Estimating the Extreme Behaviors of Students Performance using Quantile Regression—Evidences from Taiwan

Sheng-Tung Chen
Assistant Professor, Department of Public Finance,
Feng Chia University, Taichung, Taiwan

Hsiao-I Kuo
Assistant Professor, Department of Senior Citizen Service Management,
Chaoyang University of Technology, Taichung, Taiwan

Chi-Chung Chen*
Professor, Department of Applied Economics,
National Chung-Hsing University, Taichung, Taiwan
Email: mayjune@nchu.edu.tw
Phone: (886) 4-22858137
Fax: (886) 4-22860255

*Corresponding author.
Estimating the Extreme Behaviors of Students Performance using Quantile Regression—Evidences from Taiwan

Abstract

The two-stage least squares approach together with quantile regression analysis are adopted here to estimate the educational production function. Such a methodology is able to capture the extreme behaviors of the two tails of students’ performance and the estimation outcomes have important policy implications. Our empirical study is applied to the case of students’ scores in the Basic Competence Test in Taiwan. The empirical estimation outcomes between traditional OLS and quantile regression on peer-group effects, school characteristics, and family characteristics are diversity and depend on students’ ability. Such findings have important implications for parents as well as for government.

Key words: Score of Basic Competence Test, Peer-Group Effects, Quantile Regression

JEL Code: C13, C21, I21
Estimating the Extreme Behaviors of Students Performance using Quantile Regression—Evidences from Taiwan

1 Introduction

The estimation and measurement of students’ performance has received much attention in the literature in recent years. Empirical studies have found that schools, peer group effects, and individual and family background characteristics are the three factors to affect a student’s performance. For instance, the effects of the schools in terms of private versus public schools on a student’s performance have been examined using two-stage Probit least squares by Goldhaber (1996), Sander (1999) and Stevans and Sessions (2000), where the selection of a private or a public school is an endogenous dummy variable. On the other hand, peer group effects play an important role in a student’s achievement (Deller and Rudnicki, 1993; Caldas and Bankston, 1997; McEwan, 2003). They found that if family income and parents’ education are used as a group index, students with an above-average index in one group will perform better than those in the other group.

The characteristics of an individual’s family background, which include family income, the number of children, the mother’s full-time employment, the father’s education level, the student’s study, and leisure time, are highly correlated with the student’s achievement (Deller and Rudnicki, 1993; Caldas and Bankston, 1997; Sander, 1999; Okpala and Smith, 2001; Robertson and Symons, 2003). The above studies and literatures are all based on the mean effects to capture the students’ performance or achievements since all of the econometric models used in these papers provide the mean effects of the regressors on students’ performance. However, if we would like to observe the extreme performance (i.e. extreme highest or lowest score),
the above general regression approach will not be able to provide any information.

On other hand, some explanatory variables such as the choice of public or private school or educational expenditure in the estimating equation of students’ performance are endogenous. Either a traditional two-stage least square approach or instrumental variable approach is applied to solve this endogenous problem when estimating the equation of students’ performance to avoid inconsistent estimates. In here, both the Probit two-stage least squares (P2SLS) and quantile regression analysis are adopted to capture both the mean and quantile behaviors unbiased in this educational production function. Koenker and Hallock (2001) have indicated that quantile regression may be viewed as a natural extension of the classical least squares estimation of conditional mean models to the estimation of an ensemble of models for conditional quantile functions.

In Taiwan, according to the educational policy of multiple-channel school admission system, a project of the Ministry of Education, there have been three ways of entering senior high school since 1998. One is the selection admission system, another is making a formal application to enter, and the other involves a register to allocate students among the different senior high schools. Regardless of which approach is used to enter senior high school, the Basic Competence Test (BCT) score is the main reference criterion when it comes to the senior high school selecting students. A student has a priority to enter the higher quality senior school if he gets the higher BCT scores. The BCT includes five subjects, namely, Chinese, English, Mathematics, Society and Nature. The scores for each subject are determined by a “scale score”, which is a score to which raw scores are converted by numerical transformation (e.g., the conversion of raw scores to percentile ranks or standard scores). This score can actually show the related rank of every student in each subject.
The scale score for each subject is 60 points and the total score for the BCT is 300 points. The use of this scale score has made the competition among students seeking to enter a prestigious senior high school even tighter than before in Taiwan. Thus, the BCT score serves not only as a valuable source of reference for senior high schools, but also reflects the achievements of the students in studying.

The BCT score is very important when it comes to the senior high school selecting students in Taiwan. Students have the priority to enter the higher quality senior school if they get the higher test scores. On the contrary, the lower test scores give rise to decrease the opportunity of selecting the higher quality senior school. The major purpose of this paper, therefore, is to analyze which factors will affect the learning performance and subsequent achievements of the students based on the results of the BCT. A Probit two-stage least squares (P2SLS) was adopted to capture the mean behavior, while the quantile regression was implemented on the case in Taiwan to provide more information on the two-tails of students’ performance. The data set is based on a questionnaire sample from all high schools in Taiwan in 2005 that has been applied here to estimate the impacts of various factors on BCT scores in Taiwan. The remainder of this paper is organized as follows. Section 2 provides the literature review related with educational production function. Section 3 introduces the empirical model’s setting with the estimation methodology. Section 4 displays the empirical data with statistical analysis and the empirical results are reported and analyzed in Section 5, and Section 6 concludes with policy implications.

