Higher Education and Regional Innovation

——FE Analysis of China's Provincial Data

Kexin Cai

School of Mathematics and Statistics, Shandong University, 180 Wenhua West Road, Weihai, Shandong, 264209, China

E-mail:ckx_sdu@163.com

Abstract. With the growing interest in regional innovation capacity and higher education, educational equity (especially the urban-rural education gap) has emerged as a key factor affecting innovation output. Although studies have examined the impact of higher education on innovation, less attention has been paid to the moderating role of the urban-rural education gap. In this paper, we adopt a fixed-effects model to study the relationship between higher education and regional innovation capacity from a regional perspective using panel data of 31 provinces in China from 2003 to 2022, focusing on the moderating role of the urban-rural education gap. The results show that higher education has a significant positive impact on regional innovation capacity, especially in regions with concentrated educational resources. However, the urban-rural education gap significantly weakens this impact, especially in rural areas where resources are scarce.

Further analyses show that increased financial investment in education and higher tertiary education intensity have a significant contribution to innovation. Meanwhile, the moderating effect of the urban-rural education gap is heterogeneous across regions. This paper ensures the reliability of the results through endogeneity and robustness tests. This study provides new empirical evidence and highlights the need to address the urban-rural education gap. Policymakers should ensure equitable distribution of educational resources, especially in less developed regions, to promote innovation and sustainable development.

Keywords: Higher education; Urban-rural educational disparities; Regional innovation capacity; Fixed effects model; Panel data analysis

1 Introduction

Enhancing regional innovation capacity is one of the core objectives of China's high-quality development, and higher education is regarded as an important driver due to its role in manpower supply and knowledge transformation. However, higher education resources are concentrated in the eastern region, with obvious shortcomings in the central and western regions[3]; urban-rural education gaps (e.g., enrolment, educa-

H. Kannan et al. (eds.), *Proceedings of the 5th International Conference on Internet, Education and Information Technology (IEIT 2025)*, Atlantis Highlights in Social Sciences, Education and Humanities 40,
https://doi.org/10.2991/978-94-6463-803-5_56

tion funding, education quality) also constrain the accumulation of human capital and diffusion of innovation[7].

The mechanisms by which higher education promotes innovation show differences across regions, and urban-rural education inequality may weaken knowledge spillovers from HEIs[2]. Existing studies have mostly focused on the direct effect of HE on innovation and less on the moderating effect of urban-rural differences.

In this paper, based on panel data from 31 provinces in China from 2003 to 2022, we use a fixed-effects model to systematically analyse the impact of higher education on regional innovation and introduce urban-rural differences in education as a moderating variable. The empirical results find that higher education significantly enhances innovation capacity, but the promotion effect is weakened in regions with large urban-rural differences. Meanwhile, education financial investment and university density also have important positive effects on innovation. This study reveals the heterogeneous roles of finance and HEI agglomeration and emphasises the key significance of educational equity in driving innovation.

2 Literature Review

In the context of knowledge economy, higher education is widely recognised as the core driving force to enhance regional innovation capacity. As centres of human capital cultivation and knowledge production, higher education supports regional innovation systems through talent supply, technology transformation and knowledge spillover mechanisms[4]. It has been shown that higher education is significantly and positively associated with technological innovation, especially in human capital-intensive and urban agglomeration regions[1]. On the one hand, higher education improves the overall quality of regional human capital and enhances firms' R&D and technology absorption capabilities; on the other hand, universities create significant knowledge spillovers in geographic neighbourhoods through the diffusion of research results[5][6].

However, the role of higher education in promoting innovation shows significant heterogeneity across regions. Developed regions have more significant knowledge diffusion effects due to abundant research resources and close cooperation between universities and enterprises, while less developed regions such as central and western regions have insufficient investment in research resources and limited ability of enterprises to absorb innovations despite the growth in the number of universities[4]. In recent years, the role of colleges and universities in leading low-carbon development and sustainable technology transition has also become increasingly prominent in the field of green innovation.

