Eurasian Journal of Educational Research 76 (2018) 41-72

Eurasian Journal of Educational Research
www.ejer.com.tr

Investigation of School Effects on Student Achievement in Primary Education Using Value-Added Assessment[1]

Filiz KOC[1]

ARTICLE INFO

Article History:
Received: 10 Mar. 2018
Received in revised form: 21 Jun. 2018
Accepted: 11 Jul. 2018
DOI: 10.14689/ejer.2018.76.3

Keywords
Assessing School Performance,
Value-Added Assessment,
Simple Fixed-Effect Model,
Level Determination Exam

ABSTRACT

Purpose: The purpose of this study is to assess the contribution of primary schools within the metropolitan municipality of the province of Ankara, Turkey to the achievement outcomes of 7th grade students using the results of the end of year Level Determination Exam. **Research Methods:** Carried out using a casual comparative study framework, the population of this study consists of the primary schools in the province of Ankara. The sample includes the 24 primary schools whose students have attended to the same school in both the 2007-2008 and 2008-2009 educational years within the territorial jurisdiction of the metropolitan municipality of capital of Turkey. The value-added effects of the schools in the sample on the student growth are assessed using a simple fixed-effect model. Moreover, in order to determine whether or not there exists a statistically significant relationship between the rankings of the schools according to average student achievement levels and the rankings according to the value-added effects on student growth, Kendall tau rank correlation coefficients are calculated. **Findings**: The results of this study indicate that there are significant inconsistencies between the rankings of the schools according to their value-added effects on student improvement and the rankings according to the average student achievement, the latter being the method frequently used to assess the performance of the schools in Turkey. Moreover, the results demonstrate that the value-added effects of the schools on student improvement differ drastically from subject to subject. **Implications for Research and Practice:** It is expected that this research will lead to a more balanced evaluation of schools particularly given the likely emergence of more data over the years. In addition, this is the first value added assessment study carried out in Turkey. It points out that the way Turkish schools are assessed is problematic and suggests that value added methods should be considered in evaluating the effects of schools.

[1] This study is an outcome of the authors Masters Thesis under supervision of Asst. Prof. Ömer KUTLU

[1] TURKEY, e-mail: filizcinkirkoc@gmail.com, ORCID: https://orcid.org/0000-0002-0989-3461

Introduction

In today's information society, organizations have constantly reinvent??? themselves because of the international competition brought by the globalizing world. As a social institution, it is necessary for schools to adapt to changing social conditions in order to increase the quality of education and to improve their students both academically and sociall. For this reason, educating individuals who are able to use knowledge, adapt to the times and develop themselves has been the main task of schools. It is known that there are differences in the success levels of schools in terms of fulfilling these aims. Identification of the causes of the resulting success and the promotion of less successful schools to the level of successful schools are important to increase the quality of education as a whole. The beginning of this process is the determination of what makes a successful school.

Families want to enroll their children in schools that offer good education. But how is a "good school" determined? Is there a right method of identification? For many the method used is closely connected to student achievement, which in Turkey is associated with the success of the nationwide exam. This may be the case where an association and an appraisal method may not be fair and objective. This is because student success relies not only on the contribution made by the school and the teacher, but also depends on such factors as the socio-economic level of the family, the academic background of the student, and the knowledge of the student. The evaluation of the contributions of schools and teachers by scientific methods without the influence of such factors is an important research topic in educational sciences.

It is arguable that what a teacher or school is actually responsible for is the change in student attainment during their period within the school community. In this context, the real responsibility of the school or teacher should be "how much has the school/teacher contributed to student success" regardless of the student's socioeconomic level, intellectual background. It is arguable that the main objective of education is to determine the academic performance of the students and to maximize their achievement (Evergreen Freedom Foundation [EFF], 2008).

Taking this aim into consideration, successful education maximizes the academic performance of the students. When evaluating schools in this context, it would be more appropriate to assess schools based on their contribution to academic success of student (US Department of Education, 2009).

In the United States and England, the leading countries in the field of school and teacher evaluation in educational sciences, Value Added Assessment (VAA) approaches have been developed so that school and teacher influences can be assessed scientifically and objectively without any other effects. In these countries, the success of the students is followed by the year-end exams similar to Level Determination Exam (LDE, SBS in Turkish) in Turkey; but school and teacher evaluations use VAA approaches instead of average achievement scores. VAA is the general name of complex statistical techniques aimed at predicting the individual causal effects of

teachers or schools on student achievement, using test results from more than one year (McCafrey and Hamilton, 2007).

In VAA methods, each student forms his own control group. The extracurricular factors that influence the success of the pupils (socioeconomic status -SES, the educational status of the family, etc.) do not significantly change over subsequent years. The scores that a student receives from two equivalent exams (in two successive years) are influenced by both school-related and non-school factors. However, non-school factors can be considered to be equally effective in both exams, as non-school factors such as SES, do not change significantly over a couple of years. In this case, the amount of change in student achievement (development) is the result of school-related factors. This approach constitutes the basic principle of the VAA. Students are followed through repeated exams over the years and the school is held responsible for the change in student's success.

The idea of evaluating teacher effectiveness based on the development of students was first proposed by Eric Hanushek in 1971 (Hanushek, 1971). Subsequent studies have shown that the difference of teacher effectiveness is statistically significant (Hanushek, 1972, Murnane, 1975). Following the 1980s, work on VAA was accelerated, particularly in the US and UK educational sciences literature (eg, Department of Education and Science, 1983, 1984, Gray and Jesson, 1987, Gray et al., 1984, 1986; Woodhouse and Goldstein, 1988; Willms, 1987). Studies in this field show that although socioeconomic status of pupils is a significant factor for student success, it does not influence the increase in the success level significantly (Sanders, 1997). The effects of the teacher and the school are much more dominant than the student related background factors including the socioeconomic status (Sanders and Rivers, 1996; Wright, Horn and Sanders, 1997). Studies examining the effectiveness of schoosl and teachers using the VAA approach have continued to take place after 2000 (eg Kane and Staiger, 2008; Ishii and Rivkin, 2009; Hanushek and Rivkin, 2010). The most well-known example of CRB is the Tennessee Value Added Assessment System (TVAAS) (Sanders and Horn, 1994, 1998). The Dallas Independent School District has a system that examines teacher influences and uses their results in an official teacher assessment (Webster and Medro, 1997). It has become one of the most common methods used to assess the effectiveness of teachers and schools in the US, especially with the Race to The Top program, which has distributed upwards of $ 4.5 billion of federal aid to States who use VAA to assess teacher and school performances (US Department of Education, 2009).

According to McCaffrey et al. (2003), there are two reasons why more and more attention is paid to value added assessment studies. First, VAA distinguishes the impact of school and teacher, irrespective and independent of the other strong factors in student success. These non-school related factors include, but are not limited to, the socioeconomic status of the family, race, and family history. This differentiation is very important for schools and teachers to be assessed in a robust and fair way. The second reason is that VAA studies show that the teacher effect can change a lot from one teacher to another. If these differences can be embodied and some characteristics of the teachers can be connected within a cause-effect relationship, significant improvements

can be made in the education system. The success of schools in Turkey is generally determined by their performance in the exams they enter the transition from one school level to an upper level (Basaran and Cınkır, 2013). It is therefore difficult to say that school, teacher and student achievements have been correctly and fairly determined because there are different factors that influence the success of students.

In Turkey, discussions are continuing on both the high school and university entrance systems (SBS, OKS, TEOG, LGS for high school; OSS, OYS, LYS, YGS for university). Political parties offer promises in their education programs that they will either cancel these exams or apply alternative exams. It is arguable that what Turkey really needs is that as the alternative systems being developed, the assessment of teachers and schools should also be taken into account. It is envisaged that this study will contribute to the evaluation of school and teacher assessment in the search for alternative systems.

To date, in Turkey, schools are assessed with respect to the average of their students' successes in the nationwide exams. When evaluating schools in this context, it would be more appropriate to assess school students based on their contribution to academic development (US Department of Education, 2009). Many states in the United States have made it mandatory to use VAA in evaluating teachers and schools, as assessments using VAA are thought to be fairer. There is no study (in academia or practice) related with VAA in Turkey, except for the masters thesis of the author of this study. In this context, this is the first study carried out in Turkey for Turkish educational system. This study is intended to be a cornerstone for more VAA studies to carry out in Turkey.

Common Problems in Measurement and Evaluation

One of the most common problems that educators encounter and discuss in the education system is the evaluation of student achievement in the school and the objective assessment of schools by means of correct measurement.

The score that the pupil gets from a test is not simply due to the academic accumulation he earns in his current institution. A student's success in a test may be based on two different accumulations: 1) the academic background of the student, and 2) the academic background of the school in which the student is enrolled during the test. Moreover, the accumulation of the student's previous academic life may have determined the school in which he is registered. In Turkey in order for a student to qualify for enrollment in some school (e.g., Science high schools), he/she must demonstrate a certain level of success in the nationwide high school entrance examinations. Those who have high scores in these tests are enrolled in Science high schools. In other words, the academic background of students enrolled in Science high schools is much better than those studying in other high schools. This is because Science high schools use academic achievement as a prerequisite for enrollment. Therefore, in the university entrance exam Science high school students have much higher achievement than ordinary high school graduates. In this context, it is not true that the achievements of the students are assessed as a school-based achievement and that these schools provide a more successful education. This type of assessment is

largely attributed to the enrollment conditions of schools and to the high pre-knowledge of students.

