

**Student Effort and Parent
Attitude on Education
Attainment: Evidence from
Multi-year Survey in Gansu,
China**

Ridge Zhong-yuan Ren

*Professor Pengpeng Xiao, Faculty
Advisor*

*Professor Kent P. Kimbrough,
Faculty Advisor*

*Honors Thesis submitted in partial fulfillment of the requirements for Graduation with
Distinction in Economics in Trinity College of Duke University.*

Duke
University
Durham, North Carolina
2023

Acknowledgments

I would like to express my gratitude to Dr. Pengpeng Xiao for her insightful comments on my work and for helping me through the process of refining my analysis. I would like to thank Dr. Kimbrough for providing valuable feedback on my writing and helping me make my ideas clear. I would like to thank everyone in the honors economics seminar for providing feedback on my many rounds of presentations. I want to thank Emily Hannum of the University of Pennsylvania for providing the data used in this research. This experience of researching and writing a thesis has helped me grow as a student and I want to thank everyone who made that possible.

Abstract

This paper explores whether student effort and parent attitudes have varying effects at different stages of a student's life in terms of educational attainment and job outcomes. With survey data in Gansu, China, a largely rural province in Northwest China that lags behind the rest of China in education, this paper employs a multivariate regression model. This method allows me to measure the achievement or outcome of the child between each successive wave of surveys and estimate which factors held the strongest effect on the next wave. Student achievement in early waves is measured by the student's score on assessments in math and Chinese, and the later outcome is measured by the student's income and the highest level of education achieved. This paper finds that effort in Math and math achievement have a positive association with better education attainment and career outcomes later in life. In addition, I find that parental education levels also have a positive association with child outcomes.

JEL Codes: I25, I26

Keywords: Education Attainment; Student Effort; China

I. Introduction

In China, market and education reforms since the 1970s have brought dramatic economic development to the country. The quality of education in China has increased, but this improvement is largely concentrated in Eastern China. The poorer, more rural provinces in Western China have consistently lagged behind Eastern cities such as Shanghai and Beijing. In 2014, the average primary school funding in Beijing was 23,000 yuan per student, the equivalent of about 3,300 dollars per student (OECD, 2016) while primary schools in Gansu province received an average of 7000 yuan per student, or about 1000 dollars per student. The gap in funding is even bigger for middle school and high school, with the average Gansu middle school or high school student receiving the same funding as primary school students while the average Beijing middle school or high school student had an average funding of 36000 yuan and 40000 yuan per year respectively. The inequality in development has also been along a rural-urban divide, with disparities in access to health care, education opportunities, employment opportunities, and child mortality rate (Guang et al., 2010). This research seeks to understand factors that drive student outcomes in the Northwestern province of Gansu, China. This research focuses on the effects parental attention and parental style and student effort have on the future income and education attainment of the students. Utilizing data collected from 2000 to 2015 in multiple waves of surveys, this paper will trace first how different factors affect student education achievement and later on in their education, their highest level of education attained, and employment outcomes. This paper will first describe the previous research done in the field of effort and parental attitude on education. It will also describe other relevant papers that use the same data set. The paper proceeds with the description of the data and empirical regressions.

II. Literature Review

The primary research question is what are the effects of parental and student effort on education and career outcomes. This paper contributes to the existing literature in the fields of family economics and behavioral economics. Parenting style and parental attention have been studied and modeled in economics, while there are fewer studies specifically about the self-perceived effort of students and their education outcome. There is evidence from Avvisati et al. (2014) that more parental involvement has a positive effect on student behavior in schools. I seek to find relationships between the amount of attention parents invest into their children's education and the student outcome of those children. According to Doepke et al. (2019), current research in parenting economics generally categorizes parenting styles into three: permissive, authoritative, and authoritarian parents. These categories are defined by whether the parent(s) indoctrinate the child, the degree to which they restrict the child's choice, and the cost to the parent(s). I use the longitudinal and multilevel survey data to pinpoint which questions in the surveys to the children and their families can illustrate the type of parenting style the parents use. My data source, the Gansu Survey of Children and Families, provides survey responses from students and their parents about education, their life, etc. The students are asked about their attitude towards education, their effort in school, and how they perceive themselves in the classroom. The parents are asked about their own views on education and whether their child is receiving a good education. This data might also provide information on how the fathers and mothers might differ in their approaches, and the ambiguities between parenting styles that the responses to questions might create. One can imagine that the parents might be responding to the student's initial performance in school and modifying their parenting style. If a student is frequently misbehaving in school or not doing well in classes, the parent might take a greater

role in their life and instill a higher level of discipline. Early childhood investment in education is important and the parental attitude toward education early in their child's life might influence their later education outcome and income.