2 Literature Review

According to the extant education and economics literature, we find that the factors affecting a student’s achievement can be classified into three categories,
namely, school factors (Goldhaber (1996); Stevans and Sessions (2000); Sander (1999)), peer group factors (McEwan (2003); Caldas and Bankston III (1997); Schneeweis and Winter-Ebmer (2005)), and individual and family background characteristics (Deller and Rudnicki (1993); Caldas and Bankston (1997); Okpala et al. (2001); Robertson and Symons (2003); Sander (1999)).

First, as for school factors, such as the choice between public and private schools, the teacher’s devotion to teaching and the school resources devoted to studying, all of them affect the students’ achievements in studying. Goldhaber (1996) applied the survey data of the National Educational Longitudinal Study (LELS88) in 1988 to investigate the effect of choosing private and public schools on students’ grades. In addition, Stevans and Sessions (2000) used survey data from the US National Educational Longitudinal Survey in 1992 to identify the relationship between the choice of school (public or private) and the performance of students in terms of grades. They found that while White students perform marginally better in private schools compared to public schools, a performance gain for private school minority students was not realized and they concluded that “school choice” was mostly taken advantage of by White urban residents. Because the private/public choice was endogenously decided, Stevans and Sessions (2000) used a Probit model to estimate the factors that affect the choice of school and, after calculating Heckman’s inverse Mill’s ratio ($\lambda$ ), included it in the estimated equation in order solve the problem of endogeneity. This estimation procedure will also be accepted in our research to solve the problem of endogeneity arising from the private/public choice of schools.

As regards the factor of the teacher’s devotion, Deller and Rudnicki (1993) used ordinary least squares (OLS) to estimate the educational production function of
the elementary school students in Maine State. They found that the salaries of teachers had a positive impact on students’ grades. This meant that the higher the salary paid to the teacher, the higher the quality in terms of the teaching. Besides, Sander (1999) investigated the effects of expenditures per pupil and expenditure-related variables on student achievements in Illinois. By estimating third and eighth grade Illinois Goal Assessment Program (IGAP) test scores in mathematics, he found that average teacher salaries have modest positive effects on eighth grade test scores in mathematics. Stevans and Sessions (2000) also found that the teacher factor had the most significant effect on the students’ grades. In terms of the school factor, Sander (1999) found significant determinants of eighth grade scores in his estimate to include a positive college effect and negative African-American, Hispanic, low income and urban effects. He also found that students living in urban areas had low achievements in studying because of their having low income and being African-American or Hispanic. Thus, the school’s location should be considered in the estimation of the educational production function. Furthermore, we think that the result will be the opposite in Taiwan because the people living in urban areas are often those with high incomes, a situation that is much different from that in the US.

Second, the peer group factors that affect achievements include family socioeconomic composition (such as the parents’ income and level of education), the unemployment rate, as well as the adoption of an ability grouping system based on the students’ intelligence. McEwan (2003) estimated the peer effect on the achievements of eighth-grade students in Chile in 1997. He used three proxies as the peer effect factors, one being the parents’ income, another being the parents’ educational attainment and the other the extent of the parents’ involvement in the students’ activities. Caldas and Bankston III (1997) discussed the relationship between the
socioeconomic composition of the students and their scores using OLS. They used eighth-grade students in public school as the subjects. Furthermore, the socioeconomic composition was measured by the parents’ career, level of education, and the proportion of students who took advantage of the free/low price lunch. The empirical results of McEwan (2003) and Caldas and Bankston III (1997) indicated that the high economic status had a positive impact on the students’ scores. In addition, Schneeweis and Winter-Ebmer (2005) used PISA 2000 data for Austria to estimate the peer effects for 15- and 16-year-old students. They found that the quantile regressions suggested that the peer effects were asymmetric and in favor of low-ability students, meaning that students with lower skills benefited more from being exposed to clever peers, whereas those with higher skills did not seem to be affected much (Schneeweis and Winter-Ebmer (2005)). Their research tells us that the regression results can not just be based on the mean condition but should also consider different quantile distributions.

Third, as for the factors regarding individual and family background characteristics (such as family income, the parents’ educational level and careers), studies by Deller and Rudnicki (1993), Caldas and Bankston (1997), Okpala et al. (2001), Robertson and Symons (2003) and Sander (1999) have indicated that individual and family background characteristics will affect students’ achievements. The higher the level of the parents’ education and income, the higher the children’s achievements will be. The better the career that the parents have will have a positive impact on their children’s achievements. In terms of the factors based on individual characteristics, we have treated the students’ input and time distribution as the main variables. Stevans and Sessions (2000) found that the students’ input in terms of studying had a positive impact on their achievements. Borg et al. (1989) estimated the
impacts of the degree to which students studied hard on their achievements based on a sample of 195 students from the University of North Florida (UNF). They suggested that the students with high achievements could enhance their scores by spending more time on studying but the effect was not significant with low achievement ones. They explained that the students with low achievements often have low ability in studying and that their efficiency is also low in reading. Although they spend more time studying, the results they obtain are not significantly better. Caldas and Bankston III (1997) also pointed out that the time spent on a part-time job and watching television was negatively correlated with the scores obtained. However, anticipating schooling activities and doing more reading could help increase scores. In addition, they also found an interesting result in that the more time that was spent doing homework, the lower the scores the students received.

