



Urban Expansion and Provincial Economic Convergence in Thailand: A Satellite-Based Rank-Size and System GMM Analysis

Wittawat Kadthan^{1,*}

¹ School of Economics and Investment, Bangkok University,
Rangsit Campus, Pathum Thani, Thailand

Abstract

This study jointly examines Thailand's urban size distribution and provincial economic convergence over 1995–2024, using multi-epoch satellite-derived built-up area for 77 provinces and gross provincial product (GPP) records for 76 provinces. Three empirical strategies correspond to three objectives: rank-size (Zipf's Law) analysis across six epochs, a test of Gibrat's Law of proportionate urban growth, and a dynamic panel System GMM model of GPP convergence. The Zipf exponent rises modestly from 1.158 (1995) to 1.206 (2020), indicating a stable but persistently primate urban hierarchy. Gibrat's Law is rejected ($\beta = -0.0039$, $p = 0.010$): smaller urban areas expanded faster than Bangkok, whose primacy index fell from 0.493 to 0.443. Cross-sectional beta-convergence in GPP is not significant unconditionally ($p = 0.142$), and urban expansion is only marginally associated with GPP growth ($p = 0.064$); however, once structural controls (initial population, agricultural share, distance to Bangkok, education) are added, the coefficient on initial GPP reverses sign and becomes significant ($\beta = -0.0055$, $p = 0.012$ including Bangkok; $\beta = -0.0062$, $p = 0.007$ excluding Bangkok, $R^2 \approx 0.26$ – 0.28), indicating genuine conditional convergence. The System GMM estimate of GPP persistence is near unity ($\rho = 0.997$), implying a convergence speed of just 0.32% per year and a half-life exceeding two centuries. These results indicate that Thailand's urban system is deconcentrating physically, and that provinces are converging conditionally toward different structurally determined steady states even though absolute convergence is absent, with direct implications for secondary-city infrastructure investment and the use of open-access satellite data for subnational economic monitoring.

Keywords: Zipf's Law, Gibrat's Law, Urban primacy, Regional convergence, System GMM, Satellite built-up data

Article history: Received 9 July 2026, Revised 27 July 2026, Accepted 7 August 2026

1. Introduction

More than half of the world's population now lives in cities, and city-size distributions worldwide follow the rank-size, or Zipf's Law, regularity, whereby a city's rank is approximately proportional to the inverse of its size [1,2]. Because city-size distributions and urban growth are closely tied to how efficiently economic activity is spatially organized, departures from proportionate growth carry direct implications for national development [3]. Southeast Asia illustrates this tension sharply: regional income convergence coexists with persistent clustering of high-income

activity around primary cities [4,5], and cross-country evidence shows no uniform relationship between urban concentration and growth [6], motivating country-specific study. Thailand is a stark case: Bangkok has been the political, financial, and demographic core for over a century. Panel evidence identifies three distinct provincial convergence clubs rather than uniform catch-up [7], with agglomeration around Bangkok found to slow convergence relative to earlier-invested provinces [8]; remotely sensed built-up area has also been shown to track provincial GPP growth closely in

*Corresponding author; e-mail: wittawat.k@bu.ac.th

Thailand [9], validating satellite data as a high-frequency indicator of subnational activity.

This relationship is grounded in established theory. Zipf's Law is the equilibrium outcome of proportionate urban growth [10]; primate-city concentration is systematically related to a country's stage of development [11,12], and New Economic Geography explains such concentration as the result of circular causation between manufacturing location and demand, with falling transport costs eventually favoring dispersal to the periphery [13]. Empirically, provincial GPP series in developing countries are compiled at low frequency and are prone to revision, motivating remotely-sensed proxies such as built-up area or nighttime lights [14,15]; the Global Human Settlement Layer programme now provides harmonized, multi-temporal built-up grids for this purpose [16,17]. On the convergence side, standard fixed-effects panel estimators are biased when the lagged dependent variable is included and the time dimension is short [18], a problem addressed by difference and system GMM estimators [19,20], the latter well suited to persistent, near-unit-root income series. No published study has combined multi-epoch satellite built-up data, joint Zipf-Gibrat rank-size testing, and dynamic panel System GMM convergence estimation for Thailand within a single framework. This combination is not merely a methodological convenience: the three strategies test related but distinct implications of the same underlying spatial growth process, so that agreement or disagreement among them is itself informative. This paper's central empirical contribution is to show that this triangulation matters in practice: an unconditional test finds no evidence of provincial income convergence in Thailand, yet a conditional specification informed by the same rank-size and proportionate-growth framework reveals that provinces do converge once structural characteristics are held constant, a distinction with direct implications for how targeted the country's regional development policy needs to be.

This paper addresses that gap by jointly examining Thailand's urban size distribution and provincial economic convergence between 1995 and 2024, pursuing three objectives: (1) to examine the rank-size (Zipf) distribution of Thailand's urban system across six epochs from 1995 to 2020 and assess its temporal stability; (2) to test Gibrat's Law of proportionate urban growth; and (3) to evaluate provincial economic convergence through cross-sectional beta-convergence and a dynamic panel System GMM model of GPP persistence over 1995–2024, extended with a conditional convergence specification incorporating structural controls. Section

2 describes the data and empirical specifications; Section 3 reports the results; Section 4 discusses the findings and their robustness; and Section 5 concludes with policy implications and limitations.

2. Materials and Methods

This section describes the data sources, variable construction, and empirical specifications used to test the three objectives stated above. The analysis proceeds from the physical (built-up area) to the economic (GPP) dimension of Thailand's spatial development, using rank-size analysis, cross-sectional regression, and dynamic panel estimation in turn.