Another important factor that drives student performance is the student's effort in school. While the cognitive ability of the students and the education resources that a student has access to are important factors for their education attainment, another driver of student education achievement is their effort level. Previously, research has been done showing that the different score levels that students in different countries score on the PISA test can be partially attributed to effort (Zamarro et al., 2019). Zamarro et al. infer a high and low student effort level for the test through whether the student responds to questions and whether they give inconsistent responses to questions that should be highly correlated to suggest a low effort. In the Gansu Survey of Children and Families used for my research, the measure for student effort is a question directly asked to the students. The survey asks the student to answer whether they study hard for math and Chinese and choose between "Seldom study hard," "Sometimes study hard," and "Always study hard." My data has more direct measures of effort and I incorporate this measure of effort into my regression models to test if it has an effect on the student outcome and to what degree it does. Later in the paper, I also include a section describing how the effort variables change over the different survey waves.

Other researchers find that effort has no statistically significant effect on student performance. Using data from student behavior in an online economics course, the researchers use total time spent on the website as a proxy for effort (Patron and Lopez, 2011). I find this measure flawed because the amount of time spent using a website could be affected by other

factors other than effort. However, I still take into account time spent on math and Chinese in my analysis.

Other factors also affect students' education achievement in school and I take into account the level of the father and mother's highest education attainment level and also the household income level. Dubow, Boxer and Huesmann (2009) found that students with higher parental education levels experienced positive effects beyond higher academic achievement throughout the school years. They had positive effects of obtaining more prestigious occupations and higher education level into middle adulthood. They found that the long-term positive effects of the parents' education are indirect.

My research draws upon previous efforts to find long-term effects of education on an individual's life, the effect of parental attitude on a child's education, and the role self-perception plays in decision-making. There is evidence that early childhood education can have effects on future income (De Haan and Leuven, 2020). In an examination of the Head Start program in the United States, De Haan and Leuven (2020) found that the Head Start program produced statistically significant results in higher wages and more years of education received in the recipients when they are in their 30s.

There is also evidence that parental involvement in a child's life is an important factor in their future economic or educational outcome. Parental attention to their children's schooling has been found to have a positive effect on the children's attitude toward their own education (Avvisati et al., 2014). A program of parent-teacher meetings in a deprived district in France has shown results in decreasing student truancy and the number of disciplinary actions that occur at the school. However, the study did not find statistically significant increases in the children's

grades. In my data set, there are both questions to the child about their parents' attitudes towards education and also questions about the topic directly to the parents. There are also surveys of teachers and school administrators and those data might be helpful for judging the quality of the education that the children receive as well.

Previously, my data set, the Gansu Survey of Children and Families, has been used to research topics such as teacher retention, gender inequality, etc. The study most relevant to my own is Zhang et al. (2007), which studies the attitude the women surveyed toward the education of their daughters versus the attitude they held toward their sons. They found that mothers who viewed male and female children with more equality tended to have higher expectations for their daughters. This study found that the parents had higher aspirations for their sons if they were doing well in school. Therefore, it is quite likely that a child showing initial promises of educational achievement might receive greater expectations from their parents and receive more resources.

III. Data

Building on the existing research, I use a setting in Gansu, China and examine the effects of parenting style and student effort on a student's future education attainment and career outcome. The data used in this research is from the Gansu Survey of Children and Families conducted in China 5 times over a 15-year period from 2000-2015. The original sample size in the year 2000 is 2000 children from age 9 to 12 and by the 2015 wave of surveys, 1383 of the original 2000 children surveyed remain. The survey was conducted using randomized samples of counties in the province, then municipalities within each county are randomly chosen, then villages were randomly chosen within each municipality, then households with children aged

between 9 to 12 were chosen. Researchers from the University of Pennsylvania (2000) conducted this survey to research education attainment, inequality, gender inequities, etc., with field work conducted by Statistics Bureau interviewers and supervisors trained by faculty and graduate students from Northwest Normal University in Gansu. To further understand the research setting, I provide demographic and economic measurements of Gansu province.

Gansu is in the northwestern part of China and contains the upper reaches of the Yellow River. As of 2015, the population of Gansu reached 26 million, with an urban population of 11 million people and a rural population of 15 million (Cao 2015). The birth rate in 2015 was 12.36 per 1000 people. As of the 2010 census, the province contained 9.43% ethnic minorities. The ethnic minorities are defined as non-Han Chinese, the predominant ethnic group in China. Out of this, the Hui people, made up 4.92% of the province. The 98 percent of the households originally sampled were of Han Chinese descent, and the remaining 2 percent is primarily Hui. Gansu is economically impoverished as part of the lower income western region of China. In 2019, the GDP per capita was 32,995 Chinese yuan, or 4,600 U.S. dollars (HKTDC research, 2021). Gansu's main economics sectors are agriculture and local service work (Cao, 2015).

