The Impact of Preschool Education on Academic Achievement in Junior High School Students: An Analysis Based on CEPS Data

Xitong Li
Early Childhood Education (Teacher Education), Luoyang Institute of Technology, Luoyang, Henan, China

Abstract: **This study employs descriptive statistics, correlation analysis, and multiple regression methods based on China Education Panel Survey (CEPS) data to empirically examine the impact of preschool education experiences on junior high school students' academic performance.Findings reveal that students who received preschool education significantly outperformed their peers in Chinese, mathematics, and English. Parental education level exerts a significant positive influence on children's academic performance, with preschool education demonstrating particularly pronounced effects on mathematics and English scores. This study recommends increasing investment in preschool education resources, with special attention to early education opportunities for rural and disadvantaged children, to collectively promote educational equity.**

Keywords: **Preschool Education; Academic Achievement; CEPS; Parental Education; Empirical Research**

1. Problem Statement

Preschool education constitutes a crucial foundational stage in an individual's lifelong learning journey, playing a pivotal role in shaping children's physical development, cognitive growth, social-emotional skills, and learning dispositions. High-quality preschool education not only enhances early cognitive abilities and problem-solving skills but also cultivates logical thinking and learning habits. This early capital accumulation may yield sustained positive impacts throughout an individual's subsequent schooling and long-term development.

As China's socioeconomic development and educational standards advance, the importance of preschool education has gained widespread recognition among academics and policymakers. However, despite overall growth in preschool resources, disparities persist in resource allocation across regions, between urban and rural areas, and among different socioeconomic groups.Some children, due to family circumstances or geographic location, lack access to high-quality preschool education, potentially placing them at a relative disadvantage from the outset of their educational journey. Whether such early disparities persist and influence academic performance in junior high school remains an empirical question demanding urgent validation. As educator John Dewey emphasized, education is life and growth. As the starting point of this "life" and "growth," the value of preschool education warrants in-depth exploration.Building upon Dewey's pragmatic educational philosophy and adapting it to China's context, Tao Xingzhi proposed the theory of life education. Its core principles-localization, democratization, and practicality-constitute the foundational ideology for the localization of modern Chinese preschool education. Central to this approach is the educational philosophy that "life is education, society is the school, and teaching, learning, and doing are integrated." This stance opposes the "elitism and Westernization" of preschool education, laying a solid foundation for educational equity.

Building upon this, this study utilizes nationally representative data from the China Education Panel Survey (CEPS) to systematically examine the long-term effects of preschool education experiences on junior high students' academic performance. By controlling for key factors such as family background, particularly parental education levels, the research seeks to answer the following core questions: First, does preschool education significantly and positively impact junior high students' Chinese, mathematics, and English grades?Second, does this effect vary across different subjects? Third, to what extent do family background factors,

especially parental education, influence students' academic performance? By providing empirical answers to these questions, this study aims to offer scientific evidence for optimizing preschool education policies and promoting educational equity.

2. Literature Review

The long-term impact of preschool education on individual development remains a research hotspot across educational economics, developmental psychology, and sociology. Overseas research began earlier, with classic studies-particularly those based on longitudinal projects like the U.S. Perry Preschool Program-demonstrating that preschool education not only significantly enhances children's early cognitive abilities but also profoundly influences adult outcomes such as income, employment rates, and crime rates. Research by Nobel Prize-winning economist Heckman indicates that the return on investment in early childhood is far higher than that of skills training for adults. [1,2] Furthermore, using a causal identification analysis framework, Heckman conducted a systematic and in-depth study of the impact of preschool experiences on children's cognitive development during elementary school. The results show that preschool education has a significant positive effect on both academic performance and skill development during elementary school. Compared to children in the control group who did not receive preschool education, students with preschool experience have a more solid academic foundation, stronger learning skills, and achieve better academic performance upon entering elementary school. [3] Results from the OECD's PISA assessments also reveal that, after controlling for family background, 15-year-old students who received preschool education demonstrate superior academic performance.

