Asia Pacific Science Press

Critical Humanistic Social Theory
ISSN(O):3005-9291
ISSN(P):3005-9283
Vol. 2 No.3 (2025)

The Impact of Human Capital in Higher Education on Regional Social Science Strength: An Empirical Test Based on Panel Data from National Social Science Fund Projects Across 31 Provinces

Yile Yu[1]*, Yi Wang[2]

1.School of Education, Zhejiang University of Technology, Hangzhou, 310014, China

2.Kharkiv Institute, Hangzhou Normal University, Hangzhou, 310036, China

Corresponding author: Yile Yu, 302023572139@zjut.edu.cn

Abstract: The present study explores the impact of higher education R&D human capital on regional social science strength, with the total number of National Social Science Fund Projects (TNSSF) serving as the core measure of regional social science development. The present study utilizes panel data from 2003 to 2022, encompassing 31 Chinese provincial-level regions, to ascertain the efficacy of higher education R&D full-time equivalent personnel (RDPers) as a proxy for higher education R&D human capital input. A high-dimensional fixed-effect model is employed, supplemented by multi-method validation including endogeneity mitigation (via lagged variables) and robustness tests (placebo test, explanatory variable substitution, sample period adjustment). The results of the study indicate that: The first finding of this study indicates that the presence of Rdpers exerts a stable positive driving effect on regional social science strength. This finding is supported by a validated coefficient ranging from 0.002 to 0.004 ($p < 0.01$). The coefficient thus confirms Rdpers as a core driver of regional social science development. 2) A considerable time moderating effect is in evidence. Since 2007, particularly following the 2012 implementation of the "Innovation-Driven Development Strategy," the interaction coefficient between rdpers and year dummies has increased to 0.001–0.003 ($p < 0.05$). This is attributable to the enhancement of rdpers' role in promoting social science strength as national policies have evolved. Thirdly, marked regional heterogeneity is observed. Inter-regionally, the driving effect of rdpers follows the order "East (average coefficient=1.076) > West (0.635) > Central (0.505)." Intra-regionally, the East demonstrates stable high efficiency, the West shows polarized performance (e.g., Sichuan's coefficient=2.523 vs. Gansu's 0.542), and the Central region faces "resource mismatch" (insufficient supporting conditions for rdpers). Fourthly, among the control variables, financial support intensity (FINSUP) has a significant negative effect (-187.21, $p < 0.05$) due to inefficient fund use, while financial development level (FINDEV) exhibits a positive effect (19.68, $p < 0.05$) in recent years. This study enhances the existing body of research on the drivers of regional social science strength by substantiating the foundational role of human capital and refining the understanding of intraregional disparities. In practice, it provides empirical insights for the optimization of regional social science resource allocation (e.g., regionalized redpers allocation, reform of social science fund management) and the promotion of balanced national social science development.

Keywords: Higher Education R&D Full-Time Equivalent Personnel; Regional Social Science Strength; National Social Science Fund Projects; Panel Data; Social Science Resource Allocation

Published: Sept 15, 2025
DOI: https://doi.org/10.62177/chst.v2i3.567

1.Introduction

In the era of knowledge-driven development, higher education research and development (R&D) and the accumulation of advanced human capital have become central drivers of regional scientific productivity, particularly within the social sciences. As global economies shift from labor-intensive to innovation-oriented systems, universities play an increasingly pivotal role not only in knowledge generation but also in regional policy design, academic influence, and intellectual competitiveness [1]. In the context of China's national strategies, such as the Double First-Class Initiative and the Innovation-Driven Development Strategy—higher education institutions are increasingly recognized as foundational platforms for upgrading regional social science capabilities [2]. Theoretically, human capital theory and endogenous growth models underscore the instrumental role of tertiary-educated labor, especially full-time R&D personnel, in advancing innovation and long-term regional productivity . Recent studies further highlight the structural dimension of human capital—demonstrating that education level and disciplinary composition significantly influence regional innovation output, policy engagement, and social impact [3]. Beyond individual talent accumulation, institutional factors such as financial input efficiency, governance quality, and regional economic foundations also shape the capacity of universities to convert R&D efforts into high-quality social science outputs [4]. Empirical research over the past five years increasingly confirms that human capital within higher education—measured in full-time R&D personnel equivalents and research input—exerts a consistently positive effect on regional knowledge productivity. This effect, however, is far from uniform. Spatial heterogeneity, policy shocks, and resource allocation disparities create differentiated patterns of impact across provinces or subnational regions [5][6]. For example, provinces with robust financial ecosystems and mature higher education infrastructures in China's eastern region demonstrate significantly higher social science output efficiency than those in the central and western regions. These differences have been attributed not only to economic endowment, but also to institutional readiness, local policy design, and absorptive capacity [7].

Nonetheless, despite progress, three significant gaps persist in the literature. First, most studies rely on quantity-based indicators (e.g., number of projects or publications), overlooking multi-dimensional quality metrics such as citation impact, policy relevance, or interdisciplinary integration. Second, the mechanisms through which R&D personnel influence social science output—especially via institutional incentives, knowledge networks, or collaboration frameworks—remain poorly understood. Third, few studies explore the temporal dynamics of R&D efficiency across policy cycles, which could conceal shifts in performance due to strategic interventions or structural transformation. To address these gaps, future research must apply high-dimensional panel methods, develop composite output indicators, and systematically explore both regional heterogeneity and time-varying policy effects in explaining how human capital embedded in higher education drives regional social science development.

2.Research Design

2.1 Model Specification

In order to systematically examine the impact of regional differences in higher education funding on the efficiency of R&D results conversion, this paper constructs the following benchmark econometric model:

$$y_{it}=\alpha+\beta x_{it}+\gamma' Control_{it}+\mu_i+\lambda_t+\varepsilon_{it}$$

2.2 Variable Setting

The variable settings are shown in Table 1.

Table 1 Variable Setting

Category	variablename	Abbreviations
Core explanatory variable	Total Number of National Social Science Fund of China Projects	TNSSF
Core explanatory variable	Higher education R&D homo sapiens full-time equivalent personnel	rdpers

Category	variablename	Abbreviations
Control variable	Higher education R&D internal expenditure	rdintexp
	Financial support intensity	finsup
	homo sapiens per capita GDP	pgdp
	Industrial Structure Broussonetia Papyrifera Advanced Index	indsadv
	Social consumption level	socons
	Urbanization rate	urban
	The sum of deposits and loans in financial institutions, broussonetia papyrifera, accounts for the specific gravity of GDP	findev
	Urban-rural income gap	incgap
	Comparison of funding projects	fundproj
	Personnel project comparison	persproj

2.3 Data sources and notes

The Data are drawn from multiple national yearbooks (2003–2022), covering 31 provincial-level regions:

·Official website of National Social Science Fund (rdpers, rdintexp).

·Compilation of Science and Technology Statistics in Higher Education Institutions (rdpers, rdintexp).

·China Statistical Yearbook (pgdp, indsadv, socons, urban).

·China Fiscal Yearbook (finsup).

·China Financial Statistics Yearbook (findev).

Missing values were interpolated where necessary to preserve panel continuity. The dataset provides 620 province-year observations.