In 2000, Gansu's population was 24.62 million people. 14.34% of the population was illiterate, down from 27.93% in 1990 (cnstats, 2001). 75.99% of the population was rural and 24.01% of the population was urban. 0–14-year-olds made up 27% of the total population and 15–64 year-olds made up 68.00% of the population. In 2010, Gansu's population was 25.58 million people (cnstats, 2011). The illiteracy rate decreased from 14.34% in 2000 to 8.69% in 2010. In 2010, 63.88% of the population was rural and 36.12% of the population was urban. 0–14-year-olds made up 18.16% of the total population and 15–64-year-olds made up 73.61% of

the population. From the year 2000 to 2010, the population of Gansu has grown and urbanized, and the demographic is now older.

The surveys were conducted in 2000, 2004, 2007, 2009, and 2015. These surveys were spaced to reflect the stages of life of a person from adolescence to adulthood. The initial survey conducted in 2000 had a sample size of 2000 rural children in Gansu aged 9-12. In 2004, the researchers surveyed the oldest younger school-aged sibling of the original target child from those 2000 students surveyed in the year 2000. In 2007, the survey subjects were further expanded to a sample of 1,400 children in the same villages aged 9-15 that year. The surveyed children came from 100 sample villages. Beyond just the individual child, the child's household, mother, the child's primary and middle school teachers and principals, and village leaders were all surveyed. The surveys are linked based on the county, township, village, family, school, and teacher. Each child surveyed is also given a unique ID.

The dataset is from the Chinese province of Gansu on an individual level. This data was collected over a 15-year period in waves and asks a variety of questions. In the first survey, the child surveyed is 9 to 12 years old, and by the 5th wave of surveys in 2015 they were 24-27. In the intervening years between 2000 and 2015, three other surveys were conducted in 2004, 2007, and 2009. Of the most interest to my research is the self-perception of the children and parental attitudes revealed in these surveys.

The year 2000 survey includes sections on study habits, time management, views on education, environment of the school, life description, and parenting style. For a clearer description of the data, I will group the survey questions into several groups.

The first group is parenting style and self-perception. The survey refers to self-perception as life description and asks questions about whether the child is satisfied in their life, their

outlook on the future, their self-esteem, etc. The life descriptions also include their social behaviors, both positive and negative. The parenting style section asks about requirements for the child's behavior set by the parent, rules in the house, whether they are willing to communicate with their parents, how much attention the parents give the child, etc.

The second group is individual life characteristics. This includes study habits, time management, and views on education. Study habits cover their level of effort for different subjects, whether they are disciplined in school and what punishments are handed out. Time management asks how much time the student spends on various tasks such as working on the farm, playing with friends, and doing homework. It also asks about how much the child reads outside of school, their level of media consumption, and whether they receive outside tutoring. In views on education, the questionnaire asks about the level of education the person wants to achieve, their views on if school is important and what factors contribute to the success of a student, and whether their mother finds education and school related activities are important.

The third group is school characteristics. The environment of school asks questions about whether the child faces violence at school, whether the teacher shows up to teach, how well the food is, etc. The survey also asks questions to the schoolteachers and principal of the students and ask them about the education environment that the individual children are experiencing.

In addition, there are also surveys of individual education achievement scores in the subjects of math and Chinese for the years 2000 and 2004. Unfortunately, each individual child only took either the math or the Chinese assessment in the year 2000 so they did not have both scores while in the year 2004, the surveyed students who took the assessment had both a score in the math assessment and also a score in the Chinese assessment. In the year 2007, the population

of students who took the assessment were no longer the same as the original 2000 from the year 2000 and I could no longer link the data on the individual level.

Table 1: Data Linking

number of kids and households verified as the same one across all available waves							
Data Source		number of kids			among 2000 households of target kids		
		same (1)	not sure (2)	different (3)	same (4)	not sure (5)	different (6)
data in all 5 waves	00,04,07,09,15	3,470	36	306	1,383	10	97
data in 4 waves	00,04,07,09	398	5	48	163	2	15
	00,04,07,15	89	6	17	36	3	7
	00,04,09,15	274	2	41	101	1	12
	00,07,09,15	25	0	0	10	0	0
data in 3 waves	00,04,07	100	0	3	45	0	1
	00,04,09	63	3	11	23	1	5
	00,04,15	17	0	0	7	0	0
	00,07,09	2	0	0	1	0	0
	00,07,15	3	0	2	1	0	1
	00,09,15	36	10	4	16	3	1
data in 2 waves	00,04	14	0	0	6	0	0
	00,07	6	0	0	3	0	0
	00,09	8	8	8	4	3	3
	00,15	13	0	4	5	0	1
TOTAL		4,518	70	444	1,804	23	143

Notes: This table reports the number of all kids and households verified as the same one across all available waves. Column (1) reports the number of households which are verified as the same household across all the waves the household surveyed. Column (2) reports the number of households which we are not sure if they are the same families being surveyed across the waves. Column (3) reports the number of families verified as the different households being surveyed across waves. Column (4) reports the number of households which are only being surveyed once and thus the definition "same household" is not well defined. Similarly, column (5)-(8) reports the number of kids verified as the same, not sure, different or undefined. However, column (5)-(8) focus on the 2000 sampled kids from the original Cohort 2000.

