

The LIMR Computational Framework for Diagnosing Structural Imbalances in Digital Finance Systems

Chenhao Ying

Ningbo University

Lexing Xiong

xionglexing@nbu.edu.cn

Ningbo University

Research Article

Keywords: digital finance, financial inclusion, convergence analysis, Markov chains, panel data, structural resilience

Posted Date: October 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10988037/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

The LIMR Computational Framework for Diagnosing Structural Imbalances in Digital Finance Systems

Abstract

We develop LIMR (Level Inequality Mobility Resilience), a modular diagnostic framework for multidimensional digital-finance panels. Conventional composite indices condense each region into a single score and cannot reveal whether access and use evolve together, whether convergence in levels coexists with stable rank hierarchies, or whether continued growth preserves internal coordination. LIMR processes the raw coordinate panel through five deterministic closed-form modules: a triad separating absolute mismatch from relative balance, annual inequality decomposition, convergence rates and half-lives, quartile transition matrices, and joint measures of growth and coordination drift over a stress window. Applied to the Peking University Digital Financial Inclusion Index of China for 31 mainland provinces during 2011–2023 (403 province-years), the framework identifies relative convergence amid absolute divergence, slower adjustment in usage depth than in coverage breadth, persistent relative clubs despite lower dispersion, and asymmetric structural resilience from 2019 to 2023. Permutation, subsample, quantile-grid, spatial-conditioning, and fixed-effect checks support all findings. LIMR offers a transparent, reproducible diagnostic complement to aggregate scores for evaluating the structural condition of regional digital-finance systems.

Keywords: digital finance; financial inclusion; convergence analysis; Markov chains;
panel data; structural resilience

1 Introduction

China’s provincial digital finance systems expanded almost tenfold between 2011 and 2023, and this expansion has been certified routinely by composite instruments such as the Peking University Digital Financial Inclusion Index of China, abbreviated PKU-DFIIC [1]. Downstream studies feed the resulting single score into models of consumption, income distribution, entrepreneurship, and regional inequality [2–7], following established measurement practice [8, 9]. This certification is incomplete in a way that matters technically. A composite score answers whether a region grew. It cannot answer whether the components grew together. Yet the distinction between having access and making meaningful use is well documented across financial inclusion research [8, 10, 11], multidimensional index construction [9], and digital divide theory [12], and recent work ties the digital finance usage divide directly to income disparities between urban and rural households [13–16].

The blind spot produces reversals rather than mere omissions. In our data the national absolute gap between usage depth and coverage breadth more than doubles over the sample, rising from 17.74 to 34.41, which taken alone signals deepening malfunction. The relative balance index simultaneously improves from 0.707 to 0.953, which taken alone signals tightening coordination. Both statements hold at once because the system scaled almost tenfold. An analysis built on either statistic alone returns opposite conclusions about the same object. Similar tensions arise between collapsing dispersion and frozen rankings, and between continued growth and degraded internal coordination after the disturbance year of 2019. Diagnosing tensions of this

kind calls for a multi module instrument applied to the raw dimensions rather than for better aggregation.

This paper designs such an instrument and reports what it finds. We make four contributions.

- 1) **Problem formulation.** We formulate diagnosis of a regional digital finance panel as four estimation tasks labeled T1 through T4 over raw coordinate series, covering coordination state, distribution dynamics, rank mobility, and stress window structural integrity. Each task specifies inputs, estimators, and outputs and stands independent of behavioral modeling assumptions.
- 2) **Novel measurement core.** Our constructed indicator triad decouples absolute mismatch scale from relative balance. Its formal decoupling property explains why single statistic readings reverse one another and prevents such reversals by construction. To our knowledge this coordination margin has never been measured directly on released index data, and the triad supplies the instrument behind finding F1.
- 3) **Integrated algorithm.** We develop LIMR, a five module pipeline of closed form estimators connected by an explicit data contract stated in Algorithm 1. Two elements are new. The rate transform $\lambda = -\ln(1 + \beta)$ together with its half life companion renders convergence speeds commensurable across dimensions, which no prior treatment of these indices provides, and it drives finding F2. The resilience module emits growth and coordination drift side by side rather than collapsing them into one scalar verdict, and it drives finding F4.
- 4) **Validation and findings.** Executed on 31 mainland provinces from 2011 through 2023, the pipeline returns four robust findings labeled F1 through F4. Jointly they establish that aggregate scores are unreliable readiness certificates for AI era financial systems and pinpoint exactly where, and along which margin, the system is unbalanced.

Section 2 reviews related work. Section 3 defines notation, the indicator triad, and the four tasks. Section 4 details the five modules, their novelty points, and the complete algorithm. Section 5 describes the dataset and experimental protocol. Section 6 reports results organized around findings F1 through F4, and Section 7 concludes.

2 Related Work

Composite indices and their discontents. Financial inclusion measurement evolved from single participation measures toward multidimensional indices spanning availability, penetration, and use [8, 9, 11]. Composite aggregation then collapses these axes for modeling convenience [2–5]. What disappears in the collapse is the correlation structure among axes, dimensional heterogeneity, and within score composition. Prior work stops at aggregation. Our Module A begins exactly where that literature stops, measuring the internal margin directly on the released coordinates, and the remaining modules convert this measurement into portable diagnostics.

Regional inequality and convergence. Sigma convergence tracks cross sectional dispersion, beta convergence tests whether initially lower units grow faster [17], and distribution dynamics methods follow ranks to expose persistence invisible to both [18]. Their decomposition partitions inequality into between group and within group shares [19]. These tools live in separate literatures and are rarely applied jointly to digital finance panels. Our contribution couples them under one data contract and adds the rate transform of Eq. (10), an ingredient absent from the reviewed convergence treatments, which turns heterogeneous beta coefficients into comparable per dimension speeds and half lives.

System resilience. Resilience denotes the capacity to absorb disturbances while maintaining function [20, 21]. Applied studies connect fintech outcomes to technological capability, economic context, and institutional design [22, 23]. Existing operationalizations reduce resilience to sustained growth and never verify whether

the relationships among service layers survive a disturbance. Our Module E fills precisely that void by scoring growth and coordination drift jointly over a matched stress window, operationalizing the structural view suggested by socio technical accounts of digital finance [24, 25].

3 Problem Formulation

This section fixes the objects on which our contribution operates. It introduces a data model, the indicator triad carrying novelty point N1, and the four diagnostic tasks that organize the pipeline.

3.1 Data model and notation

Table 1 summarizes notation. The input is a balanced province by year panel

$$\mathcal{D} = \{(Y_{i,t}, C_{i,t}, U_{i,t}, D_{i,t}) : i = 1, \dots, n, t = 1, \dots, T\},$$

where Y denotes the aggregate inclusion index, C coverage breadth, U usage depth, and D digitalization degree, with $n = 31$ provinces and $T = 13$ years spanning 2011 through 2023, hence $|\mathcal{D}| = 403$. A mapping $r : \{1, \dots, n\} \rightarrow \{\text{East, Central, West, NE}\}$ assigns National Bureau of Statistics macro regions [26]. For year t , μ_t denotes the cross sectional mean of the coordinate under consideration, and $x_{i,t}$ denotes a generic scalar observation.

3.2 Indicator triad and the decoupling property

Three indicators describe the coverage usage relationship of province i at time t :

$$G_{i,t} = |U_{i,t} - C_{i,t}|, \tag{1}$$

$$B_{i,t} = 1 - \frac{|U_{i,t} - C_{i,t}|}{U_{i,t} + C_{i,t}}, \tag{2}$$

Table 1 Notation.

