Archipel: Journal of Indonesian Interdisciplinary Studies
E-ISSN : 3124-6621
Vol. (2) No. (3) November 2026
https://ojsap.com/index.php/archipel
DOI : https://doi.org/10.65739/archipel.v2i3.103

From Digital Transactions to Risk Scores: A Sustainability Assessment Model for MSMEs in the Era of the Platform Economy

Adrian ED Serang
Perbanas Institute, Jakarta, Indonesia

adrian@perbanas.id*; adrian.edserang@gmail.com

Article Information:
Received October 12, 2026
Revised October 28, 2026
Accepted November 16, 2026
Published November 28, 2026

Keywords: *MSMEs; platform economy; alternative credit scoring; sustainability assessment; explainable risk model*

Abstract

Micro, small, and medium-sized enterprises (MSMEs) increasingly operate through e-commerce marketplaces, digital wallets, and QR-based payment systems, which generate continuous transaction records. Credit and sustainability assessments, however, still rely largely on collateral, audited statements, and static questionnaires that most MSMEs cannot provide. This study develops a Digital Transaction-Based Sustainability Risk Score (DT-SRS), a model that converts platform transaction data into a forward-looking assessment of MSME sustainability. Using a design science research approach, the study combined a PRISMA-guided systematic literature review of 2021–2026 publications with indicator construction, hybrid weighting, and an illustrative scoring exercise on six hypothetical MSME archetypes. The model comprises four dimensions: transaction performance, financial behavior and liquidity, platform dependence and exposure, and sustainability and resilience practices. These dimensions are operationalized through 16 indicators and aggregated into a 0–100 score with five risk bands. The illustrative results show that firms with similar revenues can fall into different risk bands once revenue volatility, channel concentration, and platform dependence are considered. Platform dependence is the dimension that conventional scoring most often overlooks. The study contributes a theoretically grounded and explainable framework that links credit risk and sustainability, and it offers practical guidance for lenders, platforms, and regulators seeking inclusive and responsible MSME financing.

INTRODUCTION

Micro, small, and medium-sized enterprises (MSMEs) are the backbone of most emerging economies, yet they remain the segment least served by formal finance. The core problem is informational. Lenders cannot easily observe the quality, stability, and prospects of small firms, which often lack audited financial statements, credit histories, and pledgeable assets. For decades, banks addressed this opacity by demanding collateral, typically land or the owner's home. This practice systematically excludes young, asset-light, and women-led enterprises, which make up a large share of the MSME population in developing countries.

This is an open access article distributed under the terms of the
Creative Commons Attribution 4.0 International License (CC BY 4.0).
Copyright © 2026 The Author(s). Published by Archipel: Journal of Indonesian Interdisciplinary Studies

The resulting financing gap constrains investment, limits employment growth, and leaves many MSMEs vulnerable to shocks.

The rise of the platform economy has begun to change the informational landscape. E-commerce marketplaces, ride-hailing and food-delivery ecosystems, digital wallets, and QR-based payment systems now mediate a growing share of MSME sales and purchases. Every order, payment, refund, rating, and settlement leaves a digital trace. Gawer (2022) argues that digital platforms and their ecosystems have become the dominant organizational form of the digital age, coordinating value creation among millions of independent complementors. For MSMEs, joining these ecosystems offers access to customers, logistics, and payment infrastructure at low entry cost. As a by-product, it also produces a continuous, high-frequency record of business activity that traditional financial statements cannot match.

The COVID-19 pandemic accelerated this shift. Fu and Mishra (2022) document a sizeable increase in finance-app adoption during lockdowns, with BigTech and newer fintech providers outperforming incumbents over time. The Global Findex 2021 shows that account ownership worldwide rose from 51 percent in 2011 to 76 percent in 2021, driven substantially by mobile money and digital payments in developing economies (Demirgüç-Kunt et al., 2022). Ozili (2021) notes that financial inclusion is now shaped strongly by financial innovation, financial literacy, and regulatory frameworks, and that digital channels carry both inclusion benefits and new systemic risks. In short, MSMEs are increasingly digital in how they trade and pay, even when they remain informal in how they keep accounts.

