

Umi Mahmudah,
Suhartono,
Siti Fatimah
Indonesia

A Robust Approach to Analyzing the Factors Influencing Quality Education in Indonesia

DOI: 10.15804/tner.2017.49.3.06

Abstract

The main purpose of this study was to analyze factors that affect quality education in Indonesia. Although ordinary least square (OLS) regression is very popular in the analysis of the relationship between independent variables and dependent variable, it is less than perfect when data are contaminated by outliers. Therefore, a robust approach is applied to face the problem of outliers due to its ability to down-weight the outlier influences. Results reveal that M-estimator of robust regression is better used to investigate the factors that influence the quality of education in Indonesia because it provides better accuracy.

Keywords: *quality education, robust approach*

Introduction

Educational quality as a pillar of human resource development plays a strategic role in national development. Thus, education with high quality is an important factor in advancing the future of a nation. The definition of quality education is a complex system, which is embedded in a political, cultural and economic context. Therefore, quality education consists of several categories such as healthy learners who are ready to participate and learn as well as support in learning by family and communities; healthy environment that provides safety, is protective and gender-sensitive as well as provides adequate resources and facilities; content

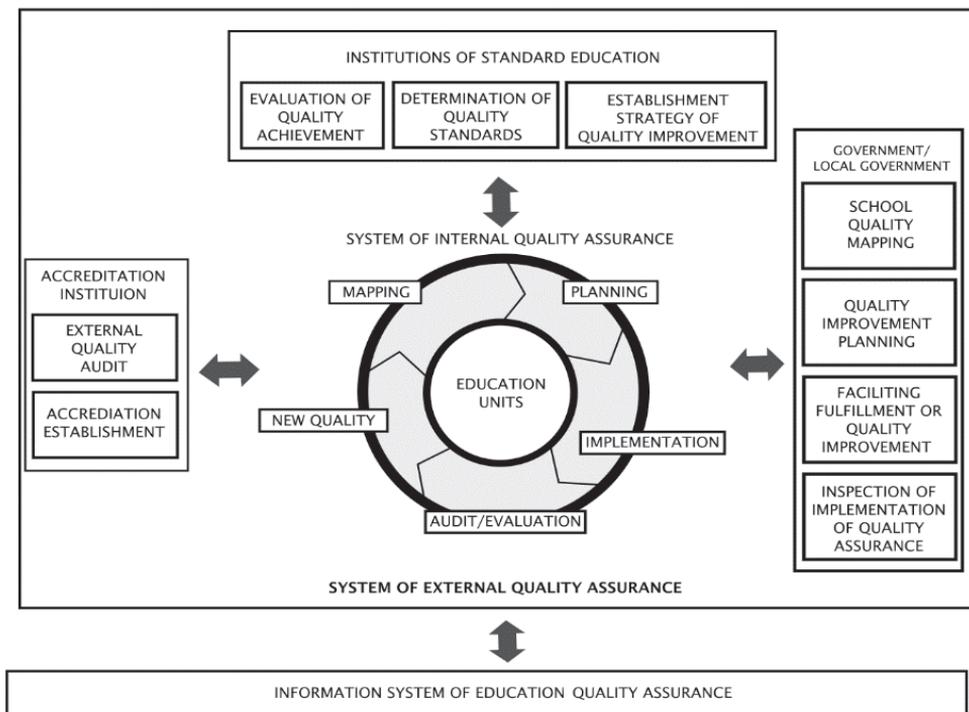
that reflects relevant curricula and materials; processes by trained teachers using child-centered teaching approaches in well-managed classrooms and schools; and outcomes that reflect knowledge, skills and attitudes, which are linked to national goals for education and positive participation in society (UNICEF, 2000). Based on the Ministry of Education and Culture (MoEC, 2013), there are roughly 236,000 schools with 55 million students and 3 million teachers in Indonesia. Even though Indonesia faces problems in quality education especially in improving education outcomes, over the past decade there has been improvement in quality alongside gains in access to primary and junior secondary education (Development Progress, 2014).

