

An Empirical Study of Science and Technology Policy Funding Effects Based on the DID Model: A Case Study of 20 Years of Earth Science Projects under the Distinguished Young Scholars Fund (Postprint)

Authors: Tian Renhe, Zhang Zhiqiang, Yu Jie, Wang Ping

Date: 2023-08-27T00:00:00+00:00

Abstract

[Purpose/Significance] Scientific and technological innovation constitutes a critical strategic pillar of the modern economic system. Following the reform and opening-up, a series of science and technology policies have been formulated and implemented to foster technological advancement and economic growth in China. However, the funding effects of these policies and their temporal evolution throughout the implementation history have garnered widespread scholarly attention. [Method/Process] This study conducts an empirical analysis of the National Science Fund for Distinguished Young Scholars (Earth Science Program), which has been implemented for two decades. To mitigate endogeneity concerns arising from sample selection bias, the Difference-in-Differences (DID) model is employed to evaluate the impact of the Distinguished Young Scholars Fund on scientists' research paper productivity. An "Environment-Motivation-Behavior" model is proposed to interpret the empirical findings. [Results/Conclusion] The empirical results demonstrate that during 1994-2008, the funding effect of the Distinguished Young Scholars Fund was statistically significant for 12 years, representing 80% of the period. Compared with scientists who did not receive the award, recipients could publish an additional 0.412-3.234 papers per year. After data conversion, when compared with other contemporaneous scientists, recipients could publish an additional 0.426-3.277 papers annually. The 2002 Distinguished Young Scholars Fund exhibited the largest funding effect, whereas the 2007 and 2008 funds showed no significant impact. This methodology can be utilized to construct control groups for evaluating the funding effects of science and technology policies, thereby achieving causal effect inference.

Full Text

Empirical Research on the Funding Effects of Science and Technology Policy Based on the DID Model: Taking the 20-Year Implementation of the Distinguished Young Scientist Fund Earth Science Projects as an Example

Tian Renhe^{1,2}, Zhang Zhiqiang¹, Yu Jie³, Wang Ping⁴ ¹Chengdu Library and Information Center, Chinese Academy of Sciences, Chengdu 610041 ²University of Chinese Academy of Sciences, Beijing 100049 ³Interdisciplinary Center for Social and Behavioral Studies, Dongbei University of Finance and Economics, Dalian 116025 ⁴School of Public Health, Jilin University, Changchun 130021

Abstract

[Purpose/Significance] Scientific and technological innovation serves as a crucial strategic pillar for modern economic systems. Following China's reform and opening-up, a series of science and technology policies were introduced and implemented to promote scientific development and economic growth. However, the funding effects of these policies and how these effects have evolved over the historical implementation cycle have attracted widespread attention. **[Method/Process]** This study conducts an empirical analysis of the National Science Fund for Distinguished Young Scholars (hereafter referred to as the "Distinguished Young Scientist Fund" or "Jie Qing Fund") in earth science, which has been implemented for 20 years. To address endogeneity issues caused by sample selection bias, we employ the Difference-in-Differences (DID) model to evaluate the impact of the Distinguished Young Scientist Fund on scientists' research paper output efficiency. We propose an "Environment-Motivation-Behavior" model to interpret the research findings. **[Result/Conclusion]** Empirical results indicate that between 1994 and 2008, the Distinguished Young Scientist Fund demonstrated significant funding effects in 12 years, accounting for 80% of the period. Compared with scientists who did not receive the fund, recipients published 0.412–3.234 additional papers per year. After data conversion, compared with other scientists in the same period as the funded scientists, recipients published 0.426–3.277 additional papers per year. The strongest funding effect occurred in 2002, while the effects in 2007 and 2008 were not statistically significant. This methodology demonstrates that constructing control groups enables causal inference in science and technology policy evaluation.

Classification Number: C93-03

Keywords: data science, DID model, National Science Fund for Distinguished Young Scholars, science and technology policy evaluation

Globally, science and technology have increasingly become the primary force driving economic and social development, with innovation-driven growth representing the prevailing trend, leading countries worldwide to continuously in-

crease their scientific and technological investments. According to statistics from the National Bureau of Statistics and the Chinese Academy of Science and Technology for Development [1-2], China's total R&D expenditure grew from 55.13 billion yuan in 1998 to 1,416.99 billion yuan in 2015. With the exception of 1998, 2014, and 2015 when growth rates were single-digit (8.2%, 9.9%, and 8.9% respectively), all other years experienced double-digit growth, peaking at 27.7% in 2004. Figure 1 [Figure 1: see original paper] shows that China's R&D expenditure as a percentage of GDP increased from less than 1% in 1998 to 1.84% in 2011, while per capita expenditure rose from 73,000 yuan per person in 1998 to 377,000 yuan per person in 2015.

