

## Introduction

With education systems in all OECD countries coming under increasing pressure to enhance their effectiveness and efficiency, there is a growing recognition of the need for accurate school performance measures. Assessments of student performance are now common in many OECD countries, and the results are often widely reported and used in public debate as well as for school improvement purposes. There are diverging views on how results from evaluation and assessment can and should be used. Some see them primarily as tools to reveal best practices and identify shared problems in order to encourage teachers and schools to improve and develop more supportive and productive learning environments. Others extend their purpose to support contestability of public services or market-mechanisms in the allocation of resources, *e.g.* by making comparative results of schools publicly available to facilitate parental choice or by having funds following students. Regardless of the objectives of measuring school performance it is important that they truly reflect the contributions which individual schools make rather than merely or partly the different socio-economic conditions under which teachers teach and schools operate. If this is not the case, resources can be misallocated and perverse incentives created if, for example, schools can receive a higher performance measure through academic selection or through selecting students from privileged socio-economic backgrounds, rather than improving outcomes through investment in better instructional methods.

This report documents state of the art methods, referred to as value-added modelling, which allow users to separate the contributions of schools to student performance from contextual factors that are outside the control of classrooms and schools. The greater accuracy they provide in measuring school performance and the role they can play in the development and implementation of education policy and school development initiatives has created a growing interest in value-added modelling. A number of studies have shown that value-added modelling provides more accurate estimates of school performance than do the comparisons of raw test scores or cross-sectional contextualised attainment models (discussed in more detail below) that are often used to provide school performance estimates (Doran & Izumi, 2004). They provide a fundamentally more accurate and valuable quantitative basis than do raw test scores and cross-sectional studies for

school improvement planning, policy development and for enacting effective school accountability arrangements.

Value-added models are statistical analyses that provide quantitative school performance measures (*e.g.* a school value-added score) that can be used to develop, monitor and evaluate schools and other aspects of the education system. In this sense, implementing a system of value-added modelling should be viewed as a means to an end rather than an end in itself. How value-added measures are used shall differ between education systems and these differences should inform decisions and actions undertaken in the development of a system of value-added modelling. Therefore, the development process should be shaped by the intended use and application of schools' value-added scores to achieve specified policy objectives.

Three broad policy objectives are identified in this report that can benefit from the use of value-added modelling: school improvement initiatives; school accountability; and school choice. The effectiveness of the use of performance data in decision-making concerning these policy objectives relies on the accuracy of the performance measures used. However, the growth of data-based decision-making to advance policy objectives has been stymied by the lack of accurate school performance data that is essential for educational improvements (Raudenbush, 2004; Vignoles et al., 2000). Raw test scores provide measures of student performance but there are clear problems with drawing inferences from these data about school performance. Cross-sectional contextualised-attainment models take into account contextual characteristics such as student background but are less useful in isolating the effects of individual schools upon students' education. Value-added measures are a significant advance, providing an accurate measure of school performance upon which to base decisions to advance policy objectives and lift school performance. This report illustrates how value-added information can be used for school improvement purposes, for individual programmes and policies and in decision-making at the system- and school-level.

For all *school improvement initiatives* it is important to recognise that improvement in a given activity or set of activities first requires an accurate evaluation of the current situation that, in turn, requires an accurate measure of performance (Sammons et al., 1994). It is difficult to effectively develop programs for the future if it is not possible to accurately analyse the current situation. At the system level, value-added information can be used to determine the areas of the education system and schools that are adding the most value and those areas in which further improvement is required. At the school level, the subjects, grades and groups of students can be identified where the school is adding most value and where improvement is needed. In this sense, value-added scores and information are most valuable if they not only document the current status of schools but also generate information that can support continuous school improvement. Statistical analyses of the

relations between school inputs and indicators of school performance can suggest which strategies are and are not working, leading to policy adjustments and the reallocation of resources.

Value-added modelling can also be used to create projections of school performance that can assist in planning, resource allocation and decision-making. Projections can be used to identify future outcomes, for example, providing estimates if current performance trajectories were to continue, and also to set performance targets. Such targets can inform decision-making at the school level of how best to utilise resources and structure the education offered to meet specified performance targets (Hill et al., 2005; Doran and Izumi, 2004). Combined with additional information collected within schools, the projections of future student performance based on value-added estimates provide a comprehensive picture of a school's performance. School personnel then have at their disposal an information base that can serve as a foundation for planning and action.

*Systems of school accountability* can benefit greatly from the use of value-added modelling. Systems of accountability identify which entities are accountable to which bodies for specific practices or outputs (McKewen, 1995). Such systems might provide information to the general public: taxpayers might be informed as to whether tax money is used efficiently, and users might be able to choose educational institutions on a more informed basis. Yet the key issue remains whether the assessment of processes and of performance is accurate and fair to individual schools. This report illustrates that value-added modelling provides a more accurate, and therefore fairer, measure of school performance (as measured by increases in student performance) that can also be used to improve the evaluation of school processes. The results of value-added modelling (*i.e.* schools' value-added scores) provide measures of the extent to which schools have succeeded in lifting student performance. When used in systems of school accountability, these measures can be used effectively in school evaluations, with fairer consequences for schools and school personnel.

