

A Quantile Regression Analysis of Family Background Factor Effects on Mathematical Achievement

Maozai Tian
Renmin University of China

Abstract: Family background factor can be a very important part of a person's life. One of the main interests of this paper is to investigate whether the family background factors alter performance on mathematical achievement of the stronger students the same way that weaker students are affected. Using large sample of 2000, 2001 and 2002 mathematics participation in Alberta, Canada, such questions have been investigated by means of quantile regression approach. The findings suggest that there may be differential family-background-factor effects at different points in the conditional distribution of mathematical achievements.

Key words: Factors, family background, mathematics education, mathematical achievement, quantile regression.

1. Introduction

Children's mathematical achievements have long been a concern of society. Mastering mathematics has become more important than ever. Previous research indicates that a senior high school student with a strong grasp of mathematics has an advantage in academics and in the job market, i.e., mathematical achievement is a key to college entrance and success in the labor force.

For several decades a number of studies have been focus on gathering and investigating information from many variables with effects on the mathematical achievement. The social, economic, and cultural factors that are either in favor of or not conducive to children are not well understood. Among our main research interests here are the variables including family background factors, such as *number of parents, number of siblings, mother's socioeconomic status, father's socioeconomic status, gender, immigrants, language problem, native and minority*, etc.

As for *gender* difference in learning mathematics, evidence shows that females are not likely to believe that mathematics has utility in their lives (Fennema and Sherman, 1978). They see mathematics as unconnected to a relationship model of thinking. Even if females continue to take mathematical courses, they are apt

to find that they themselves do not like these courses. However, liking a subject is key to succeeding at it (Lockhead, *et al.*, 1985).

Some researches on *immigrants'* school performance suggest that their performance is above averages (e.g., Rumbaut, 1996¹; Viadero, 1997; Lapin, 1998). While there is also evidence that immigrant children, especially Hispanics and others with impoverished background, suffer poor academic achievement and lower educational attainment (e.g., McPartland, 1998²; Vernez and Abrahamse, 1996). Also, more recent studies of immigrant children's academic achievement provide some insights for understanding the variation among immigrant children's academic achievement. For example, Hao *et al.* (1998) used the concept social capital to explain immigrant children's academic performance.

It's well known that *language problem* limits immigrant children's learning on key subject areas such as mathematics and science. Living in socially and linguistically isolated communities, poor immigrant children can hardly improve their new language skills and the language barriers persist over the school years. On the other hand, bilingual proficiency, defined as the mastery of both the mother tongue and a new language, is found to be a strength for immigrant children's cognitive growth (e.g., Hao and Portes, 1998).

Several authors have recognized that *minorities* may see mathematics as a White domain, are less likely than Whites to understand its future value, and are negatively influenced by the school staff's attitudes toward them and their work (Mathews 1983). Educational reform advocated by the politicians and policy makers has been performed in enhancing minority mathematical achievement including good discipline and attendance, small class size, placement in advanced tracks, and materials that affirm the important role of minorities in mathematics (Mathews 1983; Taylor 1983).

Generally speaking, studies mentioned above have primarily dependent on classical mean regression methods such as Ordinary Least Squares(OLS) or Instrumental variables (IV). These methodology may miss crucial points such as how family background factors affect mathematical achievement differently at different quantiles of the conditional test score distribution. What is worse, these approaches cannot be used to characterize the entire conditional distribution of mathematical achievement given high-dimensional covariates (family background factors) and the estimated coefficient vector (marginal effects) is not robust to outlier observation on mathematical achievement.

¹Rumbaut, R. G. (1996). The new californians: assessing the educational progress of children of immigrants. The Education Resources Information Center (ERIC). See http://eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/27/36/7/e.pdf

²McPartland, J. (1998). The adaptation of immigrant children in the American Education System Center for Research on the Education of Disadvantaged Students. Project #7126. See <http://www.ed.gov/officers/OERI/At-Rish/cds1rp22.html>

Fortunately, the drawbacks mentioned above can be overcome by combining another statistical method called quantile regression (QR), which was proposed by Koenker and Bassett (1978) and has become a comprehensive approach to the use of linear and nonlinear response models for conditional quantile functions. Roughly speaking, QR, which is based on minimizing “Check function” residual, enables us to estimate all conditional quantile functions, just as classical linear regression techniques based on least squares estimation offer a mechanism for estimating conditional mean functions. Therefore, QR is gradually emerging as a unified statistical methodology and find wide applications in education, economics, biology, ecology, finance, econometrics, statistics, and applied mathematics.

Many studies have focussed on the school background factors, such as the size of class. Using data from the High School and Beyond longitudinal study, Ehrenberg and Brewer (1994) estimated the extent to which school characteristics and teacher characteristics influence the probability that public school students drop out of high school between their sophomore and senior years. Ehrenberg and Brewer (1995) examined the effect of school quality on student achievement. They have found that verbal aptitude scores of teachers influenced synthetic grain scores for both black and white students. Corman and Chaikind (1998) examined the school performance and behavior of children age six to fifteen years who were born weighing less than 2500g, compared with a group of normal birth weight children, holding constant socio-economic characteristics of the child and family. Eide and Showalter (1998) used quantile regressions to estimate whether the relationship between school quality and performance on standardized tests differs at different points in the conditional distribution of “test score gains”. Their results suggested that there may be differential school quality effects at different points in the test score gain conditional distribution. Levin (2001) performed a quantile regression analysis of the controversial topic of class size and peer effects on scholastic achievement.

