Research on the Effects of Science and Technology Talent Concentration in the Guangdong-Hong Kong-Macao Greater Bay Area Based on Econometric Methods

Jialin Zheng [#], Xiaoying Lv [#], Canbin Wang [#], Zhongyan Li [*]

School of Statistics and Mathematics, Guangdong University of Finance and Economics, Guangzhou, China, 510320

* Corresponding Author Email: 20141230@gdufe.edu.cn

[#]These authors contributed equally.

Abstract. This study investigates the impact of the science and technology talent policy on the concentration of science and technology talents in the Guangdong, Hong Kong, and Macao Greater Bay Area from 2010 to 2017. Utilizing data from the China Urban Statistical Yearbook and various prefecture-level city statistical yearbooks, the study employs a combination of single-difference and double-difference methods to analyze the effects of the policy. The results indicate that the policy significantly enhances the concentration of science and technology talents in the region, with the double-difference method providing more accurate estimates than the single-difference method. The research further applies the Propensity Score Matching (PSM) and DID methods to address potential selection biases, confirming that the policy's impact is robust and reliable. Additionally, mediation analysis reveals that the policy's effectiveness is partially mediated by macroeconomic regulation, industrial structure, and demographic factors. The findings underscore the importance of targeted science and technology talent policies in promoting regional talent concentration, offering valuable insights for policymakers aiming to foster innovation-driven development.

Keywords: Science and technology talent policy, Talent concentration, Guangdong-Hong Kong-Macao Greater Bay Area, Double-difference method.

1. Introduction

In recent years, the Guangdong, Hong Kong, and Macao Greater Bay Area have become a focal point for China's innovation-driven development strategy[1,2]. As one of the most economically dynamic and open regions in China, the Greater Bay Area plays a critical role in the country's overall modernization and international competitiveness[3]. However, the rapid economic development of the region has also highlighted a pressing need for scientific and technological talents[4], who are essential to sustaining innovation and maintaining the region's competitive edge in a globalized economy. To address this need, the central and local governments have introduced a series of policies aimed at attracting, retaining, and cultivating high-caliber science and technology professionals[5]. The science and technology talent policy, as part of these initiatives, aims to promote the accumulation of high-quality talent by enhancing government macro-control, optimizing the industrial structure, and improving the demographic environment. While numerous studies have explored the impact of various talent policies on regional development, few have specifically examined the unique context of the Greater Bay Area, a region characterized by its distinct administrative divisions and economic integration with Hong Kong and Macao. Understanding how these policies impact the concentration of science and technology talents in the Greater Bay Area is crucial for evaluating their effectiveness and guiding future policy formulation.

Moreover,selectingf appropriate methods for evaluating policy effects is a key challenge in policy research[6]. Traditional single-difference methods may overestimate the impact due to unobserved factors, while the double-difference method with two-way fixed effects offers a more precise approach by controlling for both time-variant and time-invariant factors[7,8]. Additionally, given the potential for selection bias in the policy implementation, the use of Propensity Score Matching (PSM) combined with DID can help produce more reliable results by creating a more comparable control group. Therefore, this study aims to fill the gap by systematically analyzing the impact of the science and technology talent policy on the concentration of talent in the Greater Bay Area using a combination of rigorous quantitative methods. The study also explores the mediating mechanisms through which the policy operates, providing a comprehensive evaluation of its effectiveness and offering insights into how similar policies could be designed and implemented in other regions.

2. Evaluating the Effectiveness of Science and Technology Talent Policies

The data in this paper mainly come from the 2010-2017 China Urban Statistical Yearbook and the statistical yearbooks of each prefecture-level city. Cities with more missing data are deleted, and some of the default values are filled in by linear interpolation to finally obtain the panel data of 37 cities from 2010 to 2017, and the descriptive statistics of their main variables are shown in Table 1.

