

观点与分析

编辑日期: 2026年11月1日

当银行流水成为一份人生履历

当账户活动成为信用信息

2025年6月，FinRegLab发表了一项研究，分析两家美国科技贷款机构在2015年8月至2024年5月间发放的近3.8万笔小企业贷款。研究者没有追问创业者是否品德高尚，或是否怀有卓越梦想。他们提出了一个更克制、后果却更重要的问题：如果个人信用史不足以说明一家企业的全貌，银行账户里的资金流动还能揭示什么？在样本中，对于成立不足五年的企业，账户余额每增加一个标准差——约6.4万美元——与违约概率下降超过两个百分点相关；对经营时间更长的企业，这一幅度约为一个百分点。这不是“余额更高就给每位申请人放贷”的公式，也不能证明提高余额会防止违约；它是在控制借款人与贷款特征之后观察到的统计关联。真正重要的是：同样的信息，对经营历史较短的企业更有价值。1

这份报告暴露了银行把“经验”带进机器时继承的一项旧缺陷。新企业为了成长申请贷款，却先被问经营了多少年；所有者希望依据企业表现接受判断，系统却调来他的个人借贷史。没有谁一定在刻意制造这个悖论。个人历史确实包含偿付行为的信息，企业年龄也能预测一部分存续风险。问题是，平均而言有用的指标，可能错误描述眼前这个具体的人。银行对新公司了解不足，就用它对所有者的了解来代替；企业信息的稀缺，因而可能变成对企业本身的负面判断。

美国国家经济研究局（NBER）2025年1月发布、2026年6月修订的另一篇论文，更接近问题核心。《为年轻创业者更新信贷渠道：从FICO到现金流》在其研究环境内通过因果分析得出：当贷款人把经营账户的现金流信息纳入判断时，年轻企业所有者的申请获批概率更高。收益集中在个人信用评分较低的年轻所有者身上，也不只是因为贷款人放宽标准、承受更多风险。这个结论比“AI让每个人都得到更多贷款”更窄，却也更重要。一项实验或一种放贷环境的结果，不能授权我们期待每家银行、每个国家和每个行业都会重现它。2

现金流能说什么，又不能说什么

年轻创业者为什么会输给自己的档案？设想一个人——他不是研究参与者，也不被当作真实客户。三年前，他开了一家工业设计工作室：第一个月收到项目预付款，第二个月支付材料和分包费用，第四个月客户才付尾款。全年净收入可能为正，利润率也很健康；但为“平均月薪”设计的表格，会像在跑者每迈一步后测量脉搏，然后断言心跳没有规律。延迟不总等于无力偿付，时间有时本就是收入生成方式的一部分；在另一家企业里，同样的延迟也可能预示危险。算术外观相似，无法消除因果差异。

技术话语常把三件事混在一起：钱有多少、何时到账、是否可用于偿债。一笔十万里亚尔的入账，可能是销售额而非利润，是尚未完全赚取的客户预付款，也可能只是所有者从另一账户转来、并非新的商业活动。发工资、缴税和付供应商之前的高余额，可能只是路过的流动性，不是可支配储备。相反，金额较低但持续发生的收入，若有现金缓冲，可能更能安全支持债务。没有一个流水数字能脱离会计与时间背景，独自承载所有这些含义。

经济学家把容易在机构内部编码、传输和比较的数据称为“硬信息”（Hard Information）；“软信息”（Soft Information）则依赖难以转成字段的本地判断，例如主要客户的可靠性、工作室所有者的技能、他与供应商谈判的方式，或采购合同的期限。Jeremy Stein在2002年发表的理论研究中，把机构使用软信息的能力，同组织结构及信息跨越决策层级的难度联系起来。这里不是要浪漫化“认识每个人”的员工：人的判断也会受偏袒与错误印象影响。真正的教训是，把信息变得可计算，并不会自动补全它的意义。3

