SAFWAN ALMODHAYAN



OPINION AND ANALYSIS

Editorial date: November 1, 2026

When the Bank Statement Became a Biography

When Account Activity Became Credit Information

In June 2025, **FinRegLab** published a study of nearly **38,000 small-business loans** originated by two US technology lenders between August 2015 and May 2024. The researchers did not ask whether an entrepreneur was virtuous or possessed a brilliant dream. They asked something more modest and consequential: if personal credit history does not tell the whole story of a business, what does movement through its bank account reveal? Within the sample, a one-standard-deviation increase in account balance—about **$64,000**—was associated with a reduction of more than **two percentage points** in default probability for businesses less than five years old, compared with about **one percentage point** for older firms. This is not a formula granting a loan to every applicant with a larger balance, nor proof that raising balances prevents default. It is a statistical association conditional on borrower and loan characteristics. Its importance lies elsewhere: the same information was more useful when the business had a shorter history.1

The report exposes an old weakness in the idea of “experience” that banks carried into their machines. A new firm seeks a loan to grow and is asked how many years it has operated. Its owner asks to be judged on business performance, and personal borrowing history is summoned instead. No one necessarily intends the paradox. Personal history contains real information about repayment behavior, and firm age predicts some survival risk. The problem

is that **an indicator useful on average may misdescribe the person standing before it**. When a bank knows little about a new company, it uses what it knows about the owner. Scarcity of business information can thereby become a negative judgment about the business itself.

Another paper, issued by the US **National Bureau of Economic Research** in January 2025 and revised in June 2026, comes closer to the heart of the problem. *Modernizing Access to Credit for Younger Entrepreneurs: From FICO to Cash Flow* concludes, through causal analysis within its study setting, that applications from younger business owners had higher approval probabilities when the lender incorporated cash-flow information from the operating account. Benefits were concentrated among younger owners with lower personal credit scores and were not explained merely by a more permissive lender accepting more risk. This is narrower and more important than the headline “AI gives everyone more loans.” Results from one experiment or lending environment do not license the same expectation in every bank, country, and industry.2

What Cash Flow Says—and What It Does Not

Why might a young entrepreneur lose to their file? Imagine someone who is not a study participant and is not presented as a real customer. Three years ago, they opened an industrial-design workshop. In month one, the workshop receives an advance for a project; in month two, it pays for materials and subcontractors; in month four, the client pays the balance. Annual net income may be positive and the margin healthy, yet a table designed for “average monthly salary” would read the business like a runner's pulse after every step and conclude that the heart lacks regularity. Delay does not always mean incapacity; time may be part of how revenue is produced. In another business, the same delay may signal danger. Arithmetic resemblance does not settle causal difference.

Technology discourse often conflates three things: **the amount of money, when it arrives, and whether it is available for repayment**. A deposit of one hundred thousand riyals may be sales rather than profit, a customer advance rather than fully earned revenue, or a transfer from the owner's other account rather than new commercial activity. A high balance just before payroll, taxes, and supplier bills may be passing liquidity rather than a usable reserve. Lower but recurring income accompanied by a cash buffer may support debt more safely. No single number on a statement can carry all these meanings without accounting and temporal context.

Economists call some data “hard information” because it can be coded, transmitted, and compared across an organization. “Soft information” involves local judgment that resists conversion into a field: the reliability of a major customer, a workshop owner's skill, the way that person negotiates with a supplier, or the age of purchase contracts. In theoretical research published in 2002, **Jeremy Stein** connected an institution's capacity to use soft information

with its organizational structure and the difficulty of moving that information through decision layers. The lesson is not to romanticize an employee who knows everyone; human judgment is vulnerable to favoritism and mistaken impressions. The lesson is that making information computable does not make its meaning complete.3

What changed technically? With authorized account access, it became possible to analyze deposits, withdrawals, average balances, low-cash days, returned payments, and volatility. Extracting such patterns from long statements once consumed human time, and institutions might reduce the work to simple measures. Classification algorithms can now derive cash-flow features and use them in a default model. A lender may use machine learning, logistic regression, or fixed policy rules. “Reading the bank statement” is not necessarily synonymous with AI, and not every informational improvement is an algorithmic innovation.