Based on an explanation of the Directorate General of Primary and Secondary Education of the Indonesian Ministry of Education and Culture on Guidelines for the Implementation of Quality Assurance by Educational Unit, qualified education not only relies on results but also emphasizes its process. The measure of the successful quality assurance by an educational unit consists of processes, output, outcome, and impact indicators. The process indicators are closely related to the changes in the educational unit management such as the increased ability to monitor and evaluate the executed mechanism. Output indicators are closely related to educators' increased competence and improved management of financial infrastructure. Outcome indicators are closely related to the increased results of student learning, the competence exam results, performance assessment of educators and educational personnel, educational unit achievements, and a pleasant learning environment. Impact indicators are related to the establishment of quality culture with the continuous and sustained quality assurance in the educational unit (MoEC, 2016).

Educational quality assurance aims to satisfy and exceed the national standard of education. According to the agency of the national standard of education, there are eight quality standard indicators in organizing education; the graduate competency standard of basic and middle education, content standard of basic and middle education, process standard of basic and middle education, educational assessment standard, educators and educational personnel standard, facility and infrastructure standard, management standard, and educational financial standard.

Based on Figure 1, the quality assurance system of basic and middle education is the responsibility of every component in the education unit, which consists of two components: internal quality assurance system (SPMI) and external quality assurance system (SPME). SPME is a quality assurance system implemented by the government, local government, accreditation agency, and education standardization agency, whereas SPMI is a quality assurance system that works inside the

Figure 1. Quality assurance system of basic and middle education



(source: Directorate General of Basic and Middle Education, 2016)

education unit and is implemented by all the components in the education unit. Therefore, the implementation of the educational quality assurance system in the educational unit is executed by an approach involving all the components of the educational unit together to have a quality culture. Quality culture will encourage the educational unit to increase the educational quality continually so that educational quality will increase gradually. Based on the results from the Education for All Global Monitoring Report (EFA-GMR) in 2011, Indonesia’s Education Development Index (EDI) was 0.934 based on the data in 2008. This value places Indonesia in the 69th position out of 127 countries in the world. The EDI’s result of this year decreases considerably, especially on the students’ sustainable value of assessment category until the 5th grade in elementary school. This indicates that in this category educational quality in basic education should be improved.

Education in elementary school becomes the foundation or a basis for improvement of the educational quality in Indonesia. Also, since elementary school is an educational unit that organizes the basic educational process and underlies the next education with a purpose to prepare the basic reading skill, writing and counting, knowledge and basic skills useful for the student in their growth level. The main purpose of this study was to investigate the factors that influence quality education in Indonesia by using robust regression, which is a powerful method in dealing with the outliers that often appear in original data sets. Due to the importance of elementary schools, this study focuses on organizing elementary education by Indonesia's major political subdivisions in order to analyze the factors that influence quality education in Indonesia, especially elementary school education. However, this study relies on school accreditation as a dependent variable, which is one of the indicators of quality education in Indonesia. The analysis of the school accreditation system in the framework of internal bureaucratic reform of the Ministry of Education and Culture in 2011 showed that in order to improve the quality of national education gradually, it should be planned and measured in accordance with the mandate of Law No. 20 of 2003 CHAPTER XVI Part Two Article 60. It explains that accreditation is performed to assess the feasibility of the educational unit program, which establishes the feasibility of the educational unit program of primary and secondary education. In addition, accreditation becomes the quality assurance of education and gives recognition of the rating of the feasibility of the education unit. The evaluation process of all aspects of education should be directed at ensuring that qualified education services are organized and empowering that are being evaluated to produce graduates in accordance with the established educational standards. The accreditation process is conducted periodically and openly with the aim to assist and empower educational programs and units to be able to develop resources in achieving national education objectives. Some of the data supporting accreditation at the level of basic school with the national standard of education encompass content standard, process standard, graduate competency standards, educator and educational personnel standards, facility and infrastructure standards, management standards, financing standards, and assessment standards. The accreditation process implements objective, comprehensive, transparent, accountable, and professional principles so it can be used as an assessment to establish the feasibility or quality of the educational unit.

There are studies conducted to measure factors affecting the quality of education. A study by Taneri and Demir (2011) was conducted to measure the quality assessment of education of an elementary school in Kalecik, Ankara. The result shows that both teachers and students were satisfied with the condition

of the school building. A lack of technology and parents' involvement affect the education quality. Moreover, the teachers stress achieving a good grade rather than cooperation among students, which also has an impact on education quality. Thus, the factors that affect education quality are school facilities, the teacher's characteristics, school climate, quality education provided, parents' involvement, curriculum implementation, and the availability of technology and teaching materials. Barret, Chawla-Duggan, Lowe, Nickel, and Ukpo (2006) classify education quality into five dimensions: effectiveness, efficiency, equality, relevancy, and sustainability. Effectiveness is related to the output of education and refers to the extent to which the education system objectives are achieved. There is a relationship between efficiency and consideration of the input required to fulfill the output. Equality relates to human rights. Sustainability is related to all inputs or considerations for the future. Hence, qualified education is mandatory to build and preserve the condition of everyone regardless of gender, race, tribe, and religion. It emphasizes that a human being not only works but also learns throughout life. All the dimensions work as the basis for analyzing education quality.

