International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 4987-4998

https://doi.org/10.70517/ijhsa464342

Linear programming model of the impact of family socioeconomic resources on the distribution of educational opportunities for children

Jia Bian[1,*]
[1] School of Teacher Education, Qilu Normal University, Jinan, Shandong, 250200, China
Corresponding authors: (e-mail: bianjia202412@163.com).

Abstract The scale of higher education in China continues to expand, but the problem of uneven distribution of educational opportunities remains prominent. Differences in family background have a profound impact on children's educational development, and there are obvious gaps in access to educational resources among different social classes. Using data from the China Family Tracking Survey (CFPS) 2020-2024, this paper explores the mechanisms by which family socioeconomic resources affect the distribution of children's educational opportunities through binary logistic regression modeling and structural equation modeling. The study decomposed family socioeconomic resources into three dimensions: cultural capital, economic capital and social capital, and used entropy weighting method for comprehensive evaluation. The results showed that: the overall prediction rate of the model was 86.78% correct; parental education and father's occupation had a significant effect on children's educational opportunities ($p<0.001$); educational expectations played a mediating role in the distribution of children's educational opportunities influenced by family socio-economic resources, with the mediating effect accounting for 66.35%; and the input of school resources negatively moderated the influence of family socio-economic resources on educational opportunities ($p=0.012$). The study found that family socioeconomic resources affect children's educational opportunities mainly through the indirect paths of educational expectations and school quality, which provides empirical evidence for understanding the generation mechanism of educational inequality.

Index Terms Family socio-economic resources, Distribution of children's educational opportunities, Binary logistic regression model, Educational expectations, Mediating effect, School resource inputs

I. Introduction

Education is the main way of human capital accumulation and plays an important role in personal income and comprehensive development, while the efficiency and equity of educational resource allocation is an important factor affecting a country's long-term economic growth and social stability [1], [2]. Considering that basic education has strong positive externalities, governments have intervened extensively in basic education. Although the government sector is the decisive force for human capital investment in compulsory education, the family sector still plays a very important role in human capital investment [3].

As a socially balanced system, the family possesses the functions of social information communication and social resource conversion [4]. In terms of the family system as a provider of educational resources, the input of various forms of resources in the family environment or help parents to teach their children and help them to learn [5], [6]. For example, there are individualized differences in both the quantity and quality of private sector educational inputs in the form of extracurricular tutoring compared to more standardized schooling, especially highly standardized public education [7]-[9]. The quantity and quality of private sector educational inputs depend to a large extent on the economic conditions, parental preferences, and market environment of the student's family [10], [11]. While differentiation in educational inputs across families can better meet the heterogeneous needs of different students, it may threaten educational equity by exacerbating inequality of opportunity in education, and in the long run may exacerbate social income disparities and jeopardize sustainable socio-economic development [12]-[15]. Therefore, it is important to promote a rational and balanced distribution of family socioeconomic resources, which has a direct and profound impact on all members of the family and society as a whole.

This study constructs a theoretical framework of family socio-economic resources affecting the distribution of children's educational opportunities, and subdividing family resources into cultural, economic and social dimensions for quantitative analysis. A binary logistic regression model was developed to test the direct effect of each dimension's resources on educational opportunities. Educational aspirations and school quality are further

introduced as mediating variables, and structural equation modeling is applied to explore the influence pathways. The moderating role of school resource inputs is also examined to analyze how it changes the strength of the influence of family background. Finally, the heterogeneity of the era and region is analyzed to reveal the characteristics of differences among different groups.

II. Rationale and research design

II. A.Logistic regression model

In the field of machine learning, logistic regression models are a widely used classifier. In essence, it can be regarded as a generalized linear regression model with only two dependent variables, and at the same time, an activation function is added on the basis of the linear regression model to achieve the purpose of compressing the output range of the linear model to (0,1), so that the output value can be regarded as the probability value attributed to a certain label, when the probability value exceeds 0.5, the event $\{Y=1\}$ is considered to be true, and vice versa, the event $\{Y=0\}$ is considered to be true, so as to classify and distinguish [16].

Thus, a logistic regression model for a dichotomous problem can be expressed as:

$$P(Y=1\mid X;\beta)=\sigma(X^T\beta)=\frac{1}{1+e^{-X^T\beta}}=\frac{e^{X^T\beta}}{1+e^{X^T\beta}} \tag{1}$$

where $\sigma(z)=\frac{1}{1+e^{-z}}$ denotes the activation function, i.e., the sigmoid function, and $P(Y\mid X)$ denotes the conditional probability $X=(1,x_1,x_2,\cdots,x_n)^T,\beta=(\beta_0,\beta_1,\beta_2,\cdots,\beta_n)^T$.

