

Determinants of Home Loan Default Risk among Salaried and Self- Employed Borrowers in India: A Comparative Empirical Study Using PLS-SEM and Logistic Regression

Komal Sharma

komalsharma5369@gmail.com

Others <https://orcid.org/0009-0005-9192-2688>

Deepika Rathi

others <https://orcid.org/0009-0000-0790-2371>

Dr. Garv Nijhara

Others <https://orcid.org/0009-0001-1443-3515>

Dr. jainish Bhagat

others <https://orcid.org/0000-0003-3270-9172>

Research Article

Keywords: home loan default, credit risk, salaried and self-employed borrowers, PLS-SEM, SmartPLS, multi-group analysis, financial literacy, India, logistic regression

Posted Date: September 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11092552/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Rising home loan delinquency in India's retail credit portfolio has renewed interest in identifying the behavioural, financial, and macroeconomic determinants that distinguish borrowers who default from those who do not, particularly across the salaried and self-employed segments that Indian housing finance institutions underwrite under materially different risk protocols. This study proposes and empirically tests an integrated model of home loan default risk that combines financial-behavioural constructs — income stability, debt-servicing burden, loan-to-value adequacy, credit-history discipline, financial literacy, psychological/behavioural traits, and macroeconomic stress perception — as antecedents of borrowers' perceived default risk, with employment type (salaried vs self-employed) modelled as a moderating variable. The measurement and structural models are estimated using variance-based partial least squares structural equation modelling (PLS-SEM) in SmartPLS, complemented by a supplementary binary logistic regression on loan-level indicators (EMI-to-income ratio, loan-to-value ratio, and credit bureau score) to connect the behavioural model to observable default outcomes. The full analytical pipeline — reliability and convergent validity assessment, discriminant validity via the heterotrait-monotrait (HTMT) ratio and the Fornell-Larcker criterion, bootstrapped structural path estimation, PLS-predict out-of-sample assessment, multi-group analysis (PLS-MGA), and logistic-regression diagnostics including ROC-AUC — is demonstrated on a simulated pilot dataset (N = 400) prior to fieldwork. The results (illustrative, from simulated data) indicate that all seven hypothesised antecedents significantly predict perceived default risk, that credit-history discipline exerts a significantly stronger risk-mitigating effect for salaried borrowers than for self-employed borrowers, and that the supplementary logistic model discriminates defaulting from non-defaulting accounts with good accuracy. The paper's principal contribution is a comparative, moderation-based structural framework — not previously applied jointly to salaried and self-employed home-loan borrowers in the Indian context — together with a validated instrument and analysis pipeline ready for empirical deployment. Implications for housing finance underwriting, differentiated risk-based pricing, and financial-literacy interventions are discussed, alongside limitations arising from the pilot's simulated nature.

JEL Codes: *G21 (Banks; Depository Institutions; Mortgages), G32 (Financing Policy; Financial Risk and Risk Management), D14 (Household Saving; Personal Finance), R21 (Housing Demand), D91 (Role of Psychological, Behavioural, and Cognitive Factors in Decision Making), C38 (Classification Methods; Factor Models)*

1. Introduction

Housing finance is among the fastest-growing segments of India's retail credit market, driven by rising urbanisation, government incentives for affordable housing, and increasing formalisation of income documentation. Alongside this growth, however, housing finance institutions (HFIs) and banks continue to report a persistent, non-trivial incidence of loan delinquency and default, with implications for asset quality, provisioning costs, and financial stability. A substantial empirical literature has examined the determinants of housing loan default in India using large administrative loan-level datasets, establishing

that EMI-to-income ratio and the loan-to-value position relative to the current market value of the collateral are primary drivers of default (Bandyopadhyay & Saha, 2011), that macroeconomic conditions and country-level governance quality shape aggregate non-performing loan and wilful-default trends (Jayadev & Padma, 2019), and that behavioural and psychological factors – financial literacy, materialism, indebtedness, and risk perception – independently shape borrowers' default intentions (Thomas et al., 2023).

A recurring but under-examined distinction in this literature is employment type. Salaried borrowers typically present verifiable, relatively stable monthly income supported by employer documentation, whereas self-employed borrowers – proprietors, professionals, and small-business owners – face income volatility tied to business cycles, seasonal demand, and less standardised income verification. Prior loan-level evidence suggests employment status has a significant risk-differentiating effect, with higher default incidence historically observed among self-employed borrowers than salaried borrowers. Despite this, few studies model salaried and self-employed borrowers as a formal moderating comparison within a single structural framework that simultaneously incorporates financial, behavioural, and macroeconomic antecedents.

This study addresses that gap by developing and testing a comparative structural model of home loan default risk that (a) integrates financial statement-style indicators (EMI-to-income ratio, loan-to-value ratio, credit bureau score) with behavioural constructs (financial literacy, psychological/behavioural traits, macroeconomic stress perception) grounded in the borrowing-behaviour literature (Altaf & Shah, 2025; Mandal & Pachiyappan, 2025), and (b) formally tests employment type as a moderator of the financial/behavioural-to-default-risk relationships using multi-group PLS-SEM. The study further links the behavioural structural model to observable loan performance through a supplementary logistic regression on loan-level default status, addressing the methodological limitation that perception-based SEM constructs cannot, by themselves, model a binary default outcome.

1.1 Research Objectives

- RO1: To identify and validate the key financial, behavioural, and macroeconomic determinants of perceived home loan default risk among Indian home-loan borrowers.
- RO2: To empirically test the structural relationships between these determinants and perceived default risk using PLS-SEM.
- RO3: To examine whether employment type (salaried vs self-employed) moderates these structural relationships, using multi-group analysis (PLS-MGA).
- RO4: To complement the behavioural structural model with a supplementary logistic regression linking loan-level financial indicators and employment type to actual (observed) default status.

1.2 Significance of the Study

The findings are intended to inform (i) housing finance institutions seeking to differentiate underwriting and risk-based pricing between salaried and self-employed applicants, (ii) regulators and credit bureaus interested in behavioural early-warning indicators that complement traditional financial ratios, and (iii) financial-literacy policy design targeted at segments shown to carry disproportionate behavioural risk.