Financial intermediaries have responded by experimenting with alternative data. Evidence from China shows that big tech lenders can rely on transaction volumes and network scores rather than collateral. Gambacorta et al. (2023) find that big tech credit to more than two million Chinese firms reacts strongly to firm characteristics such as transaction volume, but hardly at all to local house prices. A related study shows that machine learning models built on non-traditional data predict losses and defaults better than traditional models, particularly during a credit-supply shock (Gambacorta et al., 2024). In the United States, finance companies and fintech lenders expanded small-business lending considerably after the 2008 crisis (Gopal & Schnabl, 2022), and fintech lenders reached underserved firms during the Paycheck Protection Program (Erel & Liebersohn, 2022). Djeundje et al. (2021) further demonstrate that alternative data can enhance credit scoring for thin-file borrowers.

In parallel, the methodological literature on credit scoring has advanced rapidly. Benchmark studies compare logistic regression with ensemble and deep learning methods. Gunnarsson et al. (2021) conclude that gradient-boosting methods such as XGBoost generally outperform deep learning architectures for credit scoring, while Moscato et al. (2021) benchmark several machine learning approaches on social lending data. Because regulators and borrowers need to understand why a score is assigned, explainable artificial intelligence has become central. Bussmann et al. (2021) combine Shapley values with correlation networks to explain the credit risk of small and medium firms on peer-to-peer platforms, and Chen et al. (2024) examine interpretable machine learning for imbalanced credit datasets. For micro and rural borrowers, Ampountolas et al. (2021) and Kumar et al. (2021) show that machine learning can support digital credit scoring when conventional records are scarce.

Despite this progress, three gaps remain. First, most alternative credit scoring research is designed to predict default within a short horizon, typically 12 months. It answers whether a firm will repay a specific loan, not whether the business is sustainable. Sustainability, in the MSME literature, is a broader concept that combines economic viability with social and environmental practices. Malesios et al. (2021) show that SMEs tend to prioritize economic performance over environmental and social dimensions, and they call for holistic frameworks to measure sustainability performance. Yet sustainability assessment frameworks for SMEs usually rely on surveys and self-reported practices, not on behavioral data. The two literatures, alternative credit scoring and SME sustainability

assessment, have developed largely in isolation.

Second, existing models treat platform data as a neutral input and overlook the risk that the platform itself creates. Cutolo and Kenney (2021) describe sellers on digital platforms as "platform-dependent entrepreneurs" whose existence is shaped by the platform's algorithms, fee structures, and rules. A change in search ranking, commission rates, or account policy can sharply reduce a seller's revenue overnight. A firm with strong sales concentrated on a single marketplace may therefore look creditworthy under a transaction-based score while being structurally fragile. To our knowledge, no published scoring model treats platform dependence as an explicit risk dimension.

Third, the fairness and governance implications of data-driven scoring are rarely built into model design. Fuster et al. (2022) show that machine learning can increase disparities between groups of borrowers even when it improves average predictive accuracy. For MSMEs, similar concerns arise for firms in rural areas, informal sectors, or women-led businesses, whose digital footprints may be thinner or differently shaped. Models that ignore these issues risk reproducing exclusion through new means. Kraus et al. (2022) observe that digital transformation research in business and management is still fragmented and calls for integrative frameworks that connect technology to organizational and societal outcomes.

This study addresses these gaps by asking one central question: how can digital transaction data from the platform economy be transformed into a transparent, forward-looking assessment of MSME sustainability risk? It pursues three objectives. The first is to identify, through a systematic literature review, the transaction-based and behavioral indicators that recent research links to MSME creditworthiness and sustainability. The second is to construct an integrated scoring model that aggregates these indicators into a single interpretable score with defined risk bands. The third is to demonstrate the model's behavior on illustrative MSME archetypes and to derive implications for lenders, platforms, and policymakers.

The novelty of this study lies in four contributions. First, it proposes the Digital Transaction-Based Sustainability Risk Score (DT-SRS), which shifts the unit of assessment from loan default to business sustainability. The model combines transaction performance and financial behavior with sustainability and resilience practices in one framework. Second, it introduces platform dependence and exposure as a distinct risk dimension, operationalized through channel concentration, platform-policy exposure, and account-health indicators. This dimension translates the theory of platform-dependent entrepreneurship into measurable credit-relevant variables. Third, it embeds explainability and fairness at the design stage through reason codes, monotonic constraints, and a thin-file pathway, rather than treating them as after-the-fact checks. Fourth, it offers a modular blueprint that lenders and platforms in emerging economies such as Indonesia can calibrate with their own data, bridging academic research and implementation.

The remainder of this article is organized as follows. The Methods section describes the design science research method, including the literature review protocol, indicator construction, weighting, and illustrative scoring. The Results and Discussion section presents and discusses the results, namely the literature synthesis, the DT-SRS framework, the scoring procedure, and its application to six MSME archetypes. The Conclusions section closes the article with implications, limitations, and directions for future research.

