ISSN: 3008-1629

Regional Path Differentiation and Convergence in the Massification of Higher Education in China Mainland: A Dual Framework of Quantitative Expansion and Structural Upgrading

Gan Weixiang[1], Xiao Mengfei[2*], Zhang Naiqian[3], Tara Ahmed Mohammed[4], Zhang Yongli[5]

[1]Zhengguigu Technology Co., Ltd., Guangyang District, Langfang City, 065000, Hebei Province, China

[2]SEGi University, Petaling Jaya, 47810, Selangor, Malaysia*

[3]Chongqing University of Arts and Sciences, Chongqing 402160, China

[4,5]Graduate School of Business, SEGi University, Petaling Jaya, 47810, Selangor, Malaysia

Corresponding Author: 282584787@qq.com

Abstract

Existing scholarship on the massification of higher education in China has typically relied on single indicators or composite indices to capture educational expansion, focusing primarily on how rising enrolment rates or increasing average years of schooling shape economic growth and regional inequality. Yet it has seldom distinguished between two qualitatively different processes, namely improvements in overall educational attainment and the structural upgrading of higher education, which limits our ability to identify concrete regional trajectories of expansion. To address this gap, this study employs panel data for thirty one provincial level regions in mainland China from 1997 to 2024, measuring attainment improvement by the cumulative change in average years of schooling and structural upgrading by the cumulative change in the share of the population holding a college degree or above. Using the full sample median as the threshold, a four quadrant framework is constructed to classify provincial expansion paths. The results show that China's higher education massification does not exhibit a single convergence pattern; instead, it forms four relatively stable trajectories, including the dual track leap type, the mass expansion type, the structural upgrading type, and the slow expansion type. Overall educational attainment and higher education structural upgrading frequently evolve at different speeds across provinces: some regions achieve coordinated progress on both dimensions, some expand broadly before upgrading, some advance structure ahead of widespread attainment gains, while others remain relatively constrained on both dimensions. These findings indicate that the most consequential differences in educational expansion lie not simply in the magnitude of improvement, but in how quantitative expansion and structural upgrading are combined. By shifting the analytical focus from scale based comparisons to trajectory identification, this study offers a clearer framework for understanding regional heterogeneity in human capital accumulation and its linkages with industrial restructuring and population mobility.

Keywords

Massification of higher education, educational expansion typology, regional disparity, path differentiation, convergence

ISSN: 3008-1629

1.Introduction

1.1 Research Background

Since the reform and opening up period, China's education system has undergone a sustained process of expansion and universalization. The compulsory education system has been progressively consolidated, secondary education has steadily expanded, and higher education has entered a stage of mass development under the combined impetus of enrolment enlargement and institutional reform (Jiang & Ke, 2021). From a macro perspective, educational development exhibits a general upward trajectory that extends from foundational access toward higher levels of attainment (Song & Xu, 2024). Average years of schooling have continued to rise, and the coverage of educational opportunities has expanded substantially (Mathew & Martin, 2025). However, this overall upward movement does not necessarily imply that higher education expansion proceeds in synchrony with general improvements in educational attainment (Cunningham & Samson, 2021). Table 1 indicates that since 1997 the average years of schooling across regions have followed a sustained upward trend, with most provinces displaying trajectories that move in a consistent direction. This pattern suggests that educational expansion along the quantitative dimension has been broadly diffused, primarily reflecting the widening of educational access and the extension of schooling duration, both of which contribute to improvements in aggregate educational attainment. In contrast, Table 2 reveals pronounced regional heterogeneity. The speed, magnitude, and temporal rhythm of growth in the proportion of the population holding a college degree or above vary substantially across provinces. Municipalities directly under the central government and several high capacity regions demonstrate earlier, faster, and steeper trajectories of structural elevation. By comparison, many provinces that exhibit steady increases in average years of schooling in Table 1 show relatively modest or gradual gains in the share of highly educated population in Table 2. This contrast clearly demonstrates that increases in average years of schooling cannot be equated with improvements in the proportion of higher education attainment. Incremental educational gains do not necessarily accumulate at the tertiary level; rather, they may be concentrated within primary and secondary education, generating broader coverage without a corresponding elevation in educational hierarchy. The juxtaposition of Table 1 and Table 2 therefore highlights the empirical distinction between quantitative expansion and structural upgrading within the broader process of educational development. The former primarily reflects the diffusion of educational participation, whereas the latter captures vertical advancement within the

ISSN: 3008-1629

educational structure. Although both dimensions may display parallel growth at the national level, they frequently evolve asynchronously or even divergently across provinces. It is therefore necessary to more precisely uncover the nonlinear evolution and differentiated regional trajectories that characterize the massification of higher education in China.

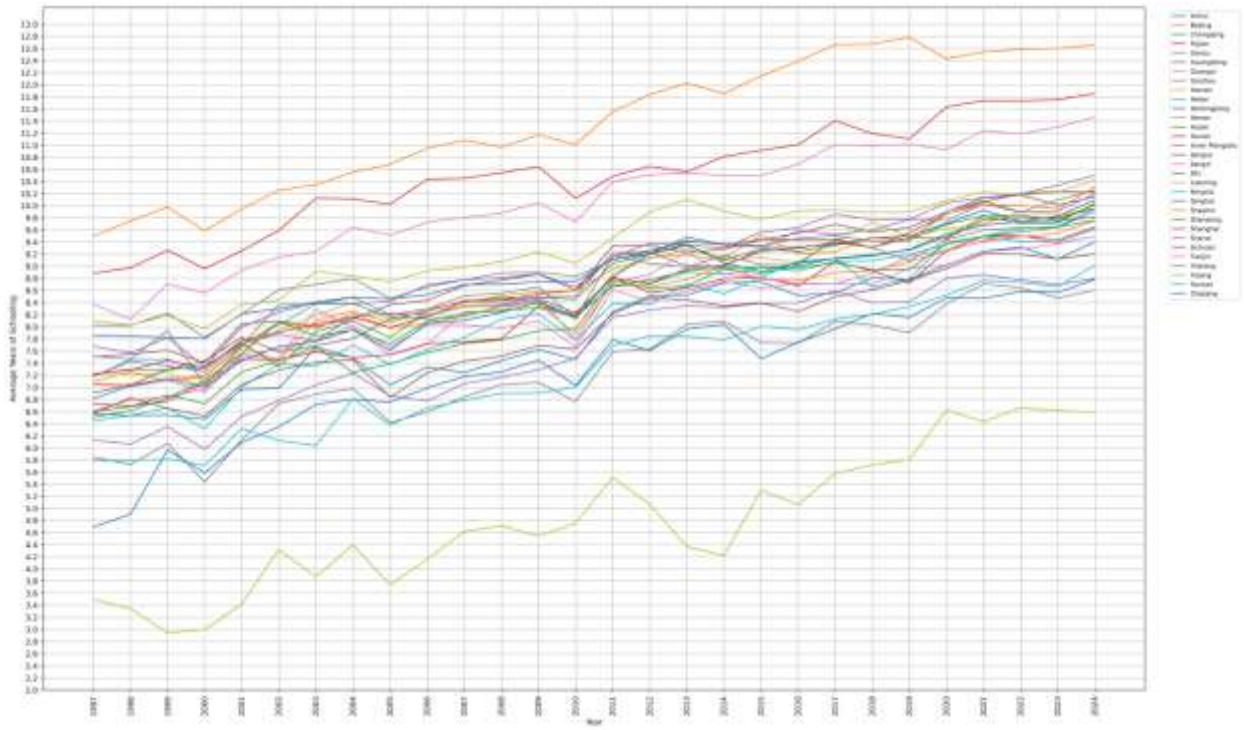
Table 1 Average Years of Schooling Across Provinces (1997–2024)

