Heterogeneity in Drivers of Higher Education Attainment: Panel Data Evidence from 21st-century China

Xiaoyi Chen

January 2025

Abstract

Education plays a crucial role in development, as it improves quality of life and enhances stability. This study investigates the key drivers of higher education attainment in China from 2005-2019 using panel data from 31 provinces. Through fixed effects regression, the study explore strength of urbanisation, economic development, education expenditure, and transportation accessibility in affecting new enrolment into higher education at a national and regional level. The study finds urbanisation has the largest positive effect in education attainment. Economic growth shows diminishing returns on education attainment over time, with positive effects primarily in less developed regions. Education expenditure yields mixed results, while transportation infrastructure demonstrates increased importance in later years, and in western and northeastern regions. The analysis highlights distinct heterogeneities in the factors that influence higher education attainment across the economic regions of China. These findings suggest policy interventions should consider granular effects of indicators, particularly at a provincial level. Nevertheless, heterogenous effects across China's economic regions and temporal stages highlight the importance of considering local contexts and development stages when formulating higher education policy.

1 Introduction

The importance of education in development is well-established, as it plays a pivotal role in improving quality of life and enhancing stability. The World Bank recognises education as a "cornerstone of economic and social development" (Haddad et al., 1995) and one of the strongest instruments to reduce poverty and improve peace. The development of human resources through education is a capital investment that reduces social inequality and fosters innovation (Galor, 2011). In addition, the emphasis on the importance of higher education continues to grow in the modern global economy. UNESCO recognises higher education as a "rich cultural and scientific asset which enables personal development and promotes economic, technological, and social change" (UNESCO, n.d.).

Since the mid-20th century, China's education system has undergone turbulent transformations, marked by radical shifts in policy and structure. In the late 1980s, the government began to stimulate reforms in education, emphasising its role in economic development. These policies enacted nine-year compulsory schooling for all children between the ages six and fifteen (Congress, 1986), and saw the near complete eradication of illiteracy amongst the Chinese population (Yeoh & Chu, 2012). However, concerns remain about inequalities across China's regions, particularly disparities in non-compulsory education. Research has examined education outcomes across China's urban-rural divide (Qian & Smyth, 2008; Yang & Qiu, 2016) and coastal-inland regions (Mok et al., 2009; Zhang & Zhang, 2003). However, factors that constrain the development of education

has largely been explored independently (Fu & Ren, 2010; Hannum, 1999; Wu, 2024). Studies such as Yang et al. (2014) explore multiple drivers of education and their importance, but their their broad focus on total education attainment may obscure significant gaps in non-compulsory education, particularly in post-secondary studies.

Given these limitations in existing literature, this study aims to provide a comprehensive analysis of the factors that influence educational outcomes and their relative influence, with a focus on regional variations in higher education attainment. By using panel data from the 31 provinces, municipalities, and autonomous regions of China over 15 years, this study hopes to offer new insights into the forces driving disparities in post-secondary education across China. In line with this, the following research questions are presented:

1. **RQ1**: What are the driving factors of higher education attainment in 21st-century China?
2. **RQ2**: What are the regional differences in factors that influence higher education attainment?

2 Theoretical Framework

2.1 Higher Education and Human Capital Formation

Education attainment is often discussed as a mechanism for human capital development, shaping both individual outcomes and aggregated productivity through the acquisition of competencies (T. P. Schultz, 1992). A substantial body of literature has identified drivers of human capital formation, including education infrastructure (Lange & Topel, 2006; T. W. Schultz, 1971), economic wellbeing (Marquez-Ramos & Mourelle, 2019; Pastor et al., 2018), transportation accessibility (Arouri et al., 2014), and urbanisation (Bertinelli & Zou, 2008). Given the fundamental relationship between human capital and educational attainment, these factors similarly shape educational development of regions. Drawing upon these frameworks, this study examine four dimensions that influence educational outcomes through a longitudinal panel-data analysis of Chinese provinces.

2.2 Economic Development

Research into economic implications of education show a strong connection between skills acquired through higher education and economic outcomes, including national economic growth and individual earnings (Hanushek & Woessmann, 2008). Additionally, an analysis of 195 countries found higher levels of expected human capital was associated with higher levels of GDP (Lim et al., 2018). The economic strength of a region also plays a crucial role in its ability to develop infrastructure that support knowledge creation (Vila et al., 2014). Previous studies find a positive correlation between capital growth and education attainment (Agasisti & Bertoletti, 2022; Paudel, 2023; Pink-Harper, 2015). However, considerations need to be made about the law of diminishing returns on investment in human capital (Gemmell, 1995), though some empirical evidence suggests investments in education can defy these laws (Arshed et al., 2021).

2.3 Investment in Education

Investment in education systems represent a critical dimension for understanding higher education attainment. Returns to investment based on human capital theory have been systematically

evaluated since the mid-20th century (Psacharopoulos & Patrinos, 2018). A study on education funding gaps in China show that inequitable funding can result in disparate resources and opportunities for students of different socioeconomic backgrounds, particularly across the urban-rural divide (Molnar et al., 2015). Additionally, economically strong provinces maintain substantial advantages in educational provision, leading to persistently higher rates of university enrolment (Zhang & Kanbur, 2009).

2.4 Transportation Accessibility

A small body of literature has examined the impact of transport accessibility on employment and education (Gurley & Bruce, 2005; Kawabata, 2002), showing that advanced transportation can significantly improve these outcomes. Through investigating the role of transportation on economic development, Chen et al. found improvement of accessibility to public transportation infrastructures improved the efficiency of allocation of resources, such as goods and human capital. Considering China's significant railway and highway reforms in the 21st-century (Shaw et al., 2014), it is worthwhile to consider how transportation access has affected higher education attainability, particularly in regions where mobility is becoming more feasible.

