

AI-native credit intelligence on digital public infrastructure: evidence from India's account aggregator ecosystem

Malcolm Athaide

Independent Researcher, Mumbai, India

Fintech and
Digital
Accounting
Review

Received 30 March 2026
Revised 27 June 2026
Accepted 17 August 2026

Abstract

Purpose – This paper examines how India's Account Aggregator (AA) framework (269 million consents, 2.12 billion enabled accounts) can support an AI-native consumer credit intelligence platform. It develops testable hypotheses for expanding financial product eligibility to 300–450 million underserved adults across housing, personal lending, MSME credit and insurance.

Design/methodology/approach – The study combines institutional analysis of India's regulatory architecture with econometric specifications: probit models, difference-in-differences exploiting staggered AA rollout, panel IV estimation, regression discontinuity and heterogeneous treatment effects. The framework extends Rochet and Tirole's (2003) two-sided market model to DPI-augmented platforms. The paper is conceptual. It sets out these identification strategies, a five-scenario impact assessment and a platform cost model as a research agenda, and does not report estimated causal effects.

Findings – India's AA data layer, UPI transaction histories and Aadhaar verification enable credit scoring for populations invisible to bureaux. Scenario analysis across five cases projects incremental demand between \$180 billion and \$480 billion over a decade, with a genuinely incremental impact of \$108–216 billion after organic growth adjustment.

Originality/value – This study develops a formal model of a credit platform on government-provided digital infrastructure, showing DPI-augmented platforms face different pricing dynamics. It proposes India-specific algorithmic fairness criteria for caste, gender and geography, and treats regulatory constraints as binding design parameters.

Keywords AI credit scoring, Account aggregator, Digital public infrastructure, Financial inclusion, India stack, Fintech platform, Alternative data, Machine learning, Two-sided markets, Algorithmic fairness

Paper type Conceptual paper

Plain language summary

Over 700 million Indian adults cannot access formal loans, insurance or investments because they lack credit records. This paper proposes using India's Account Aggregator (AA) system – which lets people share bank data securely – to build an AI platform creating credit scores for those the existing system ignores. The study estimates that this could bring financial products to 300–450 million more adults and generate \$108–216 billion in genuinely new economic activity over ten years. The paper addresses fairness concerns, noting AI scoring must account for caste, gender and geography.

Abbreviations

AA	Account Aggregator
AMFI	Association of Mutual Funds in India
AUC	Area Under the Curve

© Malcolm Athaide. Published in *Fintech and Digital Accounting Review*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/>

Funding: No external funding was received for this study.

Conflict of interest: The author declares no conflict of interest.



Fintech and Digital Accounting Review
Emerald Publishing Limited
2977-7631
DOI 10.1108/FDAR-03-2026-0027

FDAR	CapEx	Capital Expenditure
	CIBIL	Credit Information Bureau (India) Limited
	DiD	Difference-in-Differences
	DPD	Days Past Due
	DPDP	Digital Personal Data Protection (Act/Rules)
	DPI	Digital Public Infrastructure
	DPIA	Data Protection Impact Assessment
	DPO	Data Protection Officer
	FIP	Financial Information Provider
	FREE-AI	Framework for Responsible and Ethical Enablement of AI
	GST	Goods and Services Tax
	HTE	Heterogeneous Treatment Effects
	IRDAI	Insurance Regulatory and Development Authority of India
	IV	Instrumental Variables
	KFS	Key Fact Statement
	ML	Machine Learning
	MSME	Micro, Small and Medium Enterprises
	NPA	Non-Performing Asset
	OBC	Other Backward Classes
	OpEx	Operating Expenditure
	PLFS	Periodic Labour Force Survey
	RBI	Reserve Bank of India
	RDD	Regression Discontinuity Design
	RERA	Real Estate (Regulation and Development) Act
	SC/ST	Scheduled Caste/Scheduled Tribe
	SEBI	Securities and Exchange Board of India
	SHAP	SHapley Additive exPlanations
	UPI	Unified Payments Interface

1. Introduction

India’s formal credit markets exhibit persistent information asymmetry: a large share of credit-eligible adults either lack a bureau score or carry thin files, limiting lenders’ ability to distinguish risk and producing rationing dynamics consistent with the work of [Stiglitz and Weiss \(1981\)](#). Of 1.036 billion credit-eligible adults, only 277 million (27%) participate in formal credit markets, leaving more than 700 million invisible to the four licensed bureaux ([TransUnion CIBIL, 2025](#)).

The consequences span multiple verticals. India’s mortgage-to-GDP ratio stands at 12.3% versus 65% in the United States. Insurance penetration reaches only 4% against a 7.4% global average. The MSME credit gap exceeds Rs 30 lakh crore, about \$360 billion ([SIDBI and CRISIL, 2025](#)). These figures reflect a set of reinforcing constraints, including documentation and cash-flow visibility gaps among informal workers (who often cannot evidence stable income using conventional documents), limited credit information sharing and high customer acquisition costs in small-ticket lending. As [Figure 1](#) illustrates, India’s credit pyramid leaves the majority outside the formal system.

Two developments altered the structural conditions. Credit Karma (acquired by Intuit for \$7.1 billion; [Intuit, 2020](#)) attracted 130 million members through free credit transparency. Upstart demonstrated that AI underwriting with non-traditional variables approves 27% more borrowers at 16% lower rates ([Upstart, 2025](#)). Both, however, operate within the mature US bureau ecosystem.

India’s digital public infrastructure (DPI) creates conditions absent in the US context. The AA framework (269 million consents, 2.12 billion enabled accounts; [Sahamati, 2025](#)),

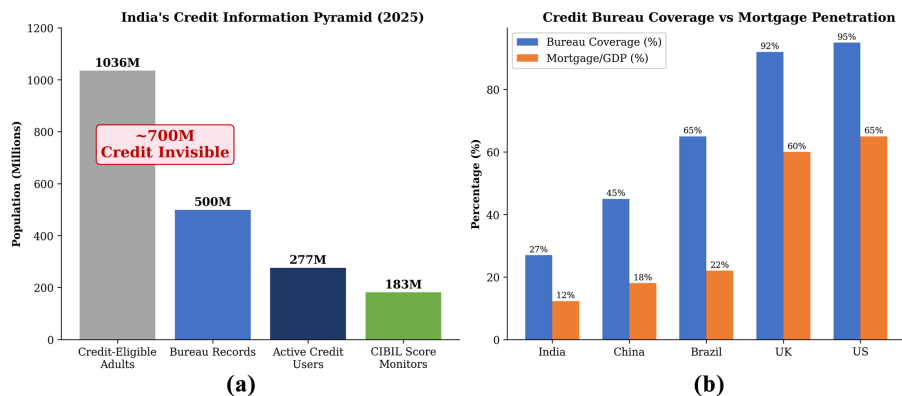


Figure 1. India's credit information gap: (a) credit pyramid showing 700 million credit-invisible adults; (b) cross-country bureau coverage versus mortgage penetration. *Source:* Author created using data from TransUnion CIBIL (2025) and World Bank

Aadhaar (approximately 1.44 billion enrolments; UIDAI, 2026) and UPI (228.3 billion annual transactions; RBI, 2025a,b,c) form a consent-based data layer at an unusually large scale relative to most open-finance regimes (Sahamati, 2025). Figure 2 traces the growth of these components, showing AA consent and account trajectories alongside a scale comparison indexed to Aadhaar.

