Backwash or Spread? Absorptive Capacity and the Nonlinear Mechanism of Regional Convergence in China

GU LUOLUO[1]

Faculty of Economics, Thammasat University, Bangkok, Thailand

MONTHIEN SATIMANON

Faculty of Economics, Thammasat University, Bangkok, Thailand

Despite prolonged government interventions, China's regional economy continues to exhibit a "convergence puzzle," characterized by persistent club solidification. Existing literature relies on the assumption of homogeneous spatial spillovers, overlooking how local absorptive capacity conditions the direction and size of external benefits. This paper incorporates endogenous technology diffusion into a neoclassical growth framework and estimates the moderating role of tech density in a Spatial Durbin specification with interaction terms, using Chinese provincial panel data and an identification-robust empirical design.

The results indicate: (1) Distribution dynamics over 1992–2024 confirm pronounced club solidification, with retention probabilities of 0.933 for the low-income club and 0.990 for the high-income club. (2) On the 2011–2019 benchmark window, fixed-effects estimates of the moderation coefficient are attenuated toward zero, a pattern consistent with measurement error in the interaction term; an instrumental-variable strategy based on 2010 baseline tech density recovers $\theta_2 = 0.048$ (cluster-robust standard error 0.011), with a Kleibergen–Paap first-stage F of 30.5, an Anderson–Rubin weak-instrument-robust p-value of 0.0007, and wild cluster bootstrap p-values between 0.018 and 0.057 across weight schemes. (3) A Hansen-type grid search finds no sharp threshold: the residual sum of squares varies by only 3.3% across 21 candidate split points, and the regime-specific coefficients are statistically indistinguishable from each other. The moderation is a continuous and era-dependent mechanism: windows extending back to the 1990s yield a significantly negative interaction, while the positive moderation concentrates in the digital-era 2010s. Policy should therefore aim at strengthening local absorptive capacity to move provinces along the moderation curve, rather than at crossing a specific numerical threshold.

Keywords: Regional convergence; Club convergence; Spatial spillovers; Absorptive capacity; Tech density; Instrumental variables; Backwash effect; Spread effect; Spatial Durbin Model; China

JEL Classification: O18, O33, R11, R12, R58

1. Introduction

1.1 The Convergence Puzzle under Strong Policy Intervention

Since the implementation of the Reform and Opening-up policy, China has achieved an economic growth miracle that has captured global attention. However, persistent regional disparity remains a critical bottleneck constraining high-quality development. According to

[1] Gu Luoluo, Corresponding Author: Thai-Chinese International School of Management, University of the Thai Chamber of Commerce and Faculty of Economics, Thammasat University, Bangkok, Thailand, gu.luo@st.econ.tu.ac.th; Satimanon: Faculty of Economics, Thammasat University, Bangkok, Thailand Email: monthien@econ.tu.ac.th
Acknowledgements: The authors thank the anonymous reviewers for their constructive comments. All remaining errors are the authors' own.

Neoclassical Growth Theory (Solow, 1956), driven by the law of diminishing marginal returns to capital, backward regions should theoretically possess higher returns on capital than developed regions, thereby achieving faster catch-up growth.

However, scrutinizing the evolutionary trajectory of China's provincial economies from 1992 to 2024, we confront a perplexing "Convergence Puzzle."

On one hand, the Chinese government has implemented the largest regional balancing strategy in human history. From the Western Development Strategy (1999) and the Rise of Central China (2004) to recent targeted poverty alleviation and fiscal transfer payments, the central government has attempted to bridge the regional divide through forceful administrative intervention. Theoretically, such a "Big Push" should have significantly accelerated the process of β-convergence.

On the other hand, the data reveals a stubborn characteristic of "Club Convergence." Both existing literature and the Kernel Density Estimation (KDE) presented in this paper show that China's provincial income distribution exhibits a pronounced right-skewed pattern, with the majority of provinces clustering below the national mean while a small elite group maintains a structural lead in the upper tail. This polarization is not unique to China but represents a common challenge for many emerging economies during rapid industrialization. For instance, Warr (2013) documents persistent Bangkok-periphery gaps in Thailand despite decades of government interventions, with structural factors including educational deficiencies and infrastructure imbalances constraining long-term regional convergence. Similarly, empirical research by Paweenawat and McNown (2014) on income inequality determinants in Thailand using synthetic cohort analysis confirms that educational dispersion significantly affects income distribution patterns and regional mobility.

This implies that mere factor accumulation has failed to drive automatic catch-up in backward provinces. Instead, some central and western provinces have fallen into a "Low-Level Equilibrium Trap," where the gap with their eastern neighbors has not narrowed but, in some periods, widened. Why have forceful fiscal equalization policies failed to break the long-term pattern of regional differentiation? Why can some inland provinces (like Anhui) successfully embed themselves into the coastal industrial chain to achieve a leap, while others with similar geographic locations suffer from continuous "siphoning" under the shadow of developed neighbors? This is not only a core puzzle in China's regional economics but also a shared challenge for developing countries striving to cross the "Middle-Income Trap."

1.2 Literature Gaps and Reflections

Existing literature attempts to explain the regional convergence dilemma from various dimensions but generally suffers from two logical disconnects that limit its explanatory power for the complex dynamics in China and other emerging economies like Thailand.

First, the assumption of "homogeneous" spatial interactions and the debate on geographic distance.

New Economic Geography (NEG) emphasizes that spatial interaction is key to driving growth (Krugman, 1991). However, traditional empirical studies (LeSage & Pace, 2009) often assume that Spatial Spillovers are linear and positive. This presumption of a "trickle-down effect" obscures the complexity of regional competition. Although recent studies have explored how the digital economy mitigates geographic constraints through enhanced connectivity and information transmission (Tao et al., 2022), for "hard tech" innovation relying on physical industrial chains, the agglomeration effect of physical space remains irreplaceable. As pointed out by Bai et al. (2012), spatial spillovers in China exhibit significant regional heterogeneity. Recent research further finds that due to the heterogeneity of absorptive capacity among different cities, spatial spillovers exhibit significant "asymmetric" and "localized" characteristics (Yang & Lin, 2012). Existing models often report only an average coefficient, ignoring the dynamic nature of the transition between "siphoning" (backwash) and "radiation" (spread). A parallel empirical literature identifies the causal effects of market access, connective infrastructure, and spatial frictions on regional outcomes (Faber, 2014; Donaldson & Hornbeck, 2016; Fan, 2019; Hsieh & Moretti, 2019; Moretti, 2021). That work establishes that spatial linkages matter causally, but it leaves open how their direction depends on the receiving region's absorptive capacity—the state-dependence examined here.

Second, the "black box" treatment of threshold mechanisms and the measurement of absorptive capacity.

The concept of absorptive capacity originates from Cohen and Levinthal (1990), who define it as a firm's ability to recognize, assimilate, and apply external knowledge. While research on "threshold effects" highlights the importance of human capital or institutional environments (Hansen, 2000), it often treats regions as isolated systems, failing to clarify how thresholds change the *nature* of spatial interactions. Hsieh and Klenow (2009) argue that resource misallocation is key to Total Factor Productivity (TFP) losses, while recent empirical evidence from Ma et al. (2023) suggests that severe misallocation of innovation resources often stems from backward regions failing to meet specific technological thresholds, preventing effective integration into regional innovation networks. For example, what critical level must a province's technological "hard power" reach to reverse the "Agglomeration Shadow" caused by the brain drain and capital flight, and instead enjoy technological radiation from neighbors? Under the current drive for coordinated regional development, precisely identifying this break-even point from "backwash" to "spread" (Chen & Partridge, 2013) remains a "black box" in existing linear or simple threshold models.

1.3 Research Logic and Marginal Contributions

Addressing these gaps, this paper constructs a growth model incorporating nonlinear spatial spillovers and endogenous technology thresholds, utilizing panel data from 31 Chinese provincial administrative regions (1992–2024) to open the "black box" of the convergence trap. The marginal contributions of this paper are primarily reflected in the following three aspects:

1. A Capacity-Moderated Spillover Framework: Endogenizing the Direction of Spatial Spillovers

This paper departs from the traditional linear convergence framework by proposing the "Absorptive Capacity Moderation Hypothesis." We characterise a break-even capacity level γ:

- When local tech density $H_{it} < \gamma$, developed neighbors act as "siphons." Due to the excessive technological gap, the local region cannot capture industrial spillovers and instead suffers from resource outflows, resulting in negative spillovers.
- When $H_{it} \geq \gamma$, the local region possesses the complementary capacity to transform developed neighbors into "technology sources," driving growth through supply chain spillovers.

This framework successfully reconciles the contradiction between Neoclassical Growth Theory (emphasizing diffusion) and Core-Periphery Theory (Krugman, 1991; Myrdal, 1957), providing a new perspective for understanding the "Center-Periphery" structure within developing countries.

2. Identification Strategy: Identification-Robust Estimation of the Moderation

This paper employs "Tech Density" as the core moderator variable and estimates the Spatial Durbin interaction specification under an identification-robust design: a baseline-interacted instrumental variable addresses the endogeneity of tech density, wild cluster bootstrap and two-way clustering guard small-sample inference, and a Hansen-type grid search tests—and rejects—a sharp-threshold representation of the mechanism. The estimates imply a continuous transition between backwash-dominant and spread-dominant spillovers rather than a knife-edge regime boundary, and they locate the positive moderation in the digital-era 2010s.