Table 1 Descriptive statistics of main variables

Variable	Obs	Mean	Std. Dev.	Min	Max.
Agg	296	1.543044	1.209877	0.1139913	5.151502
did	296	0.1216216	0.3274018	0	1
dev	296	11.21815	0.5406938	9.693198	12.15643
infor	296	0.4538958	0.2942133	0.0807799	1.890194
medi	296	0.0053001	0.002516	0.0004665	0.0137458
open	296	-4.732726	1.190909	-7.98071	-2.079475
edu	296	0.0013689	0.0016571	0.0000669	0.0077316
traf	296	15.16251	19.16204	0.5049226	149.9038

The regression results of the impact of science and technology talent policy on the level of science and technology talent concentration in Guangdong, Hong Kong, and Macao Greater Bay Area are shown in Table 2. The estimation results of the OLS regression fitting single difference method are applied to obtain model (1) and model (2) in Table 2, and the difference in the level of scientific and technological talent concentration in the sample cities of the experimental group and the control group are compared singly. In contrast, the double difference method with two-way fixed effects is applied to construct model (3) and model (4), and the corresponding estimation results are obtained.

According to the analysis of the regression results, the regression coefficients obtained from the regression of single-difference method and double-difference method are both significantly positive at the 5% level, which indicates that the scientific and technological talent policy of the Greater Bay Area significantly affects the level of scientific and technological talent pooling in the region. Comparing model (1) and model (3), it can be found that the regression coefficients fitted by the single-difference method are larger than those by the double-difference method, which indicates that selecting the single-difference method for the study of the scientific and technological talent policy on the level of scientific and technological talent concentration in the Greater Bay Area overestimates the effect of the implementation of the policy, and the estimation by the double-difference method is more accurate.

After adding six control variables such as the level of economic development set in the previous section based on model (1) and model (3), the size and significance of the regression coefficients did not change essentially, which shows that the science and technology talent policy can effectively enhance the level of science and technology talent concentration in the Greater Bay Area under the control of the influence of external factors.

Table 2 Benchmark regression results

variant	Model (1)	Models (2)	Models (3)	Models (4)
did	0.363**	0.171**	0.151***	0.112**
dev		1.457***		1.08**
traf		0.049		0.062
medi		-3.034		-2.952
infor		0.166*		0.121*
edu		6.397		5.983
open		0.115**		0.177**
_cons	0.177**	0.282*	0.155**	0.471**
vintage effect	√	√	√	√
city affect	×	×	√	√
R-square	0.137	0.163	0.156	0.234
Observations	296	296	296	296

Note: *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

3. Effect Testing with Parallel Trend and PSM-DID Methods

When conducting an assessment of policy or treatment effects, the double-difference method requires that the treatment and control groups have similar trends before the treatment[9]. This paper performs a parallel trend test for the first four periods of 2014 (2010-2013), and the test plot is shown in Figure 1. The figure shows that in the four periods before the implementation of the policy, the test coefficient is not significantly different from 0 at the 5% significance level, indicating that the treatment and control groups passed the parallel trend test before the implementation of the policy, and the double difference method can be carried out to test the effect of the science and technology talent policy.