The Sigmoid function has some nice properties:

$$\begin{cases}\sigma(z)=1-\sigma(z)\\ \sigma'(z)=\sigma(z)(1-\sigma(z))\end{cases} \tag{2}$$

Using the above properties can greatly simplify the subsequent derivation of the formula.

In general, it is assumed that the samples satisfy the independent homogeneous distribution and obey the normal distribution. The conditional distribution function of the logistic regression model can be derived from the distribution function of the two-point distribution as:

$$\begin{aligned}P(Y\mid X;\beta)&=\left(\frac{e^{X^T\beta}}{1+e^{X^T\beta}}\right)^Y\left(1-\frac{e^{X^T\beta}}{1+e^{X^T\beta}}\right)^{1-Y}\\&=\left(\frac{e^{X^T\beta}}{1+e^{X^T\beta}}\right)^Y\left(\frac{1}{1+e^{X^T\beta}}\right)^{1-Y}\end{aligned} \tag{3}$$

Let $p=P(Y=1)$, then $1-p=1-P(Y=1)=P(Y=0)$, and the probability of the event $\{Y=1\}$ occurring or not occurring is then:

$$\frac{P}{1-p}=\frac{\frac{e^{X^T\beta}}{1+e^{X^T\beta}}}{\frac{1}{1+e^{X^T\beta}}}=e^{X^T\beta} \tag{4}$$

Taking logarithms on both sides, there:

$$\ln\left(\frac{p}{1-p}\right)=\ln e^{X^T\beta}=X^T\beta \tag{5}$$

Notice that the right-hand side of the equation is a form of linear regression, so logistic regression is essentially a generalized linear regression model. Thus, it is possible to transfer many of the methods and ideas from linear modeling to logistic regression, leading to some new discoveries.

In the data set dealing with the actual binary classification problem, the sample can be represented as a binary random variable (X,Y). where $X\in\mathbb{R}^d,Y\in\{0,1\}$, $P(Y=0)$ and $P(Y=1)$ represent a priori information, which can usually be obtained based on the distribution of labels in the data set. Remember that the sample of the dataset is

a random variable (X,Y), the number of samples is N, the number of belonging to the label $\{Y=1\}$ is N_1, and the number of belonging to the label $\{Y=0\}$ is N_0, and the prior distribution of the sample is:

$$\begin{cases} P(Y=1)=\dfrac{N_1}{N} \\ P(Y=0)=\dfrac{N_0}{N} \end{cases} \tag{6}$$

II. B.Study design

II. B. 1) Data sources and processing

(1) Data Source

This study chooses to use data from the China Family Tracking Survey (CFPS), which is collected and obtained by the China Social Science Survey Center (ISSS) of Peking University. This database covers information on residents' families, economy, and education in Province A, and is characterized by wide coverage and strong representativeness. In order to obtain more accurate information about students' families before the college entrance examination, the study chose the tracking data from 2020-2024, and matched the data from the adult questionnaire and the family questionnaire.

(2) Variable Selection

a) Explained variables

The explanatory variable of the study is the distribution of children's educational opportunities, which is categorized into the quantity of access to opportunities and the quality of access to opportunities in this paper.

b) Explanatory variables

The explanatory variables selected in this paper are family socioeconomic resources, educational expectations and school quality. Family socioeconomic resources are divided into three dimensions: cultural capital, economic capital and social capital. Among them, family cultural capital is operationally defined as the highest educational attainment of both parents. Family economic capital is obtained by matching the sample ID of high school students in a given sampling year with the sample ID of families in the sampling year, and is operationally defined as net family income per capita. Family social capital is operationally defined as the ISEI Occupational Socioeconomic Status Index (ISEI), which is processed by converting the parents' occupation type to ISCO88, then converting the ISCO88 occupation code to ISEI, and taking the value of the highest value of both parents. Since the family socio-economic resources cover three indicator layers, this paper intends to use entropy weighting method to comprehensively evaluate the indicators, and the family socio-economic resources indicator design and weights are shown in Table 1. Educational expectations are generated by the question “Your desired level of education” in the adult questionnaire. School quality is generated by the question “whether it is a model/key school” in the adult questionnaire.

Table 1: Family capital index design and weight

Variable name	Index	Index definition and value	Weighting
Family social and economic resources	Family culture capital	Parents' highest education period (range 0-19)	0.183
	Family social capital	The parents' highest isei occupational socio-economic status index (range 19-90)	0.354
	Family economic capital	The per capita net income of the court is processed by 1%	0.463

(3) Control variables

The control variables selected in this paper include gender, urban-rural household, individual academic scores and the region where the family is located. The meanings and assigned values of the variables are shown in Table 2.