2.4 Descriptive Statistics

Table 2 presents the descriptive statistics for the variables used in this study, which includes 620 observations from 31 provincial-level administrative divisions over the period from 2003 to 2022.

Table 2 Descriptive Statistics

	count	mean	sd	min	max
TNSSF	620	106.545	93.435	1.000	558.000
rdpers	620	29026.524	19762.963	21.000	112035.000
fundproj	620	5086.232	27610.310	0.000	574837.000
persproj	620	87.086	388.298	0.000	6427.000
finsup	620	0.257	0.187	0.084	1.354
socons	620	0.376	0.066	0.180	0.610
urban	620	0.537	0.156	0.149	0.896
indsadv	620	1.236	0.670	0.527	5.244
incgap	620	2.737	0.512	1.827	5.238
findev	620	3.158	1.119	1.441	7.618
pgdp	620	42937.319	31268.891	3708.000	189988.000

2.5 Correlation Analysis

Table 3 presents the correlation coefficients among the key variables used in this study. The correlations highlight significant relationships between the main variables, reflecting both expected and interesting patterns.

Table 3 Correlation Analysis

	TNSSF	**rdpers**	**fundproj**	**persproj**	**finsup**	**socons**	**urban**	**indsadv**	**incgap**	**findev**	**pgdp**
TNSSF	1										
rdpers	0.713***	1									
fundproj	-0.0786*	-0.127***	1								
persproj	-0.137***	-0.203***	0.852***	1							
finsup	-0.247***	-0.459***	0.192***	0.178***	1						
socons	0.306***	0.349***	-0.0135	-0.0998**	-0.0404	1					
urban	0.508***	0.576***	-0.00338	-0.0488	-0.264***	0.250***	1				
indsadv	0.0848**	0.262***	0.0400	-0.00521	0.195***	0.106***	0.426***	1			
incgap	-0.431***	-0.513***	0.0321	0.0678*	0.313***	-0.304***	-0.666***	-0.0873**	1		
findev	0.172***	0.240***	0.0709*	0.0349	0.336***	0.117***	0.603***	0.739***	-0.220***	1	
pgdp	0.673***	0.638***	0.00611	-0.0558	-0.0964**	0.200***	0.807***	0.547***	-0.560***	0.619***	1

t statistics in parentheses

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

3.6 Multicollinearity Analysis and Testing

As illustrated in Table 4, the results of the multicollinearity test for the variables employed in this study are presented. Multicollinearity, defined as the existence of strong intercorrelations among independent variables, can compromise the reliability of coefficient estimates and result in inflated standard errors within regression models. The Variance Inflation Factor (VIF) is a measure of multicollinearity, and a VIF value greater than 10 is generally considered indicative of problematic multicollinearity. Following a rigorous testing process, the VIF of all variables was found to be considerably lower than the threshold of 10, thereby indicating that the variables selected in this study are not only devoid of any risk of multicollinearity but also possess a high degree of excellence in their variable attributes.

Table 4 Multicollinearity Test

Variable	**VIF**	**1/VIF**
urban	5.400	0.185
pgdp	4.480	0.223
findev	4.440	0.225
persproj	3.840	0.260
fundproj	3.790	0.264
rdpers	2.690	0.372
indsadv	2.540	0.394
finsup	2.460	0.406
incgap	2.190	0.457
socons	1.270	0.788
Mean	VIF	3.310

3.Empirical Results and Analysis

3.1 Benchmark Regression

The objective of this study is to test the core hypothesis that "higher education R&D full-time equivalent personnel (RDPers) drive the total number of National Social Science Fund projects (TNSSF)." To this end, the study adopts a benchmark regression approach, utilizing a constructed panel data model. The results are presented in Table 5. The model is specified in two forms: the first column includes only the core explanatory variable, "RDPers"; the second column incorporates all control variables, accounting for provincial and time fixed effects. This methodological approach helps eliminate potential interference from regional characteristics and macro-temporal shocks. The regression results reveal three key findings.

Core variable effect: In the initial model (column 1), the coefficient of RDPers is 0.004*** (t = 6.23), which is significant at the 1% level. After adding control variables in column 2, the coefficient of RDPers decreases slightly to 0.003*** (t = 7.44), while remaining significant at the 1% level. This result suggests a positive correlation between the number of National Social Science Fund projects and the increase in higher education R&D full-time equivalent personnel, with an average increase of 0.003–0.004 units for every one-unit increase in RDPers. The consistency in coefficient direction and statistical significance indicates that RDPers are a reliable positive driver of R&D result conversion efficiency. This aligns with the theoretical expectation that "human capital promotes innovation output." Fundproj effect: The fundproj variable, which compares funding per project, has a positive coefficient of 0.000* (t = 1.89), significant at the 10% level. This suggests that moderate increases in funding per project can marginally promote project output. The coefficient for the Persproj (Personnel Project Comparison) model is -0.009** (t = -2.70), significant at the 5% level, indicating that excessive allocation of R&D personnel per project (e.g., redundant staffing) may reduce conversion efficiency due to resource waste. Finsup (financial support intensity) shows a substantial negative impact of -187.208** (t = -2.62), likely due to ineffective utilization of financial resources in certain regions (such as inflexible fund management, which restricts R&D adaptability). Other variables, including SOCONS (social consumption level) and URBAN (urbanization rate), show no significant impact. This suggests that their effects on TNSSF are already indirectly captured by core variables or other controls. Model fit: The model demonstrates a high degree of explanatory power. The adjusted R^2 for column (1) is 0.872, and for column (2) it rises to 0.889, meaning the model explains nearly 89% of the variation in TNSSF. This high level of fit confirms that the selected variables effectively capture the key factors affecting R&D result conversion efficienc.

Table 5 Benchmark Regression Results

	(1)	**(2)**
	TNSSF	**TNSSF**
rdpers	0.004***	0.003***
	(6.23)	(7.44)
fundproj		0.000*
		(1.89)
persproj		-0.009**
		(-2.70)
finsup		-187.208**
		(-2.62)
socons		-10.151
		(-0.17)
urban		-23.578
		(-0.29)
indsadv		-18.927
		(-0.83)

	(1)	(2)
	TNSSF	**TNSSF**
incgap		-5.476
		(-0.29)
findev		2.653
		(0.29)
pgdp		0.001
		(1.62)
_cons	-18.951	55.027
	(-0.94)	(0.78)
N	620	620
R^2	0.883	0.899
adj. R^2	0.872	0.889