2.1 Data sources and variable construction

Built-up surface area for 77 provincial-level administrative units of Thailand is drawn from the Global Human Settlement Layer built-up surface grid (GHS-BUILT-S R2023A), produced by the European Commission's Joint Research Centre from harmonized Sentinel-2 and Landsat composites at 100-metre resolution [21]; methodology and validation are documented in [16] and the earlier GHSL Data Package release [22]. Administrative boundaries follow the geoBoundaries open database. Built-up surface area is extracted for six epochs (1995, 2000, 2005, 2010, 2015, 2020) and aggregated to the provincial level. Provincial gross provincial product (GPP) at constant 2002 prices (chain volume measures) for 76 provinces is obtained from the Office of the National Economic and Social Development Council (NESDC) for 1995–2024. The cross-sectional analyses (Sections 3.1 and 3.5, Tables 1, 6–8) use these 76 provinces with complete 1995–2020 coverage; the dynamic panel System GMM model (Section 3.6, Table 9) additionally includes Bueng Kan, whose GPP series is unbalanced but partially available, for a total panel of 77 provinces. Ten provinces are classified as regional economic hubs (Chiang Mai, Chiang Rai, Khon Kaen, Nakhon Ratchasima, Udon Thani, Surat Thani, Songkhla, Phuket, Chon Buri, and Rayong), with Bangkok treated as a separate observation given its outsized scale.

2.2 Rank-size (Zipf's Law) and urban primacy

To examine the shape of the urban size distribution, provinces are ranked by built-up area within each epoch and the standard rank-size regression is estimated separately for each of the six epochs:

$$\ln(\text{Rank}_i) = \alpha - \zeta \cdot \ln(S_i) + \varepsilon_i \quad (1)$$

where Rank_i is the descending-order rank of province i by built-up area, S_i is built-up surface area, and ζ is the Zipf coefficient. A value of $\zeta = 1$ corresponds to the

pure rank-size rule; $\zeta > 1$ indicates a system more concentrated at the top of the size distribution than pure Zipf predicts, and $\zeta < 1$ indicates a more even distribution [10]. Urban primacy is further summarized for each epoch t using the conventional four-city primacy index,

$$PI_t = \frac{S_{1,t}}{S_{2,t} + S_{3,t} + S_{4,t}} \quad (2)$$

where S_1 denotes the built-up area of the largest urban unit (Bangkok) and S_2 through S_4 denote the second-, third-, and fourth-ranked units, following [11].

2.3 Gibrat's Law of proportionate growth

Gibrat's Law of proportionate growth is tested by regressing each province's annualized 1995–2020 built-up growth rate on its initial (1995) built-up area:

$$g_i = \alpha + \beta \cdot \ln(S_{i,1995}) + \varepsilon_i \quad (3)$$

where g_i is the compound annual growth rate (CAGR) of built-up area between 1995 and 2020. Under strict Gibrat's Law, as originally formulated by [23], growth is independent of initial size and $\beta = 0$; a significant negative β implies that smaller urban areas grow disproportionately faster, consistent with size convergence in the urban system [24]; [25].

2.4 Cross-sectional convergence models

Provincial economic convergence is first assessed with two unconditional cross-sectional ordinary least squares (OLS) regressions over 1995–2020, both estimated with heteroscedasticity-robust (HC1) standard errors to accommodate the small-sample, cross-sectional design. The classical beta-convergence specification is

$$CAGR_i = \alpha + \beta \cdot \ln(GPP_{i,1995}) + \varepsilon_i \quad (4)$$

where $CAGR_i$ is the compound annual growth rate of provincial GPP, $GPP_{i,1995}$ is initial GPP, and a significant negative β indicates convergence – poorer provinces growing faster than richer ones [26]. To assess whether physical urban expansion is associated with economic growth, a second specification regresses GPP growth on urban built-up expansion directly:

$$GPPCAGR_i = \alpha + \beta \cdot UrbanCAGR_i + \varepsilon_i \quad (5)$$

where $UrbanCAGR_i$ is the 1995–2020 built-up area growth rate of province i .

2.5 Dynamic panel System GMM model

Because cross-sectional regressions cannot exploit the full time-series information in the 1995–

2024 GPP panel and are silent on year-specific shocks, a dynamic panel model is estimated using the two-step System Generalized Method of Moments (GMM) estimator of [20], implemented via Stata's `xtabond2` following [27]. Ordinary within-group or first-differenced estimators yield severely biased autoregressive coefficients when the lagged dependent variable is included and the time dimension is short [18]; System GMM [20], which augments the first-differenced moment conditions of [19] with levels equations, corrects this bias and suits the near-unit-root persistence found in provincial income series. The estimating equation is

$$\ln(GPP_{i,t}) = \rho \cdot \ln(GPP_{i,t-1}) + \sum_t \gamma_t \cdot D_t + \eta_i + u_{i,t} \quad (6)$$

where $GPP_{i,t}$ is provincial GPP in year t , ρ is the persistence parameter, D_t are year dummies (1997–2024, collapsed, GMM-style instruments lagged two to four periods), η_i is a province fixed effect removed by the System GMM transformation, and $u_{i,t}$ is an idiosyncratic error. Instrument validity is assessed with the Hansen J-test [28], noting the caution of [29] that excessive instruments can render the test uninformative. From ρ , an implied convergence speed and income half-life are computed as

$$\lambda = -\ln(\rho), \quad t_{1/2} = \frac{\ln(2)}{\lambda} \quad (7)$$

All cross-sectional regressions were estimated in Python (`statsmodels`); the dynamic panel GMM model was estimated in Stata 18 using `xtabond2`.

3. Results

This section presents the empirical findings in three parts, moving from the physical size distribution of Thailand's urban system to its economic convergence properties. Descriptive statistics are reported first, followed by the rank-size (Zipf) and proportionate-growth (Gibrat) results, and finally the cross-sectional and dynamic-panel convergence estimates.

3.1 Descriptive statistics

Table 1 summarizes the built-up area and GPP variables across the 76 provinces with complete GPP series. Mean built-up area expanded from approximately 51 km² in 1995 to 88 km² in 2020, with an average annualized growth rate of 2.26%, while mean GPP grew at 2.64% per year over the same period. GPP's standard deviation is an order of magnitude larger than its mean, reflecting the extreme right skew introduced by Bangkok.