One of the most unique and interesting aspects of this data set is the repeated survey as previously described. Data loss is a worry here. This table shows the proportion of children who were surveyed in the previous waves.

From column 4 of Table 1, we can see that out of the 2000 households from which the original 2000 children were selected in the first wave of surveys, 1383 children were retained through all the waves. This is still a good sample. 2000 households were surveyed during the first wave. By the 5th wave in 2015, 1490 households out of the original 2000 completed every

survey (2000,2004,2007,2009,2015). Out of these 1490 households, 1383 kids were confirmed to be the same kid as the one who was originally surveyed in 2000.

The dependent variables of my research are the education attainment level and career outcomes of the students who were first surveyed in 2000. In the later surveys, this information is asked. The table below summarizes the educational outcome of the students who were in the original cohort that remained from 2000 to 2015.

Table 2: Education Outcome in 2015

	Frequency	Percent	Cumulative
no school	5	0.34	0.41
did not finish primary school but is literate	25	1.72	2.14
primary school	165	11.38	13.52
middle school	406	28	41.52
high school	141	9.72	51.24
vocational school/technical school	221	15.24	66.48
community college/associate degree	278	19.17	85.66
undergraduate	196	13.52	99.17
masters	12	0.83	100
Total	1,450	100	

In our empirical analysis, I employ regression models between successive waves of surveys and see how the main independent variables of student effort and parenting style impact the achievement score or output measure in the next survey. Particularly, I am regressing the math and Chinese scores in 2004 on the student and parent survey responses and characteristics

from the 2000 survey. I then study the student outcome in 2015, the wave of surveys to students who are now adults in their mid-20s and see how the education attainment and employment outcome of the students are affected by their individual characteristics and survey responses in 2004. In Table 3, I present the outcome variables in the 2004 survey.

Table 3: 2004 Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Math Score 2004	1272	17.258	13.286	0	50
Chinese Score 2004	1272	20.85	11.067	0	49.5
Combined Score 2004	1272	38.107	20.510	0	91

The math score and Chinese score are each measured from 0 to 50. They are calculated based on how the student scores on an assessment in 2004 that is given out to all student respondents to the survey. The combined score in 2004 is created by adding together the math and Chinese scores and the maximum score for an individual is 100. The 2004 scores are a measure of the student's education achievement that is the outcome variable for the student characteristics, parent characteristics, and the baseline score of the students in the year 2000.

Table 4: 2015 Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Graduated Middle School	1401	.867	.34	0	1
Graduated High School or equiv.	1401	.587	.493	0	1
Graduated College or equiv.	1401	.338	.473	0	1
Employed or in school	1401	.859	.348	0	1
Monthly Income (CNY)	1004	3197.846	2265.077	8	31000

In 2015, the outcome variables are whether the student graduated middle school, whether the student graduated high school or equivalent, whether the student graduated college or equivalent, whether the student was still employed or in school vs not (0 means unemployment), and monthly income measured in Chinese yuan. Table 4 includes the cohort of students who were both in the 2000 survey and the 2015 survey. The first three outcome variables are dummy variables with an answer of 1 being the individual having successfully done this or in the case of “employed or in school,” a 1 means the person is employed or in school. We see that 86.7% of the students graduated middle school, 58.7% of the students graduated high school or equivalent, and 33.8 percent of the students graduated college or equivalent. 85.9 percent of the student surveyed or either employed or in school. The mean monthly income among the sampled was 3197.84.6 yuan, around \$450 per month.

IV. Empirical Analysis

In two steps, I see how education attainment level builds up between different waves of surveys. Between each wave of survey, the student brings with them what they have learned so far as a baseline of education attainment and the student effort and parental effort are the independent variables that increase the outcome level of education. Given that the first survey wave occurred in 2000 and the outcome data come from the highest level of education and income level that the students achieved in 2015, I use the data from 2000, 2004, and 2015. Each step is between successive waves. I do not use the year 2007 wave since the achievement scores collected in the year 2007 to assess gains in education on the individual level are not linked to the individual who were in the sample in 2000. This is likely due to this assessment being taken by the new cohort of 1400 students and not the original group of students in the year 2000. I do

not use the 2009 survey because the content of the 2009 survey is no longer asking about the effort of the student and the attitudes of the parents as the previous surveys had been asking. Instead, the 2009 survey was asking questions about the education attainment and the income of the students. These are outcome measures that the 2015 survey sufficiently covers. The 2015 survey asks the same question about what is the highest level of education attainment that the child achieved, which is the relevant question in the 2009 survey. Since education achievement is built up on each level, I aim to find out in which periods of the student's education do they make significant progress. The main explanatory variables are the student's self-reported efforts in school on different subjects, the parent's attitude on investing in their student's education, and the outcome variable is the education attainment level of the student at the next level.