METHODS

Research design

This study adopts a design science research (DSR) approach, which is suited to research that builds and evaluates an artifact intended to solve a practical problem. The artifact here is the DT-SRS model. The research proceeded in five phases: problem identification, systematic literature review, model design, demonstration, and evaluation (Figure 1). Evaluation results were fed back into the design phase to refine weights and decision rules.

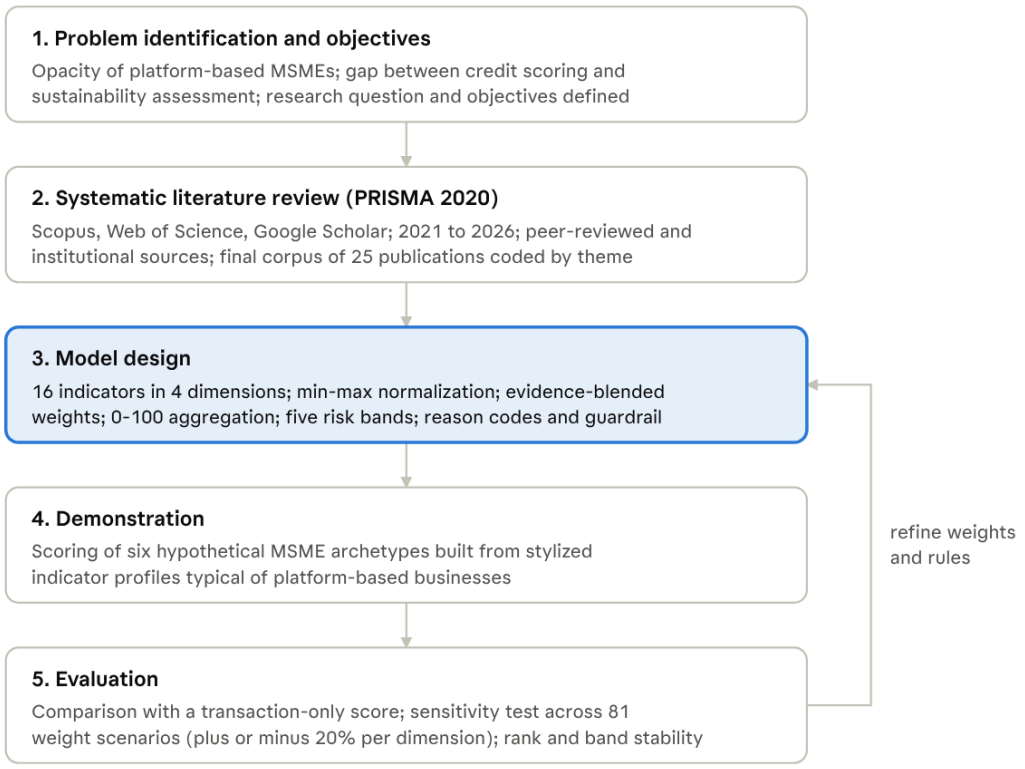


Figure 1. Research design of the DT-SRS model
Source: Author's design (2026)

Figure 1 shows that the model design phase is the core of the study, informed by the literature review and refined by the evaluation results.

Systematic literature review

The review followed the PRISMA 2020 reporting guideline (Page et al., 2021). Searches were run in Scopus, Web of Science, and Google Scholar for publications from January 2021 to 2026. The search string combined three blocks: ("MSME" OR "SME" OR "small business") AND ("alternative data" OR "digital footprint" OR "transaction data" OR "platform" OR "fintech") AND ("credit scoring" OR "credit risk" OR "sustainability performance" OR "financial inclusion"). Institutional reports from the World Bank and the OECD were added because they provide authoritative data on financial inclusion and SME conditions (Demirgüç-Kunt et al., 2022; OECD, 2023).

Studies were included if they (a) were published in 2021 or later, (b) had an active DOI, (c) addressed alternative data, credit risk modeling, platform-dependent business, or SME sustainability, and (d) reported empirical evidence, a validated method, or a well-developed conceptual framework. Studies on consumer-only lending without transferable insights, and non-peer-reviewed opinion pieces, were excluded. After title, abstract, and full-text screening, the final corpus comprised 25 publications. Each was coded for data source, analytical method, the indicators it used or proposed, and its main finding.

Indicator construction and normalization

Indicators were derived from the coded corpus and grouped into four dimensions: transaction performance (TP), financial behavior and liquidity (FB), platform dependence and exposure (PD), and sustainability and resilience practices (SR). Each dimension contains four indicators, giving 16 indicators in total. All indicators are computable from data that platforms, payment providers, or MSMEs themselves can supply, such as order logs, settlement records, wallet balances, and account-health reports.