Table 1. Descriptive statistics of built-up area and gross provincial product (GPP), 76 Thai provinces, 1995 and 2020

Variable	n	Mean	SD	Min	Max
Built-up area, 1995 (km ²)	76	51.34	37.84	8.57	212.15
Built-up area, 2020 (km ²)	76	87.87	61.52	15.17	334.61
Urban built-up CAGR, 1995–2020 (%/yr)	76	2.26	0.98	1.06	5.25
GPP, 1995 (million baht)	76	66,841	226,178	4,547	1,963,319
GPP, 2020 (million baht)	76	135,634	434,528	7,591	3,740,369
GPP CAGR, 1995–2020 (%/yr)	76	2.64	1.24	0.03	8.90

3.2 Rank-size (Zipf's Law) distribution

Table 2 reports the estimated Zipf coefficient for each of the six epochs from 1995 to 2020, obtained

by regressing the log rank of each province's built-up area on its log built-up area.

Table 2. Estimated Zipf coefficient by epoch, Thai urban built-up areas, 1995–2020

Epoch	Zipf coefficient (ζ)	SE ^a	R ²
1995	1.158	0.0903	0.819
2000	1.181	0.0892	0.819
2005	1.190	0.0896	0.820
2010	1.196	0.0897	0.821
2015	1.199	0.0894	0.821
2020	1.206	0.0886	0.817

Note: ^a SE for all six epochs derived from a bootstrap procedure (2,000 resamples per epoch, resampling provinces and re-estimating the rank-size regression), consistent with the paired bootstrap test of the 1995–2020 difference reported in Section 3.2. SEs are stable across epochs (0.089–0.090), and the corresponding 95% confidence intervals overlap substantially across adjacent epochs, reinforcing the conclusion in Section 3.2 that the rise in ζ is not statistically significant.

Table 3. Five largest Thai provinces by built-up area, 2020

Rank	Province	Built-up area, 2020 (km ²)
1	Bangkok	334.6
2	Nakhon Ratchasima	310.0
3	Chon Buri	244.1
4	Ubon Ratchathani	201.3
5	Chiang Mai	175.4

The Zipf coefficient rises steadily and monotonically from 1.158 in 1995 to 1.206 in 2020, with model fit essentially unchanged ($R^2 = 0.817$ – 0.821) across all six epochs. Because ζ exceeds unity throughout, Thailand's urban hierarchy is more concentrated at the top than a pure rank-size rule predicts, and this concentration intensified marginally over the 25-year window. This 1995–2020 change in ζ is small in magnitude and not statistically distinguishable from zero: a paired bootstrap test (2,000 resamples, resampling provinces jointly across the two epochs) yields a mean difference of 0.046 with 95% CI $[-0.055, 0.154]$, and a pooled rank-size regression with a year-interaction term gives an equivalent interaction coefficient of -0.048 (robust

SE = 0.154, $p = 0.753$). The appropriate reading is therefore a stable rather than intensifying urban hierarchy across the sample period. Table 3 lists the five largest provinces by built-up area in 2020: Bangkok remains largest by a wide margin, followed by Nakhon Ratchasima, Chon Buri, Ubon Ratchathani, and Chiang Mai. Figure 1(a) shows that while the log-linear rank-size pattern is preserved, the entire distribution shifted outward as built-up area expanded. This stability mirrors the finding of [30] for the United States (1900–1990) and is consistent with the resilience of national Zipf structures to structural change documented by [31] for Germany. The question that follows is whether the underlying growth was

proportionate across the size distribution, as required by Gibrat's Law, or size-dependent. The use of the province, rather than the city proper, as the unit of rank-size analysis follows directly from [2]'s theoretical result that Gibrat's Law holding within any sub-national partition implies Zipf's Law at that same

spatial scale, a prediction confirmed empirically at the regional level for Germany [41] and, using spatial-econometric methods that account for inter-provincial dependence, for prefectural-level regions in China [42].

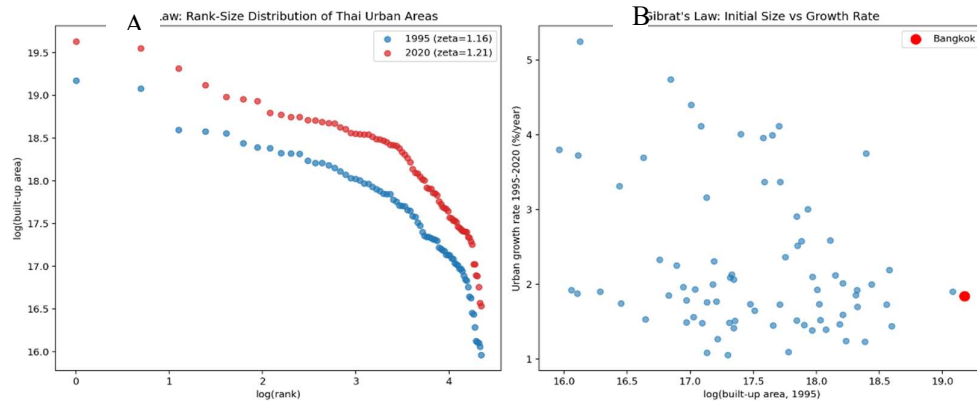


Figure 1. (A) Rank-size (Zipf) distribution of Thai urban built-up areas, 1995 and 2020; (B) Gibrat's Law scatter of initial built-up size (1995, log scale) against 1995–2020 urban growth rate, by province. Bangkok is marked separately in panel (B).

3.3 Gibrat's Law and proportionate urban growth

Table 4 reports the OLS regression of each province's 1995–2020 built-up growth rate on its

initial (log) built-up area in 1995, the standard test of Gibrat's Law.

Table 4. OLS regression of urban built-up growth rate (1995–2020) on initial built-up area, 1995 (test of Gibrat's Law)

Variable	Coefficient	Robust SE (HC1)	z	p-value
Constant	0.0906	0.027	3.362	0.001
log(built-up area, 1995)	-0.0039	0.002	-2.572	0.010

Note: $n = 77$; $R^2 = 0.082$; Dependent variable: built-up area growth rate, 1995–2020.