3. Policy Reorientation: From Gap Filling to Capacity Building

Based on the empirical results, this paper proposes that regional policy should shift from "Gap Filling" to "Capacity Building." For provinces in the backwash-prone range, "scattergun" style infrastructure investment yields only short-term effects; concentrating resources on technological absorptive capacity moves a province along the moderation curve toward positive externalities. Because the data support a continuous mechanism rather than a discrete threshold, the payoff to such investment accrues gradually rather than at a single crossing point. This conclusion applies not only to China but also holds significant implications for other economies attempting to upgrade industries to avoid the middle-income trap.

1.4 Organization of the Paper

- Section 2 develops a theoretical framework in which spatial spillover effects depend nonlinearly on local absorptive capacity, deriving testable hypotheses for the moderation mechanism.
- Section 3 introduces the econometric model specification, variable selection, and data sources, focusing on identification strategies for endogeneity.
- Section 4 reports the main empirical results, including Moran's I spatial correlation tests, the identification ladder for the spatial moderation estimates, and the implied marginal-effect profile.
- Section 5 presents a suite of identification-robust diagnostics: the instrumental-variable strategy and its first-stage strength, alternative measures of absorptive capacity, sample-window sensitivity, wild cluster bootstrap inference, leverage analysis, two-way clustering, and a Hansen-type grid search for a structural threshold.
- Section 6 summarizes the main conclusions and discusses policy implications for coordinated regional development.

2. Theoretical Framework

2.1 Baseline Setup: From Factor Accumulation to TFP Growth

Consider an open economy system consisting of N regions. For any region i, its output Y_{it} follows a standard Cobb-Douglas production function:

$$Y_{it} = A_{it} K_{it}^{\alpha} L_{it}^{1-\alpha} \tag{1}$$

where A represents generalized Total Factor Productivity (TFP). As China moves past its Lewis turning point, its growth momentum has structurally transitioned from capital accumulation to TFP-driven technological progress.

Defining output per capita as $y = Y/L$, capital accumulation follows the canonical Solow (1956) setup:

$$\dot{k} = s \cdot f(k) - (n + \delta)k \tag{2}$$

2.2 Core Innovation: Nonlinear Diffusion and the Absorptive Threshold

The core innovation of this model lies in reconstructing the dynamic evolutionary path of A_{it}. Drawing on the technology frontier model of Acemoglu et al. (2006) and incorporating recent discussions by Tao et al. (2022) on how digital connectivity reshapes spatial innovation patterns, we specify the growth rate of technology as:

$$\frac{\dot{A}it}{Ait} = g + \Phi(H_{it}) \cdot \sum_{j \neq i} w_{ij} l\, n \left(\frac{A_{jt}}{A_{it}}\right) \tag{3}$$

Traditional linear models often assume that spatial spillovers are homogeneously positive. However, Fu (2008) highlights that absorptive capacity significantly moderates regional innovation outcomes in China. Building on the seminal framework of Cohen and Levinthal (1990), we model the spillover moderation coefficient $\Phi(H_{it})$ as a smooth, monotonically increasing function of local tech density H, with a single zero crossing at the break-even level γ.

Here, γ is the break-even level at which the marginal spillover effect crosses zero. The micro-economic intuition behind this moderation function stems from fixed costs and network externalities in technology adoption:

1. The Siphon Zone ($\Phi < 0$): When $H < \gamma$, the region lacks the complementary assets to support high-tech industries. The fixed costs of technology adoption outweigh the benefits, causing high-quality factors (talent and capital) to flow towards neighbors with higher returns. This results in a "Backwash Effect," consistent with Myrdal (1957) cumulative causation theory.

2. The Spread Zone ($\Phi > 0$): Once H moves above the break-even level γ, the regional innovation ecosystem reaches a "Critical Mass." Network externalities reduce the marginal cost of learning, enabling the region to effectively embed itself in the value chain and convert the "spatial gap" into a catch-up dividend. This supports Williamson (1965)'s hypothesis about eventual regional convergence through spread effects.

2.3 Theoretical Propositions and Testable Hypotheses

Based on the framework developed above, we differentiate the growth equation with respect to the spatial interaction term to derive the marginal response function of local economic growth to neighboring growth:

$$\frac{\partial g_i}{\partial (Wy)_i} = \theta_1 + \theta_2 \cdot H_i \tag{4}$$

Given the property of the moderation function $\Phi(H)$ that $\partial\Phi/\partial H > 0$, the spillover effect is a strictly monotonically increasing function of absorptive capacity. Consequently, we translate these mathematical derivations into two specific econometric hypotheses to be tested in Section 4:

Proposition 1 (The Nonlinear Moderation Mechanism):

Spatial spillover effects are not homogeneous constants but are endogenously moderated by local absorptive capacity (H).

- **Testable Hypothesis 1:** In the Spatial Durbin Model (SDM) with interaction terms, the coefficient of the interaction term ($W \cdot y \times Tech$) should be significantly positive ($\theta_2 > 0$). This implies that as local tech density increases, the marginal contribution of neighboring economic activities to local growth improves.

Proposition 2 (Break-even Capacity and Continuous Sign Reversal):

There exists a break-even capacity level γ at which the marginal spillover effect changes sign. Because the moderation function is smooth and strictly increasing, the change of sign is continuous: the marginal effect passes through zero rather than jumping between discrete regimes.

- **Testable Hypothesis 2:**
  - **The backwash range:** When $H < \gamma$, the total marginal spillover effect is negative ($\theta_1 + \theta_2 \cdot H < 0$). This corresponds to the "Backwash Effect," where backward regions suffer from net resource outflows.
  - **The spread range:** When $H > \gamma$, the total marginal spillover effect becomes positive ($\theta_1 + \theta_2 \cdot H > 0$). At this stage, sufficient absorptive capacity activates the "Spread Effect," facilitating regional convergence. The transition between the two ranges is continuous, so a two-regime step specification should fit no better than the smooth interaction specification—an implication tested directly by the grid search in Section 5.9.

3. Methodology and Data

To comprehensively examine the evolutionary trajectory and structural sources of China's regional economic disparity, this study first employs the Coefficient of Variation (CV) and the Theil Index as foundational measures of inequality.

3.1 Inequality Measures and Decomposition

3.1.1 σ-Convergence: General Dispersion Measure

Sigma (σ) convergence aims to capture the dynamic evolution of the cross-sectional dispersion of regional income levels over time. We employ the Coefficient of Variation (CV) as the primary metric, calculated as follows:

$$CV_t = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_{it} - \overline{y}_t)^2}}{\overline{y}_t} \tag{5}$$

where y_{it} represents the real GDP per capita of province i in year t, $\overline{y_t}$ denotes the national average for that year, and $N = 31$ is the number of provincial administrative units. A declining trend in CV_t over time indicates the presence of σ-convergence.

3.1.2 Theil Index and Spatial Decomposition

To further disentangle the structural sources of regional inequality, we use the additive decomposability property of the Theil Index (Shorrocks, 1980). Following the traditional classification by the National Bureau of Statistics (NBS), China is divided into three major belts: Eastern, Central, and Western ($K = 3$). Consequently, the overall Theil Index can be decomposed into "Between-region Inequality" ($T_{between}$) and "Within-region Inequality" (T_{within}):

$$T_t = T_{between} + T_{within} \tag{6}$$

- $T_{between}$: Captures the development gradient across the Eastern, Central, and Western belts.
- T_{within}: Measures the degree of disequilibrium within each belt.

Based on the above formulas, we plot the long-term evolutionary trend of the Theil Index and its decomposed components (see Appendix Figure A1).

3.2 Spatial Econometric Specification

3.2.1 Construction and Scientific Rationale of the Spatial Weight Matrix (W)

The spatial weight matrix is the cornerstone of spatial econometrics; its validity directly determines the accuracy of spillover identification. In this study, we construct a Queen Contiguity Matrix (W_q) based on geographic adjacency as the baseline weight, strictly adhering to the following criteria:

Topological Contiguity Rule: Queen vs. Rook

We define the element $w_{ij} = 1$ if province i and province j share a common boundary or a single vertex, and 0 otherwise. Compared to the Rook contiguity matrix, which considers only shared boundaries, the Queen matrix captures complex inter-provincial economic linkages more comprehensively, offering higher identification efficiency, particularly for hub regions where multiple provinces converge.

Handling the "Isolated Node": The Hainan Case

Given that Hainan is an island economy lacking physical adjacency to the mainland, strict adherence to contiguity rules would render it an "isolated node," causing the spatial weight matrix to be non-invertible (singular) and failing to capture its economic interaction with the mainland. Following standard academic conventions in Chinese regional economics**,** Lu and Wan (2015) and the treatment methods suggested by Salvati (2025), we manually assign an adjacency link between Hainan and Guangdong provinces. This correction ensures the connectivity of the spatial weight matrix, enabling the model to effectively identify spillover paths along the coastal economic belt.

Row-standardization and Economic Interpretation

To eliminate weighting bias arising from the varying number of neighbors across provinces, we follow the guidelines of Elhorst (2014) to row-standardize the matrix, ensuring $\sum_j w_{ij} = 1$. Consequently, the spatial lag term Wy carries a clear economic interpretation: it represents the weighted average economic development level of neighboring regions, thereby enhancing the interpretability of parameter estimates.