This paper makes three contributions. First, it develops a formal extension of Rochet and Tirole's (2003) two-sided market framework to platforms on government-provided DPI, addressing gaps in Armstrong (2006) and Weyl (2010). Second, it formulates five testable hypotheses exploiting the staggered rollout of AA Financial Information Providers. Third, it proposes India-specific algorithmic fairness criteria incorporating caste, gender and geographic dimensions, addressing concerns in Sambasivan, Arnesen, Hutchinson, Doshi, and Prabhakaran (2021) and Fuster, Goldsmith-Pinkham, Ramadorai, and Walther (2022). This article contributes a model and research design with falsifiable hypotheses; empirical estimation is proposed as future work contingent on partner data access. The paper is therefore

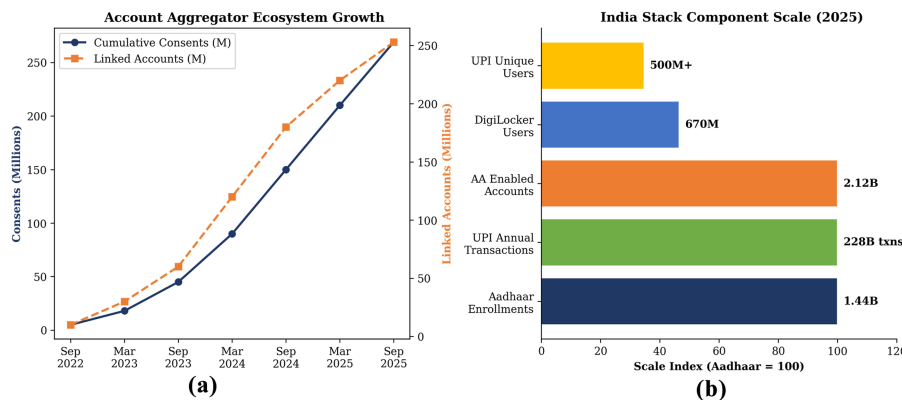


Figure 2. India Stack growth: (a) Account Aggregator consent and account trajectories; (b) component scale comparison indexed to Aadhaar. *Source:* Author created using data from Sahamati (2025), UIDAI (2026) and RBI (2025c)

best read as a conceptual and research-design contribution. The econometric specifications in [Section 4](#) set out how the hypotheses can be tested once loan-level data become available, and the projections in [Section 5](#) are scenario analyses that depend on the stated assumptions; they are not estimated treatment effects.

The paper proceeds as follows. [Section 2](#) reviews relevant literature. [Section 3](#) describes the institutional setting. [Section 4](#) presents the theoretical model and methods. [Section 5](#) provides scenario analysis. [Section 6](#) discusses implications. [Section 7](#) concludes.

2. Literature review

2.1 *Information asymmetry and credit rationing*

[Stiglitz and Weiss \(1981\)](#) established that adverse selection under imperfect information leads to credit rationing. [Jappelli and Pagano \(2002\)](#) showed credit information sharing increases lending and reduces defaults cross-nationally. [He, Huang, and Zhou \(2023\)](#) demonstrated that borrower data ownership under open banking produces complex welfare effects depending on portability costs.

2.2 *Machine learning in credit markets*

[Berg, Burg, Gombović, and Puri \(2020\)](#) found digital footprint variables match bureau scores (AUC 69.6% versus 68.3%). [Fuster et al. \(2022\)](#) showed ML models improve prediction but create distributional consequences. [Gambacorta, Huang, Qiu, and Wang \(2024\)](#) demonstrated ML superiority for thin-file borrowers. [Blattner and Nelson \(2024\)](#) found ML reduces score noise affecting underserved populations. [Di Maggio, Ratnadiwakara, and Carmichael \(2022\)](#) identified “invisible primes” misclassified by traditional scoring.

2.3 *Fintech lending and platform economics*

[Buchak, Matvos, Piskorski, and Seru \(2018\)](#) attributed 30% of shadow bank mortgage growth to technology. [Fuster, Plosser, Schnabl, and Vickery \(2019\)](#) showed fintech processes mortgages 20% faster without higher defaults. [Jagtiani and Lemieux \(2019\)](#) confirmed alternative data capture dimensions invisible to traditional scoring. [Rochet and Tirole \(2003\)](#) established the two-sided market framework; [Armstrong \(2006\)](#) extended it to competing platforms; [Weyl \(2010\)](#) provided price theory for heterogeneous agents. [Bartlett, Morse, Stanton, and Wallace \(2022\)](#) documented a 7.9 basis point minority pricing penalty.

2.4 *Digital infrastructure and financial inclusion*

[D'Silva, Filkova, Packer, and Tiwari \(2019\)](#) estimated India would not have reached 80% account ownership until 2055 without India Stack. [Jack and Suri \(2014\)](#) showed M-Pesa transformed risk-sharing behaviour. [Suri and Jack \(2016\)](#) confirmed poverty-reduction effects concentrated among female-headed households. [Philippon \(2016\)](#) documented the persistent 2% intermediation cost. [Frost \(2020\)](#) found fintech adoption highest where unbanked populations persist alongside high GDP. [Nalluri and Chen \(2024\)](#) identified regulatory uncertainty as the primary adoption barrier. [Sahay et al. \(2015\)](#) cautioned that inclusion raises stability risks without proportionate supervision. [Singh, Kumari, Sharma, and Malhotra \(2021\)](#) assessed Jan Dhan Yojana's mixed impact.

2.5 *Algorithmic fairness*

[Sambasivan et al. \(2021\)](#) argued Western fairness definitions fail in India. [Kozodoi, Jacob, and Lessmann \(2022\)](#) quantified the profit-fairness trade-off. [Hurlin, Pérignon, and Saurin \(2026\)](#) formalised fairness tests. [Field, Pande, Rigol, Schaner, and Troyer Moore \(2021\)](#) demonstrated that strengthening women's financial control shifts labour supply. [Bhukta \(2024\)](#) showed SC households benefit from branch expansion for savings but less for credit. [Dabla-Norris, Ji,](#)

Townsend, and Unsal (2015) modelled financial inclusion's GDP impacts. Mbiti and Weil (2015) documented selection effects in M-Pesa adoption.

These literatures have largely developed in isolation, and the paper sits at their intersection. The credit-rationing tradition (Stiglitz & Weiss, 1981; Jappelli & Pagano, 2002) explains why information gaps exclude borrowers, but it predates the data environment that machine learning now exploits. The machine-learning credit literature (Berg *et al.*, 2020; Fuster *et al.*, 2022; Gambacorta *et al.*, 2024) establishes predictive gains for thin-file borrowers, yet almost all of it sits in mature private-bureau economies and says little about the infrastructure that makes such data available at scale. The platform-economics tradition (Rochet & Tirole, 2003; Armstrong, 2006; Weyl, 2010) assumes private provision of that infrastructure, which publicly provisioned DPI does not require. The fairness literature (Sambasivan *et al.*, 2021; Hurlin *et al.*, 2026) shows that fairness criteria depend on context, but operational measures for caste- and geography-structured exclusion remain thin. The question that motivates the model and hypotheses below is whether the predictive gains found in private-bureau settings carry over to a publicly provisioned, consent-based data environment and on what fairness and cost terms they remain welfare-enhancing.

3. Institutional setting and data

3.1 Credit bureau ecosystem

India's four licensed bureaux operate under the Credit Information Companies (Regulation) Act, 2005. TransUnion CIBIL (2025) reports 183 million active score monitors. Over 700 million adults have no bureau record. Informal-sector workers are not inherently risky; they are undocumented.

