Available online at www.sciencedirect.com
SciVerse ScienceDirect
IERI Procedia 2 (2012) 684 – 689
Procedia IERI
www.elsevier.com/locate/procedia

2012 International Conference on Future Computer Supported Education

Application Research on Rank Return Method in Mathematics Achievement Appraisal

Yu Wang[a]*, Xiaoqin Zhang[b], Dianjun Lu[a]

[a]*Department of Mathematics of Qinghai Normal University, Xining, 810008, P. R. China*
[b]*Financial and Economic College of Qinghai University, Xining, 810008, P. R. China*

Abstract

Compared with ordinary least squares, quantile regression can more fully reflect that the dependent variable has different effects in the different parts of the distribution of the independent variables, and has a very wide range of applications. The paper makes a brief introduction to the idea of quantile regression, applying the methods into mathematics achievements achievement and having a comparative analysis of the good or bad results about quantile regression and ordinary least squares under two kinds of external pressures to mathematics achievement.

Keywords: Quantile regression; mathematics achievements; external pressure; ordinary least squares

1. introduction

Taking full advantage of the quantile regression, several recent studies have modeled the performance of public school students on standardized examinations as a function of socio-economic characteristics like parents' income and educational attainment, and policy variables like class size, school expenditures and teacher qualifications. For example, Eide and Showalte[1,2] used quantile regression to estimate whether the relation between school and performance on the standardized tests differs at different points under the condition of test score gains. We have known that among the earliest and certainly one of the most controversial studies of the effects of schooling inputs including class size on scholastic achievement was Coleman Rrport[3]. Hanushek[4,5] studied the effects of various inputs in the production of public schooling.

* Corresponding author. Tel.: +86-971-6306443
E-mail address: wangyu@qhnu.edu.cn.

 doi:10.1016/j.ieri.2012.06.154

The finding of his study is that the effects of class size reduction on achievement is ambiguous, wavering to negative depending on the study Levin[6] addressed the controversial topic of the class size. In his paper, he wrote: The “conventional wisdom” that class size reduction is a viable to increase scholastic achievement is discounted. Rather, the results point towards a far stronger peer effects through which class size reduction may an important role.

Indeed, much of the previous research has emphasized student’s economic background. In the available literature, they addressed the question: “Does money matter?” but they also asked the question: “For whom does money matter?” To my knowledge that relatively little attention has been paid, however, to the external pressure factors, such as parent’s push and peer’s push on achievements. In fact, there are two conditions in process of the growth of a student: subjective and objective conditions. And the external pressure factors play very important roles in the objective condition. This paper studies the influence of the external pressure factors on the student mathematics achievements.

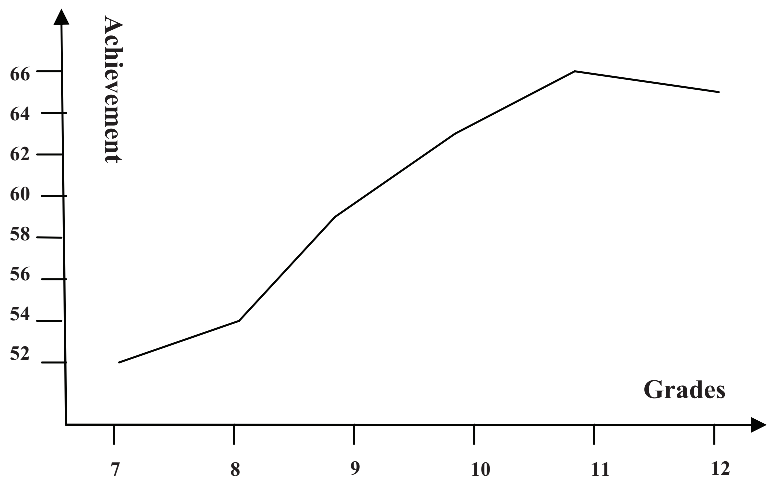
2. Methodology

2.1. Descriptive Analysis

The description for the variables used in the empirical analysis is as follows: MATHIRTR denotes mathematical achievement, KMHPH peer’s push on mathematical achievement.