Table 2: Variable meaning and assignment

Variable Type	Variable Name	Variable Meaning And Assignment	Mean	Standard Deviation
Explained Variable	Whether To Go To College	0= No,1=Yes	0.902	0.293
	Whether To Go To Undergraduate	0= No,1=Yes	0.532	0.496
Interpretation Variable	Family Capital	Entropy Value Method Weighted Synthetic Value	0.241	0.154
	Whether Demonstration/Key High School	0= No,1=Yes	0.462	0.495
	Education Expectation	The Expectation Of Education Is Converted Into The Length Of Education (9= High School, 12= High School, 15= Junior College, 16= University Undergraduate, 19= Master, 23= Doctor)	16.182	1.888
Control Variable	Hukou	0= Rural, 1= Cities	0.282	0.453
	Gender	0= Female, 1= Male	0.491	0.505
	Academic Rating	1 = Very Dissatisfied, 2 = Dissatisfied, 3= General Satisfaction, 4= Satisfied, 5= Very Satisfied	3.124	0.743
	The Area Of The Family	1= The Eastern Region, 2= The Central Region, 3= The Western Region, 4= Northeast Area	2.255	0.983

II. B. 2) Analytical methods

(1) Model selection

In this study, a binary logistic model was used to estimate the predictive effect of family socioeconomic resources on educational opportunities, in order to explore how family background affects children's educational opportunities. In this study, the dependent variable "educational opportunity" has only two values, which represent two possibilities of "educational opportunity". Educational opportunities are affected by many factors such as family economic, cultural, political, and social factors, but the final result of educational opportunities is "access to education" or "lack of access to education". Here, we define "y=1" as "access to education" and "y=0" as "no access to education". Namely:

$$y=\begin{cases} 1=\text{Access to preschool education} \\ 0=\text{No access to preschool education} \end{cases} \tag{7}$$

In a binary logistic regression model, it is common to use p to represent the probability of an event occurring and $1-p$ to represent the probability of an event not occurring, and to think of p as a linear function of the independent variable X_i. To wit:

$$p=(y=1)=F(\beta_i X_i) \quad i=1,2,...k \tag{8}$$

Different forms of $F(\cdot)$ lead to different forms of models, the most common of which are:

$$p=\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_K X_K+\varepsilon \tag{9}$$

Since the function p is insensitive and slow to changes in X_i in the neighborhood of $p=1$ or $p=0$, the Logistic transform of p, also known as the Logit transform of p, is introduced, i.e.:

$$\theta(p)=logit(p)=\ln\left(\frac{p}{1-p}\right) \tag{10}$$

Using $\theta(p)$ instead of p in the ① expression gives:

$$\theta(p)=\ln\left(\frac{p}{1-p}\right)=\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_K X_K+\varepsilon \tag{11}$$

Based on the independent and dependent variables of this study, the binary logistic regression model expression was constructed as follows:

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 economic + \beta_2 cultural + \beta_3 social + \varepsilon_i \tag{12}$$

Among them,

$$y_i = \begin{cases} 1 = \text{Access to preschool education} \\ 0 = \text{No access to preschool education} \end{cases} \tag{13}$$

$$p_i = p(y_i = 1) \tag{14}$$

In this model, y_i represents the two outcomes of the dichotomous variable "educational opportunity", denoted by 0 and 1. Economic, Cultural, and Sociall represent the four dimensions of household socio-economic resources, namely economic capital, cultural capital, and social capital, respectively, and ε_i represents error.

Second, in terms of the test of indirect effects, this paper adopts the KHB method to decompose the direct effect of family socioeconomic resources on the allocation of children's educational opportunities and the indirect effect through educational expectations and school quality. A schematic of the mediating effects is shown in Figure 1.

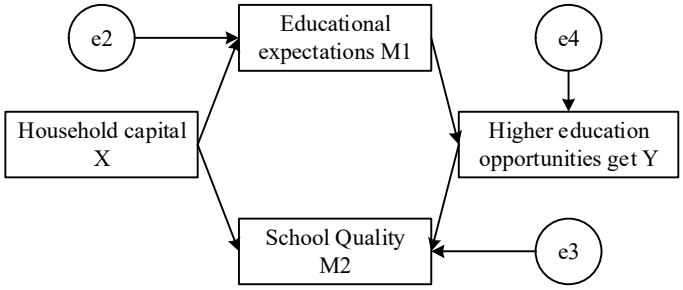

Figure 1: Schematic diagram of the mediating effect

(2) Interview design

The study will use semi-structured interviews to analyze the transmission path of family socio-economic resources, so as to shift from the data appearance of “what” to the specific situation of “why”, which is conducive to the researcher to focus on the research problem from the perspective of the subject of the educated. In this study, it is found that there are three ways in which family socio-economic resources directly affect the distribution of children's educational opportunities, namely, material support ①, motivational support ② and rule mastery ③. In these actions, there is basically a consistency between the “quantity of capital” and the “quality of behavior”.