2.5 Urbanisation

In past decades, China's urbanisation rate have shown tremendous growth (UN, 2013). Some empirical evidence suggests urbanisation is highly effective for economic growth, as it enables concentration of human capital (McGranahan & Satterthwaite, 2014). Although, inequalities in the development of urbanised areas may deepen existing socioeconomic disparities, and varying levels of available resources can further exacerbate the gap between rural and urban regions. For example, the rural-to-urban migration in early 21-century China saw massive influxes of people into eastern regions. This caused China's coastal municipalities like Beijing and Shanghai to develop rapidly. Urbanisation led to huge advancements in education in these areas, as evident in the fact that Beijing has the highest number of universities of any city in the country (Gaokao.com, 2020). Given inequalities in development across regions, investigating how these patterns influence higher education attainment is crucial for addressing disparities in access to educational resources and opportunities.

3 Methodology

3.1 Data Source and Study Design

The study analyses factors influencing the higher education enrolment in China and their relative strengths at a national and regional level. Regional boundaries were derived from definitions provided by China's National Bureau of Statistics (Figure 1). The drivers of higher education attainment were evaluated across the entire 15-year study period, and at two sub-periods. At a regional level, the assessment was conducted across the entire study period while controlling for regions to uncover patterns of heterogeneity.

Data for this study was obtained from the China Statistical Yearbooks (National Bureau of Statistics of China, 2005–2019), which provides yearly comprehensive statistical data on China's economic, social, and demographic development with data available at a national and regional level.

Data was obtained for higher education attainment and various driving factors from 2005-2019, recent years were not included in the analysis to mitigate the impacts that COVID-19 may have had on the educational development of China. Since the goal of the study was the evaluate the strength and influence of factors on higher educational attainment, data from the Greater China regions (Taiwan, Macao, Hongkong) was not included as these regions operate under distinct educational systems.

3.2 Indicator Selection

Higher education attainment was operationalized using number of entrants in regular higher education as the predicted variable. New enrolment was selected over total number of students as it provides a more dynamic measure of higher education attainment. To yield meaningful cross-regional comparisons, the number of entrants was scaled per 1000 people based on provincial population. This normalisation prevents biases that arise from differing provincial demographics and ensures meaningful interpretation given China's large population. Since population statistics from the Yearbooks are scaled per 10,000 people, the number of new entrants per 1000 people in each region (hereafter referred to as entrants) was calculated as follows:

$$\text{entrants} = \frac{\text{new\_enrolment}}{\text{population} \times 10} \tag{1}$$

Seven explanatory variables were initially chosen, each representing one of the four dimensions of China's socio-economic development conditions. All indicators were normalised by either the total population or the land area of the province (Appendix A). Due to missing education expenditure data in the 2013 Yearbooks, linear interpolation was employed to estimate the values. Regional GDP per capita was log transformed based on the assumption of diminishing returns in human development (Cahill, 2002) and evidence of heteroscedasticity and non-linearity in raw regional GDP per capita versus entrants (Appendix B). To facilitate the integration of variables with differing scales and units of measurement, all predictors were normalised using z-score standardisation. The predicted variable, entrants, was not normalised to maintain interpretability of coefficients. Next, to avoid multicollinearity among the predictor variables, an iterative selection process using Variance Inflation Factors (VIF) was employed. Following Gordon, the variables with the highest VIF value was removed, and VIFs were recalculated until all remaining predictors had VIF values below 5.0. Post-screening, four predictors were retained: log-transformed GDP, education expenditure, highway density, and urbanisation (Table 1).

Variable	**Description**	**VIF**
entrants	Entrants per 1,000 people by region	-
ln(*per_capita_GDP*)	Log transformed per capita GDP by region	4.949
per_capita_expenditure	Per capita education expenditure by region	2.440
highway_density	Highway accessibility by region	1.301
urbanisation	Regional urban population to total population	3.591

Table 1: Final selected indicators with their definitions and VIF results

3.3 Model Selection

3.3.1 National Level

A simple pooled ordinary least squares (OLS) regression was employed as a baseline model. Visual inspection of the residuals (Appendix C) show assumptions of linearity and normality generally hold, with minor deviations. Diagnostic tests were conducted to examine potential violations of other classical regression assumptions (Appendix D). Hausman test was implemented to determine the appropriate model between fixed and random effects. The results from diagnostic tests indicate violations of classical assumptions, thus Driscoll-Kraay standard errors (Hoechle, 2007) were employed in both fixed and random effect estimations to ensure robustness of results. The Hausman test indicated that the fixed effects model was the most appropriate for this study:

$$y_{it} = \alpha_i + \beta_1 \ln(\text{gdp}_{it}) + \sum_{k=2}^{4} \beta_k x_{kit} + \epsilon_{it} \tag{2}$$

where y_{it} represents the entrants in province i in year t (Equation 1), α_i is the province-specific fixed effect, $\beta_1 \ln(\text{gdp}_{it})$ is the log-transformed GDP per capita, $\sum_{k=2}^{4} \beta_k x_{kit}$ is the sum of other independent variables, and ϵ_{it} is the error term for province i and year t. The regression analysis for the relationship between the explanatory variables and higher education attainment was first evaluated at the national level as a whole (2005-2019), and across two time periods (2005-2012, 2013-2019).