3.2 Account Aggregator framework

The AA framework, operational since September 2021, covers bank accounts, mutual funds, insurance, GST returns and pension data. By September 2025, the ecosystem processed 269 million consents across 780+ institutions (Sahamati, 2025). AA-enabled personal loans constituted 10.52% of originations in H1 FY2025. FIP count expanded from 155 to 580 between August and December 2024, creating the staggered rollout exploited for identification.

3.3 Multi-vertical financial gap

Table 1 summarises the gap across seven verticals. Housing finance shows mortgage-to-GDP at 12.3%. The MSME credit gap exceeds Rs 30 lakh crore, about \$360 billion (SIDBI and CRISIL, 2025). Insurance penetration (4%) stands at roughly half the global average (IRDAI, 2025).

3.4 Regulatory constraints

Three frameworks shape the operating environment. The RBI Digital Lending Directions (RBI, 2025a) mandate Key Fact Statement disclosure and cooling-off periods. The FREE-AI Committee Report (RBI, 2025b) introduces seven governing principles. The Digital Personal Data Protection Act (Government of India, 2023) and DPDP Rules (Government of India, 2025) impose consent-first processing with penalties up to ₹250 crore. Table 7 (Section 6) maps these constraints to design requirements.

4. Theoretical model and methods

4.1 DPI-augmented platform model

Rochet and Tirole (2003) modelled platforms setting prices with cross-subsidisation incentives. Armstrong (2006) extended this to competing platforms, and Weyl (2010)

Table 1. India’s multi-vertical financial services gap (2025–2026)

Vertical	Market size	Penetration gap	Credit-invisible impact
Housing finance	\$488B outstanding	Mortgage/GDP 12.3% vs 65% US	450M workers lack income documentation
Personal loans	\$171B outstanding	25% YoY growth; bureau-concentrated	Informal borrowers pay 36–48% vs 10–18%
MSME credit	\$530B gap (IFC)	14% of 63M MSMEs have formal access	Micro-enterprises near-zero formal credit
Auto loans	\$70B outstanding	70% new-car; <20% used-vehicle	Gig workers cannot document income
Insurance	\$131B premiums	Life 3.2%, Non-life 1.0% vs 7.4%	Insurers cannot risk-price without scores
Investments	\$780B + MF AUM	4% household MF participation	Product matching absent for bottom 70%
Credit cards	\$34B outstanding	101M cards/1.4B people (7.2%)	Bureau score required; informal excluded

Source(s): Author compilation from RBI, IRDAI, SEBI, AMFI, [IFC \(2022\)](#) and [TransUnion CIBIL \(2025\)](#)

addressed agent heterogeneity. These frameworks assume platforms bear full infrastructure costs. India’s DPI removes this assumption. Consumer utility becomes $UC = \alpha C \cdot nL + \beta C \cdot s(x_i) + \gamma C \cdot \sum v_r v(x_i) - pC$. Lender utility is $UL = \alpha L \cdot nC \cdot q(\theta) - pL$. When DPI reduces marginal costs to near zero, $pC = 0$ becomes an equilibrium outcome of public infrastructure provision. Because the data and identity layers are publicly provisioned, the platform does not need to recover infrastructure costs through consumer prices, so the binding constraint moves from infrastructure recovery to lender-side monetisation and quality matching. The corner solution $pC = 0$ then holds because of public provision rather than cross-subsidy from the lender side, the mechanism in [Rochet and Tirole \(2003\)](#). This carries different welfare and entry implications, since incumbents hold no data-infrastructure advantage and entry depends instead on model quality and compliance capability. [Figure 3](#) sets out the full set of ten model specifications, covering platform utility, the AI credit-score formulation and the hypothesis-test equations.

4.2 Hypotheses

- H1. (Adoption):** A free AI credit platform achieves 15+% adoption among smartphone-enabled adults within 36 months. Instrument: staggered AA FIP onboarding. $H_0: \beta_1 = 0$.
- H2. (Credit expansion):** AI scoring expands eligibility by 40+% among informal-sector applicants. Specification (Probit with interaction). $H_0: \beta_5 = 0$.
- H3. (Default improvement):** AI-underwritten products show 20–30% lower 90+ DPD delinquency. Specification (DiD). $H_0: \beta_3 = 0$.
- H4. (Demand activation):** Higher platform penetration increases financial services demand. Specification (Panel IV). $H_0: \beta_1 = 0$.
- H5. (Welfare):** Inclusion effects concentrate among women-headed and SC/ST households. Heterogeneous treatment effects per [Chernozhukov, Demirer, Duflo, and Fernández-Val \(2025\)](#). $H_0: \beta_2 = \beta_3 = 0$.

[Table 2](#) summarises identification strategies.

Equation 1: Consumer Utility Function

$$U_C = \alpha_C \cdot n_C + \beta_C \cdot s(\mathbf{x}_C) + \gamma_C \cdot \sum_{v=1}^V r_v(\mathbf{x}_C) - \rho_C$$

Equation 2: AI-Enhanced Credit Score

$$CS_i^{AI} = f(B_i, A_i, T_i, G_i, \epsilon_i)$$

where B_i = Bureau data, A_i = AA data, T_i = UPI data, G_i = Geo-demographic features

Equation 3: H1 — Platform Adoption (Cross-Sectional)

$$Adopt_{d,t} = \alpha + \beta_1 \cdot CreditGap_d + \beta_2 \cdot SmartphonePen_d + \beta_3 \cdot UPIUsage_d + \beta_4 \cdot AAcoverage_d + \gamma' \mathbf{X}_d + \epsilon_{d,t}$$

Equation 4: H2 — Credit Expansion (Probit)

$$P(Eligible_j = 1) = \Phi(\alpha + \beta_1 CIBIL_i + \beta_2 AA_Score_i + \beta_3 UPI_Features_i + \beta_4 Informal_i + \beta_5 (AA_Score_i \times Informal_i) + \gamma' \mathbf{Z}_i)$$

Equation 5: H3 — Default Rate Improvement (DID)

$$Default_{i,t} = \alpha + \beta_1 \cdot AIScore_i + \beta_2 \cdot Post_t + \beta_3 \cdot (AIScore_i \times Post_t) + \delta' \mathbf{X}_{i,t} + \mu_i + \lambda_t + \epsilon_{i,t}$$

Equation 6: H4 — Demand Activation (Panel IV)

$$\ln(FinServices_{d,t}) = \alpha_d + \lambda_t + \beta_1 \cdot \ln(PlatformPen_{d,t}) + \gamma' \mathbf{W}_{d,t} + \epsilon_{d,t}$$

Equation 7: H5 — Welfare (Heterogeneous Treatment)

$$Fl_{h,t} = \alpha + \beta_1 Access_{h,t} + \beta_2 (Access_{h,t} \times Female_h) + \beta_3 (Access_{h,t} \times SC/ST_h) + \delta' \mathbf{V}_{h,t} + \mu_h + \lambda_t + \epsilon_{h,t}$$

Equation 8: Lender Utility Function

$$U_L = \alpha_L \cdot n_C \cdot q(\theta) - \rho_L$$

where $q(\theta)$ = lead quality, θ = AI model precision, ρ_L = per-lead fee

Equation 9: Insurance Loss Ratio (DID)

$$LossRatio_{p,t} = \alpha + \beta_1 AIPricing_p + \beta_2 Post_t + \beta_3 (AIPricing_p \times Post_t) + \gamma' \mathbf{W}_{p,t} + \epsilon_{p,t}$$