The coefficient on initial size is negative and statistically significant ($\beta = -0.0039$, robust SE = 0.0015, $z = -2.572$, $p = 0.010$), and the model explains 8.2% of cross-sectional variance in growth rates ($n = 77$). Gibrat's Law is therefore rejected: provinces with smaller initial built-up area grew systematically faster than those with larger initial area over 1995–2020. Figure 1(b) plots this relationship, with Bangkok positioned at the lower-right, growing below the sample average. This size-dependent growth pattern is reconciled with the rising Zipf coefficient explicitly in Section 4. Although statistically significant, the model explains only 8.2% of the cross-sectional variance in growth rates, indicating that initial size accounts for a modest share of why some provinces expand faster than

others; the bulk of the variation is attributable to province-specific factors outside the model, such as local infrastructure investment, sectoral composition, and proximity to transport corridors. This limited explanatory power is consistent with comparable Gibrat-law tests in other national contexts [24,25] and cautions against over-interpreting the rejection of Gibrat's Law as a complete account of urban growth dynamics; it should be read as evidence against strict proportionate growth rather than as a fully explanatory growth model.

3.4 Urban primacy and deconcentration

Table 5 reports the four-city urban primacy index for each of the six epochs.

Table 5. Urban Primacy Index (Bangkok relative to provinces ranked 2–4), 1995–2020

Epoch	Urban Primacy Index	% change vs. 1995
1995	0.493	— (base)
2000	0.530	+7.51%
2005	0.510	+3.45%
2010	0.487	–1.22%
2015	0.466	–5.48%
2020	0.443	–10.14%

The primacy index declines from 0.493 in 1995 to 0.443 in 2020, a fall of roughly ten percent, indicating that Bangkok's built-up area shrank in relative terms compared with the combined area of the second-, third-, and fourth-ranked provinces. This decline is monotonic across every epoch and is directionally consistent with the Gibrat's Law rejection above: smaller and mid-ranked units expanded faster

than the primate city, producing gradual relative, though not absolute, deconcentration.

3.5 Provincial economic convergence

Table 6 reports the two cross-sectional convergence regressions for GPP CAGR over 1995–2020.

Table 6. Cross-sectional OLS estimates of provincial GPP convergence and the urban expansion–growth relationship, 1995–2020

Variable	Model 1: Beta-convergence	Model 2: Urban expansion – growth
Constant	0.0120 (0.010)	0.0177*** (0.004)
log(GPP, 1995)	0.0014 (0.001)	–
Urban built-up CAGR	–	0.3882* (0.210)
R ²	0.014	0.093
n	76	76

Note: Robust (HCl) standard errors in parentheses. Dependent variable: GPP CAGR, 1995–2020. *** $p < 0.01$, * $p < 0.10$ (Model 1 coefficient of interest not significant, $p = 0.142$; Model 2 coefficient of interest $p = 0.064$).

Because unconditional convergence tests are known to be uninformative when provinces converge to different steady states, Table 7 reports a conditional beta-convergence specification that augments Model 1 with urban built-up CAGR and four structural control variables: log initial population (1995), the agricultural share of GPP (1995), distance to Bangkok, and mean

years of education (2010 census, the only year with complete province coverage in the available data). Investment/GFCF and FDI were confirmed unavailable at the provincial level in Thailand's official statistics and are omitted; this is disclosed as a limitation in Section 4.2.

Table 7. Conditional beta-convergence with structural controls, GPP CAGR 1995–2020

Variable	Model 3: Conditional (incl. Bangkok)	Model 4: Conditional (excl. Bangkok)
Constant	0.0317* (0.017)	0.0227 (0.018)
log(GPP, 1995)	–0.0055** (0.002)	–0.0062*** (0.002)
Urban built-up CAGR	0.5466** (0.268)	0.5157* (0.285)
log(Population, 1995)	0.0061*** (0.002)	0.0078*** (0.002)
Agricultural share, 1995	–0.0329*** (0.012)	–0.0275* (0.014)
Distance to Bangkok (km)	0.0000 (0.0000)	–0.0000 (0.0000)
Mean years of education (2010)	0.0007 (0.001)	0.0016 (0.001)
R ²	0.261	0.275
n	76	75

Note: Robust (HCl) standard errors in parentheses. Dependent variable: GPP CAGR, 1995–2020. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Condition number is large ($\sim 7.1 \times 10^4$) in both models, likely reflecting the raw-km scale of distance to Bangkok rather than substantive collinearity; rescaling to 100-km units would improve this diagnostic without altering point estimates or significance.

Once these controls are included, the coefficient on log initial GPP changes sign and becomes statistically significant: $\beta = -0.0055$ ($p = 0.012$) including Bangkok and $\beta = -0.0062$ ($p = 0.007$) excluding Bangkok, with R^2 rising to 0.261 and 0.275 respectively. This is the standard signature of conditional, rather than absolute, convergence — provinces converge toward different steady states determined by their structural characteristics, even though no evidence of absolute convergence was found

in Table 6. Initial population and agricultural share are the strongest and most robust controls (both significant at the 1% level in most specifications), while distance to Bangkok and mean education are not statistically significant, possibly reflecting the coarseness of these particular proxies. The urban built-up CAGR coefficient also strengthens under conditioning (0.547 vs. 0.388 unconditional) and regains 5%-level significance including Bangkok, though it weakens to the 10% level once Bangkok is excluded ($p = 0.070$).

Table 8. Mean built-up and GPP growth rates by province group, 1995–2020

Group	n	Urban built-up CAGR (%/yr)	GPP CAGR (%/yr)
Regional hub provinces	10	2.25	3.10
Other provinces	65	2.26	2.57
Bangkok	1	1.84	2.61