Raw values were rescaled to the 0–1 interval with min–max normalization. For indicators where a higher value indicates lower risk, the formula is:

$$\tilde{x}_{ij} = \frac{x_{ij} - \min_j}{\max_j - \min_j} \quad (1)$$

For indicators where a higher value indicates higher risk, such as revenue volatility, refund rate, channel concentration, and fee burden, the normalized value is 1 minus the expression above. In operational use, the minimum and maximum should be taken as the 5th and 95th percentiles of a reference population to limit the influence of outliers.

Weighting and aggregation

Weights were set by a hybrid rule that blends literature evidence with equal weighting. For each indicator j, an evidence score e_j was computed as the sum of evidence strengths of the studies supporting it, where conceptual support scores 1, empirical support scores 2, and large-sample empirical support scores 3. The final weight is:

$$w_j = 0.5 \cdot \frac{e_j}{\sum_{j=1}^{16} e_j} + 0.5 \cdot \frac{1}{16} \quad (2)$$

The 50:50 blend prevents well-studied financial indicators from crowding out newer but theoretically important ones, such as platform dependence and environmental practices. The composite score for MSME i is:

$$\mathrm{DT-SRS}_i = 100 \sum_{j=1}^{16} w_j \tilde{x}_{ij} \quad (3)$$

Dimension sub-scores are computed in the same way within each dimension and rescaled to 0–100. Scores are mapped to five bands: A, very low risk (80–100); B, low risk (65–79.9); C, moderate risk (50–64.9); D, high risk (35–49.9); and E, very high risk (below 35).

Two governance rules complete the model. First, a dependence guardrail caps the band at C whenever the PD sub-score falls below 40, regardless of the composite score. Second, each score is accompanied by three reason codes, which are the indicators with the largest weighted shortfall $w_j(1 - \tilde{x}_{ij})$. This follows the explainable credit-risk literature, which stresses that regulated users must understand the drivers of each score (Bussmann et al., 2021; Giudici & Raffinetti, 2021).

Demonstration and evaluation

Because the purpose is to design and explain the model, not to estimate it on a proprietary portfolio, the demonstration uses six hypothetical MSME archetypes. Their normalized indicator profiles were set to reflect business types commonly described in the platform-economy and Indonesian MSME literature, such as single-marketplace resellers, delivery-app merchants, and QRIS-based neighborhood shops (Cutolo & Kenney, 2021; Wahyudin et al., 2022). The archetypes are illustrative and do not represent real firms.

Evaluation had two parts. First, each archetype's DT-SRS was compared with a transaction-only score that uses the TP and FB dimensions alone, representing a typical alternative-data scorecard. Second, robustness was tested by varying each dimension's weight to 80, 100, and 120 percent of its base value in all combinations ($3^4 = 81$ scenarios), renormalizing the weights, and recording Spearman rank correlations and band changes against the base case. All calculations were run in Python.

RESULTS AND DISCUSSION

Synthesis of the literature

The review shows that recent research has made strong progress in using non-

traditional data to assess small firms. It also shows that this progress is uneven across the dimensions that matter for long-term business sustainability. Table 1 groups the 25 publications into six themes and states what each theme contributes to the model.

Table 1. Thematic synthesis of the reviewed literature and implications for model design

Theme	Key studies	Main finding	Implication for DT-SRS
Platform and alternative data in SME credit	Gambacorta et al. (2023, 2024); Djeundje et al. (2021); Roa et al. (2021); Stevenson et al. (2021)	Transaction volumes, app behavior, and loan-officer text carry predictive information beyond bureau data and collateral	Transaction performance and financial behavior form the core dimensions (TP, FB)
Machine learning methods for scoring	Gunnarsson et al. (2021); Moscato et al. (2021); Ampountolas et al. (2021); Kumar et al. (2021)	Ensemble methods such as XGBoost generally outperform deep learning; machine learning supports micro and rural credit	The scorecard can later be re-estimated with gradient boosting under monotonic constraints
Explainability	Bussmann et al. (2021); Giudici & Raffinetti (2021); Chen et al. (2024)	Shapley-based methods explain credit scores, but explanations become less stable with class imbalance	Reason codes are built into the output; transparent linear aggregation is used as a baseline
Fintech, inclusion, and crisis adoption	Erel & Liebersohn (2022); Gopal & Schnabl (2022); Fu & Mishra (2022); Ozili (2021); Demirgüç-Kunt et al. (2022)	Fintech expands credit supply to underserved firms; digital adoption accelerated during COVID-19	Digital traces are now widespread enough to support population-level scoring
Platform dependence	Cutolo & Kenney (2021); Gawer (2022)	Platform sellers face power asymmetries and risks from algorithm, fee, and policy changes	Platform dependence becomes a separate risk dimension (PD)
SME sustainability and fairness	Malesios et al. (2021); OECD (2023); Kraus et al. (2022); Wahyudin et al. (2022); Fuster et al. (2022)	SMEs underweight social and environmental practices; ML can widen group disparities; digital payments improve MSME sales	Sustainability and resilience practices form a dimension (SR); fairness safeguards are built into design