In terms of the interaction between parents and children, such as whether the student is living with parents or not, whether the mother’s career is part-time or full-time, and whether the father is alive or not. Stevans and Sessions (2000) indicated that students living with their parents could get high scores. Furthermore, Robertson and Symons (2003) pointed out that the variable indicating whether or not the mother had a full-time career and the father was dead had a significantly negative impact on the child’s achievements in studying. This implies that the interaction between parents and children is an important factor that affects the children’s achievements in studying.

To sum up, we find that the factors affecting the student’s achievements in studying can be classified into three categories, namely, school factors, peer group factors and individual and family background characteristics. Thus, we compiled the questionnaire on the above categories to estimate the production function for
education. Firstly, school factors include the choice of school and the school’s location. Secondly, the peer group factors include ability grouping, as well as the number of older siblings who are studying in national universities, colleges or high-ranking senior high schools. In this research, we believe that the students with low achievement may have obtained low scores because of their being grouped according to ability in Taiwan. Such an ability grouping system can be viewed as having a peer effect. Finally, factors related to individual and family background characteristics include the parents’ careers, level of education, family income and time spent studying and on leisure activities. In our research, we classify students’ time into five classifications, such as class time, time spent taking lessons after school, sports time, leisure time and time spent doing homework.

In recent years, some studies have estimated the production function for education through the methodology of quantile regression. Schneeweis and Winter-Ebmer (2005) used quantile regression to estimate the peer effects for 15- and 16-year-old students in Austria and found the peer effect on education to be asymmetric. Eide and Showalter (1998) applied the quantile regression model to investigate the effect of school quality on students’ achievements. Levin (2001) also applied the quantile regression model to discuss the effects of the peer effect and class size on the students’ achievements and found evidence of a far stronger peer effect through which a reduction in class size may play an important role in the students’ achievements. Therefore, the quantile regression model can explain the effect not only on the conditional mean but also on different quantile distributions. This will help us discuss the whole feature of the production function of education and obtain meaningful policy implications with regard to the upper quantile students who have higher achievements as well as the lower quantile ones that have lower
achievements.

3 The Methodology and Empirical Model

This study classifies three factors including school factors, peer group effect factors, and family and individual factors that will affect the BCT scores. The following function expresses the relationship:

\[
BCT\;\text{Score} = f\left(School\;\text{factors},\;Peer\;\text{Effect}\;\text{factors},\;family\;\text{and}\;individual\;\text{factors}\right)
\]

Such relationship between three factors and students’ performance on BCT score could be estimated through traditional OLS regression. However, the variable, namely, “the choice of private/public school”, is contained in the educational production function which is an endogeneity question. If traditional OLS is applied to estimate the above function, the endogeneity problem will arise and will lead to biased and inconsistent estimation results. Goldhaber (1996) and Stevans and Sessions (2000) solved the endogeneity problem by applying P2SLS in order to obtain an unbiased estimator.

Suppose the educational production function is shown in equation (1) which expresses all variables from the above three factors on students’ performance (i.e. score in BCT). However, to avoid the endogeneity problem from school choice, the selection of school choice (i.e. equation (2)) needs to be estimated before estimating equation (1). In other words, equation (2) is estimated using Probit regression approach. Later, the predicted value from equation (2) and the Heckman’s inverse Mill’s ratio (\(\lambda\)) will be incorporated into equation (1).

The specific function of equations (1) and (2) are shown as follow.

\[
SP_i = f\left(SC, UR, AP, N_i, M_{1i}, M_{2i}, FI_i, FE_{1i}, FE_{2i}, FE_{3i}, MV_i, ST_i, LT_i\right) + \epsilon_i \tag{1}
\]
\[ SC_i = f(FE_{i1}, FE_{i2}, FE_{i3}, MV_i, FI_i, N_i) + \psi_i \]  

where

\[ SP_i \]: The BCT score of student \( i \).

School Factors

\( SC_i \): School Choice. It is a dummy variable. When the parents choose a private junior high school for their children, \( SC_i = 1 \) is used; \( SC_i = 0 \) means that the parents choose a public junior high school.

\( UR_i \): Urban-Rural difference. One school, referred to as the “rural school” is defined as a school located in an area in which the population density is less than 100 persons per square kilometer and \( UR_i = 1 \). On the contrary, if the school is located in an area with a population density of more than 100 persons per square kilometer, it is defined as an “urban school” and \( UR_i = 0 \).

\( AP_i \): Ability grouping, where \( AP_i = 1 \) means that the student’s school has an ability grouping system, and \( AP_i = 0 \) means that the student’s school does not have an ability grouping system.

Peer effect factors

\( N_i \): Numbers of older siblings.

\( M_{1i} \): The number of older siblings who are studying in the national universities and colleges.

\( M_{2i} \): The number of older siblings who are studying in highly-ranked senior high schools.

Family characteristics

\( FI_i \): Family income. This is classified into 7 levels, which are, no-income, less than 20,000, 20,000~39,999, 40,000~59,999, 60,000~99,999, 100,000~149,999 and more than 150,000 NT dollars. The scale extends from 1 to 7, respectively.