FinRegLab's small-business study used account revenue and withdrawals, balance and deposit volatility, and the frequency of negative or insufficient balances alongside credit score and business age. Researchers generally associated more stable balances and deposits with less default, and some distress signals with more default. The relative value of this information was greater for younger firms and owners with lower personal scores. That does not mean every volatile business is safe. It means **an incomplete credit past should not monopolize the interpretation of the present**.1

There is a basic limit to this finding. FinRegLab's sample contains about 38,000 loans that **were actually made**. It tells us something about predicting default among borrowers who reached funding. It does not directly reveal the outcomes of everyone the lenders rejected. We cannot multiply an improvement in predictive accuracy by a country's population and announce the number of excluded people an algorithm “saved.” The study did not lend money to every rejected applicant and observe the alternative life that followed. Researchers must know the boundaries of their window even when the view is attractive.1

Small Gains Do Not Equal Fair Inclusion

In July 2025, FinRegLab published a different study of **consumer lending** comparing traditional and machine-learning models, credit-bureau and cash-flow data, and hybrid models. Two gains often combined in one announcement must be separated. Moving from logistic regression to machine learning on full data sets improved the statistical discrimination measure **ROC-AUC by about 2%** in specified settings. In a simulation using a conservative risk threshold, some machine-learning models using bureau data raised approvals by about **4%** relative to conventional counterparts while reducing the share of approved applicants expected to default by at least **9%**. But simply **adding cash-flow information** to credit data

produced narrower gains: roughly **0.3% to 0.7%** in ROC-AUC and **0.6% to 1.6%** in simulated approvals at a conservative threshold. The four percent is not the isolated effect of adding a bank statement, and simulated approvals are not loans disbursed to real people.4

The figures say something rare in an industry fond of “revolutionary”: some improvements are useful and small at the same time. A modest improvement in separating likely defaulters from non-defaulters may matter financially in a large portfolio, yet it does not prove fair inclusion. Predictive accuracy may improve by one amount while an individual applicant's approval outcome changes by another—or not at all. The study found that approval effects could shrink or become negative at higher risk thresholds. A good researcher does not ask how many points were added to a metric before asking which decision changed, for whom, and at what cost.

Irregularity Between Risk and Seasonality

Irregular rhythm is not always innocent. In a study of small-business cash-flow patterns across **25 US cities**, the **JPMorgan Chase Institute** used high-frequency transaction data and found that firms with irregular flows were more likely to exit and experienced slower revenue growth. Survival prospects were worst among those combining irregular flows with weak cash buffers. Some volatility therefore signals real fragility, not administrative misunderstanding. It would be unfair to lenders and depositors to demand that a model deny risk whenever risk resembles a social obstacle.5

The same study opens a precise distinction: disruptions are not all one kind. There is spaced but predictable seasonal revenue, random and declining revenue, sharply variable expenses, and bills arriving in a week that revenue does not recognize. Risk may reside in a deficient reserve, falling demand, or a mismatch between billing and collection dates. A model that retains only “standard deviation of deposits” seats these different illnesses in one waiting room. Standard deviation is statistically useful, but it cannot by itself explain whether disruption comes from a deliberate supply cycle or collapsing sales.

This is the paradox of alternative data. A lender brings information closer to a person's reality to escape historical generalizations, then may use it to penalize the finer detail of that reality. An entrepreneur is relieved of the requirement for a long credit history but effectively asked to maintain the account pattern of a salaried public employee: similar income each month, predictable spending, and few days under pressure. If so, the institution has not broadened creditworthiness. It has moved the old template from debt history into account activity. The constraint looks more technical because it speaks in probabilities rather than conditions, but it may still constrain the rhythm of the work itself.