III. Analysis of empirical results

III. A. Descriptive analysis of sample data

In this paper, the selected valid data is 1899, using SPSS to categorize the data by urban and rural areas and gender, the highest education level of individuals in 2024 is shown in Table 3.The highest education level of individuals* gender of individuals in 2024 is shown in Table 4. As can be seen from Table 3, among the selected samples from rural areas, 96.7% of the samples have high school and below education, of which the proportion of illiterate/semi-literate reaches 19.4%, which is still a very high proportion. In contrast, tertiary and higher education is far from reaching mass levels. It can be seen from this that the rural population is generally less educated. In contrast, the education level in towns and cities is generally higher, with tertiary education and above reaching 28.16%, which has already reached the massification level of higher education. As can be seen from Table 4, among the selected female samples, the proportion of those whose education is illiterate/semi-illiterate is 15.95%, the education of middle school and below is 69.81%, the education of high school and below reaches to account for 85.39%, while the education of college and above is only 14.6%. From the above data, it can be seen that the overall educational level of women is not high, and the proportion of illiterates/semi-literates is large. In contrast, the proportion of illiterate/semi-literate men in the sample is 9.74%, which is 6.21% less than that of women. This shows that, as a whole, the educational level of the men in the sample is generally higher than that of the women.

Table 3: The highest degree of personal education in 2024

	Country (person)	The cumulative percentage of different country degrees in the country (%)	City (person)	The cumulative percentage of different cities in cities (%)	Total (person)
Illiteracy/semi-illiteracy	200	19.40	33	3.97	233
Primary school	262	25.41	71	8.54	333
Junior high school	431	41.80	292	35.14	723
High school	104	10.09	201	24.19	305
Junior college	30	2.91	109	13.12	139
Undergraduate	2	0.19	114	13.72	116
Master	2	0.19	11	1.32	13
Total	1031	100	831	100	1862

Table 4: 2024 personal highest degree * personal nature

	Personal Nature				
	Female (person)	The percentage of women with different degrees of education (%)	Man (person)	Male sample different degree accumulated percentage (%)	Total (person)
Illiteracy/semi-illiteracy	130	15.95	102	9.74	232
Primary school	142	17.42	179	17.10	321
Junior high school	297	36.44	418	39.92	715
High school	127	15.58	187	17.86	314
Junior college	57	6.99	70	6.69	127
Undergraduate	51	6.26	77	7.35	128
Master	11	1.35	14	1.34	25
Total	815	100	1047	100	1862

III. B. Regression analysis

(1) Processing of the sample and data coding

Based on a model of individual access to education:

$$L = Ln[p/(1-p)] = \beta_0 + \beta_1 father'sedu + \beta_2 mother's\ edu + \beta_3 father's\ occupation + \beta_4 family's\ income + \mu \quad (15)$$

After organizing the data, the binary Logistic command of SPSS software was used to regress the data [17], and the following results can be obtained: the results of the sample treatment, the results of the sample treatment are shown in Table 5.

Table 5: Sample processing results

Unweighted Cases[a]		N	Percent (%)
Selected Cases	Included in Analysis	1831	100
	Missing Cases	0	0
	Total	1831	100
Unselected Cases		0	0
Total		1831	100

a) Dependent variable coded as 0 for not entering higher education and 1 for entering higher education.

b) Frequency statistics and coding of categorical variables

(a) Variables related to father's and mother's educational attainment Setting dummy variables for father's and mother's educational attainment: illiterate/semi-literate, elementary school, junior high school, high school, and college. Reference variable for father's and mother's education: undergraduate.

(b) Variables related to father's occupation (main job) Dummy variable for setting father's occupation: head of enterprise or institution, intellectual, government administrative employee, business, service worker, production worker, operative worker. Reference variables of father's occupation: others.

(2) Misclassification matrix of the model

The final prediction classification is shown in Table 6. From Table 6, it can be seen that out of 1554 people with actual personal qualifications who have not undergone higher education in this study, 1524 of them were correctly identified by the model and 30 were incorrectly identified, which is a 98.07% correct rate. Of the 277 people with actual personal qualifications who have undergone higher education, the model correctly identified 65 of them and incorrectly identified 212 of them, with a correct rate of 23.47%. In this study, the total correct prediction rate of the model was 86.78%, which is greater than 50%, and therefore the results are in the acceptable range.