In the classical beta-convergence specification (Model 1), the coefficient on log initial GPP is positive but not statistically distinguishable from zero ($\beta = 0.0014$, robust SE = 0.001, $z = 1.468$, $p = 0.142$; $R^2 = 0.014$, $n = 76$), providing no evidence of unconditional convergence. In Model 2, the coefficient linking urban expansion to GPP growth is positive and marginally significant at 10% ($\beta = 0.388$, robust SE = 0.210, $z = 1.852$, $p = 0.064$; $R^2 = 0.093$), indicating that provinces with faster physical expansion tended to also record faster GPP growth, though imprecisely estimated. Table 8 disaggregates growth by province group: the ten regional hubs recorded a mean GPP CAGR of 3.10% against 2.57% for the remaining 65 provinces, while Bangkok grew at only 2.61% despite the lowest built-up growth rate (1.84%) of the three groups. Figure 2 plots both relationships, with hub provinces and Bangkok highlighted. Two robustness

checks were carried out on Model 1 and Model 2. First, excluding Bangkok leaves the beta-convergence coefficient statistically insignificant and of similar magnitude ($\beta = 0.0019$, robust SE = 0.001, $z = 1.580$, $p = 0.114$; $R^2 = 0.018$, $n = 75$), confirming that the absence of unconditional convergence is not an artifact of Bangkok's outsized scale. Second, when initial GPP and urban built-up CAGR are entered jointly in a multivariate specification, the coefficient on initial GPP falls essentially to zero in both the full sample ($\beta = 0.0004$, $p = 0.696$, $n = 76$) and the Bangkok-excluded sample ($\beta = 0.0004$, $p = 0.755$, $n = 75$), while the coefficient on urban built-up CAGR remains positive but loses its 10%-level significance once Bangkok is excluded ($\beta = 0.375$, $p = 0.103$, $n = 75$, versus $\beta = 0.377$, $p = 0.091$, $n = 76$ including Bangkok); the urban expansion-growth relationship in Model 2 should therefore be read as suggestive rather than robust.

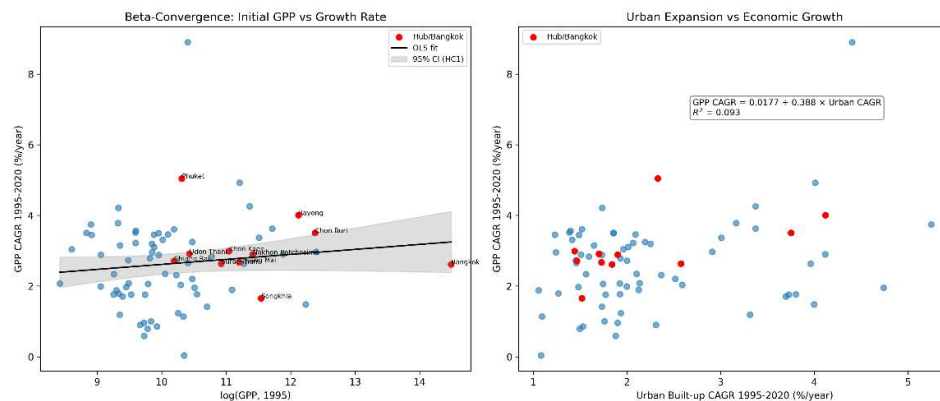


Figure 2. (A) Beta-convergence: initial (1995, log) GPP against GPP CAGR, 1995–2020, with OLS fit and 95% confidence interval (HC1); (B) urban built-up expansion against GPP growth, 1995–2020, with fitted regression equation and R^2 annotated. Red markers denote Bangkok and the ten regional hub provinces.

3.6 Dynamic panel System GMM convergence estimates

Table 9 reports the key parameters of the dynamic panel System GMM model of log GPP over the full 1995–2024 panel.

Table 9. Two-step System GMM estimates of dynamic panel log(GPP) persistence, 77 provinces, 1995–2024 (selected coefficients)

Variable	Coefficient	SE	z	p-value
L1.log(GPP)	0.996767	0.025201	39.553	<0.001
Constant	0.103792	0.254858	0.407	0.684

Note: Two-step System GMM (Blundell–Bond); GMM-style instruments lagged 2–4 periods, collapsed; year dummies 1997–2024 included (see Appendix Table A1). Dependent variable: log(GPP), 1995–2024, 77 provinces (unbalanced panel; Bueng Kan contributes 14 of 30 years), 2,140 province-year observations, 33 instruments. Specification diagnostics: Arellano-Bond AR(1) $z = -5.415$, $p < 0.001$ (expected by construction); AR(2) $z = -0.166$, $p = 0.868$ (fails to reject, supporting instrument validity); Hansen J-test of overidentifying restrictions $\chi^2(3) = 8.586$, $p = 0.035$.

The coefficient on lagged log GPP is 0.9968 (SE = 0.0252, $z = 39.55$, $p < 0.001$), indicating extremely high persistence in provincial GPP. The implied annual convergence speed is $\lambda = -\ln(0.9968) \approx 0.0032$, or roughly 0.32% per year, corresponding to an income half-life of approximately 214 years – far slower than the canonical 2% annual convergence speed reported for US states and European regions [26]. Year dummies are jointly informative: most are negative and significant relative to the 1995–1996

baseline, with the largest negative deviation in 2020 (-0.116 , $p < 0.001$), consistent with COVID-19-related GPP contraction, and smaller significant deviations around 2008–2009 and 2011, corresponding to the Global Financial Crisis and the 2011 Thailand floods. Figure 3 maps built-up levels, growth rates, and urban rank across all 77 provinces, corroborating that built-up expansion is fastest around Bangkok and in smaller, lower-rank provinces of the upper North and Northeast, rather than in Bangkok itself.

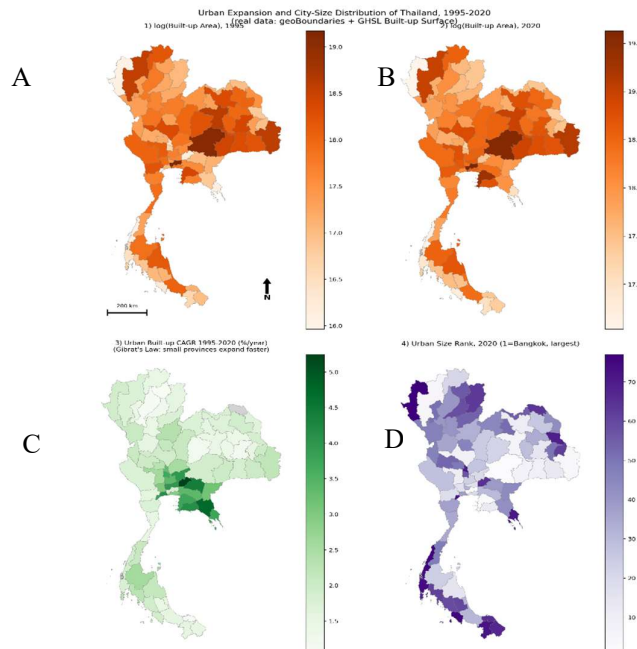


Figure 3. Spatial distribution of (A) log built-up area, 1995; (B) log built-up area, 2020; (C) built-up area CAGR, 1995–2020; and (D) urban size rank, 2020 (1 = Bangkok, largest), across 77 Thai provincial-level administrative units. Data: geoBoundaries and GHSL built-up surface.