Nor should we conflate a freelancer's life with that of a trading firm. One may earn from seasonal work, selling hours or creative output; the other may have a collection cycle, inventory, suppliers, and payroll. JPMorgan Chase data on the online platform economy show that platform income can be intermittent and often supplements other income. Reading one platform account apart from household income and obligations may fail in two directions: exaggerating risk for someone with other sources, or exaggerating safety for someone who looks active on one platform while carrying obligations outside the authorized window.6

The deeper problem is **the architecture of what is absent from the sample**. One account is not a financial life. A business owner may use personal and commercial accounts at different institutions. A transfer between their accounts may be counted twice as income if aggregation fails. A customer refund may be classified as an ordinary expense when it corrects an earlier transaction. A cash withdrawal may not identify its beneficiary. Someone who favors cash or shares a family account may look data-poor rather than financially weak. Whenever a model is said to "read the whole account," ask which account, which transactions, under which authorization, and with what classification quality.

The same data may also mean different things over time. Ramadan may change a food business's revenue and operating hours; the Hajj season may concentrate service income into weeks unlike the rest of the year; a construction business may purchase materials before receiving certified payments. These are analytical possibilities, **not published findings about Saudi borrowers**. If a model is trained on a full cycle but receives only a single month in operation, it sees more than "missing information": it may mistake the wrong part of a cycle for the whole. A longer data window is not always enough; the economic cycle the window is meant to represent must be defined.

The Loan and the Rhythm of the Business

Timing has a greater effect than the phrase "ability to repay" implies. Ability is not a static human property like height. It is a relation among available cash, obligations due, and the schedule of new debt. Two businesses may have identical annual revenue and net profit, while one collects within seven days and the other after ninety. If installments start next month, the issue is not that the second firm is morally less worthy, but that **the loan's form may fit its business cycle less well**. When a lender treats default history as independent of schedule, pricing, and collateral, it may think it measured a person when it measured part of the interaction between product and person.

The question should be divided. How much risk lies in the activity itself? How much risk do we add by imposing a schedule unlike its collection cycle? The second question is not a device for flattering creditworthiness. Flexibility can cost the lender or reduce liquidity when repayment is deferred, while fees and total costs to the borrower may rise. Seasonal installments cannot

simply be proposed and declared a triumph of fairness. Portfolio risk, funding cost, actual burden, and the chance that revenue shocks push a borrower beyond capacity must all be measured. Still, product design is a testable variable, not divine fate for which an algorithm bears no responsibility.

This article proposes an analytical concept called **credit-rhythm fit**. It is not an approved regulatory measure or a trained and tested model. It separates four layers: **the economic capacity of the activity** after costs and obligations; **the timing distribution of cash receipts**, not only their annual total; **the resilience margin** when a customer pays late or a season stops; and **the form of the loan**, including due dates and collection mechanism. The test would ask whether understanding relations among these layers improves prediction and makes financing offers safer than a model reducing income to general monthly regularity.

At the first layer, deposits are not equated with net income: transfers among accounts are removed, customer payments distinguished from new financing, and operating obligations placed correctly. At the second, volatility is separated into repeatable seasonality, unexplained fluctuation, and declining trend. At the third, “cash-buffer days”—the approximate time reserves can cover essential expenses—are monitored with conservative stress tests for delayed payments. At the fourth, where contract and controls permit, a schedule adapted to verified flows is tested against a fixed one. No layer becomes a promise of financing. The final decision must preserve the borrower's capacity to understand and not turn flexibility into hidden cost.

For an explicit numerical example, imagine a service business with average monthly operating surplus of twenty thousand riyals over a year, but six months contain most of that surplus while fixed obligations run through all twelve. Average surplus alone does not show that a large monthly installment is affordable; variation alone does not show finance is impossible. We need receipt dates, reserves, the nature of obligations, delay costs, and whether repayment can be realigned without creating more risk. These figures are **hypothetical and explanatory**. They do not represent a Saudi customer or an actual simulation result. An average removes event order, while credit defaults on particular dates, not in an annual mean.

