Value-added in higher education: ordinary least squares and quantile regression for a Colombian case

Valor agregado en educación superior: mínimos cuadrados ordinarios y regresión cuantílica para un caso colombiano

Jose D. Bogoya[1], Johan M. Bogoya[2], and Alfonso J. Peñuela[3]

ABSTRACT

Colombia applies two mandatory National State tests every year. The first, known as *Saber 11*, is applied to students who finish the high school cycle, whereas the second, called *Saber Pro*, is applied to students who finish the higher education cycle. In this paper, the result obtained by a student on the *Saber 11* exam along with his/her gender and socioeconomic stratum are our independent variables while the *Saber Pro* outcome is our dependent variable.

We compare the results of two statistical models for the *Saber Pro* exam. The first model, multi-linear regression or ordinary least squares (OLS), produces an overall well fitted result but is highly inaccurate for some students. The second model, quantile regression (QR), weights the population according to their quantile groups. OLS minimizes the errors for the students whose *Saber Pro* result is close to the mean (a process known as estimation in the mean) while QR can estimate a value in the θ-quantile for every $0 < \theta < 1$. We show that QR is more accurate than OLS and reveal the unknown behavior of the socioeconomic stratum, the gender, and the initial academic endowments (estimated by the *Saber 11* exam) for each quantile group.

Keywords: Value-added, higher education, evaluation, quantile regression, statistical model.

RESUMEN

En el sistema educativo de Colombia se realizan dos exámenes nacionales obligatorios al año. El primero, conocido como *Saber 11*, está dirigido a los estudiantes que finalizan el bachillerato, mientras que el segundo, conocido como *Saber Pro*, evalúa a los estudiantes que terminan un estudio superior. En este estudio, el resultado obtenido por un estudiante en el examen *Saber 11*, junto con su género y estrato socioeconómico, son nuestras variables independientes, mientras que el resultado del examen *Saber Pro* es nuestra variable dependiente.

Comparamos los resultados de dos modelos estadísticos para *Saber Pro*. El primer modelo, regresión multi-lineal o mínimos cuadrados (OLS, por sus siglas en inglés), produce un buen ajuste general pero es impreciso para ciertos estudiantes. El segundo modelo, regresión cuantílica (QR, por sus siglas en inglés), mide la población de acuerdo con su cuantil. El OLS minimiza los errores para los estudiantes cuyo resultado en *Saber Pro* está cercano a la media (proceso conocido como estimación en la media) mientras que el QR puede estimar un valor en el cuantil θ para cada $0 < \theta < 1$. Mostraremos que el QR es más preciso que el OLS y revelaremos el comportamiento desconocido del estrato socio económico, el género y la preparación académica inicial (estimada con el examen *Saber 11*) para cada cuantil.

Palabras clave: Valor agregado, educación superior, evaluación, regresión cuantílica, modelo estadístico.

Introduction

The representation of the educational phenomenon through mathematical models, where variables of the cognitive state participate at the beginning and at the end of the cycle, as well as, features of the accomplished process, allows developing studies of impact and the efficiency of the displayed projects by a universe of educative institutions.

Particularly, the contribution to a group of student academic achievement, conferred by the institutions and their professors, requires the employment of valid assessment tools for estimating reliably the reached states at the beginning and at the end of a period. In this way, it ensures the credibility of the calculated efficacy (Amrein-Beardsley,

[1] Chemical Engineer and M.Sc. In Computer Systems Engineering. Universidad Nacional de Colombia. E-mail: jdbogoyam@unal.edu.co.

[2] Mechanical Engineer and Pure Mathematician. Universidad de los Andes, Colombia. M.Sc. and Ph.D. in pure Mathematics, CINVESTAV México. Affiliation: Pontificia Universidad Javeriana, Colombia.
E-mail: jbogoya@javeriana.edu.co.

[3] Pure Mathematician. M.Sc. in Economics. Universidad Sergio Arboleda, Colombia. Affiliation: Universidad Sergio Arboleda, Colombia. E-mail: alfonso.penuela@usa.edu.co.

2008, p. 71). The impact of the educational project facilitates the accountability of inclusive institutions, considering feasible goals, since the cognitive state proven at the end of a cycle depends on, in a high level, the respective state that the students show at the beginning of it (Hanushek & Raymond, 2001, p. 375).