Table 6: Final prediction classification

	Observed		Predicted		
			Bachelor's degree		Percentage Correct
			0	1	
Step 1	Bachelor's degree	0	1524	30	98.07
		1	212	65	23.47
	Overall Percentage				86.78

(3) Significance test of regression model

The significance test of the regression model is shown in Table 7. The meanings of the three items in the significance test results of the regression model are: step statistic is the likelihood ratio test result of each step with the previous step, block is the likelihood ratio test result of comparing block n with block n-1, model is the likelihood ratio test result of the previous model with the model after the variables in the current model have changed, and the three statistics and hypothesis test results are completely consistent if the selection is METHOD = Backward When wald is chosen, the three statistics and hypothesis testing results are identical. From the statistics in the table above, the regression model passes the significance test.

Table 7: Significance test of regression model

		Chi - square	df	Sig.
Step 1	Step	296.642	19	0.000
	Block	296.642	19	0.000
	Model	296.642	19	0.000
Step 2	Step	-1.051	1	0.000
	Block	293.162	15	0.000
	Model	293.162	15	0.000

(4) Goodness-of-fit test

The results of the goodness of fit test are shown in Table 8. The table shows the results of Hosmer & Lemeshow goodness of fit test for the regression equation. If the value of this significance level (Sig) is less than the given significance level, the null hypothesis is rejected - there is no significant difference between the observed value of the dependent variable and the predicted value of the model, indicating that the predicted value of the model is significantly different from the observed value. If the value is greater than the given level of significance, the null hypothesis cannot be rejected, indicating that the model's estimates fit the data at an acceptable level. Based on the data in the table, it can be seen that in STEP2, the H_L goodness-of-fit test statistic for the model corresponds to an X2 value of 2.641, with a degree of freedom of 9, and a sig value of 0.895, i.e., the difference is not significant at sig > 0.05, so the hypothesis that the actual values are in agreement with the predicted values cannot be rejected, indicating that the model has a better goodness-of-fit. That is, the difference between the predicted values of the model and the actual observed values is not significant, so the hypothesis is accepted, i.e., “There is a significant correlation between children's educational opportunities and family socio-economic resources.”

Table 8: The results of the proposed excellence test

Step	Chi - square	df	Sig.
1	15.654	9	0.033
2	2.641	9	0.895

(5) Significance test of model parameters

The variables that passed the test are shown in Table 9. The variables that failed the test are shown in Table 10. It can be seen from Table 9 and Table 10: The variables that passed the Wald test shown in Table 9 are father's highest education, mother's highest education, and father's main job. The variables that do not pass the Wald test in Table 10 are the annual income of the family. From the sig value of the significance level of the regression coefficients, it can be seen that father's and mother's highest educational qualification (sig<0.001), and father's job (sig<0.01) have a very significant effect on the access to children's educational opportunities. Among them, the highest educational qualification of father and mother as illiterate/semi-literate, elementary school, and father's job as a production worker have a significant effect on the access to children's educational opportunities. The regression coefficient transforms from negative to positive as the father's and mother's educational qualifications rise above the level of specialization, at which point the parents' educational qualifications have a positive significance on the children's access to educational opportunities. It is noteworthy that on the impact of parents' education on children's access to higher education, both positively and negatively, in terms of the absolute value of B-value, the mother's impact on children's access to education is much higher than that of the father. The regression coefficient B gradually decreases from positive to negative as the father's occupation changes from government administrator, intellectual, and head of enterprise or institution to commercial or service worker, operative worker, and production worker. Among them, when the father's job is government administrators, intellectuals, heads of enterprises and institutions, the regression coefficient B is positive, which is favorable to the children's access to higher education. When the father's job is commercial or service personnel, operative personnel, production personnel, the regression coefficient B is negative, which is unfavorable to the children's opportunity to obtain higher education. The smallest regression coefficient is found when the father's job is a production worker, which is the most unfavorable for his children's access to education.

