MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages4009-4016

Exploring the Impact of Information and Communication Technology on Financial Inclusion: An Interstate Analysis for India Using Spatial-Temporal Econometric Modeling and Ensemble Machine Learning

Devashish Kushwaha[1] & Dipti Malpani[2]
[1] Department of Management, Oriental University, Indore, India
[2] Department of Management, Commerce and Media Studies, Oriental University, Indore, India

Abstract
This study investigates the heterogeneous impact of Information and Communication Technology (ICT) infrastructure and adoption on multidimensional financial inclusion across 28 Indian states and union territories during 2015–2023. Employing a novel methodological triangulation comprising Spatial Durbin Model (SDM), System Generalized Method of Moments (Sys-GMM) dynamic panel estimation, and ensemble machine learning techniques (Random Forest, Gradient Boosting, and XGBoost), we construct a composite Financial Inclusion Index (FII) incorporating penetration, availability, and usage dimensions. Results reveal significant positive spatial spillover effects ($\rho = 0.342$, $p < .01$) indicating financial inclusion convergence across contiguous states. ICT infrastructure intensity demonstrates a non-linear relationship with financial inclusion ($\beta_1 = 0.487$, $\beta_2 = -0.052$, $p < .01$), suggesting diminishing marginal returns beyond threshold levels. Mobile internet penetration emerges as the strongest predictor (importance score = 0.284) in ensemble models, surpassing traditional banking infrastructure. Spatial heterogeneity analysis reveals distinct ICT–financial inclusion nexuses across high-income (Karnataka, Maharashtra) versus low-income states (Bihar, Jharkhand), with coefficient variations ranging from 0.321 to 0.685. The study contributes theoretically by extending the Unified Theory of Acceptance and Use of Technology (UTAUT) to macro-level institutional contexts and methodologically through hybrid econometric–algorithmic approaches. Policy implications emphasize targeted ICT infrastructure investments in spatially lagging regions and addressing digital literacy gaps to maximize financial inclusion dividends.
Keywords: Financial Inclusion, Information and Communication Technology, Spatial Econometrics, Dynamic Panel Data, Ensemble Machine Learning, India

1. Introduction

Financial inclusion, defined as ensuring affordable access to formal financial services for all segments of society, has emerged as a critical policy imperative in developing economies (Demirgüç-Kunt et al., 2022). Despite significant progress through initiatives like Pradhan Mantri Jan Dhan Yojana (PMJDY), approximately 190 million Indians remained excluded from formal banking channels as of 2023, with pronounced interstate disparities (Reserve Bank of India, 2023). The confluence of Information and Communication Technology (ICT) and financial services—manifested through mobile banking, digital payment platforms, and fintech innovations—presents transformative potential for addressing persistent inclusion gaps (Ozili, 2021; Senyo & Osabutey, 2020).

Existing literature establishes ICT as a key enabler of financial inclusion through reduced transaction costs, enhanced geographic reach, and improved information asymmetries (Asongu & Odhiambo, 2019; Tchamyou, 2021). However, three critical gaps persist in current scholarship. First, most studies employ aggregated national-level analyses, obscuring substantial spatial heterogeneity across Indian states with varying developmental trajectories (Sethy & Goyari, 2022). Second, methodological approaches predominantly rely on traditional linear panel estimators, failing to account for spatial dependencies and dynamic adjustments inherent in financial inclusion processes (Sarma & Pais, 2011). Third, limited research integrates econometric rigor with predictive machine learning algorithms to capture non-linear ICT–financial inclusion relationships (Mullainathan & Spiess, 2017).This study addresses these lacunae through a comprehensive interstate analysis of 28 Indian states and union territories spanning 2015–2023, employing three complementary methodological innovations. First, we develop a multidimensional Financial Inclusion Index (FII) synthesizing 15 indicators across penetration (account ownership), availability (banking infrastructure), and usage (transaction volumes) dimensions. Second, we implement Spatial Durbin Models (SDM) to quantify direct and spillover effects while controlling for spatial autocorrelation, complemented by System-GMM dynamic panel estimators addressing endogeneity concerns (Blundell & Bond, 1998; Elhorst, 2014). Third, we deploy ensemble machine learning algorithms (Random Forest, Gradient Boosting, XGBoost) to identify non-linear patterns and variable importance hierarchies (Breiman, 2001a).

Our theoretical framework extends the Unified Theory of Acceptance and Use of Technology (UTAUT) from individual-level adoption to macro-institutional contexts, proposing that state-level ICT infrastructure and enabling conditions moderate the performance expectancy–financial inclusion nexus (Venkatesh et al., 2003). We posit spatial demonstration effects wherein neighboring states' ICT-driven financial inclusion initiatives generate knowledge spillovers and competitive diffusion pressures (Rogers, 2003).

Empirically, we construct a novel panel dataset integrating Reserve Bank of India (RBI) financial inclusion metrics, Telecom Regulatory Authority of India (TRAI) ICT statistics, and state domestic product accounts. Preliminary descriptive analysis reveals stark interstate variations: while Kerala and Goa demonstrate FII scores exceeding 0.78, Bihar and Uttar Pradesh lag below 0.42, despite comparable ICT penetration rates in urban centers—suggesting complex mediation mechanisms.

This study makes four principal contributions. Theoretically, it develops a spatial-institutional extension of UTAUT applicable to macro-level technology diffusion contexts. Methodologically, it pioneers hybrid econometric–algorithmic approaches for financial inclusion research, demonstrating complementarities between causal inference and predictive modeling. Empirically, it provides granular evidence on interstate heterogeneities, threshold effects, and spillover dynamics. From a policy perspective, findings inform spatially-targeted ICT investment strategies and identify mechanisms for accelerating convergence across lagging regions.