\( FE_i \): Level of father’s education. We set four levels and there are three dummy variables to label the degrees, i.e., \( FE_{i1}, FE_{i2}, FE_{i3} \) and \( FE_{i4} = 1 \), \( FE_{i2} = 1 \) and \( FE_{i3} = 1 \) mean the father have high school, college, and graduate degrees, respectively.

\( MV_i \): Mother’s employment status. \( MV_i = 1 \) means the mother has a job and \( MV_i = 0 \) means that the mother does not have a job.

Individual characteristics
10

ST$_i$ : Studying time (hours) means the amount of time spent studying when not in class every day.

LT$_i$ : Leisure time (hours) every day.

Since equation (2) is a binary choice function for private/public school choice, it could be estimated using Probit model as shown in equation (3).

$$SC_i^* = X_i \beta + \psi_i, \quad i = 1, 2, ..., n$$

$$SC_i = \begin{cases} 1 & \text{if } SC_i^* > 0 \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (3)

where $\psi_i$ is an independent random variable which belongs to the standard normal distribution ($\psi_i \sim N(0,1)$). Dependent variable ($SC$) will be one if parents choose a private junior school for their children and zero otherwise. Equation (4) can be expressed in the form of a probability model with a binary choice as follows:

$$P(SC_i = 0) = P(SC_i^* \leq 0) = 1 - \Phi(X_i \beta)$$

$$P(SC_i = 1) = P(SC_i^* > 0) = \Phi(X_i \beta)$$  \hspace{1cm} (4)

where $\Phi(\cdot)$ refers to the accumulated density function of the standard normal distribution.

Heckman’s inverse Mill’s ratio ($\lambda_i$) could be obtained after equation (4) is estimated and is shown in equation (5).

$$\hat{\lambda}_i = \frac{\phi(X_i \hat{\beta})}{\Phi(X_i \hat{\beta})} \quad \text{if } SC_i = 1$$

$$= \frac{\phi(X_i \hat{\beta})}{1 - \Phi(X_i \hat{\beta})} \quad \text{if } SC_i = 0$$  \hspace{1cm} (5)

where $\phi(\cdot)$ refers to the density function of the standard normal distribution.

Such Heckman’s inverse Mill’s ratio ($\hat{\lambda}_i$) as well as the predicted value from equation (3) will be incorporated into equation (1). Therefore, equation (1) will be
shown as follows.

\[ SP_i = f(\hat{SC}_i, UR_i, AP_i, N_i, M_{i1}, M_{i2}, FI_i, FE_{i1}, FE_{i2}, FE_{i3}, MV_i, ST_i, LT_i, \lambda_i) + \epsilon_i \]  

Equation (6) can be estimated using OLS. We will normally obtain OLS results that are based on the conditional mean where the focus is only on the mean performance. However, as a policy-maker or a parent, we are not merely interested in the conditional mean performance but also the extreme performance. To obtain a more complete picture of the different quantiles, we can also consider several different regression curves that correspond to the various percentage points of the distributions and not only the conditional mean distribution which neglects the extreme relationship between variables. The empirical results based on the quantile regressions lead to much more meaningful and interesting policy implications. Thus, the quantile regression methodology of Koenker and Bassett (1978) will be applied to equation (6) in this study.

Based on the study of Koenker (2005), a general method for estimating models of conditional quantile functions can be expressed as the solution to a simple optimization problem underlying the least squares model. If the least regression model is set as in equation (7):

\[ y_i = x_i \beta + u_i \]  

Then the estimation of \( \beta \) can be expressed by solving the following function:

\[ \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} (y_i - x_i \beta)^2 \]  

Similarly, Koenker and Bassett (1978) point out that the \( \tau \)th sample quantile, \( \hat{\alpha}(\tau) \), can be found by solving equation (9).

\[ \min_{\alpha \in \mathbb{R}} \sum_{i=1}^{n} \rho_{\tau}(y_i - \alpha) \]
where for any $0 < \tau < 1$, $\rho_\tau(u) = u \cdot (\tau - I(u < 0))$, $Q_y(\tau \mid x)$ is the $\tau$th linear conditional quantile function, and $Q_y(\tau \mid x) = x' \beta(\tau)$. Then, we can find $\hat{\beta}(\tau)$ by solving equation (10).

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_\tau(y_i - x_i' \beta)$$

The estimation of $\hat{\beta}(\tau)$ raises two problems. First, the estimator is constrained by the objective function, eq. (10), which is not a continuous function. Koenker and d’Orey (1987) propose using a linear programming method to facilitate the parametric estimation of the quantile regression. Second, the nuisance parameters of the covariance matrix are constrained so that different quantiles have different covariance matrices to be estimated. Moreover, for each quantile, one should estimate a sparsity function which is the reciprocal of the disturbance density function in the estimating quantile. Thus, the estimation of the quantile regression asymptotic solutions is quite involved. The following subsection briefly illustrates the method used to find the asymptotic solution.

For valid statistical inference in the quantile regression model, it is necessary to provide consistent estimators of the asymptotic variance covariance matrices (Billas et al. 2000). The latter are difficult to estimate reliably since they involve conditional densities of the error terms. Therefore, the following estimation methods provided by Koenker (2005) are employed to obtain the asymptotic variance covariance matrices.