Equation 10: Macro Housing Demand Elasticity

$$\ln(MortgageVol_{d,t}) = \alpha_d + \lambda_t + \beta_1 \ln(CreditIndiaPen_{d,t}) + \beta_2 \ln(Income_{d,t}) + \beta_3 HousePrice_{d,t} + \beta_4 InterestRate_t + \epsilon_{d,t}$$

Figure 3. Econometric model specifications: ten equations covering platform utility, AI credit score formulation and hypothesis test specifications. *Source:* Author created

4.3 Identification and robustness

For [H1/H4](#), staggered FIP onboarding (155 to 580 institutions, August–December 2024) provides exogenous variation. For [H2/H3](#), regression discontinuity at CIBIL 630–650 with McCrary density tests and bandwidth sensitivity (50–150% of optimal). For [H5](#), household panel DiD using PLFS data. Robustness checks include parallel trends tests, synthetic controls, Kleibergen-Paap $F > 10$, Lee bounds and Conley spatial HAC errors. The credibility of each

Table 2. Hypothesis identification summary

Hypothesis	Method	Instrument	Expected sign
H1 : Adoption	Probit	Staggered AA FIP onboarding	$\beta_1 > 0$
H2 : Credit expansion	Probit + interaction	Credit score threshold (RDD)	$\beta_5 > 0$
H3 : Default improvement	Difference-in-differences	UPI merchant density	$\beta_3 < 0$
H4 : Demand activation	Panel IV	Aadhaar infrastructure density	$\beta_1 > 0$
H5 : Welfare	HTE (Chernozhukov et al., 2025)	Demographic strata	$\beta_2, \beta_3 > 0$
Note(s): All specifications include district and time fixed effects. Standard errors clustered at the district level. McCrary test required for H2 RDD			

design rests on assumptions that are stated explicitly and that are, in principle, testable. The staggered-FIP strategy for [H1](#) and [H4](#) requires the timing of FIP onboarding to be uncorrelated with district-level credit-demand shocks; event-study leads and onboarding-order placebo tests are proposed to check pre-trends. The RDD for [H2](#) and [H3](#) requires no precise manipulation of the CIBIL 630 to 650 threshold, which McCrary density and covariate-continuity tests can assess. The panel-IV exclusion restriction, that Aadhaar and UPI infrastructure density affects financial-services demand only through platform penetration, cannot be tested directly and is instead supported through over-identification and bounding exercises. Stating these conditions keeps the design falsifiable rather than asserting identification in the absence of data.

4.4 Data sources

[Table 3](#) details the data architecture.

4.5 Power analysis

At 80% power ($\alpha = 0.05$): [H2](#) requires 1,500 observations per group; [H3](#) requires 3,000 loan observations; [H5](#) requires 2,000 households. All within projected data availability.

5. Scenario analysis and projected impacts

5.1 Multi-vertical eligibility expansion

[Figure 4](#) presents eligibility expansion across seven verticals. MSME credit shows the largest proportional gain (278%) because GST and bank data via AA create business credit profiles where none existed.

Table 3. Data sources for empirical analysis

Data element	Source	Frequency	Coverage
Bureau scores	TransUnion CIBIL, Experian	Monthly	Individual, subscribers
Bank transactions	Account Aggregator	Real-time	AA-consented users
UPI logs	NPCI (RBI, 2025c)	Real-time	UPI-active users
Property data	State registration, RERA	Quarterly	Property-level
Household surveys	PLFS, NSSO, Global Findex	Annual	National sample
Loan performance	Lender reports, RBI returns	Monthly	Partner institutions
District controls	Census, RBI BSR, MOSPI	Annual	District-level

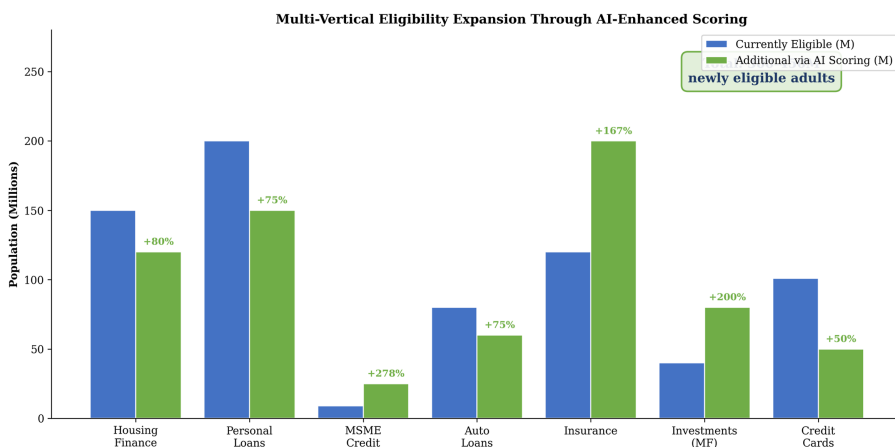


Figure 4. Multi-vertical eligibility expansion: current eligible population versus additional population enabled by AI scoring. *Source:* Author created using data from RBI, IFC (2022) and IRDAI (2025)

5.2 Five-scenario sensitivity analysis

Table 4 presents impact estimates across five scenarios with explicit assumptions, following the transparency standard of Dabla-Norris *et al.* (2015). The moderate projection (\$360 billion) assumes 60% adoption of 450 million eligible adults with 5–8% NPAs. These projections are conditional on stated assumptions and intended to support research design, not to claim realised causal impacts. The scenario parameters are tied to observed analogues rather than chosen freely. Adoption rates of 30–80% are bounded below by realised AA consent growth and above by the documented diffusion of UPI among smartphone users, with the moderate figure of 60% close to smartphone-enabled adult penetration. The non-performing-asset ranges of 3–20% span observed AA-enabled personal-loan performance and stressed informal-segment benchmarks, and the loan-size bands reflect prevailing small-ticket and MSME ticket sizes. Because each projection is a product of these parameters, Table 4 reports the full range rather than a point estimate, and the displacement- and stability-adjusted columns apply explicit haircuts so that every figure can be traced back to its assumptions. The aim is transparency of construction rather than precision of forecast.

Figure 5 presents the moderate-scenario waterfall. Housing finance (\$144 billion) and personal/MSME loans (\$96 billion) dominate.