This regional heterogeneity is further exacerbated by urban-rural education differences. Educational resources are concentrated in the east coast and urban areas, while rural areas have long suffered from a lack of both investment and quality[3]. Differences in educational opportunities not only inhibit human capital accumulation in rural areas but also reduce the resilience and scalability of regional innovation systems by limiting knowledge flows. Research has shown that regions with higher levels of educational development and more balanced structures are more likely to have sustained knowledge cycles and active innovation ecosystems. International experience also reveals that urban-rural educational disparities interact through multiple dimensions, including income level, social class, and gender, to create complex patterns of spatial inequality [2][7]. For example, empirical studies in both India and Latin America have shown that despite the expansion of overall higher education, disparities in educational opportunities between urban and rural, rich and poor groups have not narrowed.

Rural-urban differences in education are an important factor affecting the contribution of higher education to regional innovation. Most of the existing studies focus on the direct impact of higher education on innovation, but less on the moderating role of urban-rural differences. Based on the panel data of 31 provinces in China from 2003-2022, this paper investigates the changes in the marginal effects of urban-rural differences in education on the impact of higher education under different levels of educational differences. It is found that tertiary education can significantly enhance regional innovation capacity, but the impact of tertiary education on regional innovation capacity is significantly weaker in areas with large urban-rural differences. The innovation of this paper lies in the systematic introduction of urban-rural education differences as a moderating variable, which expands the research horizon of higher education and innovation capacity and provides a new theoretical basis and empirical support for optimising the allocation of education resources and narrowing the innovation gap.

3 Data and Methodology

3.1 Data

This paper selects data on higher education and innovation capacity of 31 provinces in mainland China from 2003 to 2022, mainly from China Statistical Yearbook and China Education Statistical Yearbook, covering variables such as the number of college graduates, the number of colleges and universities, education expenditure and the number of years of education in urban and rural areas. To eliminate the effects of differences in scale and extreme values, all core variables are standardised, and missing values are filled in by linear interpolation. Based on the differences in economic fundamentals, educational resources and policy environments, and in conjunction with the national regional development strategy, the 31 provinces are divided into seven regions: the Bohai Rim, the Yangtze River Delta, the Pearl River Delta, Central, Southwest, Northwest and Northeast. The detailed classification of provinces under each region is presented in Table 1.

Table 1. Regional division

Region	Province
Bohai Rim region	Beijing, Tianjin, Hebei, Shandong
Yangtze River Delta region	Shanghai, Jiangsu, Zhejiang, Anhui
Pearl River Delta region	Guangdong, Fujian, Hainan
Central region	Shanxi, Jiangxi, Henan, Hubei, Hunan, Guangxi

Southwest region	Chongqing, Sichuan, Guizhou, Yunnan, Tibet
Northwest region	Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang, Inner Mongolia
Northeast region	Liaoning, Jilin, Heilongjiang

3.2 Variables

In this paper, the regional innovation capability (INN) is used as the explanatory variable, and the comprehensive utility value is used to measure the actual output level and development potential of innovation activities in each province. The indicator system covers five first-level indicators: knowledge creation, knowledge acquisition, enterprise innovation, innovation environment and innovation performance. Specifically, they include independent R&D capability, external knowledge introduction and industry-academia-research collaboration, technology application and product development, institutional and resource protection for innovation, and the promotion effect of innovation on the economy and society. The weights are set with reference to the literature and based on the expert assignment method [8][9][10], a comprehensive indicator system with the following weights is constructed, as shown in Table 2:

Table 2. Symbol description

Level 1 indicators	Attributes	Weights
Knowledge creation	+	0.15
Knowledge acquisition	+	0.15
Enterprise Innovation	+	0.25
Innovation Environment	+	0.25
Innovation performance	+	0.20

The core explanatory variable of this paper is the number of college graduates (GRA), which is used to measure the scale of regional higher education human capital. The main control variables include:

Fiscal investment in education (ASS), defined as the ratio of regional education expenditure to total fiscal expenditure:

$$ASS_{i,t} = \frac{\text{Education\_Expenditure}_{i,t}}{\text{Fiscal\_Expenditure}_{i,t}}$$

Both components are measured in 100 million RMB, and this variable reflects the intensity of government's financial support to education and measures the material basis of education resources.