Symbol	Domain	Meaning
n, T	integer	number of provinces (31) and years (13)
i, t	–	province index and calendar year from 2011 through 2023
$Y_{i,t}$	$\mathbb{R}_+$	aggregate inclusion index of province i in year t
$C_{i,t}, U_{i,t}, D_{i,t}$	$\mathbb{R}_+$	coverage breadth, usage depth, digitalization degree
$r(i)$	$\mathcal{R}$	macro region of province i out of four classes
$G_{i,t}, B_{i,t}, K_{i,t}$	$\mathbb{R}_+, [0, 1], [0, 1]$	absolute gap, relative balance, coupling coefficient
$CV_t, \mathcal{G}_t, Th_t$	$\mathbb{R}_+$	coefficient of variation, Gini index, Theil index in year t
$x_{i,t}, \mu_t$	$\mathbb{R}_+$	generic scalar observation and its cross sectional mean
β	$\mathbb{R}$	estimated initial level coefficient from the convergence module
λ, HL	$\mathbb{R}_+$	implied annual adjustment rate and half life in years
$s_{i,t}$	$\{1, 2, 3, 4\}$	relative quartile state of province i in year t
$P = (p_{jk})$	simplex	Markov transition matrix over states
$\bar{X}_{g,t}$	$\mathbb{R}_+$	mean of coordinate X over region g in year t

$$K_{i,t} = \frac{2\sqrt{C_{i,t}U_{i,t}}}{C_{i,t} + U_{i,t}}. \quad (3)$$

Here G measures mismatch in index units and responds to scale, B normalizes the mismatch by combined size onto the unit interval and removes scale, and K is a cosine type coupling equal to one exactly when the two coordinates coincide. The economic content of the triad lies in reporting the pair (G, B) jointly:

Remark 1 (Decoupling property). Because $B_{i,t}$ divides out the common scale factor, large monotone expansion of both coordinates leaves B approximately invariant while shifting G upward in proportion to level. Sign disagreement between $\text{trend}(G)$ and $\text{trend}(B)$ is therefore not a contradiction but a signature of pure scaling. Any conclusion about divergence between coverage and usage drawn from G alone is identified

against the counterfactual supplied by B . Conversely, B alone identifies nothing about the size of operational mismatch. Every subsequent module therefore reports the pair jointly.

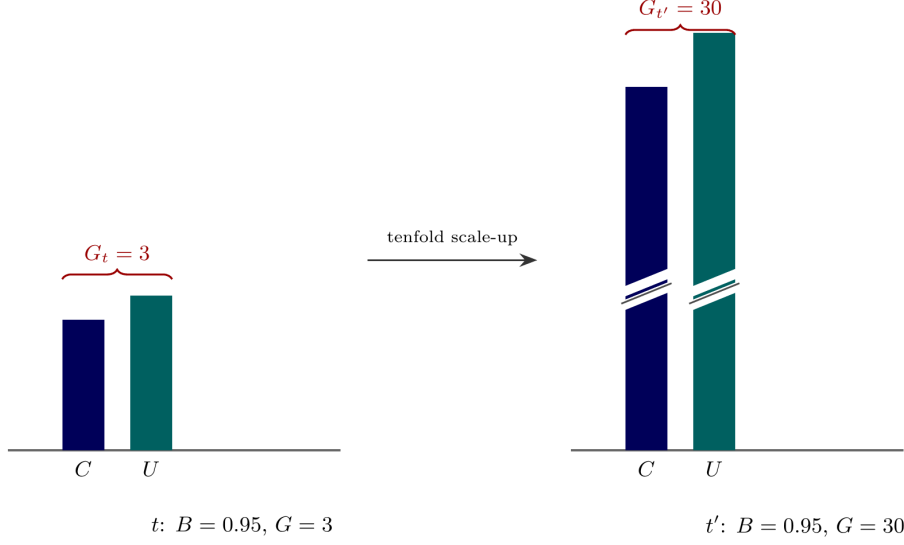


Fig. 1 The decoupling property of Remark 1, drawn schematically. A proportional tenfold expansion lifts both coordinates and stretches the absolute gap from $G_t = 3$ to $G_{t'} = 30$ while the normalized balance stays at $B = 0.95$; slashed bar segments abbreviate the scaled extent.

3.3 Diagnostic tasks

Given $\mathcal{D}$ and the region map r , LIMR solves four tasks.

- T1** *Estimate coordination state*, namely national and regional trajectories of (G, B, K) by year (Module A).
- T2** *Quantify distribution dynamics*: annual cross sectional disparity, its split between macro regions and within them, and level convergence with dimension resolved speeds (Modules B and C).
- T3** *Estimate rank mobility*, meaning relative state persistence independently of levels (Module D).

T4 *Evaluate structural resilience*: decide jointly over the stress window from 2019 through 2023 whether growth continued and whether coordination drifted, region by region (Module E).

Resilience is defined structurally rather than behaviorally, following socio technical accounts of digital finance systems [24, 25]:

Definition 1 (Structural resilience). Region g is structurally resilient over window $[\tau_1, \tau_2]$ if, conditional on nonnegative mean growth of every coordinate, its coordination quality does not deteriorate, namely $\Delta \bar{B}_g \geq 0$ and $\Delta \bar{G}_g \leq 0$, where bars denote regional means of Eq. (2) and Eq. (1) evaluated at window endpoints.

A region that merely stays balanced satisfies the definition trivially. The pair $(\Delta \bar{G}, \Delta \bar{B})$ is therefore always reported alongside growth rates rather than collapsed into a binary verdict.

4 The LIMR Algorithm

4.1 Overview and novelty points

Figure 2 shows the serpentine architecture, and Figures 1, 3, and 4 visualize the mechanisms behind novelty points N1 through N3 respectively. Modules run sequentially, and each consumes raw coordinates or artifacts of earlier stages. There are no learned parameters, no stochastic initialization, and no tuning set, so repeated executions return identical outputs. Beyond assembling known statistics, three design elements carry the novelty of the pipeline:

- N1)** the indicator triad (G, B, K) governed by the decoupling property of Remark 1, which lets one instrument measure mismatch scale and coordination quality without conflating them;
- N2)** the rate transform of Eq. (10), which renders beta coefficients from different dimensions commensurable as annual adjustment rates and half lives;

N3) the dual metric resilience output of Definition 1, which separates growth resilience from structural resilience instead of reducing both to a single number. Each subsequent section names the novelty point its output depends on, and Section 6 traces every empirical finding back to its enabling element.

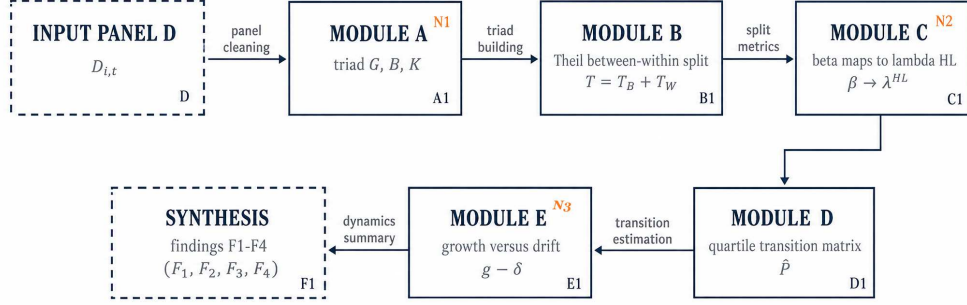


Fig. 2 Architecture and data flow of LIMR. Dashed boxes mark pipeline boundary objects; solid boxes are the five deterministic closed form modules, whose estimator fragments abbreviate Eqs. (1)–(13). Corner codes A1 through F1 index emitted artifacts, and orange badges flag the three novelty points hosted by Modules A, C, and E respectively. Only raw released coordinates enter at the upper left boundary.