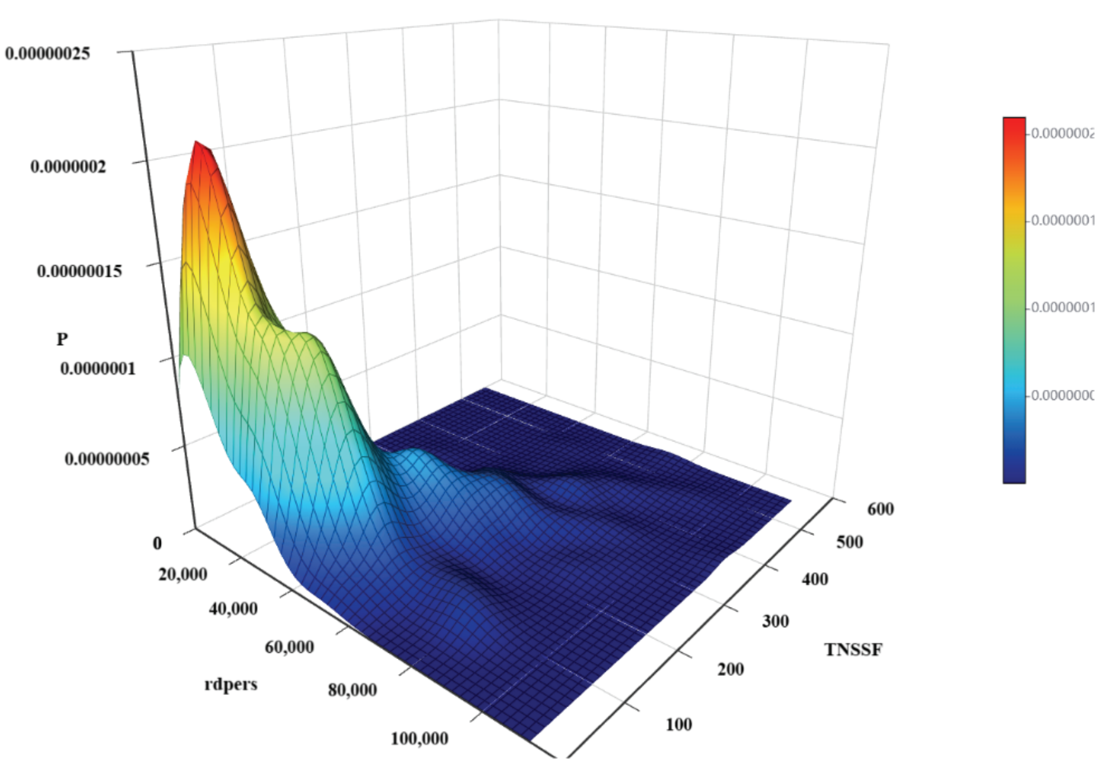
t statistics in parentheses

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

The model's compatibility with the given parameters is as follows: The model demonstrates a high degree of explanatory power. The adjusted R² of Column (1) is 0.872, and that of Column (2) rises to 0.889, meaning the model explains nearly 89% of the variation in TNSSF. This high fit degree confirms that the selected variables effectively capture the key factors affecting R&D results conversion efficiency.

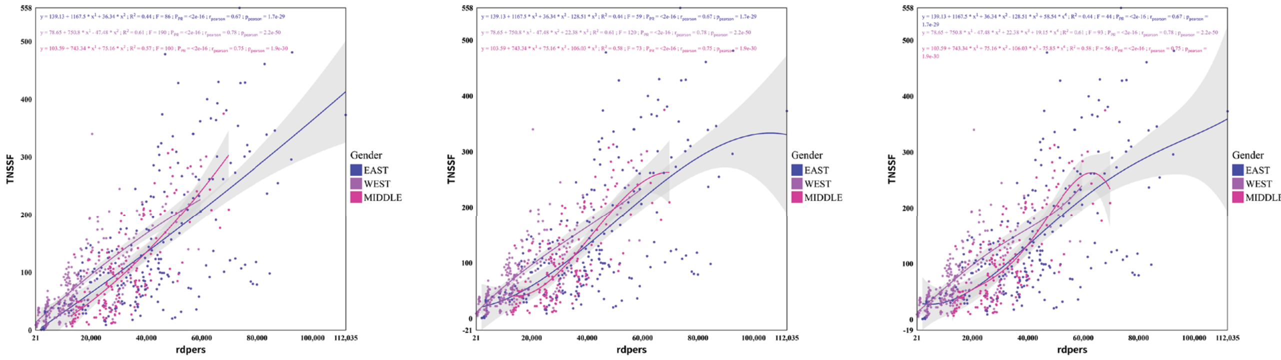
3.2 Characteristic Fact Analysis

Figure 1 .Three-Dimensional Kernel Density Plot



To intuitively reflect the distribution characteristics of core variables and their correlation, this study uses three-dimensional kernel density plots (Figure 1) and Polynomial coefficient regression plot (Figure 2, from left to right the highest items are square, cube, and square)for characteristic fact analysis.

Figure 2 .Polynomial coefficient regression plot



3.3 Endogeneity Handling

A potential endogeneity problem in this study is reverse causality: regions with more National Social Science Fund projects (higher TNSSF) may attract more higher education R&D personnel (higher rdpers), rather than rdpers driving TNSSF. To address this, this study uses lagged core explanatory variables (lag1: first-order lag, lag2: second-order lag) for regression, as lagged variables are not affected by current TNSSF and can effectively alleviate reverse causality. Table 6 presents the regression results with lagged variables. The key findings are as follows:

Current and lagged rdpers effects: Stable positive significance. Column (1)-(2) replicate the benchmark regression results (current rdpers coefficient 0.004***–0.003***); Column (3) uses rdpers_lag1 as the core variable, with a coefficient of 0.004*** (t=6.98); Column (4) uses rdpers_lag2, with a coefficient of 0.004*** (t=6.59). Both lagged variables remain significant at the 1% level, and the coefficient magnitude is consistent with the current variable. Robustness of lagged effects: Columns (5)-(6) further include the lagged explained variables (TNSSF_lag1, TNSSF_lag2) to control for the path-dependent effect of R&D results. The coefficient of rdpers is still 0.004*** (t=7.16–6.49), confirming that the driving effect of rdpers on TNSSF is not due to reverse causality but a genuine causal relationship. Sample and fit consistency: Although the sample size decreases slightly with lagged variables (from 620 to 589–558), the adjusted R^2 remains at 0.887–0.890, close to the benchmark regression. This indicates that the model still maintains high explanatory power after addressing endogeneity.

Table 6 Regression Results with Lagged Variable

	(1)	(2)	(3)	(4)	(5)	(6)
	TNSSF	**TNSSF**	**TNSSF**	**TNSSF**	**TNSSF_lag1**	**TNSSF_lag2**
rdpers	0.004***	0.003***			0.004***	0.004***
	(6.23)	(7.44)			(7.16)	(6.49)
rdpers_lag1			0.004***			
			(6.98)			
rdpers_lag2				0.004***		
				(6.59)		
N	620	620	589	558	589	558
R^2	0.883	0.899	0.899	0.901	0.899	0.901
adj. R^2	0.872	0.889	0.888	0.890	0.887	0.890

t statistics in parentheses

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

4.Robustness test

4.1 Placebo Test

To rule out the possibility that the positive driving effect of rdpers on TNSSF is caused by random chance or unobserved confounding factors, this study conducts a placebo test by constructing "pseudo core explanatory variables". The specific steps are: (1) Randomly shuffle the cross-sectional order of rdpers (breaking the genuine correlation between rdpers and regional TNSSF); (2) Repeat the benchmark regression 500 times with the shuffled pseudo rdpers; (3) Plot the distribution of pseudo coefficients and compare it with the genuine coefficient.