Student effort is an important explanatory variable for their attainment of higher levels of education. Effort is measured by the student's response to the questions: do you study hard for math and do you study hard for Chinese. The students answer, "seldom study hard," "sometimes study hard," or "always study hard." I make the assumption that student effort is not a result of doing poorly on tests, which means that those students who are measuring higher on their effort are actually those with the worst education outcomes. I tested this hypothesis by regressing student outcome directly on the effort measures in the two earliest waves of the survey. I find a positive relationship between these two measures, which means those who exert higher effort into their education are correlated with higher levels of educational attainment rather than lower. This means the students who try harder are not responding to doing badly on the test. Another concern is that students do not try in 2000 because they are too young and either do not understand the importance of their education or the subjects taught in elementary school did not require high effort. This theory is also not true since I compared the patterns of how the same

individual answered the effort variable from the 2000 survey to the 2004 survey. I actually found that between increasing in self-reported effort, staying the same, and decreasing, the majority of students actually responded they tried less by 2004. Again, this makes our effort measure stronger since students were genuinely putting in effort, even when they were young.

I also use time spent on certain subjects as a proxy for their effort in those subjects in addition to the effort variable itself. Choosing to spend more time on a subject can mean that they are trying harder on a subject. The time spent questions are asked as “Compared with your peers, how much time do you spend studying Math/Chinese.” For each subject, the student answers with “less time than most,” “about the same as most,” and “More time than Most”

The parental style is measured by the parent’s answers to 4 questions. They are asked if the following happens “never”, “sometimes”, or “often”. The questions are: “If your child has done something wrong, you or your husband just let it go.” “If your child has done something wrong you or your husband ask for an explanation and then discuss with the child the correct behavior.” “If your child has done something wrong, you or your husband point it out to him/her.” “If your child has done something wrong, you or your husband scolds the child or even beats him or her.” As discussed earlier, the predominant forms of parenting styles are permissive, authoritative, and authoritarian (Doepke et al., 2019). Given the survey questions asked by the Gansu Survey of Family and Child, parents who let their child’s mistakes go exhibit self-reported permissive parental style. Those parents who discuss with their child about the correct behavior to take or point out their mistakes exhibit a self-reported authoritative parental style. Lastly, the parents who respond to their children’s mistakes with scolding or beating exhibit an authoritarian parental style. Permissive parental style is low cost to the parents in terms of parental attention and authoritative parenting styles exhibit higher degrees of parental attention.

The authoritative parent corrects their children's behavior through persuasion, while the authoritarian parent imposes rules on their children through coercion. The child receiving authoritative parenting likely corrects their behavior with reasoning, while the child receiving authoritarian parenting changes their behavior due to fear of their parents' punishments.

Additionally, the parent is asked to answer with "never", "sometimes", or "often" to the statement "you or your husband often encourages your child to work hard at everything." I want to see if the students being motivated by their parents eventually have better education achievement and outcomes. With encouragement to the child, the child might try harder, or at least understand the signal from their parents that it is important to work hard in school.

Section 1: Survey 2000 and 2004

I begin by first understanding the factors affecting the students' education achievement in 2004. To do so, I run the following regression:

$$2004 \text{ educ achievement} = a + 2000 \text{ educ achievement} + \beta_1(\text{ParentAtti}) + \beta_2(\text{StudAtti}) + \text{Controls}_i + e_i \quad (1)$$

The 2004 education achievement is regressed on by the year 2000 education achievements, parental attitudes, student effort, parental education level, and household income. The education achievements in 2004 and 2000 are both measured by how they scored on a subject specific assessment in math and Chinese. First, I regress the math score in 2004 on the math score in 2000, parental attitudes, math specific student effort, parental education level, and household income. Second, I regress the Chinese score in 2004 on the Chinese score in 2000, parental attitudes, Chinese specific student effort, parental education level, and household income. Third, I regress the combined score of Chinese and math on the math and Chinese scores

in 2000, the effort in math and Chinese by the student, parental education level, and household income. In Table 5, regression results are presented.