4.2 Module B: inequality and its regional decomposition

Per year t and coordinate x with mean μ_t , the pipeline computes dispersion, distributional inequality, and entropy based disparity:

$$CV_t = \sigma_t / \mu_t, \quad (4)$$

$$\mathcal{G}_t = \frac{1}{2n^2\mu_t} \sum_{i=1}^n \sum_{j=1}^n |x_{i,t} - x_{j,t}|, \quad (5)$$

$$Th_t = \frac{1}{n} \sum_{i=1}^n \frac{x_{i,t}}{\mu_t} \ln \frac{x_{i,t}}{\mu_t}. \quad (6)$$

Among these only the Theil component admits exact additive decomposition. With n_g provinces and mean $\bar{x}_{g,t}$ in region g , total Theil splits as $Th_t = \mathcal{W}_t + \mathcal{B}_t$ where

$$\mathcal{B}_t = \sum_g \frac{n_g}{n} \frac{\bar{x}_{g,t}}{\mu_t} \ln\left(\frac{\bar{x}_{g,t}}{\mu_t}\right), \quad \mathcal{B}_t/Th_t \text{ is the reported between region share.} \quad (7)$$

Reporting this share annually lets later findings locate surviving disparity geographically instead of treating inequality as a featureless aggregate.

4.3 Module C: convergence estimation and the rate transform

Two specifications are fitted per coordinate $d \in \{Y, C, U, D, G\}$. The cross section model regresses average log growth from 2011 through 2023 on the initial level,

$$\frac{\ln d_{i,2023} - \ln d_{i,2011}}{12} = \alpha + \beta \ln d_{i,2011} + \boldsymbol{\gamma}^\top \mathbf{1}_{r(i)} + \epsilon_i, \quad (8)$$

with macro region dummies and heteroskedasticity robust inference, and the annual panel model adds time and region effects,

$$\Delta \ln d_{i,t} = \alpha + \beta \ln d_{i,t-1} + \phi_t + \sum_g \gamma_g \mathbf{1}[r(i)=g] + \epsilon_{i,t}, \quad \text{SE}(\hat{\beta}) : \text{ province clustered.} \quad (9)$$

A negative $\hat{\beta}$ indicates that initial disadvantage predicts faster growth. Estimates with $\hat{\beta} < -1$ would be explosive and are excluded although none arises here. Otherwise the estimate maps to

$$\lambda = -\ln(1 + \hat{\beta}), \quad \text{HL} = \frac{\ln 2}{\lambda}, \quad (10)$$

so that λ is the discrete time adjustment rate and HL is the half life of a deviation expressed in years. Raw betas are not comparable across coordinates because each coordinate carries its own growth scale, whereas the rate transform normalizes that scale away. To our knowledge no prior study uses this transform to rank convergence

speeds across index dimensions, and it is the ingredient that makes finding F2 measurable in common units. One interpretive precondition deserves statement because the panel spans thirteen years of strongly trending levels. The half life reading presumes reversion around common year effects rather than stationarity of raw log levels, and Section 6.7 verifies this presumption through trend aware unit root diagnostics together with a province fixed effect guard specification.

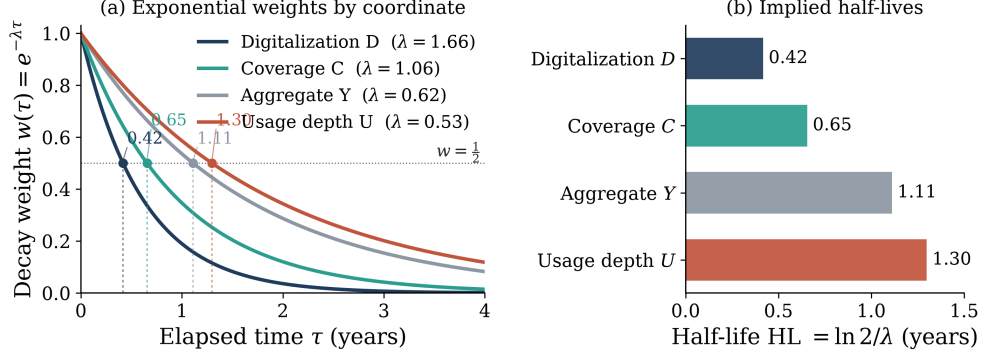


Fig. 3 Mechanics of novelty point N2 before any regression table is read. Panel (a) plots exponential weights $w(\tau) = e^{-\lambda\tau}$ at the estimated panel rates; dots on the half weight line mark half lives. Panel (b) ranks the implied $HL = \ln 2/\lambda$, exposing usage depth as the slow coordinate.

4.4 Module D: Markov mobility

Each year ranks provinces and assigns pooled quartile states $s_{i,t} \in \{1, 2, 3, 4\}$ from low to high. Adjacent year pairs then yield counts $N_{jk} = \#\{i : s_{i,t}=j, s_{i,t+1}=k\}$, and transitions follow maximum likelihood estimation,

$$\hat{p}_{jk} = \frac{N_{jk}}{\sum_{k'} N_{jk'}}, \quad j, k \in \{1, \dots, 4\}. \quad (11)$$

With 31 provinces and twelve adjacent year pairs reaching from 2011 into 2023, estimation rests on 372 transitions distributed over sixteen cells. Diagonal mass $\hat{p}_{jj}$ measures persistence and off diagonal mass measures remixing. Quartile assignment removes the

level trend entirely, so $\hat{P}$ captures pure relative mobility and complements Modules B and C in a way neither can substitute.

4.5 Module E: stress window resilience

Over $[\tau_1, \tau_2]$ set to 2019 through 2023, each region g receives stepwise and cumulative growth rates,

$$g_g^{\text{step1}} = \frac{\bar{Y}_{g,2020}}{\bar{Y}_{g,2019}} - 1, \quad g_g^{\text{step2}} = \frac{\bar{Y}_{g,2023}}{\bar{Y}_{g,2020}} - 1, \quad g_g^{\text{cum}} = \frac{\bar{Y}_{g,2023}}{\bar{Y}_{g,2019}} - 1, \quad (12)$$

and coordination drift,

$$\Delta \bar{G}_g = \bar{G}_{g,2023} - \bar{G}_{g,2019}, \quad \Delta \bar{B}_g = \bar{B}_{g,2023} - \bar{B}_{g,2019}, \quad (13)$$

evaluated against Definition 1. Because growth and drift answer different questions, Module E emits them side by side and never merges them into one score, which is the design choice separating structural resilience from mere growth continuity.

4.6 Complexity and implementation

Every module performs linear passes over the panel in order nT operations, with an additional pairwise term confined to the Gini sum worth $n^2 = 961$ additions, closed form ordinary least squares, and cellwise division for $\hat{P}$. The complete pipeline executes well under one second on commodity hardware and contains no random elements, so every printed digit reproduces exactly. Implementation uses Python 3.13 with pandas 2.2.3, NumPy 2.3.5, statsmodels 0.14.6, and Matplotlib 3.10.8. Replication scripts and the cleaned panel accompany the manuscript as described in Appendix A.

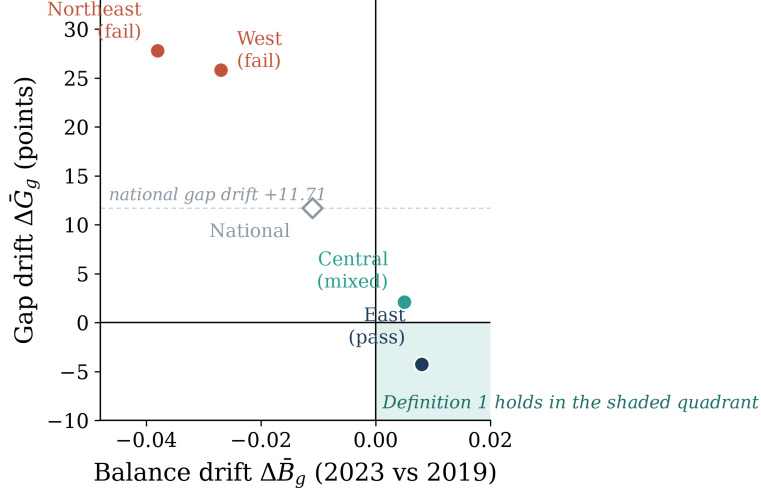


Fig. 4 The dual metric of novelty point N3 as a decision plane. Each region receives one point from its balance drift $\Delta\bar{B}_g$ and gap drift $\Delta\bar{G}_g$ over the stress window; Definition 1 holds only in the shaded quadrant, which eastern regions alone occupy here.