Table 4. Economic impact sensitivity analysis (10-year, USD)

Dimension	Pessimistic	Moderate	Optimistic	Displace-adj.	Stability-adj.
Adoption	30% of 300M	60% of 450M	80% of 450M	60% of 450M	60% of 450M
Loan size	\$900–\$2,400	\$1,800–\$6,000	\$3,600–\$9,600	Same	Same
NPA rate	15–20%	5–8%	3–5%	5–8%	5–8%
Total demand (\$B)	\$180	\$360	\$480	\$216–\$252	\$270–\$324
Incremental (\$B)	\$72–\$108	\$144–\$216	\$192–\$288	\$86–\$151	\$108–\$194
Housing units (M)	1.5	5.2	10.5	3.1–3.6	3.9–4.7
Platform rev. Y5 (\$M)	\$240	\$480	\$720	\$288–\$336	\$360–\$432

Note(s): Displacement-adjusted reduces base by 30–40% for informal substitution. Stability-adjusted applies Sahay *et al.* (2015) haircut. Incremental assumes 40–60% organic baseline

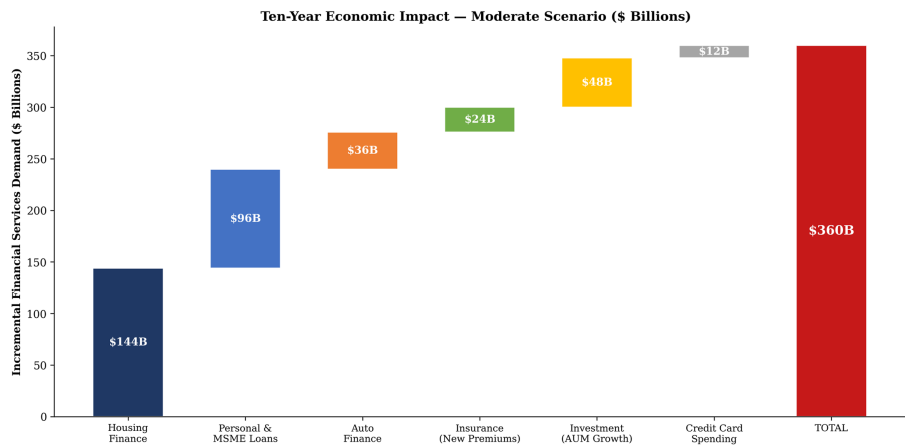


Figure 5. Economic impact waterfall: moderate scenario (\$360 billion). *Source:* Author scenario modelling

5.3 Counterfactual baseline

The organic baseline – fintech’s 25% CAGR, AA expansion, government programmes – accounts for 40–60% of projected growth. Genuinely incremental impact thus ranges from \$108 to \$216 billion. [Figure 6](#) illustrates the adoption trajectory.

5.4 Platform comparison

[Table 5](#) compares the proposed platform against existing Indian fintechs. [Figure 7](#) provides a multi-dimensional radar comparison with Credit Karma.

5.5 Platform cost structure and net impact

The projections above estimate gross and incremental demand, but they do not by themselves establish commercial viability. As both reviewers note, India’s DPI lowers the cost of acquiring data without removing the cost of building and operating a proprietary AI platform. AA, Aadhaar and UPI remove the need to build a private data and identity layer, yet the

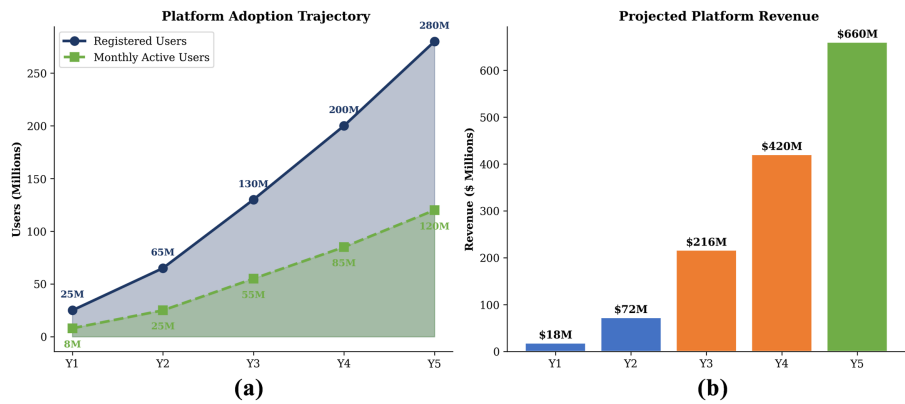


Figure 6. Platform adoption trajectory: (a) registered and monthly active users; (b) projected *platform* revenue. *Source:* Author scenario modelling

Table 5. Proposed platform versus existing Indian fintechs

Dimension	BankBazaar	PolicyBazaar	Perfios/ FinBox/Setu	Proposed platform
Function	Loan marketplace	Insurance marketplace	AA analytics (B2B)	Credit intelligence + marketplace
Target	Bureau-scored adults	Insurance buyers	Lenders/fintechs	Credit-invisible + scored
AA integration	Limited	None	Core (backend)	Full consumer + lender
AI scoring	None	None	Backend only	Consumer-visible AI score
Credit-invisible reach	Low	Low	Moderate	High (no bureau minimum)

Source(s): Author analysis of public platform disclosures (2025–2026)

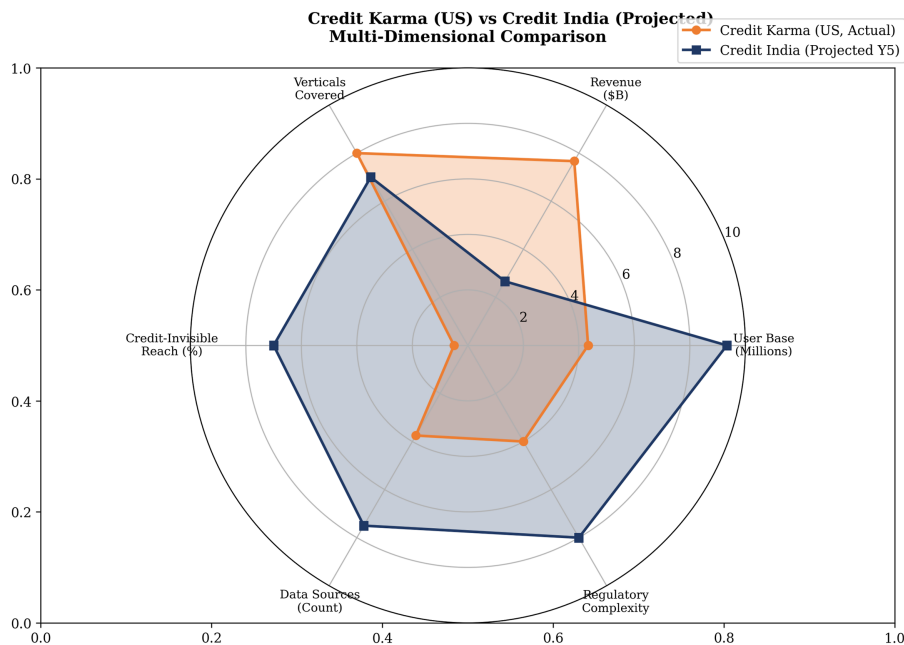


Figure 7. Multi-dimensional radar comparison: Credit Karma (US, actual) versus proposed platform (projected Year 5). *Source:* Author created using [Intuit \(2020\)](#) and [Upstart \(2025\)](#) data

platform must still fund engineering, model development, AI compute, regulatory compliance and customer acquisition. The closest pure-play analogue, Upstart, incurred operating expenses of roughly \$809 million in fiscal 2024, of which about \$127 million was sales and marketing and a further \$125 million borrower acquisition, while still posting a net loss ([Upstart Holdings, 2025](#)). This gives a tangible proof of concept for the cost envelope of an AI-native lender, and a benchmark against which an India-built platform, with lower engineering costs and DPI-subsidised data access, can be scaled.