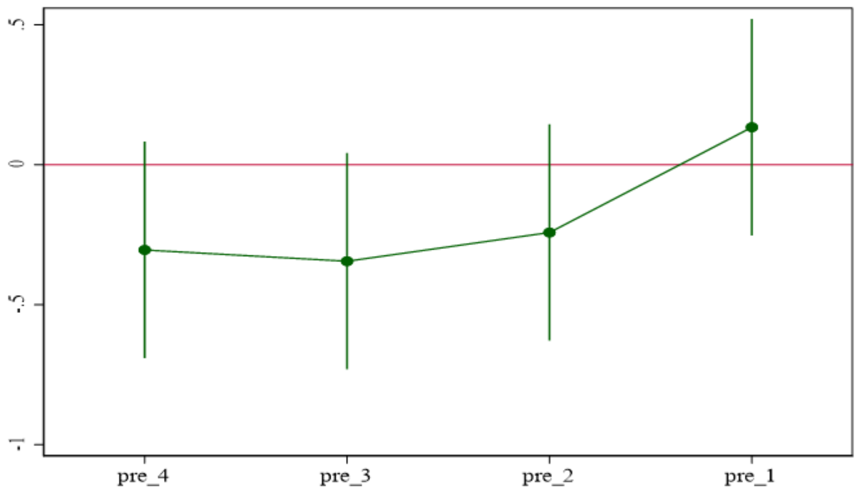

Figure 1 Parallel trend test plot

Further development of propensity score matching. The first condition that needs to be fulfilled for propensity score matching (PSM) is the balance assumption[10]. Before obtaining the results of propensity score matching scores, the collected panel data need to be tested for balance. Figure 2 and Table 3 show the results of the data balance assumption test. After the propensity score matching, the deviation of the regional GDP per capita of the sample experimental group and the control group decreases to 3.8%; the deviation of the level of public service, the level of informatization, the level of transportation, and the level of education decreases to less than 15%; the deviation of the level of public service decreases to 5.8%, the level of informatization decreases to 11.1%, the level of transportation decreases to 2.4%, and the level of education decreases to 2.4%, and the level of education decreases to 2.4%. Is 2.4%, and the education level drops to 8.8%, so it can be shown that the selection of control variables in this paper is reasonable, and the matching results can balance the data better. This also reflects the fact that this paper seeks to match the pilot cities in the Greater Bay Area that have implemented the Science and Technology Action Plan with a collection of cities that are highly similar to the pilot cities in terms of other characteristics, such as economic level, but have not implemented the Science and Technology Action Plan.

Table 3 Table of results of balanced hypothesis testing for panel data

variant	Before and after sample matching	Average value Treatment Group Control Group		Decrease in deviation (%)	Deviation (%)	t-test	
dev	prematch after matching	11.433 11.421	11.159 11.402	56.4 3.8	93.3	3.83 0.18	0.000 0.859
medi	prematch after matching	0.00757 0.00722	0.00461 0.00708	60.7 5.8	90.4	5.71 0.24	0.033 0.752
infor	prematch after matching	0.72828 0.66181	0.36857 0.62977	48.8 11.1	77.2	6.20 0.7	0.223 0.488
open	prematch after matching	-4.1906 -4.1908	-4.9054 -3.7691	68.9 -40.7	41.0	4.62 - 2.52	0.000 0.013
edu	prematch after matching	0.00223 0.00224	0.0011 0.00208	34.8 8.8	74.7	5.34 0.46	0.154 0.648
traf	prematch after matching	18.718 18.923	13.821 18.474	26.1 2.4	90.8	1.91 0.16	0.058 0.875

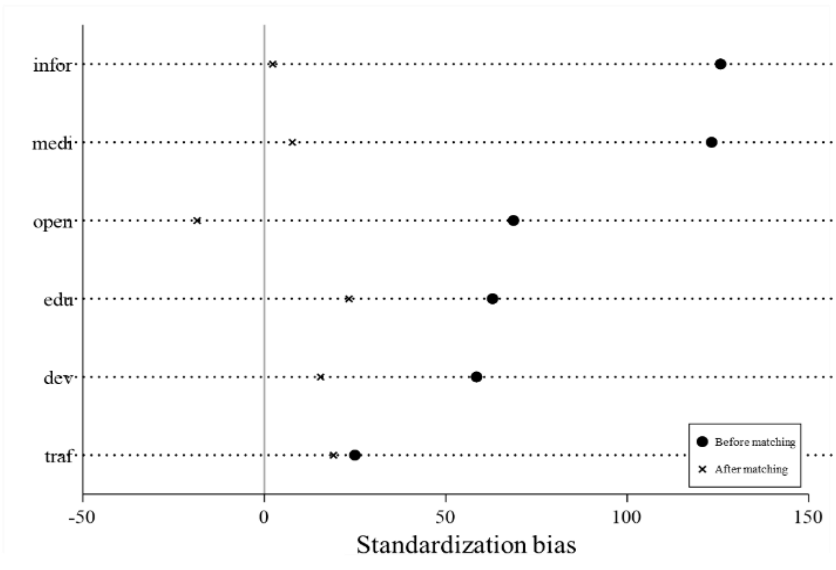

Figure 2 Plot of the effect of balanced hypothesis testing for the panel data in this paper

In addition, as shown in Figure 3, from the kernel density function plots before and after propensity score matching of the treatment and control groups, it can be found that the overlapping region of the kernel density function of the treatment and control groups increased significantly after completing the PSM matching, indicating that the propensity score matching was effective in decreasing the sampling bias of the treatment and control groups, and therefore, we can further adopt the double difference method to assess the policy effects of the experimental group and the newly matched control group.