To the best of my knowledge, no paper has ever dealt systematically with the family background factors impacting on the student’s mathematical achievement based on quantile regression approach. There are 9 main findings in the paper showing some family background factors which appear to have no effect for the average mathematical achievements may indeed matter at other points in the conditional distribution of mathematical achievement.

This paper is organized as follows: Section 2 describes the data used for the mathematical achievement regression. Section 3 introduces the quantile regression model. The estimation results are reported in Section 4. The interval concepts for the estimated effects and related interpretations are discussed in Section 5. The conclusion is presented in the last section.

2. Data

The data come from *Canadian Center for Advanced Studies of National Databases*. Their research study is entitled “A Longitudinal Study of Mathematics Participation in Alberta, Canada”. One of the purposes was to determine the relationship between mathematical achievement of students in senior high school (from Grade 10 to Grade 12) and other factors including the social, economic, and cultural ones. They mainly employed the classic mean regression method based on hierarchical linear model.

Table 1: Descriptive statistics of selected family background factors

Description	Grade 10		Grade 1		Grade 12	
	Mean	Std	Mean	Std	Mean	Std
number of parents	1.83	0.39	1.83	0.39	1.82	0.43
number of siblings	2.23	1.59	2.16	1.47	2.21	1.53
mother’s socioeconomic status	0.00	1.67	0.00	0.73	0.00	0.72
father’s socioeconomic status	0.00	1.73	0.00	0.75	0.00	0.75
female	0.50	0.50	0.50	0.50	0.50	0.50
not born in Canada	0.05	0.22	0.05	0.22	0.05	0.22
language problem	0.05	0.21	0.05	0.21	0.05	0.21
native	0.06	0.24	0.06	0.24	0.06	0.24
minority	0.08	0.27	0.08	0.27	0.08	0.27
mathematical achievement	23.14	7.06	25.77	7.80	27.92	8.25

Selected family background factors: number of parents (0=no parents, 1=single parent, 2=two parents); number of siblings (1=one sibling, ..., 9= nine siblings); female (1 if child is female); not born in Canada (1 if child not born in Canada); language problem (1 if child has language barrier); native (1 if child is native); minority (1 if child is minority). father’s socioeconomic status and mother’s socioeconomic status are measured by the International Socioeconomic Index (ISEI).

The data set was collected once a year (in May or another time arranged with each school in Alberta, Canada) for three years (from 2000 to 2002). The first instrument is a student questionnaire (30 minutes), which contains many items. Among these items are the family background factors, such as *number of parents*, *number of siblings*, *mother’s socioeconomic status*, *father’s socioeconomic status*, *female*, *not born in Canada*, *language problem*, *native* and *minority*, etc.

The resulting sample includes 1454 students in 35 schools. Table 1 contains descriptive statistics for the family background factors used in the coming analysis. The definitions for the variables used in the empirical analysis are described clearly in Table 1. The only other worthwhile mention is that *father's Socioeconomic status* and *mother's Socioeconomic status* are usually measured by the International Socioeconomic Index (ISEI), a measure based on family income, parental education level, parental occupation, and social status in the community. Usually, *parents' socioeconomic status* is an international educational indicator. Research shows that families with high socioeconomic status often have more success in preparing their young children for schools because they typically have access to a wide range of resources to promote and support young children's development. They are able to provide their young children with high-quality child care, books, and toys to encourage children in various learning activities at home. Also, they have easy access to information regarding their children's health, as well as social, emotional, and cognitive development. In addition, families with high socioeconomic status often seek out information to help them better prepare their young children for school.

A brief look at the table reveals nothing too surprising. Note that the scores of mathematics increase monotonically from Grade 10 to Grade 12.

3. Methodology

Quantile regression (QR), as introduced by Koenker and Bassett (1978), is gradually developing into a comprehensive approach to the statistical analysis of linear and non-linear response model. Useful features of QR are as follows: (a) the models can be used to characterize the entire conditional distribution of a dependent variable given regressors; (b) the resulting estimated coefficients from QR are robust, i.e., not sensitive to the outlier observations on the dependent variable; (c) the resulting QR estimators are more efficient than those from OLS in the case that the error term is non-normal; (d) potentially different solutions at different quantiles may be interpreted as differences in the response of the dependent variable to change in the regressors at various points in the conditional distribution of the dependent variable; and (e) a linear programming representation (LP) makes QR-estimation easy.

There are at least four equivalent mathematical definitions of quantile regression:

(I) definition based on the conditional quantile function.

