标。举例来说，大约有 20%的预测风险为 20%的贷款确实变为坏贷款了。但是，还应看到，围绕平均值的误差和范围比起高收入国家中的评分大得多。遗憾的是，自我就业型穷人的大部分风险与可量化的特征不相关。因此，评分是信贷员以及他们主观评估的有益补充，而不是替代。它在信贷委员会中是一个“第三种声音”，一种对信贷员和信贷经理主观判断的补充。

评分的目的是预测风险。但是，对于一个要实施评分的小额信贷机构而言，它的预测能力放在第二位，因为机构可以使用历史数据提前对评分进行测试。小额信贷机构首先要关心的是说服董事会成员、管理者以及信贷员，让他们接受评分系统。总而言之，统计的弱点不会抹杀评分项目，然而人却可以。¹⁷评分（即使它的实施就象一场梦）总会产生一些人企图抵制的变化。让大众接受它就需要不断地培训和说教不同层次的利益相关者，并且不断展示评分对当前未偿贷款风险的预测能力。

评分可能不是小额信贷的又一个突破，但是，对于一些小额信贷机构而言，评分能够减少收款的时间，因而可以提高效率、增加覆盖面以及保持可持续性。随着越来越多的组织了解评分项目并且按程序去积累足够的数据，评分将很有可能成为小额信贷领域最佳实践的一部分。

一些人也许会说，评分是个新鼓弄出来的小把戏，对小额信贷来说可有可无。“如果它没坏，就不要修理它”，这是一个共识。在高收入国家里，这曾是贷款机构几十年老生常谈的话题，但是，现在评分几乎替代了人工评估，特别是，对于那些类似于小额信贷小型、短期、无担保的贷款而言，情形更是这样。¹⁸ 小额信贷很好，但是，它仍然有待改进，而且增长和竞争的压力会逐渐增加，这意味着最好的小额信贷机构必须先行一步，积极地寻求改变。等级评分就是争取领先地位的一种方法。

¹⁷ Mona A. Mayr, "实施评分：成功，或者不成功"《等级评分手册》Elizabeth Mays编著（芝加哥：Glenlake，2000年）：第 337-52 页；Kevin J. Leonard, 《金融统计》中的"等级评分和质量管理"，由David J. Hand 和Saul D. Jacka编著（伦敦：John Wiley 和Sons，1998年）：第 105-26 页；Leonard J. McCahill, 《信贷风险模型》中的"建立和保持信贷风险模型的组织问题"，由Elizabeth Mays编著（芝加哥：Glenlake，1998年）：第 13-22 页；和D. B. Edelman, 《等级评分和信贷控制》中的"分支银行引入等级评分"，由L. C. Thomas, J. N. Crook, 和D. B. Edelman编著（牛津：Clarendon出版社，1992年）：第 161-77 页。

¹⁸ Edward M. Lewis, "等级评分的引入"。

Filename: OP07Part2_ch.doc
Directory: N:\ForNeeranToPost\Final Chinese
Template: C:\Documents and Settings\wb294243\Application
Data\Microsoft\Templates\Normal.dot
Title: 等级评分：小额信贷领域的又一次突破？
Subject:
Author: xingda
Keywords:
Comments:
Creation Date: 12/21/2005 11:02:00 AM
Change Number: 3
Last Saved On: 12/21/2005 11:03:00 AM
Last Saved By: wb296321
Total Editing Time: 1 Minute
Last Printed On: 12/21/2005 11:20:00 AM
As of Last Complete Printing
Number of Pages: 35
Number of Words: 14,336 (approx.)
Number of Characters: 15,484 (approx.)