To estimate the educational performance of a student i in a certain moment t, the Equation (1) has been formulated (Hanushek, 1979, p. 363), whose variables are: innate capacities (I_i), accumulated characteristics until the moment t (B_i^t), peer influence (P_i^t), and institution contribution (S_i^t).

$$A_{it} = f\left(B_i^t, P_i^t, S_i^t, I_i\right). \tag{1}$$

Moving towards the value-added notion, Equation (2) has been proposed to estimate the educational performance of a student i at the end of a period (Hanushek, 1979, p. 364), according to the state of the considered variables at the beginning of such period t^*.

$$A_{it} = f^*\left(B_i^{t-t^*}, P_i^{t-t^*}, S_i^{t-t^*}, I_i, A_{it^*}\right). \tag{2}$$

From the variability of the proposed models, it highlights (3) (Hanushek & Rivkin, 2012, p. 134), in which the educational performance of a student i is estimated according to the following variables: peer and scholar influence (S_i), family and neighbors incidence (X_i), and student individual capacity (μ_i).

$$A_{it} = f\left(S_i, X_i, \mu_i\right). \tag{3}$$

Concerning the connection between the familiar background and academic performance, it has been reported very low and often negative correlation values (Woessmann, 2004, p. 17). In light of the average performance of a mathematical test, based on data from TIMSS 95, the coefficient found is equal to -0,11. It shows the difference of the academic performance between students with parents without a high school degree and those with a professional degree.

Other types of models focus attention on variables that can be oriented to the educational institutions; for example, the Equation (4) estimates the quality of an educational institution (Bishop & Woessmann, 2004, p. 8). This is determined by the learning ability and the effort of the students (AE), and the quantity of resources and their effectiveness of use (IR).

$$Q = (AE)^\alpha (IR)^\beta, \ \alpha + \beta < 1. \tag{4}$$

Related to the cognitive progress of a group of students, the linear Equation (5) introduces the value-added v, in which β_1, β_2, and β_3 are real constants, x_1 and y represent the cognitive state at the beginning and at the end of an educational cycle, respectively, x_2 reflects the socioeconomic condition, s is the student program, and ε is the estimated error. The group value-added is calculated as

the average of the deviations of the observed results to the individual level (Bogoya & Bogoya, 2013, p. 78).

$$y(i) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 + v(s) + \varepsilon(i). \tag{5}$$

For a case study, the authors proposed three approximations to the student level:

- The value of the cognitive state variable at the beginning of the higher education cycle in Colombia as the *Saber 11* exam result.

- The value of the cognitive state variable at the end of the higher education cycle in Colombia as the *Saber Pro* exam result.

- The value of the socioeconomic condition variable as the socioeconomic stratum.

With the case study data, the model solution leaded to the following finding: the cognitive state at the beginning of the cycle explains one portion of the variance of the corresponding variable at the end of that cycle; it is thirteen times greater than the variance related with the socioeconomic condition (Bogoya & Bogoya, 2013, p. 81).

The use of value-added models predicted four considerations. First, it is necessary to remind that the findings significance depend on, among other variables, the number of evaluated students. The greater the population is, the more reliable the estimated value for the effectiveness of an educational institution (Ray, 2006, p. 34). Second, when conducting studies of trends the variation of the student cognitive state, at the end of a cycle, fluctuates relatively seldom among two consecutive years. It implies that volatility of the variation reduces the reliability of the estimation and thus it is important to have averages of several years in small populations (Ray, 2006, p. 34). Third, it is uncertain the variation estimation of the student cognitive state that at the beginning of a period are placed in the top of the generated ladder; in this case it is possible to take the average of several students (Tymms & Dean, 2004, p. 14,15). Finally, in a regression, the coefficient of determination is greater for aggregated data than for individual data. It must be avoided the ecological fallacy, due to that the independent effects tend to be mistaken in that aggregated and it is hard to clarify them (Hanushek, Jackson, & Kain, 1974, p. 100).