The remainder of this paper proceeds as follows. Section 2 reviews theoretical foundations and empirical literature. Section 3 outlines the multidimensional financial inclusion framework and data sources. Section 4 details spatial econometric specifications and machine learning architectures. Section 5 presents empirical results including spatial diagnostics, estimation outputs, and ensemble model predictions. Section 6 discusses theoretical implications, policy recommendations, and limitations. Section 7 concludes.

2. Theoretical Framework and Literature Review

2.1 Conceptual Foundations: ICT and Financial Inclusion Nexus: The theoretical relationship between ICT and financial inclusion operates through multiple transmission channels. Transaction cost economics posits that digital infrastructure reduces fixed costs associated with establishing physical banking presence in remote areas, enabling expanded geographic coverage (Williamson, 1981). Information asymmetry theory suggests ICT platforms facilitate credit assessment through alternative data sources (mobile usage patterns, digital footprints), particularly benefiting unbanked populations lacking traditional collateral (Stiglitz & Weiss, 1981).

Network effects theory provides additional explanatory power: as digital payment adoption reaches critical mass, merchant acceptance expands, creating positive feedback loops that accelerate financial inclusion (Katz & Shapiro, 1985). This aligns with spatial diffusion models wherein technology adoption exhibits spatial autocorrelation—neighboring regions demonstrate correlated adoption patterns due to knowledge spillovers, demonstration effects, and labor mobility (Tobler, 1970).

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages4009-4016

2.2 Technology Adoption at Macro-Institutional Levels: While Technology Acceptance Model (TAM) and UTAUT frameworks dominate individual-level adoption research, their macro-institutional extensions remain undertheorized (Davis, 1989; Dwivedi et al., 2019). We propose that at state-level analysis, UTAUT constructs translate as follows: *Performance Expectancy* manifests as aggregate economic returns from digital financial services; *Effort Expectancy* reflects population-wide digital literacy and infrastructure accessibility; *Social Influence* captures inter-state demonstration effects and policy emulation; *Facilitating Conditions* encompass regulatory frameworks, telecommunications infrastructure, and public–private partnerships. Recent scholarship emphasizes spatial interdependencies in technology diffusion (Caragliu & Del Bo, 2019). Utilizing spatial econometric frameworks, researchers document significant spillover effects in financial development, suggesting cross-border knowledge transmission and competitive dynamics (Tchamyou, 2019). However, application to ICT–financial inclusion contexts, particularly in heterogeneous developing country settings, remains limited.

2.3 Empirical Evidence: Global and Indian Contexts: Cross-country studies establish positive associations between ICT penetration and financial inclusion. FinTech technology survey during 2011–2014 documents spatial dependencies in financial inclusion levels, with internet usage and mobile subscriptions emerging as significant predictors (Gai et al., 2018). Panel data regression across 84 countries (2012–2020) reveals that digital lending exhibits context-dependent effects—positive in developing nations but statistically insignificant in aggregate samples (Aldasoro et al., 2024). India-specific research demonstrates mobile banking and Aadhaar-linked direct benefit transfers as pivotal drivers of inclusion during the Digital India era (Kaur et al., 2020; Khera, 2019). However, interstate heterogeneity analyses remain scarce. Spatial panel studies of Indian states identify regional convergence in financial development but predate widespread digital payment adoption (Sethy, 2016).

Methodologically, emerging literature advocates ensemble machine learning for financial prediction tasks. Random Forest, Gradient Boosting, and XGBoost algorithms demonstrate superior predictive accuracy compared to traditional econometric approaches in corporate distress forecasting and credit scoring applications (Chen & Guestrin, 2016; Friedman, 2001). Their capacity to capture non-linearities and interaction effects positions them as complementary tools for financial inclusion research, yet integration with causal spatial econometric frameworks remains nascent.

2.4 Research Gaps and Hypotheses: Synthesizing extant literature, we identify three critical gaps. First, limited spatial econometric analysis of interstate ICT–financial inclusion dynamics within India. Second, absence of hybrid methodologies combining causal inference (spatial models, dynamic panels) with predictive algorithms (ensemble learning). Third, insufficient attention to threshold effects and non-linearities in ICT–financial inclusion relationships.

Accordingly, we advance the following hypotheses:

H1: ICT infrastructure intensity positively influences financial inclusion, exhibiting non-linear (inverted U-shaped) relationships due to diminishing marginal returns.

H2: Financial inclusion demonstrates positive spatial autocorrelation, with significant spillover effects from neighboring states' ICT investments.

H3: The ICT–financial inclusion nexus varies systematically across states based on initial development levels, with stronger effects in lagging regions.

H4: Mobile internet penetration exerts stronger influence on financial inclusion than fixed broadband or traditional banking infrastructure, reflecting mobile-first digital adoption patterns.

3. Data and Variable Construction

3.1 Sample and Data Sources

We construct a balanced panel dataset spanning 28 Indian states and union territories over nine years (2015–2023), yielding 252 state-year observations. Data integration from multiple authoritative sources ensures comprehensiveness:

- **Financial Inclusion Metrics:** Reserve Bank of India (RBI) Financial Inclusion Reports, providing state-wise data on bank accounts, ATMs, branches, credit penetration, and digital transaction volumes
- **ICT Infrastructure:** Telecom Regulatory Authority of India (TRAI) annual reports detailing mobile subscriptions, internet penetration, broadband connections, and telecom tower density
- **Economic Controls:** Ministry of Statistics and Programme Implementation (MoSPI) state domestic product accounts, literacy rates, urbanization ratios, and industrial composition
- **Spatial Weights:** Geographic Information System (GIS) coordinates for constructing contiguity-based spatial weight matrices

3.2 Financial Inclusion Index Construction

Following established multidimensional indexing methodologies (Sarma, 2012), we construct a composite Financial Inclusion Index (FII) aggregating 15 indicators across three dimensions:

Dimension 1: Penetration (40% weight)

- Bank accounts per 1,000 adults
- Credit accounts per 1,000 adults
- Deposit accounts per 1,000 adults
- Insurance policy holders per 1,000 adults
- Mobile banking registered users per 1,000 adults

Dimension 2: Availability (30% weight)

- Bank branches per 100,000 population
- ATMs per 100,000 population
- Banking correspondents per 100,000 population
- Microfinance institution density
- Rural bank branch proportion

Dimension 3: Usage (30% weight)

- Average deposit amount per account
- Average credit amount per account
- Digital transaction volume per capita
- Credit-deposit ratio
- Active account ratio (transactions in last 90 days)

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages4009-4016

Each indicator is normalized using min-max scaling: $X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$. Dimension scores aggregate normalized indicators through weighted averaging, and the composite FII combines dimension scores. Mathematically:

$$FII_{it} = 0.40 \times Penetration_{it} + 0.30 \times Availability_{it} + 0.30 \times Usage_{it}$$

where subscripts denote state i and year t. The FII ranges from 0 (complete exclusion) to 1 (universal inclusion).

3.3 Independent Variables: ICT Infrastructure

Primary ICT Variables:

- **Mobile_Penetration:** Mobile phone subscriptions per 100 population
- **Internet_Penetration:** Internet subscribers per 100 population
- **Broadband_Density:** Fixed broadband connections per 100 population
- **Telecom_Infrastructure:** Telecom towers per 1,000 sq. km
- **Data_Usage:** Average monthly data consumption per subscriber (GB)

Composite ICT Infrastructure Index (ICTII):

Similar to FII construction, we develop ICTII aggregating the five ICT metrics using principal component analysis (PCA) weights:

$$ICTII_{it} = \sum_{j=1}^{5} w_j \times ICT_{jit}$$

where w_j represents the first principal component loading for indicator j.

3.4 Control Variables

To isolate ICT effects, we control for established financial inclusion determinants:

- **GSDP_PC:** Gross State Domestic Product per capita (log-transformed)
- **Literacy:** Adult literacy rate (%)
- **Urbanization:** Urban population share (%)
- **Bank_Competition:** Herfindahl-Hirschman Index for banking sector concentration
- **Road_Density:** Road network per 1,000 sq. km (infrastructure proxy)
- **Youth_Ratio:** Population aged 15–35 as proportion of total population

3.5 Descriptive Statistics

Table 1 presents summary statistics for key variables across the full sample and stratified by development quintiles.

Table 1: *Descriptive Statistics of Key Variables (2015–2023)*

Variable	Mean	SD	Min	Max	Q1	Q5
FII	0.586	0.148	0.312	0.834	0.423	0.782
ICTII	0.524	0.172	0.187	0.891	0.298	0.856
Mobile Penetration	78.34	16.52	42.11	118.45	56.23	106.78
Internet Penetration	44.67	18.29	12.34	89.76	19.45	78.92
Broadband Density	3.82	2.94	0.45	14.23	1.12	11.87
GSDP per capita (₹ 000s)	148.7	89.3	52.1	456.8	68.4	389.2
Literacy Rate (%)	76.4	8.7	61.2	94.6	68.5	91.3
Urbanization (%)	35.8	15.6	14.2	97.5	21.4	88.7
Digital Transactions (per capita)	18.45	12.67	2.34	67.89	4.23	58.34

Note. N = 252 (28 states × 9 years). Q1 = lowest development quintile mean; Q5 = highest development quintile mean.

Table 2: *Interstate Variation in FII and ICTII (2023 Cross-Section)*

State Group	FII Mean	ICTII Mean	Correlation
High Income (n = 6)	0.756	0.812	.847***
Upper Middle Income (n = 7)	0.642	0.634	.723***
Lower Middle Income (n = 8)	0.534	0.468	.691***
Low Income (n = 7)	0.412	0.312	.564***
Overall Sample (n = 28)	**0.586**	**0.524**	**.782*** **

Note. **p < .01. State groups based on GSDP per capita quartiles.

Notable patterns emerge: substantial cross-state heterogeneity exists in both FII (coefficient of variation = 0.253) and ICTII (CV = 0.328). High-income states demonstrate nearly double the financial inclusion levels of low-income counterparts. Positive correlations between ICT infrastructure and financial inclusion strengthen with development levels, suggesting complementarities with enabling infrastructure.

Figure 1 visualizes the geographic distribution of FII scores across Indian states in 2023, revealing pronounced regional clustering. Southern states (Karnataka, Kerala, Tamil Nadu) and western states (Maharashtra, Gujarat) form high-inclusion clusters, while eastern and central states exhibit spatial concentration of low-inclusion regions—preliminary evidence for spatial autocorrelation.

4. Empirical Methodology: Our empirical strategy employs three complementary methodological approaches: spatial econometric models for causal inference and spillover quantification, dynamic panel estimators for endogeneity correction, and ensemble machine learning for non-linear pattern detection and predictive accuracy.

4.1 Spatial Econometric Framework

4.1.1 Spatial Autocorrelation Diagnostics

We first test for spatial dependence using Global Moran's I statistic:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2}$$

where w_{ij} represents spatial weight between states i and j, y_i denotes FII for state i, and N is the number of states. We employ queen contiguity spatial weights (sharing common borders or vertices) and row-standardized inverse distance weights for robustness.

Local spatial autocorrelation is assessed through Local Indicators of Spatial Association (LISA), identifying high-high clusters (spatial hotspots), low-low clusters (cold spots), and spatial outliers (Anselin, 1988).