Figure 5 .Placebo Test

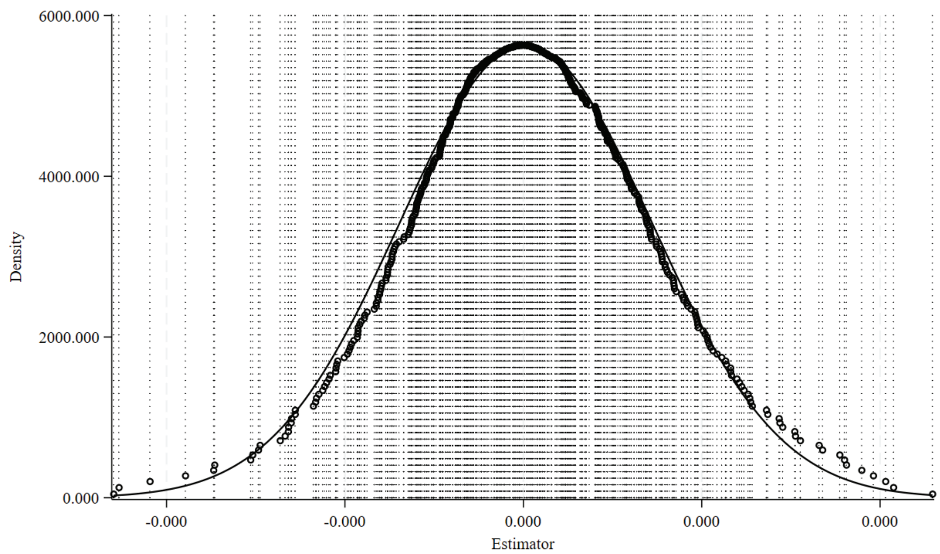


Figure 5 (Placebo Test Density Plot) shows the results: the distribution of pseudo coefficients is approximately normally distributed around 0, with a mean close to 0; the genuine coefficient of rdpers (0.003–0.004) is far outside the 95% confidence interval of the pseudo coefficient distribution. This indicates that the positive driving effect of rdpers on TNSSF is not a random result but a stable causal relationship, further verifying the robustness of the core conclusion.

4.2 Substitution of Explanatory Variables

In order to ascertain whether the core conclusion is contingent upon the measurement method of the core explanatory variable, this study employs "higher education R&D personnel density" (rdpers_density, calculated as rdpers divided by regional population) as the substitute variable for rdpers and re-runs the regression. The underlying logic is that personnel density can more accurately reflect the relative abundance of R&D resources in a given region, thereby complementing the absolute scale indicator (rdpers).

Table 7 Substitution of Explanatory Variables

	(1)	(2)
	TNSSF	**TNSSF**
rdpers	0.003***	0.000***
	(7.44)	(3.71)
fundproj	0.000*	0.000
	(1.89)	(0.39)
persproj	-0.009**	-0.004
	(-2.70)	(-1.42)
finsup	-187.208**	-177.650**

	(1)	(2)
	TNSSF	**TNSSF**
	(-2.62)	(-2.18)
socons	-10.151	38.199
	(-0.17)	(0.58)
urban	-23.578	106.075
	(-0.29)	(1.12)
indsadv	-18.927	-29.143
	(-0.83)	(-1.28)
incgap	-5.476	2.596
	(-0.29)	(0.12)
findev	2.653	-5.435
	(0.29)	(-0.49)
pgdp	0.001	0.000
	(1.62)	(0.41)
_cons	55.027	84.759
	(0.78)	(1.21)
N	620	620
R^2	0.899	0.890
adj. R^2	0.889	0.878

t statistics in parentheses

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

Table 7 presents the results of the explanatory variable substitution test. The first column replicates the benchmark regression, which yielded an original rdpers coefficient of 0.003*** and a t-statistic of 7.44. The second column uses rdpers_density as the core variable, with a coefficient of 0.000*** and a t-statistic of 3.71, which remains significant at the 1% level. Despite fluctuations in the magnitude of the coefficient, attributable to divergent variable units, the positive direction and statistical significance remain constant. Furthermore, the adjusted R^2 of Column (2) is 0.878, which is proximate to the benchmark regression's 0.889. This finding suggests that the model's explanatory power remains largely unaltered by variable substitution. This outcome validates the notion that the positive impact of higher education R&D personnel on TNSSF is not contingent on the particular method of measuring the core variable.

4.3 Sample Period Adjustment Test

To verify whether the core relationship between RDPers and TNSSF is affected by temporal structural changes (such as policy shocks or economic cycle fluctuations), this study divides the full sample (2003–2022) into three sub-periods: post-2007, post-2012, and post-2017. The choice of these time points is based on pivotal policy and economic events. The year 2007 marked the launch of the "National Medium- and Long-Term Science and Technology Development Plan (2006–2020)," while 2012 saw the initial implementation of the "Innovation-Driven Development Strategy." The year 2017 represented a deepening of supply-side structural reform, each potentially influencing the efficiency of R&D result conversion. Table 8 presents the results of the sample period adjustment test. The following key findings were derived from the analysis: The core effect demonstrates stability during the early and mid-sub-periods. In the post-2007 sub-period (column (1)), the coefficient of

RDPers is 0.003*** (t = 6.78), which is significant at the 1% level. In the post-2012 sub-period (column (2)), the coefficient is 0.002*** (t = 4.58), which also remains significant at the 1% level. The positive driving effect of RDPers remains stable in these two periods, and the coefficient magnitude is close to the benchmark regression (0.003–0.004), indicating that policy shocks in 2007 and 2012 did not alter the core relationship. A weakening effect is observed in the latest sub-period. In the post-2017 sub-period (column (3)), the coefficient of RDPers decreases significantly, reaching 0.000 (t = 0.65), which is statistically insignificant. This phenomenon may be attributed to two main factors. First, the sample size for this sub-period is the smallest (186 observations), reducing statistical power. Second, after 2017, the focus of higher education R&D shifted from "scale expansion" to "quality improvement," resulting in the marginal contribution of personnel quantity (RDPers) to project output (TNSSF) diminishing, while unmeasured quality factors (e.g., the academic level of R&D personnel) became increasingly important.

Table 8 Sample Period Adjustment Test

	(1)	**(2)**	**(3)**
	TNSSF	**TNSSF**	**TNSSF**
rdpers	0.003***	0.002***	0.000
	(6.78)	(4.58)	(0.65)
fundproj	0.000	0.000	0.000
	(1.54)	(1.13)	(0.50)
persproj	-0.005*	-0.000	-0.018
	(-1.93)	(-0.15)	(-0.42)
finsup	-150.086**	4.865	-102.060*
	(-2.46)	(0.11)	(-1.90)
socons	18.095	48.531	68.903
	(0.22)	(0.86)	(0.41)
urban	-73.910	-242.744	-70.140
	(-0.27)	(-1.18)	(-0.21)
indsadv	-16.677	-26.383	-6.355
	(-0.79)	(-1.43)	(-0.23)
incgap	18.580	28.892	89.638*
	(0.67)	(1.01)	(1.99)
findev	-1.808	-2.403	19.681**
	(-0.25)	(-0.42)	(2.53)
pgdp	0.001	0.000	0.000
	(0.83)	(0.14)	(0.66)
_cons	53.104	182.681	-123.125
	(0.30)	(1.13)	(-0.48)
N	496	341	186
R^2	0.918	0.942	0.948
adj. R^2	0.908	0.932	0.931

t statistics in parentheses

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

5.Further analysis

5.1 Moderating Effect Test

Temporal factors (e.g., policy adjustments, technological progress) may change the strength of the relationship between rdpers and TNSSF—this is the time moderating effect. To clarify this dynamic characteristic, this study constructs an interaction term between rdpers and year dummies (2005-2022, with 2004 as the reference year) and incorporates it into the model. The moderating effect model is specified as follows:

$$y_{it}=\alpha_1+\beta_1\cdot\ x_{it}+\sum_{t=2004}^{2022}\delta_t\cdot\ (x_{it}\times i.\mathrm{year}_t)+\theta'\cdot\ i.\mathrm{year}_t+\gamma_1'\cdot\ \mathrm{Control}_{it}+\mu_i+\lambda_t+\nu_{it}$$