[Table 6](#) applies the pessimistic, moderate and optimistic discipline used in [Section 5.2](#) to a ten-year cost structure, separating one-time capital expenditure (CapEx) from recurring

Table 6. Illustrative platform cost structure (10-year, USD)

Cost component	Type	Pessimistic	Moderate	Optimistic
Platform and data engineering (build)	CapEx	\$12M	\$25M	\$45M
AI/ML model development and validation	CapEx	\$8M	\$15M	\$28M
Security and compliance architecture (DPDP, DPO, DPIA)	CapEx	\$5M	\$10M	\$18M
Licensing, integration and data onboarding	CapEx	\$2M	\$5M	\$9M
Subtotal CapEx (one-time)		\$27M	\$55M	\$100M
Cloud and AI compute (training + inference)	OpEx	\$70M	\$140M	\$260M
Engineering and data-science talent	OpEx	\$90M	\$180M	\$320M
Model monitoring, re-training and explainability	OpEx	\$25M	\$50M	\$95M
Compliance, audit and fairness testing	OpEx	\$20M	\$45M	\$80M
Customer acquisition and support	OpEx	\$45M	\$110M	\$230M
Subtotal OpEx (10-year cumulative)		\$250M	\$525M	\$985M
Total 10-year cost		\$277M	\$580M	\$1,085M

Note(s): Illustrative order-of-magnitude estimates, not audited projections. Benchmarked to [Upstart Holdings \(2025\)](#) operating-cost disclosures, Indian technology-sector salary ranges, DPDP compliance requirements ([EY, 2025](#)) and prevailing cloud and AI compute and fintech customer-acquisition benchmarks ([Grand View Research, 2024](#)). CapEx is one-time build cost; OpEx is cumulative over ten years

operating expenditure (OpEx). CapEx covers platform and data engineering, AI and ML model development and validation, DPDP-aligned security and compliance architecture and licensing. Full-featured digital-lending platforms are built for roughly \$0.5 to \$2 million, but national, multi-vertical infrastructure raises this to tens of millions. OpEx is dominated by cloud and AI compute, where inference rather than training accounts for most spend at production scale, followed by engineering and data-science talent (benchmarked to Indian salaries of Rs 10 to 35 lakh per annum, well below US equivalents), model monitoring and explainability, ongoing compliance (a Significant Data Fiduciary must maintain a Data Protection Officer, an independent data auditor and periodic Data Protection Impact Assessments under the DPDP Rules; [EY, 2025](#)), and customer acquisition. Blended acquisition cost per user is assumed well below the global fintech benchmark of about \$1,450, because distribution runs on existing DPI rails and consumer-side pricing is zero by design ([Section 4.1](#)).

Under the moderate scenario, cumulative ten-year cost is about \$0.58 billion, within a range of \$0.28–\$1.09 billion, against projected platform revenue reaching \$480 million a year by Year 5 ([Table 4](#)) and a digital-lending market growing at more than 30% annually ([Grand View Research, 2024](#)). Two implications follow. First, operating-cost recovery rather than infrastructure cost is the binding viability constraint, which is what the DPI-augmented model in [Section 4.1](#) predicts. Second, setting illustrative platform costs of well under \$1 billion against the incremental economic impact of \$108 to \$216 billion ([Section 5.3](#)) leaves the welfare case intact: even under the optimistic cost scenario, total platform cost is of the order of 1% of the incremental impact it helps unlock. DPI moves the cost frontier from data acquisition to compute, talent and compliance, and these costs, not access to bureau data, are what determine commercial sustainability. The figures are illustrative, order-of-magnitude estimates meant to bound the cost question raised in review, and are not audited projections.

6. Discussion

6.1 Algorithmic fairness

Western fairness definitions prove insufficient for India ([Sambasivan et al., 2021](#)). Caste intersects with income, geography and digital access uniquely. [Bhukta \(2024\)](#) shows SC households benefit from branch expansion for savings but not credit. PIN-code features create geographic proxy discrimination. India’s reservation system may legally require favouring

SC/ST/OBC populations. Figure 8 compares traditional and AI-enhanced credit variable architectures. Putting these criteria into practice needs dedicated fairness machinery, not definitions alone. This includes pre-deployment bias audits across caste, gender and geographic strata; continuous post-deployment monitoring of approval and default rates by protected group; counterfactual and proxy-feature testing to catch PIN-code-based geographic discrimination and a documented escalation and remediation process. These measures put the FREE-AI fairness principle (RBI, 2025b) into operation and are mapped to design requirements in Table 6.

6.2 Welfare framework

Credit eligibility, access and welfare represent three stages with substantial drop-off. The Jan Dhan Yojana – 500+ million accounts, 20% dormant (Singh *et al.*, 2021) – demonstrates access does not equal welfare. The 2010 Andhra Pradesh crisis showed expansion without safeguards produces catastrophic losses. For 90% of women entrepreneurs who have never borrowed formally (IFC, 2022) and 42% of women with inactive accounts (Demirguc-Kunt, Klapper, Singer, & Ansar, 2022), thin data may penalise rather than assist.

6.3 Financial stability and scoring-error risks

Sahay *et al.* (2015) caution that inclusion raises stability risks without proportionate supervision. Correlated AI failures across verticals represent underexamined systemic risk. Selection effects (Mbiti & Weil, 2015) mean the most excluded may be reached last.

A specific and underexamined risk is scoring error in the form of false positives, where applicants receive high AI scores but are in fact high-risk. Section 6.2 considers the consumer harm of false negatives, where creditworthy applicants are wrongly excluded; false positives instead transfer risk to lenders and, through them, to the wider financial system. At the firm level, systematic over-prediction inflates approvals and feeds delinquency, eroding net interest margins and, if mispricing is large enough, capital adequacy; under the NPA ranges in Table 4, a rise of a few percentage points in realised defaults materially compresses unit economics. At the system level, because one model and a shared set of AA-derived features would underwrite borrowers across housing, personal, MSME and insurance lending, a common scoring bias

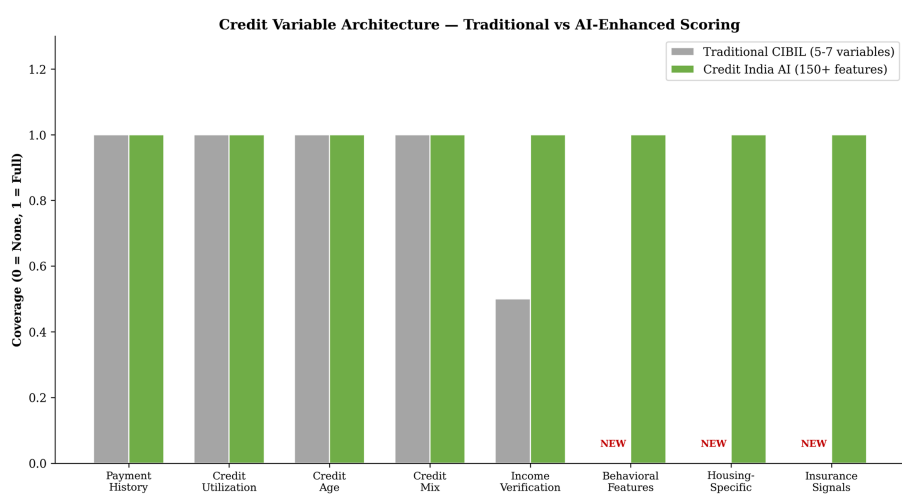


Figure 8. Credit variable architecture: traditional CIBIL (5–7 variables) versus AI-enhanced (150+ features). Source: Author created

produces correlated errors rather than isolated ones, so the multi-vertical design that drives the upside in [Section 5](#) also concentrates model risk. This points to vertical-specific score calibration, conservative thresholds and reserve buffers during model burn-in, champion-challenger deployment with held-out validation and supervisory stress-testing of AI-underwritten portfolios, in line with the work of [Sahay et al. \(2015\)](#). Controlling false positives is therefore not only a prudential concern for individual lenders but a financial-stability requirement for the ecosystem.