While the asymptotic theory of $\hat{\beta}(\tau)$ is derived from the practical non-IID settings, the limiting covariance matrix of $\sqrt{n}(\hat{\beta}(\tau) - \beta(\tau))$ takes the form of a Huber (1967) sandwich, $\sqrt{n}(\hat{\beta}(\tau) - \beta(\tau)) \sim N(0, \tau(1-\tau)H_n^{-1}J_n H_n^{-1})$, where $J_n(\tau) = n^{-1}\sum_{i=1}^n x_i x_i'$ and $H_n(\tau) = \lim_{n \to \infty} n^{-1}\sum_{i=1}^n x_i x_i' f_i(\xi_i(\tau))$, with $f_i(\xi_i(\tau))$ being the conditional density of
the response, \( y_i \), evaluated at the \( \tau \)-th conditional quantile where \( \xi_i(\tau) \) is the \( \tau \) th conditional quantile function. In the IID case, these \( f_i \)'s are identical and easy to find in asymptotic theory. However, in the non-IID case, we need to take more aspects into consideration as follows.

As Koenker (2005) points out, the preceding approach may be easily extended to the problem of estimating asymptotic covariance matrices for distinct vectors of quantile regression parameters. In these cases, Koenker provides that the asymptotic covariance matrix of \( \hat{\beta}(\tau_1), \ldots, \hat{\beta}(\tau_m) \) can be estimated in terms of the following blocks

\[
\text{Cov}(\sqrt{n}(\hat{\beta}(\tau_i) - \beta(\tau_i)), \sqrt{n}(\hat{\beta}(\tau_j) - \beta(\tau_j))) = [\tau_j \wedge \tau_j - \tau_j \wedge \tau_j]H_n(\tau_j)^{-1}J_nH_n(\tau_j)^{-1},
\]

where \( i \) and \( j \) run from 1 to \( m \).

Three approaches for estimating the matrix \( H_n \) are introduced by Hendricks and Koenker (1991), Powell (1991), and Bilias et al. (2000). Hahn (1995) shows that the bootstrapping method can be estimated from the linear equation (3). In this case, \( \hat{\beta}(\tau) \) is obtained by

\[
\hat{\beta}(\tau) = \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_i(y_i - x_i^T \beta)
\]

and \( \hat{F}_n(u) \) denotes the empirical distribution of the residuals, \( \hat{u}_i = y_i - x_i^T \hat{\beta}(\tau) \), and \( \hat{F}_n(u) = n^{-1} \sum_{i=1}^n I(\hat{u}_i < u) \). Then, by drawing bootstrapping samples \( \hat{u}_i^*, \ldots, \hat{u}_n^* \) from \( \hat{F}_n(u) \), and setting \( y_i^* = x_i^T \hat{\beta}(\tau) + \hat{u}_i^* \), \( \hat{\beta}(\tau) \) can be computed from the above equation. However, the IID error, location-shift model and the residual bootstrap are less practical for quantile regression purposes. Fortunately, the \( (x, y) \) pair form of the bootstrap provides a simple and effective alternative for the independent but non-IID setting (Koenker, 2005). Bilias et al. (2000) provide a simple bootstrapping method by convexifying Powell’s approach in the resampling stage. A major advantage of the new methods is that they can be implemented using efficient linear programming. Simulation studies
show that the methods are reliable even with moderate sample sizes. Therefore, the bootstrapping method is much more powerful for constructing tests and confidence intervals and one can compute the empirical covariance matrix directly based on this method. In the following, we employ Powell’s method to estimate the quantile regression sandwich and the reliable bootstrap method of Bilias et al. (2000) to build the confidence interval for $\hat{\beta}(\tau)$. We also accept that the above two indicators provide the significance level to test the significance of the estimation of $\hat{\beta}(\tau)$.

4 Data Description

The data were collected by using a questionnaire survey to obtain relevant data including a set of variables describing school characteristics, peer group characteristics, and family and student background characteristics. Based on these three important characteristics, the major features of our questionnaire include the BCT score, the environment of the junior high school that the student attended, as well as the student and family background characteristics of every investigated student. The questionnaire sampling data were drawn from the overall survey of 309 public and private senior high schools in 2005 in Taiwan. The random sampling method was employed to select one class in each school.

The final effective sample included 5,175 observations which were collected from 225 of the 309 senior high schools in Taiwan. Descriptive statistics of variables can be found in Table 1. The observations included 685 students from private schools, and 4,490 students from public schools. The tuition fee in a private junior high school is about four to five times than in a public school.

To achieve good scores in this BCT test in Taiwan, many junior high schools
categories students into two groups. Those who has higher performance in school will be grouped into a “Best Group” while the others go to a “Normal Group”. The “Best Group” will receive more resources than the “Normal Group”. Arranging excellent teachers in the “Best Group” is an example. On the other hand, peer-group effects are found to be significant in this ability grouping system. Students can learn more efficiently in this “Best Group” when most students in the class have same ability. Thus, students in the “Best Group” usually obtain higher scores in the BCT than the students in the other group.