In terms of research output, analysis of paper counts in the Clarivate Analytics database [3] reveals that China's paper production has grown continuously among the world's six major paper-producing countries. Figure 2 [Figure 2: see original paper] illustrates the annual distribution of paper counts and citation frequency ratios relative to the United States for major countries worldwide (1991-2011). Over the past two decades, China's paper output increased from just over 6,000 in 1990 to more than 120,000 in 2011. As shown in Figure 2-a, China surpassed the United Kingdom in 2004, and Germany and Japan in 2005 in terms of paper quantity. However, Figure 2-b indicates that when China exceeded these countries in paper counts during 2004-2005, its citation frequency remained far lower. Overall, China's average citations per paper increased from approximately 26% of the U.S. level before 1995 to about 50% after 2006.

Despite the continuous growth in total R&D expenditure, China's basic research funding has remained at approximately 5% for years, significantly lower than major countries worldwide (China: 4.7% (2013), United States: 16.5% (2013), France: 24.4% (2011), Japan: 12.3% (2011), South Korea: 18.1% (2011), Russia: 16.5% (2012)) [4]. This phenomenon has attracted attention from top international journals such as *Nature* and *Science* [5-8]. Domestic scholars including Zhang Xian'en, Tang Ting, and Peng Kefeng [9-11] have also examined this issue from various perspectives, while Wang Lizheng, Yang Aihua, and Cheng Yao [12-14] argue that basic research investment remains insufficient. Yao Changle and Gao Changlin [15] compared China's scientific investment with that of developed countries across different periods, concluding that basic research investment should be strengthened.

In summary, alongside China's rising scientific and technological development level, two major dilemmas coexist: first, the simultaneous existence of insufficient scientific investment (particularly in basic research) and 质疑 about research funding efficiency; second, the coexistence of continuously increasing S&T funding and 质疑 about research output. The former primarily comes from the scientific community and public policy scholars, pointing to issues of scientific funding effectiveness, while the latter originates from the general public. The only solution to these dilemmas is to evaluate China's science and technology policies, particularly long-standing funding programs, over extended periods to provide micro-level data support for policy formulation.

2 Literature Review

Research on research project evaluation can be categorized into three levels according to evaluation objects: (1) micro-level, referring primarily to individual researchers, research groups, and specific projects; (2) meso-level, referring mainly to research institutions such as institutes, universities, and companies, as well as research programs and disciplinary fields; and (3) macro-level, referring primarily to nations (including regions and departments), regional and international organizations. This study focuses on scientists funded by specific research projects, defined as the micro-level.

In international micro-level research, J.M. Benavente [16] used regression discontinuity design to evaluate Chile's FONDECYT fund, finding that it significantly promoted funded researchers' paper output compared to the control group. A. Arora [17] examined the relationship between NSF funding and researchers' paper output, discovering that NSF funding had a more pronounced effect on early-career economists, with each additional \$10,000 in funding increasing output by 12 papers. D. Popp [18] analyzed R&D investment data from 15 OECD countries and WOS output data from 1991-2011, finding that each \$1 million in public R&D investment generated 1-2 papers for the investing country. P. Azoulay [19] studied the impact of NIH funding on patent output, finding that NIH funding more effectively promoted private patent production, with each \$10 million increase in funding generating 2.3 additional patents.

Domestic research has focused almost exclusively on NSFC projects, using evaluation indicators including international papers, domestic papers, and questionnaire surveys. Shi Xiaomin [20] found that Distinguished Young Scientist Fund project papers had higher citation frequencies 3-4 years after publication than key projects. Hu Ping [21] discovered uneven and differential funding outcomes for the Distinguished Young Scientist Fund. Zhou Ping [22] found that NSFC projects held an absolute leading position in Chinese paper publication.

These studies share several characteristics: (1) international research primarily employs econometric models for horizontal comparison across different research projects, without longitudinal comparison before and after project implementation; (2) domestic research uses statistical methods and bibliometric analysis of publication counts and citations to evaluate research output quantity and quality, failing to assess science and technology policies from a "causal effect" perspective; (3) no studies have evaluated funding programs over extended time series from an implementation cycle perspective. Therefore, this study addresses three questions: (1) What is the funding effect of science and technology funding programs on scientists' research paper output efficiency? (2) How do funding effects change throughout the historical implementation cycle of science and technology funding programs, and what are the possible reasons? (3) Which years show the best funding effects during the implementation cycle, and what are the possible reasons? Using the 20-year anniversary of the Distinguished Young Scientist Fund in earth science as a case study, this research attempts to

answer these questions through theoretical modeling based on existing qualitative research, followed by empirical analysis using the DID model.