5 Experimental Setup

Dataset. The PKU-DFIIC province appendix for 2011 through 2023 [1] supplies, for each of 31 mainland provincial units and each year, the aggregate index, three dimensions, and functional subindices covering payment, insurance, investment, and lending. The investment subindex lacks early observations and enters only descriptive statistics with 310 available values. The result is the balanced input $\mathcal{D}$ with 403 rows. Hong Kong, Macao, and Taiwan fall outside the released appendix. Macro regions follow the National Bureau of Statistics classification [26]. No external variables, deflation, or rescaling enters any module.

Protocol. All five modules execute once over the full panel. Reported quantities are exact statistics of the released data. No train and test split applies because no prediction task is posed, and no bootstrap re estimation is required because all estimators are closed form with cluster robust inference. Determinism makes reruns identical, which we verified.

Algorithm 1 LIMR diagnostic pipeline

Require: balanced panel $\{Y_{i,t}, C_{i,t}, U_{i,t}, D_{i,t}\}$, region map $r(\cdot)$, stress window 2019 through 2023

Ensure: coordination artifact $\mathcal{C}$, inequality artifact $\mathcal{I}$, convergence table $\hat{\beta}$, transition matrix $\hat{P}$, resilience report $\mathcal{R}$

▷ **Module A**

1: **for all** (i, t) **do**
2: $(G_{i,t}, B_{i,t}, K_{i,t}) \leftarrow$ Eqs. (1) through (3)
3: **end for**
4: **return** $\mathcal{C} \leftarrow$ national and regional series of the triple

▷ **Module B**

5: **for all** t and $d \in \{Y, C, U, D\}$ **do**
6: $(CV_t, \mathcal{G}_t, Th_t) \leftarrow$ Eqs. (4) through (6), split by Eq. (7)
7: **end for**
8: **return** $\mathcal{I} \leftarrow$ annual disparity series with between shares

▷ **Module C**

9: **for all** $d \in \{Y, C, U, D, G\}$ **do**
10: fit Eqs. (8) and (9), then $(\lambda_d, HL_d) \leftarrow$ Eq. (10)
11: **end for**
12: **return** $\hat{\beta} \leftarrow$ betas with robust and cluster errors plus rates and half lives

▷ **Module D**

13: **for all** t **do**
14: $s_{i,t} \leftarrow$ pooled quartile of $Y_{i,t}$ within year t
15: **end for**
16: $N_{jk} \leftarrow$ adjacency counts; $\hat{P} \leftarrow$ Eq. (11); **return** $\hat{P}$

▷ **Module E**

17: **for all** regions $g \in \mathcal{R}$ **do**
18: growth vector $\leftarrow$ Eq. (12); $(\Delta \tilde{G}_g, \Delta \bar{B}_g) \leftarrow$ Eq. (13)
19: **end for**
20: **return** $\mathcal{R}$

▷ **Synthesis**

21: combine $\mathcal{C}, \mathcal{I}, \hat{\beta}, \hat{P}, \mathcal{R}$ into findings F1 through F4 (§6)

6 Results

Results follow the synthesis order of Algorithm 1. Section 6.1 calibrates the panel and establishes the scaling regime. Section 6.2 reports finding F1 produced by Module A. Sections 6.3 and 6.4 deliver the two halves of finding F2 from Modules B and C respectively. Section 6.5 presents finding F3 from Module D, and Section 6.6 presents finding F4 from Module E. Section 6.7 subjects every finding to permutation, sub-sample, quantile grid, spatial conditioning, and unit root aware re estimation checks.

Each table lists module outputs verbatim, and each finding names the novelty point that made it observable.

6.1 Calibration: the panel under the pipeline

Table 2 records full panel statistics computed by Module A alongside raw coordinate moments. National means rose from 40.00 for the aggregate index in 2011 to 393.70 in 2023, an increase of 884.1 percent, with coverage breadth moving from 34.28 to 398.08, usage depth from 46.93 to 364.89, and digitalization degree from 46.32 to 431.63. Shanghai held the maximum aggregate value at both endpoints, while Xizang was minimum in 2011 and Qinghai in 2023, and bottom ranked provinces grew fastest in percentage terms. Figure 5 plots national trajectories. This near tenfold scaling regime matters methodologically, because Remark 1 predicts that under such conditions level effects dominate every single scale statistic, which is why the full battery runs rather than levels being read directly.

Table 2 Full panel descriptive statistics for 31 provinces across 2011 through 2023.

Variable	N	Mean	Std. dev.	Min	Median	Max
Aggregate index	403	254.48	111.33	16.22	267.80	473.83
Coverage breadth	403	239.25	116.13	1.96	247.19	466.54
Usage depth	403	245.55	109.57	6.76	258.44	510.69
Digitalization degree	403	321.01	117.80	7.58	369.19	476.90
Usage coverage gap	403	30.01	22.05	0.18	25.88	106.31
Balance index	403	0.91	0.11	0.20	0.94	1.00
Coupling coefficient	403	0.99	0.04	0.60	1.00	1.00
Payment	403	202.31	88.68	0.00	224.18	379.51
Insurance	403	503.61	206.17	0.25	541.83	951.94
Investment	310	299.95	167.29	9.20	282.75	758.68
Lending	403	164.90	74.69	1.16	170.02	303.74

Note: index values come from PKU-DFIIC, and constructed indicators are computed from the same data. Investment lacks early observations.

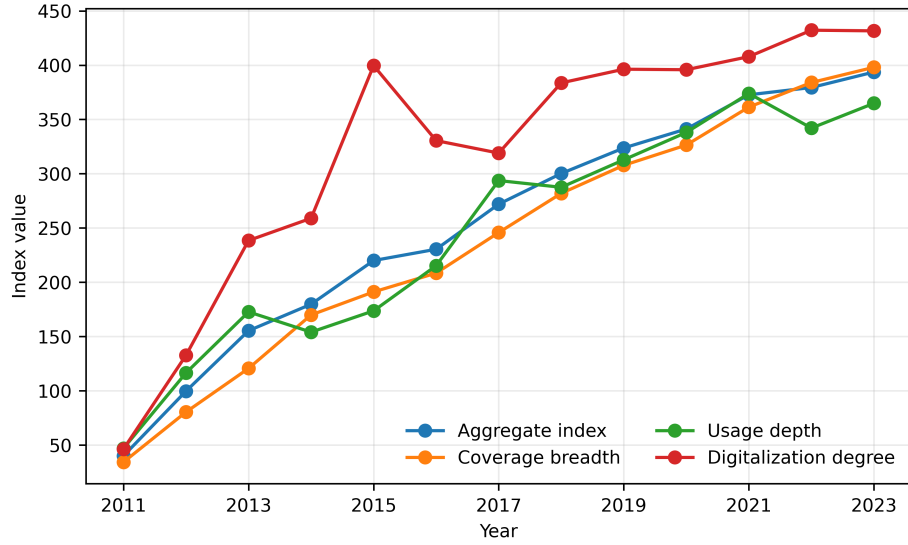


Fig. 5 National mean trajectories of the four coordinates across 2011 through 2023.