2. Literature Review

2.1 Financial Determinants of Home Loan Default in India

Using a large administrative dataset of 13,487 housing loan accounts sanctioned between 1993 and 2007 by Indian banks and housing finance companies, Bandyopadhyay and Saha (2011) show that default is primarily driven by the EMI-to-income ratio and by adverse movement in the market value of the mortgaged property relative to the outstanding loan amount, with a 10 percent increase in the EMI-to-income ratio raising delinquency odds by approximately 4.5 percent and a 10 percent adverse change in collateral value raising default odds by approximately 1.55 percent. The same study reports that self-employed borrowers exhibit a materially higher default incidence than salaried borrowers, and that metropolitan borrowers are riskier than those in Tier-II/III towns, plausibly reflecting easier access to credit and greater over-leverage in metros.

At the macro level, Jayadev and Padma (2019), analysing CIBIL-disclosed wilful-defaulter data from 2003 to 2016, find that wilful default is significantly associated with weak country-level governance quality and adverse macroeconomic conditions, consistent with the broader cross-country finding that slower GDP growth, rising unemployment, and falling asset prices are principal determinants of non-performing loans (Nkusu, 2011; Salas & Saurina, 2002).

2.2 Behavioural and Psychological Determinants

A parallel behavioural-finance literature situates default intention within borrowers' psychological and financial-literacy profiles rather than loan-level ratios alone. Thomas et al. (2023), studying young urban home-loan clients in India, find that financial literacy, materialism, emotion, indebtedness, and risk perception jointly shape default intention. Using SmartPLS-based structural equation modelling on 385 Indian respondents, Mandal and Pachiyappan (2025) show that materialism has the strongest direct effect on real-estate loan default behaviour, that financial literacy reduces financial stress, and that financial stress exerts only a limited mediating role between behavioural factors and default behaviour. In a related PLS-SEM study of 425 Indian credit users, Altaf and Shah (2025) demonstrate that credit-score literacy and psychological traits significantly shape borrowing behaviour, and that the resulting model exhibits good out-of-sample predictive validity via PLSpredict and CVPAT, providing methodological precedent for the analytical approach adopted here.

2.3 Employment Type as a Risk-Differentiating Factor

Employment type is consistently identified as a risk-relevant borrower characteristic in the Indian housing finance literature, with self-employed borrowers exhibiting higher default incidence attributable

to income volatility, less standardised income verification, and greater exposure to business-cycle and sector-specific shocks. However, existing studies typically enter employment type as a covariate or control rather than as a formal moderator of the full set of financial-behavioural antecedents, leaving open the question of whether the strength – not merely the level – of these relationships differs systematically between the two segments.

2.4 Research Gap and Contribution

Three gaps motivate the present study. First, financial-ratio-based studies (e.g., Bandyopadhyay & Saha, 2011) and behavioural PLS-SEM studies (e.g., Altaf & Shah, 2025; Mandal & Pachiyappan, 2025) have largely proceeded on separate methodological tracks; no study identified integrates both within a single comparative framework for the Indian home-loan market. Second, employment type has not been formally tested as a structural moderator across a comprehensive antecedent set spanning financial, behavioural, and macroeconomic dimensions using multi-group PLS-SEM. Third, the mechanism connecting perception-based behavioural constructs to observed, loan-level default outcomes is rarely made explicit. This study addresses all three gaps.

3. Theoretical Framework and Hypotheses Development

The conceptual model draws on the borrower-characteristics and loan-characteristics tradition in Indian housing-finance default research (Bandyopadhyay & Saha, 2011), the behavioural-finance tradition linking financial literacy and psychological traits to borrowing behaviour (Altaf & Shah, 2025), and the macro-financial vulnerability literature on systemic stress and loan performance (Nkusu, 2011; Salas & Saurina, 2002). Seven exogenous constructs are hypothesised to predict the endogenous construct, Perceived Default Risk (PDR), with Employment Type specified as a categorical moderator of all seven paths.

3.1 Hypotheses

Hypothesis	Statement
H1	Income Stability (IS) is negatively associated with Perceived Default Risk (PDR).
H2	Debt Servicing Burden (DSB) is positively associated with Perceived Default Risk (PDR).
H3	Loan-to-Value ratio concern (LTV) is positively associated with Perceived Default Risk (PDR).
H4	Credit History Discipline (CHD) is negatively associated with Perceived Default Risk (PDR).
H5	Financial Literacy (FL) is negatively associated with Perceived Default Risk (PDR).
H6	Psychological/Behavioural Traits (PBT – impulsivity, risk propensity, present bias) are positively associated with Perceived Default Risk (PDR).
H7	Macroeconomic Stress Perception (MSP) is positively associated with Perceived Default Risk (PDR).
H8a-g	Employment Type (salaried vs self-employed) moderates the strength of each of the H1-H7 relationships.
H9	The salaried and self-employed sub-groups differ significantly in overall structural path patterns (PLS-MGA / measurement invariance via MICOM).
<i>Note: H1-H7 are estimated on the pooled sample; H8a-g and H9 are tested via multi-group PLS-SEM (Section 5.5).</i>	

4. Research Methodology

4.1 Research Design

The study adopts a quantitative, cross-sectional, survey-based design combined with loan-level financial data, consistent with prior PLS-SEM studies of Indian borrowing behaviour (Altaf & Shah, 2025; Mandal & Pachiyappan, 2025). A structured questionnaire (Appendix A) captures the seven exogenous constructs and the endogenous construct via 5-point Likert-scale reflective indicators (1 = Strongly Disagree, 5 = Strongly Agree), while loan-level fields (EMI, income, loan-to-value ratio, credit bureau score, tenure, interest rate) are captured either from the respondent's loan documentation or self-reported with bank-statement verification where feasible.