3.3.2 Regional Differences

To evaluate regional differences in the factors influencing higher education attainment, provincial data was divided into four regions: eastern, central, western, and northeastern, according to the economic regions defined by the National Bureau of Statistics of China and previous works (Guan et al., 2018; Li et al., 2018). The baseline fixed effects model (Equation 2) was employed while incorporating regional heterogeneity through dummy variables and interaction terms between regional dummies and all explanatory variables:

$$y_{it} = \alpha_i + \sum_{r \in C,E,NE} \gamma_r D_r + \beta_1 \ln(gdp_{it}) + \sum_{k=2}^{4} \beta_k x_{kit} + \sum_{r \in C,E,NE} \left[\beta_1^r D_r \ln(gdp_{it}) + \sum_{k=2}^{4} \beta_k^r D_r x_{kit} \right] + \epsilon_{it} \tag{3}$$

where D_r are dummy variables for the eastern, central, and northeastern regions, with their coefficients γ_r measuring baseline differences in entrants compared to the western region. The interaction terms β_1^r and β_k^r measure how the effects of indicators differ in the other regions relative to the west, which allows for testing of whether determinants of higher education enrolment vary systematically across China's economic regions. The western region was chosen as the reference for baseline comparison due to the large number of provinces that exist in the region to allow for a more robust model.

	Pooled OLS	Fixed Effects	Random Effects
Parameter Estimates			
Constant	5.149**	-	5.149**
Urbanization rate	0.954**	1.564**	1.431**
	(0.099)	(0.212)	(0.265)
Regional GDP	0.302*	0.270*	0.294
	(0.121)	(0.179)	(0.171)
Education expenditure	-0.208*	-0.358**	-0.325*
	(0.082)	(0.125)	(0.128)
Highway density	0.125*	0.089**	0.090*
	(0.060)	(0.033)	(0.037)
Model Fit Statistics			
R^2 (within)	0.555	0.587	0.586
Hausman test statistic	-	33.470**	-

Notes: Standard errors in parentheses are reported in parentheses, Driscoll-Kraay SE were used for FE and RE models. ** and * indicate significance at 1% and 5% levels, respectively.

Table 2: Screening of models: pooled OLS, fixed effects, and random effects on final selected indicators: urbanisation, expenditure, regional gdp, and highway density

Figure 1: Economic regions based on administrative boundary data (IGISMAP)

4 Results

4.1 National Effects

The relationship between predictors and entrants (Table 3) were attained from fixed effects panel regressions across three time periods: 2005-2019, and two sub-periods (2005-2012 and 2013-2019).

The dataset comprised 465 total observations, with 248 and 217 observations for the sub-periods, respectively.

Urbanisation had the strongest effect on higher education attainment, and a significant positive effect on entrants, with a coefficient of 1.564 ($p < 0.01$). This effect showed smaller magnitudes across the two separate time periods, (0.626, $p < 0.01$, 1.092, $p < 0.01$, respectively). On the contrary, increasing education expenditure showed a statistically significant decrease in entrants at a small coefficient (-0.358, $p < 0.01$) across the entire study period, and revealed statistically insignificant and varying results (-0.278, 0.094) for both 2005-2012 and 2003-2019.

Regional GDP had a statistically significant positive effect on entrants for the entire study period (0.270, $p < 0.05$) and from 2005-2012 (0.664, $p < 0.01$), with a significant but marginal and negative effect on entrants from 2013-2019 (-0.146, $p < 0.01$). Highway density increased entrants across the entire time period with a coefficient of 0.089 ($p < 0.01$). The effect was statistically significant across the two separate time period as well, with a marginal positive effect of 0.063 ($p < 0.05$) from 2005-2013 and 2013-2019 revealing the strongest effect (1.418, $p < 0.01$).

The model for the full study period explained 58.7% of the within-province variation in entrants, while the models for 2005-2012 and 2013-2019 explained 58.8% and 43.6% of with within-province variation, respectively. The robust F-statistics were highly statistically significant $p < 0.01$ for all periods, indicating overall model appropriateness and explanatory power. However, the magnitude of the R^2 (within-province) and F-statistic decreased considerably for 2013-2019, showing decline in model strength for this time period.

	2005–2019	2005–2012	2013–2019
Urbanization rate	1.564**	0.626**	1.092**
	[7.366]	[2.856]	[3.618]
	(0.212)	(0.219)	(0.302)
Regional GDP	0.270*	0.664**	-0.146**
	[2.510]	[4.496]	[2.693]
	(0.179)	(0.148)	(0.054)
Education expenditure	-0.358**	-0.278	0.094
	[2.860]	[1.142]	[0.694]
	(0.125)	(0.243)	(0.135)
Highway density	0.089**	0.063*	1.418**
	[2.693]	[2.077]	[2.717]
	(0.033)	(0.030)	(0.522)
R^2 (within)	0.587	0.538	0.436
F-statistic (robust)	266.732**	347.113**	18.200**
Observations	465	248	217

Notes: T-statistics (abs) are reported in brackets, standard errors in parentheses
* and ** represent statistical significance at 5% and 1% levels respectively.

Table 3: Regression results for drivers of entrants at national level by time period

4.2 Regional Variations

To evaluate whether factors of higher education attainment varied across the economic regions of China, the study investigated regional-specific effects through an incorporation of regional interaction terms with all explanatory variables. The west was chosen as the reference category due it having the largest number of N (provinces), and coefficients for the other regions (central, north-

east, east) were derived by a summation of their deviation to the reference region. The results (Table 4) revealed regional heterogeneity in the strengths of drivers that influence high education attainment. The model explained 67.3% of the variations in the dependent variable, with F-statistic for poolability (59.837, $p < 0.01$) confirming that the effects of explanatory variables varied meaningfully across provinces.

Urbanisation remains the strongest overall coefficient, with a significant positive effect in the west (1.134,$p < 0.01$) and central regions (2.093, $p < 0.01$). In the northeast, urbanisation had a strong negative effect (-1.406, $p < 0.01$). For the east, urbanisation had the smallest effect overall compared to the other economic regions (1.054), and the results did not prove to be significant. Per capita regional GDP had varying effects across the economic regions, with a positive effect in the west (0.601, $p < 0.01$) and east (0.551), with statistically insignificant results in the east. In both the central and northeast regions, regional GDP had a negative but significant ($p < 0.01$) effect on higher education attainment, although the magnitudes of the effects were weaker (-0.297 and -0.425, respectively).