Model complexity is not a virtue in itself. Every added feature needs evidence that its benefit exceeds its noise and collection burden. More complex models may widen the gap between people with long, clearly classified banking histories and those whose data are fragmented. Promised fairness may create a new privilege: **the privilege of legibility**. Not everyone about whom we lack data is unworthy, and not everyone filling our fields with figures is sound. Sometimes the scientifically correct inference from incomplete data is to withhold certainty, request supplementary evidence, or offer a small safe option rather than assigning a falsely confident low score.

Fairness, Explanation, and Appeal

Fairness here divides into three questions. **Representational fairness:** do we have enough information to describe the business cycle, or do we punish people who fall outside our data-collection pattern? **Estimation fairness:** among people assigned the same risk score, are realized default rates similar across firm ages, income types, and groups the law permits us to examine? **Decision fairness:** at comparable risk, are applicants offered payable terms, a comprehensible explanation, and a chance to correct error—or does information of unknown origin become an airtight rejection? One fairness metric cannot replace these layers, and no single definition should be presented as scientific consensus.

What does it mean when a credit system says "declined because of volatility"? Sales volatility, balances, transaction timing, or frequent transfers among accounts? Did the measurement cover two unusual months or a whole season? If a bank transfer was misclassified, can the applicant seek reassessment? In the United States, the **Equal Credit Opportunity Act** and its implementing **Regulation B, section 1002.9**, require a covered creditor, as applicable, to notify an applicant of the principal specific reasons for adverse credit action; general phrases such as "internal policy" do not suffice when the regulation requires actual reasons. Older interpretive publications should not be recycled as current law: in May 2025, the Consumer Financial Protection Bureau withdrew interpretive materials concerning adverse-action notices in AI models and model forms, while the regulatory section itself remained. This is **a US framework**, not a description of Saudi law.7

Explainability is not a tale in which a person decodes a black box and sleeps peacefully. FinRegLab and Stanford researchers examined commercial and open-source tools for model explanation and fairness in consumer lending. The literature distinguishes "factors that influenced a model result" from "reasons a customer can understand, verify, or correct." A local explanation may say balance volatility pushed a decision downward, but it cannot alone establish that the volatility was recorded accurately, resulted from an ordinary season, was later resolved, or can be acted upon by the customer. A calculation explainable to an engineer may remain a mystery to its subject.8

Open Banking Is Not Open-Ended Permission

In Saudi Arabia, the **Open Banking** framework announced by the Saudi Central Bank makes it possible to share financial data securely with other service providers under defined uses, rules, and technical standards. This is important infrastructure for imagining assessment across a broader set of accounts. It does not mean every instance of data sharing is open permission to build a credit score, or that the presence of technical interfaces makes an

assessment model valid or fair for the local market. **The ability to access data**, **the right to use it for a specific purpose**, and **the quality of the inference drawn from it** are three matters that must not be collapsed.9

A useful future Saudi test—if created with lawful consent and appropriate protections—could study businesses that do not resemble a payroll ledger: small contractors, seasonal traders, creative projects, and entrepreneurs with short credit histories. It should not assume these groups are better or worse than employees, or that more data must improve fairness. It could compare baseline credit models, hybrid models, and different repayment schedules on data separated by time and sector. It should test performance outside the training period and as seasons change, and report **differences in approval, default, total cost, and classification complaints**, not average accuracy alone. Authorized institutions would need to conduct the study and protect account confidentiality. This is **a research proposal**, not an announcement that an experiment has occurred.

The familiar trick must be avoided: comparing newly approved borrowers only with previously approved ones is insufficient. The new offer can change who applies, loan size, pricing, timing, and the economy itself. Causal impact research needs a suitable design and independent review, perhaps including controlled trials or comparison strategies that do not damage applicants' rights. It must also separate the *quality of the credit decision* from the *quality of the credit product*. A system may correctly find that someone can repay and then burden that person with a badly scheduled loan. Design error often disappears behind the attractive phrase “customer risk.”