Table 9: Tested variable

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 2	Father's Highest Degree	***			5	0	
	Illiteracy/Semi-Illiteracy	-1.359**	0.511	6.779	1	0.007	0.282
	Primary School	-1.076*	0.504	4.695	1	0.029	0.379
	Junior High School	-0.644	0.487	1.757	1	0.187	0.475
	High School	-0.149	0.47	0.099	1	0.746	0.814
	Junior College	0.305	0.565	280.002	1	0.597	1.331
	Mother's Highest Degree	***		43.41	7	0	
	Illiteracy/Semi-Illiteracy	-2.353**	0.775	9.3	1	0.002	0.098
	Primary School	-1.65*	0.759	4.623	1	0.029	0.212
	Junior High School	-1.217	0.754	2.549	1	0.112	0.305
	High School	-0.793	0.751	1.119	1	0.288	0.425
	Junior College	0.396	0.979	0.163	1	0.694	1.487
	Father's Work	**		23.208	7	0.001	
	The Head Of The Enterprise And Cause Unit	0.001	0.475	0	1	0.993	1.015
	Intellectuals	0.226	0.384	0.342	1	0.548	1.241
	Government Administrator	0.492	0.408	1.505	1	0.225	1.644
	Business, Service Personnel	-0.002	0.333	0.001	1	0.973	0.979
	Production Personnel	-0.737***	0.186	17.13	1	0	0.485
	Operant	-0.343	0.283	1.52	1	0.216	0.73
	Constant	1.086	0.707	2.383	1	0.129	2.893

Table 10: Unchecked variable

			Score	df	Sig.
Step 2[a]	Variables	Family income	1.485	1	0.226
	Overall Statistics		1.485	1	0.226

III. C. Analysis of intermediation effects

In order to deeply explore the inner mechanism of the influence of family socioeconomic resources on the distribution of children's educational opportunities, this study further substitutes the mediating variables into the structural equation modeling for analysis. By using Model 4 in the SPSS macro program PROCESS to carry out the test of mediating effects, the mediating roles of educational expectations and school quality between family socio-economic resources and children's educational opportunity allocation were verified and analyzed according to the Bootstrap method. The mediation model of family socioeconomic resources affecting the distribution of children's educational opportunities based on the linear programming model is shown in Figure 2.

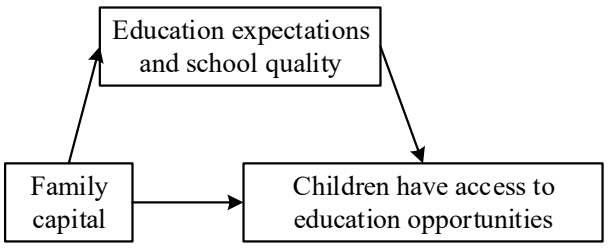

Figure 2: The intermediary model of the opportunity obtained by resources

The results of the mediation model test of family socio-economic resources are shown in Table 11, showing that the positive effect of family socio-economic resources on the distribution of children's educational opportunities is insignificant (β =0.097, t=1.801, p=>0.05), and that the positive effect of family socio-economic resources on educational expectations and school quality is significant (β =0.342, t=11.909, p<0.001). t=11.909, p<0.001). The positive effect of educational expectations and school quality on the distribution of educational opportunities for children is significant (β = 0.163, t = 3.359, p < 0.001).

Table 11: The results of the domestic socio-economic resource mediation model

Result variable	Predictor variable	Fitting index			Coefficient and significance	
		R	R^2	F value	β	t
Cumulative dominance	Family social and economic resources	0.342	0.111	51.666	0.361	11.909***
Education opportunity	Family social and economic resources	0.097	0.006	3.054	0.088	1.801
Education opportunity	Cumulative dominance	0.163	0.038	11.417	0.169	3.359***

The results of the mediating effect decomposition are shown in Table 12. It can be seen that the upper and lower limits of the Bootstrap 95% confidence interval for the mediating effect of cumulative advantage do not contain 0, indicating that family socioeconomic resources can influence the distribution of children's educational opportunities through the mediating effect of cumulative advantage. The value of the direct effect of family socioeconomic resources on the distribution of children's educational opportunities is 0.012, the value of the mediating effect of educational expectations and school quality is 0.022, and the value of the total effect is 0.041.

Table 12: Mediation effect decomposition

	Effect value	se	The lower limit of the boot cl	The upper limit of the boot cl	Relative effect
Total effect	0.041	0.0223	-0.0065	0.0861	
Direct effect	0.012	0.0251	-0.0352	0.0623	33.65%
Mediation effect	0.022	0.0092	0.0071	0.0453	66.35%

In summary, both family socioeconomic resources and educational expectations positively affect the distribution of children's educational opportunities, but the direct effect of family socioeconomic resources on educational opportunities is not significant, and the mediating effect through educational expectations is significant. According

to the analysis of the mediation model, it can be seen that educational expectation is the mediator of family socioeconomic resources affecting the distribution of children's educational opportunities, and the mediation effect accounts for 66.35%.