Based on the above problem, it is necessary to make a measurement in the analysis of factors affecting education quality. It will contribute to developing and improvement of the education system especially in Indonesia. This study uses a robust approach to deal with the outliers that often appear in the actual data in order to improve the accuracy of the results, as compared to the traditional OLS regression. This approach is widely used in various fields of studies.

Research Methodology

This study involved 34 provinces in Indonesia that organize elementary school education as samples of research with eight independent variables to analyze the factors that affect the quality of education, which is represented by the number of elementary schools with A-accreditation. They are Indonesia's major political subdivisions, namely Special Capital Region of Jakarta, West Java, Banten, Central Java, Special region of Yogyakarta, East Java, Special Region of Aceh, North Sumatra, West Sumatra, Riau, Riau Islands, Jambi, South Sumatra, Bangka Belitung Islands, Bengkulu, Lampung, West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan, North Kalimantan, North Sulawesi, Gorontalo, Central Sulawesi, South Sulawesi, West Sulawesi, Southeast Sulawesi, Maluku, North Maluku, Bali, West Nusa Tenggara, East Nusa Tenggara, Papua, and West Papua. The independent variables are the number of schools, the number of teachers, the

number of students, the number of classes, the number of libraries, the number of laboratories, the number of graduates, and the average of national exam, which are adjusted to the instruments of education quality in Indonesia as mentioned before.

This study was ex post facto research with a correlation type, which is data collection action that aims to analyze the level of relationships between two variables or more. Ordinary least squares (OLS) regression is the most popular procedure to find out the effect of an independent variable on a dependent variable. Although this method is very powerful, most researchers believe that it is very sensitive to the existence of outliers. However, a robust approach is applied to the regression method in order to overcome the weakness of OLS regression related to the outliers.

There are several methods of estimation in robust regression such as M-estimation (maximum likelihood type), MM-estimation (method of moment), S-estimation, Least Trimmed Square (LTS), and Least Mean Square (LMS), but the most commonly used parameter for estimating regression is the first method. Following a linear model by Fox (2002) below:

$$\begin{aligned} y_i &= \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i \\ &= \mathbf{x}'_i \boldsymbol{\beta} + \epsilon_i \end{aligned} \quad (1)$$

for the i th of n observations. The fitted model of the equation (1) is as follows:

$$\begin{aligned} \hat{y}_i &= \alpha + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_k x_{ik} + e_i \\ &= \mathbf{x}'_i \mathbf{b} + e_i \end{aligned} \quad (2)$$

Where $e_i = y_i - \hat{y}_i$. Let $\rho(\cdot)$ shows the contribution of each residual to the objective function. Then, the general M -estimator to obtain the \mathbf{b} estimates is to minimize the objective function:

$$\sum_{i=1}^n \rho(e_i) = \sum_{i=1}^n \rho(y_i - \mathbf{x}'_i \mathbf{b}) \quad (3)$$

When the derivative of ρ represents the influence of curve, which is denoted by $\psi = \rho'$, then a system of $k + 1$ estimating equations for the \mathbf{b} coefficients is produced by differentiating the objective function with respect to the \mathbf{b} coefficients and then setting the partial derivation to zero:

$$\sum_{i=1}^n \psi(y_i - \mathbf{x}'_i \mathbf{b}) \mathbf{x}'_i = 0 \quad (4)$$

This equation is solved by minimizing $\sum w_i^2$ using iteratively reweighted least squares (IRLS). The estimation procedure using M-estimator can be explained as follows:

1. Calculate estimates $\hat{\beta}$, by using *least square* to obtain \hat{y}_i .
2. Calculate residuals $e_i = y_i - \hat{y}_i$ for $i = 1, 2, \dots, n$. Where \hat{y}_i and e_i are treated as initial values while y_i is the result of observation.
3. From these residuals, calculate $\hat{\sigma}$ and weights $w_i = \frac{\psi(e_i^*)}{e_i^*}$ with $e_i^* = \frac{e_i}{\hat{\sigma}}$ is the standardized residual values.
4. Apply weighted least square (WLS) to calculate the estimates of robust regression coefficient $\mathbf{b} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{Y}$.
5. Iterate step 2 through step 4 until $\sum_i^n e_i$ is convergent, in other words $\sum_i^n e_i$ is small enough.
6. If a standardized value is derived from e_i then based on the simulation conducted by Huber (1967) the selected value $k = 1.345\sigma$, so the following equation is obtained:

$$\rho_H(e) = \begin{cases} \frac{1}{2}e^2 & \text{for } |e| \leq k \\ k|e| - \frac{1}{2}k^2 & \text{for } |e| > k \end{cases}$$

which is the objective function of Huber's estimation. While the weight function is as follows:

$$\rho_H(e) = \begin{cases} 1 & \text{for } |e| \leq k \\ \frac{k}{|e|} & \text{for } |e| > k \end{cases}$$

where k is tuning constant. It is important to note that smaller values of tuning constant lead to more resistance to outliers at the expense of lower efficiency when errors are normally distributed.

Research Results

The presented study used secondary data derived from primary school education data from the Ministry of Education and Culture of Indonesia in 2014/2015. It is a government ministry that organizes early childhood education, elementary education, secondary education and community education affairs and the management of culture in Indonesia. This study focuses on elementary education. To understand the factors affecting the quality of education in Indonesia, this study used all the 34 provinces in Indonesia as samples. Table 1 shows descriptive statistics of these variables.

Table 1. Descriptive statistics

Variables	Minimum	Maximum	Mean	Std. Deviation
A-Accredited (y)	11.00	3,188.00	344.82	706.52
School (x_1)	451.00	19,817.00	4,338.62	5,105.10
Teacher (x_2)	5,641.00	227,876.00	49,863.21	58,203.78
Student (x_3)	77,694.00	4,611,703.00	768,592.38	966,253.40
Classes (x_4)	128.00	139,525.00	27,246.12	31,221.76
Library (x_5)	243.00	11,973.00	2,535.09	2,869.81
Laboratory (x_6)	2.00	727.00	125.53	200.46
Graduates (x_7)	10,979.00	767,720.00	128,511.15	166,566.08
Exam (x_8)	16.98	22.66	20.10	1.54

The dependent variable (y) is the number of A-accredited schools. There are eight (8) independent variables (x) used in this study: the number of schools (x_1), the number of teachers (x_2), the number of students (x_3), the number of classes (x_4), the number of libraries (x_5), the number of laboratories (x_6), the number of graduates (x_7) and the average of national exam (x_8). Figure 2 shows the outliers that appeared on the actual data.

From Figure 2 it can be seen that most of the variables used in this study are contaminated with outliers. There are extreme data whose values are far away from the others. These values may lead to bias of estimation results that does not describe the relationship between dependent and independent variables accurately. Other than that, it has a big possibility to cause large residuals of regression estimation. Tables 1 and 2 present the coefficients estimate of OLS regression and M-estimator of robust regression.