6.2 Finding F1: relative convergence unfolds amid absolute divergence

Finding F1. Under near tenfold scaling, the national coordination pair moves oppositely in absolute and relative units. The absolute gap rises from 17.74 to 34.41 while the balance index climbs from 0.707 to 0.953 and coupling strengthens from 0.920 to 0.998. *Only the triad of novelty point N1 returns this result, because either single statistic reverses the conclusion of the other.*

Table 3 disaggregates by region. Eastern ends 2023 highest in mean level at 428.14 despite the slowest percentage gain of 604.8 percent, while central gains 1101.8 percent and western gains 1232.3 percent. Coordination tells a different story than levels. Western ends with the largest gap at 49.96 against 20.19 in eastern regions and the lowest balance at 0.929 after widening by 33.39 points. Northeastern widens by 26.36 points. Central stays flat between 28.65 and 28.38. Eastern drifts up modestly from 16.32 to 20.19 while holding the best balance at 0.975. Figures 6 through 8 display these

trajectories. An analyst holding only the absolute gap concludes that regional mal-function deepened everywhere except centrally. An analyst holding only B concludes full convergence. The paired readout shows both statements are claims about scale, exactly the behavior anticipated by Remark 1 and measured here with magnitudes.

Table 3 Regional endpoint comparison of mean index and coordination pair from Module A.

Region	Index 2011	Index 2023	Index change (%)	Gap 2011	Gap 2023	Balance 2023
East	60.74	428.14	604.8%	16.32	20.19	0.975
Central	32.85	394.82	1101.8%	28.65	28.38	0.962
West	27.85	371.05	1232.3%	16.58	49.96	0.929
Northeast	33.79	367.28	986.8%	5.26	31.62	0.955

Note: gap averages $|U - C|$ within region and year, and balance equals one minus the gap divided by the sum of U and C .

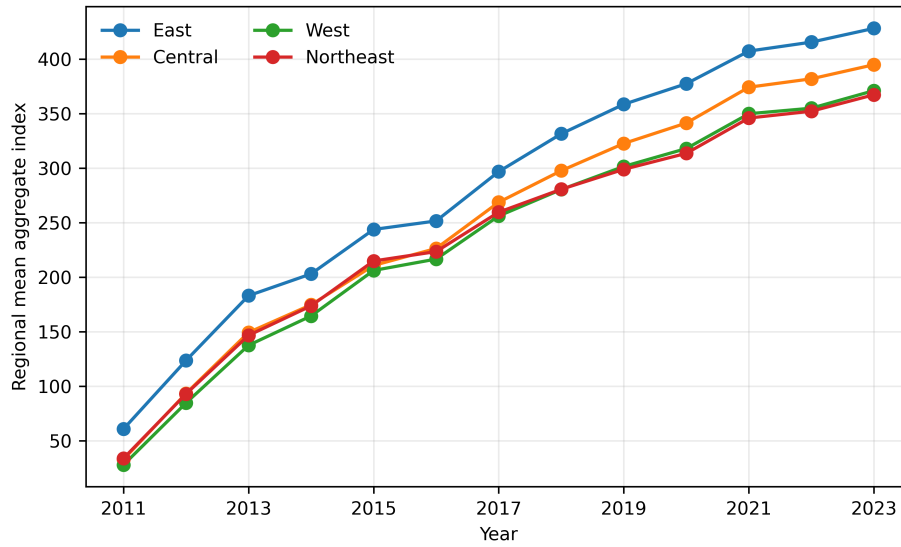


Fig. 6 Regional mean aggregate index trajectories across 2011 through 2023.

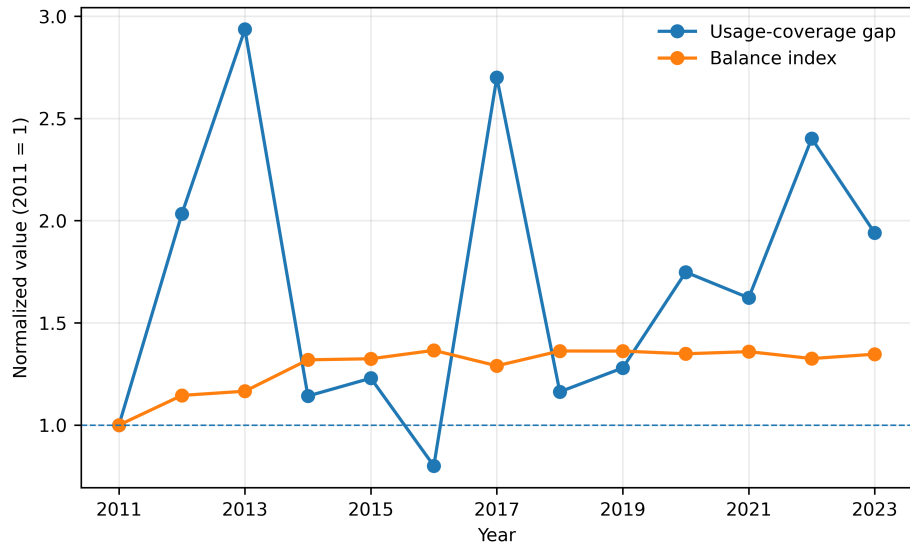


Fig. 7 National gap and balance trends normalized to a 2011 value of one.

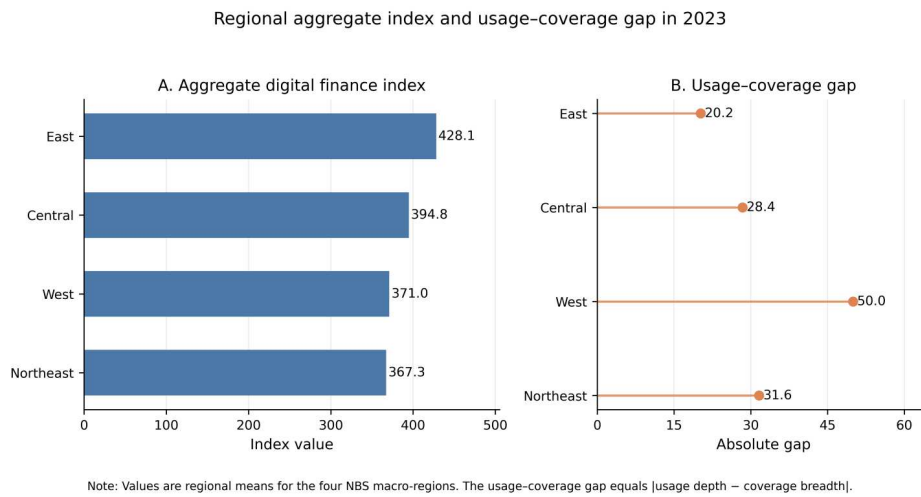


Fig. 8 Endpoint snapshot of regional mean index versus mean usage coverage gap in 2023.

6.3 Finding F2a: dispersion collapses and the survivor concentrates between regions

Module B quantifies the level story in Table 4. Dispersion falls along every metric, with the coefficient of variation declining from 0.458 to 0.082, the Gini index from

0.248 to 0.045, and the Theil index from 0.0982 to 0.0033. Agreement in direction and magnitude across three independent statistics rules out metric artifacts. Decomposition through Eq. (7) localizes survivorship geographically. The between region share stands at 63.0 percent in 2011 and 63.1 percent in 2023 with a mid sample dip to 52.9 percent in 2016. Shrinking totals did not dissolve the regional axis of inequality but concentrated upon it. Figure 9 overlays the normalized series, and leader identity oscillates only between Shanghai and Beijing throughout the panel.

Table 4 Annual inequality metrics and between region Theil share from Module B.