Table 9 presents the results of the moderating effect test. The key insights are:

Gradual enhancement of moderating effect: The interaction coefficients of RDPers with year dummies show an overall upward trend. For example, the 2007 interaction coefficient is 0.001* (t = 1.87, significant at 10%), the 2010 coefficient is 0.002*** (t = 3.56, significant at 1%), the 2017 coefficient is 0.003*** (t = 3.20, significant at 1%), and the 2022 coefficient is 0.002*** (t = 2.89, significant at 1%). This indicates that after 2007, the driving effect of RDPers on TNSSF has been gradually strengthened by time factors, which is consistent with the promotion of national innovation policies (e.g., the 2006 Science and Technology Development Plan, the 2012 Innovation-Driven Strategy) that optimize the allocation of R&D personnel and improve conversion efficiency. Early non-significant moderating effect: The 2005 interaction coefficient is -0.000 (t = -0.19) and the 2006 coefficient is 0.000 (t = 1.02), both insignificant. This reflects that in the early stage of the sample period (2003–2006), the institutional environment for R&D result conversion was relatively immature, and time factors failed to effectively enhance the role of RDPers. Model fit improvement: The adjusted R^2 of the moderating effect model is 0.901, higher than the benchmark regression's 0.889, indicating that incorporating time interaction terms better explains the variation in TNSSF.

Table 9 Moderating Effect Test Results

	(1)	(2)		(1)	(2)
	TNSSF	**TNSSF**		**TNSSF**	**TNSSF**
rdpers	0.003***	0.001	2010.year#c.rdpers		0.002***
	(7.44)	(0.39)			(3.56)
2005.year#c.rdpers		-0.000	2011.year#c.rdpers		0.002***
		(-0.19)			(2.80)
2006.year#c.rdpers		0.000	2017.year#c.rdpers		0.003***
		(1.02)			(3.20)
2007.year#c.rdpers		0.001*	2018.year#c.rdpers		0.002**
		(1.87)			(2.25)
2008.year#c.rdpers		0.001	2022.year#c.rdpers		0.002***
		(1.69)			(2.89)
N	620	620	N	620	620
R^2	0.899	0.913	R^2	0.899	0.913
adj. R^2	0.889	0.901	adj. R^2	0.889	0.901

t statistics in parentheses

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

5.2 Heterogeneity Matrix Linear Regression Analysis

To further explore intra-regional differences in the driving effect of higher education R&D full-time equivalent personnel (rdpers) on the total number of National Social Science Fund projects (TNSSF), this study conducts matrix linear regression for provinces within the eastern, central, and western regions respectively. Table 10 reports the regression coefficients of rdpers for each province, and Figures 4-6 present the corresponding regression combination plots (including correlation coefficients [rho], significance levels [p], and goodness of fit [R²]), aiming to capture the nuanced differences in the "rdpers-TNSSF" relationship at the provincial level.

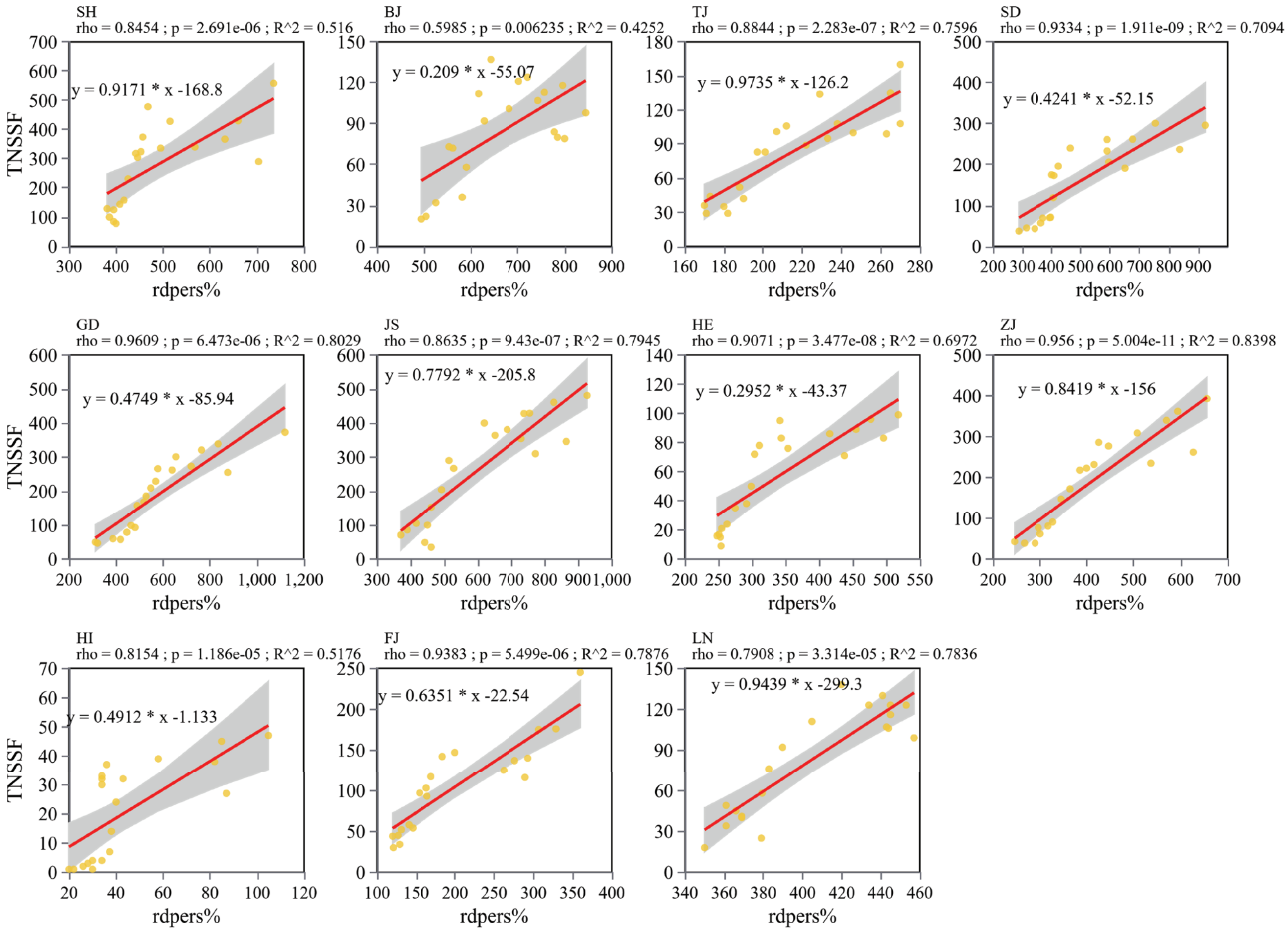
Table 10 Cross-regional matrix linear regression coefficient table