Figure 2. Outliers

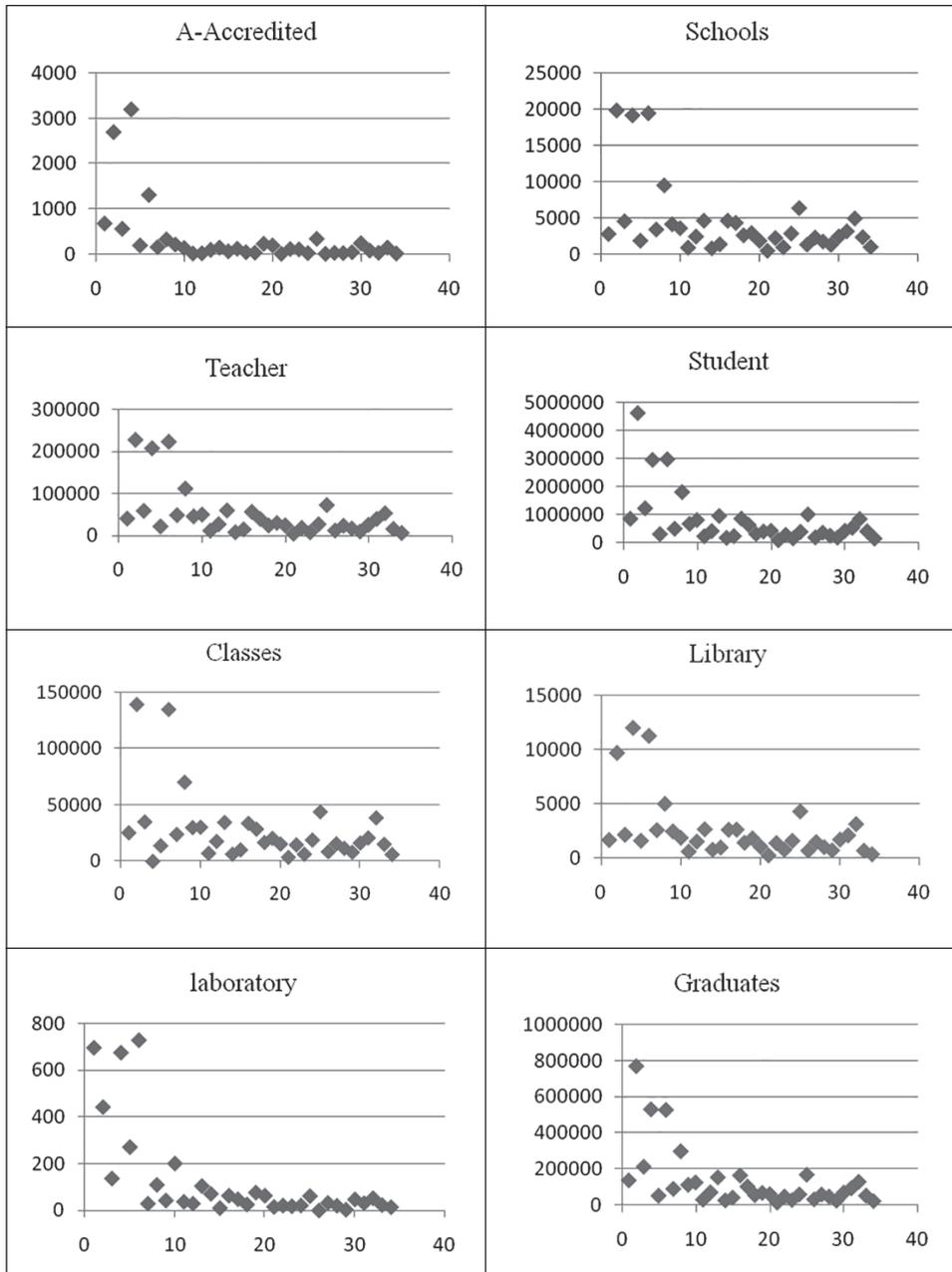


Table 2. OLS Regression

Value	Estimate	Std. Error	t-value	Pr(> t)
(Intercept)	-434.9000	343.8000	-1.2650	0.2176
x1	0.1068	0.0757	1.4110	0.1706
x2	-0.0192	0.0059	-3.2760	0.0031
x3	0.0003	0.0005	0.6750	0.5057
x4	-0.0125	0.0016	-7.6390	0.0000
x5	0.0925	0.1077	0.8590	0.3986
x6	0.6637	0.1646	4.0320	0.0005
x7	0.0050	0.0033	1.5340	0.1376
x8	19.8700	17.7100	1.1220	0.2725

This study presents the estimation results of OLS regression as a comparison to the M-estimator of robust regression in order to see clearly these two methods in analyzing the relationship between the number of A-accredited schools and these independent variables are determined. Table 2 indicates the OLS regression estimates, which show there are two independent variables with negative results, namely the number of teachers and the number of classes.

Table 3. M-Estimator of Robust Regression

Value	Estimate	Std. Error	t-value	Pr(> t)
(Intercept)	-462.7354	270.7553	-1.7091	0.0998
x1	0.1445	0.0596	2.4249	0.0229
x2	-0.0223	0.0046	-4.8266	0.0001
x3	0.0006	0.0004	1.6139	0.1191
x4	-0.0119	0.0013	-9.2926	0.0000
x5	0.0963	0.0848	1.1356	0.2669
x6	0.6439	0.1296	4.9671	0.0000
x7	0.0031	0.0026	1.2278	0.2310
x8	20.5528	13.9474	1.4736	0.1531

Table 3 shows the results of robust regression with M-estimator, which provide similar results to OLS regression where both the number of teachers and the num-

ber of classes have negative results. Further, robust regression provides smaller standard errors for all variables as well as smaller p -values.