Year	Mean index	CV	Gini	Theil	Between share	Top province	Bottom province
2011	40.00	0.458	0.248	0.0982	0.630	Shanghai	Xizang
2016	230.41	0.090	0.048	0.0040	0.529	Beijing	Qinghai
2019	323.73	0.101	0.054	0.0050	0.610	Shanghai	Qinghai
2020	341.22	0.100	0.054	0.0049	0.618	Shanghai	Qinghai
2023	393.70	0.082	0.045	0.0033	0.631	Shanghai	Qinghai

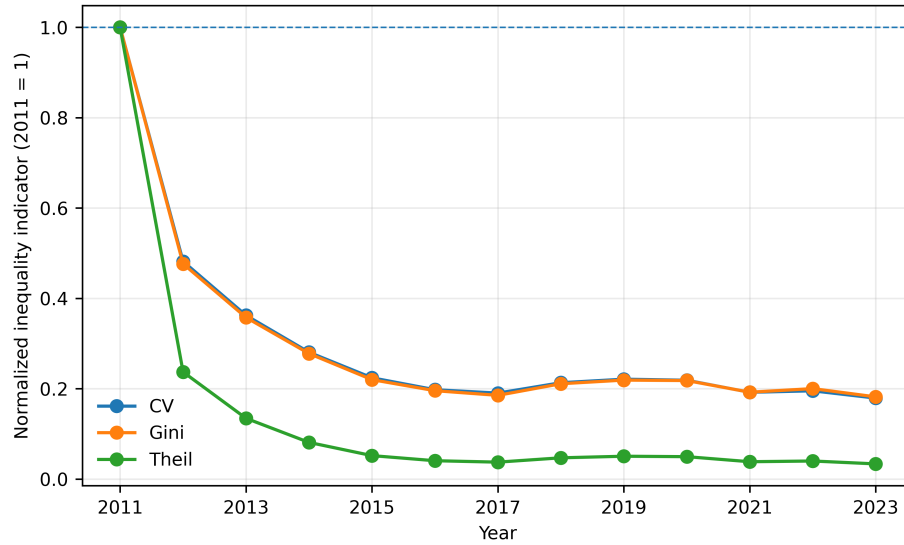


Fig. 9 Sigma convergence shown as disparity metrics normalized to a 2011 value of one.

6.4 Finding F2b: beta convergence holds for every coordinate at dimension specific speeds

Finding F2. Adjustment rates differ across coordinates by factors of two to three, and usage depth is the slow layer. *Comparability itself is an output of novelty point N2, because raw betas become commensurable only after transformation through Eq. (10).*

Both estimators reject nonconvergence for all five targets in Table 5. Cross section betas span -0.0712 for the aggregate to -0.0855 for digitalization, all significant at the 0.001 level with tiny robust errors, and panel betas with year and region effects and cluster errors agree in sign and significance. The transform of Eq. (10) then separates them sharply. Digitalization degree adjusts at rate 1.662 per year with a half life of 0.42 years. Coverage breadth adjusts at 1.059 with 0.65 years. The usage coverage gap adjusts at 0.935 with 0.74 years. Usage depth adjusts at only 0.534 with a half life of 1.30 years. The ordering admits a direct reading. Layers carried by capital and platform rollout behave like infrastructure, while usage depth, meaning behavioral adoption of credit, insurance, and investment functions, behaves like culture and moves slowly. Readiness certificates keyed to aggregate scores inherit the fast layers and silently ignore the slow one.

Table 5 Convergence estimates and rate transform from Module C.

Coordinate	Cross sect. β	Robust SE	Panel β	Cluster SE	Rate λ	Half life
Aggregate index	-0.0712***	0.0016	-0.4639***	0.0408	0.624	1.11
Coverage breadth	-0.0795***	0.0008	-0.6531***	0.0460	1.059	0.65
Usage depth	-0.0728***	0.0026	-0.4138***	0.0584	0.534	1.30
Digitalization degree	-0.0855***	0.0010	-0.8103***	0.0363	1.662	0.42
Usage coverage gap	n.a.	n.a.	-0.6073***	0.0343	0.935	0.74

Note: rate and half life derive from the panel coefficient through Eq. (10). Asterisks mark p below 0.001, 0.01, and 0.05 respectively.

6.5 Finding F3: level equality coexists with sticky relative clubs

Finding F3. Once levels are removed by quartile ranking, positions barely move. Low state persistence reaches 0.844 and high state persistence reaches 0.958. *Pairing this evidence with F2 required the joint reporting interface of the pipeline, because dispersion statistics and transition matrices historically lived in separate literatures.*

Module D delivers the complement of F2b in Table 6 and Figure 10. Direct jumps longer than one step are rare, with low to upper middle at 0.010 and low to high at 0.000, and downward analogues vanish likewise since high to lower middle equals 0.000. Diagonal dominance holds uniformly at intermediate states with 0.760 and 0.833. Read jointly with F2a, absolute catching up exhausted itself in levels and not in orderings. Each tier improved within itself while the geography of advantage persisted at higher means. Findings F2a, F2b, and F3 together falsify the certificate that falling dispersion solves the regional problem, because residual inequality remains regional, the hierarchy behind it is stable, and the slow coordinate is usage.

Table 6 Adjacent year Markov transition probabilities over quartile states from 372 transitions produced by Module D.

Current state	Next low	Next lower middle	Next upper middle	Next high
Low	0.844	0.146	0.010	0.000
Lower middle	0.146	0.760	0.094	0.000
Upper middle	0.012	0.107	0.833	0.048
High	0.000	0.000	0.042	0.958

6.6 Finding F4: growth proved resilient everywhere while structural resilience did not

Finding F4. Over the stress window every region kept growing while western and northeastern coordination degraded, with gap drift of plus 25.80 and plus 27.81 points respectively, whereas eastern coordination improved with minus 4.26 points. *This*

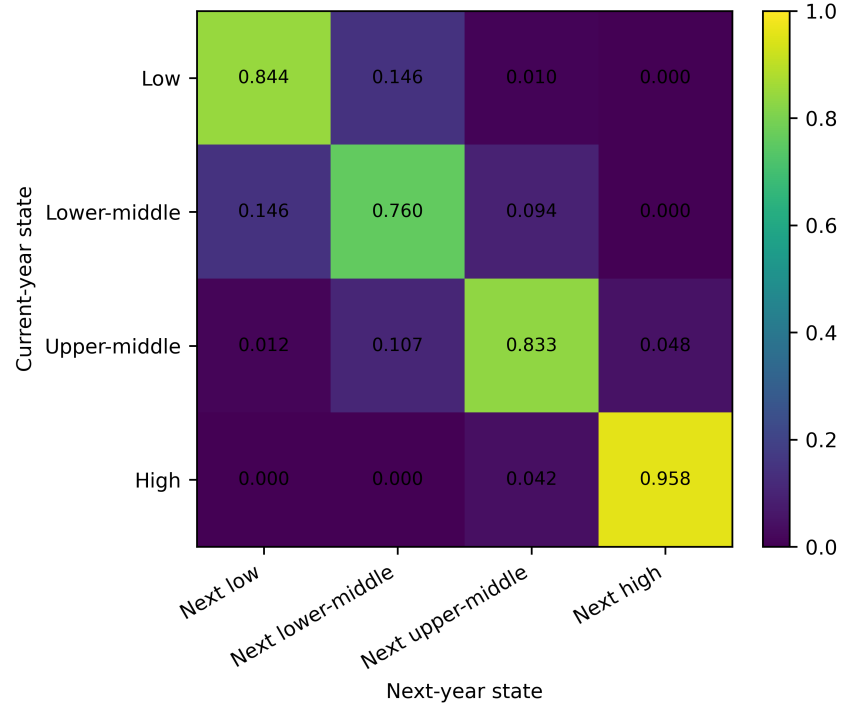


Fig. 10 Estimated transition heat map of relative states from Module D.

asymmetry became visible only through the dual metric of novelty point N3, since a single growth verdict would have passed every region.