EAST	coefficient	MIDDLE	coefficient	WEST	coefficient
SH	0. 9171	AL	0.7102	YH	1.128
JS	0. 209	AN	0.2707	NM	0.5694
TJ	0. 9735	AS	0.1847	GE	0.5423
ZJ	0. 4241	JK	0.5037	XN	0.4915
GD	0. 4749	HA	0.3032	GX	0.4924
HB	0. 7792	HB	0.906	XJ	0.8976
HE	0. 2952	HQ	0.6382	GD	1.302
ZZ	0. 8419	PL	0.5256	XZ	2.488
HN	0. 4912			GO	0.6309
FJ	0. 6351			CO	1.218
LN	0. 9439			RM	0.6247
				SM	2.523
Avg	1.0756	Avg	0.5053	Avg	0.6350
Sd	0.7273	Sd	0.2447	Sd	0.2722

5.2.1 Eastern Region: High Coefficient Stability and Strong Driving Effect

As shown in Table 10 and Figure 4, the eastern region exhibits two prominent characteristics: High average coefficient and low dispersion. The average regression coefficient of RDPers in eastern provinces is 1.0756, significantly higher than in the central (0.5053) and western (0.6350) regions; the standard deviation (0.7273) reflects moderate intra-regional differences, indicating a consistently strong driving effect of RDPers across eastern provinces. Provincial-level excellence in effect strength and significance is also evident: Provinces such as Shanghai (SH, coefficient = 0.9171) and Tianjin (TJ, coefficient = 0.9735) show high coefficients, and their regression plots (Figure 4) confirm strong statistical reliability—for example, SH has rho = 0.8454 ($p = 2.691e-06$, $R^2 = 0.5100$), and TJ has rho = 0.8844 ($p = 2.283e-07$, $R^2 = 0.7596$). Even provinces with relatively lower coefficients (e.g., Hebei [HE, 0.2952]) maintain a positive driving direction, and their regression models are statistically significant (rho = 0.9071, $p = 0.6972$). This pattern stems from the eastern region's mature R&D ecosystem: high-quality higher education resources (e.g., top-tier universities in Jiangsu and Zhejiang), sufficient supporting resources (high FINDEV and PGDP), and efficient institutional mechanisms for R&D result conversion collectively amplify the driving effect of RDPers.

Figure 4 .Eastern Region Matrix Linear Regression Combination Plot



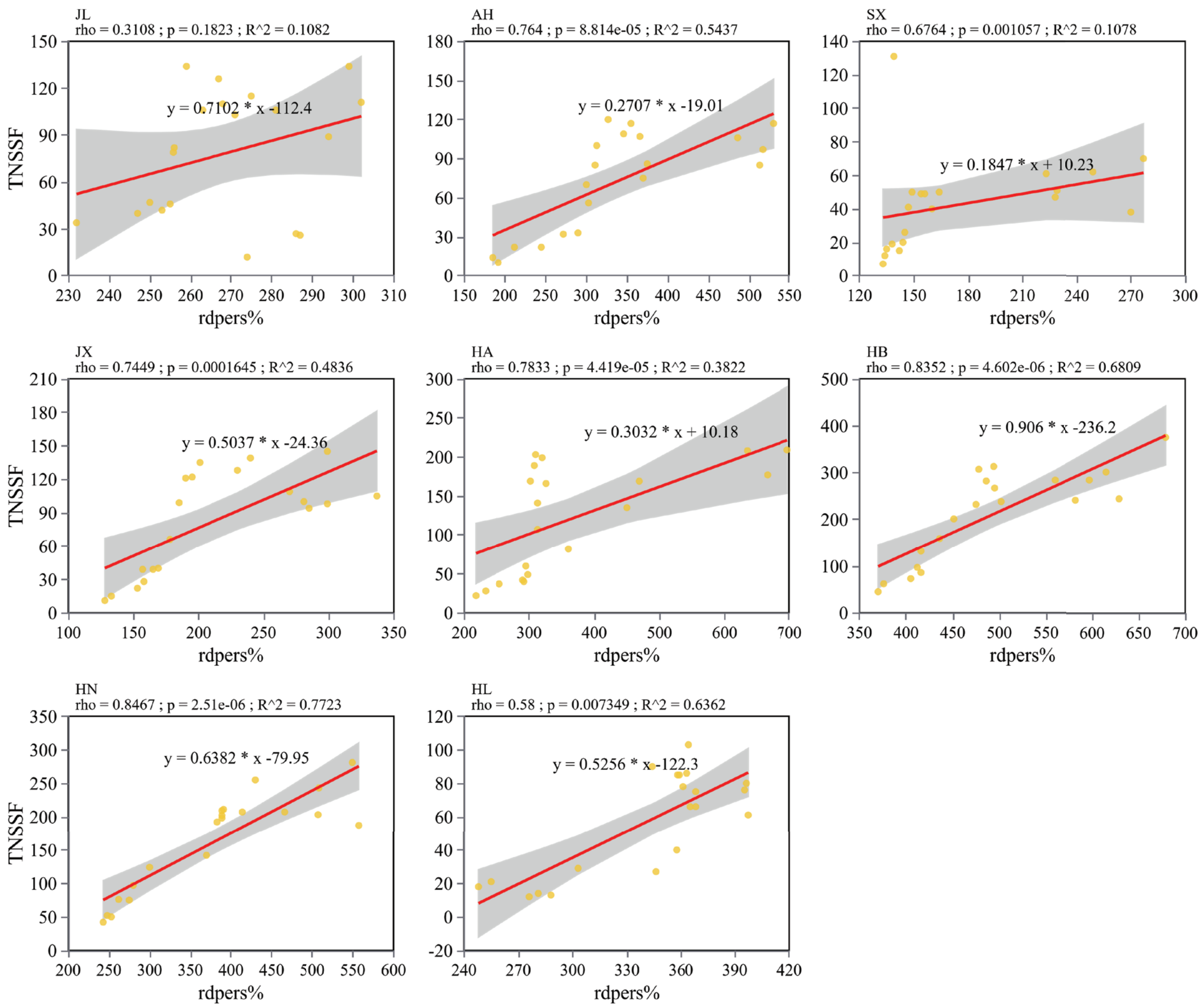
5.2.2 Central Region: Low Average Coefficient and Volatile Significance

Table 10 and Figure 5 reveal the central region's "weak and unstable" driving effect of RDPers: The lowest average coefficient and minimal dispersion. The average coefficient of central provinces is 0.5053, the lowest among the three regions, and the standard deviation (0.2447) is the smallest, indicating that intra-regional differences are not the main issue—instead, the overall driving strength is weak. Provincial-level instability in effect direction and significance is also observed: Anhui (AN, coefficient = 0.2707) and Jiangxi (JX, coefficient = 0.5037) show positive but low coefficients, with their regression plots (Figure 5) indicating weak correlation (e.g., AN: rho = 0.764, p = 8.814e-05, R^2 = 0.5437); Hubei (HB) is an outlier with a high coefficient (0.906) and strong significance (rho = 0.8352, p = 4.602e-06, R^2 = 0.6809), which may be attributed to its concentration of key universities (e.g., Wuhan University) and robust R&D infrastructure. In contrast, provinces like Heilongjiang (HL, coefficient = 0.5256) have low coefficients and marginal significance (rho = 0.58, p = 0.007349), reflecting insufficient translation of RDPers into TNSSF. The central region's predicament lies in the "resource mismatch"—while it has a certain scale of RDPers, it lacks supporting resources (e.g., lower FINSUP and FINDEV than the east) and efficient R&D collaboration mechanisms, leading to weak overall driving effects.