4.1.2 Spatial Durbin Model Specification

Given confirmed spatial dependence, we estimate the Spatial Durbin Model (SDM), which incorporates both spatially lagged dependent variables and spatially lagged independent variables (Elhorst, 2014):

$$FII_{it} = \rho \sum_j w_{ij} FII_{jt} + \beta_1 ICTII_{it} + \beta_2 ICTII_{it}^2 + \theta \sum_j w_{ij} ICTII_{jt} + \gamma' X_{it} + \delta \sum_j w_{ij} X_{jt} + \mu_i + \lambda_t + \varepsilon_{it}$$

where:

- ρ captures spatial autoregressive parameter (spillover intensity)
- β_1, β_2 represent direct effects of own-state ICT (linear and quadratic terms test H1)
- θ measures spatial spillover from neighboring states' ICT infrastructure
- X_{it} denotes control variable vector with coefficient vector γ
- μ_i and λ_t represent state and year fixed effects
- ε_{it} is the error term

SDM allows decomposition of total effects into direct (own-state) and indirect (spillover) components, addressing H2 (Vega & Elhorst, 2015).

4.1.3 Effect Decomposition

Following LeSage and Pace (2009), we compute partial derivative matrices to decompose total marginal effects:

$$\frac{\partial E(FII)}{\partial ICTII} = (I_N - \rho W)^{-1}(\beta_1 I_N + \theta W)$$

Direct effects represent average impacts of own-state ICT changes on own-state financial inclusion. Indirect effects capture average spillover impacts on neighboring states. Total effects sum both components (LeSage & Pace, 2014).

4.2 Dynamic Panel Data: System GMM

To address potential endogeneity from reverse causality (financial inclusion may stimulate ICT adoption) and unobserved heterogeneity, we employ System Generalized Method of Moments (Sys-GMM; Arellano & Bond, 1991; Blundell & Bond, 1998):

$$FII_{it} = \alpha FII_{i,t-1} + \beta_1 ICTII_{it} + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

Sys-GMM utilizes lagged levels and differences as instruments, addressing: (a) dynamic panel bias from lagged dependent variables; (b) endogeneity of regressors; (c) unobserved state-specific heterogeneity; and (d) heteroskedasticity and autocorrelation.

Validity depends on two specification tests:

- **Arellano-Bond AR(2) test:** Tests for second-order serial correlation in first-differenced residuals (null: no autocorrelation)
- **Hansen J-test:** Tests overidentifying restrictions (null: instruments valid)

The implementation follows Roodman's (2009) recommended guidelines for avoiding instrument proliferation.

4.3 Ensemble Machine Learning Framework

4.3.1 Algorithm Selection Rationale

While spatial econometric models provide causal interpretation and hypothesis testing, machine learning algorithms excel at capturing non-linear relationships and complex interactions without imposing functional form assumptions (Mullainathan & Spiess, 2017). We implement three ensemble methods:

Random Forest (RF): Bootstrap aggregation of decision trees, reducing overfitting through averaging (Breiman, 2001a):

- 500 trees with maximum depth = 15
- Minimum samples per leaf = 10
- Feature importance via mean decrease in impurity

Gradient Boosting Machine (GBM): Sequential tree construction minimizing residual errors (Friedman, 2001):

- Learning rate = 0.05; 300 boosting iterations
- Maximum depth = 8; feature importance via gain metric

Extreme Gradient Boosting (XGBoost): Regularized boosting with parallel processing (Chen et al., 2015):

- L1 (alpha = 0.1) and L2 (lambda = 0.5) regularization
- Learning rate = 0.03; 500 estimators
- Feature importance via SHAP values (Lundberg & Lee, 2017)

4.3.2 Model Training and Validation: We partition data into training (70%, 2015–2020), validation (15%, 2021), and testing (15%, 2022–2023) sets, maintaining temporal sequence to prevent data leakage. Hyperparameters are optimized through 5-fold time-series cross-validation on training data, minimizing root mean squared error (RMSE). Model selection and validation procedures follow best practices outlined in Hastie et al. (2009).

4.3.3 Feature Importance and Partial Dependence: To interpret black-box models, we compute: (a) **Permutation Importance**, measuring prediction accuracy decrease when feature values are randomly permuted; (b) **SHAP Values** (Lundberg & Lee, 2017), quantifying each feature's contribution to individual predictions based on cooperative game theory; and (c) **Partial Dependence Plots**, visualizing marginal effects of ICT variables on predicted FII holding other features constant.

4.4 Heterogeneity Analysis

To test H3, we estimate state-specific coefficients through: (a) **Quantile Regression**, examining ICT effects across the FII distribution (10th, 25th, 50th, 75th, 90th percentiles); (b) **Interaction Models**, including interaction terms between ICTII and development level indicators; and (c) **Cluster-Specific Estimations**, separate regressions for high/low development state groups.

5. Empirical Results

5.1 Spatial Autocorrelation Analysis

Table 3 reports Global Moran's *I* statistics testing spatial dependence in financial inclusion.

Table 3: *Spatial Autocorrelation Tests for Financial Inclusion Index*

Year	Moran's *I*	*Z*-score	*p*-value	Interpretation
2015	0.287	3.42	.001***	Positive SA
2017	0.314	3.78	<.001***	Positive SA
2019	0.356	4.21	<.001***	Positive SA
2021	0.389	4.67	<.001***	Positive SA
2023	0.412	4.89	<.001***	Positive SA

Note. Queen contiguity spatial weights matrix. ***p* < .01. SA = Spatial Autocorrelation.

Results strongly reject the null hypothesis of spatial randomness across all years ($p < .01$), confirming significant positive spatial autocorrelation in FII. Moran's *I* values increase over time (0.287 to 0.412), suggesting intensifying spatial clustering—states with high financial inclusion increasingly neighbor similar states, while low-inclusion regions form persistent clusters. This provides empirical support for H2 and justifies spatial econometric modeling (Anselin, 1988; Elhorst, 2014).LISA cluster maps identify distinct spatial regimes: High-High clusters in southwestern regions (Karnataka, Maharashtra, Goa) and Low-Low clusters in eastern/central India (Bihar, Jharkhand, Madhya Pradesh, Uttar Pradesh). These patterns persist throughout the analysis period, indicating spatial poverty traps in financial inclusion.