4. Discussion

4.1 Urban hierarchy and proportionate growth

The Zipf coefficient of 1.16–1.21 places Thailand's urban hierarchy within the range documented across developing countries generally, where average estimates exceed unity [32,33]; this reflects the historical concentration of administrative, financial, and manufacturing functions in Bangkok, consistent with the primate-city dynamic [11] and circular-causation logic [13] of self-reinforcing agglomeration -- though, as noted in Section 3.2, the modest increase in ζ over 1995-2020 is not itself statistically significant, so the hierarchy is best characterized as stable rather than actively intensifying over this period. The rejection of Gibrat's Law ($\beta = -0.0039$, $p = 0.010$) shows that smaller urban areas expanded systematically faster than Bangkok, plausibly reflecting diminishing returns to further agglomeration at the core as congestion and land costs rise, pushing new development toward lower-cost peripheral locations [34]. This departure from strict proportionate growth is consistent with evidence from untruncated city-size samples elsewhere [24,25] and with a comparable developing-country departure documented for Brazil [35]. The apparent tension between a mildly rising Zipf coefficient and a falling primacy index is reconcilable: the Zipf exponent summarizes the full rank-size distribution, including the lower tail of small provinces, whereas the primacy index isolates Bangkok's position relative to its closest rivals only [10,37]. A further caveat applies throughout this section: built-up surface area is a physical footprint measure, not a direct measure of urbanization or economic agglomeration. An increase in built-up area may reflect low-density suburban or peri-urban sprawl rather than the denser, more productive land use typically associated with agglomeration economies, and the elasticity linking built-up or night-light intensity to true economic output has been shown to be unstable across regions even in data-rich settings [45], with proxy performance degrading further in lower-density, more rural contexts [46]. The built-up-based findings in this section should therefore be read as evidence about the physical footprint of Thailand's urban system specifically, which need not move in lockstep with economic concentration; this is one reason the built-up-based results are triangulated against the independent GPP-based convergence evidence in Section 4.2 rather than treated as interchangeable.

4.2 Provincial economic convergence

Provincial economic convergence, by contrast, remains extremely slow. Unconditional beta-convergence is not statistically significant ($p = 0.142$), a result that should not be over-interpreted given known limitations of cross-sectional convergence tests [36], while urban expansion is only marginally associated with GPP growth ($p = 0.064$), echoing the built-up–GPP relationship previously documented for Thai provinces [9]; both results are robust to excluding Bangkok, and the urban-expansion coefficient loses even its marginal significance once initial GPP and urban expansion are modeled jointly and Bangkok is dropped (Section 3.5), indicating that this weaker part of the headline finding is not driven by the capital's outsized weight in the sample. However, the absence of unconditional convergence should not be read as an absence of convergence altogether. Because convergence and divergence regimes among sub-national units can shift once conditioned on structural characteristics [40], Table 7 augments the beta-convergence specification with province-level controls shown to matter for Thai provincial convergence [43] and for dynamic-panel convergence across ASEAN more broadly [44]: log initial population, agricultural share of GPP, distance to Bangkok, and mean education. Conditioning on these controls reverses the sign of the coefficient on initial GPP and renders it statistically significant ($\beta = -0.0055$, $p = 0.012$ including Bangkok; $\beta = -0.0062$, $p = 0.007$ excluding Bangkok) — the standard signature of conditional convergence: provinces do not converge to a common income level, but do converge toward different steady states once their structural characteristics are held constant. This is a materially stronger and more standard growth-economics finding than the unconditional result alone would suggest, and we regard it as a core contribution of the revised analysis. The System GMM estimate of near-unit-root persistence ($\rho = 0.997$) is consistent with the club-convergence pattern found for Thailand [7] and with evidence that Bangkok-centred agglomeration has historically slowed provincial income convergence [8]. That the ten regional hub provinces outgrew both Bangkok and the remaining provinces in GPP terms, despite built-up expansion rates similar to the provincial average, suggests these hubs function as intermediate growth poles rather than sites of purely physical sprawl.

4.3 Robustness and methodological considerations

Three caveats bear on these findings. The Hansen J-test ($p = 0.035$) signals some caution

regarding GMM instrument validity, though the lag-2-to-4 specification follows recommended practice and near-unit-root persistence is itself a challenging setting for any GMM estimator [20,29]; the remaining specification diagnostics are more reassuring, with the Arellano-Bond test correctly rejecting the null of no first-order serial correlation in first differences ($z = -5.415$, $p < 0.001$) while failing to reject the null of no second-order serial correlation ($z = -0.166$, $p = 0.868$), the pattern required for the moment conditions underlying System GMM to be valid. Unlike the cross-sectional regressions, the System GMM model absorbs province fixed effects rather than a small set of observable controls, so its very high estimated persistence should still be weighted heavily when drawing conclusions about long-run convergence, even though Table 7 now provides conditional cross-sectional evidence of convergence. As a further robustness check, urban built-up area was incorporated directly into the System GMM specification (interpolated to annual frequency between the six built-up epochs and necessarily truncating the panel to 1995-2020, since built-up area cannot be observed or extrapolated beyond the most recent epoch). The persistence parameter is essentially unchanged ($\rho = 0.993$ vs. 0.997 in the baseline), built-up area itself is not a significant predictor of GPP once dynamic persistence is controlled for (coefficient = -0.025 , $p = 0.618$), and AR(2) continues to pass cleanly ($z = 1.01$, $p = 0.313$); the Hansen test also rejects at 5% in this augmented specification ($\chi^2(5) = 13.335$, $p = 0.020$), confirming that the borderline Hansen result is a pre-existing feature of the GMM specification rather than one introduced by adding built-up area. Even so, the Zipf, Gibrat, and System GMM results triangulate through three independent channels on a consistent picture: gradual physical deconcentration alongside very slow unconditional (though not conditional) income convergence.