Turkey has been part of many international student assessment programs including PISA (Program for International Student Assessment), TIMSS (Trends in International Mathematics and Science Study), and PIRLS (Progress in International Reading Literacy Study. These projects enable participating countries to evaluate their own education systems and to follow the development of the knowledge and skills of students in the fields of mathematics, science and reading skills according to years. It is not intended for a competition amongst countries. It is expected from the countries to follow the effects of these reforms by ensuring the necessary education reforms and the regular participation in the projects in the country by means of the conclusions (Ministry of National Education of Turkey, 2010). Many studies have been conducted by the MoNE on contextual indicators that relate student and school characteristics to performance outcomes using the data of these studies.

There are many studies that evaluate and interpret the results of these internationally held exams in this way, and compare how the achievements of schools change with respect to each other, or over the years. In all these studies, the average of the scores of students in that school on a national or international (eg PISA) exam is used as a sign of the success of the schools. The common result of studies comparing different school types with each other is that the most successful schools are Science high schools, which are followed by the various groups of high schools and the lowest in the order of success is the vocational high schools (eg Berberoğlu and Kalender, 2005, Farmer 2006, Demir, Kilic and Depren, 2017). The main point that is ignored in these studies is that students are compared to each other without paying attention to socio-cultural characteristics, students pre-knowledge. There is no control group in these evaluations. A group of students is compared with another group of students who are completely different. The main difficulty in evaluating the schools is the fact that the students are being evaluated and sorted, and the evaluation is done by ignoring some major factors that affect student achievement. However, the most important factor in student success is the student himself or herself, with all background knowledge, SES, accumulated educational history etc.

While such a school assessment approach is applied in this way in Turkey, there are different practices in other countries. There are still debates in the US on how exactly to implement an assessment strategy that uses standardized test results. One possible approach is the cohort-to-cohort comparison method, which has gained increasing interest among people engaged in educational policy for the last two decades. In this approach, successes of students in consecutive cohorts are compared against each other (eg 6th graders in the 2017-2018 school year versus 6th graders in the 2018-2019 year) and it is expected that the success of the educators will show an increasing tendency when passing from one cohort to another. In this method students are not followed individually; instead, all of a cohort is compared with past cohorts that have already enrolled in the same class (Ehlert et al., 2013). The method of the Public Schools Accountability Act adopted in California in 1999 includes such a benchmarking approach (McCaffrey, Lockwood, Koretz and Hamilton, 2003).

Early examples of studies for school and teacher effect on student achievement through scientific methods are seen in the 1950s, when trends such as science management and behaviorism gained strength in the United States (Ellett and Teddlie, 2003). Over the next 20 years, American educational scientists have conducted intensive research to identify effective teaching methods. These studies are usually based on observable teacher behaviors, observable student outputs, and a causal relationship. As in the United States, the most commonly used toolkit for determining school effectiveness in the UK since the 1980s is VAA (Department of Education and Science, 1983, 1984; Gray and Jesson, 1987; Gray et al., 1984, 1986; Woodhouse and Goldstein, 1988, Willms, 1987, Kurtz, 2018).

Value Added Assessment Method

Arguably the most common tool used to determine the contribution of schools to student development in recent years is the value added assessment method (VAA). In this approach, the rate of students' success in the school system is monitored, and various statistical methods are used to estimate the contribution of the teacher or school to the increase in the success of students. These methods have been associated with the value-added term borrowed from the literature of economics, as it relates to the contribution of the teacher or school from the moment the student enters the school system (McCaffrey et al., 2003). In the field of value-added assessment, the school or teacher is used to refer to the cumulative effect of education or to the original contribution of student development, independent of the student's own socioeconomic status and environment.

As mentioned before, although there are many factors that affect the academic success of the students, at least it can be said that it would not be fair to hold only schools or teachers accountable for success or failure. For example, the average academic success of a class of 25 students will increase in a nationwide exam if the least successful 5 students drop out the class. Similarly, if the top 5 students leave, this reduces the average score of the class. This change in the average cannot be regarded as success or failure of the teacher or school. This system, which shows student success as a school success, has led to years of unfair evaluation of schools. As a result, schools and teachers have been held unfairly accountable for the academic success or failure of students.

Research findings show that the main factors in student *development* are school and teacher, and other characteristics (socioeconomic level, background knowledge of the student, environment, etc.) are also statistically insignificant (EFF, 2008). For example, Sanders (1997) statistically determined that low socioeconomic level and academic development amounts are irrelevant.

In recent years, VAA has also become a focus of attention particularly in the United States, Canada, the United Kingdom and Australia. Because the evaluation of education as a system is demanded important stakeholders such as parents, educators, academicians, and politicians. To make this assessment, tools that can accurately and qualitatively measure the quality of education are needed. It is not enough to measure students individually in order to find a logical way of evaluating the effectiveness of

the education system. Objective and standardized tools are needed to measure how successfully schools and teachers are meeting their objectives with regards to student attainment. Examinations designed to evaluate the quality of education should be applied regularly and provide equivalent information. The results should provide an objective comparison. VAA is a candidate for hosting all of these features (EFF, 2008).

Studies in the literature show that the effect of teachers is more dominant than the socioeconomic level and other student growth factors (Sanders and Rivers, 1996; Wright, Horn and Sanders, 1997).

Different VAA studies measured the effects of individual teachers and shared these measurements with the school management to improve teacher performance. Among these studies, the most known is TVAAS (Tennessee Value Added Assessment System). Sanders and Horn (1994, 1998) examined the effects of teachers in the 4th and 8th grades of the state of Tennessee since 1996. The Dallas Independent School District also has a system that examines teacher influences and uses their results in the evaluation of official teachers (Webster and Medro, 1997). In a study conducted by William Sanders in the state of Tennessee, state-wide achievement tests conducted across the state found that students trained by effective teachers for 3 years had a score of 50 percent higher than students who attended in the class for 3 years.

It is possible to examine the literature on teacher effects under three main headings:

1) The influence of teachers is very important, and it is indeed arguable that the most important educational input on student development is the effectiveness of the teacher (Rivkin et al., 2000; Rowan, Correnti and Miller, 2002; Wright et al., 1997, Paufler, 2014).

2) Teacher influence is long-term and cumulative (Kain, 1998; Mendro, Jordan, Gomez, Anderson, and Bembry, 1998; Rivers, 1999;

3) Teacher efficacy differs from student to student according to achievement levels of students (Sanders and Rivers, 1996, Loeb, Soland and Fox, 2014).

The studies in the first group prove teacher effectiveness using different methods. The size of the impact and the relative prevalence have been compared with other factors. The studies in the second group are about the permanence of the student's influence in the later education periods of previous teachers. The fact that all of these findings are consistent with each other is proof that the teacher's influence is permanent. However, the magnitude of this persistence may be exaggerated (McCaffrey et al., 2003). The study in the third group shows that the students with the lowest success are the ones who benefit the most from a more effective teacher.

In Turkey, the schools are assessed by the performance of their students in a nationwide exam for passing to a higher level institution. However, both the results of international success projects (eg MoNE, 2003, 2010) and studies in the VAA field literature (eg EFF, 2008; Sanders, 1997) have shown that evaluating schools in this way reflects the socioeconomic status of many students in the success of schools. VAA is a type of method to be used to objectively set the success of schools in Turkey.

VAA distinguishes the contribution of the teacher or the school from the contribution of the student. Based on the academic background of each student in VAA practice, an assessment of this year's success is forecasted. The level of success predicted by a student at a normal level of education determines the level of success. It is important to note that achievement scores that are equal to or higher than the predicted grade are effective (Hershberg, 2008, Everson, 2017). On the contrary, teaching is not effective if the actual results are below the predicted values.

As an example, if a student is in the highest 15% in a nationwide exam in the last two years, it can be assumed that she will be in the upper 13% -17% this year (with a deviation of 2%, for example). If the student becomes in the upper 10% of the cohort, it can be said that the student has developed above normal. If all students in a school or classroom have made a development above normal, it can be said that the school or teacher achieves this successful outcome. We can say that the source of development that is above or below the expectation is the school or teacher. Because the demographic impacts on student achievement are the same for all years of comparison (Hershberg, 2004, 2008, Everson, 2017).

VAA cannot, by itself, explain the reason for student failure. However, when there is as much data, the educators can ask themselves questions about the failure or where the successes are observed. VAA helps teachers to measure the results of their teaching, to help teachers understand what their teaching is focused on (which students have benefited most), and their impact (how effective the year's development is in communicating to students). The development of student achievement may show different patterns according to class, subject, school and region. It is possible to examine these patterns from the results of VAA (Hershberg, 2004, Loeb, Soland and Fox, 2014).

Like every method, VAA has strengths and weaknesses. It takes many years to use the information obtained with VAA. At least a few years of data accumulation is needed to make a sound assessment of schools and teachers. It is very difficult to make a comment about the first year as it allows evaluation based on the development values of the VAA. Despite the strong scientific bases of VAA, it is difficult to apply because the statistical models it is based on are not easy to understand and explain. There are some computer programs available for the use of VAA in the countries where VAA is used and they are evaluating students, teachers and schools using these computer programs. For this reason, schools are required to allocate money and time in order to implement the VAA.

Despite all these difficulties, however, VAA provides very useful data. Managers can make plans for the future through the data provided by the VAA. In addition, successful schools and teachers can be identified, the key factors for their success can be searched, and these factors can be made available to other schools and teachers. If schools more developed schools are compared to other schools and the difference is located in the underlying hardware, school administrations may act to complete the lack of equipment. The results of the VAA allow teachers to be evaluated and followed

up. In some states in the United States, the VAA-based reward systems are used by the school districts (EFF, 2008).

The problem of this research is to determine the contribution of primary schools to student success by means of value added methods.