Education expenditure showed a negative impact on higher education attainment in the west (-0.364), central (-0.900), and east (-0.631) regions, though only west had a statistically significant result. The opposite effect was seen in the northeast, where education expenditure had a statistically significant and the largest marginal effect on higher education attainment in comparison to the other variables (1.822,$p < 0.01$). Highway density had a positive effect on the west (0.604, $p < 0.05$), east (0.080, $p < 0.05$) and central regions, though the result was not statistically significant in the central region. Lastly, similar to education, the northeast saw a large, positive, and significant effect of highway density on higher education enrolment.

The results (Figure 2 from the fixed effects model with regional interaction terms showed differences in the influence of the chosen indicators on higher education attainment. Urbanisation had the strongest positive contribution for the west and central and east regions. Overall, all predictors varied in magnitude and direction for the four economic regions. It should be noted that besides highway density, none of the indicators proved to be statistical significant in the eastern region.

	West	Central	East	Northeast
Urbanization rate	1.134**	2.093**	1.054	-1.406**
	[5.508]	[3.059]	[0.341]	[3.212]
	(0.206)	(0.641)	(0.236)	(0.791)
Regional GDP	0.601**	-0.297**	0.551	-0.425**
	[3.995]	[5.711]	[0.212]	[3.459]
	(0.150)	(0.157)	(0.236)	(0.297)
Education expenditure	-0.364**	-0.900	-0.631	1.822**
	[8.303]	[1.459]	[1.367]	[4.766]
	(0.044)	(0.367)	(0.195)	(0.459)
Highway density	0.604*	0.105	0.080*	1.643**
	[1.886]	[1.369]	[1.839]	[2.671]
	(0.320)	(0.365)	(0.302)	(0.381)
Provinces	12	6	10	3
R^2 (within)	0.673			
F-stat (poolability)	59.837**			
Observations	465			

Notes: T-statistics (abs) are reported in brackets, standard errors in parentheses
* and ** represent statistical significance at 5% and 1% levels respectively
Coefficients derived by adding deviations to reference coefficient (west)

Table 4: Regression results for drivers of entrants with regional interactions

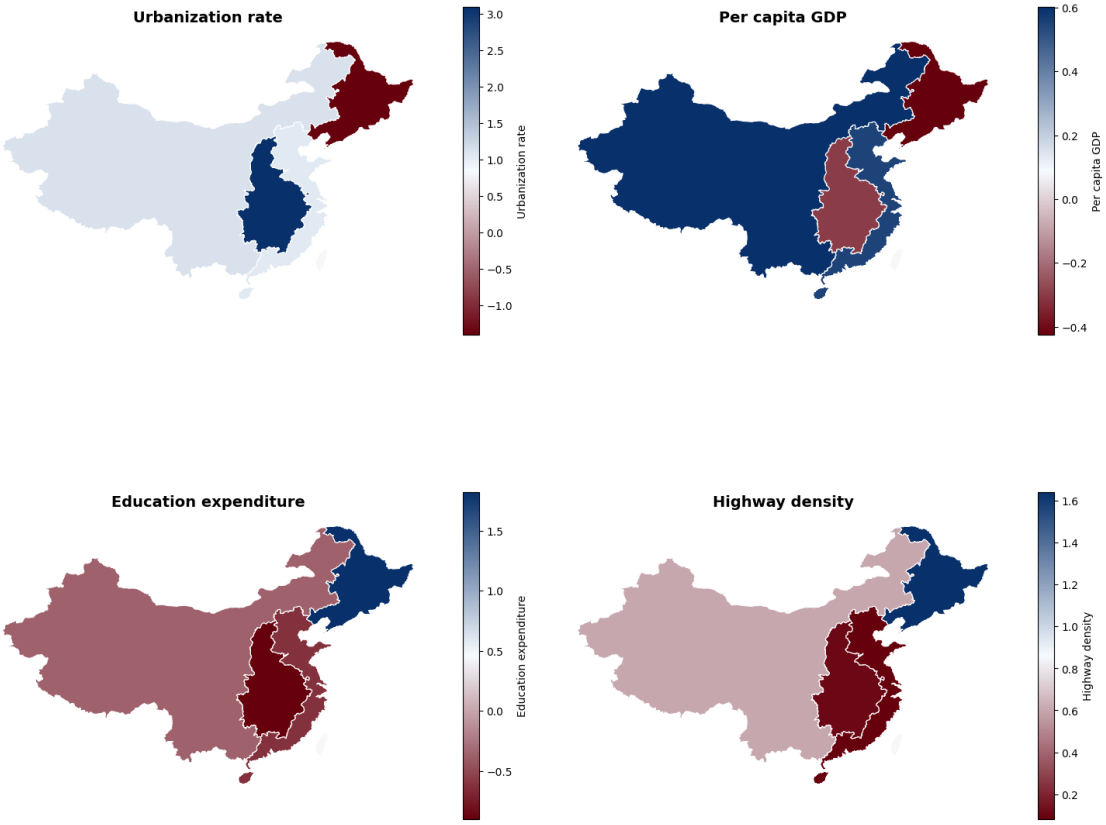

Figure 2: Explanatory variables and influence across the economic regions of China

5 Discussions

5.1 National Evaluations

The study found that urbanisation had overwhelming influence on higher education attainment over the entire study period, with one standard deviation increase in urbanisation rate in a province leading to about 1.564 more entrants (new enrolments per 1000 people for that province). Additionally, sub-period comparisons reveal that urbanisation had a larger effect on entrants in the later period. Increase in the average effect of urbanisation on higher education attainment across all provinces could be a result of Chinese leaders' attempts to bridge urban-rural divide through the Eleventh Five-Year Plan in 2006 with increased infrastructure development and ensuring equal access to education for rural students in underdeveloped regions (Fan, 2006). A similar effect can be seen in the evolution of the strength of highway density in influencing higher education attainment. The plan also outlines several initiative to improve transportation accessibility, with a focus on connecting rural areas and developing a comprehensive national transportation network(Congress, 2006). While the effect was marginal in magnitude across the entire study period and from 2005-2012, later period saw a drastic increase in the effect highway density had on entrants.