技术上究竟改变了什么？在获得授权的账户访问后，机构可以分析存款、取款、平均余额、资金紧张日、退回付款以及波动程度。过去，从冗长流水中提取这些模式需要大量人工时间，机构也可能只把工作压缩成几个简单指标。现在，分类算法可以提取现金流特征（Cash-Flow Features）并用于违约模型。贷款人可以使用机器学习（Machine Learning）、逻辑回归（Logistic Regression）或固定政策规则。“阅读银行流水”不必然等于人工智能，每一次信息改善也并非算法创新。

FinRegLab的小企业研究除信用评分与企业年龄外，还使用账户收入和支出、余额与存款波动，以及余额为负或资金不足的次数。研究者总体上发现，余额和存款越稳定，违约越少；部分困境信号则与更多违约相关。这些信息对年轻企业和个人评分较低的所有者，具有更高的相对价值。这并不说明每家波动大的企业都安全，而是说明：不完整的信用过去，不应垄断对当下的解释。1

这项结果有一个基本边界。FinRegLab约3.8万笔贷款的样本，都是已经实际发放的贷款。它告诉我们，如何预测已经获得融资者中的违约，却没有直接揭示所有被拒者后来会怎样。我们不能把预测准确度的改善乘以一国人口，就宣布算法“拯救”了多少被排除的人。研究没有向每位被拒申请人放贷，再观察另一种人生如何展开。即使窗外景色诱人，研究者也必须知道这扇窗的边界。1

小幅进步不等于公平普惠

2025年7月，FinRegLab发表另一项消费信贷研究，比较传统模型与机器学习模型、征信机构数据与现金流数据，以及混合模型。市场宣传常把两种收益合并，分析时必须拆开。在指定环境下，从逻辑回归转向对完整数据集使用机器学习，使统计区分指标ROC-AUC提高约2%。在采用保守风险阈值的模拟中，

一些使用征信数据的机器学习模型，相对传统对应模型使批准率提升约4%，同时让获批申请人中预计违约的比例至少下降9%。但是，单独把现金流信息加入信贷数据，增益更窄：ROC-AUC约提高0.3%至0.7%，在保守阈值下模拟批准率约提高0.6%至1.6%。那4%不是“加入银行流水”的独立效果，模拟批准也不等于向真人发放贷款。4

这些数字说出了一件在热衷“革命性”措辞的行业里并不常见的事：一些改善可以既有用，又很小。在大型资产组合中，模型区分可能违约者与不违约者的能力哪怕只改善一点，也有财务意义；但它不能证明公平普惠。预测准确度可能提高一个幅度，某位申请人的批准结果却改变另一个幅度，甚至完全不变。研究发现，在更高风险阈值下，批准效果可能缩小或转为负值。严谨研究者不会只问一个指标加了多少点，而会先问：哪项决定改变了、对谁改变、代价是什么。

风险与季节性之间的不规律

不规律的节奏并不总是无辜的。JPMorgan Chase Institute使用高频交易数据，研究美国25座城市的小企业现金流模式，发现现金流不规则的企业更可能退出，收入增长也更慢；同时拥有不规则流量与薄弱现金缓冲的企业，存续前景最差。因此，一部分波动确实表示真实脆弱性，而非行政误解。若一种风险恰好也像社会障碍，就要求贷款人与存款人否认它，同样不公平。5

但同一项研究也打开了一道精确的区分：扰动并非只有一种。收入可能间隔很长却有可预测的季节性，也可能随机下降；支出可能剧烈变化，账单也可能在收入尚未到来的那一周集中出现。风险可能来自储备不足、需求下滑，也可能来自开票与回款日期错配。若模型只保留“存款标准差”，就把这些不同疾病安排在同一个候诊室。标准差在统计上有用，却不能独自解释扰动究竟源于有意安排的供给周期，还是销售正在崩塌。

这正是替代数据（Alternative Data）的悖论：贷款人引入更贴近个人现实的信息，试图摆脱过去的概括，却又可能用它惩罚现实中更细的差异。企业所有者不再需要很长的信用史，却被实际要求维持类似受薪公务员的账户形态：每月收入相近，支出可预测，很少出现受压日。如果如此，机构并没有扩展“信用良好”的含义，只是把旧模板从债务历史搬到账户活动。它用概率而非条件说话，看起来更技术化，却仍可能限制工作本身的节奏。