Table 7 reports Module E outputs. Every region grew through 2019 to 2020 with rates between 4.90 and 5.81 percent and accelerated into 2020 to 2023 with rates between 13.44 and 17.14 percent. Cumulatively western at 23.05 percent and north-eastern at 22.88 percent outpace eastern at 19.42 percent. Drift from Eq. (13) then splits the verdict. Nationally gap drift is plus 11.71 with balance down 0.011. Eastern passes Definition 1 strictly with gap down 4.26 and balance up 0.008. Central is quasi neutral with gap up only 2.10 and balance up 0.005. Western and northeastern fail decisively with gap up 25.80 against balance down 0.027, and gap up 27.81 against

balance down 0.038. Combined with F3, failure concentrated exactly in the tiers farthest from the frontier. The stress window amplified rather than closed the structural margin first measured by F1.

Table 7 Stress window resilience outputs of growth percentages and coordination drift from Module E.

Region	Growth 2019 to 2020	Growth 2020 to 2023	Cumulative	Drift of $\bar{G}$	Drift of $\bar{B}$
National	5.40%	15.38%	21.61%	+11.71	-0.011
East	5.27%	13.44%	19.42%	-4.26	+0.008
Central	5.81%	15.67%	22.39%	+2.10	+0.005
West	5.44%	16.70%	23.05%	+25.80	-0.027
Northeast	4.90%	17.14%	22.88%	+27.81	-0.038

Note: growth refers to the regional mean aggregate index, and drift compares 2023 with 2019 following Eqs. (12) and (13).

6.7 Robustness diagnostics

Every finding above rests on a point estimate, and this section stress tests all of them. Four checks are run against the same cleaned panel with fixed seeds so each reported number reproduces exactly.

F4 under label permutation. The asymmetry verdict compares province level gap drift over 2019 to 2023 between the pooled western and northeastern group and the eastern group. Observed means are plus 25.80 for twelve western provinces with dispersion 18.48, plus 27.81 for three northeastern provinces with dispersion 8.07, and minus 4.26 for ten eastern provinces with dispersion 24.54, giving an observed difference of **30.46 points**. Randomly reassigning group labels across the twenty five provinces in those two groups for twenty thousand draws yields a one sided permutation probability of **0.0011**, close to the resolution floor of the design, and a bootstrap interval over within group resampling spans **13.46 to 46.79 points**, far from zero. A distribution free Mann Whitney comparison of western against eastern drift agrees

at p equal to 0.0031. Finding F4 therefore passes the decisive test with roughly forty times the conventional evidence requirement rather than merely clearing it.

F2 across sub windows. Table 8 refits Module C on nonoverlapping windows. The dimension ordering survives intact because digitalization degree remains fastest and usage depth slowest within both halves. Rates collapse in absolute size during 2017 to 2023 as gaps close toward maturity, which lengthens half lives mechanically without touching the ordering. Finding F2 is thus a property of the system rather than an artifact of full window averaging.

Table 8 Sub window adjustment rates and half lives from Module C.

Coordinate	λ 2011 to 2016	Half life	λ 2017 to 2023	Half life
Aggregate index	0.696	1.00	0.053	13.02
Coverage breadth	1.135	0.61	0.132	5.24
Usage depth	0.644	1.08	0.040	17.26
Digitalization degree	2.069	0.34	0.452	1.53

Note: year and region effects with province clustered errors, mirroring Table 5. Windows cover model years 2012 through 2016 and 2017 through 2023 on lagged levels.

Markov anchoring and grid sensitivity. At the baseline quartile grid every diagonal exceeds its strongest off diagonal competitor with z statistics of 13.50, 10.87, 13.74, and 31.77 respectively, so each relative state is anchored beyond binomial noise, and row independence is rejected decisively by a Pearson statistic of 732.8 on nine degrees of freedom. Table 9 then moves the quantile grid. Top state persistence stays above 0.93 everywhere while bottom persistence declines mildly from 0.844 to 0.764 as finer cuts make upward neighbours more numerous, a geometric effect expected of any ladder. The stickiness conclusion is therefore not a four way slicing choice.

Conditioning transitions on the prosperity context of same region peers sharpens the picture descriptively. With rich peers the high state keeps 97.7 percent of its members against 80.0 percent amid poor peers, and low states escape far more easily near rich peers, with low to low persistence falling from 88.5 percent under poor peer

contexts to 44.4 percent under rich ones. Clubs hence interact with geography instead of being pure spatial spillover ghosts, although cell sparsity cautions against reading the conditioned matrices as precise parameters.

Table 9 Markov robustness across quantile grids with calibrated assignment.

Grid q	Transitions	Low persistence (SE)	High persistence (SE)	Trace	Independence
4	372	0.844 (0.037)	0.958 (0.020)	3.40	$p < 10^{-151}$
5	372	0.810 (0.043)	0.972 (0.019)	4.02	$p < 10^{-182}$
6	372	0.764 (0.050)	0.933 (0.032)	4.40	$p < 10^{-186}$

Note: quantile boundaries use linear interpolation with left side assignment, exactly reproducing Table 6. Trace sums diagonal probabilities, whose independence benchmark equals one per column of mass.

Trend awareness and the fixed effect guard. Raw log levels trend strongly upward, so classical unit root testing carries limited information at thirteen annual observations. Constant specification Fisher statistics reject unit roots for every coordinate, and trend augmented versions do so as well except usage depth, whose log series shows no significant detrended reversion and we report this openly. The stronger guard is dynamic: regressing growth on its own lag inside provinces with year effects estimates reversion purely from within variation where common trends cannot confound it. Table 10 shows negative significant slopes for every coordinate with the fast versus slow ordering of finding F2 preserved and small sample dynamic bias of order one over T acknowledged as a few hundredths at most. Half life statements thereby stand conditional on specified dynamics, which the method section now states explicitly.

Collectively the battery converts all four findings from single shot point claims into checked results, and it strengthens the paper’s central innovation claim because the decoupling instrument N1, the rate transform N2, and the dual metric verdict N3 each earn their place under stress rather than by construction alone.

Table 10 Unit root awareness and the fixed effect guard from Section 6.7.

Coordinate	Fisher p constant	Fisher p constant plus trend	FE slope $\hat{\beta}$ (SE)
Aggregate index	$< 10^{-6}$	$< 10^{-19}$	-0.636 (0.028)
Coverage breadth	$< 10^{-31}$	$< 10^{-256}$	-0.727 (0.038)
Usage depth	$< 10^{-113}$	0.984	-0.632 (0.030)
Digitalization degree	$< 10^{-253}$	$< 10^{-88}$	-0.803 (0.038)

Note: Fisher columns combine thirty one provincewise ADF probabilities into Maddala and Wu chi square statistics with sixty two degrees of freedom. FE slope comes from province and year fixed effects with clustered errors.

7 Conclusions

We formulated diagnosis of multidimensional digital finance panels as four estimation tasks and designed LIMR, a five module deterministic pipeline that reads raw coordinates rather than composite scores. Three elements define its novelty. The indicator triad decouples absolute mismatch from relative balance and thereby prevents reversal artifacts. The rate transform makes convergence speeds commensurable across dimensions. The dual metric resilience verdict separates growth continuity from structural integrity. Execution on the released PKU-DFIIC panel for 31 provinces across 2011 through 2023 returned four mutually reinforcing findings, and every one survived label permutation, subsample splitting, quantile grid variation, spatial conditioning, and unit root aware re estimation. Absolute and relative coordination move oppositely under scaling. Dispersion collapsed, yet its surviving third concentrates between macro regions while convergence proceeds two to three times faster on infrastructural coordinates than on usage depth. Ranking order mobility stays nearly frozen regardless of those gains. The 2019 through 2023 stress window exposed asymmetric structural resilience that degraded exactly the lagging regions growing fastest. Jointly these findings demonstrate constructively that a composite index serves as a growth certificate but never as a structure certificate nor a readiness certificate for AI-era financial systems [27, 28].