6.4 Regulatory constraints as design parameters

[Table 7](#) maps regulatory provisions to platform requirements. The FREE-AI Explainability Sutra ([RBI, 2025b](#)) conflicts with complex ML models; the paper proposes SHAP-based explanations for audit and simplified natural-language summaries for consumers. Explainability works at two levels. For supervisors and internal audit, SHAP-based global and local attributions give model-level and decision-level transparency that meets the FREE-AI Explainability principle and supports adverse-action reasoning. For consumers, the same attributions are turned into simplified natural-language reason codes that stay understandable without exposing proprietary model internals. Where a vertical or a regulator needs full interpretability, the platform can deploy monotonic gradient-boosting or explainable-boosting models as interpretable challengers, accepting a small and measurable accuracy trade-off in return for auditability ([Kozodoi et al., 2022](#); [Hurlin et al., 2026](#)).

6.5 Limitations

This paper develops a platform model and testable hypotheses alongside illustrative scenario analysis. It does not present executed causal estimates from loan-level data; instead, it specifies identification strategies implementable when data access is available. The numerical projections are conditional on stated assumptions. External validity concerns arise from structural differences between India’s informal sector and US thin-file borrowers. Digital literacy constraints (36% rural internet penetration) limit addressable population. The revenue model assumes a competitive product marketplace that may not exist at the projected scale.

7. Conclusion

India’s financial exclusion reflects an information gap affecting 700+ million adults. This paper demonstrated that consumer platform models, AI underwriting and public infrastructure create conditions for a credit intelligence platform addressing this gap across multiple

Table 7. Regulatory constraints mapped to platform design

Provision	Requirement	Design implication
DL Directions: KFS (RBI, 2025a)	Disclose APR and total cost	Recommendations include Key Fact Statement
DL Directions: Cooling-off	1-day minimum exit	Enable 24-h cancellation
DL Directions: Transparency	Unbiased lender comparison	Rank by borrower benefit, not commission
FREE-AI: Explainability (RBI, 2025b)	Understandable by design	SHAP + simplified summaries
FREE-AI: Fairness	Non-discrimination	Caste/gender/geography audits
DPDP Act (Gov. of ndia, 2023)	Separate consent per purpose	Granular consent management
DPDP Rules (Gov. of India, 2025)	Mandatory DPO and DPIA	Data Protection Officer; annual assessments

Source(s): [RBI \(2025a, b\)](#), [Government of India \(2023, 2025\)](#).

verticals. The moderate projection of \$360 billion, adjusted to \$108–216 billion after organic growth accounting, remains economically significant.

Four policy priorities emerge. First, the RBI should issue safe-harbour guidance permitting AA-derived cash-flow underwriting for loans below \$42,000. Second, the National Housing Bank should create an AI-Ready refinancing window. Third, IRDAI should expand its insurtech sandbox for AI risk profiling. Fourth, SEBI should accelerate AA integration for mutual funds. The infrastructure exists and the data flows are live; what remains is the empirical evidence these hypotheses are designed to generate.

About the author

Malcolm Athaide is an independent researcher with interests that include AI in financial services, credit scoring for underserved populations and digital public infrastructure.

Ethics statement

This research does not involve human participants, animal subjects or personally identifiable individual-level data. The study relies on publicly available aggregate statistics. No ethics committee approval was required.

Data availability

This study uses publicly available aggregate data from Sahamati, TransUnion CIBIL, Reserve Bank of India, NPCI, UIDAI and published regulatory documents. No proprietary or restricted-access individual-level data were used.

References

- Armstrong, M. (2006). Competition in two-sided markets. *RAND Journal of Economics*, 37(3), 668–691. doi: [10.1111/j.1756-2171.2006.tb00037.x](https://doi.org/10.1111/j.1756-2171.2006.tb00037.x).
- Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*, 143(1), 30–56. doi: [10.1016/j.jfineco.2021.05.047](https://doi.org/10.1016/j.jfineco.2021.05.047).
- Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the rise of FinTechs: Credit scoring using digital footprints. *Review of Financial Studies*, 33(7), 2845–2897. doi: [10.1093/rfs/hhz099](https://doi.org/10.1093/rfs/hhz099).
- Bhukta, R. (2024). Bridging the caste divide: How financial inclusion drives social inclusion, World Bank Impact Evaluations Working Paper.
- Blattner, L., & Nelson, S. (2021). *How costly is noise? Data and disparities in consumer credit* (ArXiv Preprint abs/2105.07554). Available from: <https://arxiv.org/abs/2105.07554>
- Buchak, G., Matvos, G., Piskorski, T., & Seru, A. (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics*, 130(3), 453–483. doi: [10.1016/j.jfineco.2018.03.011](https://doi.org/10.1016/j.jfineco.2018.03.011).
- Chernozhukov, V., Demirer, M., Duflo, E., & Fernández-Val, I. (2025). Fisher–Schultz lecture: Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. *Econometrica*, 93(4), 1121–1164. doi: [10.3982/ECTA19303](https://doi.org/10.3982/ECTA19303).
- Dabla-Norris, E., Ji, Y., Townsend, R. M., & Unsal, D. F. (2015). Identifying constraints to financial inclusion and their impact on GDP and inequality. IMF Working Paper No. 15/22. doi: [10.5089/9781484398067.001](https://doi.org/10.5089/9781484398067.001).
- Demirguc-Kunt, A., Klapper, L., Singer, D., & Ansar, S. (2022). *The Global Findex Database 2021*. Washington, DC: World Bank. doi: [10.1596/978-1-4648-1897-4](https://doi.org/10.1596/978-1-4648-1897-4).
- Di Maggio, M., Ratnadiwakara, D., & Carmichael, D. (2022). Invisible primes: Fintech lending with alternative data. NBER Working Paper No. 29840. doi: [10.3386/w29840](https://doi.org/10.3386/w29840).