However, in order to use the quantile regression methodology to solve value-added models, we found the initial definition about quantiles of an ordered observations set sample, which are structured in a linear model. Considering $\{y_t : t = 1, \ldots, T\}$ as a sample of a random variable Y with cumulative distribution function F, any solution of (6) can be defined as the quantile sample θ, $0 < \theta < 1$ (Koenker & Basset, 1978, p. 38).

$$\min_{b \in \mathbb{R}} \left\{ \sum_{y_t \geq b} \theta \left| y_t - b \right| + \sum_{y_t < b} (1-\theta) \left| y_t - b \right| \right\}. \tag{6}$$

The adjust procedure for quantile regression has been improved in an analogous form as it happens in conventional statistics R^2 of the least squares regression (Koenker & Machado, 1999, p. 1). Simultaneously, several inferential procedures can be formulated for proving hypothesis about combined effects of covariance of a whole range of quantile conditional functions. It is stated that the quantiles are linked with ordering operations and classification of the observations that are used to define them (Koenker & Hallok, 2001, p. 145). It is possible to delimit the quantiles as an optimization problem, taking the sample mean as the solution to minimize the sum of squared residuals and the mean as the solution to minimize the sum of absolute residuals. By symmetry, the minimization of the absolute residual sum must be equal to the positive and negative residuals to guarantee the same number of observations above and under the mean.

It is important to point out that even if the quantile regression has had a considerable development and a variety of applications, there are numerous aspects for research, especially about regularization parameters (Koenker, 2004, p. 88). There are different versions of the model, which might extend the optimal structure for the fixed effects, which incorporate ordinal factors and nonparametric components. The analysis of the method performance for the samples of fixed size is equal to a research route, likewise applications to growing curves that can appear as the natural laboratory of future developments of quantile regression models for longitudinal data.

Econometric methodology

The learning outcomes of higher worldwide education programs come from several conditions and variables (Hanushek, 1979). We study, in two different ways, some possible relationships between them. These variables are approximations of certain general conditions for each individual, such as: socioeconomic and cultural environment, learning level of the students at the beginning of their university studies, and the existence of a wide variety of academic value-added elements of such projects.

We define the following input variables: the score obtained by the student on the national higher education admission exam (*Saber 11*) as a synthesis of the partial scores observed in the evaluated areas; the student socioeconomic stratum at the end of his/her university studies; and the student gender. *Saber 11* result is understood as a proxy of the initial academic level of a student when starting a university program, while the socioeconomic stratum is understood as a proxy of the family income and socioeconomic conditions. For economic decision purposes, the Colombian state uses a number between 1 and 6, called "socioeconomic stratum", to indicate the relative people wealth in certain location; we use this indicator as an input variable. On the other hand, the student gender is a frequently used control variable in this kind of studies.

The output variable is the student score on the national higher education exit exam (*Saber Pro*), understood as a proxy of the academic level when finishing a university program. Our objective is, using the same input variables, to compare two statistical models for the output. The first one is the well-known multi-linear regression or ordinary least squares (OLS) and the second one is quantile regression (QR). Generally speaking, the QR method gives us a detailed OLS-view when analyzing linear models, by supplementing focus on the estimation of the outcome variable for each possible quantile (Brennan, Cross, & Creel, 2015; Frumento & Bottai, 2016). Thus, OLS and QR are different econometric tools and we are interested in comparing them in our specific study.

Let x be a $n \times p$ matrix of independent variables (*Saber 11* outcome, socioeconomic strata, and gender) and $y \in \mathbb{R}^n$ a vector of dependent variables (*Saber Pro* outcome). We assume the following linear model

$$y = x\beta + \varepsilon, \tag{7}$$

where $\beta \in \mathbb{R}^p$ and $\varepsilon \in \mathbb{R}^n$ are constant vectors. Let $x_j \in \mathbb{R}^p$ be the j–st row of the matrix x. We can split the Equation (7) as

$$y_j = x_j \beta + \varepsilon_j, \quad j = 1, \ldots, n. \tag{8}$$

The vector β, which minimizes $\sum_{j=1}^{n} \varepsilon_j^2$, is given by the multi-linear or ordinary least squares regression (OLS) of y with respect to x; the well-known solution is

$$\beta = \left(x^T x\right)^{-1} x^T y,$$

here x^T stands for the transpose matrix of x. This solution is based on the assumption that the expected value of the errors ε_j is zero. In statistics, β is known as the *regression vector* and ε as the *error vector*. OLS minimizes the errors ε_j for the students whose *Saber Pro* result is close to the mean of y while paying less attention to the rest of the population; this behavior is known as *estimation in the mean*.