Source: Author's synthesis of the reviewed literature (2026)

Three patterns stand out. First, the evidence base for transaction and repayment indicators is large and empirical, often drawing on millions of observations (Gambacorta et al., 2023; Fuster et al., 2022). Second, platform dependence is well theorized but rarely measured: Cutolo and Kenney (2021) offer a strong conceptual account, yet none of the reviewed scoring studies operationalizes it. Third, sustainability research on SMEs relies mainly on surveys and case studies (Malesios et al., 2021). It has not yet been connected to the transaction data that platforms already hold. These patterns justify both the four-dimension structure and the blended weighting rule, which keeps the less-studied dimensions from being marginalized.

The DT-SRS indicator framework

Table 2 presents the final framework. Each indicator is defined so that it can be computed from routine platform, payment, or bookkeeping records, and each carries an explicit risk direction.

Table 2. DT-SRS dimensions, indicators, measurement, and weights

Code	Indicator	Measurement	Direction	Evidence score	Weight
TP1	Sales growth	12-month growth in gross merchandise value (GMV)	+	8	0.077
TP2	Revenue stability	Coefficient of variation of monthly sales	–	9	0.082
TP3	Transaction frequency	Share of active trading days per month	+	7	0.071

Code	Indicator	Measurement	Direction	Evidence score	Weight
TP4	Repeat-customer rate	Share of orders from returning buyers	+	4	0.054
FB1	Net cash-flow ratio	Total inflows divided by total outflows over 6 months	+	9	0.082
FB2	Cash buffer	Average balance divided by average daily outflow, in days	+	5	0.060
FB3	Payment punctuality	Share of bills, supplier invoices, and instalments paid on time	+	10	0.088
FB4	Refund and chargeback rate	Refunds and chargebacks divided by orders	−	3	0.048
PD1	Channel concentration	Herfindahl index of sales across marketplaces, social commerce, and offline QR payments	−	4	0.054
PD2	Platform tenure	Months of continuous active seller account	+	5	0.060
PD3	Account health	Seller rating, policy-violation and penalty record	+	4	0.054
PD4	Fee burden	Platform commissions, advertising, and service fees divided by GMV	−	2	0.043
SR1	Formalization and digital bookkeeping	Use of a digital ledger, business registration, and tax identification	+	5	0.060
SR2	Supplier concentration	Herfindahl index of purchases across suppliers	−	3	0.048
SR3	Environmental and social practices	Eco-packaging, certifications, local or women employment signals	+	5	0.060
SR4	Shock-recovery time	Months needed to regain 90% of pre-shock monthly sales	−	5	0.060

Note. Direction "+" means higher values lower risk; "−" means higher values raise risk and are inverted during normalization. Dimension weights are TP 0.284, FB 0.278, PD 0.210, and SR 0.227; indicator weights may not sum exactly because of rounding.

Source: Author's compilation (2026)

The weighting results reflect the state of the evidence. Payment punctuality (FB3) and revenue stability (TP2) receive the highest weights because they are supported by large-sample studies of both fintech and bank lending. The PD and SR dimensions receive smaller but still substantial weights, together about 44 percent of the total. Under pure evidence weighting, their combined share would be considerably lower, which would reproduce the blind spot identified in the literature synthesis.

Two design choices deserve comment. Channel concentration (PD1) treats offline QR payments as a separate channel. This matters in Indonesia, where QRIS adoption has been linked to better MSME sales performance (Wahyudin et al., 2022), and where many sellers combine marketplace, social-commerce, and in-store sales. Shock-recovery time (SR4) uses the firm's own history as a natural stress test. A firm that recovered quickly after a past disruption, such as the pandemic, provides direct evidence of resilience that no static ratio can capture.