III. D. Analysis of moderating effects

This paper examines the impact of family socio-economic resources on the distribution of children's educational opportunities, with school resource inputs set as a moderating variable. On the basis of factor analysis, this study utilizes SPSS as well as the PROCESS plug-in to conduct moderating effect tests. Stratified regression analysis was used to explore the moderating role of school resource inputs in family socioeconomic resources in predicting the distribution of children's educational opportunities. The regression analysis of the moderating effect of school resource inputs between family socioeconomic resources and the distribution of children's educational opportunities is shown in Table 13. As can be seen from the table, family socioeconomic resources and school resource inputs all significantly positively predicted the allocation of children's educational opportunities (p family socioeconomic resources = 0.046 < 0.05, p school resource inputs < 0.001), and the interaction term between family socioeconomic resources and school resource inputs significantly negatively predicted the allocation of children's educational opportunities (p family socioeconomic resources × school resource inputs = 0.012 < 0.05).

Table 13: Regression analysis of school resource input

Dependent Variable	Predictor Variable	Nonnormalized Coefficient	Normalized Regression Coefficient β	R^2	ΔR^2	T	P
Educational Opportunity	Gender	-0.208	-0.048	-0.001	0.009	-0.94	0.341
	Peoples	-0.115	-0.011	-0.002	1.005	-0.284	0.779
	Family Social and Economic Resources	0.223	0.094	0.013	0.008	2.038	0.046
	School Resource Input	0.569	0.265	0.084	0.063	4.999	<0.001
	Family Social And Economic Resources X School Resources Investment	-236.994	-123.988	0.09	0.021	-2.429	0.012

To further test the substance of the moderating effect, this study conducted a simple slope analysis. The moderating effect of school resources is shown in Figure 3. Based on M±SD as a criterion, two levels of school resource inputs, high and low, were distinguished, and the values of the moderating effect of school resource inputs on the distribution of family socioeconomic resources and children's educational opportunities at different levels were analyzed. The results show that family socio-economic resources significantly and positively predict the distribution of children's educational opportunities under both high and low levels of school resource inputs, in addition, the slope of the high school resource input group is smaller than that of the low school resource input group, and the effect of family socio-economic resources on the distribution of children's educational opportunities diminishes as the level of school resource inputs increases. It can be seen that school resources, such as teachers' strength, learning atmosphere, hardware conditions and learning resources, are favorable support conditions for children's competition, and key schools have a great influence on the distribution of children's educational opportunities.

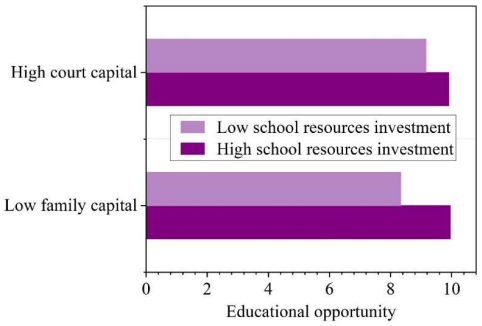

Figure 3: Regulation of school resources

III. E. Heterogeneity analysis

III. E. 1) Analysis of temporal heterogeneity

Since the expansion of higher education enrollment in 1999, the scale of higher education in China has been expanding, and the access opportunities for higher education have been increasing, so the impact of family background on higher education access opportunities will also change with the times, so this study divides the samples into two groups of "post-80s" and "post-90s" by dividing the samples into two groups of "post-80s" and "post-90s" by performing logistic regression to study the differences between generations of the impact of family background on higher education access opportunities. Descriptive statistics of epochal heterogeneity are shown in Table 14. It can be seen from the table that the average value of enrollment opportunities, enrollment quality and family socio-economic resources of the post-90s generation is higher than that of the post-80s generation, which reflects the positive impact of China's economic development and the expansion of higher education on China's higher education to a certain extent.

Table 14: Age heterogeneity descriptive statistics

Age	Statistic	Admission opportunity	Admission quality	ISEI	EDU
After 80(N=2105)	Mean	0.1550124	0.3134183	59.863514	7.965743
After 90(N=1633)	Mean	0.2485571	0.5255787	62.53135	8.728553

III. E. 2) Analysis of regional heterogeneity

Since China's higher education resources and economic development are still disparate in the three regions of East, Central and West, the impact of family socioeconomic resources and urban and rural hukou on higher education will also be different in the three regions of East, Central and West. In this paper, the provinces are derived from the permanent residential addresses of the respondents during their minor years and are divided into three regional groups, East, Central and West, and the results of the grouping are shown in the following table: the statistics of the regional heterogeneity attributes are shown in Table 15. As can be seen from the table, the mean values of access, access quality and family socio-economic resources in each region gradually decrease from east to west, which indicates that both access to higher education and overall family background, which is in line with the differences in higher education resources as well as economic differences in the East, Central and West. An ordered logistic regression was conducted on each of the three groups to examine the differences in the impact of family background on the quality of higher education access in the East, Central, and West regions.