Education expenditure had a negative effect on enrolment over the entire study period and from 2005-2012, but saw shift to positive but marginal influences on higher education attainment from 2013-2019. This finding possibly reflects improvement in resource allocation efficiency in the later period. For instance, from 2006-2012, the government launched a slew of education policies to ensure strategic investments in education, particularly in accelerating the development of rural education (Jian, 2020). According to the Ministry of Education, a national enrolment project launched in 2012 attempted to advance rural students into the country's top universities by providing scholarships and accepting lower scores from the national entrance exam (University, n.d.), which could indicate policy influences shifted education expenditure in a positive direction. The results also suggest aggregated educational expenditure across all education level may not be a strong indicator of investments in higher education, thus resulting in its low explanatory power of higher education attainment. While regional GDP had a positive effect on higher education attainment from 2005-2012, the shift to a negative effect in later periods aligns with theories of diminishing returns on human capital capital investment (Cahill, 2002), suggesting that the marginal benefits of economic growth on enrolment may decrease once China reached a certain threshold of economic development.

The national level analysis for the study period and sub-periods reveal complex dynamics between indicators and education attainment in 21st-century China. The dominant influence of urbanisation, particularly in recent years, highlights the success of policy initiative aimed at reducing the urban-rural education gap. The evolution of the impact of transportation infrastructure, as operationalised by highway density, further supports the effectiveness of these development approaches in underdeveloped regions of China. Furthermore, the shifting roles of education expenditure and GDP per capita underscore the importance of effective resource allocation and consideration for diminishing returns on economic inputs. These findings also suggest favouring targeted interventions to address specific barriers of higher education attainment over analysis of broad economic or investment trends.

5.2 Regional Heterogeneity

Evaluating these factors at a regional level shows heterogeneity in all metrics of the study. Historically, west and central regions of China were predominantly rural (Fang et al., 2024). Urbanisation having the strongest influence in entrants within these regions and compared to the east and northeast aligns with the theory that a rise in urbanisation improves human capital formation. The diminishing return of human capital investment is evident in regional comparisons, as west China has the lowest economic performance out of all regions (Quan, 1991). The positive effect of GDP indicates that west China's education outcomes still benefit significantly from economic growth, while negative coefficients of GDP in central and northeast China indicate diminished returns.

Transportation accessibility has a stronger influence in the west and northeast, indicating that mobility has a higher impact on education outcomes in these regions. Besides the northeast, education expenditure appears to have a negative effect across all regions, suggesting that the impacts of education initiatives in later years are not apparent when evaluating regions across the entire study period. It should be noted that the northeast consistently shows larger standard errors across all variables compared to other regions. The higher uncertainty in the estimates suggests that the limited number of observations (45) were insufficient to capture stable patterns in this region. The east shows remarkably weak statistical significance across all variables. This indicates that the development drivers of already established coastal provinces may differ and the lack of significant results makes it difficult to draw reliable conclusions about the factors that influence higher education attainment in this region. Although, the statistical significance of highway density in the east reveals that the bridging of urban and rural areas proves to be a reliable indicator of higher education attainment.

The study aimed to explore regional variations in the factors that influence education outcomes. While the analysis provided insights into some aspects of how different drivers affect higher education attainment in these regions, insufficient data hinders the reliability of some of these results. Additionally, the interaction term structure of the regression means all coefficients represent deviations from west, which could obscure the absolute effects in each region. Granular and longer explorations of these factors within regions through a more robust model could provide better insights when assisting policymakers with targeted interventions.

5.3 Limitations

This study is not without its limitations. While the fixed effects model provides a robust analysis for drivers of higher education attainment and their strengths at a national level, the interaction term structure of the regional analysis could have introduce multicollinearity and reduced the reliability of the model. Significant differences in the number of observations made for each region further limits the explanatory power of the model, and a longer study period may improve the explanatory power of the model for regional variations. Additionally, future works with more robust models, such as Geographically and Temporally Weighted Regression (Fotheringham et al., 2015), can ensure a more dynamic evaluation of the drivers that influence higher education attainment. The study evaluates national and regional drivers of higher education attainment, but the factors were not considered at a provincial level. Constant observations of the independent variables lack consideration for urban-rural divides within provinces, which can lead to misleading results as coefficients fail to consider intra-region variations. Recommendations for future works

include more granular analysis of these indicators through provincial level analysis, which can assist policymakers in making more targeted interventions.

6 Conclusion

Using panel data and a fixed effects regression, the study analyses factors that influence enrolment in China at a national level and regional level, identifying the main drivers of higher education attainment while controlling time invariant omitted variables. Additionally, evaluating the strength of these indicators over the entire study period and across two sub-periods allows for a more granular analysis of how these determinants have evolved over time, particularly in relation to policies that may have caused structural changed in China's socioeconomic and education framework. Regional level comparisons revealed heterogeneity in the strengths of the drivers of entrants, demonstrating that policy interventions should consider regional contexts. These findings contribute to the evaluation of drivers on higher education attainment in rapidly developing nations, offering empirical evidence of how urbanisation, economic development, transportation access and education infrastructure come together to shape education outcomes at both national and regional levels. Distinct effects across China's economic regions and temporal stages highlight the importance of considering local contexts and development stages when formulating higher education policy.