Let $q_p(x)$ be the p -th quantile of the dependent variable Y given $X = x$. In this case, $q_p(x)$ can be found by solving

$$F(q_p(x)|x) = P(Y \leq q_p(x)|X = x) = p,$$

where F is the cumulative distribution of Y ;

(II) definition based on the quantile regression model (Bailar, 1991):

$$Y = x^T \beta + \epsilon,$$

where the error term ϵ is assumed to satisfy $Quantile_p(\epsilon) = 0$. In standard linear regression model, the error term is assumed to be a Gaussian error;

(III) definition based on a check function (Koenker and Bassett, 1978):

$$\min_{\beta \in \Theta} E \{ \rho_p(\mathbf{Y} - \mathbf{X}'\beta) | \mathbf{X} = x \},$$

where $I_A(z)$ the usual indicator function of the set A , Θ is a parametric space for β ,

$$\rho_p(z) = pzI_{[0,\infty)}(z) - (1-p)I_{(-\infty,0)}(z)$$

is called check function; and

(IV) definition based on asymmetric Laplace density (Yu and Moyeed, 2001):

$$f(\epsilon) \propto \exp \left\{ - \sum_{i=1}^n \rho_p(y_i - x_i^T \beta) \right\},$$

where $f(\epsilon)$ is the probability density of the model error ϵ .

In this paper, we use definition III, in which the dependent variable \mathbf{Y} is the score of mathematics and the independent variable \mathbf{X} is the family background variable mentioned in Table 1.

4. Estimation Results

In this section, we present the results for quantile regression and linear regression conditional on family background factors in the 10-th, 11-th and 12-th grades. To do so, we may see the data imparts a highly structure pattern of interdependence among observations and the changes of the pattern over these years. Quantile regressions were estimated at five different quantiles: 5%, 25%, 50%, 75% and 95%. For comparison, the empirical results from ordinary least squares regression are also reported. Quantile regression software is now available in most modern statistical languages. We here use R which is an open source software project built on foundations of the S language of John Chambers. Capabilities for quantile regression are provided by the "quantreg" package. Once R is installed on a networked machine packages can be easily installed using the command `install.packages("quantreg")` in an R session.

4.1 The estimated effects in grade 10

Table 2 shows the comparison of quantile regression and OLS results conditional on family background factors in Grade 10. The estimated standard errors are reported in parentheses. We first note from the OLS results that those factors such as *number of siblings*, *mother's socioeconomic status*, *female*, *not born in Canada* and *minority* are all insignificantly different from zero. In contrast, the effects of both *number of parents* and *father's socioeconomic status* are significantly positive values whereas the effects of both *language problem* and *native* are significantly negative ones.

Table 2: Comparison of quantile regression and OLS results conditional on student background variables in grade 10

Description	Quantile regression results					OLS
	5%	25%	50%	75%	95%	
number of parents	1.25 (1.01)	2.16* (0.58)	1.88* (0.63)	1.47* (0.67)	1.04 (1.25)	1.57* (0.47)
number of siblings	-0.14 (0.22)	0.11 (0.15)	0.05 (0.13)	-0.08 (0.17)	-0.29 (0.20)	0.01 (0.12)
mother's ses ^a	0.11 (0.32)	0.36 (0.24)	0.33 (0.22)	0.44 (0.27)	0.01 (0.30)	0.25 (0.18)
father's ses ^a	1.65* (0.27)	0.98* (0.24)	1.06* (0.21)	0.82* (0.26)	0.92* (0.30)	1.07* (0.18)
female	0.05 (0.57)	-0.04 (0.48)	-0.26 (0.43)	-1.49* (0.52)	-1.98* (0.65)	-0.58 (0.36)
not born in Canada	0.18 (0.62)	-0.84 (1.15)	-0.53 (1.29)	-1.17 (1.31)	-1.94* (0.85)	-1.06 (0.86)
language problem	1.06 (0.75)	-2.00 (1.02)	-2.80* (1.30)	-2.34* (1.01)	-1.94* (1.06)	-2.07* (0.90)
native	-2.64 (2.47)	-3.45* (0.99)	-4.44* (1.01)	-4.32* (1.04)	-4.48* (0.80)	-3.98* (0.76)
minority	-1.40 (0.79)	-1.04 (1.15)	-1.61 (1.00)	-0.91 (1.14)	-0.72 (0.86)	-1.16 (0.69)

^a ses = socioeconomic status

From the quantile regression results, we find out several important differences. The effects of the factor *number of parents* are positive significant in the middle of

the conditional distribution of mathematical achievement changes, i.e., 25%, 50% and 75% quantiles. However, the effects are insignificant at both the lower and upper end of the distribution, which suggests that those whose performances are in the middle of the conditional distribution appear to benefit from living with their parents whereas those whose performances are at the lower and upper end of the conditional distribution do not. Furthermore, those who are living with parents are superior to those who are only living with single parent. The factors *female*, *not born in Canada* and *language problem* are significantly negative at the top of the conditional distribution of mathematical achievement changes; however, they are insignificant at the lower end of the distribution.

In general, *father's socioeconomic status* plays a more important role in the student's mathematical achievement than *mother's socioeconomic status* does. Four family factors, i.e., *number of parents*, *father's socioeconomic status*, *language problem* and *native*, play prominent parts in mathematical achievement in grade 10. The first two factors have positive effects on mathematical achievement whereas the last two factors have negative ones.