Now, the quantile regression (QR), as the second model that we study, will be described as follows. For each real $\theta \in [0,1]$, QR consists of determining the vector $\beta^{(\theta)} \in \mathbb{R}^p$ which minimizes

$$\sum_{y_j \geq x_j \beta^{(\theta)}} \theta \left| y_j - x_j \beta^{(\theta)} \right| + \sum_{y_j < x_j \beta^{(\theta)}} (1-\theta) \left| y_j - x_j \beta^{(\theta)} \right|. \tag{9}$$

Note that $y_j \in \mathbb{R}$ and x_j is a $1 \times p$ real vector, thus $x_j \beta^{(\theta)}$ stands for the matrix (inner) product between x_j and $\beta^{(\theta)}$. Assuming (2.2), with $\beta = \beta^{(\theta)}$, we can write (9) as,

$$\sum_{\varepsilon_j \geq 0} \theta \left| \varepsilon_j \right| + \sum_{\varepsilon_j < 0} (1-\theta) \left| \varepsilon_j \right| = \sum_{j=1}^{n} \left| \tilde{\varepsilon}_j \right|,$$

where $\tilde{\varepsilon}_j := \theta \varepsilon_j$ if $\varepsilon_j \geq 0$, and $\tilde{\varepsilon}_j := (1-0)$ if $\varepsilon_j < 0$. Then $\beta^{(\theta)}$ minimizes $\sum_{j=1}^{n} |\tilde{\varepsilon}_j|$, i.e. minimizes the sum of the absolute values of the errors with certain weights. In our case, a θ quantile is a value for the outcome variable y that is bigger than the θ portion of the observations and less than the remaining $1 - \theta$ portion. Additionally, some authors give a nice step-by-step explanation of how to run QR in *Stata* software (Cameron & Trivedi, 2010).

Business administration

An extensive data mining results for 160.207 students which presented the 2009 *Saber Pro* exam in Colombia was used. For these students we know their *Saber 11* result, socioeconomic stratum, gender, and the selected higher education program. From this universe, the set of students evaluated through the business administration *Saber Pro* exam (the largest) is considered. Because of reliability issues, only programs with 20 or more students are taken into account. The database used[4] reports of 10.783 students. The socioeconomic stratum and the gender variables, being categorical, are treated as dummies.
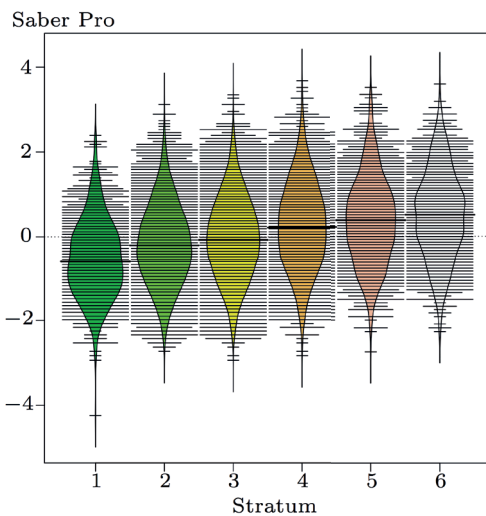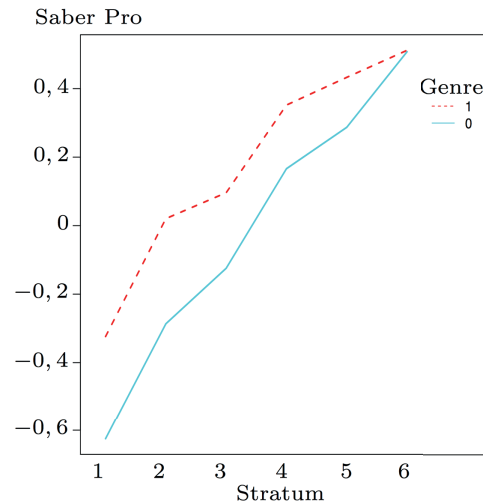

Figure 1. The stratum and *Saber Pro* interplay.
Source: Authors.

Saber Pro is scaled with mean 100 and standard deviation 10, while *Saber 11* has mean 330 and standard deviation 30. In order to simplify the analysis of the models outcomes, we normalized them both, i.e. mean 0 and standard deviation 1.

Figures 1 and 2 show some characteristics of the behavior of the main variables. With colored regions, Figure 1 shows the close to normal distribution of *Saber Pro* in each stratum. Figure 1 also reveals the linear relation between the two variables: in general, the higher the stratum of an individual, the higher his/her *Saber Pro* result will be. Figure 2 shows also the linear relation between *Saber Pro*

and stratum that means, in general, gender 1 (male) gets higher *Saber Pro* results than gender 0 (female) but the difference decreases as the stratum increases.