4.2 Population, Sampling, and Sample Size

The target population is home-loan borrowers in India, stratified into two employment-type segments: salaried employees and self-employed individuals (professionals and business owners). A purposive, quota-based sampling strategy is recommended, targeting approximately equal sub-samples of salaried and self-employed borrowers across Tier-1, Tier-2, and Tier-3 cities to support multi-group comparison. Minimum sample size was determined via G*Power a-priori power analysis for multiple regression (7 predictors, medium effect size $f^2 = 0.15$, $\alpha = 0.05$, power = 0.95), yielding a minimum requirement of

approximately 153 respondents overall and at least 100 per employment-type sub-group to support stable multi-group PLS-SEM estimation, consistent with the sample sizes used in comparable Indian PLS-SEM borrowing-behaviour studies (385–425 respondents; Altaf & Shah, 2025; Mandal & Pachiyappan, 2025).

For methodological validation prior to fieldwork, the full pipeline below was executed on a simulated pilot dataset of N = 400 (Salaried n = 210; Self-Employed n = 190), generated under a documented data-generating process reported in the accompanying workbook (Simulated_Dataset_and_Results.xlsx, sheet "Simulation_Method"). This simulated dataset must be replaced with actual field data before submission.

4.3 Measures and Constructs

Construct	Role	Items	Representative Content
Income Stability (IS)	Exogenous, reflective	4	Regularity and predictability of income over the past 12 months
Debt Servicing Burden (DSB)	Exogenous, reflective	4	Perceived strain of EMI obligations relative to income
Loan-to-Value Concern (LTV)	Exogenous, reflective	4	Perceived adequacy of collateral value relative to loan outstanding
Credit History Discipline (CHD)	Exogenous, reflective	4	Self-reported discipline in credit card and prior loan repayment
Financial Literacy (FL)	Exogenous, reflective	4	Understanding of interest rates, credit scores, and loan terms
Psychological/Behavioural Traits (PBT)	Exogenous, reflective	4	Impulsivity, present bias, and risk propensity in financial decisions
Macroeconomic Stress Perception (MSP)	Exogenous, reflective	4	Perceived vulnerability to interest-rate and income shocks
Perceived Default Risk (PDR)	Endogenous, reflective	4	Self-assessed likelihood of missing or delaying EMI payments
Employment Type	Categorical moderator	1	Salaried / Self-Employed (grouping variable for PLS-MGA)

Full item wording is provided in Appendix A. All reflective indicators use a 5-point Likert scale; items were adapted from validated scales in the cited Indian borrowing-behaviour literature and refined for the home-loan context.

4.4 Common Method Bias and Data Quality

Because all Likert-scale constructs are collected from a single respondent at one point in time, common method bias (CMB) is a concern. Procedural remedies recommended for the fieldwork stage include: psychological separation of predictor and criterion items in the questionnaire, guaranteed respondent

anonymity, and reverse-worded attention-check items. Statistically, a full collinearity VIF test (Kock, 2015) should be run across all constructs in SmartPLS; VIF values below 3.3 support the absence of severe CMB, alongside Harman's single-factor test as a supplementary check.

4.5 Data Analysis Strategy

Data analysis proceeds in three stages using SmartPLS 4 and supplementary statistical software (Python/R or SPSS for the logistic-regression arm):

- Stage 1 - Measurement model assessment: internal consistency reliability (Cronbach's alpha, Composite Reliability), convergent validity (indicator loadings > 0.70 , Average Variance Extracted > 0.50), and discriminant validity via the Heterotrait-Monotrait ratio (HTMT < 0.90 ; Henseler et al., 2015) and the Fornell-Larcker criterion (Fornell & Larcker, 1981).
- Stage 2 - Structural model assessment: path coefficients estimated via PLS algorithm with bootstrapping (5,000 resamples, bias-corrected and accelerated confidence intervals), coefficient of determination (R^2), effect size (f^2), predictive relevance (Q^2 , blindfolding) and out-of-sample predictive power (PLSpredict; Hair et al., 2019).
- Stage 3 - Multi-group analysis (PLS-MGA): measurement invariance is first established via the Measurement Invariance of Composite Models (MICOM) procedure; if partial or full invariance holds, structural paths are compared across the Salaried and Self-Employed sub-groups using permutation-based PLS-MGA (5,000 permutations) to test H8a-g and H9.
- Supplementary Stage - Logistic regression: a binary logistic regression of actual/observed default status on EMI-to-income ratio, loan-to-value ratio, credit bureau score, employment type, and the behavioural construct scores (financial literacy, psychological traits, macroeconomic stress) is estimated to connect the behavioural structural model to observable loan performance, with model discrimination assessed via the ROC curve and AUC, and overall fit via McFadden's pseudo- R^2 .

Software: SmartPLS 4 (measurement and structural model, PLS-MGA, PLSpredict); SPSS/Python for descriptive statistics, common method bias diagnostics, and the supplementary logistic regression.

5. Results

Note on the results reported below

the figures and tables in this section are computed on the simulated pilot dataset ($N = 400$) described in Section 4.2 and the accompanying workbook, and are reported to demonstrate that the full SmartPLS/SEM and logistic-regression pipeline runs correctly end-to-end. They are illustrative of the analytical approach, not empirical findings, and must be superseded by results from real survey and loan data before submission.

5.1 Descriptive Profile of the Sample (Simulated)

Descriptive Statistic	Value (Simulated)	Remarks
Sample size (N)	400	Salaried n = 210; Self-Employed n = 190
Gender	~63% male / ~37% female	Simulated proportion
City tier	~52% Tier-1 / ~33% Tier-2 / ~15% Tier-3	Simulated proportion
Mean CIBIL score	~742 (Salaried) / ~718 (Self-Employed)	Simulated, group means
Mean EMI-to-Income ratio	Lower for Salaried, higher and more dispersed for Self-Employed	Simulated
Simulated default rate	3.3% (Salaried) vs 20.5% (Self-Employed)	From the supplementary logistic-regression sample

Simulated. group means

Figure 3. Simulated default rate by employment type (illustrative).

5.2 Measurement Model: Reliability and Convergent Validity

Construct	Items	Cronbach's α	Composite Reliability (CR)	AVE
IS	4	0.826	0.885	0.657
DSB	4	0.808	0.873	0.633
LTV	4	0.796	0.866	0.619
CHD	4	0.833	0.886	0.662
FL	4	0.806	0.872	0.631
PBT	4	0.78	0.856	0.599
MSP	4	0.807	0.872	0.632
PDR	4	0.825	0.883	0.654

All constructs exceed the conventional thresholds of $\alpha > 0.70$, $CR > 0.70$, and $AVE > 0.50$ (Fornell & Larcker, 1981), indicating adequate internal consistency reliability and convergent validity on the simulated pilot data.