Province	1997	2001	2005	2009	2013	2017	2021	2024
Beijing	9.5013	9.948	10.6858	11.1726	12.0284	12.6651	12.543	12.6538
Shanghai	8.8851	9.2569	10.0257	10.6469	10.5615	11.4061	11.7377	11.8562
Tianjin	8.3756	8.9388	9.5129	10.0515	10.5386	11.0099	11.2412	11.4658
Chongqing	6.601	7.2575	7.3918	7.9341	8.676	9.136	9.8258	10.0203
Hebei	7.174	7.7206	8.1695	8.4249	8.9017	9.0858	9.4984	9.9567
Shanxi	7.6823	7.9969	8.4171	8.8764	9.3566	9.8621	10.1368	10.5041
Inner Mongolia	7.1816	7.7305	8.2234	8.4946	9.0102	9.5245	10.0304	10.1955
Liaoning	8.1017	8.3705	8.7461	9.2373	10.1049	9.9279	10.2315	10.431
Jilin	8.0274	8.2103	8.4683	8.9028	9.4033	9.5062	10.0781	10.2561
Heilongjiang	7.8562	8.2172	8.46	8.7457	9.4806	9.3632	9.9269	10.0949
Jiangsu	6.9104	7.829	8.134	8.5461	9.4223	9.439	10.061	10.2227
Zhejiang	6.8148	7.4429	7.6137	8.4043	9.3675	9.1275	9.6799	9.8108
Anhui	6.5571	6.9678	7.0386	7.6213	8.5245	8.5577	9.2466	9.4051
Fujian	6.7342	7.4709	7.543	8.3464	8.6471	9.0826	9.4976	9.6471
Jiangxi	7.0467	7.5326	7.5313	8.5229	9.2385	8.7164	9.4292	9.4968
Shandong	6.5031	7.5603	7.7219	8.3128	8.9248	9.0614	9.4931	9.7586
Henan	7.1014	7.6986	7.9855	8.3874	8.7833	8.8942	9.413	9.7561
Hubei	7.2215	7.7407	7.8223	8.4887	9.3448	9.3457	9.8388	10.0422
Hunan	7.2243	7.7768	7.9909	8.4653	8.9576	9.3958	9.7708	9.9565
Guangdong	7.5044	8.0515	8.365	8.8717	9.2265	9.6969	10.1287	10.1459
Guangxi	6.613	7.5534	7.6595	8.0957	8.5933	8.7144	9.2374	9.4962
Hainan	7.2121	7.6533	8.1086	8.4369	9.1877	9.4158	9.6974	9.95
Sichuan	6.5695	7.0482	6.8372	7.694	8.4475	8.4976	9.2086	9.2017
Guizhou	5.8494	6.1338	6.4182	7.0815	8.0427	8.0933	8.7123	8.6069
Yunnan	5.7889	6.3202	6.3778	6.9053	7.8422	8.1267	8.765	9.0191
Shaanxi	7.0661	7.6875	8.0624	8.5845	9.283	9.2424	10.0187	10.3126
Gansu	6.1273	6.5202	6.86	7.2911	8.3465	8.5972	8.8625	8.7975
Qinghai	4.6926	6.0914	6.7581	7.4479	7.9641	7.9674	8.4689	8.7829

ISSN: 3008-1629

Ningxia	6.4475	7.0043	7.3749	8.2166	8.7074	9.128	9.4903	9.6184
Xinjiang	7.512	7.6865	8.2033	8.658	8.9875	9.458	9.6	9.909
Xizang	3.4998	3.4199	3.7384	4.5478	4.3674	5.577	6.4362	6.5889

Graphic 1. Trend of Average Years of Schooling across Provinces and Municipalities in Mainland China (Range from 1997-2024)



Note*: The data used in this study are obtained from the CNPD database.

Table 2 Provincial Higher Education Structure (1997-2024)

Province	1997	2001	2005	2009	2013	2017	2021	2024
Beijing	0.1344	0.1754	0.2449	0.3077	0.4121	0.4761	0.4914	0.5083
Tianjin	0.0727	0.0944	0.1408	0.1701	0.2305	0.2885	0.3295	0.3616
Hebei	0.021	0.0289	0.0473	0.0562	0.0773	0.0999	0.1417	0.1902
Shanxi	0.0336	0.0374	0.0557	0.0766	0.1072	0.1492	0.2011	0.2449
Inner Mongolia	0.036	0.0408	0.0793	0.0795	0.1008	0.1813	0.2282	0.2555
Liaoning	0.0602	0.0652	0.0834	0.1182	0.1983	0.1734	0.2127	0.2403
Jilin	0.0499	0.0519	0.0668	0.0822	0.1156	0.1482	0.2084	0.2365

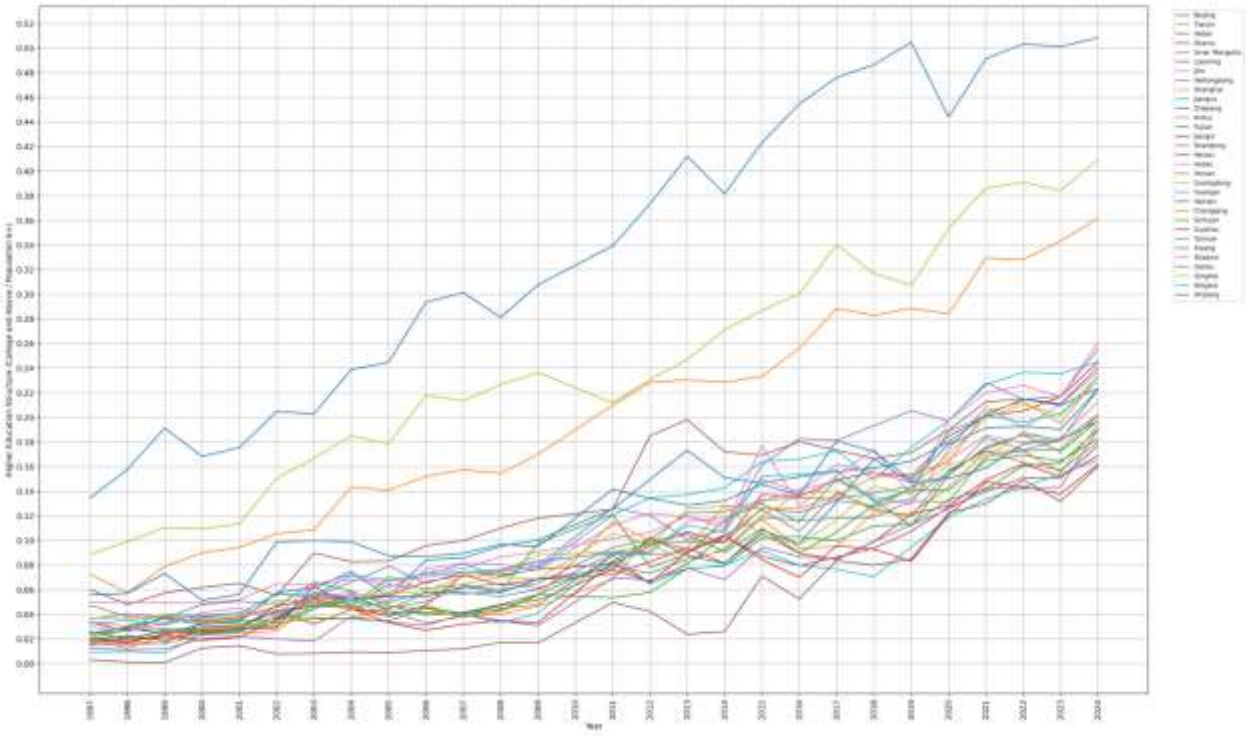
ISSN: 3008-1629

Heilongjiang	0.0475	0.0507	0.0642	0.0655	0.1224	0.1338	0.1838	0.2228
Shanghai	0.0889	0.1136	0.1784	0.2366	0.2469	0.3403	0.3865	0.4093
Jiangsu	0.0205	0.0414	0.068	0.0776	0.1373	0.1729	0.2266	0.2453
Zhejiang	0.025	0.034	0.0542	0.1004	0.1733	0.1558	0.202	0.2233
Anhui	0.0183	0.0249	0.0385	0.0466	0.0915	0.0952	0.1725	0.1939
Fujian	0.0263	0.0316	0.0498	0.098	0.0889	0.1508	0.176	0.1974
Jiangxi	0.0188	0.0282	0.0385	0.0686	0.0939	0.0843	0.1492	0.169
Shandong	0.0149	0.0355	0.0444	0.0601	0.0989	0.1326	0.1741	0.2011
Henan	0.016	0.0288	0.0422	0.0516	0.0809	0.0866	0.1397	0.1767
Hubei	0.0339	0.0412	0.0507	0.0763	0.1192	0.162	0.1853	0.2124
Hunan	0.0202	0.0312	0.0449	0.0614	0.085	0.1181	0.1685	0.2016
Guangdong	0.0367	0.0388	0.0581	0.0687	0.082	0.1404	0.1994	0.2287
Guangxi	0.0093	0.0259	0.04	0.041	0.077	0.0765	0.1339	0.1651
Hainan	0.0232	0.0349	0.0545	0.0688	0.0878	0.1312	0.1588	0.1971
Chongqing	0.0197	0.0304	0.0463	0.0549	0.0936	0.1365	0.2042	0.2336
Sichuan	0.0199	0.0268	0.0348	0.0562	0.1054	0.1069	0.1649	0.1827
Guizhou	0.0201	0.0216	0.0332	0.0331	0.0909	0.0957	0.1475	0.1618
Yunnan	0.0124	0.0223	0.0337	0.0306	0.0776	0.0851	0.1431	0.1901
Xizang	0.0032	0.0145	0.0089	0.0168	0.0239	0.0837	0.1299	0.1603
Shaanxi	0.0299	0.045	0.0617	0.091	0.1199	0.1528	0.2201	0.2612
Gansu	0.0165	0.0292	0.0426	0.0479	0.0902	0.139	0.1726	0.1799
Qinghai	0.0201	0.0359	0.0711	0.0884	0.1257	0.1165	0.1672	0.2024
Ningxia	0.0327	0.0411	0.0682	0.0836	0.1125	0.1572	0.2017	0.2328
Xinjiang	0.0558	0.0563	0.0875	0.0951	0.1285	0.1809	0.1916	0.2227

Data Source*: The data used in this study are obtained from the CNPD database.