5.2.3 Western Region: Moderate Average Coefficient and Extreme Intra-Regional Differences

The western region's performance (Table 10, Figure 6) is characterized by a "moderate average but extreme intra-regional polarization": Moderate average coefficient, high dispersion. The average coefficient (0.6350) is higher than in the central region but lower than in the east, while the standard deviation (0.2722) is higher than in the central region, indicating pronounced intra-regional differences in the driving effect of RDPers. There is polarization between provinces with high and low coefficients: Tibet (XZ, 2.488) and Sichuan (SM, 2.523) have extremely high coefficients, with their regression plots showing strong significance (e.g., XZ: rho = 0.8504, p = 2.042e-06, R^2 = 0.6641)—this may be due to targeted national support policies (e.g., special R&D funds for ethnic regions) that amplify the role of RDPers. In contrast, Gansu (GE, 0.5423) and Ningxia (NX, 0.4915) have low coefficients and weak significance (e.g., NX: rho = 0.3935, p = 0.08607, $R^2 = 0.3286$), as their underdeveloped economic foundations and scarce R&D resources limit the effectiveness of RDPers. This polarization reflects the western region's uneven development of higher education R&D—provinces with policy support or resource concentration (e.g., Sichuan, Tibet) can leverage RDPers effectively, while most provinces struggle with resource constraints.

Figure 5 .Central Region Matrix Linear Regression Combination Plot

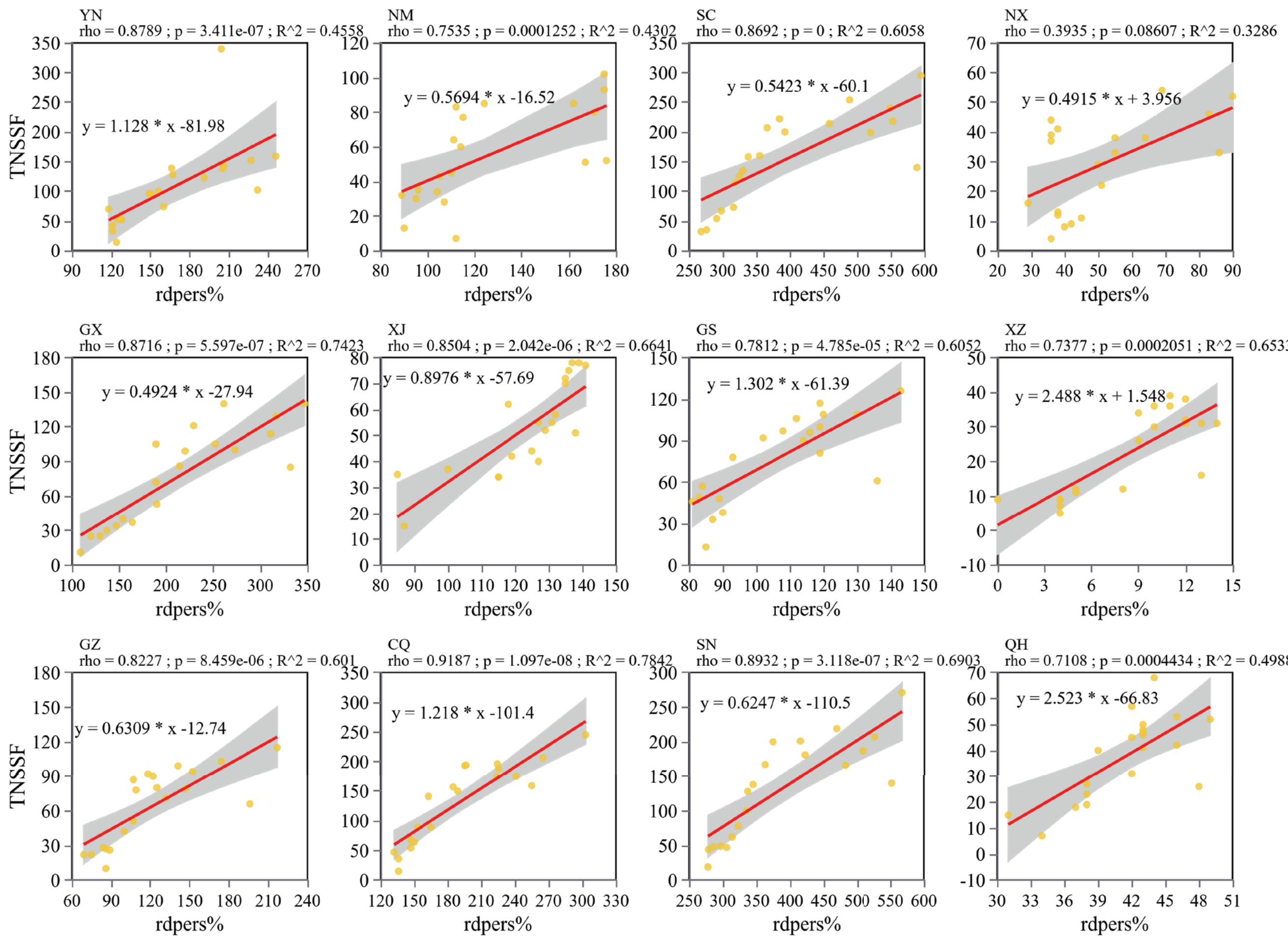


6. Discussion

The empirical results indicate that higher education R&D full-time equivalent personnel (RDpers) function as a stable and significant driver of the total number of National Social Science Fund projects (TNSSF). This finding supports the core premise of human capital theory, which asserts that R&D personnel, as carriers of professional knowledge and technical capabilities, directly facilitate the generation and transformation of research outcomes [8][9]. However, this positive impact exhibits marked temporal and spatial heterogeneity. Regarding temporal heterogeneity, time-interaction analyses demonstrate that the driving effect of RDpers has progressively strengthened since 2007, notably after the introduction of the 2012 "Innovation-Driven Development Strategy." The shift is largely attributable to improved policy-level optimization of R&D resource allocation and enhanced institutional environments promoting research translation, university-industry linkages, and streamlined project approval processes [10][11]. Regionally, the data reveal a clear "east > west > central" pattern in the marginal effects of RDpers on TNSSF outcomes. The eastern region benefits from a mature innovation ecosystem, with advanced support infrastructure and effective mechanisms for outcome transformation [12]. In contrast, the western region demonstrates intermediate performance, often buoyed by targeted national policy incentives in select provinces. The central region consistently underperforms, a condition attributed to "resource mismatch," where R&D personnel are present but poorly supported by complementary resources such as institutional incentives, financial autonomy, and administrative capacity [13][14].