Figure 2. The *Saber Pro*, stratum, and gender interplay. Genders 0 and 1 stand for female and male, respectively.
Source: Authors.

We assume the model (7) where y_j is the *Saber Pro* outcome for the student j, the row vector x_j is $(x_{j,1}, x_{j,2}, \ldots x_{j,7})$ where the entry $x_{j,1}$ is the *Saber 11* test outcome for the student j, $x_{j,1}$ for $k = 2, \ldots, 6$ takes the value 1 if the student j lives in a socioeconomic stratum k area and the value 0 otherwise, finally $x_{j,7}$ takes the value 0 for a male and the value 1 for a female. The previous description means that socioeconomic stratum 1 and male gender play the role of base variables.

Results

Table 1 shows the numerical results produced by the OLS model. The data were obtained with *Stata*. We get $R^2 = 0{,}842$ showing an accurate model.

Tables 2, 3 and 4 show the numerical results produced by the QR model. The data were obtained with *Stata*. For $\theta = \frac{1}{4}, \frac{1}{2}$ and $\frac{3}{4}$ we show the coefficient β_ℓ, the standard deviation σ_ℓ, and the 95 % confidence interval for each variable $\left(x_{j,\ell}\right)_{j=1}^{n} (\ell = 1, \ldots, 7)$.

Figures 3 through 9 show the QR parameter behavior $\beta_\ell^{(\theta)}$, for $\ell = 1, \ldots, 7$ versus θ. In each case the solid (green) line represents $\beta_\ell^{(\theta)}$, the gray band is the 95 % confidence interval, and the dotted (black) line is the OLS value for β_ℓ. Thus, when the dotted line falls outside the gray band (see Figure 3) the OLS model will generate big errors.

---

4 The public database found at ftp://ftp.icfes.gov.co was used.

Table 1. OLS results. All the variables are meaningful. β_ℓ is the ℓ-th component of the vector β and σ_ℓ is the standard deviation of the column vector $\left(x_{j,\ell}\right)_{j=1}^n (\ell = 1,\ldots,7)$, see (2.2).

Saber Pro	β_ℓ	σ_ℓ	95 % conf. int.
Saber 11	0,65	0,008	[0,64; 0,67]
Stratum 2	0,15	0,030	[0,095; 0,21]
Stratum 3	0,20	0,030	[0,14; 0,26]
Stratum 4	0,24	0,033	[0,18; 0,31]
Stratum 5	0,26	0,036	[0,19}; 0,33]
Stratum 6	0,13	0,041	[0,05; 0,21]
Gender	0,15	0,015	[0,13; 0,18]

Source: Authors.

Table 2. QR results for $\theta = \dfrac{1}{4}$. The variable stratum 6 is not meaningful in this quantile.

Saber Pro	$\beta_\ell^{(1/4)}$	σ_ℓ	95 % conf. int.
Saber 11	0,66	0,0099	[0,64; 0,68]
Stratum 2	0,11	0,039	[0,035; 0,19]
Stratum 3	0,15	0,039	[0,074; 0,23]
Stratum 4	0,16	0,043	[0,074; 0,24]
Stratum 5	0,20	0,047	[0,11; 0,29]
Stratum 6	0,057	0,053	[-0,0047; 0,16]
Gender	0,14	0,019	[0,097; 0,17]

Source: Authors.

**Table 3.** QR results for $\theta = \dfrac{1}{2}$. All the variables are meaningful in this quantile.

Saber Pro	$\beta_\ell^{(1/2)}$	σ_ℓ	95 % conf. int.
Saber 11	0,67	0,010	[0,65; 0,69]
Stratum 2	0,14	0,041	[0,057; 0,22]
Stratum 3	0,20	0,040	[0,12; 0,28]
Stratum 4	0,24	0,045	[0,15; 0,32]
Stratum 5	0,26	0,050	[0,17; 0,36]
Stratum 6	0,16	0,056	[0,049; 0,27]
Gender	0,18	0,020	[0,14; 0,22]

Source: Authors.

Table 4. QR results for $\theta = \dfrac{3}{4}$. All the variables are meaningful in this quantile.