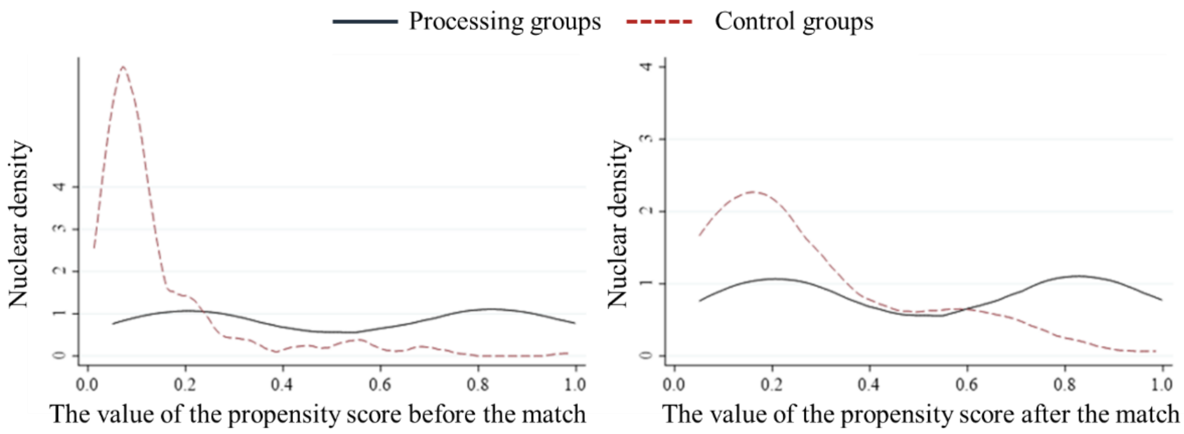

Figure 3 Kernel density function plot for the panel data in this paper

The use of the DID double difference method can effectively identify whether the S&T talent policy affects the level of S&T talent agglomeration, but due to the reality that the S&T talent policy tends to take the lead in the implementation of the more economically developed cities, such as the municipalities and provincial capitals, i.e., there exists a certain degree of selective bias, which is not able to satisfy the assumption condition of random grouping. To make the treatment group and the control group cities as similar as possible in all aspects of the characteristics, and to eliminate the selection bias, the choice of the PSM-DID method can weaken the characteristic differences of the sample, to more accurately assess the scientific and technological talent policy effect. In this paper, we set the width of the caliper 0.005 (1:2) matching method to match the treatment group and control group, as shown in the figure, the sample difference between the treatment group and the control group is larger before matching, and the difference between the two is narrowed and close to 0 after matching, which indicates that PSM reduces the systematic difference between the treatment group and the control group, and the regression results of using PSM-DID will be more reliable. The test of PSM-DID The results are shown in table 4, model (1) contains only policy dummy variables, and model (2) adds control variables, the regression coefficient is still significant at a 5% significance level, indicating that the policy of scientific and technological talents improves the level of scientific and technological talent pooling.

Table 4 Results of PSM-DID versus placebo test

variant	PSM-DID		Placebo testing	
	(1)	(2)	(3)	(4)
did	0.1103^{**} (0.0523)	0.0836^{**} (0.0421)	0.0092 (0.0201)	0.0097 (0.0193)
control variable	×	√	√	√
vintage effect	√	√	√	√
city affect	√	√	√	√
R^2	0.2893	0.3073	0.0604	0.0613

Considering that the change in trends in the treatment and control groups after the time point of the policy intervention may not be caused by the policy, but by other policies in the same period. Drawing on Topalova's approach, select the years before the implementation of the policy pilot, assume that the STT policy was started in one of those years, and conduct DID estimation of the effect of the virtual policy shocks; if the estimated coefficient of the virtual DID is not significant, then it indicates that the virtual policy pilot does not have a significant effect on the level of STT aggregation in the city, which suggests that the increase in the level of STT aggregation in the pilot city is indeed caused by the policy of scientific and technological talents that has nothing to do with other factors. This paper assumes that the implementation of the S&T talent policy starts in 2011 and 2012, and estimates the effect of the virtual policy shock by the double difference method. The results are shown in model (3) and model (4) of Table 2. Setting the policy implementation time in 2011 and 2012, the virtual estimated coefficients are not significant. This indicates that the virtual policy pilot has no significant effect on the level of scientific and technological talent pooling, and from the side, it shows that the increase in the level of scientific and technological talent pooling in the pilot cities is indeed caused by the scientific and technological talent policy, rather than other factors.