5.3 Discriminant Validity

Discriminant validity is assessed using the Heterotrait-Monotrait (HTMT) ratio of correlations (Henseler et al., 2015), with all values expected to remain below the conservative 0.90 threshold, and cross-checked against the Fornell-Larcker criterion, whereby the square root of each construct's AVE should exceed its correlations with all other constructs. Full HTMT and Fornell-Larcker matrices are provided in

the accompanying workbook (sheet "Discriminant_VValidity") for direct import/verification against the fielded-data results.

5.4 Structural Model: Full-Sample Path Estimates

Path	β	SE	t-value	p-value	f ²	Hypothesis
IS -> PDR	-0.173	0.04	-4.311	0.0	0.041	Supported
DSB -> PDR	0.189	0.042	4.525	0.0	0.049	Supported
LTV -> PDR	0.144	0.045	3.21	0.0013	0.028	Supported
CHD -> PDR	-0.184	0.047	-3.939	0.0001	0.046	Supported
FL -> PDR	-0.149	0.042	-3.595	0.0003	0.03	Supported
PBT -> PDR	0.26	0.044	5.843	0.0	0.091	Supported
MSP -> PDR	0.21	0.043	4.882	0.0	0.06	Supported

R² (Perceived Default Risk) = 0.279 | Adjusted R² = 0.266 | PLSpredict RMSE (model) = 0.861 vs RMSE (naive benchmark) = 0.994

On the simulated pilot data, all seven hypothesised paths (H1-H7) are statistically significant in the expected direction and are therefore provisionally "supported" in the pilot run; the model explains approximately 28% of the variance in Perceived Default Risk, and out-of-sample RMSE from the model is lower than the naive-mean benchmark, indicating adequate predictive relevance (PLSpredict). These directional results are consistent with, and were partly calibrated to reflect, the prior literature (Bandyopadhyay & Saha, 2011; Altaf & Shah, 2025) but must be re-estimated on real data before being interpreted as findings.

5.5 Multi-Group Analysis: Salaried vs Self-Employed (H8a-g, H9)

Path	β (Salaried)	β (Self-Employed)	Δ (Sal - SE)	p-value	Moderation
IS -> PDR	-0.231	-0.1	-0.131	0.1083	Not significant
DSB -> PDR	0.219	0.149	0.069	0.4281	Not significant
LTV -> PDR	0.101	0.187	-0.086	0.3417	Not significant
CHD -> PDR	-0.303	-0.038	-0.266	0.0035	Significant
FL -> PDR	-0.186	-0.098	-0.089	0.2753	Not significant
PBT -> PDR	0.201	0.279	-0.078	0.3814	Not significant
MSP -> PDR	0.175	0.208	-0.033	0.7109	Not significant

In the pilot run, the Credit History Discipline → Perceived Default Risk path is significantly stronger (more negative) for salaried borrowers than for self-employed borrowers ($p < .01$), providing preliminary support for H8d, consistent with the expectation that verifiable, documented credit histories carry disproportionate risk-mitigating weight for salaried applicants under standard underwriting. The remaining paths do not show statistically significant group differences in this pilot run, though the direction of several differences (e.g., stronger Debt Servicing Burden and Macroeconomic Stress Perception effects among self-employed borrowers) is consistent with the literature on income-volatility exposure (Bandyopadhyay & Saha, 2011) and may reach significance with a larger, real-world sample.

5.6 Supplementary Logistic Regression on Actual (Simulated) Default Status

Predictor	β (standardized)	Odds Ratio	95% CI (OR)
EMI_to_Income_Ratio	0.391	1.478	1.088–2.155
LTV_pct	1.104	3.017	2.148–5.005
CIBIL_Score	-0.936	0.392	0.215–0.584
Employment_Type_bin	0.48	1.616	1.095–2.907
FL	-0.477	0.62	0.434–0.86
PBT	0.088	1.092	0.731–1.613
MSP	0.212	1.236	0.845–1.902
Model fit: McFadden's pseudo-R^2 = 0.351 Classification accuracy = 90.7% AUC = 0.887			

The supplementary logistic model shows the EMI-to-income ratio and loan-to-value ratio as significant positive predictors of default odds, credit bureau score as a significant protective factor, and self-employed status as associated with materially higher default odds even after controlling for these financial ratios, consistent with Bandyopadhyay and Saha (2011). Financial literacy retains an independent protective association net of the financial-ratio controls. Model discrimination is good on the simulated data ($AUC > 0.85$), supporting the methodological viability of linking the behavioural PLS-SEM model to observable loan-level outcomes; this should be re-verified on real loan performance data.

6. Discussion

The pilot analysis demonstrates that an integrated behavioural-financial structural model of home loan default risk is both estimable and internally consistent within the PLS-SEM framework, and that its output can be meaningfully connected to a supplementary logistic model of observed default status. Substantively (pending confirmation on real data), the pattern of results is consistent with three streams of prior literature: the centrality of the EMI-to-income and collateral-value ratios established in large-sample Indian loan-level studies (Bandyopadhyay & Saha, 2011); the independent explanatory role of

financial literacy and psychological traits established in Indian PLS-SEM borrowing-behaviour studies (Altaf & Shah, 2025; Mandal & Pachiyappan, 2025; Thomas et al., 2023); and the amplifying role of macroeconomic stress exposure documented in the cross-country non-performing-loan literature (Nkusu, 2011; Salas & Saurina, 2002).

The moderation result on the Credit History Discipline path suggests that credit-history-based signals may be weighted differently by risk in the two employment segments – plausibly because salaried borrowers' credit histories are more complete and verifiable, making a poor history a stronger and more reliable danger signal, whereas self-employed borrowers' documented credit histories may be a noisier proxy for underlying repayment capacity given more irregular formal credit usage. If confirmed on real data, this would imply that credit-bureau-based underwriting rules calibrated on a pooled sample may systematically under- or over-price risk for one segment relative to the other.