Higher education intensity (SCH), defined as the number of colleges and universities in a region divided by the total population (expressed per million people):

$$SCH_{i,t} = \frac{\text{Number\_of\_schools}_{i,t}}{\text{Population}_{i,t}}$$

This variable reflects the regional distribution of higher education resources and accessibility.

Urban-rural education gap (DIF), defined as the difference between urban and rural residents' years of education and reflects the equity of education in the region.

Table 3 provides the summary statistics of the core explanatory variables and control variables. The data show that there are significant differences in the innovation capacity of provinces, the number of graduates fluctuates greatly, and the distribution of educational resources is uneven. Developed regions are densely populated with colleges and universities and have outstanding resource advantages, while urban-rural education differences are still evident in some regions, and the balance of education resources needs to be further optimised.

Table 3. Summary statistics

Variables	Obs	Mean	SD	Min	Median	Max
INN	140	29.90	7.88	18.98	26.79	46.15
GRA	140	19.60	9.17	2.83	19.71	44.80
ASS	140	0.24	0.05	0.16	0.24	0.41
SCH	140	1.39	0.80	0.26	1.20	4.10
DIF	140	2.19	0.35	1.38	2.16	3.25

Table 4 demonstrates the correlations between the explanatory variables. The data indicates that the correlation between these variables is low, which suggests that potential covariance problems within the model are minimal.

Table 4. Results of correlation analysis

Variables	(1)	(2)	(3)	(4)	(5)
(1)INN	1.000***	0.453***	0.426***	0.173**	-0.383***
(2)GRA		1.000***	-0.215**	0.419***	-0.175**
(3)ASS			1.000***	-0.300***	-0.536***
(4)SCH				1.000***	0.095
(5)DIF					1.000***

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.3 Model

In this study, the fixed effects model is used to analyze the panel data for two main reasons: on the one hand, the fixed effects model is able to control for unobservable heterogeneity factors at the provincial level that do not change over time during the study period, such as historical development paths, institutional environments, geographic locations, and cultural traditions, etc., which may affect both educational inputs and innovation outputs, and which, if left unchecked, may lead to biased model estimation. On the other hand, from an empirical point of view, this paper conducted the Hausman test on the fixed-effects and random-effects models, and the results rejected the original hypothesis that the individual effects are not correlated with the explanatory variables ($p < 0.05$), indicating that the fixed-effects model is more suitable for the data structure of this study, and it can provide more consistent estimation results. Therefore, this paper finally adopts a two-way fixed-effects model, introducing area and time fixed effects to improve the robustness and explanatory power of the estimation results.

Based on the above considerations, this paper constructs the following panel regression model:

$$INN_{i,t} = \alpha_0 + \alpha_1 GRA_{i,t} + \alpha_2 \sum Col_{i,t} + Province_i + Year_t + \varepsilon_{i,t}$$

Where $INN_{i,t}$ is the utility value of innovation capacity in region i in period t, $GRA_{i,t}$ refers to the number of college graduates in region i in period t, and $Col_{i,t}$ refers to control variables such as ratio of education expenditure to fiscal expenditure in region i in period t. α_i is a coefficient to be estimated, $Province_i$ is the individual fixed effect, $Year_t$ is the time fixed effect, and $\varepsilon_{i,t}$ is the error term.