4. Mediation Analysis and Policy Mechanisms

Through the DID and PSM-DID models, it can be proved that the Science and Technology Action Plan can have a significant positive impact on the concentration of scientific and technological talents in the Guangdong, Hong Kong, and Macao Greater Bay Area, and as mentioned before, the policy on scientific and technological talents can promote the level of scientific and technological talents concentration by carrying out the government's macro-control, optimizing the level of industrial structure as well as enhancing the quality and quantity of the population, therefore, this paper will introduce the test of mediating effect to further So this paper will introduce the mediation effect test to further analyze the influence mechanism of scientific and technological talent policy on the level of scientific and technological talent concentration. As shown in Table 5.

Table 5: Selection table of relevant variables for the mediation effect test

variable name	Variable Meaning	Data sources
Macro-regulation (C)	Share of investment in science and technology personnel in total investment expenditure	China Statistical Yearbook
Industrial structure (S)	Share of regional secondary and tertiary output in total regional output value	China Statistical Yearbook
Demographic environment (E)	Both the quantity and quality of the population	China Statistical Yearbook

Construct the following mediated effects model, whereMed_{it} is the mediating variable:

$$Agg_{it} = \beta_0 + \beta_1 did_{it} + \beta_2 X_{it} + u_i + \lambda_t + \varepsilon_{it} \tag{1}$$

$$Med_{it} = \alpha_0 + \alpha_1 did_{it} + \alpha_2 X_{it} + u_i + \lambda_t + \varepsilon_{it} \tag{2}$$

$$Agg_{it} = \varphi_0 + \varphi_1 did_{it} + \varphi_2 Med_{it} + \varphi_3 X_{it} + u_i + \lambda_t + \varepsilon_{it} \tag{3}$$

Table 6 shows the results of the mediation effect test of industrial structure: the explanatory variable did and the explanatory variable Agg are significantly correlated at the 5% level, with a regression coefficient of 0.1124; macro-control and the explanatory variable did are significantly correlated at the 5% level, with a coefficient of 0.0622; when the mediator variable macro-control is added, the independent variable Med and dependent variable Agg are significantly correlated at the 10% level, which indicates that the government macro-control plays the role of "bridge" between S&T talent policy and S&T talent pooling. The government's macro-control plays the role of "bridge" between science and technology talent policy and science and technology talent pooling. Therefore, under the premise of significant effect, we can get the total utility of scientific and technological talent policy on scientific and technological talent concentration is 0.1124, and the direct effect of scientific and technological talent policy on scientific and technological talent concentration is 0.1087, which indicates that the government's macroeconomic regulation plays a part of the mediating role between the scientific and technological talent policy and the level of scientific and technological talent concentration.

Table 6 Macro-control mediation effect test table

	(1)	(2)	(3)
VARIABLES	Agg	C	Agg
C			0.1277**
			(0.8304)
did	0.1124**	0.0622*	0.1087*
	(11.2213)	(5.6221)	(12.9521)
dev	1.0835**	0.0421***	0.9723**
	(1.2287)	(0.5351)	(3.7324)
traf	0.0623	0.0000	0.0007
	(0.5398)	(0.0634)	(0.5516)
infor	0.1212*	0.0753**	0.1091**
	(-0.6228)	(1.4155)	(-0.6015)
medi	-2.9516	18.9189	-9.7345
	(-0.2253)	(0.9144)	(-0.1925)
open	0.1772**	-0.0391**	0.1738**
	(1.3780)	(-2.2672)	(1.3925)
edu	5.9831	76.6247	-58.1348
	(-0.5177)	(0.9028)	(-0.4590)
N	296	296	296
Adj. R^2	0.2785	0.2467	0.1945