Robustness Checks: Alternative Spatial Weight Matrices

Acknowledging that geographic adjacency may not capture all relevant linkage channels, we construct two alternative matrices for robustness checks, referencing LeSage and Pace (2009): an inverse geographic distance matrix and a five-nearest-neighbour matrix. Appendix Table A1 reports the moderation estimate under both alternatives, together with an SLX specification and a Hainan-exclusion check.

3.2.2 The Spatial Moderation Model: Identifying Interaction Effects

To empirically test the threshold characteristics of "Absorptive Capacity," we construct a Spatial Durbin Model (SDM) augmented with interaction terms:

$$g_{it} = \alpha + \beta ln\left(y_{i,t-1}\right) + \rho W g_{it} + \theta^1 W ln\left(y_{i,t-1}\right) + \theta^2\left(W ln\left(y_{i,t-1}\right) \times Tech_{it}\right) + \delta \mathrm{Tech}_{it} + \mu_{\mathrm{i}} + \eta_{\mathrm{t}} + \varepsilon_{it} \quad (7)$$

Where:

g_{it} denotes the economic growth rate.

$\ln(y_{i,t-1})$represents the initial income level (log-form).

$Tech_{it}$ denotes the standardized Tech Density (Absorptive Capacity).

W is the row-standardized spatial weight matrix.

μ_i and η_t represent province and year fixed effects, respectively.

Economic Interpretation:

This specification unpacks the "black box" of spatial spillovers. The composite coefficient $(\theta_1 + \theta_2 \cdot Tech_{it})$ explicitly captures the state-dependent nature of spatial interactions.

- If $\theta_1 < 0$ and $\theta_2 > 0$, the sign of the total spillover effect depends on whether the local tech density Tech exceeds the break-even level $-\theta_1/\theta_2$.
- Whether such an interior tipping point exists in a given sample and era is an empirical question rather than a mechanical implication of the specification; Sections 4.2.2, 5.3, and 5.9 take it up directly.

3.3 Distribution Dynamics Analysis

3.3.1 Kernel Density Estimation (KDE)

Standard mean regression techniques often mask the underlying heterogeneity and polarization within the data by focusing solely on average convergence trends. To comprehensively capture the morphological characteristics and dynamic evolution of China's provincial income distribution, this study employs the non-parametric Kernel Density Estimation (KDE) method.

Estimation Formula and Kernel Selection

We utilize a Gaussian Kernel to estimate the continuous distribution of relative per capita $GDP(y_{it}/\bar{y}_t)$ for each year. By smoothing discrete histograms, KDE reveals critical distributional features such as skewness, kurtosis, and multi-modality. The estimation formula is defined as:

$$\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^{N} K(\frac{x - X_i}{h}) \tag{8}$$

where N denotes the number of provinces, $K(\cdot)$ represents the Gaussian kernel function, and h is the bandwidth.

Scientific Selection of Bandwidth

The selection of bandwidth h is critical as it determines the smoothness of the density curve. Following the classic rule of thumb proposed by Silverman (1986), we select an optimal bandwidth to balance the trade-off between bias and variance. This ensures that key distributional features are not obscured by over-smoothing, nor is the distribution distorted by spurious noise due to under-smoothing.

Logic of Dynamic Identification

By systematically comparing cross-sectional density estimates across key time nodes (2000, 2010, 2019, 2024), we scrutinize the distributional evolution along three dimensions:

- Location (Shift): We examine the lateral movement of the distribution's center to identify the overall improvement in development levels.
- Modality (Shape): We observe the morphological characteristics of the distribution, including skewness and the formation of distinct clusters, which are crucial for testing the "Club Convergence" hypothesis proposed by Quah (1996).
- Tail Characteristics: We analyze the extensiveness of the Right Tail to identify the degree to which high-income provinces are pulling away from the national average and their potential for spillover.

3.3.2 Markov Transition Matrix and Club Convergence Measurement

To deeply dissect the dynamic mobility paths of China's provincial economies across different development tiers, we employ the Markov chain approach to construct discrete state transition matrices.

State Space Discretization

Based on relative per capita GDP levels ($y_{it}/\bar{y}_t$), we discretize the sample into five income state spaces using the quintile method: ranging from the lowest income group (Q1) to the highest income group (Q5).

Construction of the Transition Matrix

Assuming that regional economic growth follows a first-order time-homogeneous Markov process, we calculate the transition probabilities p_{rs} (the probability of moving from state r at time t to state s at time $t+1$) using Maximum Likelihood Estimation (MLE). This allows us to construct a 5×5 transition matrix M.

Quantification of Club Lock-in

To quantify the overall degree of solidification across the system, we construct the Mobility Index (MI) following Shorrocks (1980):

$$MI = \frac{K - \mathrm{tr}(M)}{K - 1} \tag{9}$$

where $K = 5$ represents the number of income groups. The interpretation is straightforward: as $MI \rightarrow 0$, it indicates extremely high diagonal probabilities, suggesting that the regional economy is trapped in severe "Club Lock-in," with minimal vertical mobility between income classes.

3.4 Data Sources and Variable Descriptions

3.4.1 Sample Selection and the "Dual-Track" Time Window Strategy

This study utilizes a panel dataset covering 31 provincial administrative regions in China. Given the phased characteristics of China's economic evolution, we adopt a rigorous "Dual-Track" strategy regarding the time dimension:

1. Panoramic Evolution Window (1992–2024): For the analysis of Theil Index trends, Kernel Density dynamics, and Markov transition probabilities, the sample covers the entire period up to 2024. This design aims to capture the complete long-term evolutionary logic of China's regional economy and objectively observe the latest divergence trends in the "post-pandemic era."

Benchmark Regression Window (2011–2019, N = 279): For the parameter estimation of the Spatial Durbin interaction specification and the instrumental-variable analysis, we use the 2011–2019 window. The window concentrates the benchmark estimates on the most recent complete decade, within which the 2010 pre-window value of tech density provides a credible excluded instrument, and it ends in 2019 to isolate the model from the asymmetric economic shocks of the public health emergency (2020–2022). The longer 1993–2019 panel is retained for the sample-window sensitivity analysis in Section 5.3, which traces how the moderation estimate evolves as earlier reform-era years enter the sample.

3.4.2 Regional Division and Heterogeneity Strategy

In the descriptive statistics, we follow the standard classification by the National Bureau of Statistics (East, Central, and West). However, for the subsequent heterogeneity regression analysis, we merge the Central and Western provinces into a single "Inland Region" (Central & West). The logic is twofold:

- Statistical Validity: Spatial econometric models require sufficient cross-sectional sample size (N) and degrees of freedom. Merging the regions significantly enhances the statistical power of the test and reduces the probability of Type II errors.
- Mechanism Homogeneity: Calculations show that the tech density in the vast majority of both Central and Western provinces has long remained well below the national mean. According to New Structural Economics (Lin, 2011), these regions face similar constraints regarding factor endowments and are exposed to the same backwash-prone position in spatial interactions.

3.4.3 Variable Definitions and Measures

Dependent Variable: Economic Growth (g_{it})

Following standard literature, we measure economic growth as the log-difference of real GDP per capita. To eliminate inflationary factors, all nominal values are deflated using provincial GDP deflators with 1992 as the base year.

Core Moderator: Tech Density ($Tech_{it}$)

We define Tech Density as the ratio of "number of corporate units in scientific research and technical services" to the "year-end permanent population" (units per 10,000 people). Compared to R&D expenditure, this proxy better captures the density of market-based innovation actors and the ecosystem's actual capacity to absorb technology (Fu, 2008; Cohen & Levinthal, 1990). Before entering the regression model, this variable is Z-score standardized to facilitate the interpretation of interaction terms. The underlying firm count is consistently available from 2011 onward; in earlier years the variable carries its 2010 baseline value, a construction used by the sample-window sensitivity analysis in Section 5.3.

Data Sources

Raw data are sourced from the China Statistical Yearbook, China Regional Economic Statistical Yearbook, and various provincial statistical yearbooks. The estimating specification conditions on the variables of equation (7)—initial income, its spatial lag, the spatial lag of contemporaneous growth, and the tech density main effect—with province fixed effects absorbing time-invariant provincial characteristics and year fixed effects absorbing common shocks. Additional time-varying provincial controls are deliberately excluded: covariates such as openness or investment intensity are plausibly downstream of the spillover process itself, and conditioning on them would risk blocking the channels the moderation is meant to capture.

The descriptive statistics for the main variables are presented in Table 1.

Table 1. Descriptive Statistics of Main Variables (2011–2019)

Region	Variable	Obs.	Mean	Std. Dev.	Min	Max
Full Sample	Growth rate (g)	279	0.091	0.042	−0.021	0.219
	$\ln(PGDP)$	279	10.744	0.434	9.682	11.994
	$W \cdot \ln(PGDP_{t-1})$	279	10.639	0.333	9.745	11.528
	Tech Density	279	5.839	8.498	0.589	70.297
Eastern	Growth rate (g)	99	0.079	0.030	0.012	0.183
	$\ln(PGDP)$	99	11.105	0.414	10.240	11.994
	Tech Density	99	10.690	12.711	0.813	70.297
Central	Growth rate (g)	72	0.093	0.049	−0.021	0.219
	$\ln(PGDP)$	72	10.590	0.244	10.163	11.248
	Tech Density	72	3.536	2.179	0.691	10.480
Western	Growth rate (g)	108	0.102	0.044	−0.017	0.218
	$\ln(PGDP)$	108	10.517	0.324	9.682	11.216

Tech Density	108	2.928	1.609	0.589	7.157

Note. Tech Density is measured in units per 10,000 people. Average Tech Density: Eastern 10.690, Central 3.536, Western 2.928. Z-score standardization applied before regression.