As to the educational policy of each school in terms of implementing either normal grouping or ability grouping, the data set comprised 2,646 students attending normal grouping schools, and 2,529 students attending ability grouping schools. For a descriptive description of these test scores, the mean value was 207.90, and the maximum and minimum score were 300 and 23, respectively. The test score distribution is provided in Table 2. The table shows that the score distribution appeared to be skewed to the right, and that 44.164% of the observations were distributed between the scores of 201 and 250 while 26.280% were distributed between the scores of 151 and 200.

As for the levels of these factors in terms of the student’s family peer group effects, family experience was measured by conventional inputs such as the number of older siblings, and the number of older siblings who had been educated in better universities or senior high schools. The national universities (i.e. public universities) were defined as better universities while the good senior high schools were ranked in the upper level in Taiwan. The descriptions of our observations according to the family peer effect are illustrated as follows. First, the data consisted of 509 students without elder siblings, and of 2,343, 1,885, and 438 students having one, two and
more than three older siblings, respectively. Secondly, students without older siblings attending better university or good senior high school is about 2,604, which accounted for 50.32% of all observations. However, 1,775 students have one older brother or sister in a better university or high school while 796 students who have two older brother or sister in a better university or high school.

The descriptions of family characteristics are given below. For the point of view on father’s educational levels, students’ fathers with a high school educational level is about 2,144 while there are 1,555 and 272 students whose fathers had a college degree and master degree. On other hand, 1,790 students’ mothers are house-keepers while the others have full time jobs.

5 Empirical Results

The Probit 2SLS regression approach was implemented to deal with the potential endogenous of public-private school choice. After addressing the endogeneity problems, the estimated outcomes of the variable, namely, public-private school choice, was substituted by its predicted value and inverse Mill’s ratios in the educational production function of equation (6) to avoid biased estimators. Later, the educational production function is implemented by using the conventional OLS and quantile regression approaches. The econometric software applied to complete the quantile regression estimation is the R language. The estimates for the quantiles were $\tau = 0.1, 0.3, 0.5, 0.7, 0.9$, and the reliable bootstrap method of Bilias et al. (2000) was employed to build the confidence interval for $\hat{\beta}(\tau)$. We employed 1,000 bootstrap replication samples in this study.

Table 3 reports the first-step results of school choice regressions based on completing the Probit regression. Family income, father’s educational level and
mother’s employment status affect student school choice positively. In other words, the empirical results indicate that the family with higher income as well as higher father educational level is found to have a higher probability to select a private school while students’ mother with a full time job are found to be more likely to select a private school. This finding agrees with the study by Goldhaber (1996) which shows that families with more education, higher income, and spending more money on their children’s education, are more likely to send their children to a private school.

The estimation results of educational production function using OLS and quantile regressions are shown in Table 4. Several dimensions including school characteristics, peer-group effects, and family and student individual characteristics are addressed and compared between the OLS and Quantile approaches.

5.1 School Characteristics

Two dimensions of school characteristics could be compared here. The first one is to compare the effects of private and public schools on students’ performances while the second one is effects of school location. In the OLS estimates, the mean effect of Mill’s ratio ($\hat{\lambda}$) is found to be significantly positive which indicates that average achievement of private school students is higher than that in public schools. Such finding is consistent with the study by Stevans and Sessions (2000). Similar results are found in quantile regression. However, such estimated coefficients of Mill’s ratio vary along with the different quantiles. We found that the magnitude of Mill’s ratio exhibits a decreasing trend with respect to quantile as shown in Table 4 and subfigure LD in Figure 1. This finding implies that the gap of student’s performance between private and public school is reduced as student’s ability is increased. In other words, there is no much difference for a higher ability student on
selecting a private or a public school.

A particularly interesting question is related to the issue of urban-rural difference which could be found from school location variable (UR). The OLS estimate shows a significant coefficient of 59.866 for the indicator of urban-rural difference, indicates that the mean test scores of the urban students are 59.866 significantly higher than those of the rural students when the maximum score is 300 point. In addition, the quantile estimate of the test scores shows that the maximum effect of the urban-rural difference lies in the 0.1 quantile, while the minimum effect is in the 0.9 quantile. Moreover, the subfigure of UR in Figure 1 shows that the magnitude of the urban-rural difference exhibits a decreasing trend as quantile increases. The policy implication for school characteristics indicates that a good ability student is not affected by school characteristics including school location and private or public school. However, such school characteristics do affect students whose ability is not an excellent.

5.2. Peer Group Effects

Two peer group effects will be analyzed here. The first one is the ability grouping system while the second one is family peer effects from older sisters or brothers. As mentioned above, there is ability grouping system in Taiwan that groups higher ability students into a “Best Group” and the others go to a “Normal Group”. Some schools adopt this ability grouping system but some schools not. Therefore, it is interested to compare student performance of this grouping system. It shows that scores are not significantly affected by ability grouping (AP) system from OLS regression. However, the results from quantile regressions indicate that the effects of ability grouping are both positive and significant in the 0.7 and 0.90
quantiles. This result provides an interesting point which implies that the ability grouping influences the higher-ability students, but does not significantly affect the lower-ability students. In other words, students with higher-ability are grouped in a class will have a better performance, but such outcome is not suitable for other students.