The purpose of this research is to determine the added value of the education given to the students who are studying in the 7th grade of the primary schools within the borders of the Ankara Metropolitan Municipality with the data obtained from the year-end Level Determination Examinations (SBS). For this purpose, we try to answer the following questions.

1. How much relative value is added by primary schools to their students?

2. Is the order of schools consistent in terms of student success and student development?

The aim of the study is to prove that there are inconsistencies between how the schools are assessed by the authorities and the public and the actual value added by the schools to its students.

Method

Research Design

The research uses a causal comparison model. In the study, the effects of primary schools in Ankara on student performance measured by the SBS is assessed by means of a Value Added Method.

Population and sample

The population of this research is the primary schools in Ankara- capital of Turkey, and the sample is composed of 24 primary schools within the borders of the Metropolitan Municipality, in which the students who studied both in the 2007-2008 and 2008-2009 years are in the same school.

Provinces that formed a sample in the first phase were Çankaya, Altındağ, Yenimahalle, Mamak, Keçiören, Gölbaşı, Sincan and Etimesgut. One of the purposes of the research is to understand the relationship between socioeconomic status of the students and the value added by the schools to these students. In the second phase, the three socioeconomic levels in these provinces are low, medium and high, determined by random selection method. In the third stage, the schools were ranked according to the socioeconomic level. In this order, schools with low socioeconomic level 3, schools with medium socioeconomic level 2, and schools with high socioeconomic level are numbered 1. The selection of the socioeconomic levels of the schools was made with the help of an expert working in the provincial directorate of national education in Ankara. One of the main reasons to select this city was that it fairly represents Turkey. Data for the 24 aforementioned schools have been requested,

with 22 schools complying. The study was conducted on data from these 22 schools. The names of the schools are provided in Table 6 in the Appendix.

The Level Determination Exam (SBS)

In Turkey, the level determination exam (SBS) is a central exam applied by the ministry of national education. Students are allocated to high schools with respect to the weighted average of multiple SBS exams applied at the end of 6th, 7th, and 8th grade. SBS exams started in 2008 to be applied only to 6th and 7th graders, and finalized in 2013. Other than the SBS system, high school entrance exams consist of a single exam at the end of 8th grade and the data is not suitable to use in a VAA study. The SBS exam consists of five subject areas: Turkish, mathematics, science, social sciences, and foreign Language-English. The number of right answers in each subject area is refereed to as *number right (NR)* score and the NM score which is calculated as NR minus the number of wrong answers/3 for each subject area is referred to as *negative marking (NM)* score.

Data Collection

The data for this study comprises the students who enrolled in the 6th and 7th grades in the schools within the borders of Ankara Metropolitan Municipality over the years 2008 and 2009. This data has been taken in the electronic environment with the necessary permission as a result of the correspondences made with the Ministry of National Education, EĞİTEK.

In VAA, data are not experimental but observational (McCaffrey et al., 2003). For this reason, the observational data are not encountered except for the problems to be experienced in collecting. Observational data is the fact that the largest problem data is not complete. Despite the fact that 24 of the records were requested from the EĞİTEK, 22 studies could be reached. The research group was selected as students who were in grades 6 and 7 in the same school. In this respect, data incompleteness problems that may arise from the students who are transported are prevented. It is also considered that the actual contributions of schools can be determined independently of the effect of school change on the student.

Analysis of the data was done using "SPSS 13.0 for Windows" and "Microsoft Excel 2007" programs. There are known true and false numbers in the SBS scores and subtests that all students in each school receive from two exams. The change in NM scores and true / NM scores of these students are calculated using Microsoft Excel. Then the average scores of the students in each school were calculated, which was stated as the school average. Some outliers were detected using z-scores approach. For example, students who did not attend one of the two exams did not participate in the analysis. A student may not have been able to raise any sub-test questions in the second test and may have shown very low success because he did not use the exam. This change is not caused entirely by the influence of the school.

Data Analysis

The main statistical models used in the VAA are examined under three main headings: auto-regressive approaches (Hanushek, 1972, Murnane, 1975, Rowan et al., 2002, Sanders and Rivers, 1996, Webster and Medro, 1997), development-based approaches (Allison, 1990, Bryk and Weisberg, 1976; Lord, 1969; Rogosa, 1995; Rowan et al., 2002; Thum, 2001) and multivariate approaches (Sanders, Saxton and Horn, 1997; Ballou, Sanders and Wright, 2004).

In the most well-known examples of VAA, a multivariate model that uses many years of accumulated data is used. In this study, we only used two years' data. This was mainly due to difficulty in obtaining data in Turkey.

In this study, a development based approach: simple fixed effect model is used in the study. The reasons for selecting this approach is because:

- this model contains fewer variables,
- the solution is easy to understand and explain,
- some studies (Tekwe et al., 2004) show that more complex models with this model give similar results,
- the availability of data that does not allow the use of more detailed models

There are many factors, both in-school and out-of-school, that affect a student's score from exams or subtest scores. The objective of the VAA approach is to determine the contribution of the school by eliminating the effects of extracurricular factors. A student's score can be modeled linearly as follows.

$$y_{ijst} = b_{st} + m_{ijs} + SED_{ij} + O_{ist} + e_{ijst} \quad (6)$$

In this model, the score of student j in the i-th school in subtest s at year t is related to b_{st} (a constant for test s, year t), to the influence of the socioeconomic level (SED), the school effect (O), and an error term (e), as a constant number, the effect of the incremental knowledge of the student on this subject (m). The error term here assumes a normal distribution with a mean of 0 and a standard deviation of σ_{ijst}. The following is the score obtained by a student in the sequential tests conducted in 2008.

$$y_{ijs,2008} = b_{s,2008} + m_{ijs} + SED_{ij} + O_{is,2008} + e_{ijs,2008} \quad (7)$$

$$y_{ijs,2009} = b_{s,2009} + m_{ijs} + SED_{ij} + O_{is,2009} + e_{ijs,2009} \quad (8)$$

It is not expected that the background knowledge and socioeconomic level of a student will change very much from one year to the next, and such a change is considered as an outlier. Information on the determination of such outliers and their discarding from the analysis is given in the section on analysis of data. Apart from this, the difference between the scores of students who have taken two years of top-up examinations will be as follows.

$$y_{ijs,2009} - y_{ijs,2008} = \left(b_{s,2009} - b_{s,2008}\right) + \left(O_{is,2009} - O_{is,2008}\right) + \left(e_{ijs,2009} - e_{ijs,2008}\right) \quad (9)$$

In the above equation, the SED and the incremental contributions (m) in the two-year period are not included in this equation. The reason is that it is assumed that the SED and the incremental contributions of the students do not change from year to year. Based on the knowledge that the increasing knowledge of students and their socioeconomic levels have not changed, it is observed that this difference is not related to the SED or the information (m) when the students score difference in the following years. Thus, each student forms his own control group and the contribution of the school is revealed in a way free from SED and student background. The above equation can be edited as follows.

$$\left(O_{is,2009}-O_{is,2008}\right)=\left(y_{ijs,2009}-y_{ijs,2008}\right)-\left(b_{s,2009}-b_{s,2008}\right)-\left(e_{ijs,2009}-e_{ijs,2008}\right) \quad (10)$$

By solving this equation system for all students, the contribution each school makes to its students in 2009 can be calculated. Because, in the above equation, the left side of the equation is the contribution of the school up to 2009 with the contribution up to 2008, which is the contribution the school has made to its students in 2009.

By expanding this method to include more years, a multi variable structure can be obtained. However, in this study, the simple fixed effect model, which is the most appropriate method, has been used.

Fixed Simple Effects Model:

In this method, in a sample of n schools, the change in student scores is modeled as follows:

$$d_{ijs}=\beta_{0s}+\sum_{k=1}^{n-1}\beta_{1ks}s_{kij}+e_{ijs} \quad (11)$$

Here,

$$d_{ijs}=y_{ijs2}-y_{ijs1} \quad (12)$$

y_{ijst}, is the score from test *s* of student *j* in school *i* at year 2009 (*t*=2)

$$s_{kij}=\begin{cases}1, & i=k, i\neq n\\ 0, & i\neq k, i\neq n\\ -1, & i=n, k=1,2,\cdots,n-1\end{cases} \quad (13)$$

$e_{ijs}\sim N(0,\sigma_{es})$, d_{ijs} is the score difference for student *j* at school *i* from test *s*.

The β_{0s} constant at Equation (11) corresponds to $\left(b_{s,2009}-b_{s,2008}\right)$ in Equation (9), and β_{1ks} is the value added to students of school k for test s, which corresponds $\left(O_{is,2009}-O_{is,2008}\right)$ in equation (9). In summary, if equations (11) are to be written

explicitly, the following equation system is obtained, assuming that the average contribution of all schools is 0.

$$\begin{aligned} d_{1js} &= \beta_{0s} + \beta_{1,1s} + e_{1js} \\ d_{2js} &= \beta_{0s} + \beta_{1,2s} + e_{2js} \\ &\vdots \\ d_{22js} &= \beta_{0s} + \beta_{1,22s} + e_{22js} \end{aligned} \quad (14)$$

Adding all equations side by side and taking the expected value results in

$$\beta_{0s} = \bar{d}_s - \mathrm{E}[e_s] \quad (15)$$

And as the expected error value is 0, $\beta_{0s} = \bar{d}_s$ .In other words, it is found as the mean changes of all schools in the sample in *s* test. In this case, the average contribution of all schools indicates the relative contribution of the school to this average. In other words, the contribution of some schools will be positive, others will be negative; because the average contribution of all schools is assumed to be zero. This is mainly because this method calculated the contribution of individual schools relative to the others. Hence, as the analysis is in relative terms there is no need for vertical scaling.