自由职业者的生活也不能和贸易企业混为一谈。前者可能依靠季节性项目出售工时或创意成果，后者则有回款周期、库存、供应商和工资。JPMorgan Chase关于在线平台经济的数据表明，平台收入可能间歇发生，而且常常只是其他收入的补充。把一个平台账户与家庭收入和义务割裂开来，会向两个方向误判：对另有收入来源的人夸大风险，或对一个平台上看似活跃、授权窗口之外却背负义务的人夸大安全性。6

更深的问题是样本缺失部分的架构。一个账户并不等于完整的金融生活。经营者可能在不同机构使用个人账户和商业账户；如果聚合失败，账户间转账可能被重复计为收入。客户退款可能被归作普通支出，尽管它是在修正更早的交易。现金取款未必显示最终收款人。偏好现金或与家人共用账户的人，可能只显得“数据贫乏”，并非财务薄弱。每当模型声称“读完整个账户”，都应追问：哪个账户、哪些交易、依据什么授权，以及分类质量如何。

相同数据在不同时间也可能有不同含义。斋月会改变餐饮企业的收入与营业时间；朝觐季可能把服务收入集中在几周，而全年其他时段完全不同；建筑企业可能先购入材料，再收到经认证的工程款。这些只是分析上的可能性，不是关于沙特借款人的已发表研究结果。如果模型以完整周期训练，运行时却只收到一个月数据，它看到的不只是“信息缺失”，还可能把周期中的错误片段当成全部。延长数据窗口也不总能解决问题；还必须界定这个窗口意在代表哪一种经济周期。

贷款与经营节奏

时间的影响，远大于“偿付能力”一词所暗示的程度。能力不像身高那样是人的静态属性，而是可用现金、到期义务与新债安排之间的关系。两家企业年收入与净利润可以完全相同，一家七天回款，另一家九十天后才收款。若分期从下个月开始，问题不是第二家在道德上更不值得信任，而是贷款形式可能更不适合它的经营周期。当贷款人把违约史视为与还款安排、定价和担保无关的独立事实，它以为测量了一个人，实际测到的却是产品与人互动的一部分。

问题应拆成两个：经营活动本身有多大风险？我们强加一个不同于回款周期的安排，又增加了多少风险？第二问不是美化信用状况的技巧。延后还款会让贷款人付出成本或减少流动性，借款人的费用和总成本也可能上升。不能提出季节性分期，就宣布公平取得胜利。资产组合风险、资金成本、实际负担，以及收入冲击将借款人推过承受边界的概率，都必须衡量。但产品设计仍是一项可检验变量，不是算法无需承担责任的天命。

本文提出一个分析概念：“信用节奏匹配”（Credit-Rhythm Fit）。它不是获监管认可的指标，也不是已经训练和测试的模型。它把问题分成四层：扣除成本与义务后的经营经济能力；现金收入的时间分布，而不只是年度总额；客户延迟付款或旺季中断时的韧性余量；以及贷款形式，包括到期日和收款机制。检验要问的是：理解这些层次之间的关系，是否能比把收入压缩成一般月度规律的模型更好地预测风险，并给出更安全的融资方案。

第一层不把入账等同于净收入：剔除账户间转账，区分客户付款与新增融资，并正确放置经营义务。第二层不把所有波动当成一件事，而是分开可重复季节性、无法解释的起伏和下降趋势。第三层用保守的延迟付款压力测试，监测“现金缓冲天数”——储备大致能覆盖必要支出的时间。第四层在合同和控制允许时，把适配已验证现金流的还款安排与固定安排比较。任何一层都不构成融资承诺；最终决定必须保持借款人理解方案的能力，也不能把灵活性变成隐藏成本。