In turning to the family’s peer effect, the OLS estimation results indicate that the number of older siblings who had attended better junior high schools (M2) had a significantly positive effect on their test scores. Similar outcomes are found from quantile regression but the magnitude of this effect decreases as student’s ability increases.

5.3. Individual and Family Characteristics

With respect to individual characteristics, there are two interesting points that should be addressed. First of all, the OLS estimate shows that the input of the student’s study time (ST) has a significant positive effect on test scores. However, the quantile regression results indicate that the effect of study time exhibits a decreasing trend from the 0.1 to 0.9 quantiles. The subfigure of ST in Figure 1 also shows that the positive impacts of study time on score for lower-performance students are stronger than those for higher-performance students. This means that the efficiency of studying time for the lower-ability students is higher than the higher-ability students which implies lower-ability students need to study more. The second result of particular interest is how the leisure time (LT) influences the student’s test scores. The first column of Table 4 presents the OLS results, where we find that the average performances were reduced by a score of about 1.580 as the student increased his or her leisure time by one hour. However, the quantile regression estimates of Table 4
shows that there is no significantly effects on higher-ability students which implies that these students grades will not be affected when they have more leisure time.

The students’ family background indicators reveal some important findings, especially in terms of the father’s educational level and the mother’s employment status. Firstly, the estimated results from the OLS models indicate that the mean effects of the father’s educational levels are significantly positive. The effects are 8.322, 15.548, and 26.013 when the fathers have high school, college, and graduate degrees, respectively. Moreover, the quantile regression results indicate that the impact of the father’s educational level is both positive and decreasing along with increasing alternative quantiles. The clearly declining trend can be found in the $\mathcal{HE}_3,\mathcal{HE}_2,\mathcal{HE}_1$ subfigure in Figure 1, which shows that the effect of a father’s education level is stronger in the lower quantiles. Secondly, a mother with a full-time job does not influence her children’s performance based on the estimation from OLS regression while there is a significant negative relationship between the mother’s employment status and her children’s test performance in the 0.7 and 0.9 quantiles based on the quantile regressions. This finding implies that a higher-ability student may need a mother without full-time job to improve his performance.

6 Concluding Remarks and Policy Implications

The main purpose of this study is to estimate the traditional educational production function using both the Probit two stage least square as well as quantile regression approach to catch the behaviors of two tails students’ performances. Such methodology is applied into the study of BCT test in Taiwan. There are certain implications for policy based on the above empirical estimation results and findings.

First, both the choice of private/public school and location of school affect
students’ performances from OLS estimation outcome as well as from quantile regression outcomes. However, we found that the gap of student’s performance between private and public school is reduced as student’s ability is increased. In other words, there is no much difference for a higher ability student on selecting a private or a public school. Similar estimation result is also found for school location. Such findings for the policy implication on school characteristics indicate that a good ability student is not affected by school characteristics including school location and private or public school. However, such school characteristics do affect students whose ability is not an excellent.

The second major findings in this study indicates the current ability grouping system for some schools in Taiwan is not necessary since the impacts of such ability grouping system on student’s average performance is not significant. The ability grouping system in Taiwan has created educational inequality problem since many parents in Taiwan have adopted many ways to push their children being arrangement in the “Best Group” in order to have a higher score in BCT test. However, the empirical finding does not support the students’ performance in such ability grouping school have better performance than the normal grouping school. Furthermore, in the cases of both the higher- and lower-ability students, the marginal effects of the family peer group always exceed those of the school peer group, which means that the peer effects arising in the students’ families deserve to receive more attention.

Third, we found that the father’s educational level and the presence of higher-ability older siblings have significant and positive impacts on test performances which is consistent with the study by Coleman et al. (1966) which pointed out that the major differences in terms of student performances arose from family inputs, and not school inputs. From these results one may say that family
characteristics have more noticeable influences than that of school factors.

Fourth, in turning to the individual inputs, the estimates of the quantile regressions show that the higher-ability students are less adversely affected by the increase in leisure time, and less positively influenced by increasing their study time. We therefore conclude that the higher-ability students could increase their leisure time to learn and develop other facilities and interests without harming their academic performance. On the contrary, if the lower-ability students were to increase their study time they would perform better and this would lead to a significant positive influence on their overall achievement.

Finally, one of the most important purposes of this study is to investigate the impacts of urban-rural differences on test scores. The mean effect of 59.866 for the urban-rural difference indicator suggests that the average test scores of urban students are higher than those of rural students. As we noted, there are original and contrived differences in educational resources between urban and rural areas in Taiwan, especially in terms of the incentives provided by learning resources, cultural communication, and educational quality. Since many parents have moved away from rural areas into the urban regions to register their children in urban schools in order into obtain more learning resources and opportunities, therefore, the urban-rural difference has subsequently worsened. In addition, the quantile estimates indicate that the urban-rural difference has a stronger effect in the lower quantile which means that lower-ability students are more significantly affected by the urban-rural difference. Thus we propose that the government should not only improve the rural educational resources, but should also develop the social and economic environment, including the reduction in the poverty gap and the enhancement of the learning environment in the rural areas, thereby narrowing the educational differences between the urban and
rural areas.
References