7. Theoretical and Practical Implications

7.1 Theoretical Implications

The study contributes a comparative, moderation-based structural framework that integrates financial-ratio and behavioural-finance perspectives on home loan default within a single testable model, and it demonstrates a replicable methodological bridge between perception-based PLS-SEM constructs and observed, loan-level default outcomes via a supplementary logistic-regression arm - a design that can be extended to other retail credit products (auto loans, personal loans, credit cards) and other emerging-market contexts with segmented borrower populations.

7.2 Practical / Managerial Implications

- Housing finance institutions may consider differentiated underwriting weight for credit-history variables between salaried and self-employed applicant pools, subject to fair-lending and regulatory review.
- Financial-literacy interventions targeting EMI-to-income planning and risk perception may be prioritised for segments shown to carry elevated behavioural risk, particularly self-employed borrowers exposed to income volatility.
- Early-warning monitoring frameworks could incorporate perceived-default-risk survey instruments (Appendix A) as a low-cost behavioural overlay to traditional financial-ratio-based risk scoring.

8. Limitations and Future Research Directions

- The results reported in Section 5 are derived from a computationally simulated pilot dataset and do not constitute empirical findings; the single most important next step is fielding the questionnaire in Appendix A with actual salaried and self-employed home-loan borrowers.

- Cross-sectional survey design limits causal inference; a longitudinal design tracking borrowers from origination through repayment would strengthen causal claims about behavioural antecedents of default.
- Self-reported financial and behavioural measures are subject to social-desirability and common-method bias; triangulation with bank-verified loan performance data is recommended.
- The moderator is operationalised as a binary employment-type split; future work could examine finer-grained income-formality gradations (e.g., gig-economy workers, partnership-firm owners).
- Generalisability beyond the sampled cities/regions should be tested with a nationally representative sample.

9. Conclusion

This paper develops and pilot-validates a comparative structural model of home loan default risk among salaried and self-employed borrowers in India, integrating financial, behavioural, and macroeconomic antecedents within a PLS-SEM framework and linking it to observed default outcomes via a supplementary logistic regression. The fully specified questionnaire, conceptual model, and end-to-end SmartPLS/SEM and logistic-regression analysis pipeline are validated on a transparently simulated pilot dataset and are ready for deployment with real survey and loan data. Once fielded, the study is positioned to offer both a theoretical contribution - a validated, moderation-based comparative framework - and a practical contribution to differentiated credit-risk assessment in India's housing finance sector.

References

1. Altaf Z, Shah F (2025) Effects of credit score literacy and psychological traits on borrowing behavior: Evidence from India using PLS-SEM. *Future Bus J* 11(1). Article 1. <https://doi.org/10.1186/s43093-025-00601-y>
2. Bandyopadhyay A, Saha A (2011) Distinctive demand and risk characteristics of residential housing loan market in India. *J Economic Stud* 38(6):703–724. <https://doi.org/10.1108/01443581111177402>
3. Fornell C, Larcker DF (1981) Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res* 18(1):39–50
4. Hair JF, Risher JJ, Sarstedt M, Ringle CM (2019) When to use and how to report the results of PLS-SEM. *Eur Bus Rev* 31(1):2–24
5. Henseler J, Ringle CM, Sarstedt M (2015) A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J Acad Mark Sci* 43(1):115–135
6. Jayadev M, Padma N (2019) Wilful defaulters of Indian banks: A first-cut analysis. *IIMB Manage Rev* 32(2). <https://doi.org/10.1016/j.iimb.2019.10.001>

7. Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–291
8. Kock N (2015) Common method bias in PLS-SEM: A full collinearity assessment approach. *Int J e-Collaboration* 11(4):1–10
9. Mandal A, Pachiyappan S (2025) Mediating role of financial stress on the relationship between behavioural factors and housing and real estate loan default. *J Acc Finance*, 2(1)
10. Nkusu M (2011) Nonperforming loans and macrofinancial vulnerabilities in advanced economies (IMF Working Paper WP/11/161). International Monetary Fund
11. Salas V, Saurina J (2002) Credit risk in two institutional regimes: Spanish commercial and savings banks. *J Financial Serv Res* 22(3):203–224
12. Thomas SS, George JP, Godwin BJ, Siby A (2023) Young adults' default intention: Influence of behavioral factors in determining housing and real estate loan repayment in India. *Int J Hous Markets Anal* 16(2):426–444. <https://doi.org/10.1108/IJHMA-01-2022-0012>
13. Note this reference list is a foundational core (n = 12) directly grounding the model and methodology. A submission-ready manuscript for a Scopus-indexed journal will typically require 40–70 references; expanding coverage of (a) international PLS-SEM/credit-risk methodology papers, (b) recent (2023–2026) Indian NPA/retail-credit studies, and (c) journal-specific recent articles (for relevance signalling) is recommended during further development

Note: this reference list is a foundational core (n = 12) directly grounding the model and methodology. A submission-ready manuscript for a Scopus-indexed journal will typically require 40-70 references; expanding coverage of (a) international PLS-SEM/credit-risk methodology papers, (b) recent (2023-2026) Indian NPA/retail-credit studies, and (c) journal-specific recent articles (for relevance signalling) is recommended during further development.