3 “Environment-Motivation-Behavior” Model

3.1 Research Object

With support from the Earth Science Division of NSFC, this study examines the National Science Fund for Distinguished Young Scholars in Earth Science (hereafter “Earth Science Distinguished Young Scientist Fund” or “Distinguished Young Scientist Fund”). Established in 1994, the Distinguished Young Scientist Fund represents China’s first high-intensity youth talent cultivation fund, supporting young scientists who have achieved outstanding results in basic research to independently choose research directions and conduct innovative research, thereby cultivating excellent academic leaders at the forefront of global science and technology [23]. By 2013, the fund had operated for 20 years, supporting 3,004 individuals with total funding of 4.41 billion yuan across eight disciplinary fields. Among them, the Earth Science Distinguished Young Scientist Fund (hereafter “Distinguished Young Scientist Fund”) supported 311 individuals, accounting for 10.35% of total recipients. This study groups Distinguished Young Scientist Fund recipients by funding year, resulting in 20 groups denoted as i (e.g., $i=1$ for scientists funded in 1994, and so on). The analysis involves both the historical timeline of the Distinguished Young Scientist Fund implementation and the output timeline of funded scientists. Since scientists received funding in different years, the output timeline was preprocessed by setting the funding year as time point 0, incrementing by 1 for each subsequent year, and decrementing by 1 for each preceding year.

3.2 “Environment-Motivation-Behavior” Analytical Framework

3.2.1 Review of Qualitative Research on the Distinguished Young Scientist Fund This study analyzed six existing qualitative research papers on Distinguished Young Scientist Fund recipients, chronologically labeled as Literature 1-Literature 6 [24-29], with details presented in Table 1. Literature 1 [24] reviewed the early implementation of the Distinguished Young Scientist Fund and detailed nine requirements that recipients had for the research environment. Literature 2 [25] found that recipients generally categorized key factors influencing their innovative achievements as: external environment (research environment, social environment, living environment), significant others (mentors, leaders, family, academic peers, friends, academic authorities), and personal factors (research ideas and individual capabilities). Literature 3 [26] identified problems in the current research environment based on scientists’ statements: (1) emphasis on engineering applications over basic research; (2) overly broad evaluation criteria requiring not only research papers but also engineering projects and teaching tasks; (3) management incentives using “one-size-fits-all” evaluation schemes focusing on paper quantity rather than quality; (4)

insufficient supporting research funding; (5) inadequate cultivation of young talent, hindering the formation of cohesive research teams. Literature 4 [27] found that the research environment affected Distinguished Young Scientist Fund recipients' work through physical, social, and membership factors. Literature 5 [28] analyzed recipients' curricula vitae and research output to examine growth characteristics of management division recipients. Literature 6 [29] used bibliometric methods to study recipients' research output and interpreted findings in relation to the research environment.

3.2.2 “Environment-Motivation-Behavior” Analytical Framework

Based on the above analysis, research on the effectiveness of the Distinguished Young Scientist Fund can be summarized into three aspects: scientists' research motivation (motivation), the research environment (environment), and scientists' research behavior (behavior). This study proposes the “Environment-Motivation-Behavior” theoretical framework (Figure 3 [Figure 3: see original paper]). The framework comprises three elements: motivation, environment, and behavior. Specifically, scientists' research behavior is determined by their research motivation, which is constrained and influenced by internal and external factors collectively categorized as the research environment. The environment can be divided into internal and external components. The Distinguished Young Scientist Fund is treated as an external incentive within this framework, belonging to external environmental factors. Scientists in different environments perceive the Distinguished Young Scientist Fund as different incentives, leading to changes in their internal research environment, which in turn alters their research motivation and ultimately affects their research behavior.

The research environment refers to the social environment in which researchers operate, including economic development level, local education level, science and technology policies, and cultural atmosphere. To clearly explain relevant assumptions, the research environment most directly and closely related to Distinguished Young Scientist Fund recipients is summarized into four aspects: S&T policy incentives, research evaluation mechanisms, administrative leadership mechanisms, and salary and benefits. S&T policy incentives primarily refer to the Distinguished Young Scientist Fund policy itself. In subsequent empirical research, scientists are divided into treatment and control groups, where the treatment group receives Distinguished Young Scientist Fund support. This funding constitutes an external incentive factor that changes scientists' external research environment, thereby affecting their research motivation and ultimately manifesting as changes in research behavior.

Research motivation refers to scientists' purposes for conducting research. Based on existing literature, scientists' research motivation is categorized into two types: scientific motivation and career motivation. Scientific motivation refers to research driven by curiosity to solve scientific problems, while career motivation refers to treating research as a means of livelihood to obtain living security,

career advancement, and social status. Theoretically, from Maslow's hierarchy of needs perspective, career motivation represents lower-level needs (salary, benefits, healthcare, housing, children's education, institutional matching funds) compared to scientific motivation (social status, self-actualization). According to Herzberg's two-factor theory, motivation factors related to work itself generate job satisfaction, while hygiene factors related to work environment prevent dissatisfaction. Scientific and career motivations are not independent but both constrained by the research environment, such as evaluation mechanisms and administrative leadership. Lower-level career motivation constrains scientific motivation; only when basic needs are met can scientists pursue higher-level scientific motivation under scientific motivation drive.