ISSN: 3008-1629

Graphic 2. Trend of Educational Structure across Provinces and Municipalities in Mainland China
(Proportion of Population with Diploma or Above 1997-2024)



Note*: The data used in this study are obtained from the CNPD database.

1.2 Research Objectives

Provincial educational development does not proceed along a single linear trajectory; rather, it evolves simultaneously along two dimensions, namely the improvement of educational attainment and the deepening of structural upgrading, and these processes frequently unfold at different paces across regions (Mathew & Martin, 2025). When educational development is assessed solely through a single indicator such as average years of schooling or the proportion of the population with higher education, fundamentally distinct developmental pathways are easily reduced to simplified comparisons, thereby obscuring the stage specific differences and structural differentiation that characterize provincial expansion processes (Jiang & Ke, 2021). For instance, some regions may complete the universalization of basic education at an early stage while lagging in higher education development. Others may achieve rapid expansion in tertiary education while registering only modest gains in overall years of schooling. Still others may advance in both broad access and structural upgrading in a relatively coordinated manner (Song & Xu, 2024). A typological approach to educational expansion is therefore essential for identifying how different regions accumulate human capital and optimize educational hierarchy. Such an approach provides a more realistic and

ISSN: 3008-1629

analytically robust framework for understanding regional development capacity, industrial transformation potential, and the effectiveness of education policy implementation, while also offering an empirical foundation for more targeted regional education strategies (Sims et al., 2025). At a deeper level, the separation between quantitative expansion and structural upgrading at the regional scale is not incidental but arises from the interaction of multiple mechanisms. From the supply side, the regional carrying capacity of higher education resources, including institutional distribution, enrolment scale, concentration of high quality resources, and the intensity of local fiscal support, determines whether newly educated cohorts can effectively transition into higher education (Hu & Cao, 2023). From the demand side, the configuration of industrial structure and skill demand shapes both the returns to highly educated labour and the absorptive capacity of local economies. Manufacturing based or resource dependent regions often prioritize the expansion of secondary education to match prevailing labour market needs, whereas regions dominated by knowledge intensive industries and modern services are more likely to generate endogenous incentives for sustained increases in the share of higher education attainment (Song & Xu, 2024). In spatial terms, highly educated populations tend to exhibit stronger interregional mobility. Even where local higher education provision expands, outward migration of degree holders may dilute the statistical share of highly educated population, while core metropolitan areas may achieve rapid structural upgrading through net talent inflows (Mathew & Martin, 2025). It is through the interaction of supply mechanisms, demand structures, and population mobility dynamics that educational expansion assumes a dual track pattern in which quantitative diffusion and structural deepening operate along partially distinct trajectories. These two dimensions may advance in coordination, but they may also diverge at particular stages of development (Cunningham & Samson, 2021). On this basis, the present study situates provincial educational development within an analytical framework defined by quantitative expansion and structural upgrading, and constructs a typology comprising four representative paths: a coordinated advancement type in which both dimensions progress simultaneously; a quantitatively driven type characterized by substantial expansion with limited structural upgrading; a structurally leading type in which upgrading outpaces overall attainment gains; and a relatively constrained type in which both dimensions advance slowly. This classification provides a clear theoretical foundation and empirical basis for the subsequent analysis of path differentiation and convergence testing.

ISSN: 3008-1629

1.3 Research Significance

At the theoretical level, this study draws upon provincial and municipal data on average years of schooling and educational attainment structure in mainland China from 1997 to 2024 to identify empirically grounded differences in educational development. By explicitly distinguishing between quantitative expansion and structural upgrading, it demonstrates that increases in average years of schooling do not necessarily correspond to proportional growth in the share of the population with college education or above. In doing so, the study moves beyond interpretative frameworks that equate educational development with improvements in a single aggregate indicator. By conceptualizing educational diffusion and higher education deepening as two evolutionary trajectories that may either advance in coordination or diverge over time, the analysis more precisely captures the nonlinear and path dependent character of higher education massification at the regional level. This perspective provides a more analytically robust foundation for understanding interprovincial differences in human capital accumulation and for examining how such differences relate to regional development capacity and industrial transformation potential. At the methodological and practical levels, the construction of a dual dimensional framework encompassing quantitative expansion and structural upgrading, together with a systematic typological classification, enables a clearer representation of the stage specific characteristics and structural differentiation of provincial educational expansion. Compared with approaches based solely on indicator rankings or mean comparisons, this framework offers more explicit analytical units for subsequent examination of path differentiation and convergence dynamics. Furthermore, the resulting classification yields a more precise diagnostic logic for regional education policy and resource allocation. In regions where quantitative expansion proceeds rapidly while structural upgrading lags behind, policy priorities should shift toward strengthening the capacity and quality of higher education provision. In contrast, in regions where structural upgrading advances more rapidly than broad based educational attainment, greater attention should be devoted to reinforcing equitable development and foundational support within primary and secondary education. Through this differentiated policy orientation, educational intervention can achieve greater precision and effectiveness.

2 Literature Review

2.1 Current State of Research

First, theoretical explanations of higher education expansion are commonly anchored in the framework of massification. The intellectual foundation of this perspective can be traced to

ISSN: 3008-1629

Trow's stage theory, which conceptualizes the evolution of higher education from elite to mass and subsequently to universal participation. Trow argues that once participation rates surpass certain thresholds, profound transformations occur in institutional structures, organizational forms, and social functions (Trow, 1973). In the Chinese context, the large scale enrolment expansion initiated around 1999 is widely regarded as a critical institutional turning point. Wu and Zhang interpret this reform from the perspective of the interaction between macroeconomic cycles and policy adjustment, suggesting that the expansion simultaneously served the purposes of educational reform and macroeconomic regulation (Wu & Zhang, 2010). Li, Whalley, Zhang, and Zhao examine the reform through the lens of labour market restructuring and factor composition, demonstrating that expansion significantly altered the supply structure of highly educated labour and exerted sustained effects on returns to education and income distribution (Li et al., 2011). As participation levels continued to rise, Mok emphasizes that massification in East Asia is frequently accompanied by intensified institutional stratification and increasingly complex governance arrangements (Mok, 2016). Marginson further argues that in high participation systems, expansion and stratification coexist in structurally embedded and enduring forms (Marginson, 2016). Second, research on regional disparities and spatial differentiation in educational development constitutes an important strand for understanding the uneven character of expansion in China. Zhang and Kanbur propose that regional inequality in China is institutionally embedded, closely linked to the allocation of public resources and development structures (Zhang & Kanbur, 2005). From a human capital perspective, Fleisher, Li, and Zhao provide empirical evidence that regional differences in educational attainment are significantly associated with disparities in economic growth (Fleisher et al., 2010). At the methodological level, Han and Liu, among others, apply spatial econometric techniques and regional indicators to reveal pronounced spatial clustering and concentration of higher education advantages in eastern China (Han et al., 2023). Guo and colleagues, using composite indices and interregional comparisons, find that although provincial educational gaps exhibit dynamic adjustment, structural differences remain significant and persistent (Guo et al., 2024). Overall, this body of literature effectively identifies where disparities exist and how they are spatially manifested, yet it offers limited direct examination of the internal composition and typological pathways of regional educational expansion. Third, studies on the socioeconomic consequences of educational expansion are predominantly grounded in human capital and growth theories, with a focus on macro level outcomes. Fleisher, Li, and Zhao demonstrate that human capital accumulation significantly promotes regional economic