- D'Silva, D., Filkova, Z., Packer, F., & Tiwari, S. (2019). The design of digital financial infrastructure: Lessons from India, BIS Papers No. 106, Available from: <https://www.bis.org/publ/bppdf/bispap106.pdf> [Accessed 28 March 2026].
- EY (2025). Decoding the Digital Personal Data Protection Act, 2023 and DPDP Rules, 2025, EY India, Available from: https://www.ey.com/en_in/insights/cybersecurity/decoding-the-digital-personal-data-protection-act-2023 [Accessed 27 June 2026].
- Field, E., Pande, R., Rigol, N., Schaner, S., & Troyer Moore, C. (2021). On her own account: How strengthening women's financial control impacts labor supply and gender norms. *American Economic Review*, 111(7), 2342–2375. doi: [10.1257/aer.20200705](https://doi.org/10.1257/aer.20200705).
- Frost, J. (2020). The economic forces driving fintech adoption. BIS Working Paper No. 838, Available from: <https://www.bis.org/publ/work838.pdf> [Accessed 28 March 2026].
- Fuster, A., Plosser, M., Schnabl, P., & Vickery, J. (2019). The role of technology in mortgage lending. *Review of Financial Studies*, 32(5), 1854–1899. doi: [10.1093/rfs/hhz018](https://doi.org/10.1093/rfs/hhz018).
- Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. *Journal of Finance*, 77(1), 5–47. doi: [10.1111/jofi.13090](https://doi.org/10.1111/jofi.13090).
- Gambacorta, L., Huang, Y., Qiu, H., & Wang, J. (2024). How do machine learning and non-traditional data affect credit scoring?. *Journal of Financial Stability*, 73, 101284. doi: [10.1016/j.jfs.2024.101284](https://doi.org/10.1016/j.jfs.2024.101284).
- Government of India (2023). The Digital Personal Data Protection Act, 2023, Act No. 22 of 2023, Available from: <https://www.indiacode.nic.in/bitstream/123456789/22037/1/a2023-22.pdf> [Accessed 28 March 2026].
- Government of India (2025). Digital personal data protection Rules, 2025, Press Information Bureau, 3 January, Available from: <https://www.pib.gov.in/PressReleasePage.aspx?PRID=2190655> [Accessed 28 March 2026].
- Grand View Research (2024). India digital lending platform market size & outlook, 2024-2030, Available from: <https://www.grandviewresearch.com/horizon/outlook/digital-lending-platform-market/india> [Accessed 27 June 2026].
- He, Z., Huang, J., & Zhou, J. (2023). Open banking: Credit market competition when borrowers own the data. *Journal of Financial Economics*, 147(2), 449–474. doi: [10.1016/j.jfineco.2022.12.003](https://doi.org/10.1016/j.jfineco.2022.12.003).
- Hurlin, C., Pérignon, C., & Saurin, S. (2026). The fairness of credit scoring models. *Management Science*, 72(1), 406–425.
- IFC (2022). *Financing India's MSMEs: Estimation of debt requirement*. Washington, DC: International Finance Corporation.
- Intuit (2020). *Intuit to acquire credit Karma*. Investor Relations Press Release. 24 February, Available from: <https://investors.intuit.com/news-events/press-releases/detail/237/intuit-to-acquire-credit-karma> [Accessed 28 March 2026].
- IRDAI (2025). Annual Report 2024–25, insurance regulatory and development Authority of India, Hyderabad.
- Jack, W., & Suri, T. (2014). Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *American Economic Review*, 104(1), 183–223. doi: [10.1257/aer.104.1.183](https://doi.org/10.1257/aer.104.1.183).
- Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending. *Financial Management*, 48(4), 1009–1029. doi: [10.1111/fima.12297](https://doi.org/10.1111/fima.12297).
- Jappelli, T., & Pagano, M. (2002). Information sharing, lending and defaults: Cross-country evidence. *Journal of Banking and Finance*, 26(10), 2017–2045. doi: [10.1016/S0378-4266\(01\)00185-6](https://doi.org/10.1016/S0378-4266(01)00185-6).
- Kozodoi, N., Jacob, J., & Lessmann, S. (2022). Fairness in credit scoring: Assessment, implementation and profit implications. *European Journal of Operational Research*, 297(3), 1083–1098. doi: [10.1016/j.ejor.2021.06.023](https://doi.org/10.1016/j.ejor.2021.06.023).
- Mbiti, I., & Weil, D. N. (2015). Mobile banking: The impact of M-Pesa in Kenya. In S. Edwards, S. Johnson, & D. N. Weil (Eds.), *African Successes* (Vol. III, pp. 247–293). Chicago, IL: University of Chicago Press.

- Nalluri, V., & Chen, L. -S. (2024). Modelling the FinTech adoption barriers in emerging economies. *Technological Forecasting and Social Change*, 199, 123049. doi: [10.1016/j.techfore.2023.123049](https://doi.org/10.1016/j.techfore.2023.123049).
- Philippon, T. (2016). The FinTech opportunity. NBER Working Paper No. 22476. doi: [10.3386/w22476](https://doi.org/10.3386/w22476).
- RBI (2025a). Reserve Bank of India (digital lending) Directions, 2025, Available from: https://www.rbi.org.in/scripts/BS_CircularIndexDisplay.aspx?Id=12848 [Accessed 28 March 2026].
- RBI (2025b). Framework for responsible and ethical enablement of AI (FREE-AI), Press Release, Available from: https://www.rbi.org.in/scripts/bs_pressreleasedisplay.aspx/BS_PressReleaseDisplay.aspx?prid=61018 [Accessed 28 March 2026].
- RBI (2025c). Payment system indicators, Available from: <https://www.rbi.org.in/Scripts/PSIUserView.aspx?Id=49> [Accessed 28 March 2026].
- Rochet, J. -C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990–1029. doi: [10.1162/154247603322493212](https://doi.org/10.1162/154247603322493212).
- Sahamati (2025). Sahamati marks 4th account aggregator Foundation Day, Available from: <https://sahamati.org.in/sahamati-marks-4th-account-aggregator-foundation-day-indias-data-empowerment-revolution-scales-new-heights/> [Accessed 28 March 2026].
- Sahay, R., Cihák, M., N'Diaye, P., Barajas, A., Mitra, S., Kyobe, A., . . . Yousefi, S. R. (2015). Financial inclusion: Can it meet multiple macroeconomic goals?, IMF Staff Discussion Note No. 15/17, 2015(17), 1. doi: [10.5089/9781513585154.006](https://doi.org/10.5089/9781513585154.006).
- Sambasivan, N., Arnesen, E., Hutchinson, B., Doshi, T., & Prabhakaran, V. (2021). Re-imagining algorithmic fairness in India and beyond. In *Proceedings of FAccT '21* (pp. 315–328). ACM. doi: [10.1145/3442188.3445896](https://doi.org/10.1145/3442188.3445896).
- SIDBI and CRISIL (2025). MSME pulse: Assessment of micro, small and medium enterprises debt demand and credit gap. *Small Industries Development Bank of India and CRISIL Intelligence*, Available from: <https://www.sidbi.in/msme-pulse> [Accessed 27 June 2026].
- Singh, B. P., Kumari, A., Sharma, T., & Malhotra, A. (2021). Financial inclusion, Pradhan Mantri Jan Dhan Yojana scheme and economic growth. *Economic Notes*, 50(3), e12186. doi: [10.1111/ecno.12186](https://doi.org/10.1111/ecno.12186).
- Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *American Economic Review*, 71(3), 393–410.
- Suri, T., & Jack, W. (2016). The long-run poverty and gender impacts of mobile money. *Science*, 354(6317), 1288–1292. doi: [10.1126/science.aah5309](https://doi.org/10.1126/science.aah5309).
- TransUnion CIBIL (2025). Annual credit market summary, Available from: <https://www.transunioncibil.com/> [Accessed 28 March 2026].
- UIDAI (2026). State/UT wise Aadhaar saturation, Available from: <https://uidai.gov.in/images/AadhaarSaturationReport.pdf> [Accessed 28 March 2026].
- Upstart (2025). Access to credit report, Available from: <https://www.upstart.com/lenders/regulatory-compliance/access-to-credit-report/> [Accessed 28 March 2026].
- Upstart Holdings (2025). Form 10-K: Annual report for the fiscal year ended December 31, 2024, U.S. Securities and Exchange Commission, Available from: <https://www.sec.gov/Archives/edgar/data/0001647639/000164763925000018/upst-20241231.htm> [Accessed 27 June 2026].
- Weyl, E. G. (2010). A price theory of two-sided platforms. *American Economic Review*, 100(4), 1642–1672. doi: [10.1257/aer.100.4.1642](https://doi.org/10.1257/aer.100.4.1642).

Corresponding author

Malcolm Athaide can be contacted at: malcolm.athaide@gmail.com