Table 7 shows the results of the mediation effect test of industrial structure: the explanatory variable did and the explanatory variable Agg are significantly correlated at the 5% level, with a regression coefficient of 0.1124; the industrial structure and the explanatory variable did are significantly correlated at the 5% level, with a coefficient of 0.0632; when the mediator variable industrial structure is added, it and the dependent variable Agg are still significantly correlated at the 5% level, with a coefficient of 0.0353; indicating that the regional industrial structure is linked to the policy of scientific and technological talents and has a driving influence on the agglomeration of scientific and technological talents. 0.0353; it shows that the regional industrial structure is linked to the policy of scientific and technological talents, and plays a driving influence on the concentration of scientific and technological talents. We can get the total utility of scientific and technological talent policy on scientific and technological talent concentration for 0.1124, in the case of the significance test passed, the direct utility of scientific and technological talent policy on the level of scientific and technological talent concentration for 0.0932, indicating that the level of industrial structure in the scientific and technological talent policy and the level of scientific and technological talent concentration to play a part of the intermediary role between the level of scientific and technological talent.

Table 7 Test table for the mediating effect of level of industrial structure

	(1)	(2)	(3)
VARIABLES	Agg	S	Agg
S			0.0353*
			(0.8812)
did	0.1124**	0.0632**	0.0932**
	(11.2213)	(4.4231)	(12.2422)
dev	1.0835**	2.1516*	0.9523**
	(1.2287)	(0.2242)	(3.4762)
traf	0.0623	0.4244	0.3123
	(0.5398)	(0.2411)	(0.8452)
infor	0.1212*	0.0753*	0.3451**
	(-0.6228)	(1.3121)	(-0.6452)
medi	-2.9516*	13.1321	0.3923
	(-0.2253)	(0.1567)	(-0.8613)
open	0.1772**	-0.0213**	0.4234**
	(1.3780)	(-3.3242)	(1.4242)
edu	5.9831	14.3127	12.4348
	(-0.5177)	(0.4324)	(-0.7490)
N	296	296	296
Adj. R^2	0.2785	0.2313	0.2972

In this paper, the demographic environment is categorized into population quantity and population quality, and tests for the mediating effects of each are discussed below.

A test of the mediating effect of population size:

Table 8 shows the results of the mediating effect test of population size can be concluded that the mediating variable of population size and the dependent variable Agg are not correlated at the level of significance, therefore, the scientific and technological talent policy can not have a mediating effect on the scientific and technological talent agglomeration through the population size.

Table 8 Tests for the mediating effect of population size

	(1)	(2)	(3)
VARIABLES	Agg	E1	Agg
E1			-0.1324
			(0.8314)
did	0.1124**	-0.0423	0.1733*
	(11.2213)	(6.5234)	(17.34)
dev	1.0835**	0.3424**	0.7252**
	(1.2287)	(0.2214)	(3.4141)
traf	0.0623	0.0131*	0.8583
	(0.5398)	(0.0414)	(0.5516)
infor	0.1212*	-0.0628**	0.1132**
	(-0.6228)	(1.6241)	(-0.6412)
medi	-2.9516	1.6234*	-9.3215
	(-0.2253)	(0.7614)	(-0.1375)
open	0.1772**	-0.2313*	0.1838**
	(1.3780)	(-2.2831)	(1.3411)
edu	5.9831	15.2311	12.5341
	(-0.5177)	(0.9413)	(-0.4524)
N	296	296	296
Adj. R^2	0.2785	0.2825	0.1495

Table 9 shows the results of the mediation effect test of population quality: the explanatory variable did is significantly correlated with the explanatory variable Agg at the 5% level, with a correlation coefficient of 0.1124; the population quality and the explanatory variable did are significantly correlated at the 5% level, with a coefficient of 0.4242; when the mediator variable industry structure is added, it is significantly correlated with the dependent variable Agg at the 10% level, with a coefficient of 0.1069; the mediator variable population quality and the dependent variable Agg are significantly correlated at the 5% level, with a coefficient of 0.1927. We can get the total utility of scientific and technological talent policy on scientific and technological talent agglomeration is 0.1124, and the direct effect of scientific and technological talent policy on scientific and technological talent agglomeration is 0.1069, which indicates that the quality of population plays a part of mediating role between the scientific and technological talent policy and scientific and technological talent agglomeration level. Intermediary role.