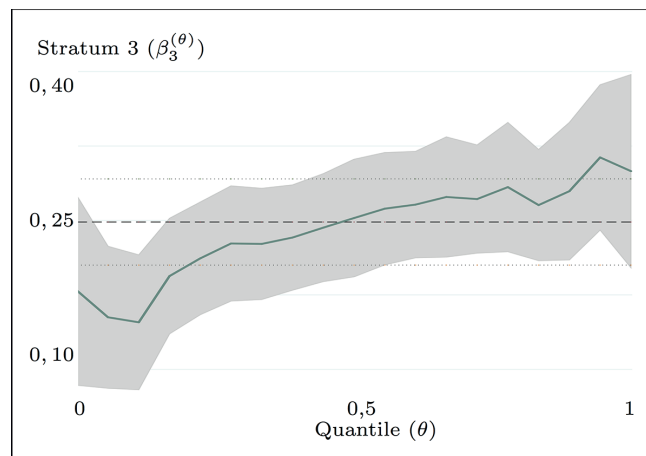| Saber Pro	$\beta_\ell^{(3/4)}$	σ_ℓ	95%s conf. int.
Saber 11	0,68	0,011	[0,65; 0,70]
Stratum 2	0,18	0,045	[0,087; 0,26]
Stratum 3	0,24	0,044	[0,16; 0,33]
Stratum 4	0,31	0,049	[0,21; 0,40]
Stratum 5	0,32	0,055	[0,21; 0,42]
Stratum 6	0,24	0,062	[0,12; 0,37]
Gender	0,17	0,022	[0,13; 0,21]

Source: Authors.
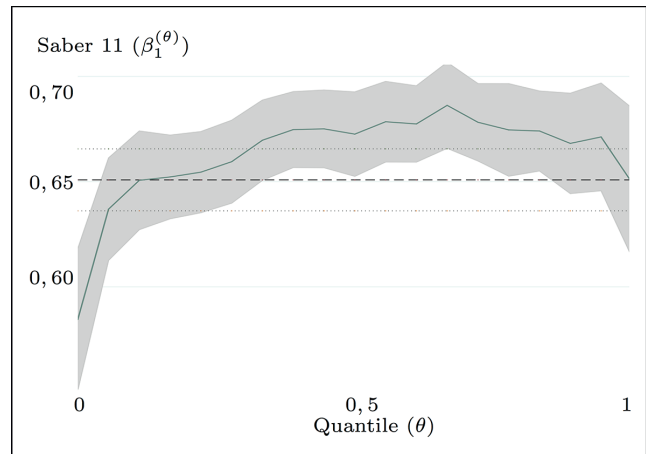
Figure 3. The behavior of $\beta_1^{(\theta)}$ versus θ.
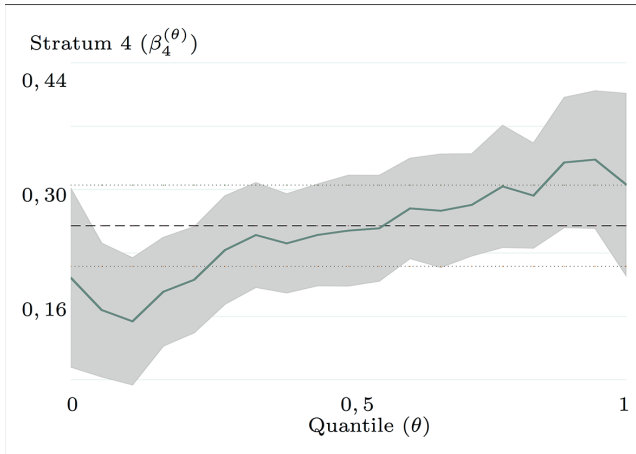**Source:** Authors.
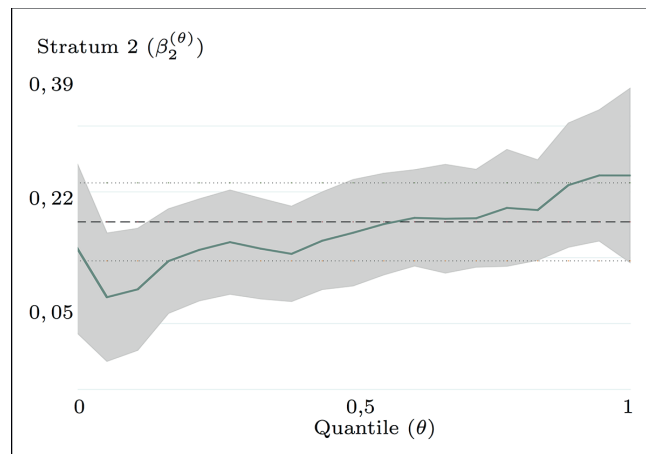
Figure 4. The behavior of $\beta_2^{(\theta)}$ versus θ.
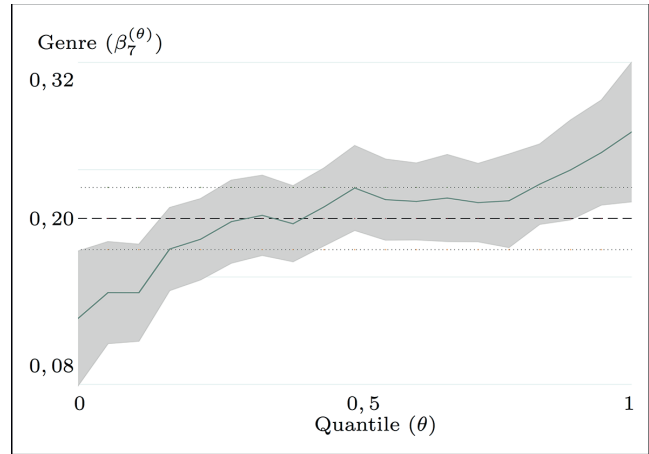**Source:** Authors.

