mgr Aleksandra Jasińska-Maciążek

Summary of the doctoral thesis
„Measuring school effectiveness"

Moving away from the bureaucratic model of education management to increase the autonomy of schools and teachers as well as the decentralisation of management and financing of education are increasing the demand for a reliable measurement of learning outcomes in schools. The results of achievement tests are not sufficient because, to a large extent, they depend on factors that are beyond the school's control. Therefore researchers are developing methods of assessing school effectiveness that take into account the fact that schools differ in terms of student characteristics as well as a wider context in which the schools are operating. The goal of these methods is to capture the part of the learning outcomes which can be attributed to the activities of the school and exclude the part that is independent of the school. Value-added models (VAM) are the most popular. They allow for estimating the student achievement growth.

The aim of the doctoral thesis, which is a series of thematically coherent publications, is to discuss selected aspects of measuring school effectiveness in the context of the Polish education system. My studies were conducted within the research tradition known as the *school effectiveness research.* The publications have been written as part of my work on the projects dedicated to the development of the value-added method in Poland (*EWD project*) as well as the SUEK research project (*Szkolne Uwarunkowania Efektywności Kształcenia; School Effectiveness Research Project*). The data that my analyses have been based on come from national examination system as well as three major nationwide longitudinal studies conducted in primary and lower secondary schools in Poland within the EWD and SUEK research projects (each of the studies included data sourced from more than 5,000 students from over 150 schools). In my research I have used advanced quantitative data analysis methods, such as hierarchical linear models, item response theory models, exploratory factor analyses, and linear correlation.

In the series of my publications I have focused on five research problems. The first research problem concerned investigating how to measure student achievement in such a way that the results of these measurements could be used as a basis for school effectiveness models, as well as how to examine whether these measurements meet the requirements. Based on the set of achievement tests for grade three TOS 3 (*Testy osiągnięć szkolnych dla klas 3*) that I have co-created, in the first publication I showed the stages of test development that every researcher must utilise in order to create a tool with the right psychometric properties. In two subsequent publications I concentrated on the requirements that need to be met by the school achievement tests so that they could be used as a basis for valid and reliable measures of

school effectiveness. I showed that some of the national examinations used in Poland meet these requirements at a satisfactory level. At the same time, the results from the low-stake assessments, for instance Nationwide Testing of Skills of Third Graders (*Ogólnopolskie Badanie Umiejętności Trzecioklasistów*), give rise to justifiable doubts. The analyses showed that it is problematic to use the low-stake tests to evaluate school effectiveness. In my last publication in this area I analysed the possibility of using quasi-observation scales completed by the teachers to assess the effectiveness of schools. I demonstrated that while this measurement can provide valuable information about the within-class variance of students' characteristics, it ceases to be useful if we want to grasp the differences between classes or schools, which is crucial for assessing school effectiveness.

My second research problem focused on how much Polish schools differ in terms of the results and effectiveness of teaching and how to correctly estimate this diversity based on the hierarchical structure of educational data. In one of the publications I discussed the hierarchical linear models in this context. The analyses based on these models, presented in my subsequent publications, have shown that the between-school variance in student performance after the third grade of primary school varies in the Polish education system between 9% and 13% (depending on sample and the area of skills), and after the sixth grade it is 3 to 6 percentage points higher. This explains how much the opportunities for students to receive high test scores depend on which school they attend. Part of this variance is explained by an uneven distribution of student characteristics between schools (those characteristics that are relevant to the learning outcomes). When we take this into account in the analyses, the differences between primary schools become significantly smaller, but still remain significant. My analyses showed that while the between-school variance of school effectiveness after the first stage of education is not large, it is growing in the second stage of education. Between-school variance in student performance in lower secondary schools is larger and reaches over 20%, and, in addition, the analyses show significant differences in this area depending of the size of the city or village.

The third research problem focused on developing models of assessing school effectiveness that take into account the different goals, also when the data availability is limited. The main research question referred to the consequences of the adoption of a specific model for the estimated indicators and their interpretation. I focused my research on the models for primary schools. I examined the possibility of developing models for the first level of education where only one measurement of student achievement seems to be realistic, at the end of grade three. I proposed two contextualised attainment models (CAM) in which the achievement test results at the end of third grade are controlled for the main contextual variables which are important for the learning outcomes and, at the same time, are beyond school's influence. The indicators estimated from those models were highly consistent with indicators estimated from comprehensive models, i.e. those that included all important determinants of student

achievements that are beyond school's control and that were confirmed in the study. I documented also other characteristics of these models. In my next publication I showed how to develop different models of assessing school effectiveness, and I compared the four most commonly used models: a contextualised attainment model, as well as three value-added models that had different sets of the additional control variables. The analyses showed first of all that it is dangerous to assign simple labels such as "effective school" – "ineffective school" because the interpretation depends heavily on whether we use the indicators estimated from VAM or CAM. The differences between various value-added indicators were not that big. In addition, I documented the relationship between estimated indicators and selected factors that are beyond school's influence, as well as other characteristics of the models.

The fourth research problem focused on the assessment of validity of school effectiveness models. Searching for the evidence supporting or undermining the validity of school effectiveness indicators used in Poland, I showed that the lack of the information about students' age in the value-added models did not bias the estimated indicators. Additionally, I demonstrated that although lower secondary schools vary significantly in terms of the level of intelligence of the students, the exam results in the sixth grade used in this model allow for sufficient control of the between-school variance. It turned out that including the information about students' intelligence in the model did not have a significant influence on the estimated value-added indicators. Moreover, I indicated that, contrary to popular opinion, it cannot be said that the schools in which students achieve additional support in learning are gaining when it comes to value-added indicators. I found a negative correlation between receiving such assistance and the relative achievement gains. As a result of analysing the criterion validity of the value-added method I have shown that not only do schools significantly differ in terms of students' intelligence gains, but also the students' intelligence grows faster in the schools with higher added value. The conclusions of my research support the validity of the Polish school effectiveness models.

The last of the research problems that I focused on referred to the possibility of using the school effectiveness indicators in an education policy. I focused on the less obvious area, i.e. using these indicators for monitoring and evaluation of educational processes at the local level by analysing the process of differentiation of lower secondary schools in big cities. The analyses showed that within cities we are dealing with a dynamic process of schools differentiation with regard to educational outcomes. The article also describes the possibilities of using value-added indicators to monitor and evaluate those processes at the local level from the perspective of effectiveness of individual schools.