Research behavior refers to the choices scientists make during research. This study categorizes scientists' ultimate research behavior as publishing papers or other output forms. Specifically, scientists face decisions about whether to pursue paper publication, other research activities, or non-research activities. Alternative behaviors include engineering applications, technology transfer, patent applications, social services, or teaching focus, each yielding different benefits. Publishing papers satisfies scientists' curiosity, gains peer recognition, and achieves self-actualization.

The relationship between environment, motivation, and behavior assumes that research motivation determines research behavior. Within their existing research environment, scientists receiving Distinguished Young Scientist Fund support experience changed external environments, which different scientists perceive differently based on their original environments, ultimately leading to different research behaviors. Since the analyzed scientists span 1994-2013—a period of rapid economic and S&T development in China—scientists funded in different years faced substantially different research environments. Consequently, they perceived the Distinguished Young Scientist Fund differently, affecting their motivation and behavior variably. Therefore, empirical analysis of each year's funding effectiveness may show that some years support the hypothesis while others do not, with subsequent sections analyzing each year's specific circumstances.

3.4 Theoretical Hypotheses

Based on the “Environment-Motivation-Behavior” framework, we propose: (1) Hypothesis H -a: The Distinguished Young Scientist Fund significantly improves the research paper output efficiency of group i recipients; (2) Hypothesis H -b: The Distinguished Young Scientist Fund does not significantly improve the research paper output efficiency of group i recipients. Notably, because the analyzed scientists span 1994-2013—a period of rapid Chinese economic and S&T development following 16 years of reform and opening-up—scientists from different years faced vastly different research environments. For instance, 1994 recipients experienced dramatically different environments than 2008 recipients, leading to different perceptions of the fund's incentive, varied impacts on moti-

vation, and ultimately different research behaviors. Therefore, annual empirical analysis of funding effectiveness may yield results supporting the hypothesis for some years but not others, with subsequent sections providing detailed analysis for each year.

4 Research Data and Empirical Scheme

4.1 Conceptual Definition

G.A. Crespi [30] discussed scientific research outputs, establishing models to analyze inputs and outputs. Research findings indicate that scientific research produces multiple outputs broadly categorized into three types: (1) new knowledge; (2) high-quality human resources; (3) new technologies and other knowledge with socioeconomic impact. This study focuses on the first output type. For scientists engaged in basic research, primary output consists of scientific papers, which serve as important carriers of research 成果. When examining the growth of Distinguished Young Scientist Fund recipients, this study selects scientists' research paper output as the metric for measuring research paper output efficiency. Drawing on existing methods for measuring funding effects [31], we define the number of SCI and SSCI papers published annually by Distinguished Young Scientist Fund recipients (denoted as *Pubs*) as the research paper output efficiency indicator. We select SCI and SSCI journals indexed in Clarivate's Web of Science as the output source for Distinguished Young Scientist Fund recipients.

4.2 Data Preparation

After one and a half years of retrieval, verification, screening, and organization, we constructed the following databases:

Database 1 contains basic information and output data for all 311 Distinguished Young Scientist Fund recipients in the Earth Science Division from 1994-2013. With support from the Earth Science Division of NSFC, we reviewed 47 project proposals from 1994-1998 and 240 final reports from 1994-2010 (2011-2013 recipients had not yet completed their projects). Four individuals were excluded due to withdrawn funding, death during funding period, death one year after funding ended, or terminated funding, leaving 307 subjects for analysis. Through content analysis of proposals and final reports and review of recipients' institutional homepages, we constructed a basic database including birth date, education, and work experience.

Database 2 contains output data for Distinguished Young Scientist Fund recipients. H.F. Moed [32] reviewed bibliometric methods for evaluating Chinese research, recommending the use of the international Web of Science database [33] while excluding domestic journals. This study adopts Moed's approach. Among the secondary disciplines in geography, human geography research appears primarily in SSCI, while other recipients' outputs are mainly in SCI.

Using “name + institution” search criteria, we retrieved paper data from the Web of Science Core Collection [34] through May 31, 2015. Retrieved records were further screened using discipline, institutional department, and co-author information. After multiple rounds of verification, the final output database comprised 31,098 papers.

4.3 Empirical Scheme

4.3.1 Selection Bias The fundamental question of interest is the causal effect of the Distinguished Young Scientist Fund—in other words, whether funded scientists would have achieved the same outcomes without funding. If recipient selection were randomized, we could directly compare outcomes between funded and unfunded scientists. However, selection is based on “choosing the best among the best,” creating selection bias that amplifies treatment effects. Selection bias refers to the discrepancy between what we truly need to study and what we can actually measure, known as the “First Question” in economics. Current methods to address selection bias include randomized experiments, instrumental variables, difference-in-differences, regression discontinuity, matching, and synthetic control methods. Since Distinguished Young Scientist Fund selection is not lottery-based, randomized experiments are infeasible. While instrumental variables and regression discontinuity have strong theoretical foundations, suitable instruments require extensive data and trial-and-error, and regression discontinuity requires final report scores meeting strict discontinuity conditions that are difficult to satisfy. Matching and synthetic control methods require lists of unsuccessful applicants who never reapply—unavailable since top scientists can apply annually until age 45. Therefore, this study uses the difference-in-differences method to evaluate funding effectiveness.