ISSN: 3008-1629

growth and interacts with the evolution of regional inequality (Fleisher et al., 2010). Li and coauthors further argue that higher education expansion reshapes wage structures and educational returns by altering the composition of labour supply, thereby exerting structural effects on income distribution (Li et al., 2011). Using provincial panel data, Qi and colleagues show that higher education expansion has a positive long term effect on economic growth, although the magnitude and conditions of this effect vary substantially across regions (Qi et al., 2022). From a political economy perspective, Marginson cautions that massification does not necessarily enhance social mobility, as expansion may coincide with new forms of institutional stratification (Marginson, 2016). Collectively, this strand of research tends to treat expansion as a homogeneous explanatory variable within empirical models, thereby overlooking the possibility that internal structural differences within expansion processes may generate heterogeneous outcomes. Fourth, with regard to measurement strategies, the literature has developed several representative indicators and empirical approaches. The global educational attainment database constructed by Barro and Lee provides a standardized measure of average years of schooling and has been widely applied in research on human capital and macroeconomic growth (Barro & Lee, 2013). In studies of massification, Mok employs participation related indicators to characterize expansion trajectories and underscores the parallel evolution of participation growth and governance transformation (Mok, 2016). Guo and colleagues construct a composite index based on weighted multidimensional indicators to assess regional educational development and analyze evolving disparities (Guo et al., 2024). While these approaches offer operational clarity and cross study comparability, they generally rely on single axis measurement or aggregated indices. Such strategies make it difficult to conceptually and empirically distinguish between improvements in educational diffusion and upgrading in higher education structure, and they are less capable of capturing potential asynchrony and combinational variation between these dimensions at the regional level (Barro & Lee, 2013; Mok, 2016). Consequently, if the objective is to identify typological differences and structural configurations in provincial educational expansion, a measurement framework with stronger capacity for structural decomposition and dual dimensional analysis is required.

2.2 Research Contributions

A synthesis of the existing literature on higher education massification, regional educational disparities, and the socioeconomic consequences of expansion reveals that although current scholarship has developed relatively mature analytical frameworks in relation to institutional

ISSN: 3008-1629

evolution, spatial distribution, and growth effects, educational expansion continues to be treated primarily as a single scale variable or as an aggregated composite index. Conceptual distinctions between internal structural dimensions of expansion remain underdeveloped, and there has been limited effort to identify typological development paths at the provincial level. Against this backdrop, the primary contribution of the present study lies in its explicit conceptual decomposition of educational expansion into two relatively independent yet interrelated dimensions, namely improvement in educational diffusion and upgrading in higher education structure. Methodologically, the study constructs cumulative change indicators based on long term provincial panel data and employs a dual dimensional classification framework to identify distinct regional expansion trajectories. In doing so, it advances the analytical paradigm from comparisons of developmental levels to the identification of developmental pathways. This path oriented typology addresses the limitations of single axis measurement approaches that struggle to capture asynchronous development processes. It also establishes a structured grouping foundation for subsequent convergence testing and mechanism analysis. By moving beyond scale based discussion toward a systematic examination of typological structure and dynamic evolution, the study deepens the analytical scope of educational expansion research and enhances its explanatory capacity.

3. Research Methodology

The provincial educational data employed in this study are drawn from authoritative statistical sources published by the National Bureau of Statistics of China, primarily the China Statistical Yearbook, and are cross validated with corresponding provincial statistical yearbooks. Average Years of Schooling, abbreviated as AYS, is calculated on the basis of the population distribution by educational attainment. The Higher Education Share, abbreviated as HES, is constructed as the proportion of individuals holding a college degree or above relative to the permanent population aged six and above in each province. All data are standardized according to the current provincial administrative divisions to ensure interregional comparability. Given that consistent statistical reporting for higher education structural data has been available since 1997, this study designates 1997 as the base year and 2024 as the terminal year. A reproducible panel dataset covering thirty one provincial level administrative units in mainland China is constructed to support subsequent identification and classification of educational expansion trajectories.

ISSN: 3008-1629

3.1 Variable Definition and Measurement Framework

Let i denote province and t denote year. This study characterizes provincial educational development in China along two analytical dimensions, namely quantitative educational expansion and structural upgrading of educational attainment. First, Average Years of Schooling is denoted as AYS_{it}, which represents the mean years of formal schooling among residents in province i in year t. This indicator captures the extent of educational diffusion and corresponds to the quantitative dimension of educational development. Second, educational structure is operationalized using the Higher Education Share as the indicator of structural upgrading. The Higher Education Share is defined as follows: $HES_{it}=\frac{N_{it}^{College+}}{Pop_{it}^{6+}}$. In this definition, $N_{it}^{College+}$ refers to the number of individuals in province i in year t who have attained college level education or above, and Pop_{it}^{6+} denotes the total permanent population aged six and above in the same province and year. This measure reflects the extent to which higher education attainment has increased within the overall population and serves as the structural dimension of educational development.

3.2 Research Time Scope

As data on educational structure, measured by the Higher Education Share, are consistently available beginning in 1997, this study adopts a unified temporal framework to ensure data comparability. The base year is designated as t_0=1997, and the terminal year as t_1=2024, corresponding to 1997 and 2024 respectively. This treatment avoids potential discontinuities in the dataset and mitigates comparative bias that may arise from differing starting points across indicators, thereby ensuring temporal consistency in subsequent analysis.

3.3 Construction of Educational Expansion Indices

Given that growth rates are highly sensitive to low initial baselines, particularly in the case of the Higher Education Share, this study adopts absolute change rather than growth rates to capture long term trajectories of educational development. This approach allows for a more stable and comparable assessment of cumulative expansion across provinces. Accordingly, the index of quantitative educational expansion is defined as $\Delta AYS_i = AYS_{i,2024} - AYS_{i,1997}$, which represents the cumulative increase in educational attainment over the study period. Similarly, the index of structural upgrading in higher education is defined as $\Delta HES_i = HES_{i,2024} - HES_{i,1997}$which reflects the magnitude of change in the proportion of the population holding college level education or above between the base year and the terminal year.

ISSN: 3008-1629

3.4 Threshold Determination

To reduce the influence of extreme values and enhance replicability, this study employs the full sample median as the classification threshold.

The calculation procedure is defined as follows:

$\widetilde{\Delta AYS} = \text{median}(\Delta AYS_i), \quad i = 1, \ldots, 31$ and $\widetilde{\Delta HES} = \text{median}(\Delta HES_i), \quad i = 1, \ldots, 31$, where the median values are computed across all thirty one provincial level units.

Based on the original data, the median value of the quantitative expansion index is 2.8480304, and the median value of the structural upgrading index is 0.1800612.

3.5 Classification of Educational Expansion Pathways

Based on the relative positions of ΔAYS_i and ΔHES_i, a two dimensional four quadrant classification framework is constructed. The first category is defined as the dual track leap type, corresponding to $\Delta AYS_i \geq \widetilde{\Delta AYS} \wedge \Delta HES_i \geq \widetilde{\Delta HES}$, which indicates that educational diffusion and structural upgrading advance simultaneously. The second category is the mass expansion type, defined as: $\Delta AYS_i \geq \widetilde{\Delta AYS} \wedge \Delta HES_i < \widetilde{\Delta HES}$, which reflects substantial growth in overall educational attainment while the increase in the share of higher education remains comparatively limited. The third category is the structural upgrading type, defined as: $\Delta AYS_i < \widetilde{\Delta AYS} \wedge \Delta HES_i \geq \widetilde{\Delta HES}$, which indicates relatively rapid development in higher education accompanied by more modest gains in overall years of schooling. The fourth category is the slow expansion type, defined as: $\Delta AYS_i < \widetilde{\Delta AYS} \wedge \Delta HES_i < \widetilde{\Delta HES}$, representing provinces in which both quantitative expansion and structural upgrading progress at relatively slower rates. This classification is grounded in the recognition that provincial educational expansion does not follow a single linear trajectory. Rather, it typically unfolds along two partially independent tracks, namely educational diffusion in terms of quantitative attainment and structural upgrading in terms of higher education deepening. Reliance on a single indicator, such as average years of schooling or the share of higher education attainment, risks obscuring structural differences across provinces, including patterns characterized by sequential diffusion followed by upgrading, quality led transformation, or coordinated dual advancement. By adopting a four quadrant framework, the analysis captures the combinational characteristics of quantitative expansion and structural upgrading within a unified analytical space. This approach enables a clear typological distinction among provinces, enhances the explanatory power of classification, strengthens the replicability of regional comparison, and provides a structured analytical foundation for subsequent mechanism testing and policy interpretation.