Table 1. Descriptive Statistics of Variables (N=5,175)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP</td>
<td>209.70</td>
<td>44.84</td>
<td>23.00</td>
<td>300.00</td>
</tr>
<tr>
<td>N</td>
<td>1.45</td>
<td>0.83</td>
<td>0.00</td>
<td>6.00</td>
</tr>
<tr>
<td>M1</td>
<td>0.68</td>
<td>0.82</td>
<td>0.00</td>
<td>6.00</td>
</tr>
<tr>
<td>M2</td>
<td>0.15</td>
<td>0.42</td>
<td>0.00</td>
<td>4.00</td>
</tr>
<tr>
<td>ST</td>
<td>2.51</td>
<td>1.68</td>
<td>0.04</td>
<td>20.00</td>
</tr>
<tr>
<td>LT</td>
<td>1.95</td>
<td>1.45</td>
<td>0.00</td>
<td>16.00</td>
</tr>
</tbody>
</table>

Table 2. Test Score Distribution

<table>
<thead>
<tr>
<th>Scores</th>
<th>Number of Observations</th>
<th>Percentage (%)</th>
<th>Accumulative Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-100</td>
<td>128</td>
<td>2.473</td>
<td>2.473</td>
</tr>
<tr>
<td>101-150</td>
<td>389</td>
<td>7.517</td>
<td>9.988</td>
</tr>
<tr>
<td>151-200</td>
<td>1360</td>
<td>26.280</td>
<td>36.268</td>
</tr>
<tr>
<td>201-250</td>
<td>2389</td>
<td>46.164</td>
<td>82.462</td>
</tr>
<tr>
<td>251 above</td>
<td>909</td>
<td>17.565</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>5175</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

* The percentage (%) means the ratio of grouped samples to whole observations collected by answering the total test score.
Table 3. The Results of Public-Private School Choice

<table>
<thead>
<tr>
<th>Variable</th>
<th>Public-private school choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interception</td>
<td>-1.652**</td>
</tr>
<tr>
<td></td>
<td>(-16.129)</td>
</tr>
<tr>
<td>Father has a high school education (FE₁)</td>
<td>0.216**</td>
</tr>
<tr>
<td></td>
<td>(3.481)</td>
</tr>
<tr>
<td>Father has a college degree (FE₂)</td>
<td>0.237**</td>
</tr>
<tr>
<td></td>
<td>(3.508)</td>
</tr>
<tr>
<td>Father has a graduate degree (FE₃)</td>
<td>0.319**</td>
</tr>
<tr>
<td></td>
<td>(2.976)</td>
</tr>
<tr>
<td>Mother’s employment status (MV)</td>
<td>0.085*</td>
</tr>
<tr>
<td></td>
<td>(1.761)</td>
</tr>
<tr>
<td>Family income (FI)</td>
<td>0.082**</td>
</tr>
<tr>
<td></td>
<td>(4.286)</td>
</tr>
<tr>
<td>Number of older siblings (N)</td>
<td>-0.037</td>
</tr>
<tr>
<td></td>
<td>(-1.336)</td>
</tr>
</tbody>
</table>

Note: ** and * denote significance at the 5% and 10% levels, respectively. The numbers in parentheses are t-values.
### Table 4. Estimates of Test Scores with respect to OLS and Quantile Regressions

<table>
<thead>
<tr>
<th>Variable</th>
<th>OLS</th>
<th>Quantile Regressions</th>
<th>( \tau ) 0.10</th>
<th>0.30</th>
<th>0.50</th>
<th>0.70</th>
<th>0.90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>130.025</td>
<td></td>
<td>(19.034)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School choice (SC)</td>
<td>-66.527</td>
<td></td>
<td>(-1.744)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban-rural difference (UR)</td>
<td>59.866</td>
<td></td>
<td>(6.205)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ability grouping (AP)</td>
<td>1.595</td>
<td></td>
<td>(1.312)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of older siblings (N)</td>
<td>-2.671</td>
<td></td>
<td>(-2.929)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The number of older siblings who are studying in national universities or colleges (M1)</td>
<td>-2.852</td>
<td></td>
<td>(-3.008)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The number of older siblings studying in high-ranking senior high schools (M2)</td>
<td>8.764</td>
<td></td>
<td>(5.411)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family Income (FI)</td>
<td>4.608</td>
<td></td>
<td>(5.306)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father has a high school education (FE₁)</td>
<td>8.322</td>
<td></td>
<td>(3.822)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father has a college degree (FE₂)</td>
<td>15.548</td>
<td></td>
<td>(6.352)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father has a college degree (FE₃)</td>
<td>26.013</td>
<td></td>
<td>(6.524)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother’s employment status (MV)</td>
<td>-1.605</td>
<td></td>
<td>(-1.123)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study time (ST)</td>
<td>3.338</td>
<td></td>
<td>(9.238)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leisure time (LT)</td>
<td>-1.580</td>
<td></td>
<td>(-3.764)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{\lambda} ) (LD)</td>
<td>36.707</td>
<td></td>
<td>(1.790)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: ** and * denote significance at the 5% and 10% levels, respectively. The numbers in parentheses are t-values. The t-values of the quantile regression are obtained by employing the bootstrapped standard errors of Bilias et al. (2000)
Figure 1. The Coefficients and 95% Confidence Intervals of the Quantile Regression (blue line) and the OLS Regression (red line) (the X-axis consists of the quantile, and the Y-axis the coefficient)