5. Conclusions

This study examined the joint evolution of Thailand's urban size distribution and provincial economic convergence between 1995 and 2024 using multi-epoch satellite built-up data, rank-size analysis, and dynamic panel GMM estimation. Three findings correspond to the study's three objectives. First, Thailand's urban hierarchy is stable but mildly intensifying, with the Zipf coefficient rising from 1.158 (1995) to 1.206 (2020) while model fit remains essentially unchanged ($R^2 \approx 0.82$). Second, Gibrat's Law is rejected: smaller urban areas grew significantly faster than larger ones ($\beta = -0.0039$, $p = 0.010$), and Bangkok's primacy index fell from 0.493 to 0.443. Third, provincial income convergence is absent

unconditionally but present conditionally: unconditional beta-convergence is insignificant and the System GMM model implies an income half-life exceeding two centuries ($\rho = 0.997$), yet once initial population, agricultural share, distance to Bangkok, and education are held constant, the coefficient on initial GPP reverses sign and becomes significant ($\beta = -0.0055$ to -0.0062 , $p < 0.05$), indicating genuine conditional convergence toward structurally-determined steady states, even as regional hub provinces outgrew both Bangkok and the remaining provinces.

These results carry three policy implications. First, because physical expansion has not yet translated into commensurate unconditional income convergence, though the conditional convergence evidence in Table 7 shows that provinces with more favorable initial population, industrial structure, and urban expansion do converge toward higher steady states, regional development policy should prioritize complementary investments in human capital, secondary-city industrial capacity, and logistics connectivity, so that urban growth in smaller provinces yields durable income gains rather than built-up area alone [38], noting that mean education was not itself a significant control in Table 7, so this recommendation rests on the general growth literature rather than on direct evidence from this specification. Second, the ten regional hub provinces, which recorded the fastest GPP growth despite average built-up expansion, warrant continued policy attention as intermediate growth poles that can absorb activity that would otherwise concentrate further in Bangkok, consistent with the New Economic Geography prediction that falling transport costs favor gradual dispersal from a dominant core [13,39]. Third, given that provincial GPP statistics are compiled at low frequency and are subject to revision, national planning agencies should consider open-access satellite-derived built-up and nighttime-light indicators as a complementary, higher-frequency monitoring tool, particularly where administrative data capacity is limited [15,16].

Two limitations qualify these conclusions: the conditional convergence model in Table 7 should still be read as descriptive rather than strictly causal, since the control variables are not plausibly exogenous and the education control is limited to a single 2010 cross-section rather than a true 1995 initial condition; investment and FDI, two theoretically important controls, could not be included because they are not published at the provincial level in Thailand's official statistics; and the System GMM model's very high estimated persistence leaves limited scope to distinguish genuine slow convergence from a unit-root

process, a concern only partly mitigated by confirming that adding built-up area to the dynamic panel leaves this persistence estimate essentially unchanged ($\alpha = 0.993$). Future research should extend the conditional convergence framework with instrumented or panel-based controls, update the satellite panel beyond 2020 as new GHSL epochs become available, and apply spatial econometric methods to model inter-provincial spillovers [7]. Overall, four independent empirical strategies converge on a nuanced narrative: Thailand's urban system is deconcentrating gradually in physical terms, absolute provincial income convergence is absent, yet conditional convergence toward structurally-determined steady states is present and statistically robust, underscoring the continued relevance of Bangkok-centred agglomeration to the country's regional economic geography.

6. Acknowledgements

The author gratefully acknowledges the use of open-access data from the Global Human Settlement Layer programme of the European Commission Joint Research Centre, provincial gross provincial product statistics published by the Office of the National Economic and Social Development Council of Thailand, and the institutional support of Bangkok University. The complete set of year-fixed-effect coefficients from the System GMM model summarized in Table 9 is available from the corresponding author upon request.

References

- [1] G. K. Zipf, *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*, Addison-Wesley Press, Cambridge, MA, 1949.
- [2] X. Gabaix, Zipf's law for cities: An explanation, *Quarterly Journal of Economics* 114(3) (1999) 739 – 767.
- [3] G. Duranton, Growing through cities in developing countries, *World Bank Research Observer* 30(1) (2015) 39 – 73.
- [4] C. Mendez, F. Santos-Marquez, Regional convergence and spatial dependence across subnational regions of ASEAN: Evidence from satellite nighttime light data, *Regional Science Policy & Practice* 13(6) (2021) 1750 – 1778.
- [5] W. Lang, M. Pan, J. Wu, T. Chen, X. Li, The patterns and driving forces of uneven regional growth in ASEAN countries: A tale of two Thailand's path toward regional coordinated development, *Growth and Change* 52(1) (2021) 130 – 149.
- [6] S. A. Frick, A. Rodríguez-Pose, Change in urban concentration and economic growth, *World Development* 105 (2018) 156 – 170.
- [7] K. Tipayalai, C. Mendez, Regional convergence and spatial dependence in Thailand: Global and local assessments, *Journal of the Asia Pacific Economy* 29(2) (2024) 693 – 720.
- [8] S. Limpanonda, Provincial disparities, convergence and effects on poverty in Thailand, *The Singapore Economic Review* 60(2) (2015) 1550015.
- [9] S. Teerarojanarat, Remotely-sensed derived built-up area as an alternative indicator in the study of Thailand's regional development, *Indonesian Journal of Geography* 54(2) (2022) 240 – 253.
- [10] X. Gabaix, Y. M. Ioannides, The evolution of city size distributions, in: J. V. Henderson, J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Vol. 4, Elsevier, Amsterdam, 2004, pp. 2341 – 2378.
- [11] M. Jefferson, The law of the primate city, *Geographical Review* 29(2) (1939) 226 – 232.
- [12] R. L. Moomaw, A. M. Shatter, Urbanization and economic development: A bias toward large cities?, *Journal of Urban Economics* 40(1) (1996) 13 – 37.
- [13] P. Krugman, Increasing returns and economic geography, *Journal of Political Economy* 99(3) (1991) 483 – 499.
- [14] J. V. Henderson, A. Storeygard, D. N. Weil, Measuring economic growth from outer space, *American Economic Review* 102(2) (2012) 994 – 1028.
- [15] X. Chen, W. D. Nordhaus, Using luminosity data as a proxy for economic statistics, *Proceedings of the National Academy of Sciences* 108(21) (2011) 8589 – 8594.
- [16] M. Pesaresi, M. Schiavina, P. Politis, S. Freire, K. Krasnodebska, J. H. Uhl, A. Carioli, C. Corbane, L. Dijkstra, P. Florio, H. K. Friedrich, J. Gao, S. Leyk, L. Lu, L. Maffenini, I. Mari-Rivero, M. Melchiorri, V. Syrris, J. Van Den Hoek, T. Kemper, Advances on the Global Human Settlement Layer by joint assessment of Earth observation and population survey data, *International Journal of Digital Earth* 17(1) (2024) 2390454.
- [17] K. Faisal, A. Shaker, S. Habbani, Modelling the relationship between the gross domestic product and built-up area using remote sensing and GIS data: A case study of seven major cities in Canada, *ISPRS International Journal of Geo-Information* 5(3) (2016) 23.