ISSN: 3008-1629

3.6 Data Processing Procedure

Prior to formal computation, this study first clarifies the construction logic of the educational expansion indicators and, on this basis, implements the full data processing and typological identification procedure using Stata. Specifically, educational expansion is decomposed into two core dimensions. The first dimension concerns educational diffusion, measured by average years of schooling as a representative indicator of overall educational attainment. Values for 1997 and 2024 are extracted and the difference between the two periods is calculated to obtain a long term cumulative quantitative expansion index, which captures the magnitude of improvement in overall educational attainment during the study period. The second dimension concerns structural upgrading in higher education, measured by the proportion of the population holding a college degree or above as a representative indicator of structural deepening. Values for 1997 and 2024 are similarly extracted and differenced to construct a structural upgrading index, which reflects the extent to which higher education has increased within the overall educational composition of the population. Both indicators are based on cumulative change rather than growth rates in order to reduce distortions arising from baseline effects and short term fluctuations, thereby enabling more accurate cross provincial comparison of long term development trajectories. From a technical perspective, the entire computational procedure is executed through Stata programming to ensure transparency and replicability. The process begins with importing the original Excel dataset and standardizing variable names and data types to ensure that year and indicator variables are properly formatted as numeric values. Data for 1997 and 2024 are then extracted separately, and a one to one provincial matching procedure is performed using the merge command to generate difference variables for quantitative expansion and structural upgrading. After integrating the two dimensions into a unified dataset, the summarize, detail command is employed to calculate the median values of the two indicators, which are stored as scalar thresholds for classification. Finally, conditional statements are used to construct the four quadrant typology variable, assigning each province to a specific expansion pathway based on its relative position to the median values along both dimensions. This programming workflow not only enhances efficiency and precision in data processing but also ensures transparency and reproducibility in typology construction, thereby providing a stable empirical foundation for subsequent convergence testing and mechanism analysis.

ISSN: 3008-1629

Table 4 Stata Data Analysis Source Code

```
************************************************************
* Provincial Education Expansion Typology Construction
* Replication-ready Stata Master Script (All-in-One)
************************************************************

clear all
set more off

*------------------------------------------------
* Step 0: Define file path
*------------------------------------------------
global data "/mnt/data/Provincial_Education_Data.xlsx"

************************************************************
* STEP 1: LOAD AND CLEAN AYS DATA (LONG FORMAT)
* Expected variables: city | year | AYS
************************************************************
import excel "$data", sheet("Average Years of Schooling") firstrow clear

* Ensure consistent variable names (adjust if needed)
rename (city year AYS) (city year AYS)

* Drop empty rows and any duplicated header row accidentally read as data
drop if missing(city)
drop if city == "city"

* Enforce numeric types
destring year, replace
destring AYS, replace

************************************************************
* STEP 2: EXTRACT AYS FOR 1997 AND 2024
************************************************************
preserve
keep if year == 1997
rename AYS AYS1997
keep city AYS1997
tempfile ays97
save `ays97'
restore

keep if year == 2024
rename AYS AYS2024
```

ISSN: 3008-1629

```
keep city AYS2024
tempfile ays24
save `ays24'

*********************************************************
* STEP 3: COMPUTE QUANTITY EXPANSION (ΔAYS)
*********************************************************
use `ays97', clear
merge 1:1 city using `ays24'
drop _merge

gen dAYS = AYS2024 - AYS1997

tempfile ays_delta
save `ays_delta'

*********************************************************
* STEP 4: LOAD AND CLEAN HES DATA (WIDE FORMAT)
* Expected: city in first column; year columns include 1997 and 2024
*********************************************************
import excel "$data", sheet("Higher Education Structure") clear

* Assume the first column is the province/city name
rename A city

* Rename year columns (Stata typically imports numeric headers as _1997, _2024)
rename _1997 HES1997
rename _2024 HES2024

* Enforce numeric types
destring HES1997 HES2024, replace

* Compute structural upgrading (ΔHES)
gen dHES = HES2024 - HES1997

keep city HES1997 HES2024 dHES
tempfile hes_delta
save `hes_delta'

*********************************************************
* STEP 5: MERGE AYS AND HES DIMENSIONS
*********************************************************
use `ays_delta', clear
```

ISSN: 3008-1629

```
merge 1:1 city using `hes_delta'
drop _merge

*******************************************************
* STEP 6: COMPUTE MEDIAN THRESHOLDS (ROBUST TO OUTLIERS)
*******************************************************
summarize dAYS, detail
scalar med_dAYS = r(p50)

summarize dHES, detail
scalar med_dHES = r(p50)

display "Median ΔAYS = " med_dAYS
display "Median ΔHES = " med_dHES

*******************************************************
* STEP 7: ASSIGN EXPANSION TYPOLOGIES (FOUR QUADRANTS)
*******************************************************
gen Type = ""

replace Type = "Dual-track Leap" ///
   if dAYS >= med_dAYS & dHES >= med_dHES

replace Type = "Mass Expansion" ///
   if dAYS >= med_dAYS & dHES <  med_dHES

replace Type = "Structural Upgrading" ///
   if dAYS <  med_dAYS & dHES >= med_dHES

replace Type = "Slow Expansion" ///
   if dAYS <  med_dAYS & dHES <  med_dHES

*******************************************************
* STEP 8: EXPORT FINAL RESULTS
*******************************************************
export excel using "Provincial_Education_Expansion_Typology.xlsx", ///
   firstrow(variables) replace

display "=== Typology construction completed successfully ==="
```

ISSN: 3008-1629

4 Empirical Analysis

4.1 Data Analysis Results

Table 5 Provincial Educational Expansion Indicators (1997–2024)

Province	AYS1997	AYS2024	ΔAYS	HES1997	HES2024	ΔHES	Expansion Type
Shanghai	8.8851	11.8562	2.9711	0.0889	0.4093	0.3204	Dual-track Leap
Yunnan	5.7889	9.0191	3.2302	0.0124	0.1901	0.1777	Mass Expansion
Inner Mongolia	7.1816	10.1955	3.0139	0.036	0.2555	0.2195	Dual-track Leap
Beijing	9.5013	12.6538	3.1525	0.1344	0.5083	0.3739	Dual-track Leap
Jilin	8.0274	10.2561	2.2287	0.0499	0.2365	0.1866	Upgrading
Sichuan	6.5695	9.2017	2.6322	0.0199	0.1827	0.1628	Slow Expansion
Tianjin	8.3756	11.4658	3.0902	0.0727	0.3616	0.2889	Dual-track Leap
Ningxia	6.4475	9.6184	3.1709	0.0327	0.2328	0.2001	Dual-track Leap
Anhui	6.5571	9.4051	2.848	0.0183	0.1939	0.1756	Mass Expansion
Shandong	6.5031	9.7586	3.2555	0.0149	0.2011	0.1862	Dual-track Leap
Shanxi	7.6823	10.5041	2.8218	0.0336	0.2449	0.2113	Upgrading
Guangdong	7.5044	10.1459	2.6415	0.0367	0.2287	0.192	Upgrading
Guangxi	6.613	9.4962	2.8832	0.0093	0.1651	0.1558	Mass Expansion
Xinjiang	7.512	9.909	2.397	0.0558	0.2227	0.1669	Slow Expansion
Jiangsu	6.9104	10.2227	3.3123	0.0205	0.2453	0.2248	Dual-track Leap
Jiangxi	7.0467	9.4968	2.4501	0.0188	0.169	0.1502	Slow Expansion
Hebei	7.174	9.9567	2.7827	0.021	0.1902	0.1692	Slow Expansion
Henan	7.1014	9.7561	2.6547	0.016	0.1767	0.1607	Slow Expansion
Zhejiang	6.8148	9.8108	2.996	0.025	0.2233	0.1983	Dual-track Leap
Hainan	7.2121	9.95	2.7379	0.0232	0.1971	0.1739	Slow Expansion
Hubei	7.2215	10.0422	2.8207	0.0339	0.2124	0.1785	Slow Expansion

ISSN: 3008-1629

Hunan	7.2243	9.9565	2.7322	0.0202	0.2016	0.1814	Upgrading
Gansu	6.1273	8.7975	2.6702	0.0165	0.1799	0.1634	Slow Expansion
Fujian	6.7342	9.6471	2.9129	0.0263	0.1974	0.1711	Mass Expansion
Xizang	3.4998	6.5889	3.0891	0.0032	0.1603	0.1571	Mass Expansion
Guizhou	5.8494	8.6069	2.7575	0.0201	0.1618	0.1417	Slow Expansion
Liaoning	8.1017	10.431	2.3293	0.0602	0.2403	0.1801	Upgrading
Chongqing	6.601	10.0203	3.4193	0.0197	0.2336	0.2139	Dual-track Leap
Shaanxi	7.0661	10.3126	3.2465	0.0299	0.2612	0.2313	Dual-track Leap
Qinghai	4.6926	8.7829	4.0903	0.0201	0.2024	0.1823	Dual-track Leap
Heilongjiang	7.8562	10.0949	2.2387	0.0475	0.2228	0.1753	Slow Expansion

***Notes on Indicators:**

1. AYS1997 refers to Average Years of Schooling in 1997, measuring the average number of completed years of formal education among the population aged 6 and above in each province at the baseline year.
2. AYS2024 refers to Average Years of Schooling in 2024, representing the corresponding educational attainment level at the end of the observation period.
3. ΔAYS denotes the cumulative change in average years of schooling between 1997 and 2024, calculated as AYS2024 minus AYS1997. It captures the magnitude of quantity-oriented educational expansion, reflecting the overall improvement in educational attainment.
4. HES1997 refers to the Share of Population with College Education or Above in 1997. This indicator measures the proportion of the provincial population holding tertiary-level qualifications at the baseline year.
5. HES2024 refers to the Share of Population with College Education or Above in 2024, representing the tertiary educational structure at the end of the study period.
6. ΔHES denotes the cumulative change in the share of population with college education or above between 1997 and 2024, calculated as HES2024 minus HES1997. It reflects structure-oriented educational upgrading, capturing the deepening of tertiary education within the overall educational composition.
7. Expansion Type represents the provincial educational expansion typology derived from the two-dimensional framework. Provinces are classified into four categories based on whether their ΔAYS and ΔHES are above or below the median thresholds across all provinces.