Table 1 reveals a stark polarization in technological endowments. The average Tech Density in the Eastern region (10.69) is approximately 3 to 3.5 times higher than that in the Central and Western regions. The high standard deviation in the full sample (8.498) relative to the mean confirms the severe spatial mismatch of innovation resources, providing evidence for the club stratification phenomenon.

4. Empirical Results

4.1 Spatial Correlation Tests and Model Specification

Before proceeding with parameter estimation, we first verify whether China's provincial economic growth exhibits significant spatial dependence. This verification constitutes the prerequisite for employing spatial econometric models. Based on the Queen Contiguity Weight Matrix (W_q) constructed in the previous section, we calculated the Global Moran's I index for provincial per capita GDP over the 2011–2019 benchmark window.

Table 2 presents the results for selected years. As shown, the Global Moran's I indices are consistently positive and significant at the 1% level across the window, stabilizing in a high range between 0.35 and 0.45.

Table 2. Global Moran's I Index of Per Capita GDP (Selected Years)

Year	Moran's I	z-Value	p-Value
2011	0.437	4.70	<.001
2012	0.421	4.54	<.001
2013	0.399	4.33	<.001
2014	0.383	4.16	<.001
2015	0.381	4.14	<.001
2016	0.382	4.15	<.001

2017	0.370	4.03	<.001
2018	0.359	3.92	<.001
2019	0.357	3.91	<.001

Note. *** denotes statistical significance at the 1% level. All tests based on Queen contiguity weight matrix (row-standardized). z-values computed under randomization assumption.

Interpretation of Results:

The statistics reveal a typical feature of "Spatial Agglomeration" in China's regional economic development. Specifically:

- High-High (HH) Clustering: Provinces with high growth rates tend to be bordered by other high-growth provinces.
- Low-Low (LL) Clustering: Low-growth provinces often fall into contiguous "depression zones" of low growth.

Given this significant spatial dependence, employing traditional OLS estimation by ignoring spatial lag terms would lead to model misspecification and biased parameter estimates. Therefore, the Spatial Durbin Model (SDM), which is capable of simultaneously capturing spatial spillover effects from both the dependent variable (endogenous interaction) and independent variables (exogenous interaction), is the optimal choice for identifying the nonlinear convergence mechanism of the regional economy.

4.2 Benchmark Regression: An Identification Ladder (Core Findings)

Table 3 reports the estimation of the spatial moderation specification (7) as an identification ladder. Column (1) is pooled OLS with year effects; column (2) adds province and year fixed effects; column (3) is two-stage least squares in which the interaction term is instrumented by the 2010 baseline value of tech density interacted with the spatial lag of lagged log income. All columns use the 2011–2019 benchmark window (N = 279), with standard errors clustered by province.

Table 3. The Identification Ladder: Pooled OLS, Fixed Effects, and Instrumental Variables

Variable	(1) Pooled OLS	(2) SDM+FE	(3) IV 2SLS
W×lnPGDP(t−1)×Tech	0.0036 (0.0057)	0.0048 (0.0113)	0.0483*** (0.0106)
W×lnPGDP(t−1)	0.0103 (0.0126)	0.1340** (0.0642)	0.1782** (0.0703)
W×g	0.6433*** (0.1643)	0.4083** (0.1904)	0.3199* (0.1811)
lnPGDP(t−1)	−0.0246** (0.0105)	−0.1796*** (0.0326)	−0.1774*** (0.0350)
Tech Density (Z-score)	−0.0359 (0.0632)	−0.0488 (0.1277)	−0.5347*** (0.1189)
Observations	279	279	279
Province fixed effects	No	Yes	Yes
Kleibergen–Paap rk Wald F	—	—	30.52
Anderson–Rubin p-value	—	—	0.0007
R^2	0.678	0.797	—

Note. Dependent variable is the annual growth rate; year effects included in all columns; columns (2)–(3) include province fixed effects; cluster-robust standard errors (31 province clusters) in parentheses. Column (3) instruments W×lnPGDP(t−1)×Tech with its 2010 baseline counterpart: Kleibergen–Paap rk Wald F = 30.52 (Stock–Yogo 10% critical value 16.38), Cragg–Donald F = 96.33, Anderson–Rubin weak-instrument-robust p = 0.0007. Spatial weight matrix: Queen contiguity (row-standardized). *** p<0.01, ** p<0.05, * p<0.10.

4.2.1 Convergence and Asymmetric Spillovers: An Absorptive Capacity Perspective

Based on the regression coefficients in Table 3, we derive the following key inferences regarding the mechanism of regional growth:

1. Confirmation of Conditional Convergence: The coefficient of initial income is significantly negative in every column (−0.025 in pooled OLS; −0.180 and −0.177 with fixed effects and under IV). After controlling for spatial factors and technological endowments, China's provincial economies exhibit a significant intrinsic tendency for conditional β-convergence.

2. The Attenuation–Recovery Pattern of the Moderation: The interaction coefficient θ_2 is small and statistically insignificant in the pooled and fixed-effects columns (0.0036 and 0.0048), but rises by an order of magnitude to 0.0483 ($p < 0.001$) once the interaction is instrumented. This pattern is what a contaminated proxy produces—attenuation under least squares, recovery once the predetermined component is isolated—and Section 5.1 grounds the reading in the noise structure of the measure, an exact decomposition of the identifying variation, and a weak-instrument-robust confidence set that excludes the least-squares magnitudes. The IV estimate is therefore our preferred estimate of the moderation.

3. A Positive Average Neighbour Effect in the 2010s: The baseline spillover coefficient θ_1 is positive and significant under IV (0.178, $p = 0.011$): at mean tech density, neighbouring income raises local growth in the 2011–2019 window, and absorptive capacity amplifies this effect. The backwash side of the mechanism does not vanish from the data—it appears in longer sample windows, as Section 5.3 documents.

4.2.2 The Marginal Effect and the Question of a Threshold

The marginal spatial spillover effect implied by the preferred IV estimates is $\partial g/\partial(W\cdot \ln y) = 0.178 + 0.048 \times \text{Tech}$. Over the observed 2011–2019 support of standardized tech density (−0.51 to 11.76) this marginal effect is positive throughout; its implied zero-crossing lies at −3.69, far outside the support. Within the benchmark window, therefore, the data do not locate an interior break-even point: the 0.23 value derived in the original submission rested on coefficients that do not survive the identification ladder of Table 3. Whether and when a break-even point exists is an era-dependent question, taken up by the sample-window analysis in Section 5.3; whether the moderation takes the form of a sharp threshold at all is tested—and rejected—by the grid search in Section 5.9.

Figure 1 plots the marginal spillover effect implied by the preferred IV estimates, together with its 95% confidence interval, as a function of tech density.

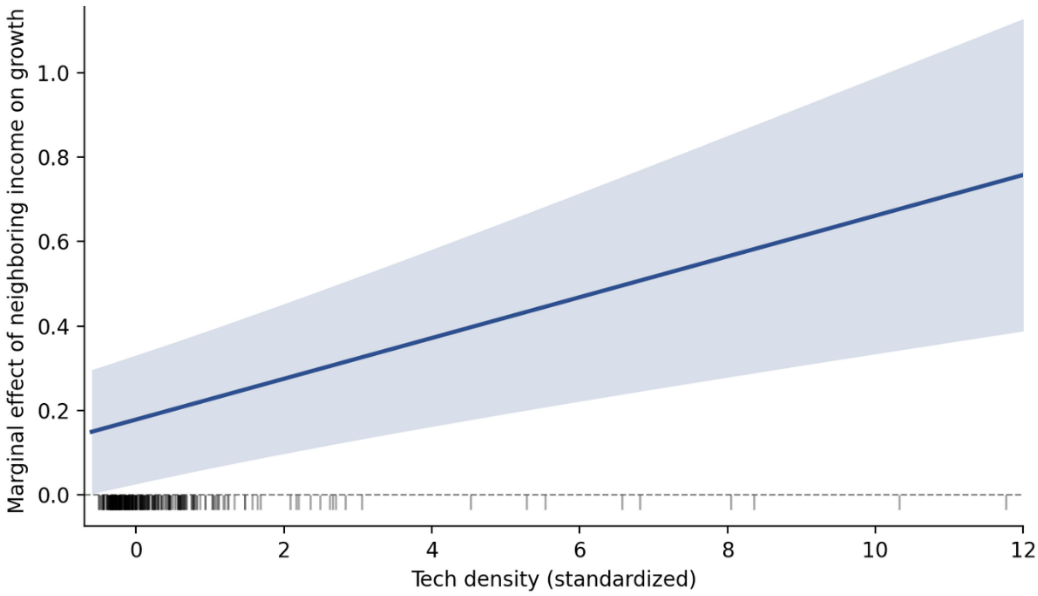


Figure 1. Conditional Spatial Spillover Effect

*Note.*The solid line represents the marginal effect of neighbouring income on local growth, computed from the IV estimates in Table 3, column (3), as $\theta_1 + \theta_2 \times$ Tech. The shaded area indicates the 95% confidence interval using the delta method with the cluster-robust covariance of (θ_1, θ_2). Rug plot at bottom shows the distribution of tech density observations, 2011–2019.