Descriptive statistics of mathematics achievements and the corresponding plots are shown in Figure 1. The figure clearly reveals the tendency of the average scores from Grade 7 to Grade 12. The significant changes in average achievements occurred at grade 8 and 11 for both courses: mathematics. As can be seen from the slopes that change sharply. Furthermore, significant change also occurred at grade 9 for science. From this point of view, grade 8 and 11 are crucial to the high school students in their mathematics achievements.

Mathematics achievements for six grades



It is well known than quantile regression, as introduced by Koenker and Bassett[7], has found many applications in longitudinal studies because of its useful features: (1)given predictors, it characterizes the entire condition distribution of a response variable; (2)both the recent advances in computing resources and the ready availability of linear programming algorithms make the estimation easy; (3)the resulting estimated coefficients are robust; (4)quantile regression estimation may be more efficient than those from least squares in the case that the error term is non-normal. Details referred to Buchinsky[8].

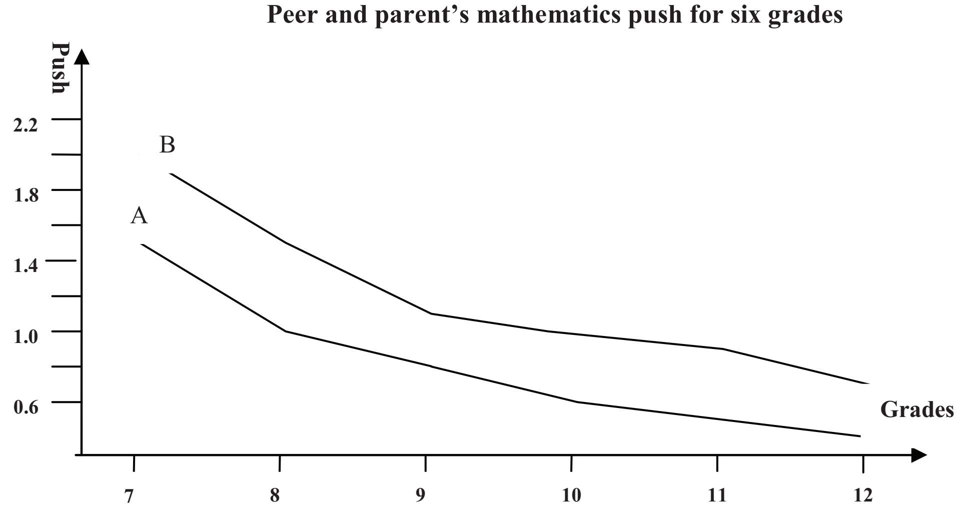
2.2. Peer and parents' push

Peer is a person who has equal standing with another or others, as in rank, class, or age. It is often said than children are easily influenced by their peers. Here kinds of peer and parent pushes are shown in Figure 2 (upper part). Specifically, the solid line with filled dots marked by the capital letter "A" and the dashed line with filled dots marked by the capital letter "B" represent peer and parent pushes to students' mathematics achievement, respectively. This finding shows that peers and parents have an effect on improving mathematics achievements.

Consistently, for the whole high school and that all the push –curves decrease monotonically across all the 6 grades in high school.

This finding that peer and parent push on high school student is also becoming weaker and weaker.

Figure 2 Peer and parent push on mathematics achievements



2.3. Quantile regression Methodology

Quantile regression, as introduced by Koenker and Bassett[7] is gradually developing into a comprehensive approach to the statistical analysis of linear and non-linear response model. There at least four equivalent mathematical definitions of quantile regression:

Definition based on the conditional quantile function

Let $s_p(x)$ be the p-th quantile of the dependent variable η given $\xi = x$. In this case $s_p(x)$ can be found by solving $Q(s_p(x)|x) = P(\eta \le s_p(x)|\xi = x) = p$, where Q is the cumulative distribution of η ;

Definition based on the quantile regression model (Bailar[9]) $\eta = \xi^{T}\theta + \delta$, where the error term ε is assumed to satisfy Quantile $_p(\varepsilon) = 0$. In standard linear regression model, the error term is assumed to be a Guassian error;