- [18] S. Nickell, Biases in dynamic models with fixed effects, *Econometrica* 49(6) (1981) 1417 – 1426.
- [19] M. Arellano, S. Bond, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies* 58(2) (1991) 277 – 297.
- [20] R. Blundell, S. Bond, Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics* 87(1) (1998) 115 – 143.
- [21] [dataset] M. Pesaresi, P. Politis, GHS-BUILT-S R2023A – GHS built-up surface grid, derived from Sentinel2 composite and Landsat, multitemporal (1975–2030), European Commission, Joint Research Centre, 2023.
- [22] [dataset] A. J. Florczyk, C. Corbane, D. Ehrlich, S. Freire, T. Kemper, L. Maffenini, M. Melchiorri, M. Pesaresi, P. Politis, M. Schiavina, V. Syrris, L. Zanchetta, GHSL Data Package 2019, Publications Office of the European Union, Luxembourg, 2019.
- [23] R. Gibrat, *Les Inégalités Économiques*, Recueil Sirey, Paris, 1931.
- [24] J. Eeckhout, Gibrat's Law for (all) cities, *American Economic Review* 94(5) (2004) 1429 – 1451.
- [25] R. González-Val, L. Lanaspa, F. Sanz-Gracia, New evidence on Gibrat's Law for cities, *Urban Studies* 51(1) (2014) 93 – 115.
- [26] R. J. Barro, X. Sala-i-Martin, Convergence, *Journal of Political Economy* 100(2) (1992) 223 – 251.
- [27] D. Roodman, How to do xtabond2: An introduction to difference and system GMM in Stata, *Stata Journal* 9(1) (2009) 86 – 136.
- [28] L. P. Hansen, Large sample properties of generalized method of moments estimators, *Econometrica* 50(4) (1982) 1029 – 1054.
- [29] D. Roodman, A note on the theme of too many instruments, *Oxford Bulletin of Economics and Statistics* 71(1) (2009) 135 – 158.
- [30] Y. M. Ioannides, H. G. Overman, Zipf's law for cities: An empirical examination, *Regional Science and Urban Economics* 33(2) (2003) 127 – 137.
- [31] M. Bosker, S. Brakman, H. Garretsen, M. Schramm, A century of shocks: The evolution of the German city size distribution 1925–1999, *Regional Science and Urban Economics* 38(4) (2008) 330 – 347.
- [32] V. Nitsch, Zipf zipped, *Journal of Urban Economics* 57(1) (2005) 86 – 100.
- [33] K. T. Soo, Zipf's law for cities: A cross-country investigation, *Regional Science and Urban Economics* 35(3) (2005) 239 – 263.
- [34] M. Brülhart, F. Sbergami, Agglomeration and growth: Cross-country evidence, *Journal of Urban Economics* 65(1) (2009) 48 – 63.
- [35] M. Resende, Gibrat's Law and the growth of cities in Brazil: A panel data investigation, *Urban Studies* 41(8) (2004) 1537 – 1549.
- [36] N. Islam, What have we learnt from the convergence debate?, *Journal of Economic Surveys* 17(3) (2003) 309 – 362.
- [37] J. C. Córdoba, On the distribution of city sizes, *Journal of Urban Economics* 63(1) (2008) 177 – 197.
- [38] S. A. Frick, A. Rodríguez-Pose, Big or small cities? On city size and economic growth, *Growth and Change* 49(1) (2018) 4 – 32.
- [39] P. C. Melo, D. J. Graham, R. B. Noland, A meta-analysis of estimates of urban agglomeration economies, *Regional Science and Urban Economics* 39(3) (2009) 332 – 342.
- [40] K. Desmet, J. Rappaport, The settlement of the United States, 1800–2000: The long transition towards Gibrat's law, *Journal of Urban Economics* 98 (2017) 50 – 68.
- [41] K. Giesen, J. Südekum, Zipf's law for cities in the regions and the country, *Journal of Economic Geography* 11(4) (2011) 667 – 686.
- [42] K. Cheng, Y. Zhuang, Spatial econometric analysis of the rank-size rule for urban system: A case of prefectural-level cities in China's middle area, *Scientia Geographica Sinica* 32(8) (2012) 903 – 912.
- [43] T. Potipiti, Output convergence among provinces in Thailand, *Thailand and the World Economy* 27(4) (2018) 1 – 21.
- [44] N. W. Ismail, Growth and convergence in ASEAN: A dynamic panel approach, *International Journal of Economics and Management* 2(1) (2008) 127 – 140.
- [45] F. Bickenbach, E. Bode, P. Nunnenkamp, M. Söder, Night lights and regional GDP, *Review of World Economics* 152(2) (2016) 425 – 447.
- [46] X. S. Pérez-Sindín, T. H. K. Chen, A. V. Prishchepov, Are night-time lights a good proxy of economic activity in rural areas in middle and low-income countries? Examining the empirical evidence from Colombia, *Remote Sensing Applications: Society and Environment* 24 (2021) 100647.