Table 15: Regional heterogeneity property statistics

Region	Statistic	Admission opportunity	Admission quality	ISEI	EDU
East	Mean	0.2273235	0.4645626	65.36971	8.826624
Middle	Mean	0.2176244	0.4418325	62.14132	8.575553
West	Mean	0.1453655	0.293428	55.45728	7.373224

IV. Conclusion

Family socio-economic resources have a significant effect on the distribution of children's educational opportunities, but this effect is mainly realized through indirect paths. The regression coefficient of -0.737 for the father's occupation as a production worker has the most unfavorable effect on children's access to higher education. When the mother's education is illiterate/semi-literate, the probability of children's access to higher education is only 9.8% of the probability of families with a mother with a bachelor's degree under the same conditions.

Educational expectations and school quality play a key role in the transmission process, and the cumulative advantage of both amplifies the family resource effect. The moderating role of school resource inputs shows that quality educational resources can mitigate the effects of family background to a certain extent, with the coefficient of the impact of family socioeconomic resources under the condition of high school resource inputs being -123.988. The complexity of the distribution of educational opportunities is further corroborated by the differences between regions and eras, with the tertiary education enrollment rate in the eastern region being 22.73%, significantly higher than that in the western region, which is 14.54%. These findings reveal the deep-rooted mechanisms of educational inequality and provide a basis for the formulation of targeted educational policies.

References

[1] Della Sala, M. R., Knoeppel, R. C., & Marion, R. (2017). Modeling the effects of educational resources on student achievement: Implications for resource allocation policies. Education and Urban Society, 49(2), 180-202.

[2] Witte, K. D., & López-Torres, L. (2017). Efficiency in education: A review of literature and a way forward. Journal of the operational research society, 68(4), 339-363.
[3] Li, Z., & Qiu, Z. (2018). How does family background affect children's educational achievement? Evidence from Contemporary China. The Journal of Chinese Sociology, 5(1), 1-21.
[4] Cui, Y., Liu, H., & Zhao, L. (2019). Mother's education and child development: Evidence from the compulsory school reform in China. Journal of Comparative Economics, 47(3), 669-692.
[5] Roksa, J., & Kinsley, P. (2019). The role of family support in facilitating academic success of low-income students. Research in Higher Education, 60, 415-436.
[6] Weiss, H. B. (2017). Family support and education programs: Working through ecological theories of human development. Evaluating family programs, 3-36.
[7] Đurišić, M., & Bunijevac, M. (2017). Parental involvement as a important factor for successful education. Center for Educational Policy Studies Journal, 7(3), 137-153.
[8] Thomson, S. (2018). Achievement at school and socioeconomic background—an educational perspective. npj Science of Learning, 3(1), 5.
[9] Qianqian, S. H. E. N., & Peihua, T. S. A. I. (2023). Effects of Family Socioeconomic Status and Educational Resources on Children's Cognitive Ability Performance Based on Preschool Education Perspectives. Revista de Cercetare si Interventie Sociala, 82.
[10] Broer, M., Bai, Y., Fonseca, F., Broer, M., Bai, Y., & Fonseca, F. (2019). A review of the literature on socioeconomic status and educational achievement. Socioeconomic inequality and educational outcomes: Evidence from twenty years of TIMSS, 7-17.
[11] Andersson, M. A. (2018). Higher education, bigger networks? Differences by family socioeconomic background and network measures. Socius, 4, 2378023118797217.
[12] Hu, S., Hood, M., Creed, P. A., & Shen, X. (2022). The relationship between family socioeconomic status and career outcomes: A life history perspective. Journal of Career Development, 49(3), 600-615.
[13] Li, W., & Xie, Y. (2020). The influence of family background on educational expectations: A comparative study. Chinese Sociological Review, 52(3), 269-294.
[14] Song, Q., & Tan, C. Y. (2022). The association between family socioeconomic status and urban–rural and high-school attainment gaps: A logistic regression analysis of the China Family Panel Studies data. British Educational Research Journal, 48(6), 1102-1124.
[15] Yang, K. (2023). The Influence of Family Socioeconomic Status, Cultural Capital and Extra Curricular Study on Students' Academic Achievement. Journal of Advanced Research in Education, 2(4), 34-40.
[16] Lu Qianqian,Chai Yongxiang,Ren Lihui,Ren Pengyu,Zhou Junhui & Lin Chunlei. (2023). Research on quality evaluation of innovation and entrepreneurship education for college students based on random forest algorithm and logistic regression model..PeerJ. Computer science,9,e1329-e1329.
[17] Aibanskhem Nongspung & Veronica Pala. (2025). Heterogeneity in Returns to Education in India: A Quantile Regression Analysis. The Indian Journal of Labour Economics,(prepublish),1-27.