4 Empirical Results

4.1 Regression Results

In the benchmark regressions, this paper uses stepwise regression to control for province fixed effects and time effects. Table 5 reports the stepwise regression results, confirming that higher education significantly enhances regional innovation, while urban-rural disparities exert a negative moderating effect. Column (1) examines only the effect of the number of higher education graduates (GRA) on regional innovation capacity (INN) without introducing other control variables. The regression results show that there is a significant positive correlation between GRA and INN, indicating that the growth of higher education human capital can effectively promote regional innovation output. After the introduction of education expenditure share (ASS) in Column (2), the regression coefficient of GRA further rises and remains significant, and ASS also exhibits positive and significant, indicating that financial investment has a facilitating effect on innovation. Columns (3) and (4) further incorporate the variables of density of colleges and universities (SCH) and rural-urban educational differences (DIF), respectively, and the results show that SCH exhibits a significant positive effect on innovation, suggesting that the clustering of colleges and universities contributes to the diffusion of knowledge and collaboration of innovations, whereas the coefficient of DIF is negative and significant at the 5% level, suggesting that the rural-urban differences inhibit innovation to a certain extent, but its effect is relatively limited. Overall, GRA is consistently significant across models, validating the central role of higher education in human capital expansion and knowledge diffusion. The model fit R^2 improved from 0.073 to 0.342, and the introduction of control variables significantly enhanced the explanatory power of the model.

The above results suggest that higher education fosters regional innovation by enhancing human capital and knowledge flow, while financial investment and university resource allocation offer institutional support for a favorable innovation environment.

However, urban-rural educational disparities hinder this effect, showing that educational equity is vital to fully unlock regional innovation potential.

Table 5. Regression result

Variables	(1) INN	(2) INN	(3) INN	(4) INN
GRA	0.237*** (0.082)	0.344*** (0.079)	0.410*** (0.079)	0.404*** (0.073)
ASS		0.313*** (0.070)	0.332*** (0.065)	0.327*** (0.066)
SCH			0.247*** (0.056)	0.249*** (0.056)
DIF				-0.076** (0.033)
Province	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
N	140	140	140	140
R^2	0.073	0.213	0.307	0.342

Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.2 Endogeneity Test

The relationship between the number of college graduates and regional innovation capacity may have potential endogeneity problems, such as reverse causality or omitted variable bias. To mitigate these problems, this paper constructs a lagged term regression model using the number of college graduates lagged one period as an instrumental variable. Table 6 reports the results of the endogeneity analysis, with columns (1) to (4) presenting the regression results with the gradual introduction of control variables, respectively. Compared to the baseline model, the lagged first-order value of the core variable, the number of college graduates (GRA), is positive in all models and is significant at either the 5 % or 1 % significance level, suggesting that the use of the instrumental variable effectively mitigates the potential endogeneity problem.

Table 6. Endogeneity test

Variables	(1) INN	(2) INN	(3) INN	(4) INN
GRA	0.202** (0.090)	0.297*** (0.080)	0.360*** (0.084)	0.354*** (0.076)
ASS		0.275*** (0.075)	0.313*** (0.067)	0.320*** (0.069)
SCH			0.246*** (0.061)	0.253*** (0.061)

DIF				-0.078** (0.038)
Province Year	Yes Yes	Yes Yes	Yes Yes	Yes Yes
N R^2	133 0.051	133 0.150	133 0.247	133 0.283

Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.3 Moderating Effect

To test whether the urban-rural education difference moderates the impact of higher education on regional innovation capacity, this paper constructs an interaction term between the number of college graduates (GRA) and the urban-rural education difference (DIF). Where DIF represents the difference in years of education between urban and rural residents, the larger value represents the more unbalanced allocation of educational resources within the region, reflecting the degree of structural educational inequality. The moderated effects regression model used is as follows:

$$INN_{i,t} = \alpha_0 + \alpha_1 GRA_{i,t} + \alpha_2 DIF_{i,t} + \alpha_3 (GRA_{i,t} \times DIF_{i,t}) + \alpha_4 \sum Col_{i,t} + Province_i + Year_t + \varepsilon_{i,t}$$

The regression results are detailed in Table 7, which show that the number of higher education graduates (GRA) has a significant positive effect on regional innovation capacity (INN), indicating that higher education plays a positive role in enhancing regional innovation. After introducing the urban-rural education difference (DIF), its coefficient is -0.079, which is significant at the 5% level, indicating that the widening of education inequality has a dampening effect on innovation capacity. After further introducing the interaction term, the coefficient of the interaction term between GRA and DIF is -0.076, which is also significant at the 5% level and has a negative relationship, suggesting that urban-rural educational disparities significantly weaken the role of higher education in promoting regional innovation. That is, in regions with unbalanced allocation of educational resources, the human capital created by higher education is difficult to be effectively diffused and utilised within the region, thus weakening the innovation spillover effect.