Discussion

The analysis of the original data shows that there are outliers where these values are far away from the other set of sample data (cf., Figure 2). In fact, the data contaminated by outliers tend to cause a large residual, which can influence the estimation results of OLS regression due to the presence of biasness on the coefficients of regression as well as inconsistency. Therefore, this study uses robust regression to down-weight the outlier influences. However, this study provides the M-estimator method, which is the most popular among other methods of robust regression.

The empirical results reveal that the residual standard error in OLS regression is 114, while the robust residual standard error is 60.06. The difference of the values between OLS and robust regression is quite high, which indicates that the model parameters are being highly influenced by outliers. Basically, a residual standard error is the estimated standard deviation of the residuals. Further, R -squared is 0.9803 in OLS regression with the degree of freedom 25, while it is 0.9980 in M-estimator of robust regression. R -squared value represents the coefficient of determination, which is defined as the ability of all independent variables in explaining the variance of the dependent variable. In other words, it is a statistical measure of how close the data are to the fitted regression line. The results show that the ability of independent variables explaining the variance of its dependent variable amounts to 98.03% in OLS regression whereas M-estimator provides 99.80% of independent variables ability to explain the variance of its dependent variable. In general, the higher the R -squared indicates the better model fits the data. Therefore, it can be said that the M-estimator of robust regression gives the best observed outcomes that are replicated by the model. Further, M-estimator indicates that only 0.20% of the variance of the dependent variable is explained by other factors.

Based on Tables 2 and 3, it can be seen that the values of standard error in the M-estimator of robust regression for all independent variables are smaller than the values in OLS regression. The coefficient standard error is used to measure the average amount that coefficient estimates vary from the actual average value of the response variable so the lower number of its coefficients means better results. Therefore, M-estimator is better used to address the outliers in order to

analyze the effect of independent variables (the number of schools, the number of teachers, the number of students, the number of classes, the number of libraries, the number of laboratories, the number of graduates and the average of national exam) on dependent variable (the number of A-accredited schools). Moreover, M-estimator provides a greater degree of the accuracy of the estimate parameters. The t -value represents the ratio of the coefficient estimate to the standard error of the respective predictor, which is used to test against a two-sided alternative hypothesis that the coefficient is not equal to zero. Then, p -value is the probability of the t -value that is observed under the null hypothesis. When p -value is less than the confidence level, then the null hypothesis is rejected, which indicates that the parameter estimate is considered to be statistically significant at that alpha level. For example, the t -value for the independent variable of the number of schools from M-estimator analysis is 2.4249 (0.1445/0.0596) whereas the p -value of its variable is 0.0229. By using the alpha level at 0.05 on 25 degree of freedom, it can be safely concluded that the regression coefficient for the number of schools was found to be statistically different from zero due to rejection of the null hypothesis.

Turning to the coefficients estimates, which indicate the unknown constants that represent the terms of intercept and slope in the linear model; these coefficients represent the values for the regression equation for predicting the dependent variables from the independent variable. The coefficients estimates of M-estimator are as follows the coefficient of the number of schools is 0.1445, which means that by holding all other independent variables constant, for every increase in this variable, it is expected to increase by a 0.1445 unit in A-accredited schools. Then, the coefficient of the number of A-accredited schools is expected to decrease by a 0.0080 unit for every unit increase in the variable of the number of teachers, whereas every unit increase in the number of students is expected to increase a 0.0006 unit in the number of A-accredited schools. Further, it is expected to decrease a 0.0119 unit in the number of A-accredited schools for every unit increase in the number of classes, while every unit increase in the number of libraries is expected to increase a 0.0963 unit in the number of A-accredited schools. Every unit increase in the number of laboratories is expected to increase a 0.0963 unit in the number of A-accredited schools, whereas it is expected to increase a 0.6439 unit for every unit increase in the number of laboratories. Then, every unit increase in the number of graduates is expected to positively influence the number of A-accredited schools by a 0.0031 unit. Meanwhile, every unit increase in the average of national exam has the biggest influence on the dependent variable, where every unit increase in this independent variable is expected to increase the number of A-accredited schools by a 20.5528 unit. Thus, there

are only two independent variables (the number of teachers and the number of classes) that have a negative effect on the quality of education in Indonesia, which is represented by a number of A-accredited schools.