4.3.2 Control Group Construction Since unsuccessful applicants in the same year cannot serve as a control group, this study uses the same physical individuals at different time points as controls. Using 1994 recipients as an example: funding ran from January 1995–December 1997 (3 years), ending in January 1998. Therefore, pre-1998 recipients were unfunded during 1995–1997, allowing us to use scientists funded in 1998 and later as the control group for 1994 recipients. This logic extends to other years, yielding the treatment and control group divisions shown in Table 2. In Table 2, i , j , and t represent the i th DID model, the j th scientist, and year t , respectively. For the 1994 fund (DID_{1}): (1) $i=1$, denoted as DID_{1}; (2) in DID_{1}, $D_{\{1\}}=1$ if scientist j belongs to the 1994 group, otherwise $D_{\{1\}}=0$; (3) since 1994 funding covered 1995–1997, $T_{\{1\}}=1$ if scientist j in DID_{1} is in year t during 1995–1997, otherwise $T_{\{1\}}=0$. Other years follow similarly.

4.3.3 Regression Model This study uses regression analysis to estimate DID estimators:

$$Y_{ijt} = \beta_{0i} + \beta_{1i}D_{ij} + \beta_{2i}T_{ijt} + \beta_{3i}D_{ij} \cdot T_{ijt} + \sum \beta_{ij} \cdot X_{ijt} + \eta_{ij} + u_{it} + \varepsilon_{ijt} \quad (1)$$

In equation (1), i , j , and t represent the j th scientist in the i th DID model (DID) in year t . The dependent variable Y represents the total number of SCI and SSCI papers published by scientist j in year t in the i th DID model (denoted as $Pubs$). D is a treatment group dummy variable ($D = 1$ for treatment group, 0 otherwise). T is a time dummy variable ($T = 1$ during funding period, 0 otherwise). The interaction term $D \cdot T$ represents the DID estimator and is the core variable of interest, with coefficient β_{3i} measuring the funding effect—the average change in the dependent variable for funded versus unfunded scientists. X represents covariates characterizing scientist attributes, η denotes unobservable individual effects, u denotes unobservable time effects, and ε is the random error term. Notably: (1) funding amount, being policy-determined, cannot be included as a control variable [35] to avoid sample selection bias; (2) R&D conditions (research teams, environment) that vary across both time and individuals are incorporated into unobservable individual and time effects. Detailed variable definitions appear in Table 3.

In practice, we control for individual effects using dummy variables for all scientists in each treatment-control combination and for time effects using dummy variables for each output year. Since the number of dummy variables differs across DID models, and given that our dependent variable is research paper output efficiency, the model expands to:

$$\begin{aligned} Pubs_{ijt} = & \beta_{0i} + \beta_{1i}D_{ij} + \beta_{2i}T_{ijt} + \beta_{3i}D_{ij} \cdot T_{ijt} + \beta_{4i}Age_{ijt} \\ & + \beta_{4i}Male_{ij} + \beta_{4i}Div_{ij} + \beta_{4i}IDoctor_{ij} + \beta_{4i}Doctype_{ij} \\ & + \beta_{4i}Worktype_{ij} + \beta_{4i}Workcity_{ij} + \text{individual effects} + \text{time effects} + \varepsilon_{ijt} \end{aligned} \quad (2)$$

5 Empirical Results

5.1 DID Model Group Discussion Rules

Since 2009-2013 recipients had not completed their projects when the database was constructed, analysis covers 1994-2008 (15 years), yielding 15 DID models (DID_{1}, DID_{2}, ..., DID_{15}) (Figure 4 [Figure 4: see original paper]). Each annual analysis includes nine steps: (1) empirical model, (2) variable definitions, (3) theoretical hypotheses, (4) overall descriptive statistics, (5) control variable descriptive statistics, (6) dependent variable mean descriptive statistics, (7) funding effect empirical results, (8) hypothesis verification, and (9) parallel trend verification. With 15 models requiring 135 steps and involving up to 20 tables, detailed discussion of all models would be excessively lengthy. Therefore, we present detailed analysis of DID_{8} as a representative example, summarizing the funding effect coefficients β_{3i} from other models.