5.2 Spatial Durbin Model Results

Table 4 presents SDM estimation results with varying specifications.

Table 4: *Spatial Durbin Model Estimates: ICT Impact on Financial Inclusion*

Variables	(1) Baseline	(2) Quadratic	(3) Full Controls	(4) Robust
Direct Effects				
ICTII	0.521***	0.487***	0.468***	0.452***
	(0.067)	(0.071)	(0.074)	(0.078)
ICTII²	—	-0.052**	-0.048**	-0.051**
		(0.021)	(0.019)	(0.023)
GSDP per capita (log)	—	—	0.134***	0.128***
			(0.032)	(0.035)
Literacy Rate	—	—	0.003**	0.002*
			(0.001)	(0.001)
Urbanization	—	—	0.004***	0.003**
			(0.001)	(0.001)
Bank Competition	—	—	-0.089*	-0.082
			(0.048)	(0.051)
Spatial Effects				
W × ICTII	0.237**	0.221**	0.198**	0.183*
	(0.094)	(0.098)	(0.089)	(0.095)
ρ (spatial lag)	0.356***	0.342***	0.328***	0.314***
	(0.086)	(0.089)	(0.091)	(0.094)
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	252	252	252	252
R^2	0.714	0.731	0.768	0.755
Log-Likelihood	348.2	356.7	374.8	368.3

Note. Standard errors in parentheses. ***$p < .01$. *$p < .05$. $p < .10$. Spatial weight matrix: Queen contiguity, row-normalized. Effects decomposition follows LeSage and Pace (2009).

Key findings:

1. **Non-linear ICT effects (H1 confirmed):** The quadratic term coefficient ($\beta_2 = -0.052$, $p < .05$) is negative and significant, confirming inverted U-shaped relationships. Inflection point calculation ($-\beta_1/(2\beta_2)$) suggests optimal ICTII around 0.68—beyond this threshold, additional ICT investments yield declining inclusion dividends.
2. **Strong spatial spillovers (H2 confirmed):** The spatial autoregressive parameter ($\rho = 0.342$, $p < .01$) indicates substantial positive spillovers, supporting spatial demonstration and competitive diffusion mechanisms hypothesized in Rogers (2003).
3. **Neighbor ICT spillovers:** The spatially lagged ICTII coefficient ($\theta = 0.198$, $p < .05$) reveals that neighboring states' ICT infrastructure positively affects own-state financial inclusion.
4. **Control variables:** GSDP per capita, literacy, and urbanization demonstrate expected positive associations, consistent with Sethy and Goyari (2022).

5.3 Direct, Indirect, and Total Effects Decomposition

Table 5 decomposes marginal effects into direct, indirect, and total components—the appropriate interpretation framework for spatial models (Elhorst, 2010; LeSage & Pace, 2009).

Table 5: *Decomposition of ICT Effects on Financial Inclusion*

Variable	Direct Effect	Indirect Effect	Total Effect
ICTII (Linear)	0.468***	0.285**	0.753***
	(0.074)	(0.112)	(0.158)
ICTII (Quadratic)	-0.048**	-0.029*	-0.077**
	(0.019)	(0.015)	(0.031)
GSDP per capita (log)	0.134***	0.072**	0.206***
	(0.032)	(0.029)	(0.053)
Literacy Rate	0.003**	0.002*	0.005**
	(0.001)	(0.001)	(0.002)
Urbanization	0.004***	0.002**	0.006***
	(0.001)	(0.001)	(0.002)

Note. Standard errors in parentheses. ***$p < .01$. *$p < .05$. $p < .10$. Effects calculated following the LeSage and Pace (2009) methodology. Indirect effects constitute approximately 38% of total ICTII impacts. A 0.1-unit increase in a state's ICTII directly raises FII by 0.047 units while indirectly increasing neighboring states' FII by 0.029 units—total system-wide effect of 0.075 units.

5.4 Dynamic Panel GMM Results

Table 6 reports System-GMM estimates addressing endogeneity concerns (Arellano & Bond, 1991; Blundell & Bond, 1998).

Table 6: *System GMM Dynamic Panel Estimates*

Variables	(1) One-Step	(2) Two-Step Robust
FII (*t*-1)	0.342***	0.328***
	(0.067)	(0.072)
ICTII	0.394***	0.412***
	(0.089)	(0.094)
ICTII²	-0.041**	-0.045**
	(0.018)	(0.021)

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages4009-4016

GSDP per capita (log)	0.156***	0.148***
	(0.038)	(0.041)
Literacy Rate	0.004**	0.003**
	(0.002)	(0.001)
Urbanization	0.005***	0.004***
	(0.001)	(0.001)
Year FE	Yes	Yes
Observations	224	224
Number of States	28	28
Number of Instruments	24	24
AR(1) *p*-value	.021	.018
AR(2) *p*-value	.347	.412
Hansen *J*-test *p*-value	.564	.623

Note. Standard errors in parentheses. $^{**}p < .01$. $^{*}p < .05$. $p < .10$. Instruments: Lags 2–4 of endogenous variables in differences, following Roodman (2009).

Specification tests confirm model validity: AR(2) test fails to reject the null hypothesis ($p > .10$), indicating no second-order serial correlation. Hansen *J*-test *p*-values exceed 0.5, suggesting instrument validity. After addressing reverse causality, ICTII coefficients ($\beta_1 = 0.412$, $\beta_2 = -0.045$) remain significant with quadratic attenuation, consistent with SDM results. The lagged FII coefficient (0.328, $p < .01$) indicates moderate persistence—half-life calculation $\ln(0.5)/\ln(0.328) \approx 0.62$ years implying relatively rapid adjustment.