Notice that the estimated effects of *native* on mathematical achievement are -3.45 , -4.44 , -4.32 and -4.48 at 25%, 50%, 75% and 95%, respectively, which suggests that the aboriginal inhabitants perform worst in mathematics test in grade 10.

As for the factor *number of parents*, ordinary least squares underestimates the magnitude of these effects at the 25% and 50% quantiles, but overestimates the magnitude of these effects at the 75%. For the factor *father's socioeconomic status*, ordinary least squares underestimates the magnitude of these effects at the 5% quantiles, but overestimates the magnitude of these effects at all other quantiles. In contrast, it overestimates the magnitude of effects for the factor *native* at the 25% and underestimates the magnitude of the effects at all other quantiles.

The results for the *number of siblings*, the *mother's socioeconomic status* and the *minority* tend to have little effect on mathematical achievement in grade 10.

4.2 The estimated effects in grade 11

Table 3 reports the quantile regression and OLS results conditional on family background factors in Grade 11. The estimated standard errors are reported in parentheses. It is obvious that ordinary least squares estimates are quite similar to the median (50%-quantile) regression estimates and that the effects of *number of parents*, *father's socioeconomic status* and *native* are significantly positive. However, the effects of both *language problem* and *minority* are significantly negative. It suggests that the aboriginal inhabitants perform best in mathematics test of grade 11. Language problem is a disadvantage factor to those whose native

Table 3: Comparison of quantile regression and OLS results conditional on student background variables in grade 10

Description	Quantile regression results					OLS
	5%	25%	50%	75%	95%	
number of parents	2.87* (0.90)	4.21* (0.66)	1.02 (0.72)	-0.84 (0.62)	0.45 (0.93)	1.57* (0.51)
number of siblings	-0.02 (0.20)	-0.15 (0.12)	-0.07* (0.02)	-0.28 (0.18)	-0.13 (0.25)	-0.08 (0.14)
mother's ses ^a	-0.36 (0.35)	0.03 (0.24)	0.24* (0.03)	0.02 (0.30)	0.61 (0.40)	0.14 (0.17)
father's ses ^a	1.67* (0.37)	1.60* (0.26)	1.41* (0.03)	1.59* (0.29)	0.65 (0.40)	1.41* (0.17)
female	-0.24 (0.81)	0.23 (0.53)	-0.47* (0.05)	-0.98 (0.61)	-0.87 (0.78)	-0.46 (0.34)
not born in Canada	1.65 (0.96)	-0.16 (1.42)	1.25* (0.17)	2.23 (1.77)	-0.60 (1.38)	1.23 (0.80)
language problem	-4.98 (2.54)	-2.01 (1.09)	-1.67* (0.15)	-3.76* (1.40)	-3.36* (1.16)	-1.85* (0.85)
native	2.54* (1.13)	1.33 (0.91)	1.99* (0.25)	3.58* (1.04)	1.21 (2.21)	1.91* (0.70)
minority	2.34* (0.98)	-0.75 (0.71)	-1.55* (0.13)	-1.26 (0.82)	-1.08 (1.51)	-1.43* (0.65)

^a ses = socioeconomic status

language is not English.

Note the quantile regression results. The differentials among parents, single parent and no parents in the mathematical achievement at the 5% and 25% quantiles are 2.87 and 4.21, respectively. That is, holding all other factors equal, the 5% quantile of mathematical achievement for a student living with a single parent in grade 11 is 2.87 above those who live without parents, but below those who live with their parents. In the case of the factor *number of parents*, ordinary least squares underestimates the magnitude of these effects at 5% and 25% quantiles.

Here the implication is that any factors resulting in the reduction of the number of parents must be responsible for the bad mathematical achievement of a senior high school student.

Table 4: Comparison of quantile regression and OLS results conditional on student background variables in grade 12

Description	Quantile regression results					OLS
	5%	25%	50%	75%	95%	
number of parents	4.08* (0.65)	5.49 * (0.70)	2.20* (0.64)	-0.99 (0.60)	0.01 (0.45)	2.51* (0.47)
number of siblings	-0.73* (0.19)	-0.44* (0.13)	-0.25* (0.06)	0.03 (0.25)	0.08 (0.22)	-0.21 (0.14)
mother's ses ^a	0.03 (0.49)	0.14 (0.15)	0.10* (0.04)	-0.18 (0.28)	-0.15 (0.27)	-0.04 (0.16)
father's ses ^a	0.89 (0.45)	0.95* (0.15)	1.01* (0.03)	0.87* (0.28)	0.27 (0.24)	0.64* (0.16)
female	0.48 (0.97)	0.90* (0.30)	0.95* (0.07)	1.25* (0.56)	0.57 (0.73)	0.91* (0.32)
not born in Canada	2.02 (1.22)	2.99* (0.33)	2.45* (0.19)	1.72 (1.05)	0.88 (1.78)	2.62* 0.77
language problem	0.48 (1.88)	-2.34* (0.75)	-2.15* (0.10)	-2.33* (1.18)	-2.12 (2.61)	-2.21* (0.81)
native	-3.44 (2.51)	-1.93 (1.51)	-1.75* (0.25)	-1.57 (1.26)	0.93 (1.95)	-1.77* (0.67)
minority	-2.31 (1.70)	-0.81 (0.47)	-0.61* (0.12)	-0.76 (1.00)	1.26 (1.63)	-1.01 (0.62)

^a ses = socioeconomic status

Family factors *number of siblings*, *mother's socioeconomic status*, *female* and *not born in Canada* have significant effects on mathematical achievement at 50% quantile (median) regression estimates and all these factors have very little effect on other quantiles.