After further introduction of control variables such as education expenditure share (ASS) and higher education density (SCH), the coefficient of the interaction term remains significantly negative, and all the control variables show positive and significant effects, indicating that financial investment and higher education agglomeration help to enhance regional innovation capacity, and the results have good robustness.

In summary, urban-rural education differences play a significant moderating role in the process of higher education influencing regional innovation. Optimising the allocation of regional educational resources and enhancing the accessibility of education in rural and less developed areas will help to further unleash the potential driving force of higher education on regional innovation.

Table 7. Moderating effect

Variables	(1) INN	(2) INN	(3) INN	(4) INN
GRA	0.237*** (0.082)	0.232*** (0.076)	0.123 (0.099)	0.310*** (0.096)
DIF		-0.079** (0.039)	-0.109** (0.043)	0.100*** (0.035)
GRA*DIF			-0.076** (0.037)	0.062** (0.031)
ASS				0.334*** (0.061)
SCH				0.218*** (0.058)
Province	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
N	140	140	140	140
R^2	0.959	0.961	0.963	0.972

Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5 Conclusion

The study found that the urban-rural education gap significantly reduces the role of higher education in promoting regional innovation, especially in rural areas, where the lack of higher education resources limits the innovation effect. To achieve balanced development of regional innovation, it is necessary to optimise the allocation of educational resources and enhance educational equity. For policymakers, they should increase investment in education in rural and less developed areas, narrow the gap between urban and rural areas, and fully unleash the potential of higher education to promote innovation. Local governments should combine local resources and formulate differentiated policies to promote the flow of talents and the concentration of innovation resources and improve the regional innovation system. Enterprises should pay attention to the quality of tertiary education resources in their regions, especially in areas with weak resources, and strengthen university-enterprise cooperation to promote technological innovation and enhance enterprise competitiveness.

Future research can further explore the role of different types of higher education institutions in regional innovation and analyse the long-term impact of educational resource allocation on innovation in the context of policy scenarios.

References

1. Cowan, R., & Zinovyeva, N. (2013). University effects on regional innovation. *Research Policy*, 42(3), 788–800.
2. España-Eljaiek, I., Mejía, D., & Gaviria Melendez, N. S. (2023). Reproducing inequality? Elite-biased policy and the rural–urban educational gap in Colombia and its regions, 1920–2019. *Revista de Historia Económica*, 41(3), 375–409.
3. Guo, Y., & Li, X. (2024). Regional inequality in China's educational development: An urban–rural comparison. *Heliyon*, 10, e26249.
4. Li, H., & Liu, S. (2021). Higher education, technological innovation, and regional sustainable development: Insights from a VAR model. *Discrete Dynamics in Nature and Society*, 2021, Article ID 8434528.
5. Peng, C., & Xu, H. (2024). Universities and cities: The impact of higher education on urban innovation, entrepreneurship, and economic growth. *Asian Economic Papers*, 23(2), 33–56.
6. Schenkenhofer, J., Block, J., & Vismara, S. (2024). University knowledge spillovers and innovation of hidden champions: Evidence from Italy. *R&D Management*, 55(2), 598–614.
7. Zahl-Thanem, A., & Rye, J. F. (2024). Spatial inequality in higher education: A growing urban–rural educational gap? *European Sociological Review*, 40(6), 1067–1081.
8. Xu, N. (2013). The promoting effect of executive equity incentives on R&D investment in high-tech firms: An empirical study from a nonlinear perspective. Science of Science and Management of S&T, 34(2), 12–19.
9. Zhou, X., Bai, J., & Wu, J. (2010). Valuation and pricing of high-tech companies listed on the Growth Enterprise Market. Statistics and Information Forum, 25(11), 46–50.
10. Zhou, X., Lu, L., & Tang, J. (2010). IPO pricing of high-tech companies based on EVA. Economic Management Research, (2), 53–57.