The average score change of each school and their averages were calculated by Microsoft Excel program and the variables in the simple fixed effect model were calculated after eliminating the outliers. Then, the standard deviation of the school contributions is calculated, and the "z-score" indicating how many standard points each school has deviated from the average is calculated for each school and subtest as follows.

$$z_{ks} = \frac{\beta_{1ks}}{\sigma_s} \quad (16)$$

Here, □ is the school effect for test *s*, and □s is the standard deviation of school effects. The z-scores are then used to classify the schools according to the following table (Tekwe et al., 2004).

Table 1

Using z-scores for Classification

z-score	Classification
$z \geq 2$	A
$2 \geq z \geq 1$	B
$1 \geq z \geq -1$	C
$-1 \geq z \geq -2$	D
$-2 \geq z$	F

After the calculation of the school contributions, the schools were ranked according to these contributions and the relationship between success rankings and contribution (development) rankings was analyzed. For this purpose, the degree of correlation between achievement orders and contribution (development) orders according to the 6th and 7th grade NR / NM numbers for each course was calculated according to kendall's tau rank correlation coefficient. SPSS 13.00 for Windows program was used to draw scatter diagrams that reflect the relationship between achievement rankings (NR / NM change for 6th and 7th grade) and contribution rankings (NR / NM change) for all courses separately.

Results

In this section, the analysis made based on the questions that are sought in the general purpose frame of the research, the findings and interpretations belonging to these analyzes are included. In the presentation of the findings, the order of the questions to be answered in the study was followed and the results of the analysis were given according to the courses.

How Much Relative Value is added by Primary Schools with respect to Their Students?

In this section, the interpretation of findings and interpretations are handled according to Turkish and Mathematics courses, respectively. There are two test scores for a student, 1: the number of question that the student *correctly* answered (NR score) and 2: the net score which is calculated by subtracting 1/3 of the wrong answers from the correct answers (NM score). The average NR and NM scores of schools for 2008 and 2009, calculated separately for each course, β (Beta) values calculated according to the fixed effect model, transformed into z-points and classified. Findings about social studies, foreign language and science courses can be found in the author's masters thesis.

Turkish Course

In Table 2, the NR and NM scores for Turkish test are calculated and analyzed. The first column of the table shows the socioeconomic status (SES) of the region where the schools are located. It was determined as 1: high, 2: medium and 3: low in this column. In the second and third columns, the students average NR and NM scores for 6th grade test is provided. The NR and NM scores in the 7th grade SBS are given in the fourth and fifth columns. The score difference between grade 7 and grade 6 was calculated as the score of change and was given in columns six and seven. Beta (β) school contributions are given in the eighth and ninth columns, and z-scores in the tenth and eleventh columns. The z-scores have been translated into the classification as given in Table 2 and shared in the last two columns of classification of schools in terms of NR and NM scores for Turkish course. The last two lines of the chart provide averages of school average scores and standard deviations.

Table 2

Turkish Course Statistics and Classification

		6th grade		*7th grade*		*Change*		*(β) School Contribution*		*z-score*		*Classification*	
SED	School number	NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score
1	School 1	16.934	16.325	16.231	14.791	-0.702	-1.534	0.682	0.914	1.762	1.807	B	B
2	School 2	14.511	13.179	13.538	11.240	-1.027	-2.011	0.357	0.438	0.923	0.865	C	C
3	School 3	12.127	10.135	10.363	7.208	-1.747	-2.903	-0.363	-0.454	-0.938	-0.898	C	C
1	School 4	13.801	12.362	11.590	8.768	-2.200	-3.581	-0.816	-1.133	-2.109	-2.239	F	F
2	School 5	12.560	10.716	11.197	8.250	-1.387	-2.498	-0.003	-0.049	-0.007	-0.097	C	C
3	School 6	10.160	7.510	8.532	4.728	-1.515	-2.603	-0.131	-0.154	-0.339	-0.305	C	C
1	School 7	13.775	12.308	12.050	9.500	-1.725	-2.808	-0.341	-0.360	-0.881	-0.711	C	C
2	School 8	16.236	15.455	15.554	13.878	-0.755	-1.676	0.629	0.772	1.625	1.526	B	B
1	School 9	15.370	14.353	14.096	12.075	-1.312	-2.327	0.072	0.121	0.187	0.240	C	C
2	School 10	13.582	11.941	11.963	9.280	-1.456	-2.464	-0.072	-0.016	-0.185	-0.031	C	C
3	School 11	11.706	9.833	10.486	7.352	-1.235	-2.490	0.149	-0.042	0.385	-0.082	C	C
1	School 12	15.379	14.422	14.235	12.183	-1.113	-2.197	0.271	0.251	0.700	0.497	C	C
2	School 13	12.347	10.427	11.180	8.322	-1.176	-2.055	0.208	0.394	0.539	0.778	C	C
1	School 14	13.931	12.561	12.502	10.040	-1.416	-2.503	-0.032	-0.055	-0.082	-0.108	C	C
2	School 15	11.789	9.649	10.895	7.886	-0.895	-1.763	0.489	0.685	1.265	1.355	B	B
3	School 16	13.527	11.938	11.450	8.592	-2.081	-3.347	-0.696	-0.899	-1.800	-1.777	D	D
1	School 17	14.339	13.018	12.875	10.423	-1.464	-2.595	-0.080	-0.147	-0.207	-0.290	C	C
2	School 18	12.391	10.465	11.166	8.300	-1.258	-2.205	0.126	0.243	0.327	0.481	C	C
3	School 19	13.378	11.685	11.824	9.069	-1.537	-2.515	-0.152	-0.067	-0.394	-0.132	C	C
1	School 20	15.386	14.426	14.261	12.285	-1.091	-2.102	0.293	0.346	0.758	0.684	C	C
2	School 21	14.638	13.438	12.813	10.401	-1.827	-3.037	-0.443	-0.588	-1.144	-1.163	D	D
3	School 22	12.301	10.393	10.829	7.824	-1.533	-2.650	-0.149	-0.202	-0.384	-0.399	C	C
Average of Averages		13.644	12.115	12.256	9.654	-1.384	-2.449	0.000	0.000				
Std. dev. of Averages		1.646	2.135	1.822	2.363	0.387	0.506	0.387	0.506				

NM Score: Correct answers, NR Score: Correctly answered

In Table 2, the 6th grade average Turkish NR score is 13.644 and the 7th grade average Turkish NR score is 12.256. There is a decrease of 1.384 points in the Turkish NR score for the sample. Similarly, while the 6th grade Turkish NM score was 12.115, it dropped to 9.654 with a decrease of 2.449 points. The standard deviation of the change scores is 0.387 and 0.506 for the NR and NM scores. The classification of schools is summarized in Figure 1. As shown in the figure, many schools are considered to be the average accepted class C by looking at their z-scores. While the average NR of schools in the sample decreased by 1.384, the average NR score of students with high SED School #1 decreased by only 0.702, which caused the school to be classified as class B.

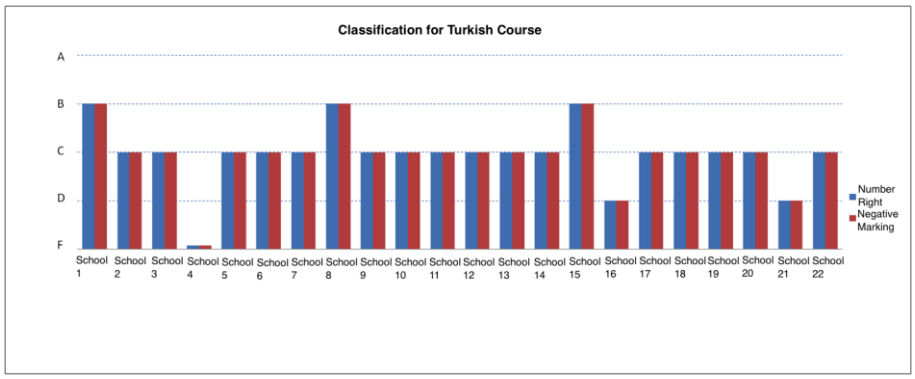

Figure 1. Classification for Turkish Course

Likewise, School #8 and #15, both of which are composed of students with high SES, have been classified as B. For School #21, which is low in SES and School #16 with medium SES, NR and NM scores in Turkish showed a decline more than the average. For these reasons, these schools were classified as class D. School #4, who had a high SES, was evaluated as class F. This school did not provide adequate improvement for its students with a NR average score of 13.801 and NM average score of 12.362 in the 6th grade. The NR score dropped by 2.720 in the 7th grade and to a 11.590 average (3.581 NM point reduction to 8.768). This decline was attributed to the fact that the school was classified as class F, with a considerable decline compared to all the schools in the sample. School #14, which has similar average 6th grade scores and same SES, provided an average improvement (-0.032 in the NR score, -0.055 in the NM score) whereas the contribution of the School #4 was -0.816 for the NR score and -1.133 for the NM score, respectively.

Mathematics Course

Table 3 corresponds to results of Mathematics course. The average NR score was 7.778 (NM score of 5.436) for the 6th grade and it is reduced by 2.396 points to 6.065 in the 7th grade. The average NM score reduced by 1.712 points to a level of 3.051. The

standard deviation of the change scores was 0.440 for the NR score and 0.499 for the NM score. The classification of schools is summarized in Figure 2. As shown in the figure, many schools are considered to be the average class C by looking at their z-scores. The average NR of all schools decreased by 1.712. The average NR score of School #6 students dropped by only 0.889, which led to this school being classified as class A.