再给出一个明确的数字例子。假设一家服务企业全年平均每月经营盈余为两万里亚尔，但大部分盈余集中在六个月，而固定义务覆盖十二个月。平均盈余不能单独证明一笔较高的月供可承受；波动也不能单独证明融资不可行。我们还需要收入日期、储备、义务性质、延误成本，以及能否在不增加风险的情况下重新安排偿还。上述数字只是用于解释的假设，不代表沙特客户或真实模拟结果。平均数抹去了事件顺序，而信用违约发生在具体日期，不发生在年度均值里。

模型复杂并非美德。每增加一个特征，都要证明收益大于噪声和收集负担。复杂模型可能扩大两类人之间的差距：一类拥有长期且分类清晰的银行记录，另一类的数据零散。对公平的承诺可能创造一种新特权：可被清晰读取的特权。缺少数据的人不一定不值得信任，能把每个字段填满数字的人也不一定稳

健。有时，面对不完整数据，科学上正确的推断是暂缓确定、要求补充证据，或提供小额安全方案，而不是给出虚假自信的低评分。

公平、解释与申诉

这里的公平可以拆成三个问题。第一，表征公平：我们是否有足够信息真实描述经营周期，还是惩罚那些不符合数据采集形态的人？第二，估计公平：在得到相同风险评分的人中，实际违约率是否在不同企业年龄、收入类型，以及法律允许研究的群体之间大致一致？第三，决策公平：在风险可比时，申请人能否得到可负担的条款、可理解的解释与纠错机会，还是来源不明的信息会变成密不透风的拒绝？一个公平指标无法替代这三层，也不能把任何单一定义包装成科学共识。

当信贷系统说“因波动而拒绝”，究竟是什么意思？是销售额、余额、交易时间，还是账户间转账频繁？测量覆盖的是两个异常月份，还是完整旺季？银行转账被错误分类后，申请人能否要求重新评估？在美国，《平等信贷机会法》（Equal Credit Opportunity Act）及其实施规则Regulation B第1002.9条，依适用情形要求受规制信贷方把不利信贷行动的主要具体原因告知申请人；当规则要求真实原因时，“内部政策”之类泛称并不足够。旧的解释性材料也不能当作现行法律反复引用：2025年5月，美国消费者金融保护局撤回了有关AI模型不利行动通知与表格的解释材料，但规章条文本身仍然有效。这是美国制度，不是对沙特法律的描述。7

可解释性并不是“某人破解黑箱后安然入睡”的故事。FinRegLab与斯坦福研究者考察了消费信贷中用于模型解释与公平的商业和开源工具。相关文献区分了“哪些因素影响模型结果”与“客户能否理解、核验或纠正的理由”。局部解释可以说余额波动压低了结果，却不能独自证明波动记录准确、源于正常季节、后来已经消失，或客户能够采取行动。对工程师可解释的计算，对被计算的人仍可能是谜。8

开放银行不等于无限授权

在沙特阿拉伯，沙特中央银行公布的开放银行（Open Banking）框架，使金融数据能够依照明确用途、规则与技术标准，安全地共享给其他服务提供者。它为跨多个账户想象更全面的评估提供了重要基础设施，却不意味着每一次数据共享都自动授权建立信用评分，也不意味着只要技术接口存在，评估模型就适合本地市场或具备公平性。访问数据的能力、为特定目的使用数据的权利，以及由数据得出的推断质量，是三件不能折叠成一件的事。9

未来若在合法同意与适当保护下开展，较有价值的沙特检验可以研究那些不像工资表的企业：小型承包商、季节性商户、创意项目和信用史较短的创业者。研究不能预设这些群体比雇员更好或更坏，也不能预设数据越多公平就越好。它可以在按时间和行业分隔的数据上，比较基准信用模型、混合模型与不同还款安排；应测试训练期之外以及季节改变时的表现，并报告批准、违约、总成本和分类申诉的差异，而不只是平均准确率。研究必须由获授权机构执行，并保护账户机密。这是一项研究建议，不是某个实验已经发生的公告。