Moreover, control variables such as financial support intensity (finsup) exhibit a statistically significant negative relationship with TNSSF, which may reflect inefficiencies in fund utilization. In some regions, rigid fund management systems or suboptimal project allocation may undermine the potential of R&D funding—creating a "resource curse" effect where inputs fail to convert into substantive outputs [15]. This study contributes to the literature by addressing endogeneity through lagged variable design and performing robustness checks including placebo tests and time-sample segmentation. It also explores intra-regional disparities via matrix linear regression, illustrating patterns such as polarization in the western region and sustained efficiency in the east [16]. Nonetheless, limitations remain. The TNSSF variable captures only the quantity—not quality—of social science output. Furthermore, while the use of year dummies captures macro-temporal trends, it fails to distinguish the causal effects of individual policy shocks. Future research may consider richer output indicators (e.g., citations, completion rates), integrate mediating variables (e.g., knowledge collaboration networks), and develop explicit policy intensity indices to improve explanatory power and causal inference [17][18].

Figure 6 .Western Region Matrix Linear Regression Combination Plot



Conclusion

To address the gap of insufficient attention to human capital and regional disparities in studies of regional social science strength—using the total number of National Social Science Fund Projects (TNSSF) as the core measure—this study analyzes panel data from 2003 to 2022 for 31 Chinese provinces. TNSSF is used as the key indicator of regional social science development, while higher education R&D full-time equivalent personnel (RDPers) serves as the main proxy for social science human capital input. Employing a high-dimensional fixed-effect model and multi-method validation (including endogeneity mitigation via lagged variables and robustness tests such as placebo and variable substitution), the key findings are as follows: First, RDPers consistently drives regional social science strength in a positive manner, with a validated coefficient of 0.002–0.004 ($p < 0.01$), confirming its role as a core driver. Second, a time-moderating effect emerges: after 2007 (especially following the 2012 "Innovation-Driven Development Strategy"), the interaction coefficient of RDPers and year dummies rises to 0.001–0.003 ($p < 0.05$), reflecting the enhanced impact of policies on RDPers' role. Third, regional heterogeneity is significant: across regions, the driving effect follows "East (avg. coefficient = 1.076) > West (0.635) > Central (0.505)"; within regions, the East demonstrates stable efficiency, the West shows polarization (e.g., Sichuan: 2.523 vs. Gansu: 0.542), and the Central region faces a "resource mismatch." Fourth, financial support intensity (FINSUP) has a negative effect (–187.21, $p < 0.05$) due to inefficient use, while financial development (FINDEV) has recently shown a positive effect (19.68, $p < 0.05$). Theoretically, this study enriches research on the drivers of regional social science by confirming the role of human capital and refining understanding of intra-regional disparities. Practically, it proposes targeted strategies such as regionalized personnel allocation, reform of social science fund management, and inter-regional collaboration. Limitations include TNSSF's focus on quantity rather than quality and unexplored mechanisms; future research may develop multi-dimensional indices and introduce mediating variables. This study provides empirical insights for optimizing social science resource allocation and promoting balanced national development.

Funding

Acknowledgments This research is supported by the project 'Path of Government Financial Input Restructuring to Promote High Quality Development of Education'（Project No.2025279）of the Canal Cup Extracurricular Academic Science and Technology Fund for College Students of Zhejiang University of Technology.

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Chen, B., Sun, Y., Tong, Y., & Liu, L. (2024). The measurement, level, and influence of resource allocation efficiency in universities: Empirical evidence from 13 "double first class" universities in China. Humanities and Social Sciences Communications. https://doi.org/10.1057/s41599-024-03461-z

[2] Wang, X., & Yang, J. (2025). Spatial–temporal heterogeneity and driving factors of resource allocation efficiency in regular higher education in China. Studies in Higher Education. https://doi.org/10.1080/03075079.2025.2537292

[3] Xi, X., Zhao, Q., Liu, L., & Yang, G. (2024). Role of Educational Level in Innovation Performance: An Exploration Study of Higher Education Graduates in China. Journal of the Knowledge Economy. https://doi.org/10.1007/s13132-024-02578-5

[4] Liang, Q., & Yin, F. (2024). Spatiotemporal Coupling of Higher Education and Economic Development in China. Sustainability, 16(16), 7198. https://www.mdpi.com/2071-1050/16/16/7198

[5] Liu, H., & Lin, W. (2022). Uneven Innovation Geography in China: Institutional Capacity and R&D Fragmentation. Economic Geography, 98(1), 44–65.

[6] Wang, L., & Zhao, J. (2023). R&D Input, Efficiency and Regional Knowledge Output: Evidence from Provincial Data. China Economic Review, 79, 101945.

[7] Xu, B., & Yang, R. (2024). Incentive Mechanisms and the Transformation of University R&D in China. Research in Higher Education, 65(2), 311–328.

[8] Xiao, S., Sheng, J., & Zhang, G. (2024). Rising tides of knowledge: Exploring China's higher education landscape and human capital growth. Journal of the Knowledge Economy, 15. https://doi.org/10.1007/s13132-024-02102-9

[9] Luo, M., & Ye, D. (2023). Measuring the contribution of knowledge-based capital to social science productivity in Chinese universities. Higher Education Quarterly. https://doi.org/10.1111/hequ.12400

[10] Wen, F., Yang, S., & Huang, D. (2023). Heterogeneous human capital, spatial spillovers and regional innovation: Evidence from the Yangtze River Economic Belt. Humanities and Social Sciences Communications, 10(1), 195. https://doi.org/10.1057/s41599-023-01809-5

[11] He, Y., Li, R., & Song, H. (2022). Government R&D investment, innovation performance, and regional development: Evidence from China's central policy reform. Economic Analysis and Policy, 74, 47–60. https://doi.org/10.1016/j.eap.2022.01.009

[12] Zhou, Y., & Fan, Z. (2021). University research ecosystem maturity and scientific output: Regional evidence from China. Frontiers in Education, 6, 634201. https://doi.org/10.3389/feduc.2021.634201

[13] Xiao, H., & You, J. (2021). The heterogeneous impacts of human capital on green total factor productivity: A regional diversity perspective. Frontiers in Environmental Science, 9, 713562. https://doi.org/10.3389/fenvs.2021.713562

[14] Cheng, L., & Su, C. (2022). Institutional constraints and academic productivity in Chinese social sciences. International Journal of Educational Development, 88, 102526. https://doi.org/10.1016/j.ijedudev.2021.102526

[15] Wang, Y., Li, Y., & Zhang, X. (2023). Misaligned incentives and inefficient R&D allocation: Evidence from subnational education funds. Asia Pacific Journal of Education, 43(2), 165–182.

[16] Liu, Y., & Wang, T. (2022). Spatial polarization of scientific productivity across Chinese provinces: A matrix regression approach. Regional Science Policy & Practice, 14(5), 1102–1120.

[17] Gao, M., Zhang, T., & Liu, Y. (2024). Evaluating policy-driven innovation productivity in higher education: A mediation analysis approach. Education Economics. https://doi.org/10.1080/09645292.2024.2298204

[18] Zhang, J., & Hu, Y. (2023). From Funding to Productivity: The Moderating Role of Institutional Incentives in University R&D. Asian Journal of Innovation & Policy, 12(2), 145–169.