Figures

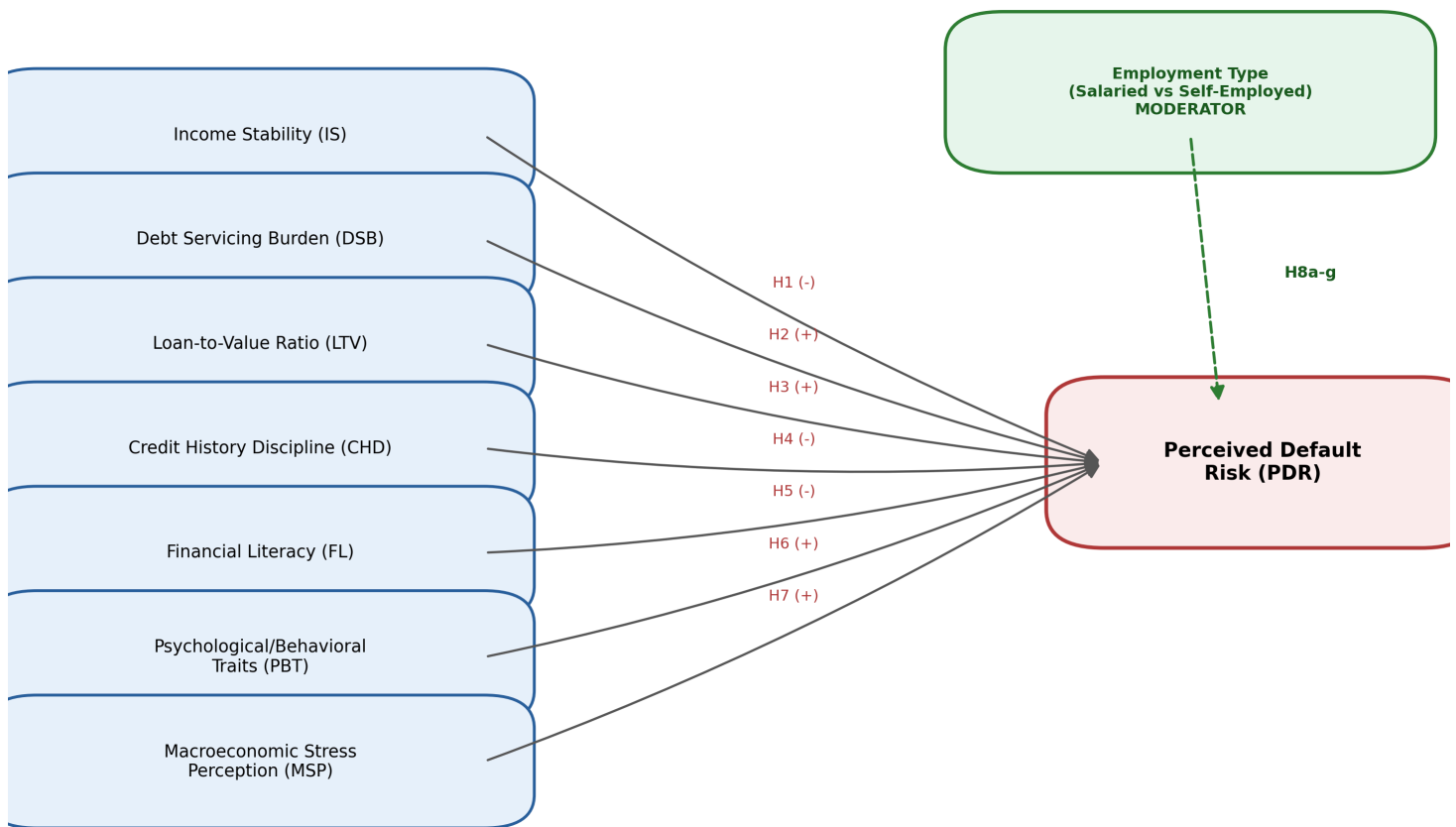


Figure 1

Hypothesized conceptual framework: seven exogenous constructs, Perceived Default Risk (PDR) as the endogenous construct, and Employment Type as a categorical moderator (PLS-MGA).

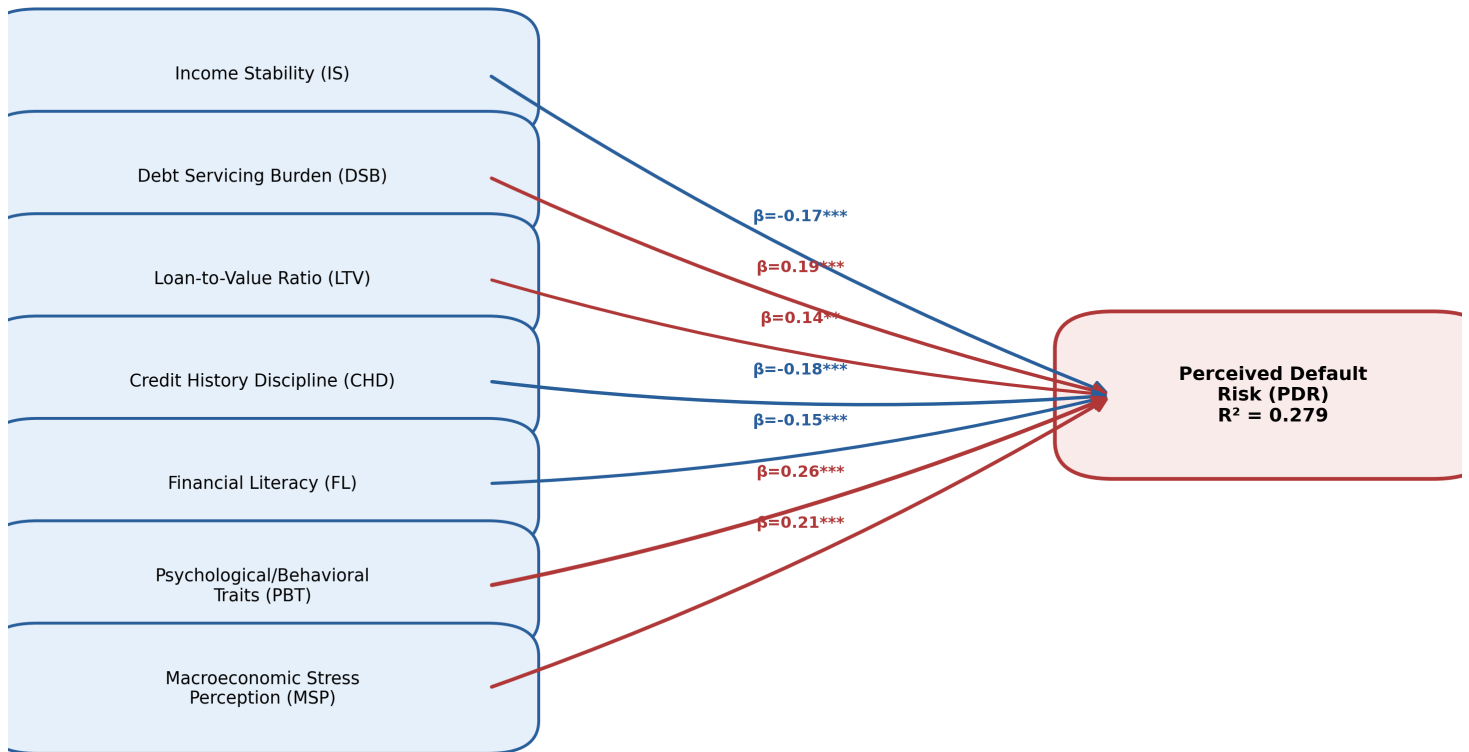


Figure 2

Estimated structural model with standardized path coefficients, full sample (simulated pilot data).