5.5 Ensemble Machine Learning Results

5.5.1 Predictive Performance Comparison

Table 7 compares predictive accuracy across ensemble algorithms and benchmark linear regression.

Table 7: *Predictive Performance: Machine Learning Models*

Model	R^2	RMSE	MAPE (%)	CCC
Training Set (2015–2020)				
Linear Regression	0.721	0.082	12.34	0.846
Random Forest	0.894	0.051	7.62	0.943
Gradient Boosting	0.912	0.046	6.89	0.954
XGBoost	0.927	0.042	6.21	0.962
Validation Set (2021)				
Linear Regression	0.698	0.089	13.45	0.828
Random Forest	0.836	0.065	9.82	0.911
Gradient Boosting	0.854	0.061	9.14	0.921
XGBoost	0.871	0.057	8.67	0.932
Test Set (2022–2023)				
Linear Regression	0.684	0.093	14.12	0.815
Random Forest	0.812	0.072	10.89	0.897
Gradient Boosting	0.829	0.068	10.23	0.907
XGBoost	0.847	0.064	9.56	0.918

Note. R^2 = coefficient of determination; RMSE = root mean squared error; MAPE = mean absolute percentage error; CCC = concordance correlation coefficient. Ensemble methods follow Hastie et al. (2009) validation protocols.

XGBoost achieves the highest predictive accuracy (test $R^2 = 0.847$), followed by Gradient Boosting (0.829) and Random Forest (0.812)—all substantially outperforming linear regression (0.684), consistent with Gu et al.'s (2020) findings on ensemble superiority in complex financial modeling contexts.

5.5.2 Feature Importance Analysis

Table 8 quantifies variable importance rankings from the XGBoost model using SHAP values (Lundberg & Lee, 2017).

Table 8: *Feature Importance Rankings: XGBoost Model (SHAP Values)*

Feature	SHAP Importance	Rank
Mobile Internet Penetration	0.284	1
GSDP per capita (log)	0.197	2
Digital Transaction Volume	0.156	3
Data Usage per Subscriber	0.143	4
Literacy Rate	0.089	5
Urbanization	0.067	6
Broadband Density	0.052	7
Telecom Infrastructure	0.041	8
Bank Branch Density	0.038	9
Road Density	0.033	10

Note. SHAP importance normalized to sum to 1.0. SHAP values computed following Lundberg and Lee (2017).

Mobile internet penetration emerges as the single most important predictor (importance = 0.284), confirming H4 and validating India's mobile-first digital strategy (Kaur et al., 2020). Bank branch density ranks ninth (0.038), confirming digital channels now supersede physical infrastructure as determinants of financial inclusion.

5.5.3 Partial Dependence Analysis: Partial dependence plots reveal non-linear relationships: mobile internet penetration exhibits an S-shaped curve with steep positive gradient between 40–80 subscriptions per 100 population, flattening beyond 90, suggesting threshold saturation effects. The ICTII composite confirms the inverted U-shape identified in econometric models, with optimal range 0.60–0.75 maximizing FII predictions—consistent with Breiman's (2001b) argument that algorithmic approaches reveal structural non-linearities obscured by parametric specifications.

5.6 Heterogeneity Analysis: State Group Differences: Table 9 presents estimates from subgroup and quantile regression analyses testing H3.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages4009-4016

Table 9: *Heterogeneous ICT Effects by Development Level*

Variables	High Dev.	Low Dev.	Q10 FII	Q90 FII
ICTII	0.321***	0.685***	0.742***	0.287***
	(0.082)	(0.104)	(0.134)	(0.089)
ICTII²	-0.038*	-0.071**	-0.089**	-0.029
	(0.021)	(0.031)	(0.038)	(0.024)
GSDP per capita (log)	0.087**	0.223***	0.268***	0.064*
	(0.038)	(0.056)	(0.072)	(0.037)
Literacy Rate	0.002	0.006***	0.007***	0.001
	(0.001)	(0.002)	(0.002)	(0.001)
Observations	126	126	252	252
R^2	0.689	0.743	—	—

Note. Standard errors in parentheses. **$p < .01$. *$p < .05$. $p < .10$. High/Low Dev.: above/below median GSDP per capita. Q10/Q90: quantile regression at 10th and 90th percentiles.

ICT impact in low-development states (β = 0.685) exceeds high-development states (β = 0.321) by over twofold—consistent with catch-up growth dynamics and diminishing returns theory (Williamson, 1981). Quantile regression reveals ICT effects are strongest at the 10th FII percentile (β = 0.742), declining monotonically to the 90th percentile (β = 0.287), suggesting ICT functions as an equalizing force disproportionately benefiting financially excluded populations (Tchamyou, 2019).

5.7 Robustness Checks

Extensive robustness analyses confirm main findings: (a) results remain stable across alternative spatial weight matrices (inverse distance, k-nearest neighbors $k = 5$, economic similarity); (b) PCA-based FII construction yields qualitatively identical conclusions; (c) exclusion of Delhi and Goa (city-state characteristics) preserves coefficient signs and significance; (d) subsample analyses (2015–2019 vs. 2020–2023) reveal stable relationships; and (e) two-stage least squares using lagged ICT infrastructure and neighboring states' telecom policy changes as instruments confirms causal interpretation, consistent with Blundell and Bond (1998).