Figure 1 shows a marginal effect that is positive across the observed range and increasing in tech density: in the 2011–2019 window, provinces with deeper absorptive capacity extract systematically more growth from neighbouring prosperity, and no province sits in a negative-spillover region. The backwash configuration—negative spillovers concentrated among low-capacity provinces—belongs to the earlier reform decades: the long-window and subperiod estimates in Section 5.3 and Appendix Table A2 show a significantly negative interaction for samples reaching back to the 1990s and for the 2001–2010 subperiod. The mechanism is therefore best read as era-dependent moderation along a continuous curve, not as a fixed cross-sectional threshold separating a “siphon zone” from a “spread zone.”

4.3 Discussion: The Micro-foundations of Club Convergence

Combining the regression results with distribution dynamics, this section further elucidates the formation mechanism of "Club Convergence."

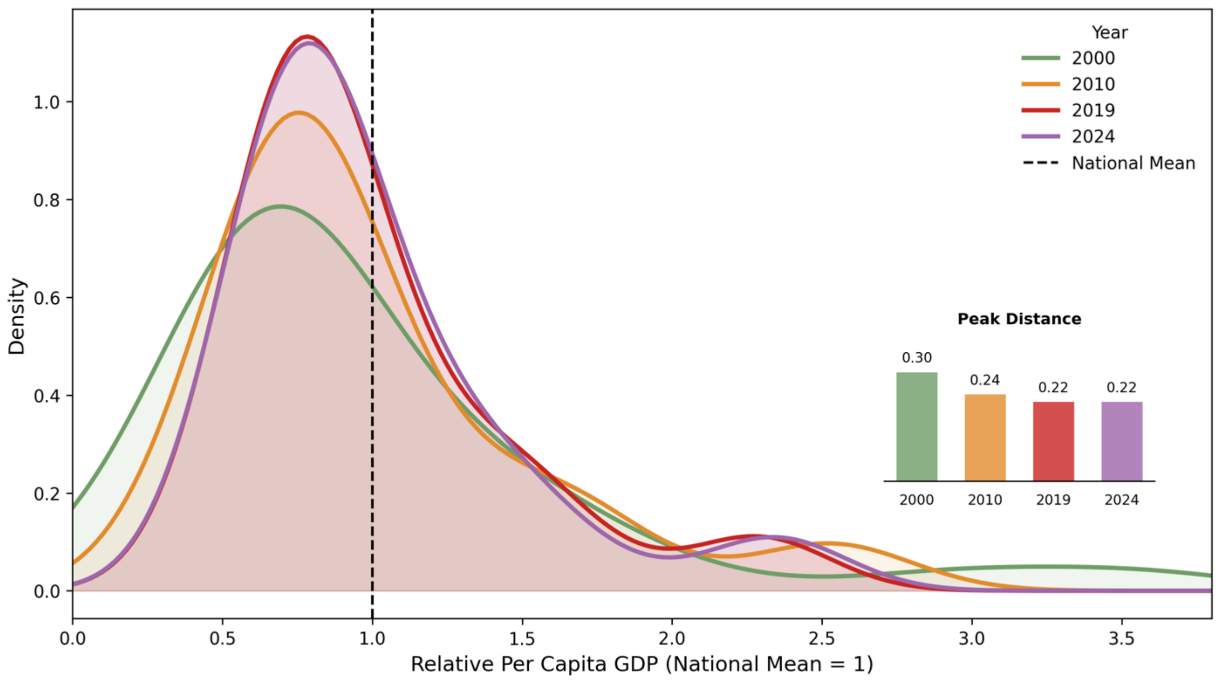


Figure 2. Evolution of Kernel Density Estimation (2000–2024)

Note. Kernel density estimates of relative per capita GDP (normalized by national mean) at four time points. Bandwidth selected using the Silverman (1986) rule of thumb.

To quantify the fluidity between these clubs, Table 6 presents the Markov transition probability matrix calculated over the full sample period.

Table 6. Markov Transition Probability Matrix (Five-Year Step, 1992–2024)

$t / t + 5$	Q1 (Low)	Q2	Q3 (Mid)	Q4	Q5 (High)
Q1 (Low)	**93.3%**	6.7%	0.0%	0.0%	0.0%
Q2	7.8%	**80.7%**	11.5%	0.0%	0.0%
Q3 (Mid)	0.0%	11.5%	**82.8%**	5.7%	0.0%
Q4	0.0%	0.0%	5.7%	**93.2%**	1.0%
Q5 (High)	0.0%	0.0%	0.0%	1.0%	**99.0%**

Note. Cell entries show probability (in %) of transitioning from row state (t) to column state (t+5). States defined by quintiles of relative per capita GDP. Diagonal elements (boldface) indicate persistence probability. Mobility Index = 0.047. Based on 31 provinces × 7 five-year transitions = 217 observations.

Analysis of the Club Convergence Mechanism:

Establishment of the "Polarized" Distribution Morphology:

As visualized in Figure 2, the distribution of provincial per capita GDP exhibits a characteristic right-skewed unimodal pattern with an extended upper tail. The primary mode is located below the national mean (at approximately 0.7–0.8 of the national average), indicating that the majority of provinces cluster in the lower-income bracket. Meanwhile, a distinct upper tail extends beyond 2.0 times the national mean, representing the handful of high-income provinces (primarily coastal) that have structurally pulled away from the pack. Notably, the peak distance metric shows a decline from 0.30 (2000) to 0.22 (2024), indicating convergence in the main body of the distribution. However, the persistence of the long upper tail reveals that elite provinces maintain their structural lead, creating a "catch-up ceiling" for the majority.

The Lock-in Effect of Stratification:

Table 6 provides the statistical evidence for this hardening. The retention probability for the Low-Income Group (Q1) is as high as 93.3%, meaning a backward province has less than a 7% chance of moving up the ladder in the short term. For the High-Income Group (Q5), the retention probability reaches 99.0%.

The Vicious Cycle: Lacking the absorptive capacity to convert neighbouring spillovers into local diffusion, backward regions fall into a self-reinforcing loop: *backwardness invites siphoning by neighbours, capital and talent flow out, and capacity becomes still harder to build.*

Synthesizing the Logic Chain:

The "Convergence Puzzle" in China's regional economy is essentially the aggregate outcome of threshold-dependent spillovers. Over the long reform era, the negative long-window moderation estimates of Section 5.3 indicate that the backwash mechanism dominated nationwide for most of the sample period.

The Virtuous Circle: Developed regions, using high tech density, enjoy positive radiation effects, continuously reinforcing their leading advantages (The Spread Effect).

- **The Matthew Effect:** Backward regions, with shallow absorptive capacity, suffered chronic siphoning over the earlier reform decades—a position shared by the majority of Chinese provinces.

- **Policy Implication:** This finding points to a policy implication: mere fiscal transfer payments, if not translated into local technological absorptive capacity, will fail to break the solidified club stratification pattern.

5. Robustness and Identification-Robust Diagnostics

Section 4 established the attenuation–recovery pattern on the 2011–2019 benchmark window and identified the IV estimate as the preferred estimate of the moderation. This section submits that estimate to a battery of complementary diagnostics: alternative measures of absorptive capacity, sample-window sensitivity, wild cluster bootstrap inference, leverage analysis, two-way clustering, and a formal grid search for a structural threshold. Throughout, the instrument is the 2010 baseline value of tech density interacted with the spatial lag of lagged log per capita output; the exclusion restriction is that this pre-window baseline affects 2011–2019 growth only through the contemporaneous moderator–spillover interaction, conditional on province and year fixed effects.

5.1 IV Strategy and First-Stage Strength

The first-stage Kleibergen–Paap rank Wald F statistic is 30.5, comfortably above the Stock–Yogo 10% maximal-bias critical value of 16.38, and the Cragg–Donald F is 96.3. The IV 2SLS estimate of the moderation coefficient is $\theta_2 = 0.0483$ (cluster-robust standard error 0.0106, $p < 0.001$), with 95% confidence interval [0.028, 0.069]. Inverting the Anderson–Rubin test yields a weak-instrument-robust 95% confidence set of [0.025, 0.077]; the set excludes not only zero but also the pooled and fixed-effects magnitudes of Table 3, columns (1)–(2). The estimate is an order of magnitude larger than those least-squares values, and a gap of that size is a diagnostic to be explained, not a result to be celebrated.

The explanation we commit to is that the contemporaneous count is a contaminated measurement of predetermined structural capacity, and the contamination has two parts, both visible in the data. The first is transitory recording noise. What the proxy registers each year is a firm count, and the within-province movement of firm counts mean-reverts: successive annual changes in measured tech density are negatively correlated (first-difference autocorrelation −0.13 over 2011–2019), whereas every alternative measure in Table 4 shows strongly positive change persistence (+0.32 for R&D intensity, +0.39 for the composite, +0.42 for patents, +0.61 for higher-education density). Tech density is the only measure whose within variation carries the footprint of entry, exit, and reclassification churn layered on a quasi-fixed stock (Griliches & Hausman, 1986); three quarters of its 2011–2019 variance lies between provinces, where the capacity ranking lives. The second part is contemporaneous response: registration activity reacts to local conditions within the year—endogeneity of exactly the kind that motivates the instrument, with a direction, it turns out, toward zero.