Table 5 2000 to 2004 Regressions

	(1) Math Score in 2004	(2) Chinese Score in 2004	(4) Combined Score in 2004
Student Baseline:			
math score in year 2000	.018 (.021)		
Chinese score in year 2000		.04** (.02)	
Combined math and Chinese score for 2000			.03 (.027)
Student Effort:			
Time spent on Math in 2000	1.957** (.859)		-.346 (1.089)
Effort on Math in 2000	1.731** (.805)		1.614 (1.033)
Time spent on Chinese in 2000		-.317 (.724)	.809 (1.006)
Effort on Chinese in 2000		.336 (.671)	1.026 (1.023)
Parent Attitude:			
do nothing when your child has done something wrong	.654 (.621)	.404 (.518)	.693 (.696)
discuss with your child the right thing to do when your child does wrong	1.368 (.864)	.133 (.721)	1.786** (.885)
point out what your child has done wrong	-1.867* (1.041)	.855 (.82)	-1.489 (1.185)
encourage your child to work hard	1.147 (.988)	-.84 (.788)	.462 (1.078)
blame or beat him/her when your child has done something wrong	-.996 (1.015)	.885 (.641)	.89 (1)
Father's highest level of education was middle school	1.126 (1.199)	-.326 (.982)	.16 (1.362)
Father's highest level of education was trade school	11.735** (5.683)	.053 (2.621)	15.147** (6.79)
Father's highest level of education was high school	2.256 (1.502)	.267 (1.265)	3.831** (1.725)
Father received higher education	-.998 (1.458)	5.841* (3.295)	3.004 (8.942)
Mother's highest level of education was middle school	3.711*** (1.353)	-2.084** (1.059)	1.995 (1.539)
Mother's highest level of education was high school	-.781 (2.853)	-4.759** (2.032)	-1.687 (3.008)
Household income measure in 2000	0 (0)	0 (0)	0 (0)
_cons	5.506 (4.194)	17.217*** (3.497)	19.97*** (5.048)
Observations	622	650	1401
R-squared	.069	.028	.024

The 3 regressions in Table 5 test the relationship between the baseline achievement level of the students in the year 2000, their effort in math and Chinese, their parents' parenting style, parental education level and how much effects those factors had on their outcome in the year 2004. The outcome variable is the math score in 2004, the Chinese score in 2004, and the combined score of their assessment score in math and Chinese in the year 2004 respectively. Math score in 2000 is positively correlated with math score in 2004 and Chinese score in 2000 is positively correlated with Chinese score in 2004. The effect is small, a 1-point increase on the Chinese score in 2000 is expected to increase the individual's Chinese score in 2004 by 0.04. Math scores and Chinese scores are both recorded as between 0 to 50. I find a statistically significant effect of student effort as measured by the time they spent on math in 2000 on their math score in 2004 and their reported effort on math in 2000 on their math score in 2004. The math effort variable is recorded as "seldom work hard," "sometimes work hard," and "always work hard." Students who answer one category higher in the effort variable are expected to have a 1.73-point increase in the math score, or 10.02% of the mean 2004 math score of 17.26 points. The time spent on math variable is recorded as compared to your classmates, how much time do you spend on studying math. The answers are "less time than most," "about the same as most," and "more time than most." Students who answer one category higher on their time spent on math are expected to score 2 points higher than those who scored one category lower. A 2-point increase is 11.34% of the mean 2004 math score of 17.26 points. Meanwhile, I do not find the same statistical significance for effort measures in Chinese. In terms of the parental attitude variables, I find that the questions that indicate authoritative parenting only has a positive and significant effect on the child's score in the combined score column. This is not true for the other variables. Additionally, I find that children with mother and father with higher levels of

education are positively correlated with the child's outcome. I do not find household income to be significant.

Section Two: 2004 to 2015

The next model is seeing how the results of the 2004 survey affected the outcome variables measured in 2015. I skip the 2007 wave since the education achievement measures I need to link the 2004 wave of survey data to 2007 does not link to the surveyed population of the 2000 cohort. Additionally, the survey in 2009 only targeted the surveyed children so I no longer have the opportunity to estimate the effect of the parental attitude on the student's career and education outcomes. We will see the longer-term effect of the 2004 child characteristics on their outcome in 2015. The surveyed children are about 14 in the year 2004 and 25 in 2015. By the year 2015, most students have finished their highest level of education and we have their employment data as well. The education outcome measures that I use are dummy variables for whether the student graduated middle school, whether the student graduated high school or equivalent, and whether the student graduated college or equivalent. The equivalent to high school here refers to vocational schools, and equivalents to college refer to vocational colleges that provide higher vocational training. They are roughly the same length as their corresponding counterparts in high schools and college respectively. I use a logit model for the dummy variable outcomes of education. In addition to the education dummy variables, I also use a variable that is 1 if the student is employed or going to school and 0 otherwise. I use this variable as a measure whether the surveyed is employed. Lastly, I use the natural log of the income of those who were working in 2015. This allows me to assess the percent increase in income of the surveyed as a percent change of 2004 achievement scores, also log transformed. The general empirical specification is the following:

$$2015 \text{ outcome} = a + 2004 \text{ educachievement} + \beta_1(\text{ParentAtti}) + \beta_2(\text{StudAtti}) + \text{Controls}_i + e_i$$

(2)