Indeed, the *language problem* is still a serious problem in obtaining satisfactory score of mathematics for those who was not born in Canada and his/her native language is not English. When in grade 11, all the family factors have different effects on various quantiles.

It is worthwhile mentioning that the factor *native* performs very well, and the *father's socioeconomic status* follows it up. The factor *minority*, on the other hand, has a significantly positive effect on the 5% quantile regression estimates, that is, 2.34, which is far away from other effects at various quantiles. From

which, we can see that quantile regression estimators may be more efficient than least squares estimators.

4.3 The estimated effects in grade 12

Table 4 provides the quantile regression and OLS results conditional on family background factors in Grade 12.

Ordinary least squares estimates are also similar to the median (50%-quantile) regression estimates. The effect of factor *not born in Canada* is the highest and that of the factor of *number of parents* follows it, and the next is *father's socioeconomic status*. The factor *language problem*, i.e., English is not his or her native language, is still a disadvantage factor to mathematical achievement and the factor *native* follows it.

The parents-differentials are evident at 5%, 25% and 50% quantiles, i.e., 4.08, 5.49 and 2.20, respectively. Here the ordinary least squares underestimates the magnitude of these effects at the 5% and 25% quantiles. The factor *number of siblings* has significantly negative effects at 5%, 25% and 50% quantiles, i.e., -0.73 , -0.44 and -0.25 , respectively, which suggests that at 5%, 25% and 50% quantiles, the more siblings one has, the worse mathematical achievement he gets. The factor of *father's socioeconomic status* is still important to 25%, 50% and 75% quantiles and so is the factor of *female*. Both *native* and *minority* are significantly negative at the median.

5. Confidence Intervals and Related Interpretations

In Figure 1 we present a visual summary of confidence intervals for the effects of family background factors. Each plot depicts one of the 27 coefficients in the quantile regression model for three years (from Grade 10 to Grade 12). The solid line with filled dots (marked by the capital letter "E") represents the 5 points estimates of the coefficient for quantile p ranging from 0.05 to 0.95. Two dashed lines with filled dots marked by the capital letter "U" and "L", respectively, consist the lower and upper confidence bound. The area between the lower and upper confidence bound is a 90% pointwise confidence band. The horizontal dotted line with filled dots marked astrast "*" indicates the ordinary least squares estimates of the mean effects.

Which is the best? Two Parents, single parent or no parent ?

For years, researchers have wondered whether children from single parent family can perform like kids whose parents stay together. The first row of Figure 1 fuels the debate. In our study, results show that the effect of the factor *two parents* is superior to that of *single parent* and so is that of single parent to that of no

parent for the three years (from 10th Grade to 12th Grade). Obviously, the family factor tends to have little effect on mathematical achievement with generally insignificant coefficients at 95% quantile for all the three years. The effect of the factor *number of parents* reaches its maximum value at 25% quantile and then decreases monotonically from 25% quantile to 75% quantile. The phenomena may be interpreted as the dependency of the children on their parents is gradually decreasing.

Why should we care about SIBLING relationship?

Sibling relationship is one of humanity's oldest problems. Both in Eastern and Western cultures, "brotherly love" is assumed to be the "ideal" type of love. In fact, the natural condition is the opposite one. It is particularly difficult for parents to get their children to love one another as brothers. Does the *number of sibling* contribute to the mathematics attainment? Answer to this question is negative. The second row of Figure 1 confirmed this point. Results presented in Figure 1 show that the *number of sibling* seems to have little effects on mathematical achievement when in grade 10, because all coefficients of the *number of sibling* at the five quantiles (5%, 25%, 50%, 75% and 95%) are insignificant. The results in grade 11 are qualitatively similar to those in grade 10 for the same variable *number of sibling*, except for the effect at median in grade 11, which is -0.07 . Until grade 12, the negative effects of *number of sibling* at the three lower tail quantiles (5%, 25% and 50%) are significant, i.e., -0.73 , -0.44 and -0.25 . In short, the negative role of the *number of siblings* becomes more obvious with the growth of grade.

What's the influential difference between mother and father?

For any specific quantile we may want to know the different influence of socioeconomic status between mother and father on the mathematical achievement. The third and fourth row of Figure 1 answer this question. Obviously, father's influence is superior to mother's in senior high school. In grade 10, *mother's socioeconomic status* has little effect on mathematics attainment with generally insignificant coefficients across the various quantiles. In contrast, the effects of the *father's socioeconomic status* are all positive and statistically significant across various quantiles. The dominant position goes straight on until the last year of high senior high school.

Is there a gender difference?