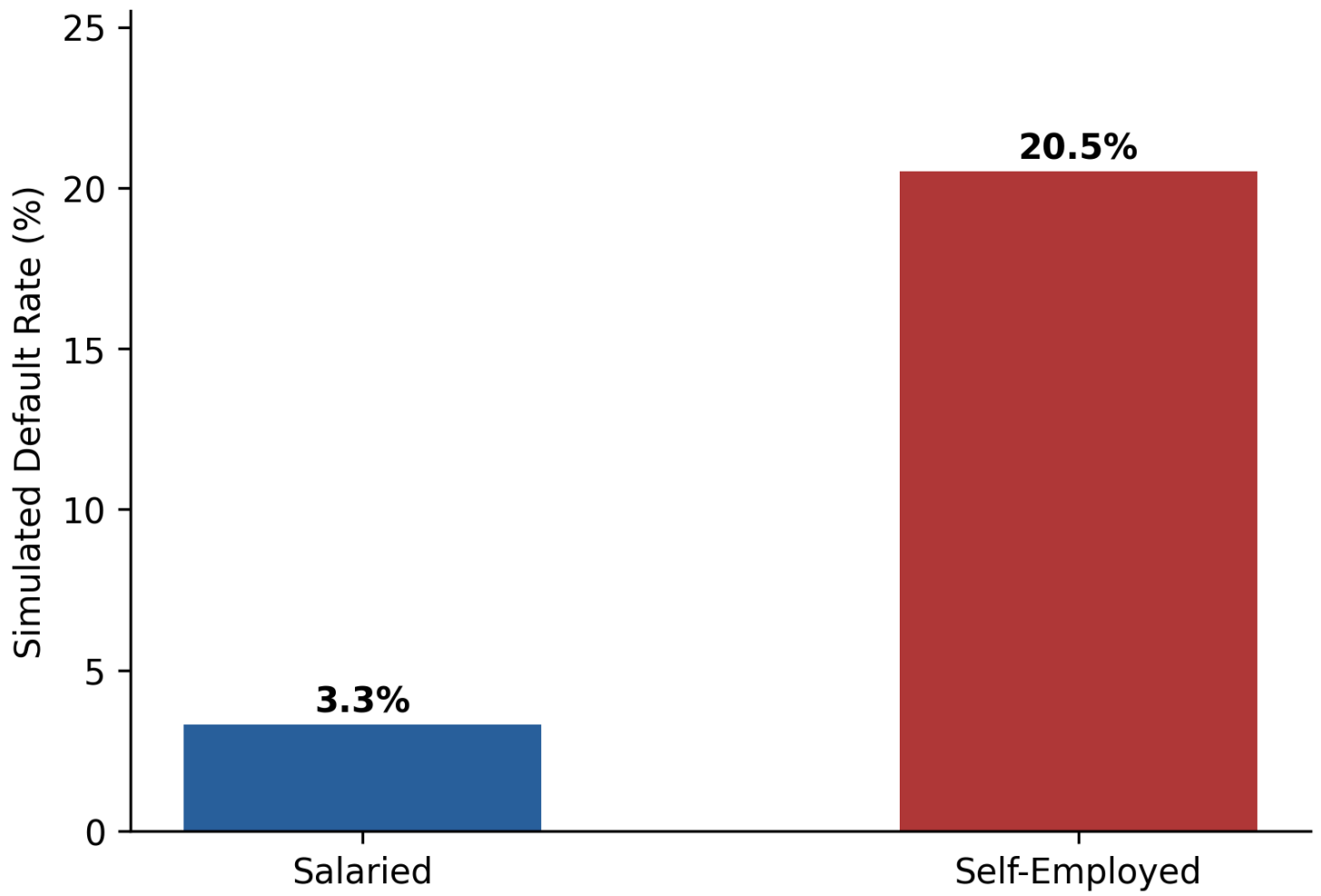


Figure 3

Simulated default rate by employment type (illustrative).