References

Agasisti, T., & Bertoletti, A. (2022). Higher education and economic growth: A longitudinal study of european regions 2000–2017. *Socio-Economic Planning Sciences*, *81*, 100940.

Arouri, M. E. H., Youssef, A. B., Nguyen-Viet, C., & Soucat, A. (2014). Effects of urbanization on economic growth and human capital formation in africa. Retrieved from `https://shs.hal.science/halshs-01068271` (⟨halshs-01068271⟩)

Arshed, N., ur Rehman, H., Nazim, M., & Saher, A. (2021). Evading law of diminishing returns, a case of human capital development. *The journal of contemporary issues in business and government*, *27*(5), 2569–2584.

Bertinelli, L., & Zou, B. (2008). Does urbanization foster human capital accumulation? *The Journal of Developing Areas*, 171–184.

Cahill, M. B. (2002). Diminishing returns to gdp and the human development index. *Applied economics letters*, *9*(13), 885–887.

Chen, Z., Li, Y., & Wang, P. (2020). Transportation accessibility and regional growth in the greater bay area of china. *Transportation Research Part D: Transport and Environment*, *86*, 102453.

Congress, N. P. (1986). *Compulsory education law of the people's republic of china.* National People's Congress of the People's Republic of China.

Congress, N. P. (2006). *Guidelines of the eleventh five-year plan for national economic and social development.* National People's Congress of the People's Republic of China.

Fan, C. C. (2006). China's eleventh five-year plan (2006-2010): from" getting rich first" to" common prosperity". *Eurasian geography and economics*, *47*(6), 708–723.

Fang, C., Chen, Z., Liao, X., Sun, B., & Meng, L. (2024). Urban-rural digitalization evolves from divide to inclusion: empirical evidence from china. *npj Urban Sustainability*, *4*(1), 51.

Fotheringham, A. S., Crespo, R., & Yao, J. (2015). Geographical and temporal weighted regression (gtwr). *Geographical Analysis*, *47*(4), 431–452.

Fu, Q., & Ren, Q. (2010). Educational inequality under china's rural–urban divide: The hukou system and return to education. *Environment and Planning a*, *42*(3), 592–610.

Galor, O. (2011). Inequality, human capital formation, and the process of development. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 4, pp. 441–493). Elsevier. Retrieved from `https://www.sciencedirect.com/science/article/pii/B9780444534446000055` doi: 10.1016/B978-0-444-53444-6.00005-5

Gaokao.com. (2020). *2020 list of universities in beijing.* Retrieved from `https://www.gaokao.com/`

Gemmell, N. (1995). Endogenous growth, the solow model and human capital. *Economics of Planning*, *28*, 169–183.

Gordon, R. A. (2015). *Regression analysis for the social sciences*. Routledge.

Guan, X., Wei, H., Lu, S., Dai, Q., & Su, H. (2018). Assessment on the urbanization strategy in china: Achievements, challenges and reflections. *Habitat International*, *71*, 97–109.

Gurley, T., & Bruce, D. (2005). The effects of car access on employment outcomes for welfare recipients. *Journal of urban Economics*, *58*(2), 250–272.

Haddad, W. D., Carnoy, M., Rinaldi, R., & Regel, O. (1995). *Education and development: Evidence for new priorities* (World Bank Discussion Papers). Washington, DC: World Bank.

Hannum, E. (1999). Political change and the urban-rural gap in basic education in china, 1949-1990. *Comparative education review*, *43*(2), 193–211.

Hanushek, E. A., & Woessmann, L. (2008, 09). The role of cognitive skills in economic development. *Journal of Economic Literature*, *46*(3), 607-668. Retrieved from `https://www.proquest.com/scholarly-journals/role-cognitive-skills-economic-development/docview/213262842/se-2`

Hoechle, D. (2007). Robust standard errors for panel regressions with cross-sectional dependence. *The stata journal*, *7*(3), 281–312.

Jian, L. (2020). Compulsory educational policies in rural china since 1978: A macro perspective. *Beijing International Review of Education*, *2*(1), 159–164.

Kawabata, M. (2002). Job access and work among autoless adults in welfare in los angeles.

Lange, F., & Topel, R. (2006). The social value of education and human capital. *Handbook of the Economics of Education*, *1*, 459–509.

Li, M., He, B., Guo, R., Li, Y., Chen, Y., & Fan, Y. (2018). Study on population distribution pattern at the county level of china. *Sustainability*, *10*(10), 3598.

Lim, S. S., Updike, R. L., Kaldjian, A. S., Barber, R. M., Cowling, K., York, H., ... others (2018). Measuring human capital: a systematic analysis of 195 countries and territories, 1990–2016. *The Lancet*, *392*(10154), 1217–1234.

Marquez-Ramos, L., & Mourelle, E. (2019). Education and economic growth: an empirical analysis of nonlinearities. *Applied Economic Analysis*, *27*(79), 21–45.

McGranahan, G., & Satterthwaite, D. (2014). *Urbanisation concepts and trends* (Vol. 220). JSTOR.

Mok, K. H., Wong, Y. C., & Zhang, X. (2009). When marketisation and privatisation clash with socialist ideals: Educational inequality in urban china. *International Journal of Educational Development*, *29*(5), 505–512.

Molnar, M., Wang, B., & Gao, R. (2015). Assessing china's skills gap and inequalities in education.

National Bureau of Statistics of China. (2005–2019). *China Statistical Yearbook*. Beijing, China: China Statistics Press. Retrieved from `http://www.stats.gov.cn/tjsj/ndsj/` (Accessed: 2024-12-07)

Pastor, J. M., Peraita, C., Serrano, L., & Soler, Á. (2018). Higher education institutions, economic growth and gdp per capita in european union countries. *European Planning Studies*, *26*(8), 1616–1637.