还要避免常见的障眼法：只把新近获批者与过去获批者比较远远不够。新产品会改变谁来申请、贷款规模、价格、时点，乃至经济环境本身。因果影响研究需要适当设计与独立审查，可能包括受控试验或不损害申请人权利的比较策略。它还必须区分信贷决策的质量与信贷产品的质量。系统可能正确判断一个人能够偿还，随后却用不合理的时间表给他增加负担。设计错误往往消失在“客户风险”这个悦耳短语背后。

实验计分表可以很简单：样本外违约预测改善多少？短信用史企业中，多少比例的决定真正改变，贷款后来表现如何？在群体、行业与季节之间，预测风险和实际风险的差距有多大？多少决定依赖错误分类的交易，其中多少在放款或拒绝前得到纠正？总借款成本与分期负担是否改变？多少逾期源于贷款时间安排，而不是企业本身无力偿付？这些问题把“金融普惠”从道德标签变成可以证伪的主张。

贷款人的责任与借款人的权利之间

反方有一项值得尊重的论点：若贷款人为改善创业者形象而损失存款人或投资者的资金，并没有真正帮助创业者；与违约相关的不规律信号，也不会仅因伤害表现出这些信号的人，就变成隐性歧视。这是正确的。负责任的机构有权衡量信用损失、资金、欺诈与流动性。本文并非要求从方程中删除波动，也不是把同情心当作监管资本，而是要求在波动变成定价惩罚或排除决定之前，先拆解它的含义；并检验调整贷款时点，是否能比彻底拒绝更低成本地解决部分风险。证据也可能显示某个群体根本无法承担债务。此时，诚实而有解释的拒绝是一项服务，而非普惠失败。

旧日“银行经理认识客户”的浪漫想象同样救不了我们。经理可能在与自己世界相似的人身上看见前景，却忽略一家更自律、但不熟悉的企业。留有记录的测量可以减少一部分个人任性和偏袒，并留下可审计轨迹。选择不是人或机器，而是如何分工：计算用于捕捉模式，却不混淆因果与关联；政策对应已经声明的风险；真正的人类复核负责缺失信息和后果重大的争议；当到期日被证明是问题时，金融产品也有能力调整。

几个世纪前，商人把熟人的证词带给贷款人；后来，还款历史成为主要文件。今天，账户活动被要求写出历史遗漏的一章。每一种新证词都显得离真相更近：从声誉，到义务，再到收入与支出的细节。但更接近生活细节，并不赋予我们随意解释它的权利。大额入账不一定是利润，沉默的月份不一定意味着企业已经死亡，规律数字也不一定代表一项义务可持续。

真正值得欢迎的变化，不是银行终于知道客户星期四买了什么，或余额多少次跌破某个门槛；而是银行或许能够提出一个比“你的生活是否像我们过去偏爱融资的借款人？”更精确的信贷问题。只要评估受检验结果以及当事人申诉权约束，这就是科学成熟的开始。如果银行流水变成一份最终判决式的人生履历，我们只会走到一个更尖锐的悖论：新开一扇窗去看见每个具体的人，随后却因为他的生活没有排成月度表格，而把门关上。

参考资料

1. FinRegLab，聚焦：使用现金流数据为融资受限企业承销，2025年6月。
2. Christopher M. Hair、Sabrina T. Howell、Mark J. Johnson与Siena Matsumoto，为年轻创业者更新信贷渠道：从FICO到现金流，NBER工作论文33367号，2025年1月，2026年6月修订。该文为工作论文，并非同行评审期刊论文。
3. Jeremy C. Stein，信息生产与资本配置：分散型企业与层级型企业，《金融学期刊》，2002年10月。
4. FinRegLab，推进信贷生态：消费信贷承销中的机器学习与现金流数据，2025年7月。
5. JPMorgan Chase Institute，面对不确定性：美国25座城市的小企业现金流模式。
6. JPMorgan Chase Institute，工资、发薪日与在线平台经济。
7. 美国消费者金融保护局，Regulation B第1002.9条：通知；已撤回指引，包括2025年5月12日撤回的2022-03与2023-03号通告。
8. FinRegLab，机器学习的可解释性与公平：来自消费信贷的洞见，2023年7月更新。
9. 沙特中央银行，沙特阿拉伯的开放银行。