Illustrative scoring of six MSME archetypes

Table 3 reports the dimension scores, composite DT-SRS, risk band, and reason codes for the six hypothetical archetypes. It also shows the score each archetype would receive from a transaction-only model using the TP and FB dimensions.

Table 3. Illustrative DT-SRS results for six hypothetical MSME archetypes

Archetype	TP	FB	PD	SR	DT-SRS (band)	Transaction-only score (band)	Top reason codes
A1 Single-marketplace fashion reseller	74.1	65.9	35.1	32.6	54.2 (C)	70.1 (B)	PD1, SR3, TP2
A2 Multichannel food producer	71.4	73.5	75.5	67.4	71.9 (B)	72.4 (B)	TP1, SR3, FB1
A3 QRIS-based neighborhood grocer	71.5	71.9	78.3	50.6	68.3 (B)	71.7 (B)	TP1, SR3, SR1
A4 Seasonal craft exporter	39.8	51.9	65.1	51.6	51.2 (C)	45.8 (D)	TP2, TP3, FB1
A5 Delivery-app food merchant	76.0	57.4	29.8	33.4	51.5 (C)	66.8 (B)	PD1, SR3, FB2
A6 Eco-certified household products	58.5	70.9	73.9	78.9	69.8 (B)	64.6 (C)	TP1, FB1, TP3

Note. Archetypes are hypothetical and built from stylized normalized indicator values; they do not represent real firms. All scores are on a 0–100 scale.

Source: Author's calculation based on hypothetical archetypes (2026)

The central result is that the two models disagree for four of the six archetypes. The fashion reseller (A1) and the delivery-app merchant (A5) look attractive under a transaction-only lens. They have the highest transaction performance scores in the set, at 74.1 and 76.0, and both receive band B. Once platform dependence and sustainability practices are included, both fall to band C, with scores of 54.2 and 51.5. Their top reason code is channel concentration (PD1). In practical terms, almost all their revenue flows through one platform, so a ranking change, a commission increase, or an account suspension could remove most of their income at once. This is exactly the vulnerability that Cutolo and Kenney (2021) describe for platform-dependent entrepreneurs, now expressed as a measurable credit signal.

Both A1 and A5 also have PD sub-scores below 40, which activates the dependence guardrail. In this demonstration the guardrail does not change their band, because their composite scores already fall in band C. Its role is preventive: it stops a firm with very strong sales but extreme dependence from reaching band A or B through high transaction scores alone.

The opposite movement occurs for the seasonal craft exporter (A4) and the eco-certified household-products firm (A6). A4 is penalized by a transaction-only model for its volatile, seasonal revenue and falls to band D. The DT-SRS recognizes its diversified channels, solid account health, and moderate sustainability practices, and places it in band C. A6 has modest growth but the strongest sustainability profile in the set (SR = 78.9) and good platform diversification. It moves from band C to band B. These shifts show that the model does not simply penalize digital firms. It rewards resilience and diversification wherever they exist.

The multichannel food producer (A2) and the QRIS-based grocer (A3) keep band B under both models. Their reason codes still differ in useful ways. For A3, the codes point to low growth and weak formalization (SR1), suggesting that a lender or program could support the grocer through digital bookkeeping and business registration rather than through credit alone.

Robustness of the model

The sensitivity test produced 81 weight scenarios. Across all of them, the Spearman rank correlation between base-case and perturbed scores ranged from 0.829 to 1.000, with a mean of 0.957. Only 4.3 percent of the 486 firm-scenario combinations changed risk band, and all changes were to an adjacent band. The ranking of archetypes is therefore stable under reasonable disagreement about weights. This matters because weights in any expert or evidence-based scorecard involve judgment. A model whose conclusions reverse under small weight changes would be difficult for lenders and regulators to trust.

Discussion

The results support the study's main argument: transaction data are necessary but not sufficient for assessing MSME sustainability in the platform economy. The findings agree with evidence that transaction and behavioral data improve on collateral-based assessment (Gambacorta et al., 2023, 2024; Djeundje et al., 2021; Roa et al., 2021). They extend that evidence by showing that the same data, read differently, can reveal structural fragility. High sales recorded on a single platform are a sign of both current strength and concentrated exposure. A sustainability-oriented model has to read both signals.

Theoretically, the model links three literatures that have so far developed separately. It brings the information-asymmetry logic of alternative credit scoring together with the power-dependence view of platform entrepreneurship (Cutolo & Kenney, 2021; Gawer, 2022). It then adds the holistic view of SME sustainability performance (Malesios et al., 2021). The PD dimension is the main theoretical bridge. It treats the platform not as a neutral data source but as an actor whose decisions shape the borrower's risk.