Conclusions

The presented study investigated the factors that affect the quality of education in Indonesia by using robust regression. This method is very powerful to down-weight the influence of outliers that often appear in original data. However, OLS regression result is presented as a comparison with the M-estimator of robust regression. The empirical results reveal that M-estimator improves the accuracy of the estimate parameters of regression. Therefore, this method is used for analyzing the independent variables that may affect the quality of education in Indonesia.

Acknowledgements

The authors would like to thank the Secretariat Directorate General of Basic Education and the Directorate of Primary School of the Ministry of Education and Culture of the Republic of Indonesia, who allowed the authors to access data for analysis purposes. They would also like to thank the Editors for their generous comments and support during the review process.

References:

- Barret, A.M., Chawla-Duggan, R., Lowe, J., Nikel, J., & Ukpo, E. (2006). The Concept of Quality in Education: A Review of the International Literature on the Concept of Quality in Education. EdQual Working Paper No 3. From http://www.edqual.org/publications/workingpaper/edqualwp3.pdf/at_download/file.pdf
- Development Progress. (2014). Towards better Education Quality: Indonesia's Promising Path. London: Development Progress. <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9066.pdf>
- Directorate General of Basic and Middle Education. (2016). Petunjuk Pelaksanaan Penjaminan Mutu Pendidikan Oleh Pemerintah Daerah. <http://pmp.dikdasmen.kemdikbud.go.id/files/docs/05.pdf>. Accessed on 22 October 2016.
- National Education Standards. <http://bsnp-indonesia.org/>
- Education For All (EFA). A Report on Education for All (EFA) Achievements Period 2000-2015. Educational For All Coordination Forum Republic of Indonesia.
- Fox, J. (2002). Bootstrapping Robust regression models: Appendix to an R and S-PLUS.

- Companion to applied Regression [online]. <<http://cran.r-project.org/doc/contrib/FoxCompanion/appendix-bootstrapping.pdf>>.
- Huber, P. J. (1967). The Behaviour of Maximum Likelihood Estimates under Nonstandard Conditions. *Proceedings Fifth Berkeley Symposium Mathematic, Statistics and Probability 1*, 221-233. <http://econ.ucsb.edu/~doug/245a/Papers/Huber%20Behavior%20of%20ML%20Estimates.pdf>
- Lord, H. L. (1973). *Ex Post Facto Studies As a Research Method*. Special Report No 7320. <http://files.eric.ed.gov/fulltext/ED090962.pdf>. Accessed on 30 May 2017.
- Ministry of Education and Culture of the Republic of Indonesia (MoEC). (2006). *Indonesia educational statistics in brief 2005/2006*. Jakarta: Ministry of National Education.
- MoEC. (2011). *Kajian Analisis Sistem Akreditasi Sekolah/Madrasah Dalam Rangka Reformasi Birokrasi Internal*. <https://luk.staff.ugm.ac.id/atur/rbi/AkreditasiSekolah-Madrasah.pdf>. Accessed on 22 October 2016.
- MoEC. (2013). *Overview of the Education Sector in Indonesia 2012: Achievements and Challenges*. Jakarta: MoEC.
- MoEC. (2016). *Petunjuk Pelaksanaan Penjaminan Mutu Pendidikan oleh Satuan Pendidikan*. <http://pmp.dikdasmen.kemdikbud.go.id/files/docs/03.pdf>. Accessed on 22 October 2016.
- Organization for Economic Co-operation and Development [OECD]. (2012). *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*. Paris: OECD. <http://www.oecd.org/education/school/49620025.pdf>
- OECD. (2012). *PISA 2012 Results in Focus What 15 Year Olds Know and What they Can Do With What They Know*. <https://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf>. Accessed on 30 May 2017.
- Taneri, P.O., and Demir, C.E. (2011). *Quality of Education in Rural Schools: A needs Assessment Study (Ankara-Kalecik Sample)*. *International Online Journal of Educational Sciences*, 3 (1), pp: 91-112. http://iojes.net/userfiles/Article/IOJES_356.pdf
- UNESCO. (2014). *Education For All Global Monitoring Report*. Paris: UNESCO. <http://en.unesco.org/gem-report/allreports>
- UNICEF. (2000). *Defining Quality in Education*. The International Working Group on Education, Florence, Italy, June 2000. <https://www.unicef.org/education/files/QualityEducation.PDF>