Paudel, R. C. (2023). Capital expenditure and economic growth: A disaggregated analysis for nepal. *Cogent Economics & Finance*, *11*(1), 2191449.

Pink-Harper, S. A. (2015). Educational attainment: An examination of its impact on regional economic growth. *Economic Development Quarterly*, *29*(2), 167–179.

Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to investment in education: a decennial review of the global literature. *Education Economics*, *26*(5), 445–458.

Qian, X., & Smyth, R. (2008). Measuring regional inequality of education in china: widening coast–inland gap or widening rural–urban gap? *Journal of international development: The journal of the development studies association*, *20*(2), 132–144.

Quan, Z. X. (1991). Urbanisation in china. *Urban Studies*, 41–51.

Schultz, T. P. (1992, June). *The role of education and human capital in economic development: An empirical assessment* (Tech. Rep. No. IPR38). Washington, D.C.: Institute for Policy Reform. (Economic Growth Center, Yale University, and Senior Research Fellow, Institute for Policy Reform)

Schultz, T. W. (1971). *Investment in human capital: The role of education and of research*. New York: Free Press. (Print)

Shaw, S.-L., Fang, Z., Lu, S., & Tao, R. (2014). Impacts of high speed rail on railroad network accessibility in china. *Journal of Transport Geography*, *40*, 112–122.

UN, D. (2013). World population prospects: the 2012 revision. *UN Department of Economic and Social Affairs*.

UNESCO. (n.d.). *The need to know about higher education.* Retrieved from `https://www.unesco.org/en/higher-education/need-know` (Accessed: 2024-12-15)

University, T. (n.d.). *Rural education given bright future.* Retrieved from `https://www.tsinghua.edu.cn/en/info/1320/11704.htm` (Accessed: 2025-1-15)

Vila, L. E., Cabrer, B., & Pavía, J. M. (2014). On the relationship between knowledge creation and economic performance. *Technological and Economic Development of Economy*, *21*(4), 539–556. Retrieved from `https://doi.org/10.3846/20294913.2013.876687` doi: 10.3846/20294913.2013.876687

Wu, X. (2024). The household registration system and rural-urban educational inequality in contemporary china. In *Understanding inequality in china* (pp. 35–55). Routledge.

Yang, J., Huang, X., & Liu, X. (2014). An analysis of education inequality in china. *International Journal of Educational Development*, *37*, 2–10.

Yang, J., & Qiu, M. (2016). The impact of education on income inequality and intergenerational mobility. *China Economic Review*, *37*, 110–125.

Yeoh, E. K., & Chu, K. M. (2012). Literacy, education and economic development in contemporary china. *SSRN Electronic Journal*, *2*(1), 11–83. Retrieved from `https://doi.org/10.2139/ssrn.2207559` doi: 10.2139/ssrn.2207559

Zhang, X., & Kanbur, R. (2009). Spatial inequality in education and health care in china. In *Regional inequality in china* (pp. 92–110). Routledge.

Zhang, X., & Zhang, K. H. (2003). How does globalisation affect regional inequality within a developing country? evidence from china. *Journal of Development Studies*, *39*(4), 47–67.

A Initial indicators and their definitions

Variable	**Implication**	**Description**
Entrants	New enrolment in higher education institutions by region	Number of entrants / total population × 100,000
GDP per capita	Per capita Gross Domestic Product by region	Regional GDP / total population of the region (rmb/person)
Unemployment Rate	Unemployment rate by region	Unemployment rate by region in urban areas
Education Expenditure	Per capita education expenditure by region	Regional education expenditure / total population of the region
Highway Density	Highway accessibility by region	Total length of highways / regional land area (km/km^2)
Railway Density	Public transportation accessibility by region	Length of railways in operation / regional land area
Population Density	Population distribution by region	Population of the region / land area
Urbanisation	Regional urbanisation rate	Ratio of urban population to total population of the region