For lenders, the DT-SRS offers a way to serve thin-file MSMEs while avoiding a predictable failure mode, namely over-lending to fast-growing but platform-captive sellers. The reason codes also support a more developmental relationship with borrowers. Instead of a simple rejection, a firm can be told which factors lowered its score, such as channel concentration or missing bookkeeping, and what it could change.

For platforms, the model suggests that sharing data with lenders creates obligations as well as benefits. Because platform policies directly affect PD scores, transparency about fee changes, ranking rules, and account suspensions becomes a matter of financial stability for sellers. For regulators and policymakers, the framework provides a template for responsible alternative-data scoring. It requires stated directions for each indicator, published weights, reason codes, and a thin-file pathway that allows a firm with limited data to be scored on the dimensions available rather than being rejected outright.

Fairness remains a central concern. Fuster et al. (2022) show that more flexible models can widen disparities between borrower groups. The DT-SRS reduces this risk through transparent linear aggregation and explicit monotonic directions. It does not remove the risk. Indicators such as platform tenure or repeat-customer rate may disadvantage newer, rural, or women-led businesses whose digital histories are shorter. Before deployment, the model should be audited for score differences across these groups. Where machine learning is later used to re-estimate weights, explanation stability should also be checked, given the evidence that explanations degrade with class imbalance (Chen et al., 2024).

Finally, the results should be read with the method in mind. The archetypes demonstrate how the model behaves; they do not test its predictive accuracy. The weights rest on a structured reading of the literature rather than on estimated default or survival outcomes. The model is therefore best understood as a validated design blueprint that lenders and platforms can calibrate with their own data.

CONCLUSIONS

This study developed the Digital Transaction-Based Sustainability Risk Score (DT-SRS), a model that turns the transaction data generated in the platform economy into a transparent assessment of MSME sustainability risk. Built through a design science approach and grounded in a systematic review of 2021–2026 literature, the model combines 16 indicators across four dimensions: transaction performance, financial behavior and liquidity, platform dependence and exposure, and sustainability and resilience practices. The illustrative scoring of six MSME archetypes shows that a transaction-only score and the DT-SRS place four of the six firms in different risk bands. Firms with strong sales concentrated on a single platform move to a higher-risk band, while diversified and resilient firms move to a lower-risk band. A sensitivity test across 81 weight scenarios confirms that these rankings are stable.

The study makes three contributions. Theoretically, it connects alternative credit scoring, platform-dependent entrepreneurship, and SME sustainability assessment in a single framework, with platform dependence as the main bridge. Methodologically, it offers an evidence-blended weighting rule and builds explainability and governance into the model through reason codes and a dependence guardrail. Practically, it gives lenders, platforms, and regulators in emerging economies such as Indonesia a blueprint for inclusive and responsible MSME financing.

The study has limitations that point to future research. The demonstration uses hypothetical archetypes, so the model's predictive validity has not yet been tested. Future studies should calibrate the DT-SRS on real platform and payment data, linking scores to observed outcomes such as default, business closure, and revenue recovery after shocks. Weights could then be re-estimated with interpretable machine learning under monotonic constraints and compared with the evidence-blended weights used here. Further work should also audit the model for fairness across firm size, location, sector, and owner gender, and should examine how data-sharing arrangements between platforms and lenders can protect MSMEs' interests.

REFERENCES

Ampountolas, A., Nyarko Nde, T., Date, P., & Constantinescu, C. (2021). A machine learning approach for micro-credit scoring. *Risks, 9*(3), 50. https://doi.org/10.3390/risks9030050

Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. *Computational Economics, 57*(1), 203–216. https://doi.org/10.1007/s10614-020-10042-0

Chen, Y., Calabrese, R., & Martin-Barragan, B. (2024). Interpretable machine learning for imbalanced credit scoring datasets. *European Journal of Operational Research, 312*(1), 357–372. https://doi.org/10.1016/j.ejor.2023.06.036

Cutolo, D., & Kenney, M. (2021). Platform-dependent entrepreneurs: Power asymmetries, risks, and strategies in the platform economy. *Academy of Management Perspectives, 35*(4), 584–605. https://doi.org/10.5465/amp.2019.0103

Demirgüç-Kunt, A., Klapper, L., Singer, D., & Ansar, S. (2022). *The Global Findex Database 2021: Financial inclusion, digital payments, and resilience in the age of COVID-19*. World Bank. https://doi.org/10.1596/978-1-4648-1897-4