The decomposition is exact arithmetic. Splitting the measured interaction into its baseline-driven component and the orthogonal remainder, and entering both in the growth equation, yields 0.048 on the capacity-driven part—the IV estimate, by construction—and -0.013 (standard error 0.008) on the remainder: transitory registration movements carry a small, statistically zero, slightly negative association with growth. The first-stage partial R^2 is 0.29, so predetermined capacity drives under a third of the identifying variation while least squares prices all of it; the variance-weighted blend, $0.29 \times 0.048 + 0.71 \times (-0.013)$, is 0.0048—the fixed-effects estimate of Table 3, column (2), to the digit. The 2010 baseline is fixed before the window, strongly correlated with the persistent capacity stock, and by construction free of both contaminants, so instrumenting with the baseline-interacted term removes the attenuation rather than inflating the parameter. Two rival readings of a large IV-to-least-squares ratio remain—a local average effect carried by a handful of provinces, or an instrument loading on something other than capacity. Section 5.6 confronts the first; the second we address directly here and again in Section 5.7.

The most natural version of the second reading is that 2010 tech density merely relabels 2010 development—rich provinces, not capable provinces, extracting more from neighbouring growth. A horse race tests this directly: we add the rival interaction, the 2010 income baseline interacted with the same spatial lag, as a control. The moderation estimate does not shrink; it edges up to 0.059 ($p = 0.02$). Substituting the 2010 R&D baseline as the rival gives 0.055 ($p < 0.001$). An estimate that survives its rivals at full magnitude is not a relabelled development effect. What the thirty-one-province cross-section cannot do is razor-separate channels that co-move this tightly—2010 capacity and 2010 income correlate at 0.70—so the income control cuts the first stage to KP F = 10.0, and entering both rivals at once dilutes precision (0.039, $p = 0.08$) without moving the order of magnitude. We report the dilution rather than hide it: capacity and development are entangled in China's economic geography, and a sharper separation requires the richer cross-section of prefecture-level data.

5.2 Alternative Measures of Absorptive Capacity

Firm counts in scientific services may capture industrial structure rather than technological capability. To test whether the moderation survives alternative measurements of absorptive capacity, we re-estimate the IV specification of Section 5.1 with four alternatives: R&D expenditure intensity, patents per capita, higher-education population density, and the first principal component of the measure set, each instrumented by its own pre-window baseline counterpart. Table 4 reports the results. Every point estimate is positive, but none is statistically significant; the first stage is weak for patents (F = 3.1) and higher education (F = 6.4), while R&D intensity and the composite have strong first stages (F = 61.6 and 53.1) with small positive coefficients. We read Table 4 as full disclosure that the identification-robust moderation result is specific to tech density: the density of market-based innovation actors is where both the instrument and the moderation have power. The alternative measures neither contradict the mechanism—no negative estimate appears—nor independently confirm it. The cross-measure pattern also corroborates the contaminated-proxy reading of Section 5.1: for R&D intensity, whose within variation is persistent rather than mean-reverting, the fixed-effects and IV estimates are of the same small order (0.0025 and 0.0064, both insignificant)—there is no attenuation gap to undo—while the order-of-magnitude gap appears precisely for the one measure with a transitory recording component.

Table 4. Alternative Measures of Absorptive Capacity (IV 2SLS, 2011–2019)

Measure	θ_2	Cluster-robust SE	KP first-stage F
Tech density (firms per 10,000 population)	0.0483***	0.0106	30.52
R&D expenditure / GDP	0.0064	0.0101	61.56
Patents per capita	0.0121	0.0396	3.05
Higher-education population density	0.0186	0.0268	6.32
First principal component of the measure set	0.0106	0.0112	53.13

Note. Each row instruments W×lnPGDP(t−1)×H with its pre-window baseline counterpart under province and year fixed effects; N = 279; standard errors clustered by province (31 clusters). First-stage F values below the Stock–Yogo 10% critical value (16.38) indicate weak identification for patents and higher education. *** $p < 0.01$.

5.3 Sample-Window Sensitivity

The benchmark window is the most recent decade; the mechanism's history is longer. Table 5 re-estimates the IV specification on progressively longer windows ending in 2019. The moderation coefficient is +0.048 ($p < 0.001$) on 2011–2019, statistically indistinguishable from zero on 2006–2019, and significantly negative on windows reaching back to 2001 and to the early 1990s (−0.013 and −0.011). Fixed-effects estimates by subperiod (Appendix Table A2) show the same reversal, with the most negative interaction in 2001–2010.

Because the underlying firm count is consistently available only from 2011, tech density enters the pre-2011 years at its 2010 baseline value; the longer windows therefore measure how a fixed baseline-capacity ranking moderated spillovers across eras, free of contemporaneous reverse causality though unable to reflect within-era capacity changes. Read this way, Table 5 indicates that the positive moderation identified in Section 5.1 is a property of the digital-era 2010s rather than a time-invariant structural constant, and that the backwash configuration—negative moderation along the same capacity ranking—dominated the earlier reform decades. Two mechanisms are consistent with the reversal, and we do not adjudicate between them here: digital infrastructure may have lowered the dependence of knowledge diffusion on local physical capacity, or the composition of spatial interactions may have shifted from competition for mobile factors toward technology diffusion.

Table 5. Sample-Window Sensitivity of the Moderation Coefficient (IV 2SLS)

Window	θ_2	Cluster-robust SE	KP first-stage F	N
2011–2019	0.0483***	0.0106	30.52	279
2006–2019	0.0073	0.0124	37.16	434
2001–2019	−0.0133*	0.0081	237.93	589
1993–2019	−0.0107***	0.0032	1251.31	837

Note. Each row estimates the Section 5.1 IV specification on the indicated window with province and year fixed effects; standard errors clustered by province. Tech density enters years before 2011 at its 2010 baseline value (Section 3.4.3). * $p < 0.10$, *** $p < 0.01$.

5.4 Wild Cluster Bootstrap Inference

With only 31 clusters, asymptotic cluster-robust standard errors may understate sampling variation. Following Roodman, MacKinnon, Nielsen, and Webb (2019), we apply the wild cluster bootstrap with 9,999 replications and four weight distributions. The p-values are 0.019 (Rademacher), 0.018 (Webb), 0.022 (Normal), and 0.057 (Mammen). Three of the four weights retain significance at the conventional 5% level; the Mammen specification places the moderation at the 10% level. The wild bootstrap produces inference more conservative than the asymptotic cluster-robust standard errors but does not overturn the central finding.

5.5 Province-Specific Linear Trends

To probe whether the instrument absorbs province-specific long-run trends rather than identifying a structural relationship, we augment the baseline IV specification with province-specific linear time trends. With nine years per province, the trends absorb nearly all of the variation on which the baseline-interacted instrument relies: the first-stage Kleibergen–Paap F collapses to 0.86, the point estimate becomes uninformative (−0.092), and the wild bootstrap p-value is 0.587. We report this transparently as a boundary of the design rather than as evidence against the moderation: the short benchmark window cannot separately identify the moderation and province trends. The long-panel window analysis of Section 5.3, which draws identifying variation from across eras rather than within the nine-year window, provides the complementary low-frequency evidence.

5.6 Rotemberg-Style Leverage

The Goldsmith-Pinkham, Sorkin, and Swift (2020) framework identifies which observations contribute most to identification. Adapted to the leave-one-out (LOO) setting, we re-estimate the IV coefficient on each 30-province subsample. The five largest single-province contributors are: Beijing (θ_2 falls to 0.0303, a 37% decline), Shanghai (0.0385, 20% decline), Tianjin (0.0414, 14% decline), Jiangsu (0.0447, 7% decline), and Guangdong (0.0478, 1% decline). Across all thirty-one leave-one-out estimates the coefficient remains positive, ranging from 0.030 to 0.055. No single province exceeds the 40% concentration threshold that Goldsmith-Pinkham et al. treat as indicative of identification fragility.

5.7 Cumulative Leverage: Excluding Directly-Controlled Municipalities

The LOO diagnostic in Section 5.6 considers single-province exclusions. A stronger test considers cumulative leverage: the simultaneous exclusion of the four directly-controlled municipalities (Beijing, Tianjin, Shanghai, Chongqing). On the 27-province subsample (N =

243), the IV moderation coefficient is −0.021 (cluster-robust standard error 0.042) and the Kleibergen-Paap F statistic falls to 6.2, below the Stock-Yogo 10% maximal-bias threshold of 16.38. The Anderson-Rubin weak-instrument-robust p-value is 0.65.

This result is consistent with the theoretical framework rather than a refutation of it. The directly-controlled municipalities have the most concentrated scientific and technical services infrastructure in China. The IV identifies the moderation effect from variation in baseline absorptive capacity; on the 27-province subsample, the cross-sectional variation in 2010 baseline tech density is substantially narrower, and the first-stage relationship attenuates accordingly. This complements the sample-window analysis in Section 5.3: the two exercises draw on different variation and answer different questions, and the moderation should be understood as identified primarily from the high-capacity tail of the 2010s cross-section. The direction of the failure is itself diagnostic: with the first stage gone, the estimate collapses toward an uninformative zero whose wide confidence interval comfortably contains the full-sample 0.048, rather than flipping sign—the behaviour expected of a noise-purging instrument, not of one loading on an omitted confounder.