Table 5: Initial indicators and explanation of their derivation

B Visualisation of GDP transformation

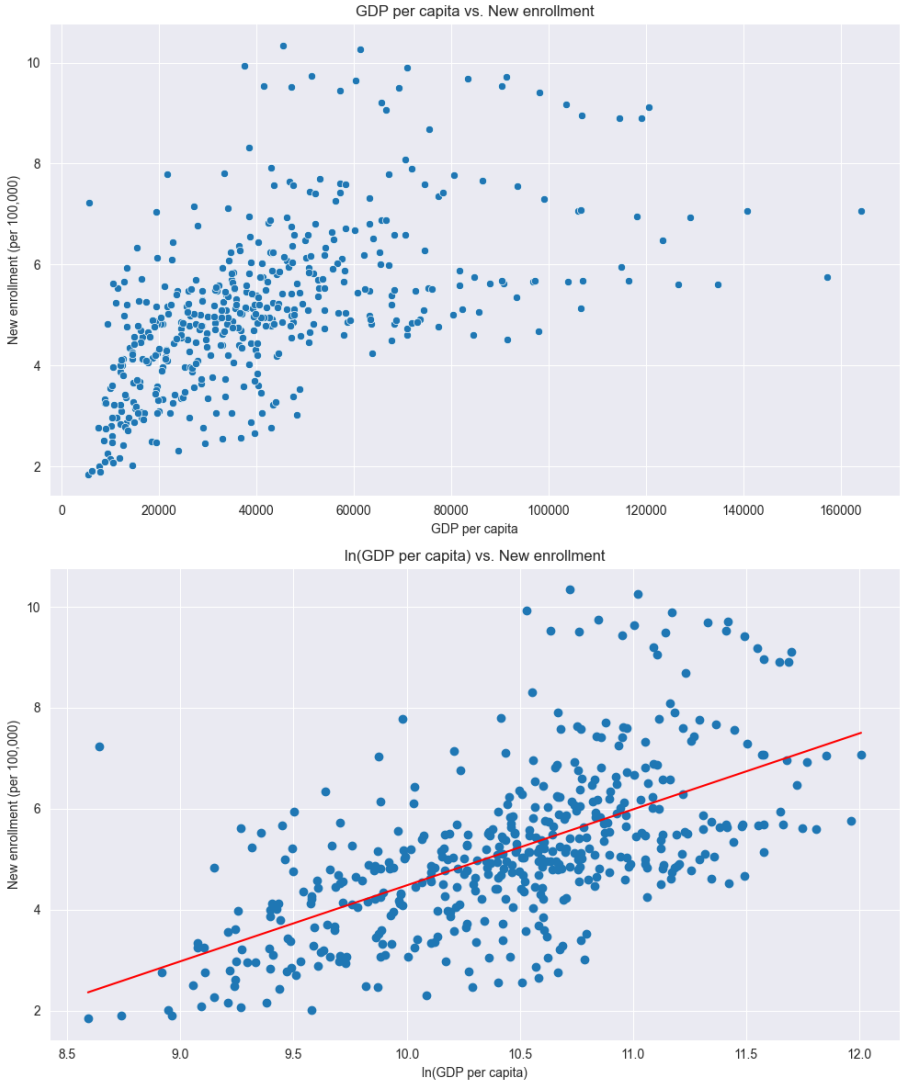

Figure 3: GDP per capita vs. Entrants (Raw and Log-transformed)

C Visual inspection of residuals from baseline model

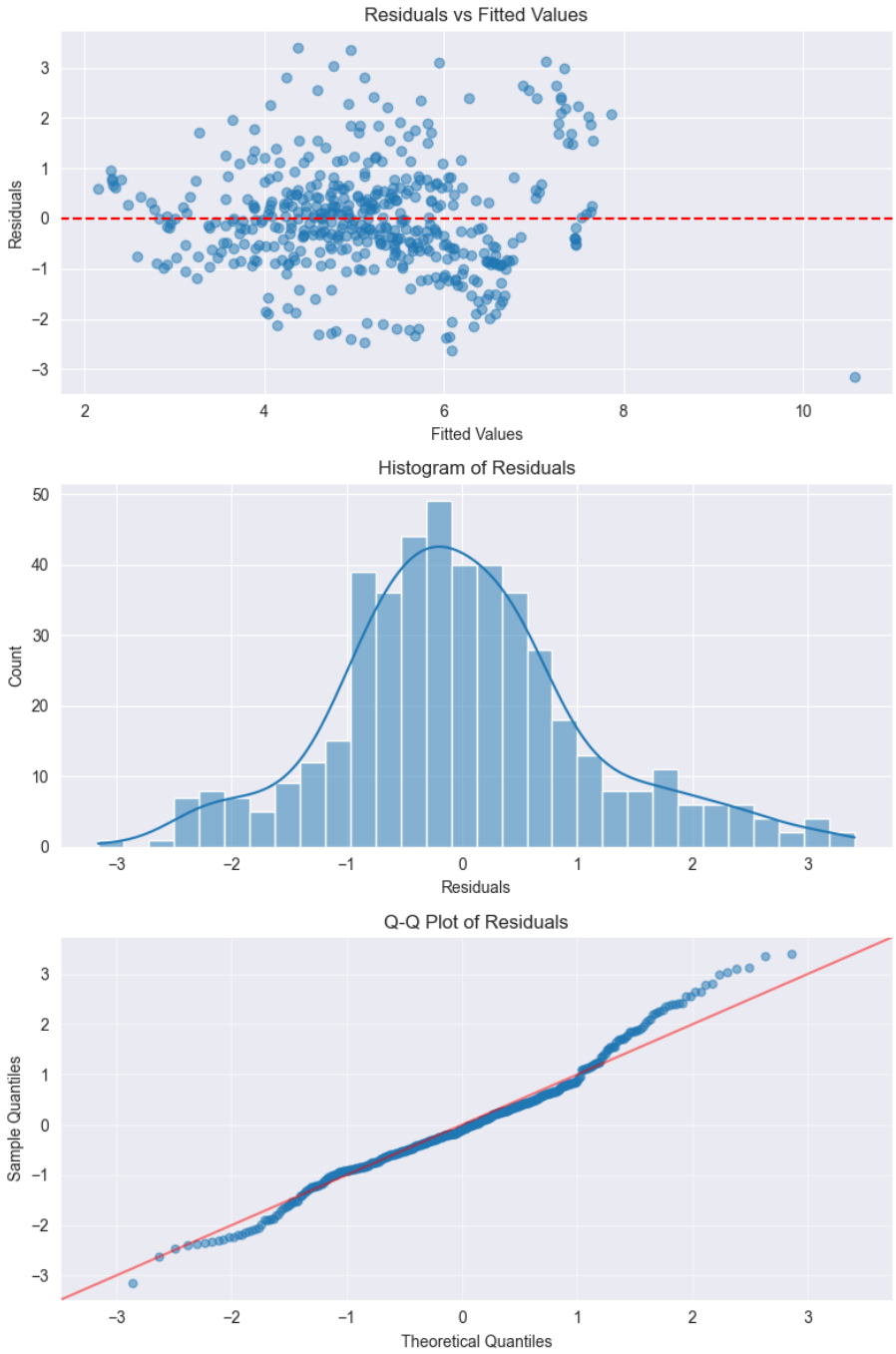

Figure 4: Visualisations of residuals

D Model assumptions test results

Test	**Statistic**	**P-value**	**Implication**
White	105.724	0.000	Heteroskedasticity
Breusch-Godfrey	310.078	2.099×10^{-68}	Serial Correlation
Pesaran CD	13.59	0.000	Cross-sectional Dependence

Table 6: Test results of classical assumptions from Pooled OLS model