5.2 DID_{8} Empirical Analysis and Results

5.2.1 Average Funding Effect of DID_{8} (1) DID_{8} Empirical Model and Variable Definitions. Based on equation (1), the DID_{8} regression model is:

$$Y_{8jt} = \beta_{08} + \beta_{18}D_{8j} + \beta_{28}T_{8jt} + \beta_{38}D_{8j} \cdot T_{8jt} + \sum \beta_{8j} \cdot X_{8jt} + \eta_{8j} + u_{8t} + \varepsilon_{8jt} \quad (3)$$

DID_{8} analyzes the 2001 Distinguished Young Scientist Fund funding effect ($i=8$). Variable definitions follow equation (1) with i replaced by 8, and D_{i} and T_{i} values follow Table 3. Coefficient β_{38} represents the 2001 funding effect. Based on the “Environment-Motivation-Behavior” framework, we propose: (1) Hypothesis H_{8} -a: The 2001 Distinguished Young Scientist Fund improves scientists’ research paper output efficiency; (2) Hypothesis H_{8} -b: The 2001 Distinguished Young Scientist Fund does not improve scientists’ research paper output efficiency.

(2) DID_{8} Descriptive Analysis. Before regression analysis, we examine overall characteristics. Table 4 shows treatment group scientists were born between 1956-1969, while control group scientists were born between 1962-1978. Notably, the youngest control group scientist was 23 years old in 2001 (denoted scientist a), potentially a recent graduate with zero paper output during 2002-2005. The DID modeling program automatically handles such cases by treating zero-output observations as 0 and effectively excluding unqualified scientists like a —a consistent approach across all DID models.

Table 5 analyzes six aspects of the treatment and control groups: (1) Gender: 87.5% male in the treatment group versus 94.8% in the control group; (2) International doctoral education: 56.2% of treatment group held foreign PhDs versus 83% of controls; (3) Primary discipline: Scientists were divided into six earth science sub-disciplines, with proportions ranging 6%-30% in both groups, covering all six sub-disciplines; (4) Doctoral institution type: Treatment group PhDs concentrated in the Chinese Academy of Sciences (CAS) and foreign institutions, with none from ordinary universities, while control group PhDs concentrated in CAS, 985 universities, and foreign institutions; (5) Host institution type: All treatment group scientists were affiliated with CAS or 985 universities, while controls were primarily in CAS and 985 universities with fewer in 211 or ordinary institutions; (6) Host institution location: Both groups concentrated in Beijing, with Nanjing and Guangzhou as other major locations.

(3) DID_{8} Empirical Results and Discussion. Table 6 presents regression results. Models 1-4 vary in controlling for individual and time effects. Results show the 2001 Distinguished Young Scientist Fund had a highly significant and robust effect on paper output efficiency, with recipients publishing 0.934 more papers annually than the control group—equivalent to 1.14 times the contemporaneous mean of 0.822 papers per scientist. Hypothesis H_{8} -a is supported.

5.2.2 DID_{8} Parallel Trend Test The DID method addresses endogeneity from time-invariant factors, eliminating unobserved confounding effects to identify causal parameters [35]. The primary requirement is the parallel trend assumption [36]. Table 7 shows DID_{8} parallel trend test results, with time series normalized such that 0 represents the 2001 funding year, 1 represents 2002 (post-funding year 1), -1 represents 2000 (pre-funding year 1), etc. Model 1 satisfies the parallel trend assumption, with non-significant differences between treatment and control groups before 2001 and highly significant differences thereafter.

5.3 Summary of DID_{1}-DID_{15} Empirical Analysis Results

Table 8 summarizes the 15 models' funding effects on research paper output efficiency (1994-2008). Column 5 shows each year's funding effect relative to the mean research output. Notably, each annual DID model analyzed four specifications: controlling for neither individual nor time effects, controlling only individual effects, controlling only time effects, and controlling both. Results showed identical β values and significance across all four specifications for each dimension, with only slight standard error differences. Therefore, we report coefficients from the full specification (controlling both effects) for all dimensions.

6 Summary and Discussion

To address endogeneity from sample selection bias, this study evaluated the impact of the 20-year Distinguished Young Scientist Fund Earth Science program on scientists' research paper output efficiency using the DID model. Empirical findings demonstrate:

- (1) The Distinguished Young Scientist Fund showed significant effects in 80% of years. From the "Environment-Motivation-Behavior" perspective, the fund serves as an incentive factor that changes recipients' research environment, affects their motivation, and ultimately determines research behavior choices. On one hand, funded scientists gain more adequate research support, can independently choose research fields, and have stronger motivation to conduct research, leading them to choose paper publication to satisfy curiosity and gain peer recognition. The significant positive effects in most years indicate the fund serves as a positive incentive. On the other hand, receiving the fund itself represents significant peer recognition and honor that could potentially create a negative incentive if scientists consider themselves "accomplished" and reduce effort. However, only three years showed non-significant effects, with all others showing significant positive effects.
- (2) Compared with similar international funding programs, the Distinguished Young Scientist Fund demonstrates more significant and larger effects on research paper output efficiency. Table 8 shows that funded scientists

published 0.412-3.234 additional papers annually (2.06-16.17 papers over 5 years) compared to controls, and 0.426-3.277 additional papers compared to same-period scientists. B. Jacob [37] found that NIH R01 funding increased paper output by only 1.2 papers over 5 years, indicating the Distinguished Young Scientist Fund's effect substantially exceeds that of NIH funding.