Our research on the senior high school presented in the fifth row of Figure 1 shows that male and female have differences in mathematical achievement. Generally speaking, the female lags behind the male in the first two academic years, but perform better in the last academic year of senior high school. The results

seem to be surprising. Traditionally, females have found mathematical achievement elusive in senior high school although their mathematic achievement in the elementary grades is equal to male's.

Where is the performance gap?

In this study, we examined academic mathematical achievement of the senior high school students and gauged the performance gaps relating to those *not born in Canada* and *born in Canada*. The results given in the sixth row of Figure 1 indicate that *not-born-in-Canada*-children's math achievement seems to be lower than the others only in grade 10 but better in the last two academic years. The effects of *not born in Canada* in grade 12, for example, are 2.99 and 2.45 at 25% and 50% quantiles, respectively, which are statistically significant.

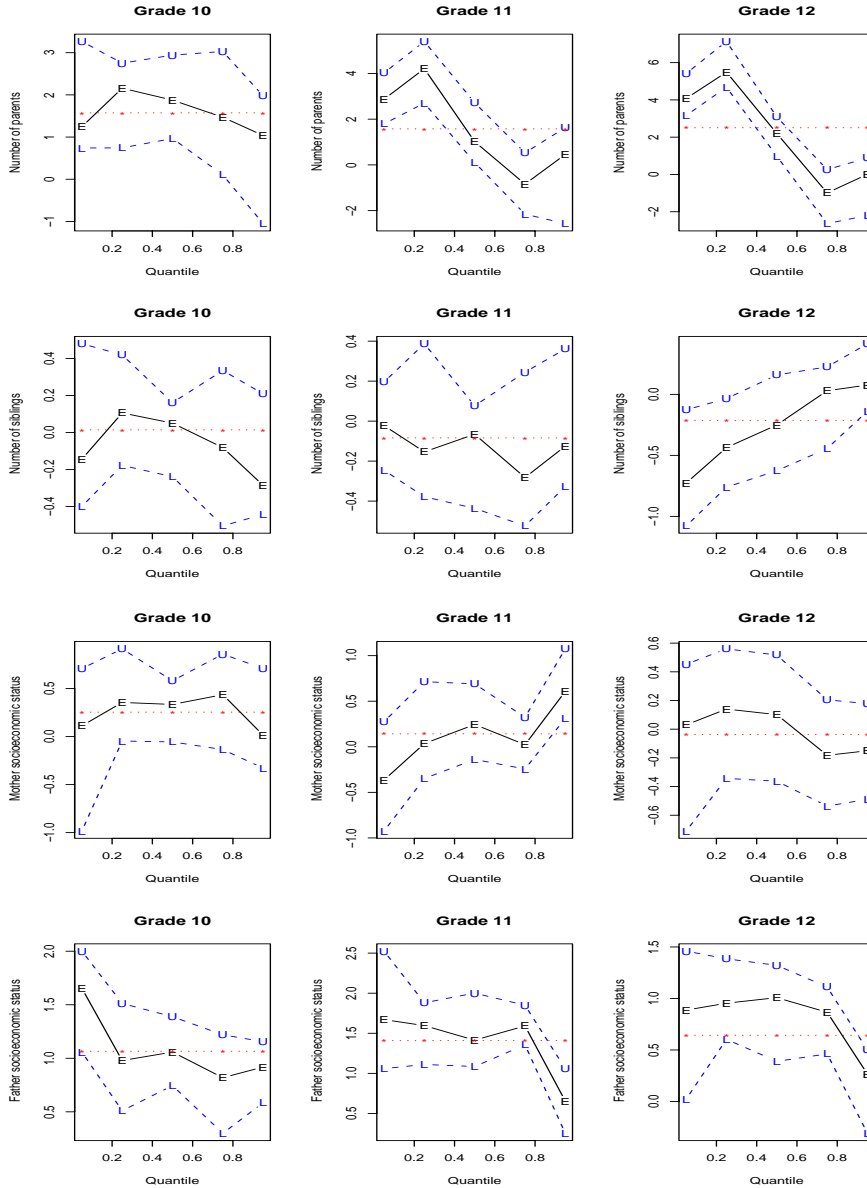
Is the language problem a serious problem?

The results presented in the seventh row of Figure 1 are as follows: At the middle and upper parts of the conditional distribution, the coefficients are all statistically significant negative, i.e., -2.80 , -2.34 and -1.94 . in grade 10; The coefficients of *language problem* are still all statistically significant negative, i.e., -1.67 , -3.76 and -3.36 in grade 11; In the last year of senior high school, the effects of factor *language problem* are -2.34 , -2.15 and -2.33 at 25%, 50% and 75%, respectively. All these suggest that limited English proficiency handicaps *not-born-in-Canada* -children's mathematical achievements. In fact, language barriers are often more detrimental for those who were not born in Canada. Due to the socially and linguistically isolated communities, poor *not-born-in-Canada* children can hardly improve their new language skills and the language barriers persist over the school years until the whole senior high school.