Based on the two core cumulative change indicators for the period 1997 to 2024, provincial educational expansion exhibits a clear fourfold path differentiation, and each category can be directly verified through the specific combination of ΔAYS and ΔHES.

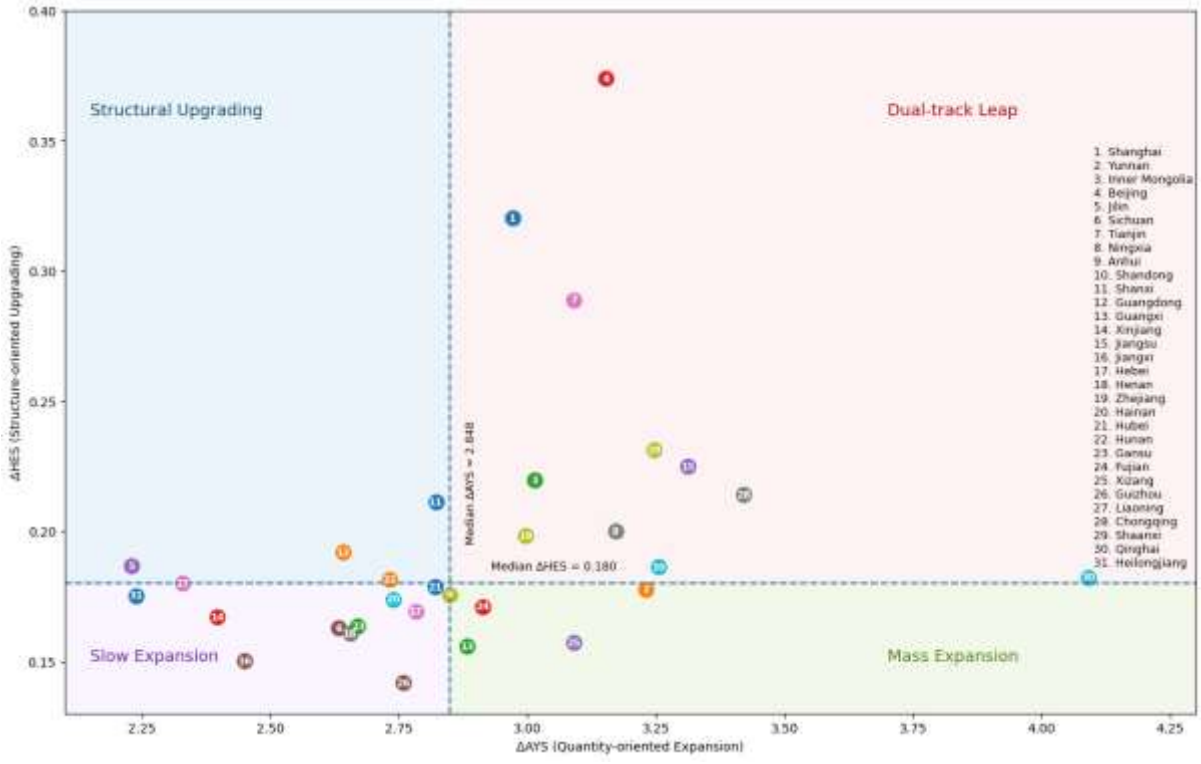
ISSN: 3008-1629

First, the dual track leap type includes Beijing, Shanghai, Tianjin, Jiangsu, Zhejiang, Shandong, Shaanxi, Chongqing, Inner Mongolia, Ningxia, and Qinghai. The defining feature of this group is the simultaneous and substantial increase in both educational diffusion and higher education structural upgrading, reflecting a pattern of coordinated advancement in both quantity and quality. For example, Beijing records a ΔAYS of 3.1525 and a ΔHES as high as 0.3739. Shanghai achieves a ΔHES of 0.3204, with its HES reaching 0.4093 in 2024, indicating that municipalities not only possess a stronger initial higher education base but also sustain continued structural deepening. Eastern coastal provinces display similarly stable coordinated growth. Jiangsu records a ΔAYS of 3.3123 and a ΔHES of 0.2248; Zhejiang records a ΔAYS of 2.9960 and a ΔHES of 0.1983; and Shandong records a ΔAYS of 3.2555 and a ΔHES of 0.1862. These patterns are consistent with a developmental logic in which industrial upgrading and upward shifts in skill demand simultaneously stimulate broader educational diffusion and higher education deepening. Notably, several central and western provinces also exhibit pronounced catch up dynamics. Chongqing records a ΔAYS of 3.4193 and a ΔHES of 0.2139, while Qinghai reaches a ΔAYS of 4.0903 and a ΔHES of 0.1823, suggesting that later developing regions may achieve dual dimensional advancement under strengthened policy support and resource investment. Second, the mass expansion type includes Yunnan, Anhui, Guangxi, Fujian, and Xizang. The salient characteristic of this group is relatively rapid improvement in ΔAYS combined with comparatively modest increases in ΔHES, reflecting a stage characterized by diffusion preceding structural upgrading. Yunnan records a ΔAYS of 3.2302 and a ΔHES of 0.1777; Xizang records a ΔAYS of 3.0891 and a ΔHES of 0.1571; Anhui records a ΔAYS of 2.8480 and a ΔHES of 0.1756; Guangxi records a ΔAYS of 2.8832 and a ΔHES of 0.1558; and Fujian records a ΔAYS of 2.9129 and a ΔHES of 0.1711. These figures indicate that while average years of schooling have increased substantially, suggesting rapid expansion in basic education coverage and general educational access, the proportion of the population with higher education has not risen at a comparable pace. Educational expansion in these provinces therefore reflects broad based human capital accumulation, whereas the transition toward a more advanced talent supply structure remains in progress. Third, the structural upgrading type includes Jilin, Shanxi, Guangdong, Hunan, and Liaoning. This group is characterized by higher education structural upgrading outpacing overall gains in average years of schooling, indicating a trajectory in which qualitative transformation precedes broad based diffusion. Shanxi records a ΔHES of 0.2113, Guangdong 0.1920, Hunan 0.1814, Jilin 0.1866, and Liaoning 0.1801, all indicating relatively rapid increases in higher education attainment. However, corresponding ΔAYS values do not rise to the same extent. Liaoning

ISSN: 3008-1629

records a ΔAYS of 2.3293, Jilin 2.2287, and Guangdong 2.6415. This configuration suggests that these provinces are achieving advancement primarily through internal restructuring of the educational hierarchy, possibly driven by intensified demand for high level skills associated with industrial transformation or by a relatively high pre existing diffusion level that constrains further marginal increases in average years of schooling. The resulting pattern reflects structural upgrading that advances ahead of broad based attainment gains. Finally, the slow expansion type includes Sichuan, Xinjiang, Jiangxi, Hebei, Henan, Hainan, Hubei, Gansu, Guizhou, and Heilongjiang. The common feature of this group is comparatively low values for both ΔAYS and ΔHES, indicating lagging performance along both dimensions. For example, Sichuan records a ΔAYS of 2.6322 and a ΔHES of 0.1628; Xinjiang records a ΔAYS of 2.3970 and a ΔHES of 0.1669; Jiangxi records a ΔAYS of 2.4501 and a ΔHES of 0.1502; Guizhou records a ΔHES of only 0.1417; and Heilongjiang records a ΔAYS of 2.2387 and a ΔHES of 0.1753. Overall, this group has neither achieved rapid improvements in average years of schooling nor significant breakthroughs in higher education structural upgrading. Such a pattern suggests the possible coexistence of multiple constraints, including pressures on basic education provision, limited industrial upgrading that weakens demand for highly educated labour, restricted higher education supply capacity, and outward migration of talent. Together, these factors contribute to a relatively gradual expansion trajectory.