- (3) Temporally, the strongest funding effect occurred in 2002 (3.234 additional papers annually), while the weakest was in 1995 (0.412 additional papers). The non-significant effects in 2007-2008 may reflect China's substantial R&D investment increase during 2006-2009, providing alternative funding sources and academic resources that reduced the Distinguished Young Scientist Fund's marginal effect.
- (4) This study provides a new method for science and technology policy evaluation. By constructing control groups and using econometric models for causal inference, it advances bibliometric research from simple measurement to "causal effect inference," enabling assessment of counterfactual scenarios.

Three areas require further research: (1) Analysis of 2009-2013 funding effects once projects are completed; (2) Expansion to other NSFC divisions to generalize findings beyond earth science; (3) Examination of top-tier journal output, citation impact, heterogeneity across recipient types, and cost-effectiveness analysis.

References

- [1] China Science and Technology Statistics [EB/OL]. [2017-09-10]. <http://www.sts.org.cn/>.
- [2] National Bureau of Statistics of the People's Republic of China [EB/OL]. [2017-09-11]. <http://www.stats.gov.cn/>.
- [3] Web of Science All Databases Homepage [EB/OL]. [2017-10-05]. http://apps.webofknowledge.com/UA_{{GeneralSearch}}_{{input}}.do?product=UA&search{mode}=General
- [4] Basic Research Funding: The "5% Mystery" or a "Pattern" with Chinese Characteristics? [EB/OL]. [2017-10-09]. http://www.sohu.com/a/197079231_753093.
- [5] Wei Y. Policy: boost basic research in China [J]. *Nature*, 2016, 534: 467-469.
- [6] Sun YT, Cao C. China: standardize R&D costing [J]. *Nature*, 2016, 536(4): 30-31.
- [7] Noorden RV. China by the numbers [J]. *Nature*, 2016, 534(1): 452-453.
- [8] Xin H. Five-year plan boosts basic research funding [J]. *Science*, 2016, 351(6280): 1382-1382.
- [9] Peng Kefeng. Basic research funding: academic calls for increase [EB/OL]. [2017-12-15]. <http://news.sciencenet.cn/htmlnews/2014/4/291179.shtm>.
- [10] Tang T, Liu YY, Wang C. Toward a science and technology powerhouse: how to strengthen basic research [EB/OL]. [2017-12-15]. <http://news.sciencenet.cn/htmlnews/2017/9/389420.shtm>.
- [11] Zhang Xian'en, Liu Y, Zhou C, et al. International comparison of basic research connotation and investment statistics [J]. *China Soft Science*, 2017, 5(1): 131-138.
- [12] Cheng Y. Analysis of current basic research investment in