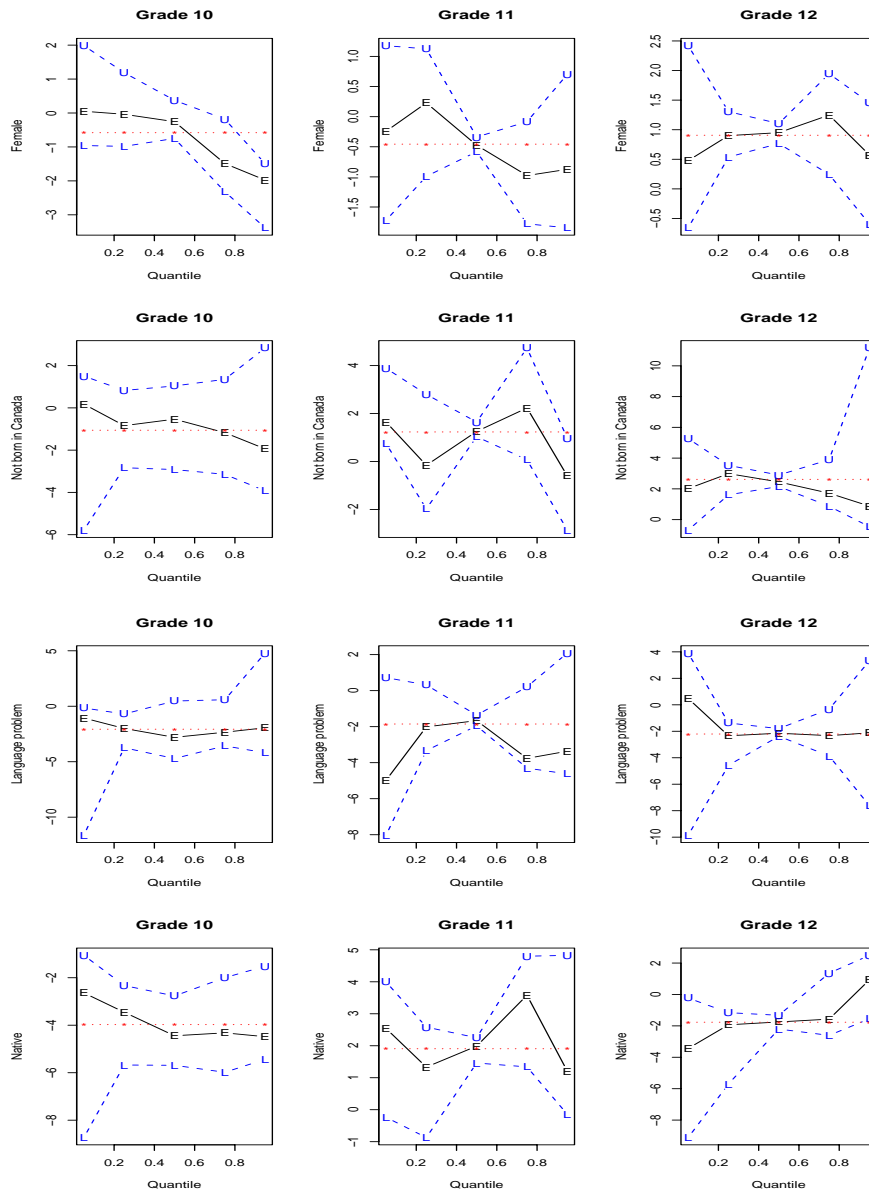
Do native students benefit from mathematical education?

Taking the research results of *native* mathematical achievements listed in the eighth row of Figure 1, we can easily find that, in grade 10, the effects of *native* at 25%, 50%, 75% and 95% are -3.45 , -4.44 , -4.32 and -4.48 , respectively. And the performance of *native* student in grade 12 appears not to be better than that in grade 10, Fortunately, the performance of *native* student in grade 11 is improved, i.e., the coefficients are 2.54, 1.33, 1.99, 3.58 and 1.21 at 5%, 25%, 50%, 75% and 95% quantiles, respectively, where 2.54, 1.99 and 3.58 are statistically significant. The finding shows that a Western approach to mathematics education directly conflicts with *native* ways of knowing and learning mathematics. It is said that *native* children are most often taught complex tasks through lengthy observation followed by practice in private, rather than through the "trial and error" approach of traditional classrooms.

At last, it is evident that the performance of students in senior high school mathematical achievements is bad with the exception of the effect 2.34 at 5% in grade 11.



Continued to next page



Continued to next page

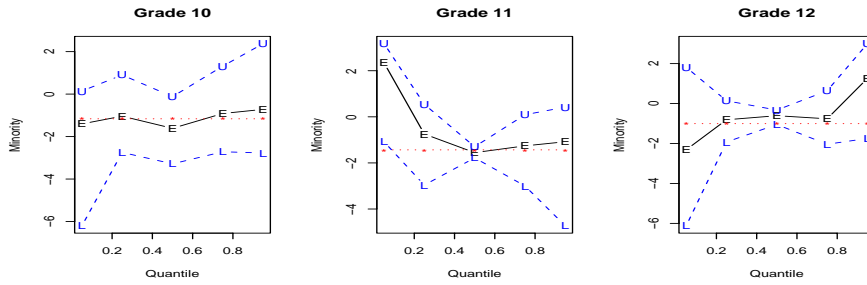


Figure 1: OLS and quantile regression estimates for mathematical achievement model. The solid line with filled dots (marked by the capital letter “E”) represents the 5 points estimates of the effects of “intercept” for quantile p ranging from 0.05 to 0.95. Two dashed lines with filled dots marked by the capital letter “U” and “L”, respectively, consist the lower and upper confidence bound. The area between the lower and upper confidence bound is a 90% pointwise confidence band. The horizontal dotted line with filled dots marked astrak “*” indicates the ordinary least squares estimates of the mean effects and is also superimposed