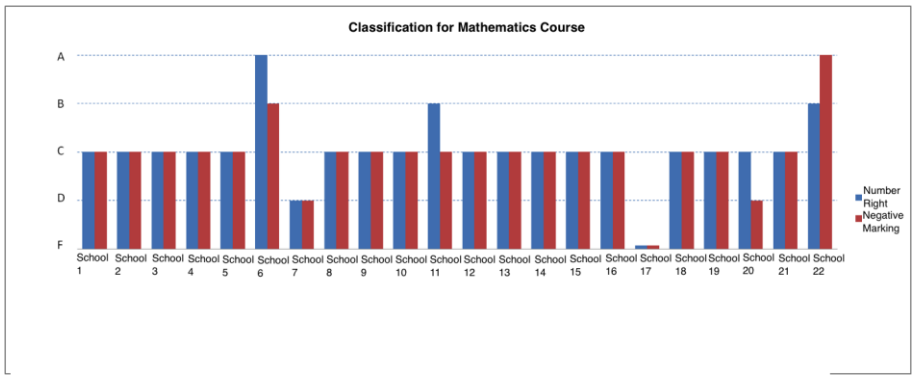

Figure 2. Classification for Mathematics Course

Likewise, School #11 and School #22, whose socioeconomic status are low, have been evaluated as A or B class in terms of their NR or NM scores. For School #7 and School #20, who have high SES, the NR and NM scores in Mathematics has declined above the average. For this reason, these schools were classified as class D. School #17, which is composed of students with high SES, is considered as F class. This school did not provide adequate contribution to the students in the 6th grade. The average NR score reduced from 9.018 to 6.268, and NM scores reduced from 7.042 to 3.494 in the 7th grade. These reductions (2.750 for NR and 3.548 points for NM scores) are at least two standard deviations away from the mean. This extremely high reduction in the points resulted in an F classification for this school in Mathematics course. The average scores of School #9, which is composed of students with the same SES, was close to School #17, in the 6th grade test for mathematics test. This school provided an average improvement for its students, whereas School #17, couldn't provide this improvement.

Is the Order of Schools Consistent in terms of Student Success and Student Development?

The differences between the NR and the NM scores of the schools in 2008 and 2009 for Turkish and Mathematics courses were calculated. In addition, a correlation analysis was performed between 2008 and 2009 student scores. The aim of the analysis is to investigate if successful students in first year continue to be successful in the second year. We also analyzed the relationship between average student success and average improvement of the students. Moreover, we compared the order of schools in terms of their average points in the first and second year with the average improvement they provide to their students.

The correlation between Turkish test NR scores for all the students in the 6th and 7th grades is 0.754. The correlation for the NM Turkish scores of both classes is found to be 0.756. These results show that there is a high positive correlation between Turkish NR and NM scores for the two consecutive years. The students who have high NM or NR score in the first year have a high NM or NR score in the second year. This is expected as students with relatively better background (who scored high in 6th grade) tend to be more successful in the 7th grade.

In the correlation analysis to determine the relationship between grade 6 points and change points, the change score of each student was found by subtracting grade 6 from grade 7 score. Correlation between students' sixth grades and change points (grades 7 - 6) was found to be -0.26 for the Turkish NR score and -0.28 for the Turkish NM score. This indicates that there is almost no (if not slightly negative) relationship between students' change scores and grade 6 scores. When combined with the high correlation between 6th and 7th grade scores, this result is extremely important. This shows that the change points are independent of students background.

In Figure 3, scatter diagrams are given for all pairs of Turkish scores 6th and 7th grade NR and NM averages and change averages. As can be seen in the figure, while there is a high level linear relationship between NR and NM scores in both 6th and 7th grade, there is almost no relation between those with the average change (improvement).

Table 4 examines the schools' consistency in terms of student average student success and student improvement. For the Turkish course, each school is ranked in the order of the 6th grade NR and NM average score, the 7th grade NR and NM average score, and the average NR and NM score improvement among the 22 schools in the sample.

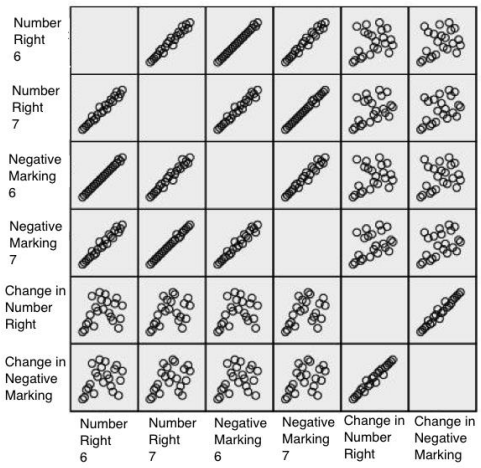

Figure 3. Average success and improvement for Turkish Course (all pairs)

Table 3

Mathematics Course Statistics and Classification

		6th grade		*7th grade*		*Change*		*(β) School Contribution*		*z-score*		*Classification*	
SED	School number	NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score
1	School 1	11.702	10.499	9.901	7.774	-1.802	-2.725	-0.090	-0.329	-0.223	-0.659	C	C
2	School 2	7.823	5.418	6.237	3.222	-1.643	-2.268	0.068	0.127	0.170	0.255	C	C
3	School 3	6.722	3.937	5.05	1.563	-1.646	-2.367	0.066	0.029	0.164	0.057	C	C
1	School 4	7.719	5.311	6.021	2.889	-1.677	-2.393	0.035	0.002	0.087	0.005	C	C
2	School 5	6.813	4.169	5.276	2.048	-1.392	-2.098	0.320	0.298	0.795	0.597	C	C
3	School 6	5.570	2.453	4.642	0.890	-0.889	-1.458	0.823	0.938	2.045	1.879	A	B
1	School 7	8.300	6.200	6.150	3.292	-2.410	-3.239	-0.699	-0.844	-1.736	-1.691	D	D
2	School 8	10.15	8.543	7.993	5.631	-2.094	-2.826	-0.382	-0.430	-0.951	-0.863	C	C
1	School 9	9.014	7.065	7.023	4.275	-1.991	-2.790	-0.279	-0.394	-0.694	-0.790	C	C
2	School 10	7.203	4.662	5.519	2.436	-1.582	-2.148	0.129	0.248	0.322	0.497	C	C
3	School 11	6.176	3.725	4.886	1.562	-1.294	-2.167	0.418	0.229	1.038	0.459	B	C
1	School 12	8.919	6.936	7.167	4.572	-1.739	-2.34	-0.027	0.056	-0.068	0.112	C	C
2	School 13	7.096	4.487	5.314	2.107	-1.817	-2.439	-0.105	-0.043	-0.262	-0.087	C	C
1	School 14	7.318	4.973	5.712	2.730	-1.598	-2.227	0.114	0.168	0.282	0.337	C	C
2	School 15	6.776	4.149	5.368	2.114	-1.408	-2.035	0.304	0.361	0.755	0.723	C	C
3	School 16	7.363	4.838	5.777	2.546	-1.545	-2.234	0.167	0.162	0.415	0.324	C	C
1	School 17	9.018	7.042	6.268	3.494	-2.750	-3.548	-1.038	-1.152	-2.581	-2.309	F	F
2	School 18	6.913	4.265	5.279	2.098	-1.693	-2.246	0.019	0.150	0.047	0.301	C	C
3	School 19	7.128	4.543	5.430	2.364	-1.695	-2.169	0.017	0.227	0.041	0.455	C	C
1	School 20	9.109	7.170	7.037	4.285	-2.075	-2.918	-0.363	-0.522	-0.902	-1.046	C	D
2	School 21	8.307	6.144	6.523	3.542	-1.787	-2.685	-0.076	-0.290	-0.188	-0.580	C	C
3	School 22	5.967	3.054	4.854	1.686	-1.131	-1.385	0.581	1.010	1.443	2.025	B	A
Average of Averages		7.778	5.436	6.065	3.051	-1.712	-2.396	0	0				
Std. dev. of Averages		1.445	1.879	1.213	1.554	0.402	0.499	0.402	0.499				

NM Score: Correct answers, NR Score: Correctly answered

When Table 4 is examined, it is clear that there are serious discrepancies between the successes and improvement orders of schools. For example, School #11, School #13 and School #15, which are left behind in the average success order, are in the higher rankings in terms of improvement. On the other hand, for the schools with high order rank with respect to average student success (School #9, School #17 and School #21), their order in terms of student improvement is very low. Some schools, such as School #1 and School #8, did not show any difference in terms of success and improvement. To analyze the difference of ranking on success and improvement, Kendall tau rank correlation coefficient between success and improvement rankings was calculated as 0.290 in terms of the 7th grade NR scores and 0.325 for NM scores. These values also indicate that there is a statistically low level of correlation between average student success and improvement ranks, and therefore these ranks are not consistent with each other.

Table 4

Turkish Course Success and Improvement Comparison

		6th grade		*7th grade*		*Change*	
SED	School number	NR score	NM score	NR score	NM score	NR score	NM score
1	School 1	1	1	1	1	1	1
2	School 2	7	7	6	6	4	4
3	School 3	19	19	21	21	19	19
1	School 4	10	10	13	13	22	22
2	School 5	15	15	15	17	11	12
3	School 6	22	22	22	22	15	16
1	School 7	11	11	10	10	18	18
2	School 8	2	2	2	2	2	2
1	School 9	5	5	5	5	10	9
2	School 10	12	12	11	11	13	10
3	School 11	21	20	20	20	8	11
1	School 12	4	4	4	4	6	7
2	School 13	17	17	16	15	7	5
1	School 14	9	9	9	9	12	13
2	School 15	20	21	18	18	3	3
3	School 16	13	13	14	14	21	21
1	School 17	8	8	7	7	14	15
2	School 18	16	16	17	16	9	8
3	School 19	14	14	12	12	17	14
1	School 20	3	3	3	3	5	6
2	School 21	6	6	8	8	20	20
3	School 22	18	18	19	19	16	17

NM Score: Correct answers, NR Score: Correctly answered

The correlation between the Mathematics course NR scores of all the students in the 6th and 7th grades is 0.651, and the correlation between the Mathematics course NM scores is 0.659. These results show that there is a moderate positive correlation between mathematics NR and NM scores made in both years. Students with high NR or NM scores in the first year have had a high NR or NM scores in the second year. This is expected as student background is one of the main drivers of student success.