Definition based on a check function (Koenker and Bassett[7]) $\min_{\theta \in \Theta} E\{\rho_p(\eta - \xi'\theta)|\xi = x\}$, where $\tau_A(z)$ the usual indicator function of the set A, Θ is a parametric space for θ , $\rho_p(z) = pz\tau_{(0,\infty)}(z) - (1-p)\tau_{(-\infty,o)}(z)$ is called check function;

Definition based on asymmetric Laplace density (Yu and Moyeed[10]) $\varphi(\varepsilon) \propto \exp\left\{-\sum_{i=1}^{n}\rho_p\left(y_i - x_i^T\theta\right)\right\}$, where $\varphi(\varepsilon)$ is the probability density of the model error ε .

In this paper, we use the definition based on Koenker and Bassett's definition, in which the dependent variable η is the mathematics and science s achievement and the independent variable ξ is the external expressive variables mentioned before.

2.4. Double kernel approach

We have seen that the condition is a vital ingredient for quantile regression. However, there are some problems associated with the existing kernel-weighting estimation of the condition distribution .For example, some estimators of the condition distribution are not a distribution function see Stone [11] for example, the others are dissatisfactory for the quantile curve based on these estimators may cross one another, which is of course absurd, see Hall et al [12] for example. To avoid these problems, we employ the "double kernel" approach of Yu and Jones [13] in this paper.

We know that the condition distribution is a vital ingredient for quantile regression. Yu and Jones [13] and Hall et al [12] have recently considered several methods for estimating conditional distribution. In this article, we employ the local linear double-kernel smoothing method proposed by Yu and Jones [13]. Specifically, suppose that(ξ_1 , η_1), $\cdots$,(ξ_n , η_2) is a set of independent observations from some underlying distribution $Q(x, y)$ with density $\varphi(x, y)$ and interest centers on the responses η_i considered to be realizations from the condition distribution $Q(y|x)$ or density $\varphi(y|x)$ of η given $\xi = x$.

Define $\hat{Q}_{l_1,l_2}(y|x) = \hat{c}$, where $(\hat{c}, \hat{d}) = \arg\min \sum_i \left(A\left(\frac{y - Y_i}{h_2}\right) - c - d(\xi_i - x)^2 \right) \times B\left(\frac{x - \xi_i}{h_1}\right)$, where l_1 and l_2 are the bandwidth in x and y directions, respectively. The functions B and A are two kernel functions.

Define $\hat{q}_p(x)$ to satisfy $\hat{Q}_{h_1,h_2}(\hat{q}_p(x)|x) = p$ so that $\hat{q}_p(x) = \hat{Q}^{-1}_{l_1,l_2}(p|x)$.

The important issue with the kernel fitting approach is the bandwidth selection. There are several different ways to select the bandwidth in the x direction. Here one rule for it simply modifies the bandwidth l_{mean} that would be used for mean regression and can be implemented as follows:

(1) Employing the Ruppert, Sheater and Wand's[14] technique to obtain l_{mean} . The technique is based on the asymptotic mean square error (AMSE) together with the "plug in" rule to replace any unknown quantity in the AMSE by its estimator.

(2) Calculate $l_p = l_{mean}\left[p(1-p)/g\{F^{-1}(p)\}^2 \right]$, where g and F are the standard normal density and distributions.

Similarly, from minimizing the AMSE of estimator over the bandwidth d_p in the y direction, the d_p can be chosen according to $\frac{d_p l_p^3}{d_{1/2} l^3_{1/2}} = \frac{\sqrt{2\pi} g\left(F^{-1}(p)\right)}{2\{(1-p)\tau(p \ge 1/2) + p(p < 1/2)\}}$, where $d_{1/2}$ is taken to be $l_{1/2}$ and $\tau(\cdot)$ is a ordinary indicator function, For further details see Yu and Jones[13].

3. Double-Kernel Quantile Regression on Mathematics Achievement

3.1. Parent's push and children's mathematics achievements

We'll further consider various external pressure variables which include MATHIRTR, PMHPH, KMHPH in this section.