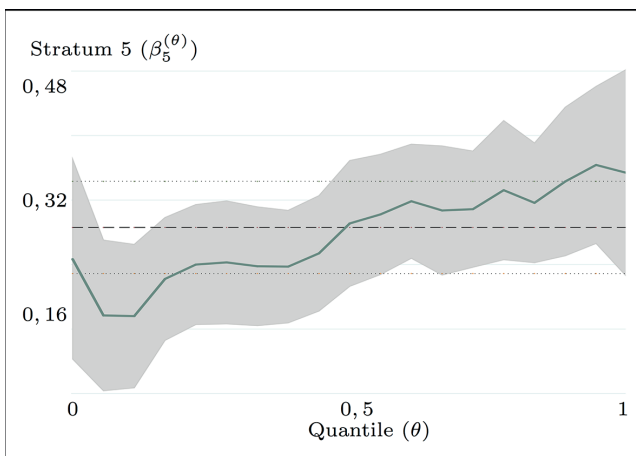**Figure 5.** The behavior of $\beta_3^{(\theta)}$ versus θ.
Source: Authors.

Figure 6. The behavior of $\beta_4^{(\theta)}$ versus θ.
Source: Authors.

Figure 7. The behavior of $\beta_5^{(\theta)}$ versus θ.
Source: Authors.
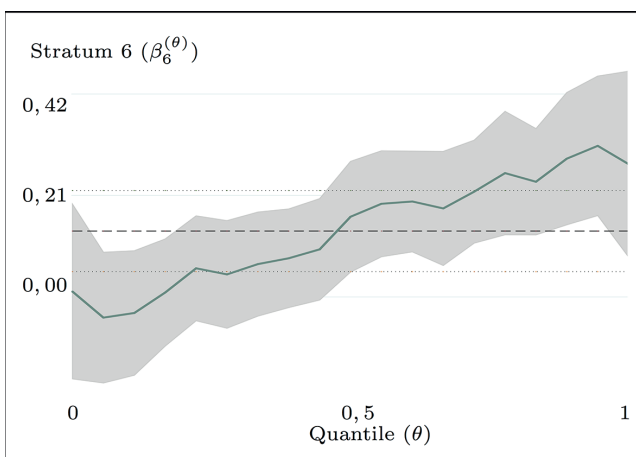

Figure 8. The behavior of $\beta_6^{(\theta)}$ versus θ.
Source: Authors.

Figure 9. The behavior of $\beta_7^{(\theta)}$ versus θ.
Source: Authors.

Conclusions

1. On one side, with the OLS method we can predict the *Saber Pro* outcome for an individual or a group; in the second case, we are minimizing the sum of the error square and modeling the conditional sample mean which is inadequate for some individuals. Thus, this method only gives us information about the individuals located close to the mean. On the other side, the QR method minimizes the absolute sum of the quantile weighted errors, that is, proportionally weighted for each individual not taking into account his/her dispersion. The achieved improvement with the QR method comes with a price, the minimized absolute sum is not differentiable hence; in order to use it, we need the numerical methods offered by most statistical packages.