In the correlation analysis to determine the relationship between grade points and change points, the change score of each student was found by subtracting grade 6 score from grade 7 score. Correlation of the change points with grade 6 scores was examined. The correlation between students' sixth grades and their exchange scores (grade 7 - grade 6) was found to be -0.437 for the mathematics course NR score and -0.474 for the NM scores. These value indicates that there is a moderate negative correlation between the change scores of the students and the 6th grade points. This suggests that students with high academic achievement in the 6th grade showed more decline than students with low academic achievement.

In Figure 4, scatter diagrams are given for all pairs of Mathematics course 6th and 7th grade NR and NM averages and 7th grade NR and NM improvement averages of schools. As seen in Figure 4, there is a high linear relationship between success levels in 6th and 7th grade for both NR and NM scores. However, there is a moderate *negative* linear relationship between 6th grade scores and improvement (change) scores. Similarly, there is a negative linear relationship between 7th grade scores and improvement scores.

Table 5 examines whether schools are consistent in terms of student achievement and student improvement. In the Mathematics class, each school is ranked in the order of the 6th grade NR and NM score average, the 7th grade NR and NM score average, and the average NR and NM change scores.

On examining Table 5, it is clear that there are serious inconsistencies between success and development sequences of schools. For example, School #3, School #5, School #6, School #10, School #11, School #15 and School #22 are in higher rank in terms of change, while School #1, School #8, School #9, School #12, School #20 and School #21 are in lower orders in terms of change. Some schools such as School #2 and School #13 have not observed a serious difference in terms of their success and improvement. The statistical comparison of the rankings with respect to each other, the Kendall tau rank correlation coefficient between success and development ranks in terms of the 7th grade NR averages was calculated to be -0.576. The same coefficient is calculated as -0.550 between success and improvement orders in terms of 7th grade NM score averages. These values also indicate that there is a moderate negative relationship between success and improvement. Therefore, there is a serious inconsistency between the improvement and success sequences of schools.

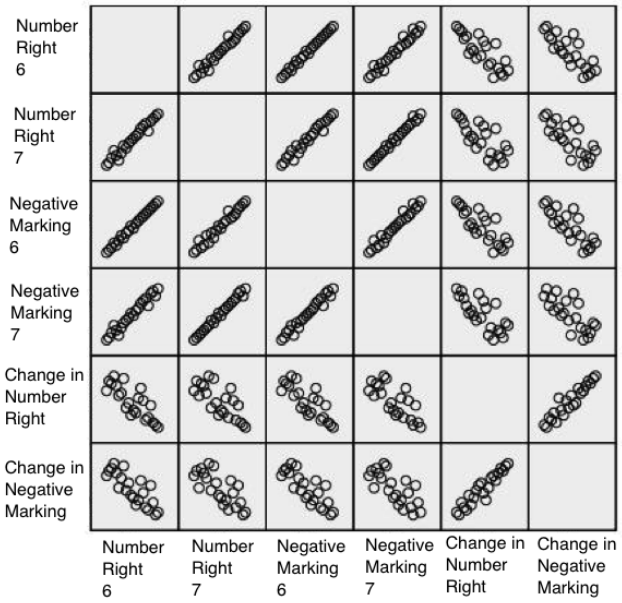

Figure 4. Average success and improvement for Mathematics Course (all pairs)

In summary, the added value of schools for the first research question is given under the Beta (β) column in the relevant schedule for each course. Schools are also categorized in terms of the contribution they provide for each course, and in the same charts, this classification is shared with the reader.

As regards the second research objective, it has been observed that schools with high student achievement are classified as F class in terms of student development in some courses. It was also determined that schools with low student achievement contributed significantly to some courses. In this case, the order of schools in terms of student success and student development is inconsistent. The level of this inconsistency varies according to the courses. However, there is no clear relationship and alignment between the success and the development order of the schools.

These findings do not provide information on the *reasons* for the differences in the development of schools indeedthere can be many reasons why the schools are different from each other. These include, for example, the physical conditions of schools, lack of teachers, the personal abilities of teachers in teaching, and out-of-school assignments of teachers. It can be determined whether there are problems with schools by using the VAA method, but the type or solution method of the problem is not of interest to the VAA. Detailed and long-term monitoring of schools is required to clarify the causes of differences between schools.

Table 5
Mathematics Course Success and Improvement Comparison

		6th grade		*7th grade*		*Change*	
SED	School number	NR score	NM score	NR score	NM score	NR score	NM score
1	School 1	1	1	1	1	16	17
2	School 2	9	9	8	9	9	11
3	School 3	19	19	19	20	10	13
1	School 4	10	10	10	10	11	14
2	School 5	17	17	18	18	4	4
3	School 6	22	22	22	22	1	2
1	School 7	8	7	9	8	21	21
2	School 8	2	2	2	2	20	19
1	School 9	5	4	5	5	18	18
2	School 10	13	13	13	13	7	5
3	School 11	20	20	20	21	3	6
1	School 12	6	6	3	3	14	12
2	School 13	15	15	16	16	17	15
1	School 14	12	11	12	11	8	8
2	School 15	18	18	15	15	5	3
3	School 16	11	12	11	12	6	9
1	School 17	4	5	7	7	22	22
2	School 18	16	16	17	17	12	10
3	School 19	14	14	14	14	13	7
1	School 20	3	3	4	4	19	20
2	School 21	7	8	6	6	15	16
3	School 22	21	21	21	19	2	1

NM Score: Correct answers, NR Score: Correctly answered

Discussion, Conclusion and Recommendations

In this study, we assessed the value-added effects of the education received by 7th graders that attended primary schools within the territorial jurisdiction of the metropolitan municipality of the province of Ankara using the results of the end of year Level Determination Exam.

We determined that the contribution of the school to student success is independent of the academic achievement of the students in the school. Moreover, school contribution is also independent of the socioeconomic status of the students. It has been determined that high academic achievement and schools with students with SES do not always make a high contribution to these students. These results are consistent with many VAA studies (see for example Sanders, 1997; Sanders and Rivers, 1996; Wright, Horn and Sanders, 1997).

In some countries including the US, UK, Canada, and Australia, VAA methods are used to evaluate school and teacher performance. A similar approach can also be

applied in Turkey by the Ministry. VAA is a statistical approach that allows schools to be evaluated in a fairer and objective manner.

Efficient schools and their activities will be a model for less effective schools, which will lead the education system to improve as a whole. Hershberg (2004) draws an analogy between VAA methods for education and a stop watch for a running team. A stop watch cannot make a runner faster; however, it can help the team to understand which members should run in different races. Likewise, VAA methods can be used to improve the effectiveness of an educational system. In order for this to happen, first of all effective and less effective schools need to be identified. With this in mind, in this study, we classified the schools with respect to their contribution to the student success using VAA methods.

We have shown that some schools which are known to be successful (School #4 in Turkish course, School #17 in Mathematics course,), were not able to contribute as much to their students as schools that are thought to be unsuccessful. There are also cases where low academic achievement and SES schools contribute greatly to their students. This also supports the findings in the literature in saying that the effect of the teacher differs with respect to the achievement levels of students (Sanders and Rivers, 1996, Loeb, Soland and Fox, 2014).

For example, School #6, School #11 and School #22 have low academic achievement and SES but have been contributed more to their students than other schools in the Mathematics course. In addition to these, there are also cases where high academic achievement and high SES schools contribute greatly to their students. For example School #1 and School #8 have made serious contributions to their students in Turkish course.

It has also been determined that the change in student scores does not depend on previous academic achievement and SES. Although the academic success of the students is low, it has been seen that a good education may make a great contribution to the student.

Successful and less successful schools have been identified by the VAA method. Less successful schools should benefit from the experience of high contributing schools. The positive impact of the facilities of teachers and schools who contribute more to their students and the experience of these teachers can also be used by other teachers. Letting less effective schools to analyze and imitate the more effective schools is one of the significant benefits of VAA methods (McCaffrey et al, 2003).

While VAA methods are used heavily for accountability reasons, there are opposing views on the value added methods. Haertel (2013) criticizes VAA models in terms of validity and reliability in evaluating teacher VAM estimates. They conclude that scores must be based in appropriate and sound test scores, comparisons should be made amongst homogeneous group of teachers, and VAM should be used flexibly to evaluate teachers. They also mention that users should be well trained to interpret the scores and there should be a clear understanding of the inherent uncertainty in the analysis.

Analysis in future studies should be expanded to include many more years, and exams. Teachers and schools that provide superior added value for a certain period of time can be rewarded according to the amount of contribution calculated by the VAA method. This reward system can also increase the effort of other teachers and schools, which can lead to an overall increase in the quality of education. Complex models can be used if more detailed information is provided. In these complex models, it can be determined whether the contributions of the school or teacher are dependent on other variables, and if so, how these dependencies are shaped. In addition, analyzes can be done on a school-by-teacher basis. This ensures that both school and teacher contributions can be distinguished in student development.

For a more successful VAA application, a large database of long-term exam results of students, including school, classroom and teacher changes is needed. In order for schools and teachers to be judged fairly, it is necessary to establish a standardized examination system that will last for many years. The exams should be used to evaluate not only students, but also teachers and schools.