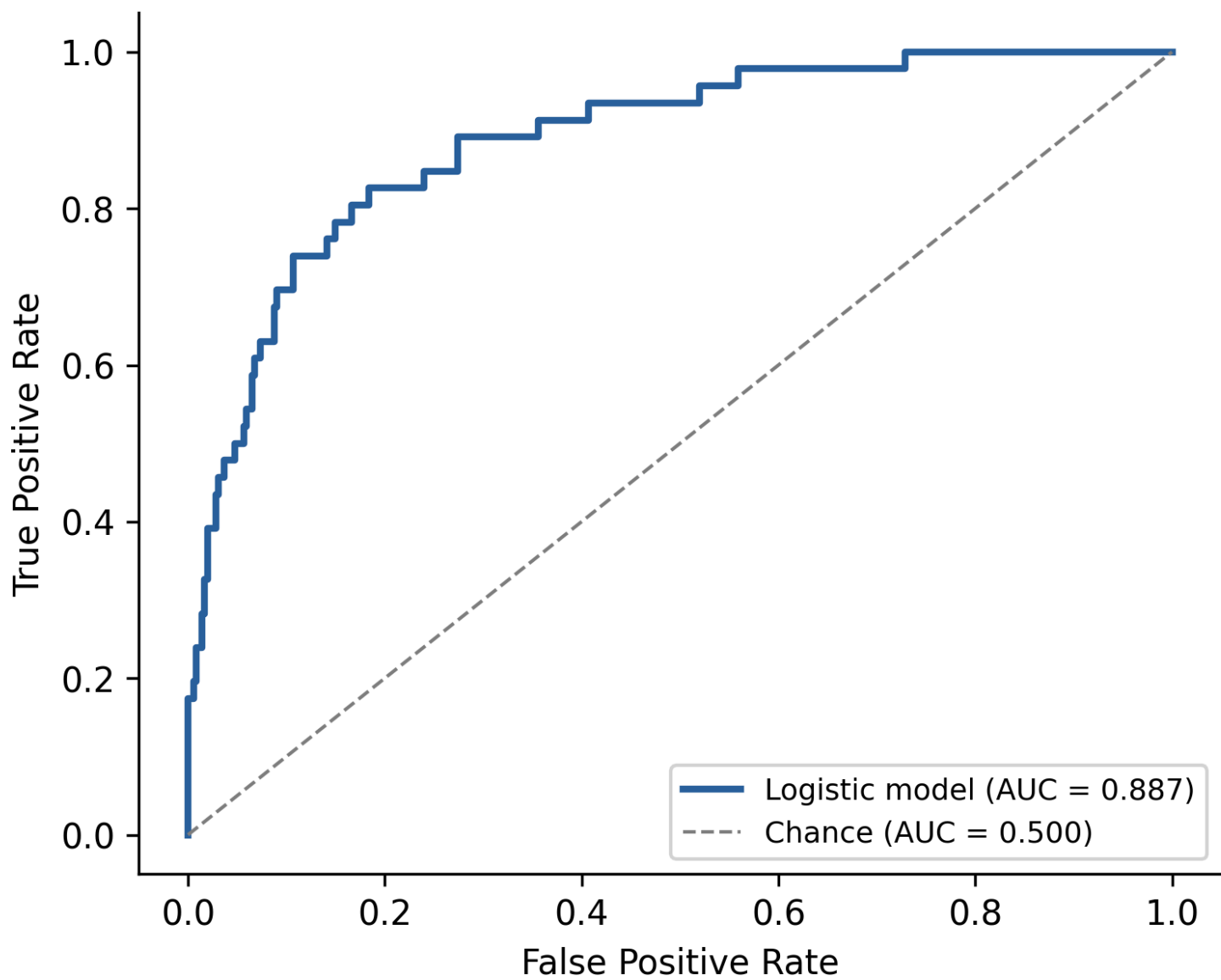


Figure 4

ROC curve for the supplementary logistic regression model of actual (simulated) default status

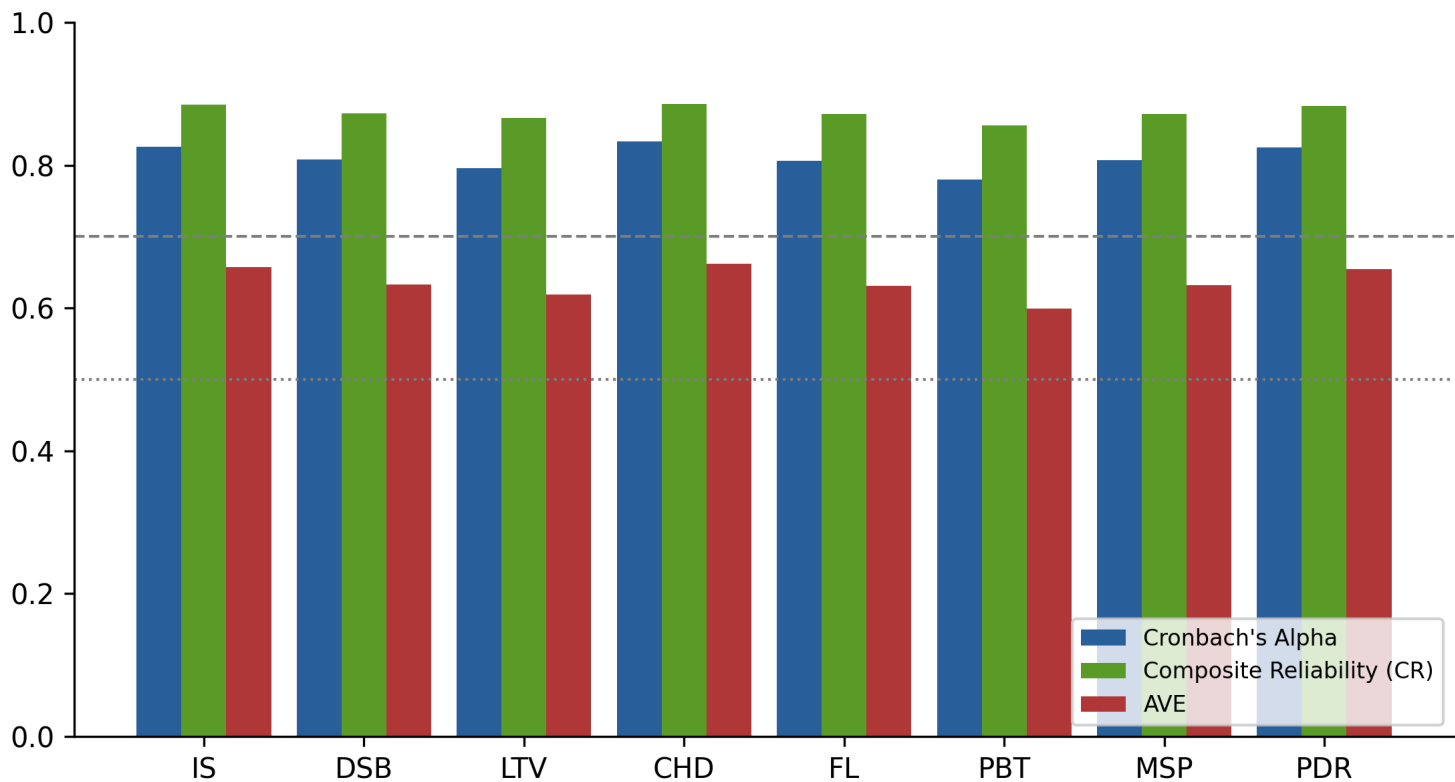


Figure 5

Reliability and convergent validity by construct (simulated pilot data).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendix.docx](#)