    With the *Saber 11* outcome of an individual we can know the corresponding quantile θ and we can apply the QR method to obtain a more accurate prediction of his/her *Saber Pro* result.

2. According to Table 1, *Saber 11* $\beta_1^{(\theta)}$ presents by far, the highest regression coefficient (65 % at least) with *Saber Pro* (y). Nevertheless Figure 3 reveals that OLS actually overestimates $\beta_1^{(\theta)}$ for the first quantiles $\left(\theta < 0{,}15\right)$ and underestimates it for the remaining ones $\left(0{,}15 < \theta \leq 1\right)$. The OLS method gives us a constant average value for the regression coefficient ignoring the dependent-independent variable interplay, while the QR method gives us the regression coefficient as a function of θ showing a closer and finer look to the study case at hand. Thus we obtained a poor $\beta_1^{(\theta)}$-OLS accuracy while the QR method shows his advantage by considering all the population in a differentiated way.

3. In our modeling procedure, we take stratum 1 as a base variable; thus Figures 4 to 8 are actually showing the strata-quantile behavior in relation with stratum 1. Taking a look at the vertical axis labels, we can see a progressive increasing value for the regression coefficient (see also Figure 1), which reveals an academic inequality related with the socioeconomic stratum. Additionally, note that all the strata have a similar behavior and that the OLS model is accurate enough for these variables.

4. Figure 9 shows the male-female regression coefficient for the different quantile groups. It reveals that we can expect higher *Saber Pro* results from males in every quantile group and that the OLS-method overestimates the lowest *Saber Pro* outcomes and underestimates the highest one. Showing again that the QR method gives us a different regression coefficient value for each segment of the population, taking into account the dependent-independent variable interplay. The OLS method forbids us to note the changes in the respective variable association.

References

Amrein-Beardsley, A. (2008). *Methodological concerns about the education value-added assessment system*. American Educational Researcher, 37(2), 65-75.

Bishop, J. H. & Woessmann, L. (2004). *Institutional effects in a simple model of educational production.* Cornell University ILR School.

Bogoya, J. D. & Bogoya, J. M. (2013). *An academic value-added mathematical model for higher education in Colombia*. Ingeniería e Investigación, 33(2), 76-81.

Brennan, A., Cross, P.C., & Creel, S. (2015). *Managing more than the mean: using quantile regression to identify factors related to large elk groups*. Journal of Applied Ecology, 52(6), 1656-1664.

Cameron, A. C. & Trivedi, P. K. (2010). *Microeconometrics using Stata*. Stata Press, Texas 2010.

Frumento, P. & Bottai, M. (2016). *Parametric modelling of quantile regression coefficient functions*. Biometrics, 72(1), 74-84.

Hanushek, E. A. (1979). *Conceptual and empirical issues in the estimation of educational production functions*. The Journal of Human Resources, 14(3), 351-388.

Hanushek, E. A., Jackson, J. E., & Kain J. F. (1974). *Model specification, use of aggregate data, and the ecological correlation fallacy*. Political Methodology, 1(1), 89-107.

Hanushek, E. A. & Raymond, M. E. (2001). *The confusing world of educational accountability*. National Tax Journal, 54(2), 365-384.

Hanushek, E. A. & Rivkin, S. G. (2012). *The distribution of the teacher quality and implications for policy*. Annual Review of Economics, 4, 131-157.

Koenker, R. (2004). *Quantile regression for longitudinal data*. Journal of Multivariate Analysis, 91, 74-89.

Koenker, R. & Bassett, G. (1978). *Regression quantiles*. Econome trica, 46(1), 33-50.

Koenker, R. & Machado, J. A. F. (1999). *Goodness of fit and related inference processes for quantile regression*. Journal of the American Statistical Association, 94(448), 1296-1310.

Koenker, R. & Hallock, K. (2001). *Quantile regression*. Journal of Econometrics Perspectives, 15(4), 143-156.

Ray, A. (2006). *School value added measures in England*. A paper for the OECD Project on the development of value-added models in educations systems.

Tymms, P. & Dean, C. (2004). *Value-added in the primary school league tables*. A report for the National Association of the Head Teachers.

Woessmann, L. (2004). *How equal are educational opportunities? Family background and student achievement in Europe and the US*. CESIFO Working Paper No. 1162.