Graphic 3. Provincial Educational Expansion Typology (1997–2024)



ISSN: 3008-1629

4.2 Robustness Tests

Within the rule based classification framework that adopts a four quadrant division based on the median threshold, the incorporation of threshold substitution tests, extreme value treatment tests, and clustering comparison tests carries substantial methodological significance. First, the threshold substitution test replaces the median threshold with the mean threshold in order to examine whether the classification results are highly sensitive to the chosen cut off criterion, thereby assessing the parametric robustness of the classification rule. In research involving quantile partitioning and rule based grouping, sensitivity analysis of threshold selection is widely regarded as a crucial procedure for preventing structural distortion arising from arbitrary boundary specification (Imbens, 2020; Angrist & Pischke, 2009). Second, the extreme value treatment test applies one percent winsorisation to ΔAYS and ΔHES to determine whether the identified typological structure is disproportionately influenced by a small number of outlying observations. In regional development studies and panel data analysis, extreme values may generate systematic shifts in distributional location and classification boundaries. Robust treatment of tail observations therefore represents a standard methodological practice to safeguard the reliability of statistical inference (Rousseeuw & Hubert, 2018). Third, the clustering comparison test employs the K means data driven grouping method to conduct cross validation against the theoretically defined four quadrant classification, thereby examining whether the rule based typology is consistent with the intrinsic structure of the data. In typological research and regional comparative analysis, contrasting rule based classification with unsupervised clustering techniques is widely considered an effective strategy for evaluating structural endogeneity (Hastie, Tibshirani, & Friedman, 2009; Everitt et al., 2011). Taken together, these three robustness procedures assess the stability and explanatory credibility of the educational expansion typology from the perspectives of parameter specification, distributional characteristics, and methodological choice. Through this multi layered validation, the study enhances the statistical credibility and methodological rigor of its empirical conclusions.

Table 6 Threshold Replacement Robustness Test

Item	Median Threshold (Baseline)	Mean Threshold (Alternative)	Agreement Rate
ΔAYS Threshold	2.848	2.9184	
ΔHES Threshold	0.1801	0.1974	
Classification Consistency	—	—	74.20%

ISSN: 3008-1629

Table 6 shows that after replacing the baseline median thresholds of ΔAYS = 2.848 and ΔHES = 0.1801 with mean based thresholds of ΔAYS = 2.9184 and ΔHES = 0.1974, the consistency rate of provincial educational expansion classification reaches 74.20 percent. This finding indicates that although the mean thresholds shift upward relative to the medians, reflecting a certain degree of right skewness in the sample distribution, particularly due to the influence of high growth provinces on the mean, the overall typological structure remains largely stable. Only a limited number of provinces located near the original boundary values experience category reassignment. From the perspective of parameter sensitivity, the classification framework therefore demonstrates moderate to strong robustness with respect to threshold specification. Educational expansion types are not mechanically generated by a specific cut off point; rather, they reflect relatively stable grouping patterns embedded in the underlying distributional structure of the sample. This result suggests that the four quadrant framework is capable of capturing the substantive combinational differences between quantitative educational expansion and structural upgrading across provinces under alternative partition criteria.

Table 7 Outlier Treatment Robustness Test

Item	Baseline Classification	1% Winsorisation	Agreement Rate
ΔAYS	Raw values	Winsorised	
ΔHES	Raw values	Winsorised	
Classification Consistency	—	—	100%

Table 7 indicates that after applying one percent winsorisation to ΔAYS and ΔHES, the consistency rate of provincial educational expansion classification reaches 100 percent, meaning that the typological assignment of all provinces remains identical to the baseline results. This outcome demonstrates that the original classification structure is not substantively influenced by extreme values or tail distributions. Although certain provinces, particularly those with relatively high ΔAYS or markedly large ΔHES, exert some influence at the numerical level, the contraction of tail observations does not alter the four quadrant configuration in any way. This finding suggests that the typology is determined primarily by the overall distributional pattern rather than being driven by a small number of outlying cases. From the perspective of statistical robustness, the structural characteristics of educational expansion types exhibit a high degree of distributional stability. The classification results therefore do not suffer from dependence on extreme observations, which substantially strengthens the empirical credibility and interpretive reliability of the study's conclusions.

ISSN: 3008-1629

Table 8 Clustering Comparison Test (K-means, k = 4)

Original Type	Cluster 0	Cluster 1	Cluster 2	Cluster 3
Dual-track Leap	0	10	0	1
Mass Expansion	3	2	0	0
Slow Expansion	7	0	3	0
Upgrading	3	0	2	0

Table 8 conducts a cross validation between the theoretically constructed four quadrant classification based on median thresholds and the results derived from K means unsupervised clustering, in order to assess whether the typology aligns with the intrinsic structure of the data. The results show that the Dual track Leap category is highly concentrated in Cluster 1, which contains ten provinces, with only one province distributed in other clusters. This indicates that the dual track leap type forms a distinct high value concentration area within the two dimensional space defined by ΔAYS and ΔHES, demonstrating strong structural homogeneity. The Slow Expansion category is primarily located in Cluster 0 and Cluster 2, reflecting the characteristics of low increment groupings. Although the Mass Expansion and Upgrading categories exhibit some degree of dispersion, they nonetheless display relatively clear clustering tendencies. Overall, there is a high level of correspondence between the rule based theoretical classification and the data driven clustering results. This consistency suggests that the four quadrant framework is not an artificially imposed grouping scheme, but rather one that effectively captures the natural grouping patterns embedded in the joint distribution of ΔAYS and ΔHES. From a methodological perspective, this finding provides strong support for the structural endogeneity and classification validity of the proposed typology.

5 Conclusion

5.1 Main Findings

Drawing upon the cumulative changes in average years of schooling and in the proportion of the population holding college education or above during the period from 1997 to 2024, the analysis demonstrates that provincial educational expansion in China has not followed a single linear trajectory. Instead, it has differentiated into four relatively stable pathway types, reflecting pronounced asynchrony and combinational variation between improvements in educational diffusion and structural upgrading across regions. The dual track leap type achieves substantial gains on both indicators simultaneously. This group includes high baseline municipalities such as Beijing, Shanghai, and Tianjin, as well as eastern coastal

ISSN: 3008-1629

provinces including Jiangsu, Zhejiang, and Shandong, and extends to central and western regions such as Chongqing and Qinghai that exhibit strong catch up dynamics. These patterns suggest that coordinated dual dimensional advancement may arise either from the agglomeration of knowledge and innovation resources that sustain higher education deepening, or from intensified policy support and the driving force of regional growth poles that enable rapid transformation. The mass expansion type is characterized by relatively rapid increases in average years of schooling alongside more limited structural deepening in higher education. In these regions, expansion is primarily concentrated in the broadening of basic education coverage and the general advancement of educational diffusion, while the transition toward a more advanced talent structure remains at an intermediate stage. The structural upgrading type exhibits faster growth in the share of the population with college education or above, combined with comparatively moderate increases in average years of schooling. This configuration indicates a tendency to achieve breakthroughs within the internal hierarchy of the education system. Such a pattern may be associated with heightened demand for high skilled labour driven by industrial transformation, or with diminishing marginal space for further diffusion in regions that have already attained relatively high baseline levels of educational attainment. The slow expansion type records comparatively low gains on both indicators, reflecting lagging performance in both educational diffusion and structural upgrading. This outcome suggests the possible coexistence of multiple constraints, including pressure on basic education provision, insufficient momentum in industrial upgrading, limited higher education supply capacity, and outward migration of talent. Overall, these findings reveal that the fundamental differences in provincial educational expansion lie not merely in variations in developmental level, but in the distinct coupling patterns between quantitative expansion and structural upgrading. It is this differentiated coupling that generates a multi pathway evolutionary landscape across regions.

5.2 Policy Implications

The differentiated pathways identified above indicate that provincial variation in educational expansion is not merely a matter of developmental level, but more fundamentally reflects distinct coupling patterns between educational diffusion and higher education structural upgrading. Accordingly, educational policy should no longer proceed under a single objective or uniform pace. Instead, it should adopt differentiated and context specific strategies aligned with the characteristics of each pathway type.