Table 6: 2004 to 2015 Regressions

	(5) Logit: Graduated Middle School	(6) Logit: Graduated High School	(7) Logit: Graduated College	(8) Logit: Employed or Studying in 2015	(9) ln(monthly income)
Student Baseline:					
Chinese score in year 2004	.007 (.01)	-.006 (.006)	-.001 (.006)	.007 (.009)	
math score in year 2004	.023*** (.008)	.031*** (.006)	.028*** (.005)	.007 (.008)	
ln(Chinese Score in 2004)					-.056 (.039)
ln(math score in 2004)					.049 (.032)
ln(household income in 2004)					-.036 (.027)
Student Effort:					
Effort on Chinese in 2004	.117 (.186)	-.087 (.123)	-.103 (.121)	-.336* (.175)	-.003 (.045)
Effort on Math in 2004	.019 (.182)	.292** (.119)	.395*** (.124)	.166 (.155)	.014 (.051)
Parent Attitude:					
ask for explanation when child does wrong	.298 (.204)	.341*** (.124)	.168 (.123)	.308* (.172)	.145** (.06)
If your child has done something wrong, you correct them	.074 (.215)	.136 (.137)	-.063 (.144)	-.15 (.181)	.003 (.071)
You (or your husband) often encourage your child	.179 (.184)	.054 (.13)	.319** (.138)	.176 (.164)	-.023 (.068)
Controls:					
Father's highest level of education was middle school	.556*** (.214)	.528*** (.138)	.44*** (.145)	.233 (.201)	.006 (.058)
Father's highest level of education was high school	.884*** (.329)	.881*** (.181)	.76*** (.174)	.176 (.256)	-.023 (.069)
Mother's highest level of education was middle school	.461* (.276)	.439*** (.16)	.474*** (.149)	-.035 (.221)	-.014 (.059)
Mother's highest level of education was high school		.821** (.389)	.506 (.326)	.434 (.546)	.022 (.08)
Household Income in 2004	0* (0)	0* (0)	0 (0)	0 (0)	
_cons	-.613 (.688)	-2.141*** (.487)	-3.306*** (.531)	1.106 (.681)	7.924*** (.257)
Observations	1192	1243	1243	1243	439
R-squared					.032
Pseudo R-squared	0.0607	0.0798	0.0743	0.0167	

Table 6 reports the results of the regressions I described earlier. Math score in the year 2004 and the effort on math in 2004 have positive and statistically significant effects on the education outcome of the student. Holding other variables constant, a one-unit increase in math score is expected to raise the chance of a student graduating middle school by 2.29% through calculating odds ratios based on the logit model results. A one-unit increase in math score in the year 2004 is expected to raise the chance of a student graduating high school by 3.12% and one unit increase in math score in the year 2004 is expected to raise the chance of a student graduating college by 2.85%. It is interesting to note that the Chinese score in year 2004 does not have a positive or statistically significant association between the math score in 2004 and the effort on math in 2004 on the future outcome of the student. Students who answered one unit higher in the effort in math variable in the year 2004 were 34% more likely to graduate high school and 48.4% more likely to graduate college than their peers who did not, holding other variables constant. In terms of parenting style, I find sporadic statistical significance for a positive effect of parents responding affirmatively to “ask for explanation when child does wrong” and “your (or your husband) often encourage your child” on the chance of graduating high school and graduating college respectively. The father and mother’s education levels also have a positive effect on the student’s education outcome. Holding other variables constant, individuals with a father who graduated high school are 142% more likely to graduate middle school than their peers whose father did not, 141% more likely to graduate high school than their peers whose father did not, and 113% more likely to graduate college than their peers whose parents did not, through the calculations of odds ratios. Similarly, holding other variables constant, individuals with a mother who graduated middle school were 58.5% more likely to graduate middle school than students whose mother did not graduate middle school, 55% more

likely to graduate high school than students whose mother did not graduate middle school, and 61% more likely to graduate college than students whose mother did not. Household income still does not have an effect on the household income in 2004. I do not find significant effects of parental education on the income of the child.

By analyzing the pattern of education achievement retention from wave to wave and seeing how the effort from the student and the parent factor in, there is an emerging picture of how education achievement eventually leads to future education and income attainment. The major finding is that the effort and achievement in math seems to have much more positive effect on the student's education attainment than effort in Chinese or score in Chinese. In both groupings, the household income does not have a significant impact on the education attainment as previous research suggests. This suggests that the income of households does not allow them to make a big difference in how much they can invest in their children. Even households that are slightly richer are not able to switch to schools with more resources or have access to initial help that might have put their children ahead of other children. The parental style of the children do not seem to have strong and consistent effects on the student education achievement or measures of outcome. The level of education achieved by the parents has significant effects on whether the child graduates middle school, high school, and college. Both the father and the mother's education level seem to be important, especially later on in life for the child's outcome. For income outcomes and employment status, the effects mentioned above are less apparent. Instead, it is likely that the extra few years beyond middle school do not materialize into higher pay for the students who continued with their education. The occupations that students who did not graduate middle school engage in are either well-paying or award their experience. By 2015,

these students would have been on the job market for about 10 years and could well be paid more than their former classmates who entered the job market later.