Djeundje, V. B., Crook, J., Calabrese, R., & Hamid, M. (2021). Enhancing credit scoring with alternative data. *Expert Systems with Applications, 163*, 113766. https://doi.org/10.1016/j.eswa.2020.113766

Erel, I., & Liebersohn, J. (2022). Can FinTech reduce disparities in access to finance? Evidence from the Paycheck Protection Program. *Journal of Financial Economics, 146*(1), 90–118. https://doi.org/10.1016/j.jfineco.2022.05.004

Fu, J., & Mishra, M. (2022). Fintech in the time of COVID-19: Technological adoption during crises. *Journal of Financial Intermediation, 50*, 100945. https://doi.org/10.1016/j.jfi.2021.100945

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance, 77*(1), 5–47. https://doi.org/10.1111/jofi.13090

Gambacorta, L., Huang, Y., Li, Z., Qiu, H., & Chen, S. (2023). Data versus collateral. *Review of Finance, 27*(2), 369–398. https://doi.org/10.1093/rof/rfac022

Gambacorta, L., Huang, Y., Qiu, H., & Wang, J. (2024). How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. *Journal of Financial Stability, 73*, 101284. https://doi.org/10.1016/j.jfs.2024.101284

Gawer, A. (2022). Digital platforms and ecosystems: Remarks on the dominant organizational forms of the digital age. *Innovation: Organization & Management, 24*(1), 110–124. https://doi.org/10.1080/14479338.2021.1965888

Giudici, P., & Raffinetti, E. (2021). Shapley-Lorenz eXplainable artificial intelligence. *Expert*

Systems with Applications, 167, 114104. https://doi.org/10.1016/j.eswa.2020.114104
Gopal, M., & Schnabl, P. (2022). The rise of finance companies and FinTech lenders in small business lending. *The Review of Financial Studies, 35*(11), 4859–4901. https://doi.org/10.1093/rfs/hhac034
Gunnarsson, B. R., vanden Broucke, S., Baesens, B., Óskarsdóttir, M., & Lemahieu, W. (2021). Deep learning for credit scoring: Do or don't? *European Journal of Operational Research, 295*(1), 292–305. https://doi.org/10.1016/j.ejor.2021.03.006
Kraus, S., Durst, S., Ferreira, J. J., Veiga, P., Kailer, N., & Weinmann, A. (2022). Digital transformation in business and management research: An overview of the current status quo. *International Journal of Information Management, 63*, 102466. https://doi.org/10.1016/j.ijinfomgt.2021.102466
Kumar, A., Sharma, S., & Mahdavi, M. (2021). Machine learning (ML) technologies for digital credit scoring in rural finance: A literature review. *Risks, 9*(11), 192. https://doi.org/10.3390/risks9110192
Malesios, C., De, D., Moursellas, A., Dey, P. K., & Evangelinos, K. (2021). Sustainability performance analysis of small and medium sized enterprises: Criteria, methods and framework. *Socio-Economic Planning Sciences, 75*, 100993. https://doi.org/10.1016/j.seps.2020.100993
Moscato, V., Picariello, A., & Sperlí, G. (2021). A benchmark of machine learning approaches for credit score prediction. *Expert Systems with Applications, 165*, 113986. https://doi.org/10.1016/j.eswa.2020.113986
OECD. (2023). *OECD SME and entrepreneurship outlook 2023*. OECD Publishing. https://doi.org/10.1787/342b8564-en
Ozili, P. K. (2021). Financial inclusion research around the world: A review. *Forum for Social Economics, 50*(4), 457–479. https://doi.org/10.1080/07360932.2020.1715238
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71
Roa, L., Correa-Bahnsen, A., Suarez, G., Cortés-Tejada, F., Luque, M. A., & Bravo, C. (2021). Super-app behavioral patterns in credit risk models: Financial, statistical and regulatory implications. *Expert Systems with Applications, 169*, 114486. https://doi.org/10.1016/j.eswa.2020.114486
Stevenson, M., Mues, C., & Bravo, C. (2021). The value of text for small business default prediction: A deep learning approach. *European Journal of Operational Research, 295*(2), 758–771. https://doi.org/10.1016/j.ejor.2021.03.008
Wahyudin, N., Herlissha, N., Christianingrum, & Aldiesi, D. R. (2022). The utilization of e-commerce and QRIS as digital payment tools to improve sales performance through competitive advantage in MSME. *Journal of Consumer Sciences, 7*(2), 134–147. https://doi.org/10.29244/jcs.7.2.134-147