6. Discussion

6.1 Theoretical Implications: Our findings extend existing theoretical frameworks in three dimensions. First, we provide macro-level validation of UTAUT constructs (Venkatesh et al., 2003): state ICT infrastructure (facilitating conditions) positively moderates financial service adoption (performance expectancy), with spatial social influence manifesting through inter-state demonstration effects. The inverted U-shaped relationship suggests effort expectancy increases non-linearly—initial ICT deployments reduce adoption barriers, but infrastructure saturation without complementary digital literacy yields diminishing returns, extending Dwivedi et al.'s (2019) revised UTAUT model to macro-institutional contexts. Second, spatial spillover mechanisms validate technology diffusion theories emphasizing network externalities and knowledge transmission (Hägerstrand, 1967; Rogers, 2003). The spatial autoregressive coefficient (ρ = 0.342) quantifies demonstration effects: neighboring states' financial inclusion success signals ICT feasibility, triggering policy emulation and competitive diffusion. This corroborates Caragliu and Del Bo's (2019) findings on spatial innovation spillovers.Third, heterogeneous treatment effects across development quintiles reconcile conflicting literature on ICT–financial inclusion relationships. ICT functions as skill-complementary technology—effectiveness conditional on minimum human capital (literacy), infrastructure (roads, electricity), and economic capacity (GSDP per capita). Lagging states benefit most from marginal ICT investments, but absolute gains require complementary factor accumulation, consistent with Tchamyou (2021) and Asongu and Odhiambo (2019).

6.2 Methodological Contributions: This study pioneers methodological triangulation integrating spatial econometrics, dynamic panel causal inference, and machine learning prediction (Mullainathan & Spiess, 2017). Convergent findings across methods strengthen causal claims. The demonstration that ensemble algorithms achieve 24% improved predictive accuracy over linear specifications (test R^2 from 0.684 to 0.847) validates algorithmic approaches in development finance contexts, extending Gu et al.'s (2020) empirical asset pricing findings to financial inclusion. A critical methodological insight concerns spatial effect interpretation: direct/indirect effect decomposition (LeSage & Pace, 2009, 2014) reveals spillovers constitute 38% of total ICT impacts—a finding invisible to standard panel estimators and underscoring why spatial models must accompany interstate analyses.

6.3 Policy Implications

1. Targeted ICT Infrastructure Investment. Low-development states (Bihar, Jharkhand, Uttar Pradesh) demonstrate highest marginal returns (coefficient = 0.685 vs. 0.321 in high-development states). Prioritizing mobile internet expansion maximizes aggregate inclusion gains, but non-linearity cautions against infrastructure-only strategies (Senyo & Osabutey, 2020).

2. Mobile-First Digital Financial Services. Mobile internet penetration's dominance (SHAP importance = 0.284) validates India's Unified Payments Interface (UPI) and mobile wallet strategies. Policy should accelerate mobile broadband coverage in rural areas through spectrum allocation reforms and universal service obligation funds (Reserve Bank of India, 2023).

3. Leveraging Spatial Spillovers. Positive spatial autocorrelation (Moran's $I = 0.412$) and spillover coefficients (θ = 0.198) justify regional cooperation frameworks. Interstate compacts facilitating knowledge exchange and shared digital public infrastructure amplify individual state efforts (Hägerstrand, 1967; Rogers, 2003).

4. Addressing Spatial Poverty Traps. Persistent Low-Low clusters in eastern/central India require coordinated interventions bundling ICT infrastructure, financial literacy, banking correspondent networks, and livelihood support, using LISA cluster maps for spatial targeting (Anselin, 1988).

5. Threshold-Aware Investment Strategies. The inflection point (ICTII ≈ 0.68) implies optimal infrastructure targets. States below thresholds require continued investment; states above (Karnataka, Kerala, Maharashtra) should pivot toward usage deepening and service diversification (Demirgüç-Kunt et al., 2022).

6.4 Limitations and Future Research

Several limitations warrant acknowledgment. FII construction relies on supply-side indicators due to limited demand-side data; household-level financial diaries would enable more granular measurement. Despite GMM and spatial econometric rigor, randomized controlled trials or quasi-experimental designs exploiting telecom tower placement discontinuities would provide stronger causal identification. The nine-year panel provides medium-term perspective; extending to 15–20 years would capture equilibrium effects amenable to cointegration analysis. Spatial models establish spillover existence but not mechanisms—structural equation models incorporating inter-state migration flows, media coverage, and policy adoption timing would disentangle channels. Finally, macro-level analyses obscure within-state heterogeneity by caste, gender, and occupation; integration with household surveys (NSSO, IHDS) through multilevel modeling would reveal distributional consequences (Demirgüç-Kunt et al., 2018).

7. Conclusion

This study provides comprehensive evidence on ICT's transformative role in financial inclusion across Indian states during 2015–2023, employing methodological triangulation spanning spatial econometrics, dynamic panel estimation, and ensemble machine learning. Three principal conclusions emerge. First, ICT infrastructure demonstrates robust positive impacts on multidimensional financial inclusion, exhibiting non-linear inverted U-shaped relationships with diminishing marginal returns beyond optimal thresholds (ICTII ≈ 0.68). Mobile internet penetration emerges as the dominant driver (SHAP importance = 0.284), surpassing traditional banking infrastructure and validating mobile-first digital strategies (Kaur et al., 2020; Reserve Bank of India, 2023).Second, financial inclusion exhibits strong positive spatial autocorrelation and spillover dynamics. Neighboring states' ICT investments and inclusion levels significantly influence own-state outcomes, constituting 38% of total effects (LeSage & Pace, 2009). These spillovers justify regional coordination frameworks and spatial targeting of lagging clusters to break poverty traps (Hägerstrand, 1967).Third, substantial heterogeneity characterizes ICT–financial inclusion relationships across development levels. Low-income states derive over double the marginal benefits (coefficient = 0.685) compared to high-income counterparts (0.321), suggesting ICT as equalizing force contingent on complementary factors—literacy, economic capacity, physical infrastructure (Tchamyou, 2021; Williamson, 1981).As India advances toward universal financial inclusion targets under the National Strategy for Financial Inclusion 2019–2024, evidence-based spatial prioritization and mobile-centric digital strategies grounded in rigorous causal inference and predictive analytics offer pathways to accelerate progress while addressing persistent regional disparities (Demirgüç-Kunt et al., 2022; Sethy & Goyari, 2022).