The experiment's scorecard can be simple. How much does out-of-sample default prediction improve? What share of short-history firms actually receive a different decision, and how do their loans later perform? How wide is the gap between predicted and realized risk across groups, sectors, and seasons? How many decisions relied on misclassified transactions, and how many were corrected before disbursement or denial? Did total borrowing cost and installment burden change? How much delinquency results from loan scheduling rather than incapacity in the business? These questions turn “financial inclusion” from a moral label into falsifiable claims.

Between the Lender's Responsibility and the Borrower's Right

There is a counterargument that deserves respect. A lender that loses depositors' or investors' money to improve an entrepreneur's image does that entrepreneur no service, and irregularity signals associated with default are not covert discrimination merely because they hurt those who display them. Correct. A responsible institution is entitled to weigh credit loss, funding, fraud, and liquidity. I am not calling for volatility to be removed from the equation or

for sympathy to become regulatory capital. I am calling for **the meaning of volatility to be decomposed before it becomes a pricing penalty or an exclusion decision**, and for tests of whether changing the loan's timing can address part of the risk more cheaply than rejecting it altogether. Evidence may show that a group cannot bear debt at all. Honest, explained refusal is then a service, not a failure of inclusion.

Nor will the old romance of “a bank manager who knows the customer” save us. A manager may see promise in someone who resembles the manager's world while overlooking a more disciplined business that does not. Documented measurement can reduce some personal whim and favoritism and leave an auditable trace. The choice is not humans or machines, but a distribution of their work: calculations that capture patterns without confusing causation and association; policies tied to declared risks; genuine human review of missing information and consequential disputes; and a financial product capable of changing when the due date proves to be the problem.

Centuries ago, a merchant carried testimony from acquaintances to a lender. Repayment history later became the main document. Today, account activity is asked to fill the gaps left by repayment history. Each new testimony seems closer to truth: from reputation, to obligations, to the details of income and spending. Yet proximity to life's detail does not grant a right to interpret it however we please. A large deposit is not necessarily profit. A silent month is not necessarily a dead business. Regular figures do not necessarily show a sustainable obligation.

The transformation worth welcoming is not that a bank can learn what a customer bought on Thursday or how often a balance crossed a threshold. It is that the bank may become able to **ask a more precise credit question than: does your life resemble the lives of borrowers we preferred to finance before?** That is the beginning of scientific maturity if assessment remains constrained by what testing establishes and what a person can contest. If a bank statement becomes a final biography, we end with a harsher paradox: a new window is opened to see each person, then the door is closed because that person's life failed to arrange itself as a monthly table.

References

1. FinRegLab, Sharpening the Focus: Using Cash-Flow Data to Underwrite Financially Constrained Businesses, June 2025.
2. Christopher M. Hair, Sabrina T. Howell, Mark J. Johnson, and Siena Matsumoto, Modernizing Access to Credit for Younger Entrepreneurs: From FICO to Cash Flow, NBER Working Paper 33367, January 2025, revised June 2026. This is a working paper, not research published in a peer-reviewed journal.
3. Jeremy C. Stein, Information Production and Capital Allocation: Decentralized versus Hierarchical Firms, *The Journal of Finance*, October 2002.
4. FinRegLab, Advancing the Credit Ecosystem: Machine Learning & Cash Flow Data in Consumer Underwriting, July 2025.

5. JPMorgan Chase Institute, Facing Uncertainty: Small Business Cash Flow Patterns in 25 US Cities.
6. JPMorgan Chase Institute, Paychecks, Paydays, and the Online Platform Economy.
7. US Consumer Financial Protection Bureau, Regulation B §1002.9: Notifications; withdrawn guidance, including the withdrawal of Circulars 2022-03 and 2023-03 on May 12, 2025.
8. FinRegLab, Machine Learning Explainability & Fairness: Insights from Consumer Lending, updated July 2023.
9. Saudi Central Bank, Open Banking in Saudi Arabia.