Table 9 Tests for the mediating effect of population size

	(1)	(2)	(3)
VARIABLES	Agg	E2	Agg
E2			0.1927**
			(0.3924)
did	0.1124**	0.4242**	0.1069*
	(11.2213)	(5.8558)	(14.2545)
dev	1.0835**	0.0302**	0.3138**
	(1.2287)	(0.6464)	(3.7324)
traf	0.0623	0.0525	0.3137
	(0.5398)	(0.07524)	(0.5316)
infor	0.1212*	0.07472**	0.1131**
	(-0.6228)	(1.8624)	(-0.5145)
medi	-2.9516	1.2556	-7.2315
	(-0.2253)	(0.9138)	(-0.7133)
open	0.1772**	-0.0425**	0.5255**
	(1.3780)	(-2.2672)	(1.3143)
edu	5.9831	13.1314**	1.3742*
	(-0.5177)	(0.9128)	(-0.7675)
N	296	296	296
Adj. R^2	0.2785	0.2441	0.1258

5. Conclusion

This study evaluates the effect of the science and technology talent policy on the concentration of science and technology talents in the Guangdong, Hong Kong, and Macao Greater Bay Area. Through the use of single-difference and double-difference methods, it is demonstrated that the policy significantly improves the level of science and technology talent concentration. The results indicate that the estimation using the double-difference method is more accurate compared to the single-difference method. Furthermore, the analysis conducted using the Propensity Score Matching (PSM) and DID methods confirms that the policy has a substantial positive effect on talent concentration. The conclusion strengthens the understanding that the policy's implementation is an effective means of enhancing the regional accumulation of science and technology talents.

The findings suggest that the adoption of science and technology talent policies, especially those targeted at enhancing macro-control, optimizing industrial structures, and improving population quality, can play a vital role in promoting regional talent concentration. The study highlights that targeted interventions can yield positive outcomes even under conditions of selective bias. Thus, policymakers should continue to employ similar approaches to foster a conducive environment for talent accumulation.

Acknowledgements

The authors gratefully acknowledge financial support from National University Student Innovation Training Project (S202410592011).

References

[1] HU Z. Research status of Guangdong-Hong Kong-Macao Greater Bay Area[J]. Open Journal of Social Sciences, 2019,7(02): 275.

[2] YANG Z, LI X, WANG F, et al. Spatial structure and network characteristics of the coupling coordination innovation ecosystems in the Guangdong–Hong Kong–Macao Greater Bay area[J]. Scientific Reports, 2024,14(1): 395.

[3] ZHAO Y. Bay Area Economic Integration in China[J]. LUNDS Universitet PhD Thesis, 2022.

[4] MELLANDER C, FLORIDA R. The rise of skills: Human capital, the creative class, and regional development[J]. Handbook of Regional Science, 2021: 707-719.

[5] ZHAO C, COOKE F L, CHEN L, et al. Between market and mayor: talent management in Chinese private firms and the role of local governments[J]. Asia Pacific Journal of Human Resources, 2023,61(2): 393-422.

[6] MERGONI A, De WITTE K. Policy evaluation and efficiency: A systematic literature review[J]. International transactions in operational research, 2022,29(3): 1337-1359.

[7] QI S, CHEN Y, CHU K. Lifting the Veil: Using a Spatial Discontinuity to Assess the Impact of National Development Zones on Upgrading of Chinese Manufacturing[J]. Open Economies Review, 2024: 1-32.

[8] MALIK M A, ROHM L R, van BAAL P, et al. Improving maternal and child health in Pakistan: a program evaluation using a difference in difference analysis[J]. BMJ Global Health, 2021,6(12): e6453.

[9] EGAMI N, YAMAUCHI S. Using multiple pretreatment periods to improve difference-in-differences and staggered adoption designs[J]. Political Analysis, 2023,31(2): 195-212.

[10] GUO S, FRASER M, CHEN Q. Propensity score analysis: recent debate and discussion[J]. Journal of the Society for Social Work and Research, 2020,11(3): 463-482.