5.8 Two-Way Clustering

Cluster-robust standard errors under one-way province clustering assume that residuals are independent across years. To address potential within-year cross-province dependence, we also report two-way clustering by province and year, following Cameron, Gelbach, and Miller (2011). The point estimate is unchanged at 0.0483, but the standard error widens from 0.0110 to 0.0255, producing a p-value of 0.094. The Kleibergen-Paap F statistic under two-way clustering is 14.74, marginally below the Stock-Yogo threshold. We report this as a candid measure of how the headline significance changes under a more stringent assumption. A caveat applies to the estimator class itself: quasi-maximum-likelihood estimation of spatial Durbin panels with fixed effects carries incidental-parameter bias (Lee and Yu, 2010). The least-squares and instrumental-variable estimates reported here enter the spatial lag as a conditioning variable rather than estimating the full model by quasi-maximum likelihood; a bias-corrected re-estimation in the spirit of Lee and Yu (2010) is left for future work.

5.9 Hansen Grid Search

The 0.23 value reported in the original submission was derived mechanically from the coefficients of the linear interaction specification. To test whether a sharp threshold exists statistically, we apply a Hansen (2000)-type grid search across 21 candidate threshold values uniformly spaced between the 10th and 90th percentiles of the standardized tech density distribution ([−0.347, 1.199]). The residual sum of squares varies by only 3.3% across the entire grid—a nearly flat objective with no pronounced minimum. At the RSS-minimizing split ($\hat{\gamma} = -0.269$, below which 21.9% of observations fall), the low-regime and high-regime moderation coefficients are −0.038 (cluster-robust standard error 0.052) and +0.0005 (standard error 0.011): each is statistically indistinguishable from zero, and they are statistically indistinguishable from each other. The full 21-point grid is reported in the replication package.

The 0.23 value should therefore be retired as a structural quantity: the data favour a continuously-conditional moderation curve rather than a sharp regime boundary. The policy discussion in Section 6 should be understood accordingly: shifting a province's position along the moderation curve through targeted investment in scientific and technical services infrastructure is plausible, but the data do not support claims that crossing a specific numerical value triggers qualitative changes in growth dynamics.

Table 7. Summary of Identification-Robust Diagnostics (2011–2019 IV Specification)

Diagnostic	Coefficient (θ_2)	SE / p-value	Significance
Baseline IV (one-way cluster)	0.0483	SE 0.0106	***
Wild bootstrap (Rademacher)	0.0483	p = 0.019	**
Wild bootstrap (Webb)	0.0483	p = 0.018	**
Wild bootstrap (Mammen)	0.0483	p = 0.057	*
Wild bootstrap (Normal)	0.0483	p = 0.022	**
Province-specific linear trends	−0.0922	KP F = 0.86	Not identified
Wild bootstrap on trended spec	—	p = 0.587	—
Two-way clustering (province, year)	0.0483	SE 0.0255; p = 0.094	*
LOO max single drop (Beijing)	0.0303	−37.3% change	—
Excluding 4 directly-controlled muni.	−0.0209	KP F = 6.2	Below threshold
Hansen grid ($\hat{\gamma} = -0.269$)	$\Delta\beta = 0.038$	RSS range 3.3%	No regime switch

Notes. *** p < 0.01, ** p < 0.05, * p < 0.10. All specifications use the 2011–2019 panel (N = 279, except the 27-province row with N = 243) with province and year fixed effects unless otherwise noted. The baseline IV instruments W × ln(y_{t-1}) × tech_density with its 2010 baseline value. Wild cluster bootstrap implementations follow Roodman, MacKinnon, Nielsen, and Webb (2019) with 9,999 replications.

Table 7 consolidates the diagnostics. The moderation coefficient is statistically significant under cluster-robust inference ($p < 0.001$) and under the wild cluster bootstrap with three of four weight schemes ($p = 0.018$–0.022); it is marginally significant under Mammen weights ($p = 0.057$) and under two-way clustering ($p = 0.094$). It is not identified under province-specific linear trends, where the instrument's first stage collapses, and its identification on the 2011–2019 window depends on the directly-controlled municipalities. The Hansen grid search rejects a sharp-threshold representation, and a rival-channel horse race (Section 5.1) leaves the point estimate intact at the cost of first-stage power. Taken together, the diagnostics support a positive but conditionally-identified moderation effect, concentrated in the digital-era cross-section.

6. Conclusion

6.1 Main Conclusions

After more than three decades of rapid growth, China's regional economy still faces a convergence puzzle characterized by the coexistence of "aggregate growth and widening relative disparities." Departing from the traditional assumption of "homogeneous spatial spillovers," this paper constructs a spatial growth model incorporating capacity-moderated technology diffusion and conducts identification-robust empirical tests using panel data from 31 Chinese provinces. The main conclusions are as follows:

First, the characteristics of "Club Convergence" are significant, indicating severe path dependence. Both Kernel Density Estimation and Markov transition matrix analysis confirm that the distribution of provincial economies exhibits a pronounced right-skewed pattern with club stratification: the majority of provinces cluster below the national mean while a small elite group maintains a structural lead in the upper tail. The diagonal solidification probabilities for the low-income and high-income clubs are as high as 93.3% and 99.0%, respectively. Without strong exogenous intervention, backward regions are unlikely to break path dependence and achieve upward mobility solely through spontaneous market forces.

Second, spatial spillovers exhibit a significant nonlinear moderation mechanism, and the moderation is era-dependent. The empirical results support the "Absorptive Capacity Moderation Hypothesis" in identification-robust form: on the 2011–2019 benchmark window, the instrumented moderation coefficient is $\theta_2 = 0.048$ ($p < 0.001$), an order of magnitude above the attenuated least-squares estimates; it survives wild cluster bootstrap inference ($p = 0.019$

under Rademacher weights), and the average neighbour effect at mean capacity is positive. On windows reaching back to the 1990s the interaction turns significantly negative, indicating that the backwash configuration dominated the earlier reform era. The Hansen grid search finds no sharp threshold—the regime-coefficient gap is smaller than its sampling uncertainty and the grid objective is nearly flat—so the 0.23 value of the original submission is retired in favour of a continuously-conditional moderation curve, and policy claims about discrete regime switches are tempered accordingly.

Third, the inland provinces' historical position was backwash-prone. Over the earlier reform decades, the significantly negative long-window moderation estimates (Section 5.3) indicate that provinces with shallow absorptive capacity—disproportionately in the Central and Western belts—were exposed to net resource outflows in their interactions with developed neighbours. This explains why infrastructure connectivity alone (such as the high-speed rail network) did not narrow the gap and, in some periods, accelerated the outflow of factors from the inland to the coast. The 2010s pattern is different: with positive average spillovers and capacity-amplified diffusion, the binding constraint has shifted from escaping siphoning to building the absorptive capacity that converts neighbouring prosperity into local growth.

6.2 Policy Implications: From "Gap Filling" to "Capacity Building"

Based on the above conclusions, this paper argues that the key to resolving unbalanced regional development lies not in simple fiscal transfers (filling the income gap) but in reshaping the absorptive capacity of backward regions (building the capacity that converts spillovers into growth).

6.2.1 Concentrating Investment in Absorptive Capacity

For inland provinces with shallow absorptive capacity, egalitarian but fragmented resource allocation (the "sprinkling pepper" approach) often results in sunk costs without activating a positive feedback loop. We recommend a strategy of "Focused Breakthroughs":

- **Capacity-Building Special Action:** Concentrate fiscal and talent policy on raising the density of scientific and technical services through sustained, multi-year investment, moving the province along the moderation curve toward stronger positive spillovers. As noted in Section 5.9, the mechanism is continuous: the payoff accrues gradually rather than at a single crossing point.

- **Prioritize "New Quality Productive Forces":** Following Lin (2011)'s new structural economics framework, backward regions should not attempt to replicate the entire industrial structure of developed regions. Instead, they should identify and cultivate industries where they have latent comparative advantages, particularly in frontier sectors such as low-altitude economy, commercial aerospace, biomanufacturing, and artificial intelligence applications.

6.2.2 Anchoring Digital Infrastructure

Addressing the disadvantage of physical distance from markets, the strategic focus of inland regions should shift from accepting obsolete capacity from the East to cultivating technological absorptive capacity through digital means:

- **Avoiding Traditional Capacity Traps:** Backward regions must not become "recycling bins" for low-end manufacturing; otherwise, they will be permanently locked at the bottom of the value chain.

- **Building a Digital Absorptive Base:** Regions should draw on national strategic opportunities such as the "Eastern Data, Western Computing" project to build high-level data centers and computing hubs. By using digital infrastructure to break physical spatial barriers and exploiting the strong permeability of digital technologies, inland regions can rapidly elevate their technological absorptive capacity to compensate for geographic disadvantages.

6.2.3 Establishing a Horizontal Compensation Mechanism for "Benefit Sharing and Cost Sharing"

Given that developed regions objectively utilize agglomeration advantages to "siphon talent" and "extract resources" from backward regions, a market-based horizontal compensation mechanism should be established:

- Talent Cultivation Compensation: For high-skilled talent flowing from the Central/West to the East, a "Talent Cultivation Compensation Fund" system is recommended, involving transfer payments from the fiscal revenue of inflow regions to outflow regions.

- Two-way Industrial Embedding: Encourage Eastern provinces to build "Reverse Enclaves" or "Co-constructed Industrial Parks" in the Central/West, ensuring that a higher proportion of output value and tax revenue remains in the inland regions, transforming one-way siphoning into two-way embedding. This approach aligns with insights from the industrial cluster literature.