Firstly, we consider the relationship between the mathematics achievement and the parents' push on mathematics. The dependent variable is the mathematics achievement and covariate is the parent mathematics push.

In Table 1 we present the double-kernel quantile regression results which were estimated at five different quantiles (5 %, 25%, 50%, 75% and 95%). It is clearly that the parent-math-push value 2 is the turning point.

Table 1 Double-kernel regression results for MATHIRTR-PMHPH

Quantile	Parents' push on mathematics achievement 0	1	2	3
5%	33.84948	35.40249	36.76398	34.29599
25%	42.79999	47.15748	48.98999	44.04999
50%	53.07489	57.35996	62.8998	52.26999
75%	62.90245	66.99841	71.32489	60.32895
95%	80.08475	83.09878	86.9889	70.09983

3.2. Peer's push and student's mathematics achievement

Next, we'll consider the relationship between the mathematics achievement MATHIRTR and the peer math push KMHPH. The dependent variable is the mathematics achievement MATHIRTR and covariate is now the peer math push KMHPH. The double-kernel regression results are reported in Table 2

Table 2 Double-kernel regression results for MATHIRTR-KMHPH

Quantile	**Peer's push on mathematics 0**	**1**	**2**	**3**	**4**
5%	32.26298	36.22648	36.25499	35.84499	33.64597
25%	45.34998	47.15748	45.81000	43.46498	41.04998
50%	57.38998	59.46997	54.88998	52.26999	50.38998
75%	68.90249	70.41998	66.23499	64.32498	60.79498
95%	83.15847	86.29696	82.81000	82.39996	79.63298

Table 2 depicts the double-kernel quantile regression results which were made at five different quantiles (5%, 25%, 50%, 75% and 95%). As for the estimated curve with quantile 0.05, the lowest one in our analysis, the optimal value of the peer math push is 2, but for the rest four quantile regression curves, the optimal value is 1.The finding implies that the student whose mathematics achievement is at the bottom of his class needs more peer's math-push if he wants to get the maximum mathematical achievement than those who do better in mathematics.

References

[1] Eide E, Showalter M H, The effect of school quality on student performance: A quantile Regression approach, Economics Letters, 1998, 58: 345-350.
[2] Su Y, Wan Y, The Idea and Application of Quantile Regression, Statistical Thinktank, 2009, 10: 58-62.
[3] Coleman J, Equality of educational opportunity, US Department of Health, Education, and welfare, Office of Education, Washington D.C., 1966.
[4] Hanushek E, The economics of schooling: Production and efficiency in public schools, Journal of Economic Literature, 1986, 24: 1141-1177.
[5] Hanusshek E, School resources and students performance, In: Burtless, G.(Ed), Does Money Matter? The Effect of school Resources on Student Achievement and Adult Succcess, Brookings Institution, Washington, D.C., 1996, 43-73.
[6] Levin J, For whom the reductions count: A quantile regression analysis of class size and peer effects on scholastic achievement, Empirical Economics, 2001, 26: 221-246
[7] Koenker R, Regression quantiles, Econnometrica, 1978, 46: 33-50.
[8] Buchinsky M, Recent advances in quantile regression model, J. Human Res., 1998, 38: 88-126.
[9] Bailar B, Salary survey of U. S. colleges and universities offering degrees in statistics, Amstat News, 1991, 182: 3-10.
[10] Yu K, Moyeed R A, Bayian quantile regression, Statist. Probab. Lett., 2001, 54: 437-447.
[11] Stone C J, Consistent nonparametric regression, The Annals of Statistics, 1977, 4: 595-645.
[12] Hall P, Wolff R, Yao C L, Methods for estimating a conditional distribution, J. Am. Statist. Assoc., 1999, 94: 154-163.
[13] Yu K, Jones M C, Local linear regression quantile estimation, J. Am. Statist. Ass., 1998, 93: 228-238.
[14] Ruppert D, Sheather S J, Wand M P, An effective bandwidth selector for local least squares regression, J. Am. Statist. Assoc., 1995, 90: 1257-1270.