In recent years, with the release of large-scale educational survey data in China, empirical research on the impact of preschool education has increased domestically. Most studies confirm the positive effects of preschool education on children's cognitive abilities and academic performance. For example, Gong Xin et al. [4], using data from the China Family Panel Studies (CFPS), found that preschool education has a significant long-term positive effect on the social-emotional skills of rural children.Yuan Yuzhi and Zhao Yi [5] used CEPS data and propensity score matching (PSM) to estimate that preschool education increases cognitive abilities by approximately 0.22 standard deviations among junior high school students. Similarly, Ye Shengjun [6], also using CEPS data and PSM, found that preschool education not only enhances academic performance and cognitive abilities but also contributes to improved mental health.

However, access to preschool education is unequal, and its effects are significantly moderated by family background. Tu Rongzhen et al. [7] noted that family economic capital, cultural capital, and social capital not only influence children's access to preschool education but also weaken its positive effects on academic achievement.Wang Dian et al. [8] found in their study of visually impaired students that preschool education effectively promotes cognitive and agreeableness development, but access to preschool education for this group is constrained by factors such as age and household size. Gong Xin and Li Zhenyi [4] focused on noncognitive abilities, discovering that preschool education helps enhance junior high students' open-mindedness and self-discipline while reducing negative emotions.

In summary, existing research has demonstrated the long-term value of preschool education from multiple perspectives-cognitive, non-cognitive, and educational equity. However, most studies either focus on specific groups or fail to systematically compare differences across subjects.Furthermore, while causal inference methods like PSM are increasingly adopted, there remains room for deepening research through multivariate regression studies that systematically examine preschool education's impact on academic performance across different subjects using national-level samples while controlling for family background factors.Building upon existing literature, this study utilizes CEPS data to systematically examine the effects of preschool education on junior high students' Chinese, mathematics, and English performance through descriptive statistics, correlation analysis, and multiple regression. It also investigates the moderating role of family background, aiming to provide more detailed empirical evidence for advancing preschool education equity and quality improvement.

Based on the above discussion, this study proposes the following hypotheses:

Hypothesis 1: Controlling for students' family background, preschool education significantly and positively impacts junior high students' Chinese language performance.
Hypothesis 2: Controlling for family background, preschool education significantly and positively affects junior high students' mathematics performance.
Hypothesis 3: Controlling for students' family background, preschool education significantly and positively affects junior high students' English performance.

3. Research Methodology

3.1 Data Source

This study utilizes baseline data from the China Education Panel Survey (CEPS) for the 2013–2014 academic year. Employing a multistage stratified random sampling method, the survey covers 112 junior high schools across 28 counties and districts nationwide, yielding a nationally representative sample.The original sample size was 19,486 junior high school students. After data cleaning and missing value handling, the effective sample size for each analysis model was approximately 18,800.

3.2 Variable Definitions

· Dependent Variables: Students' 2013 midterm exam scores, including raw scores for Chinese (tr_chn), Mathematics (tr_mat), and English (tr_eng).
· Core independent variable: Preschool participation (c01), defined as "whether preschool education was received after age 3," where 1 indicates "yes" and 2 indicates "no." For interpretability, it was converted to a dummy variable (1=yes, 0=no) in the analysis.
· Control Variables: Mother's educational attainment (stmedu) and Father's educational attainment (stfedu), both measured on a 9-level scale (1 = "no schooling," 9 = "graduate degree"). To simplify the model and focus on the aggregate effect of parental education, following the approach of Tu Rongzhen et al.[1], these variables are subsequently merged into "highest parental educational attainment" and converted into years of education for control purposes.

3.3 Analytical Strategy

This study employs a three-step progressive analysis strategy:
1. Descriptive Statistics: Present the distribution characteristics of core variables and preliminarily compare differences in subject scores and family backgrounds between students who received and did not receive preschool education.
2. Correlation Analysis ( ): Calculate Pearson correlation coefficients between preschool participation, parental education level, and the three subject scores to preliminarily explore the direction and strength of associations among variables.
3. Multivariate Linear Regression: Construct three regression models with Chinese, Math, and English scores as dependent variables. The core independent variable is preschool participation, while controlling for parental education level to estimate the "net effect" of preschool education on student academic performance after accounting for family cultural capital. Regression coefficients will be used to test the three research hypotheses proposed earlier.
Model 1 (Dependent variable: Chinese language scores)
$Chinese_i = \beta_0 + \beta_1 \times Preschool_i + \beta_2 \times MotherEducation_i + \beta_3 \times FatherEducation_i + \varepsilon_i$
Model 2 (Dependent variable: Math scores)
$Math_i = \beta_0 + \beta_1 \times Preschool_i + \beta_2 \times MotherEducation_i + \beta_3 \times FatherEducation_i + \varepsilon_i$
Model 3 (Dependent variable: English score)
$English\ Score_i = \beta_0 + \beta_1 \times Preschool\ Education_i + \beta_2 \times Mother's\ Education\ Level_i + \beta_3 \times Father's\ Education\ Level_i + \varepsilon_i$
Variable Descriptions
• Preschool Education_i: Core independent variable
• Mother's Education Level_i, Father's Education Level_i: Control variables
• β_0: Constant term; β_1, β_2, β_3: Regression coefficients, where β_1 is the core coefficient reflecting the net effect of preschool education on academic performance;
• ε_i: Random error term, following a normal distribution with mean 0 and constant variance.