ISSN: 3008-1629

For regions classified as dual track leap, policy priorities should shift from further scale expansion toward qualitative deepening and structural optimization. Emphasis should be placed on enhancing the alignment between higher education provision, technological innovation systems, and the demands of industrial upgrading, thereby preventing structural redundancy or resource misallocation that may arise from continued expansion at already high baseline levels. For mass expansion regions, while consolidating achievements in basic education coverage and broad based human capital diffusion, greater efforts should be directed toward strengthening higher education and vocational education systems. This includes expanding institutional carrying capacity, optimizing disciplinary and program structures, reinforcing integration between education and industry, and enhancing regional employment absorption capacity. Such measures would facilitate a gradual transition from extensive human capital accumulation toward a more advanced talent supply structure. For structural upgrading regions, maintaining the momentum of higher education deepening must be accompanied by improvements in basic education quality and stronger articulation between secondary and tertiary education. Enhancing balance across all educational stages is essential to mitigate the structural risk of rising higher education participation without corresponding gains in overall educational attainment. For slow expansion regions, comprehensive policy packages are required. These should simultaneously address equitable investment in basic education, allocation of higher education resources, talent retention mechanisms, and demand side coordination linked to industrial transformation. Only through coordinated intervention across these dimensions can the mutually reinforcing cycle of low diffusion and limited upgrading be effectively disrupted. Beyond governance within the education system itself, the identified typology also underscores the close coupling between educational expansion, regional industrial structure, innovation agglomeration, and population mobility patterns. Educational policy should therefore extend beyond sectoral adjustment and become integrated within broader regional development strategies. For later developing regions, rather than distributing resources evenly, greater emphasis should be placed on constructing higher education supply systems that align with dominant regional industries and urban functional positioning. This may include establishing specialized universities and application oriented institutions, strengthening high quality vocational education and technical skill training, and building collaborative platforms among universities, local governments, and enterprises to enhance alignment between talent cultivation and industrial demand while improving local retention capacity. For regions that have entered the dual track leap stage, the next phase of development should prioritize

ISSN: 3008-1629

efficiency and innovation output rather than further scale enlargement. By optimizing disciplinary layouts, improving the efficiency of research and industry transformation chains, and strengthening interregional collaboration and high level talent mobility mechanisms, these regions can convert scale advantages into quality and innovation advantages. Ultimately, whether educational expansion can be transformed into sustainable developmental momentum depends on the dynamic coordination between quantitative expansion and structural upgrading, as well as on the degree to which educational supply aligns coherently with regional development strategies.

5.3 Limitations and Future Directions

Although this study conceptually distinguishes between quantitative educational expansion and structural upgrading and constructs a provincial typological framework based on long term panel data, several issues warrant further refinement. First, at the measurement level, the use of average years of schooling and the proportion of the population holding college education or above effectively captures macro level diffusion and structural elevation, yet the framework remains centered on hierarchical proportions rather than educational quality. Structural upgrading in higher education does not necessarily imply improvements in the quality of provision. Substantial regional variation may exist in institutional stratification, research capacity, disciplinary configuration, and graduate employment outcomes, dimensions that are not incorporated into the present analysis. Second, the adoption of the provincial administrative unit as the analytical scale facilitates identification of broad patterns of regional differentiation but does not capture intra provincial disparities, metropolitan variation, or the deeper influence of interregional talent mobility on structural transformation. Given the high mobility of highly educated populations, reliance on household registration or permanent residence based statistics may underestimate the moderating effect of net talent inflows on structural upgrading. Finally, while this study identifies distinct pathways and provides typological characterization, it does not conduct rigorous econometric testing of the causal mechanisms underlying pathway formation. The dynamic interactions among supply side resource allocation, demand side industrial structure, population mobility, and policy intensity remain to be systematically examined. Future research may advance in at least three directions. First, at the level of indicator construction, the integration of educational quality alongside hierarchical distribution would allow the development of a more comprehensive framework. Indicators such as the number of nationally designated leading universities, research funding intensity, the proportion of high skill employment, and

ISSN: 3008-1629

the share of vocational education could be incorporated to reconstruct the meaning of educational upgrading from a multidimensional perspective encompassing quantity, structure, and quality. Second, methodologically, the application of spatial econometric models, club convergence tests, and dynamic panel estimators would enable systematic examination of whether convergence or divergence dynamics exist across pathway types and would facilitate identification of the causal roles of industrial upgrading, fiscal investment, and talent mobility in pathway transformation. Third, with respect to spatial and temporal scale, future studies could disaggregate the unit of analysis to the prefectural city or metropolitan cluster level and integrate data on migration flows and university distribution to explore the interaction between educational expansion and the agglomeration of regional core cities. Such analysis would help determine whether structural upgrading exhibits a spatial pattern characterized by core reinforcement and peripheral lagging. By extending measurement depth, methodological precision, and spatial scope, future research will be better positioned to uncover the long term evolutionary logic linking quantitative expansion and structural upgrading within China's higher education massification process, as well as the institutional and economic mechanisms underlying differentiated regional trajectories.

References

[1] Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton University Press.

[2] Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). Cluster analysis (5th ed.). Wiley.

[3] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer.

[4] Imbens, G. W. (2020). Potential outcome frameworks for causal inference. Journal of Economic Perspectives, 34(2), 65–92. https://doi.org/10.1257/jep.34.2.65

[5] Rousseeuw, P. J., & Hubert, M. (2018). Anomaly detection by robust statistics. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(2), e1236. https://doi.org/10.1002/widm.1236

[6] Cunningham, C. A., & Samson, C. (2021). Neoliberal meritocracy: How 'widening participation' to universities in England reinforces class divisions. On Education, 4(10). https://doi.org/10.17899/on_ed.2021.10.8

[7] Hu, W., & Cao, Y. (2023). Advancement of pharmacy accreditation in the field of Chinese higher education. Pharmacology & Pharmacy. https://doi.org/10.4236/pp.2023.149024

[8] Jiang, J., & Ke, G. (2021). China's move to mass higher education since 1998: Analysis of higher education expansion policies. Higher Education Quarterly. https://doi.org/10.1111/hequ.12313

ISSN: 3008-1629

[9] Mathew, L., & Martin, F. (2025). Gender and education mobilities in China and India: Towards an inter-Asia comparative framework. Inter-Asia Cultural Studies, 26, 231–253. https://doi.org/10.1080/14649373.2025.2470598

[10] Sims, D., Zingela, Z., Mokhachane, M., Botha, G., Mawela, D., Singaram, V., Baatjes, K., Green-Thompson, L., & Begg, K. (2025). Medical education, reflections and perspectives from South Africa: A review. BMC Medical Education, 25. https://doi.org/10.1186/s12909-025-06910-8

[11] Song, X., & Xu, D. (2024). More graduates, fewer skills? Vocational education expansion and skilled labour shortages in China. The China Quarterly, 260, 970–985. https://doi.org/10.1017/s0305741023001856

[12] Jiang, J., & Ke, G. (2021). China's move to mass higher education since 1998: Analysis of higher education expansion policies. Higher Education Quarterly. https://doi.org/10.1111/hequ.12313

[13] Song, X., & Xu, D. (2024). More graduates, fewer skills? Vocational education expansion and skilled labour shortages in China. The China Quarterly, 260, 970–985. https://doi.org/10.1017/s0305741023001856

[14] Mathew, L., & Martin, F. (2025). Gender and education mobilities in China and India: Towards an inter-Asia comparative framework. Inter-Asia Cultural Studies, 26, 231–253. https://doi.org/10.1080/14649373.2025.2470598

[15] Cunningham, C. A., & Samson, C. (2021). Neoliberal meritocracy: How 'widening participation' to universities in England reinforces class divisions. On Education, 4(10). https://doi.org/10.17899/on_ed.2021.10.8

[16] Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. Journal of Development Economics, 104, 184–198. https://doi.org/10.1016/j.jdeveco.2012.10.001

[17] Fleisher, B., Li, H., & Zhao, M. Q. (2010). Human capital, economic growth, and regional inequality in China. Journal of Development Economics, 92(2), 215–231. https://doi.org/10.1016/j.jdeveco.2009.01.010

[18] Guo, Y., & Li, X. (2024). Regional inequality in China's educational development An urban-rural comparison. Heliyon, 10(4), e26249. https://doi.org/10.1016/j.heliyon.2024.e26249

[19] Han, Y., Ni, R., & Gao, J. (2023). Regional inequality of higher education development in China Comprehensive evaluation and geographical representation. Sustainability, 15(3), 1824. https://doi.org/10.3390/su15031824

[20] Li, Y. A., Whalley, J., Zhang, S., & Zhao, X. (2011). The higher educational transformation of China and its global implications. The World Economy, 34(4), 516–545. https://doi.org/10.1111/j.1467-9701.2011.01344.x

[21] Marginson, S. (2016). The worldwide trend to high participation higher education Dynamics of social stratification in inclusive systems. Higher Education, 72, 413–434. https://doi.org/10.1007/s10734-016-0016-x

[22] Mok, K. H. (2016). Massification of higher education, graduate employment and social mobility in East Asia. Journal of Education Policy. https://doi.org/10.1080/01425692.2015.1111751

ISSN: 3008-1629

[23] Trow, M. (1973). Problems in the transition from elite to mass higher education. Carnegie Commission on Higher Education.