V. Conclusion and Discussion

Approaching education attainment and outcomes with a method of two step regression, I find that the baseline score in math has an insignificant effect on the math achievement early in a child's education career but has a positive and significant effect on the child's highest level of education attained. The effort in math has a positive effect on the student score and outcome but the effort in Chinese does not. The subject specific differentiation is interesting, as Chinese does not seem to have the same statistical effect. I find that parents with authoritative parenting styles and encourage their child to study have positive correlations with better child outcome but do not find it consistently statistically significant. I find that parental education attainment is important for the education attainment of children, The internal factors of a student that drive their education attainment seem to be more important later on in their life. This research points to the significance of math in the Gansu education setting and the importance of parental education level on their children's education.

References

- Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2016). School Quality and the Gender Gap in Educational Achievement. *The American Economic Review*, *106*(5), 289–295. <http://www.jstor.org/stable/43861031>
- Avvisati, F., Gurgand, M., Guyon, N., & Maurin, E. (2014). Getting Parents Involved: A Field Experiment in Deprived Schools. *The Review of Economic Studies*, *81*(1 (286)), 57–83. <http://www.jstor.org/stable/43551666>
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191.
- Bertrand, M., & Pan, J. (2013). The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. *American Economic Journal: Applied Economics*, *5*(1), 32–64. <http://www.jstor.org/stable/43189418>
- Cao, X. (2015). *Gansu*. People's Republic of China Government. Retrieved September 29, 2022, from http://www.gov.cn/guoqing/2016-11/21/content_5135218.htm
- Chevalier, A., Harmon, C., O'Sullivan, V., & Walker, I. (2013). The impact of parental income and education on the schooling of their children. *IZA Journal of Labor Economics*, *2*(1), 1-22.
- cnstats. (2001). *Gansu Province 2000 5th National Census*. Chinese Census Statistics. Retrieved September 29, 2022, from <https://www.cnstats.org/rkpc/201112/gss-2000-kiuqw.html>

- cnstats. (2011). *Gansu Province 2010 6th National Census*. Chinese Census Statistics. Retrieved September 29, 2022, from http://www.stats.gov.cn/tjsj/tjgb/rkpcgb/dfrkpcgb/201202/t20120228_30383.html
- Collins, J.L. (1982, March). Self-efficacy and ability in achievement behavior. Paper presented at the annual meeting of the American Educational Research Association, New York.
- De Haan, M. and Leuven, E., 2020. Head Start and the Distribution of Long-Term Education and Labor Market Outcomes. *Journal of Labor Economics*, 38(3), pp.727-765.
- Doepke, M., Sorrenti, G., & Zilibotti, F. (2019). The economics of parenting. *Annual Review of Economics*, 11, 55-84.
- Dubow, E. F., Boxer, P., & Huesmann, L. R. (2009). Long-term Effects of Parents' Education on Children's Educational and Occupational Success: Mediation by Family Interactions, Child Aggression, and Teenage Aspirations. *Merrill-Palmer Quarterly*, 55(3), 224–249. <http://www.jstor.org/stable/23096257>
- Guang, L., Adams, J., Gaetano, A., Whyte, M. K., Gustafsson, B., Hannum, E., ... & Ximing, Y. (2010). *One Country, Two Societies: Rural-urban Inequality in Contemporary China* (Vol. 16). Harvard University Press.
- HKTDC research. (2021, May 17). *Gansu Market Summary*. HKTDC research. Retrieved September 29, 2022, from <https://research.hktdc.com/sc/data-and-profiles/mcpc/provinces/gansu>
- OECD. *Education in China: A Snapshot*. OECD, 2016.

- Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into practice, 41*(2), 116-125.
- University of Pennsylvania and Hong Kong University of Science and Technology. (2015). Gansu Survey of Children and Families. Retrieved from <https://web.sas.upenn.edu/gansu-survey/>.
- Wuepper, D., & Lybbert, T. J. (2017). Perceived self-efficacy, poverty, and Economic Development. *Annual Review of Resource Economics, 9*(1), 383–404.
<https://doi.org/10.1146/annurev-resource-100516-053709/>.
- Zamarro, G., Hitt, C., & Mendez, I. (2019). When students don't care: Reexamining international differences in achievement and student effort. *Journal of Human Capital, 13*(4), 519-552.
- Zhang, Y., Kao, G., & Hannum, E. (2007). Do mothers in rural China practice gender equality in educational aspirations for their children? *Comparative Education Review, 51*(2), 131-157.