4. Empirical Analysis Results

4.1 Descriptive Statistics

Table 1 shows that approximately 80% of students in the sample received preschool education after age 3. The average scores for Chinese language were 82.97, mathematics 78.12, and English 79.29. Parental education

levels were moderately high, with fathers' education levels slightly higher than mothers'.

4.2 Correlation Analysis

Table 2 shows that preschool participation (with reversed coding, where lower values indicate attendance) exhibits significant negative correlations with all three subject scores (Chinese: -0.084; Math: -0.106; English: -0.116), preliminarily supporting the research hypothesis that students who attend preschool achieve better academic performance. Parental education level shows significant positive correlations with children's scores across all subjects.

Table 1. Descriptive Statistics of Core Variables

Variable	Mean	Standard Deviation	Minimum	Maximum	Observed Value
Preschool Experience (Attended Preschool)	1.20	0.40	1	2	19,335
Chinese Language Score	82.97	20.67	1	143.50	19,001
Math Score	78.12	31.70	1	404.00	18,990
English Score	79.29	30.34	3	152.00	18,996
Mother’s Education Level	3.81	1.97	1	9	19,441
Father’s Education Level	4.18	1.99	1	9	19,441
Total	–	–	–	–	19,486

Table 2. Correlation Matrix of Core Variables

	c01	tr_chn	tr_mat	tr_eng	stmedu	stfedu
c01	1	-0.084***	-0.106***	-0.116***	-0.184***	-0.153***
tr_chn	-0.089***	1	0.662***	0.693***	0.187***	0.190***
tr_mat	-0.108***	0.681***	1	0.708***	0.222***	0.222***
tr_eng	-0.118***	0.706***	0.719***	1	0.263***	0.262***
stmedu	-0.165***	0.190***	0.222***	0.260***	1	0.644***
stfedu	-0.143***	0.202***	0.228***	0.267***	0.673***	1

Note: *p < 0.1, **p < 0.05, ***p < 0.01.

4.3 Regression Analysis Results

Table 3 presents the results of multiple linear regression. After controlling for parental education level:

Preschool experience significantly and positively influenced Chinese, mathematics, and English scores (all coefficients negative; dependent variables were raw scores, independent variable was "did not attend = 1"). Specifically, students who attended preschool scored approximately 2.83 points higher on average in Chinese, 5.33 points higher in mathematics, and 5.38 points higher in English.

· Hypotheses 1, 2, and 3 are all supported. The magnitude of the effect varies across subjects, with the boost to math and English scores being significantly greater than that for Chinese. This may relate to preschool education's emphasis on early logical thinking and language development.

Parental education level exerts a significant positive influence on children's performance across all three subjects, with the impact coefficient for fathers generally slightly higher than that for mothers. This aligns with relevant analytical findings and underscores the importance of family cultural capital.

Table 3. Regression Results of Preschool Education on Student Academic Performance

	(1)	(2)	(3)
	Student 2013 Chinese	Student 2013 Math	Student 2013 English
Did you attend preschool after age 3?	-2.830***	-5.334***	-5.381***
	(0.375)	(0.575)	(0.529)
Mother's Education Level	0.935***	1.897***	2.136***
	(0.102)	(0.153)	(0.148)
Father's Education Level	1.385***	2.197***	2.492***
	(0.101)	(0.152)	(0.146)
Constant	77.09***	68.20***	67.30***
	(0.619)	(0.943)	(0.883)
Observations	18821	18810	18815
Standard errors in parentheses			
* p<0.1, ** p<0.05, *** p<0.01			

5. Conclusions and Discussion

Based on nationally representative CEPS data, this study confirms that preschool education experiences exert a significant long-term positive impact on junior high students' academic performance, with this effect remaining robust after controlling for parental education levels. Key findings include:

First, students who received preschool education significantly outperformed non-recipients in Chinese, mathematics, and English, supporting the research hypothesis. The enhancement effect was particularly pronounced in mathematics and English.