The instrument carries deliberate limits that define next steps. All estimators are unconditional, so the pipeline attributes structure without explaining causes. Extending Module C with covariate shifts or Module E with matched window identification over the post 2023 policy period proceeds within the same interface. Panel granularity of 31 units constrains subgroup resolution, and city or county replications require no code changes beyond the input loader. Functional subindices entered descriptives here and can move into per function convergence and mobility stacks as a natural extension. We release the cleaned panel and scripts so that each extension reproduces from data alone.

A Reproducibility notes and annual artifacts

All modules are deterministic and depend only on the cleaned panel constructed from the PKU-DFIIC province appendix for 2011 through 2023, comprising 403 rows that carry province, year, region, the four coordinates, functional subindices, and constructed gap, balance, coupling, and annual growth columns. Python 3.13 with pandas 2.2.3, NumPy 2.3.5, statsmodels 0.14.6, and Matplotlib 3.10.8 reproduces every printed digit. The robustness battery of Section 6.7 is produced by companion scripts covering label permutation with twenty thousand seeded draws, subwindow refits, calibrated quantile grids, spatially conditioned chains, block weighted Moran statistics, Fisher type unit root combination tests, and the fixed effect guard. Because the sample precedes formal deepening of the AI Plus agenda, the artifact set forms a pre policy reference baseline. Table 11 lists the annual series behind Figures 5 through 9 for direct verification.

References

- [1] Guo, F., Wang, J., Cao, Y., Cheng, Z., Li, Y., Wang, F.: Peking University Digital Financial Inclusion Index of China (2011-2023) (2024). <https://idf.pku.edu.cn/>

Table 11 Annual system level artifacts from Modules A and B.

Year	Mean index	Mean gap	Mean balance	CV	Gini	Theil
2011	40.00	17.74	0.707	0.458	0.248	0.0982
2012	99.69	36.08	0.810	0.220	0.118	0.0233
2013	155.35	52.07	0.825	0.166	0.089	0.0132
2014	179.75	20.26	0.933	0.128	0.069	0.0080
2015	220.01	21.80	0.937	0.103	0.055	0.0051
2016	230.41	14.19	0.966	0.090	0.048	0.0040
2017	271.98	47.90	0.913	0.087	0.046	0.0037
2018	300.21	20.62	0.964	0.098	0.052	0.0046
2019	323.73	22.70	0.963	0.101	0.054	0.0050
2020	341.22	31.01	0.954	0.100	0.054	0.0049
2021	372.72	28.78	0.961	0.088	0.048	0.0038
2022	379.44	42.60	0.938	0.089	0.050	0.0039
2023	393.70	34.41	0.953	0.082	0.045	0.0033

<docs/2026-03/064a456918cf4cb789d33d6e49952e5a.pdf> Accessed 2026-05-12

- [2] Li, J., Wu, Y., Xiao, J.J.: The impact of digital finance on household consumption: Evidence from china. *Economic Modelling* **86**, 317–326 (2020)
- [3] Zhang, X., Tan, Y., Hu, Z., Wang, C., Wan, G.: The trickle-down effect of fintech development: From the perspective of urbanization. *China & World Economy* **28**, 23–40 (2020)
- [4] Beck, T., Demirguc-Kunt, A., Levine, R.: Finance, inequality and the poor. *Journal of Economic Growth* **12**, 27–49 (2007)
- [5] Jiang, Q., Li, Y., Si, H.: Digital economy development and the urban-rural income gap: Intensifying or reducing. *Land* **11**, 1980 (2022)
- [6] Hau, H., Huang, Y., Lin, C., Shan, H., Sheng, Z., Wei, L.: Fintech credit and entrepreneurial growth. *Journal of Finance* **79**, 3309–3359 (2024)
- [7] Xu, Q., Zhong, M., Dong, Y.: Digital finance and rural revitalization: Empirical test and mechanism discussion. *Technological Forecasting and Social Change* **201**, 123248 (2024)

- [8] Allen, F., Demirguc-Kunt, A., Klapper, L., Martinez Peria, M.S.: The foundations of financial inclusion: Understanding ownership and use of formal accounts. *Journal of Financial Intermediation* **27**, 1–30 (2016)
- [9] Sarma, M.: Index of financial inclusion: A measure of financial sector inclusiveness. Technical Report 07/2012, Berlin Working Papers on Money, Finance, Trade and Development (2012)
- [10] Ozili, P.K.: Impact of digital finance on financial inclusion and stability. *Borsa Istanbul Review* **18**, 329–340 (2018)
- [11] Demirguc-Kunt, A., Klapper, L., Singer, D., Ansar, S., Hess, J.: The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution. World Bank, Washington, DC, USA (2018)
- [12] Scheerder, A., Deursen, A., Dijk, J.: Determinants of internet skills, uses and outcomes: A systematic review of the second- and third-level digital divide. *Telematics and Informatics* **34**, 1607–1624 (2017)
- [13] Xiao, Y., Yin, M., Wang, H., Xiang, Y.: Digital finance, digital usage divide, and urban-rural income gap: Evidence from china. *Systems* **13**, 145 (2025)
- [14] Ji, X., Wang, K., Xu, H., Li, M.: Has digital financial inclusion narrowed the urban-rural income gap? the role of entrepreneurship in china. *Sustainability* **13**, 8292 (2021)
- [15] Peng, Z., Dan, T.: Digital dividend or digital divide? digital economy and urban-rural income inequality in china. *Telecommunications Policy* **47**, 102616 (2023)
- [16] Dong, Y., Luo, W., Zhang, X.: Information and communication technology diffusion and the urban-rural income gap in china. *Pacific Economic Review* **29**,

159–186 (2023)

- [17] Barro, R.J., Sala-i-Martin, X.: Convergence. *Journal of Political Economy* **100**, 223–251 (1992)
- [18] Quah, D.: Empirical cross-section dynamics in economic growth. *European Economic Review* **37**, 426–434 (1993)
- [19] Theil, H.: *Economics and Information Theory*. North-Holland, Amsterdam, The Netherlands (1967)
- [20] Martin, R.: Regional economic resilience, hysteresis and recessionary shocks. *Journal of Economic Geography* **12**, 1–32 (2012)
- [21] Folke, C.: Resilience: The emergence of a perspective for social-ecological systems analyses. *Global Environmental Change* **16**, 253–267 (2006)
- [22] Haddad, C., Hornuf, L.: The emergence of the global fintech market: Economic and technological determinants. *Small Business Economics* **53**, 81–105 (2019)
- [23] Luo, D., Luo, M., Lv, J.: Can digital finance contribute to the promotion of financial sustainability? a financial efficiency perspective. *Sustainability* **14**, 3979 (2022)
- [24] Gomber, P., Koch, J.-A., Siering, M.: Digital finance and fintech: Current research and future research directions. *Journal of Business Economics* **87**, 537–580 (2017)
- [25] Sahay, R., Allmen, U.E., Lahreche, A., Khera, P., Ogawa, S., Bazarbash, M., Beaton, K.: The promise of fintech: Financial inclusion in the post covid-19 era. Technical Report 20/09, International Monetary Fund (2020)
- [26] National Bureau of Statistics of China: How Are China’s Economic Regions

Divided? (2023). https://www.stats.gov.cn/zt_18555/zthd/sjtjr/d12kfr/tjzsqzs/202302/t20230216_1908940.html Accessed 2026-05-12

- [27] State Council of the People's Republic of China: Opinions of the State Council on Deepening the Implementation of the AI Plus Initiative (2025). https://www.cac.gov.cn/2025-08/27/c_1758018277755538.htm Accessed 2026-05-12
- [28] State Council of the People's Republic of China: Report on the Work of the Government (2026). <https://english.www.gov.cn/atts/stream/files/69b4a963c6d0c788099000f2> Accessed 2026-05-12