As a result of this study, there are important decisions that need to be taken in terms of educational policies. It is inevitable that investments for schools with a low level of development should be increased rapidly, and education opportunities should be distributed equally among all schools.

One of the main limitations of this study is that it uses only two years' data. Using more data which is spread over multiple years would be helpful to estimate the schools' effects in a better way. Moreover, we assume that the schools are the only responsible parties for the change in the student scores. However, we did not calculate the effect of teachers independent of the schools. This is mainly because data provided by the authorities includesß only the school information. There was no data available for teacher and/or class codes.

References

Allison, P. D. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology*, *20*, 93-114.

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics, 29* (1), 37-65.

Başaran, İ., & Çınkır. Ş. (2013). *Türkiye Eğitim Sistemi ve okul yönetimi.* Ankara: Ekinoks Yayınları.

Batyra, A. (2017). Gender gaps in student achievement in Turkey.

Berberoğlu, G. (2006). *Sınıf İçi ölçme değerlendirme teknikleri.* İstanbul: Morpa.

Berberoğlu, G., & Kalender, İ. (2005). Öğrenci başarısının, yıllara, okul türlerine, bölgelere göre incelenmesi: ÖSS ve PISA analizi. *Eğitim Bilimleri ve Uygulama, 22,* 21-35.

Bryk, A., & Weisberg, H. (1976). Value-added analysis: A dynamic approach to the estimation of treatment effects. *Journal of Educational Statistics*, *1*, 127-155.

Ciftci, A. (2006). *PISA 2003 Matematik alt testi sonuçlarına göre Türkiye'deki öğrencilerin başarılarını etkileyen bazı faktörlerin incelenmesi.* Yayımlanmamış Yüksek Lisans Tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.

Demir, İ., Kılıç, S., & Depren, Ö. (2009). Factors affecting Turkish students' achievement in Mathematics. *US-China Education Review, 6* (6), 47-53.

EFF. (2008). *School directors' handbook.* Evergreen Freedom Foundation. Web: http://www.effwa.org/pdfs/Value-Added.pdf adresinden 11 Eylül 2008'de alınmıştır.

Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2013). Selecting growth measures for school and teacher evaluations: Should propor- tionality matter? *National Center for Analysis of Longitudinal Data in Education Research*, 21, 1–33. [1,5,7]

Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, *17*(1), 101-128.

Everson, K. C. (2017). Value-added modeling and educational accountability: Are we answering the real questions? *Review of Educational Research*, *87*(1), 35-70.

Gray, J., & Jesson, D. (1987). Exam results and local authority league tables. In A. Harrison & J. Gretton (ed.), *Examination and training UK,* (pp. 33-41).

Gray, J., Jesson, D., & Jones, B. (1984). Predicting differences in examination results between local education authorities: Does school organisation matter?. *Oxford Review of Education 10*(1) 45-68.

Gray, J., Jesson, D., &Jones, B. (1986). The search for a fairer way of comparing schools' examination results. *Research Papers in Education, 1*(2) 91-122.

Haertel, E. H., (2013) *Reliability and validity of inferences about teachers based on student test scores*. ETS Research and Development Princeton, NJ 08541-0001

Hanushek, E. (1971). Teacher characteristics and gains in student achievement: Estimation using micro-data. *American Economic Review, 61*(2), 280-288.

Hanushek, E. (1972). *Education and race.* Lexington, MA: D.C. Heath and Company.

Hanushek, E., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review, 100*(2), 267-271.

Hershberg, T. (2004) *Operation public education*. Retrieved from

http://www.cgp. upenn.edu/ope_value.html last access: September 11, 2008.

Hershberg, T. (2008). *An overview of value-added assessment.* Web: http://www.cgp.upenn.edu/pdf/Value-Added%20for%20Web.pdf last access September 11, 2008.

Ishii, J., & Rivkin, S. G. (2009). Impediments to the estimation of teacher value added. *Education Finance and Policy, 4*(4), 520-536.

Kain, J. F. (1998). *The impact of individual teachers and peers on individual student achievement.* New York: paper presented at the Association for Public Policy Analysis and Management 20th Annual Reseach Conference.

Kane, T. J., & Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. *National Bureau of Economic Research Working Paper 14607.*

Kurtz, M. D. (2018). Value-added and student growth percentile models: What drives differences in estimated classroom effects?. *Statistics and Public Policy*, *5*(1), 1-8.

Lord, F. M. (1969). Statistical adjustments when comparing preexisting groups. *Psychological Bulletin, 72* (5), 336-337.

Loeb, S., Soland, J., & Fox, L. (2014). Is a good teacher a good teacher for all? Comparing value-added of teachers with their english learners and Non-English learners. *Educational Evaluation and Pol- icy Analysis*, 36, 457–475. [1,4]

Maxwell, W. S. (1987). Teachers' attitudes towards disruptive behaviour in secondary schools. *Educational Review, 39*(3), 203-216.

McCaffrey, D. F., & Hamilton, L. S. (2007). *Value-added assessment in practice: Lessons from the Pennsylvania value-added assessment system pilot project*. RAND Corporation.

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. RAND Corporation.

MoNE. (2003). *TIMSS 1999 Üçüncü Uluslar Arası Matematik ve Fen Bilgisi Çalışması Ulusal Rapor*. Ankara: T.C Millî Eğitim Bakanlığı Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı.

MoNE. (2010). *PISA 2006 Projesi Ulusal Nihaî Rapor.* Ankara: T.C Millî Eğitim Bakanlığı Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı.

Murnane, R. J. (1975). *The impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger Publishing Co.

Rivers, J. C. (1999). *The impact of teacher effect on student Math competency achievements*. Knoxville: University of Tennessee.

Rivkin, S. G., Hanushek, E. A. & Kain, J. F. (2000). *Teachers, schools, and academic achievement.* Cambridge, MA: National Bureau of Economic Research, NBER Working Paper #W6691.

Rogosa, D. R. (1995). Myths and methods: "Myths about longitudinal research" plus supplemental questions. In J. M. Gottman (Ed.), *The analysis of change* (pp. 3-66). Mahwah, NJ: Lawrence Erlbaum Associates.

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the "prospects" study of elementary schools. *Teachers College Record, 104*, 1525-1567.

Sanders, W. L. (1997). *Graphical summary of educational findings from the Tennessee Value-Added Assessment System (TVAAS).* Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Sanders, W. L., & Horn, S. P. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8*, 299-311.

Sanders, W., & Horn, S. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12* (3), 247-256.

Sanders, W., & Rivers, J. C. (1996). *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement.* Knoxville, TN: University of Tennessee Value-Added Research Center.

Sanders, W., Saxton, A., & Horn, B. (1997). The Tennessee Value-Added Assessment System: A Quantitative Outcomes-based Approach to Educational Assessment. In M. J (Dü.), *Grading teachers, grading schools: Is student achievement a valid evaluational measure?* (pp. 137-162). Thousand Oaks, CA: Corwin Press Inc.

School Directors' Handbook. (2008). Evergreen Freedom Foundation: 11 09, 2008 Retrieved from

http://www.effwa.org/pdfs/Value-Added.pdf

Tekwe, C. D., Carter, R. L., Ma, C.-X., Algina, J., Lucas, Roth, J., Ariet M., Fisher, T., & Resnick, M. B. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics, 29,* 11-36.

Thum, Y. M. (2001). Measuring progress towards a goal: Estimating teacher productivity using a multivariate multilevel model for value-added analysis. *Sociological Methods and Research, 32*, 153-207.

US Department of Education (2009). *Race to the top fund: Final rule*. Web: http://edocket.access.gpo.gov/2009/pdf/E9-27426.pdf last access August 21, 2010.

Webster, W., & Medro, R. (1997). The Dallas Value-Added Accountability System. In J. Millman (Dü.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 81-99). Thousand Oaks, CA: Corwin Press, Inc.

Wikipedia. (2008). No Child Left Behind. 11 09, 2008 retrieved from http://en.wikipedia.org/wiki/NCLB

Willms, J. D. (1987). Difference between Scottish educational authorities in their educational attainment. *Oxford Review of Education 13*(2), 211-232.

Woodhourse, G., & Goldstein, P. (1988). Educational performance indicators and LEA league tables. *Oxford Review of Education, 14*(3), 301-320.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11* (1), 57-67.

İlköğretimde Öğrenci Başarısında Okulun Etkisinin Katma Değer Belirleme Yöntemiyle İncelenmesi (Ankara İli Örneği)

Atıf:

Koc, F. (2018). Investigation of school effects on student achievement in primary education using value-added assessment. *Eurasian Journal of Educational Research, 76, 41-72,* DOI: 10.14689/ejer.2018.76.3

Özet

Problem Durumu: Türkiye'de okulların başarısı genellikle öğrencilerin bir öğrenim düzeyinden üst öğretim düzeyine geçişte girdikleri sınavlarda gösterdikleri başarıya göre belirlenmektedir. Bu haliyle okul, öğretmen, öğrenci başarılarının doğru ve nesnel olarak belirlendiğini söylemek güçtür. Oysa okul, öğretmen, öğrenci başarılarına etki eden değişik faktörler bulunmaktadır. Türkiye'de gerek ilköğretim düzeyinde SBS gerekse ortaöğretim düzeyinde ÖSS sınavları üzerine tartışmalar süregelmektedir. Siyasi iktidarlar eğitim programlarına bu sınavları kaldıracakları ya da alternatif sınavlar uygulayacakları yönünde görüşler koymaktadır. ABD, İngiltere, Kanada ve Avustralya gibi gelişmiş ülkelerde okul, öğretmen, öğrenci başarılarının belirlenmesinde Türkiye'de yapılan sınavlara alternatif olan KDB yaklaşımları kullanılmaktadır. Bu ülkelerde öğrencilerin yıllık gelişimlerinin tespitinde ve bu gelişime okul, öğretmenin katkılarının hesaplanmasında SBS'ye benzer sınavlar kullanılmaktadır. Ancak okul, öğretmen değerlendirmelerinde ortalama başarı puanları yerine KDB yaklaşımları kullanılmaktadır. Benzer yaklaşım MEB tarafından Türkiye'de de uygulanabilir. Çünkü, KDB okulları adil, akla yatkın ve nesnel bir biçimde değerlendirmeyi sağlayan istatistiksel bir yöntemdir.