Second, parental education level, as a core indicator of family cultural capital, exerts an independent and significant positive influence on children's academic performance. This finding aligns with the moderating role of family capital identified by Tu Rongzhen et al. [7].

Third, access to preschool education itself is unequal. Descriptive statistics reveal that students with rural household registration, from economically disadvantaged families, and with fathers of lower educational attainment have relatively lower rates of preschool enrollment. This exposes the reality of unequal educational starting points, consistent with findings by Yuan Yuzhi [5] and Wang Dian et al. [8] on the lack of preschool opportunities for various disadvantaged groups.

The policy implications of this study are as follows:

1. Continue to increase public investment in preschool education, focusing on expanding the supply of affordable resources. Governments should prioritize establishing affordable kindergartens in rural areas, underdeveloped regions in central and western China, and low-income urban communities. Through funding and subsidies, they should effectively guarantee access for vulnerable groups such as economically disadvantaged families, families with multiple children, and children with disabilities.

2. While expanding access, place high importance on preschool education quality. Research confirms the long-term benefits of preschool education, but these benefits depend on educational quality. Strengthen teacher training, improve curriculum systems, and prevent "elementary school-like" tendencies to ensure preschool education genuinely promotes the comprehensive development of children's cognitive and non-cognitive abilities[4].

3. Promote home-school-community collaboration to compensate for insufficient family capital. For students from families with relatively weak cultural capital, preschool education serves as a crucial compensatory channel. Effective home-school cooperation mechanisms should be established, enhancing family educational capacity through parent schools and community education services to form a synergistic force for child development[6].

This study has certain limitations. First, cross-sectional data inherently carry selection bias. Although key family variables were controlled, future research could employ quasi-experimental designs such as CEPS longitudinal data or propensity score matching [5,6] to better approximate causal inference.Second, the primary focus on academic achievement as a cognitive outcome could be expanded in future studies to include multidimensional developmental indicators such as non-cognitive abilities and mental health. Finally, the study did not differentiate between types and quality of preschool education, which represents an area for future research refinement.

References

[1] James J Heckman. (2000). Policies to foster human capital. *Research in Economics, 54* (1), 3-56. https://doi.org/10.1006/reec.1999.0225.

[2] Heckman James J, Moon Seong Hyeok, Pinto Rodrigo, Savelyev Peter A & Yavitz Adam. (2010). The Rate of Return to the High/Scope Perry Preschool Program.. *Journal of public economics, 94* (1-2), 114-128.

[3] Heckman James J. (2008). Schools, Skills, and Synapses.. *Economic inquiry, 46* (3), 289. https://doi.org/10.1111/J.1465-7295.2008.00163.X.

[4] Gong, X., & Li, Z. Y. (2018). The Impact of Preschool Education Experience on Non-Cognitive Abilities Among Junior High School Students: An Empirical Study Based on CEPS. Education and Economy, (04), 37-45.

[5] Yuan, Y. Z., & Zhao, Y. (2019). The Impact of Preschool Education on Cognitive Abilities Among Junior High School

Students: An Empirical Analysis Based on CEPS Data. Educational Science Research, (11), 43-50+57.

[6]Ye Shengjun. (2022). The Impact of Preschool Education on Academic Achievement and Development Among Junior High School Students: A Counterfactual Estimation Based on Propensity Score Matching (PSM). Journal of Shaanxi Preschool Teachers College, 38(04), 1-6+39.

[7] Tu, R. Z., Zhang, W. W., & Huang, D. Q. (2017). The value and inequality of preschool education: An empirical study based on CEPS data. Preschool Education Research, (10), 15-27. https://doi.org/10.13861/j.cnki.sece.2017.10.002.

[8] Wang Dian, Cai Hefei, Xu Guoqing & Liu Xinxue. (2024). The Impact of Preschool Education Experience on the Competency Development of Junior High School Students with Visual Impairments: An Empirical Study Based on CEPS Data. Modern Special Education, (18), 57-64.