6. Conclusion

A great number of previous studies focused only on mean performance of the family background factors. It seems rather implausible that such factor effects should all act so as to shift the entire distribution of test results by a fixed amount. It is of obvious interest to know whether the family background factors alter performance of the strongest students in the same way that weaker students are affected. Such questions have been investigated in this paper by the means of quantile regression approach. Our quantile regression results suggest that there may be differential family background factor effects at different points in mathematical achievement conditional distribution.

Results show that the number-of-parents factor tends to have significant effects on attainment when moving from 5% quantile to 50% quantile in the mathematical achievement conditional distribution. The finding suggests that the rising divorce problem leading to the changing of the number of parents would yield disastrous results in mathematical achievement from 5% quantile to 50% quantile in the conditional distribution. Generally speaking, the number of siblings has negative effects on the mathematical achievement of student that are at the lower and median quantiles of the last two years in senior high school. We have also noticed the fact that the effect of father’s socioeconomic status on the student’s mathematical achievement is superior to that of mother all the three years in se-

nior high school. Roughly saying, the female lags behind the male in the first two years in senior high school but perform better in the last one year. The factor *not born in Canada* has effects on the mathematics almost in the same way as that the factor *female* does. Indeed, the language barrier is a serious problem. natives students benefit from the current senior high school mathematics education only in grade 11. At last, the evidence shows that minority student performance bad with the exception of the effect at 5% quantile in grade 11.

Acknowledgments

The author gratefully acknowledges Prof. Min-Te Chao, the editor of this journal, for his constructive comments, suggestions and encouragement. He is also grateful to Canadian research chair Xin Ma for his real data, to McDonald, Linda M., Professor and Chair for her help, and to participants in Roundtable 27, American Educational Research Association 2004 Annual Meeting.

Part of the work was performed while the author was a postdoctoral fellow at Canadian Center for Advanced Studies of National Databases, University of Alberta, Canada.

References

- Bailar, B. (1991) Salary survey of U.S. colleges and universities offering degrees in statistics. *Amstat News* **182**, 3-10.
- Corman, H. and Chaikind, S. (1998). The effect of low birthweight on the school performance and behavior of school-aged children. *Economics of Education Review* **17**, 307-316.
- Eide, E. and Showalter, M. H. (1998). The effect of school quality on student performance: A quantile regression approach. *Economics Letters* **58**, 345-350.
- Ehrenberg, R. G. and Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of Education Review* **13**, 1-17.
- Ehrenberg, R. G. and Brewer, D. J. (1995). Did teachers' verbal ability and race matter in 1960s? Coleman revisited. *Economics of Education Review* **14**, 1-21.
- Fennema, E., and Sherman, J. (1978). Sex related differences in mathematics achievement and related factors: A further study. *Journal for Research in Mathematics, Education* **9**, 189-203.
- Hao, L. and Portes, (1998). Bilingualism and loss of language in the second generation. *Sociology of Education* **71**, 4, 269-294.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica* **47**, 153-161.

- Koenker, R. and Bassett, G. S. (1978) Regression quantiles. *Econometrica* **46**, 33-50.
- Lapin, L. (1998). Study: immigrants' kids learn better. *The Sacramento Bee* **23**, 1-1.
- Levin, J. (2001). For whom the reductions count: A quantile regression analysis of class size and peer effects on scholastic achievement. *Empirical Economics* **26**, 221-246.
- Lockhead, M., Thorpe, M., Brooks-Gunn, J., Casserly, P., and McAloon, A. (1985). *Understanding Sex-ethnic Differences in Mathematics, Science and Computer Science for Students in Grades Four to Eight*. Educational Testing Service.
- Mathews, W. (1983). Influences on the learning and participation of minorities in mathematics. A Report from the Postdoctoral Program, Program Report 83-85. Madison, WI: Wisconsin Center for Educational Research.
- Taylor, B. R. (1983). "Equity in mathematics: a Case study." *Mathematics Teacher* **76**, 12-17.
- Vernez, G. and Abrahamse, A. (1996). *How Immigrants Fare in U.S. Education*. RAND Center for Research on Immigration Policy.
- Viadero, D. (1997). Immigrant children succeed despite barriers, report says. *Education Week* **17**, 14-14.
- Yu, K. and Moyeed, R. A. (2001) Bayesian quantile regression. *Statist. Probab. Letters* **54**, 437-447.

Received March 25, 2005; accepted December 6, 2005.

Maozai Tian
School of Statistics
Renmin University of China
Beijing, 100872, China
mztian@ruc.edu.cn