References

Aldasoro, I., Gambacorta, L., Giudici, P., & Leach, T. (2024). Promise (un)kept? Fintech and financial inclusion (Working Paper No. 2024/131). International Monetary Fund. https://doi.org/10.5089/9798400279195.001
Anselin, L. (1988). *Spatial econometrics: Methods and models*. Kluwer Academic Publishers. https://doi.org/10.1007/978-94-015-7799-1
Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, *58*(2), 277–297. https://doi.org/10.2307/2297968
Asongu, S. A., & Odhiambo, N. M. (2019). Challenges of doing business in Africa: A systematic review. *Journal of African Business*, *20*(2), 259–268. https://doi.org/10.1080/15228916.2019.1582294
Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, *87*(1), 115–143. https://doi.org/10.1016/S0304-4076(98)00009-8
Breiman, L. (2001a). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324
Breiman, L. (2001b). Statistical modeling: The two cultures. *Statistical Science*, *16*(3), 199–231. https://doi.org/10.1214/ss/1009213726
Caragliu, A., & Del Bo, C. F. (2019). Smart innovative cities: The impact of Smart City policies on urban innovation. *Technological Forecasting and Social Change*, *142*, 373–383. https://doi.org/10.1016/j.techfore.2018.07.022
Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785
Chen, T., He, T., Benesty, M., Khotilovich, V., & Tang, Y. (2015). *XGBoost: Extreme gradient boosting* (R package version 0.4-2).
Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008
Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., & Hess, J. (2018). *The Global Findex Database 2017: Measuring financial inclusion and the fintech revolution*. World Bank. https://doi.org/10.1596/978-1-4648-1259-0
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 Publications.
Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2019). Re-examining the unified theory of acceptance and use of technology (UTAUT): Towards a revised theoretical model. *Information Systems Frontiers*, *21*(3), 719–734. https://doi.org/10.1007/s10796-017-9774-y
Elhorst, J. P. (2010). Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis*, *5*(1), 9–28. https://doi.org/10.1080/17421770903541772
Elhorst, J. P. (2014). *Spatial econometrics: From cross-sectional data to spatial panels*. Springer. https://doi.org/10.1007/978-3-642-40340-8
Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451
Gai, K., Qiu, M., & Sun, X. (2018). A survey on FinTech. *Journal of Network and Computer Applications*, *103*, 262–273. https://doi.org/10.1016/j.jnca.2017.10.011
Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009
Hägerstrand, T. (1967). *Innovation diffusion as a spatial process*. University of Chicago Press.
Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7
Huang, Y., & Zhang, Y. (2020). Financial inclusion and urban–rural income inequality: Long-run and short-run relationships. *Emerging Markets Finance and Trade*, *56*(2), 457–471. https://doi.org/10.1080/1540496X.2018.1564904
Katz, M. L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. *American Economic Review*, *75*(3), 424–440.
Kaur, P., Dhir, A., Bodhi, R., Singh, T., & Almotairi, M. (2020). Why do people use and recommend m-wallets? *Journal of Retailing and Consumer Services*, *56*, Article 102091. https://doi.org/10.1016/j.jretconser.2020.102091
Khera, R. (2019). Aadhaar failures: A tragedy of errors. *Economic and Political Weekly*, *54*(14), 21–28.
LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. CRC Press. https://doi.org/10.1201/9781420064254
LeSage, J. P., & Pace, R. K. (2014). The biggest myth in spatial econometrics. *Econometrics*, *2*(4), 217–249. https://doi.org/10.3390/econometrics2040217
Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30* (pp. 4765–4774).
Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87–106. https://doi.org/10.1257/jep.31.2.87
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
Reserve Bank of India. (2023). *Financial inclusion report 2022–23*. Reserve Bank of India.
Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, *9*(1), 86–136. https://doi.org/10.1177/1536867X0900900106
Sarma, M. (2012). *Index of financial inclusion—A measure of financial sector inclusiveness* (Working Paper No. 07/2012). Hochschule für Technik und Wirtschaft.
Sarma, M., & Pais, J. (2011). Financial inclusion and development. *Journal of International Development*, *23*(5), 613–628. https://doi.org/10.1002/jid.1698
Senyo, P. K., & Osabutey, E. L. C. (2020). Unearthing antecedents to financial inclusion through FinTech innovations. *Technovation*, *98*, Article 102155. https://doi.org/10.1016/j.technovation.2020.102155
Sethy, S. K. (2016). Developing a financial inclusion index and inclusive growth in India. *Theoretical and Applied Economics*, *23*(2), 187–206.
Sethy, S. K., & Goyari, P. (2022). Measuring financial inclusion and its present status in South Asian and Southeast Asian countries: An empirical study. *Theoretical and Applied Economics*, *29*(3), 131–150.
Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *American Economic Review*, *71*(3), 393–410.
Tchamyou, V. S. (2019). The role of information sharing in modulating the effect of financial access on inequality. *Journal of African Business*, *20*(3), 317–338. https://doi.org/10.1080/15228916.2019.1584262
Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*(Supplement), 234–240. https://doi.org/10.2307/143141
Vega, S. H., & Elhorst, J. P. (2015). The SLX model. *Journal of Regional Science*, *55*(3), 339–363. https://doi.org/10.1111/jors.12188
Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540
Williamson, O. E. (1981). The economics of organization: The transaction cost approach. *American Journal of Sociology*, *87*(3), 548–577. https://doi.org/10.1086/227496