KDB yönteminde, öğrenci okul sistemine girdiği andan itibaren öğretmen veya okulun katkısıyla ilgilendiği için, ekonomi literatüründen ödünç alınan katma değer terimiyle ilişkilendirilmiştir (McCaffrey, Lockwood, Koretz, & Hamilton, 2003). Katma değer eğitim literatüründe okulun veya öğretmenin, eğitimin birikimli

etkisinden veya öğrencinin kendi sosyo-ekonomik durumu ve çevresinden arınmış olarak, öğrenci gelişimindeki özgün katkıları anlamında kullanılmaktadır. Eğitimde katma değer gerek akademik çevrelerde, gerekse eğitim politikalarını şekillendiren eğitim çalışanları ve siyasetçiler arasında rağbet gören bir kavramdır. Birleşik Devletlerdeki eğitimde reform çabalarının en önemli parçalarından biri de okulların, öğretmenlerin ve öğrencilerin performanslarının standardize sınavlarla ölçülmesidir. Hemen hemen bütün eyaletlerde eğitim alanında kullanılan takip sistemleri (accountability system) o ya da bu şekilde standardize sınavların sonuçlarını kullanır. Özellikle bütün öğrencilerin akademik başarımının asgari bir seviyenin üstünde olmasını amaçlayan "Hiçbir Çocuk Geride Kalmasın" (No Child Left Behind – NCLB) yasasının yürürlüğe girdiği 2001 yılından itibaren standardize sınav tabanlı takip sistemleri, Amerikan eğitim politikasının merkezine oturmuştur (Wikipedia, 2008). Devre devreye karşılaştırma ve Yıllık Yeterli Gelişim yöntemlerinin takip sistemlerindeki hızla artan baskınlığına karşın Katma Değer Belirlemesi yöntemi hem eğitim politikacıları hem de akademik çevrelerde giderek artan bir ilgi çekmeye başlamıştır. Örneğin Ohio, Pennsylvania ve Tennessee eyaletlerinde kabul edilen bazı yasalar, eğitim bölgesi yöneticilerinin, okul müdürlerinin ve öğretmenlerin, KDBdeki başarılarına göre terfi ve maaş anlamında ödüllendirilmesine veya cezalandırılmasına yönelik hükümler içermektedir (School Directors' Handbook, 2008).

Araştırmanın Amacı: Bu araştırmanın amacı, Ankara İl Büyükşehir Belediyesi sınırları içindeki ilköğretim okullarının, 7. sınıf öğrencilerinin katkısını, yıl sonu Seviye Belirleme Sınavı sonuçlarını kullanarak değerlendirmektir. Bu amaç doğrultusunda aşağıdaki sorulara yanıt aranmıştır. (1) İlköğretim okulları, öğrencilerine ne kadar katma değer sağlamıştır? (2) Öğrenci başarısı ile öğrenci gelişimi açısından ilköğretim okullarının sıralamaları tutarlı mıdır?

Araştırmanın Yöntemi: Nedensel karşılaştırma çalışma çerçevesine göre yürütülen bu çalışmanın popülasyonu, Ankara ilindeki ilköğretim okullarından oluşurken, örneklemi 24 ilköğretim okulunu kapsamaktadır. Bu okullardaki öğrencilerden 2007-2008 ve 2008-2009 eğitim yıllarında aynı okula giden öğrenciler analize dâhil edilmiştir. Örneklemdeki okulların öğrenci gelişimi üzerindeki katma değer etkileri, basit bir sabit etki modeli kullanılarak değerlendirilmiştir. Bununla birlikte okulların ortalama öğrenci başarıları ve okul katkılarının sıralaması arasında istatistiksel bir ilişki olup olmadığı da incelenmiştir.

Araştırmanın Bulguları: Bu çalışmanın sonuçları, okulların sıralamasında öğrenci gelişimine kattığı katma değer ve öğrenci başarısına göre sıralamalar arasında önemli bir tutarsızlık olduğunu göstermektedir. Bu, Türkiye'deki okulların bugüne kadar performansın değerlendirilmesinde sıklıkla kullanılan yöntemin eksikliğini göstermektedir. Ayrıca, sonuçlar, okulların öğrenci gelişimi üzerindeki katma değer etkilerinin konudan konuya büyük ölçüde farklı olduğunu göstermektedir. Bu araştırmada okulun öğrenci başarısına olan katkısının okuldaki öğrencilerin geçmiş akademik başarılarından ve bulundukları bölgenin sosyoekonomik düzeyinden bağımsız olduğu saptanmıştır. Yüksek akademik başarıya ve sosyoekonomik düzeye sahip öğrencilerden oluşan okulların her zaman bu öğrencilere yüksek katkı sağlamadıkları belirlenmiştir. Ayrıca, Başarılı olarak bilinen bazı okulların aslında

öğrencilerine düşük başarılı olarak düşünülen okullar kadar katkı sağlayamadıkları saptanmıştır. Tersine düşük akademik başarıya ve sosyoekonomik düzeye sahip okulların öğrencilerine yüksek katkı sağladığı durumlar da vardır. Öte yandan, öğrenci puanlarındaki değişimin önceki akademik başarıya ve sosyoekonomik düzeye bağlı olmadığı saptanmıştır. Öğrencilerin akademik başarısı düşük olsa da iyi bir eğitimin öğrenciye büyük katkı sağladığı görülmüştür.

Araştırmanın Sonuç ve Önerileri: Bu araştırmada yapılan çözümlemelerle başarılı ve daha az başarılı okullar KDB yöntemiyle belirlenmiştir. Daha az başarılı okullar yüksek katkı sağlayan okulların birikimlerinden faydalanmalıdırlar. Üst düzey il ve ilçe Milli Eğitim Müdürlüklerinin koordinasyonuyla öğretmenlerin bilgi alışverişinde bulunmaları sağlanabilir. Öğrencilerine daha fazla katkı sağlayan öğretmenlerin sahip olduğu olanakların olumlu etkisi ve bu öğretmenlerin deneyimlerinden diğer öğretmenlerin de faydalanması sağlanabilir. Bu araştırmanın, yıllar boyunca daha fazla verinin ortaya çıkmasıyla okulların adil bir değerlendirmesine yol açması beklenmektedir. Okul öğrencilerinin ortalama başarısı Türkiye'deki okulların performansını değerlendirmek için sıklıkla kullanılan yöntemdir. Bu çalışma, Türk okullarının değerlendirilme biçiminin sorunlu olduğunu ve okulların etkilerinin değerlendirilmesinde katma değer yöntemlerinin dikkate alınması gerektiğine işaret etmektedir. Ayrıca, bu çalışma Türkiye'de gerçekleştirilen ilk katma değer değerlendirmesi çalışması olması sebebiyle önem arz etmektedir. İleriki çalışmalarda çözümleme hem tüm Türkiye'yi içerecek şekilde hem de daha uzun yılları ve sınavları içerecek şekilde genişletilmelidir. KDB yöntemiyle hesaplanan katkı miktarına göre belirli bir süre üstün katma değer sağlayan öğretmenler ve okullar ödüllendirilebilir. Benzer şekilde KDB uygulamaları için uzun yıllara ait sınav sonuçlarının bulunduğu, öğrencilerin okul, sınıf ve öğretmen değişikliklerini de içeren geniş bir veritabanına ihtiyaç vardır. Okul ve öğretmenlerin adil bir şekilde değerlendirilebilmesi için uzun yıllar devam edecek standart bir sınav sisteminin oluşturulması gerekir. Sınavlar öğrencilere ek olarak öğretmen ve okulları değerlendirmek amacıyla da kullanılmalıdır.

Anahtar Kavramlar: Okul performansının değerlendirilmesi, katma değer belirlenmesi, basit sabit etki modeli, seviye belirleme sınavı.

APPENDIX

Table 6

Municiaplities and Names of Schools in the Study

Region	School Name
Çankaya	Yasemin Karakaya İlköğretim Okulu
	Turhan Feyzioğlu İlköğretim Okulu
	Mohaç İlköğretim Okulu
Altındağ	Cebeci İlköğretim Okulu
	Atıfbey İlköğretim Okulu
	Polis Amca İlköğretim Okulu
Yenimahalle	Türkkonut İlköğretim Okulu
	Refika Aksoy İlköğretim Okulu
Mamak	29 Ekim İlköğretim Okulu
	Şehitlik İlköğretim Okulu
	Fatih Sultan Mehmet İlköğretim Okulu
Gölbaşı	T.E.K İlköğretim Okulu
	Atatürk İlköğretim Okulu
Keçiören	Çizmeci İlköğretim Okulu
	Atapark İlköğretim Okulu
	Şenlik İlköğretim Okulu
Sincan	Andiçen İlköğretim Okulu
	Burak Reis İlköğretim Okulu
	Özkent Akbilek İlköğretim Okulu
Etimesgut	Hasan Ali Yücel İlköğretim Okulu
	Atatürk İlköğretim Okulu
	Layika Akbilek İlköğretim Okulu