China [J]. Journal of Huazhong University of Science and Technology (Social Science Edition), 2006, 20(5): 62-66. [13] Yang AH. Comparison and analysis of basic research investment in China (2001-2010) [J]. Journal of Changsha University of Science & Technology (Social Science), 2012, 27(3): 34-38. [14] Wang LZ. Analysis and policy recommendations for basic research funding sources in China [J]. Science of Science and Management of S&T, 2011, 32(12): 26-31. [15] Yao CL, Gao CL. Current situation analysis and policy recommendations for basic research funding in China [J]. China Science and Technology Forum, 2011(3): 5-9. [16] BENAVENTE JM, CRESPI G, GARONE LF, et al. The impact of national research funds: a regression discontinuity approach to the Chilean FONDECYT [J]. Research policy, 2012, 41(8): 1461-1475. [17] ASHISH A, ALFONSO G. The impact of NSF support for basic research in economics [R]. Econpapers, 2005(79/80): 1-28. [18] POPP D. Using Scientific publications to evaluate government R&D spending: the case of energy [R]. Cambridge: Cesifo working paper, 2015: 8-13. [19] AZOULAY P, GRAFF ZIVIN JS, LI D, et al. Public R&D investments and private-sector patenting: evidence from NIH funding rules [R]. Camb ridge: NBER working paper, 2015: 1-30. [20] Shi XM, Peng J, Guan JC. Comparison of output performance between NSFC key projects and distinguished young scientist projects [J]. Science and Technology Management, 2004, 6(1): 128-130. [21] Hu P, Wu SC, Li C, et al. Evaluation research on funding outcomes of distinguished young scientists in China [J]. Science of Science and Management of S&T, 2009, 3(1): 190-194. [22] Zhou P, Zhang X, Zhou DM. Comparison of Chinese paper output performance of major Chinese funding programs [J]. Science and Technology Management Research, 2012(19): 43-48. [23] National Natural Science Foundation of China. National Science Fund for Distinguished Young Scholars online [EB/OL]. [2017-09-10]. <http://www.nsf.gov.cn/portal0/tab313/>. [24] CAO C, Suttmeier RP. China's new scientific elite: distinguished young scientists, the research environment and hopes for Chinese science [J/OL]. [2018-02-25]. <https://darkwing.uoregon.edu/~chinaus/publications/China%20new%20scientific%20elite.pdf>. [25] Zhao L, Jin SH, Sun L, et al. Factors influencing creativity development of young innovative talents—a qualitative analysis based on interviews with 25 Distinguished Young Scientist Fund recipients [J]. Journal of China Youth University for Political Sciences, 2011, 30(3): 68-73. [26] Gao SH. Research on the cultivation model of Distinguished Young Scientist Fund recipients at Harbin Engineering University [D]. Harbin: Harbin Engineering University, 2011. [27] Wan Y. The impact of individual growth and research environment on research output of Distinguished Young Scientist Fund recipients—taking life science recipients as an example [D]. Beijing: Institute of Psychology, Chinese Academy of Sciences, 2014. [28] Zhang Y. Research on academic growth characteristics of high-level talents from a bibliometric perspective—taking management division Distinguished Young Scientist Fund recipients as an example [D]. Nanjing: Southeast University, 2016. [29] Zhang Y. Research on scientific output of top young Chinese biomedical scientists—taking Distinguished Young Scientist Fund recipients and Changjiang Scholars as examples [D]. Taiyuan: Shanxi Medical University, 2016. [30] CRESPI GA, GEUNA. An empirical study of

scientific production: a cross country analysis, 1981-2002 [J]. Research policy, 2008, 37(4): 565-579. [31] GUSH J, JAFFE A, LARSEN V, et al. The effect of public funding on research output: the New Zealand Marsden Fund [R]. NBER WORKING PAPER No. 216522015, (3): 1-43. [32] MOED HF. Measuring China's research performance using the Science Citation Index [J]. Scientometrics, 2002, 53(3): 281-296. [33] GARFIELD E. Citation analysis as a tool in journal evaluation [J]. Science, 1972, 178(4060): 471-479. [34] Web of Science Core Collection help: searching the "grant number" field [EB/OL]. [2017-09-10]. http://images.webofknowledge.com/WOKRS521R5/help/zh_{CN}/WOS/hs_{grant}_{number}.htm [35] Zhao XL. Basically useful econometrics [M]. Beijing: Peking University Press, 2017: 161-162. [36] ANGRIST JD. Mostly harmless econometrics: an empiricist's companion [M]. Princeton: Princeton University Press, 2010: 165-169. [37] JACOB BA, LEFGREN L. The impact of research grant funding on scientific productivity [J]. Journal of public economics, 2011, 95(9): 1168-1177.

Author Contributions: Tian Renhe: Responsible for framework design, data analysis, manuscript writing, and viewpoint development;
Zhang Zhiqiang: Responsible for conceptual guidance and writing supervision;
Yu Jie: Assisted with mathematical modeling and participated in manuscript revision;
Wang Ping: Assisted with mathematical preprocessing and modeling.

Abstract: [Purpose/significance] Science and technology innovation is an important strategic support for modern economic systems. After the reform and opening-up, in order to promote China's economic development, a series of science and technology policies were introduced and implemented to promote the development of science and technology. However, the funding effects of these science and technology policies and the changes in funding effects during the historical cycle of science and technology policy implementation have caused widespread concern. [Method/process] This paper conducted an empirical research of the National Science Fund for Distinguished Young Scholars Earth Science Project that has been in operation for 20 years (Jie Qing fund) as an example. In order to overcome the endogeneity problem caused by sample selection bias, the study used DID model to assess the effect of fund for distinguished young on the research papers output efficiency of scientists, and put forward the "environment-motivation-behavior" model to explain the research result. [Result/conclusion] The empirical results show that between 1994 to 2008, the significant funding effect of Jie Qing fund reached 12 years, with a significant proportion 80%. Scientists who have received Jie Qing fund (Jie Qing Scientists) were able to published more 0.412 to 3.234 papers per year than those who have not received Jie Qing fund. After data conversion, compared with other scientists who in the same period (Same period Scientists) of Jie Qing Scientists, Jie Qing scientists can published more 0.426-3.277 papers per year. The largest funded effect of Jie Qing fund was in 2002, but the funding effect was not significant in 2007 and 2008. When evaluating the effect of science and technology policy support, the control group can be constructed by this method in this paper to achieve the research purpose of causal effect inference.

Keywords: data science, DID model, National Science Fund for Distinguished Young Scholars, science and technology policy evaluation

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.