6.3 Limitations and Future Prospects

Subject to data availability and research perspective, this paper has the following limitations that offer avenues for future research:

First, the refinement of spatial scale. This study is based on provincial-level data, which may mask the micro-level "Core-Periphery" structures within provinces (e.g., between provincial capitals and peripheral cities). Future research could utilize prefecture-level or county-level data for more granular threshold measurement.

Second, the decomposition of spillover channels. This paper examines aggregate spatial spillovers without specifically distinguishing the heterogeneous mechanisms of capital spillovers, knowledge spillovers, and institutional spillovers. Future studies could deconstruct these channels using multi-dimensional spatial weight matrices.

Third, cross-country comparative analysis. While we draw parallels with Thailand's experience, systematic cross-country analysis could test whether the identified threshold mechanism generalizes to other large emerging economies facing similar regional development challenges, such as India, Indonesia, or Brazil.

Appendix

This appendix contains the consolidated robustness estimates for the spatial moderation coefficient, the subperiod heterogeneity analysis, and the Theil index inequality decomposition. These results support but are not central to the main empirical findings.

Appendix Table A1. Consolidated Robustness: The Moderation Coefficient under Alternative Specifications (2011–2019)

Specification	θ_2	Cluster-robust SE	N
W1 Queen contiguity, two-way fixed effects	0.0048	0.0113	279
W3 inverse geographic distance	0.0136	0.0130	279
W4 five nearest neighbours	0.0100	0.0111	279

SLX specification (no spatial lag of growth)	0.0087	0.0120	279
Excluding Hainan (W1)	0.0052	0.0119	270
Pooled OLS with year effects	0.0036	0.0057	279

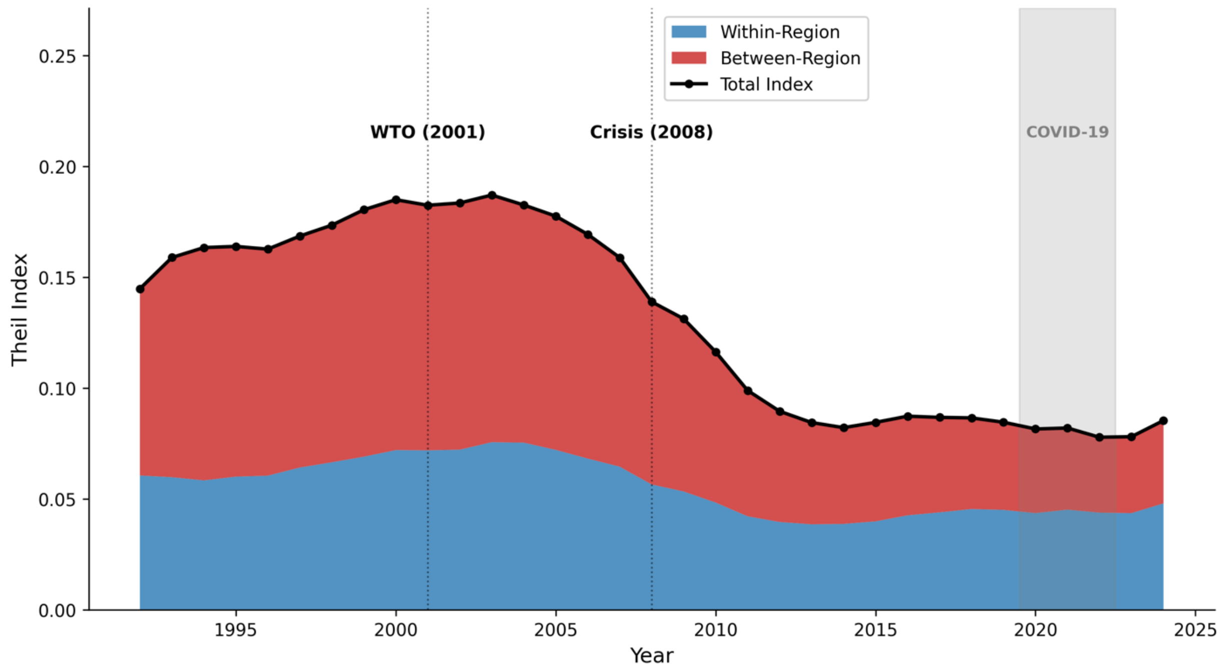
Note. Least-squares estimates of the interaction coefficient W×lnPGDP(t−1)×Tech under the indicated specification, 2011–2019. None is statistically significant, consistent with the attenuation documented in Table 3, columns (1)–(2); the identification-robust counterpart is the IV estimate in Table 3, column (3). Standard errors clustered by province.

Appendix Table A2. Subperiod Heterogeneity of the Moderation Coefficient

Subperiod	**θ_2**	**Cluster-robust SE**	**N**
1993–2000	0.0253**	0.0111	248
2001–2010	−0.0437***	0.0083	310
2011–2019	0.0048	0.0120	279

Note. Two-way fixed-effects estimates with standard errors clustered by province. Tech density enters years before 2011 at its 2010 baseline value. The sign reversal across subperiods corresponds to the window sensitivity in Table 5. ** $p < 0.05$, *** $p < 0.01$.

Appendix Figure A1. Regional Inequality Decomposition & Trends (1992–2024)



***Note.* The figure displays the evolution of total regional inequality (Theil Index) and its decomposition into between-region (Eastern, Central, Western) and within-region components, computed from real per capita GDP, 1992–2024.**

References

Acemoglu, D., Aghion, P., & Zilibotti, F. (2006). Distance to frontier, selection, and economic growth. *Journal of the European Economic Association*, *4*(1), 37-74.

Anderson, T. W., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics*, *20*(1), 46-63.

Bai, C.-E., Ma, H., & Pan, W. (2012). Spatial spillover and regional economic growth in China. *China Economic Review*, *23*(4), 982-990.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, *29*(2), 238-249.

Chen, A., & Partridge, M. D. (2013). When are cities engines of growth in China? Spread and backwash effects across the urban hierarchy. *Regional Studies*, *47*(8), 1313-1331.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, *35*(1), 128-152.

Donaldson, D., & Hornbeck, R. (2016). Railroads and American economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2), 799-858.

Elhorst, J. P. (2014). *Spatial econometrics: From cross-sectional data to spatial panels*. Springer.

Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway System. *The Review of Economic Studies*, 81(3), 1046-1070.

Fan, J. (2019). Internal geography, labor mobility, and the distributional impacts of trade. *American Economic Journal: Macroeconomics*, 11(3), 252-288.

Fu, X. (2008). Foreign direct investment, absorptive capacity and regional innovation capabilities: Evidence from China. *Oxford Development Studies*, *36*(1), 89-110.

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, *110*(8), 2586-2624.

Griliches, Z., & Hausman, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics*, 31(1), 93-118.

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica*, *68*(3), 575-603.

Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, *124*(4), 1403-1448.

Hsieh, C.-T., & Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2), 1-39.

Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, *133*(1), 97-126.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, *99*(3), 483-499.

Lee, L.-F., & Yu, J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, *154*(2), 165–185.

LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.

Lin, J. Y. (2011). New structural economics: A framework for rethinking development. *The World Bank Research Observer*, *26*(2), 193-221.

Lu, M., & Wan, G. (2015). Urbanization and urban systems in the People's Republic of China: Research findings and policy recommendations. In *China's Economy: A Collection of Surveys* (pp. 83-98).

Ma, H., Wei, Y. D., Huang, X., & Zhang, W. (2023). The innovation networks shaped by large innovative enterprises in urban China. *Journal of Geographical Sciences*, *33*(3), 599-617.

Moretti, E. (2021). The effect of high-tech clusters on the productivity of top inventors. *American Economic Review*, 111(10), 3328-3375.

Myrdal, G. (1957). *Economic theory and under-developed regions*. Duckworth.

Paweenawat, S. W., & McNown, R. (2014). The determinants of income inequality in Thailand: A synthetic cohort analysis. *Journal of Asian Economics*, *31*, 10-21.

Quah, D. T. (1996). Twin peaks: Growth and convergence in models of distribution dynamics. *The Economic Journal*, *106*(437), 1045-1055.

Roodman, D., MacKinnon, J. G., Nielsen, M. Ø., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal*, *19*(1), 4-60.

Salvati, L. (2025). Regional development, spatial disparities, and urbanization: A comparative approach. In *Economic Growth, Regional Disparities, and Urban Sprawl: The Mediterranean Experience* (pp. 61-97). Springer Nature Switzerland.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica*, *48*(3), 613-625.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. Routledge.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, *70*(1), 65-94.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 80-108). Cambridge University Press.

Tao, Z., Zhang, Z., & Shangkun, L. (2022). Digital economy, entrepreneurship, and high-quality economic development: Empirical evidence from urban China. *Frontiers of Economics in China*, *17*(3), 393.

Warr, P. (2013). Thailand's development strategy and growth performance. In A. Fosu (Ed.), *Achieving Development Success: Strategies and Lessons from the Developing World* (pp. 72-94). Oxford University Press.

Williamson, J. G. (1965). Regional inequality and the process of national development: A description of the patterns. *Economic Development and Cultural Change*, *13*(4, Part 2), 1-84.

Yang, C.-H., & Lin, H.-L. (2012). Openness, absorptive capacity, and regional innovation in China. *Environment and Planning A*, *44*(2), 333-355.