*School choice* is the third key policy objective discussed in this report that benefits from the use of value-added modelling. This data is intended to inform parents and families of the performance of different schools to aid their decision-making in choosing their school. This requires publishing the data on school results (Gorard, Fitz, and Taylor, 2001). While this does not occur in all countries, it is a growing trend among OECD member countries (OECD, 2007a). As is discussed in Part I of this report, there are numerous benefits from improved levels of school choice within an education system. Parents are able to choose schools that are better suited to their needs and resources can then flow to those schools best meeting those needs (Hoxby, 2003). However, such benefits depend upon an accurate measure of school performance, otherwise families' choices are misinformed and resources are misallocated. The greater accuracy of value-added modelling is essential to

the effectiveness of a system of school choice. It allows parents a more accurate measure of school performance upon which to base their decisions and allows schools a fairer opportunity to improve their performance.

The policy considerations and political issues surrounding systems of value-added modelling can differ. Given such differences, it can be beneficial to structure the development and implementation of a system of value-added modelling to suit the prescribed policy objectives. The use of value-added modelling to advance specific policy objectives is discussed in Part I of this report and are also detailed in Part III that deals with implementation issues.

The greater accuracy inherent in value-added modelling creates greater confidence in the use of performance measures to further the three policy objectives outlined above. The greater confidence stems from the improvements made in this modelling over time and the advantages compared to other methods of estimating school performance. The modern era of ‘school effects’ research began, at least in the USA, with the so-called Coleman Report that studied the relationships of schools and families to student academic attainment (Coleman, 1966). This complemented a number of European studies that looked at issues of inequality in terms of intergenerational analyses that compared outcomes over generations (Carlsson, 1958; Glass, 1954). Subsequent school effectiveness studies also carried out quantitative comparisons of schools. In the initial phase, high-achieving schools were identified by comparing the average test scores of the students. The next step for researchers was often to select a small number of such schools for further analysis with the hope of identifying the elements of their practice that were responsible for their success. The ultimate goal was to disseminate the findings in order to effect broader school improvement. Early work in this area is reviewed in Madaus, Airasian and Kellaghan (1980).

It was recognised early on that school rankings based on students’ ‘raw’ test score were highly correlated with their students’ socio-economic status (McCall, Kingsbury and Olson, 2004). Bethell (2005), for example, discusses some of the controversies arising from the use of tables comparing raw test scores in England. Multivariate cross-sectional analyses have been used to try and overcome these problems. In the simplest version of these analyses, school average test scores were regressed on a number of (aggregate) relevant demographic characteristics of the schools’ students. The idea was to rank schools on the basis of their residuals from the regression. These residuals were often termed ‘school effects’. Schools with large positive residuals were considered to be exemplary and worthy of further study. Schools with large negative residuals were considered to be problematic and also requiring further study, although for different reasons. Alternative adjustment strategies have been proposed and the resulting

differences in school rankings compared (Dyer, Linn and Patton, 1969; Burstein, 1980).

More sophisticated cross-sectional models have subsequently gained in popularity and use with methods that take into account the hierarchical structure of school systems, with students nested within classes, classes nested within schools and schools nested within districts/local areas (Aitkin and Longford, 1986; Goldstein, 1986; Willms and Raudenbush, 1989). The estimates provided by these models have grown in sophistication and have been commonly used in education analyses across OECD member countries. These cross-sectional estimations have been categorised in this report as *contextualised attainment models*. These multivariate models can be used to provide a measure of school performance but it was considered that such analyses did not contain the required analytic framework to be classified as value-added models. Contextualised attainment models estimate the magnitude of contributing factors to student performance or attainment at a particular point in time. A typical example is a regression model that regresses a vector of students' socio-economic backgrounds or contextual characteristics and a variable identifying the school each student attends against some achievement measure. The adjustment to raw scores made with the inclusion of contextual characteristics provides measures that better reflect the contribution of schools to student learning than the use of 'raw' test scores to measure school performance. The results of these cross-sectional models build upon theoretical analyses of the role of the family in shaping people's socio-economic outcomes and often find that the main contributor to the level of student attainment is parental socio-economic background (OECD, 2007b; Haveman and Wolfe, 1995; Becker, 1964). Information on the role of student socio-economic background in educational attainment, while interesting and important, often does not yield sufficient information to enable policy makers to make decisions on school accountability and school choice and to drive school improvement reforms. Nevertheless, these contextualised attainment models are a clear improvement on the use of unadjusted results and raw attainment scores to assess school performance.

A significant advance was made with the development of value-added modelling that utilised multiple measures of student performance to estimate the impact (or value-adding) of individual schools upon those student performance measures. An important assessment of value-added modelling was provided by Fitz-Gibbon (1997) who was asked to advise the British Government on the development of a system of value-added modelling. Fitz-Gibbon concluded that such a model could be the basis for a statistically valid and readily understood national value-added system. Value-added models employ data that tracks the test score trajectories of individual students in one or more subjects over one or more years (Mortimer et al., 1988; Goldstein et al., 1993; SCAA, 1994; Sanders, Saxton

and Horn, 1997; Webster and Mendro, 1997; Rowan, Correnti and Miller, 2002; Ponisciak and Bryk, 2005; Choi and Seltzer, 2005; McCaffrey et al., 2004; McCaffrey et al., 2003; McCaffrey et al., 2005). Through various kinds of adjustments, student growth data is transformed into indicators of school value-added. Examples are discussed of the main types of value-added models in Chapter Five of this report.

Value-added models are a substantial improvement on many current measures of school performance. Comparisons of raw test scores provide some important information but are poor measures of school performance. They fail to take account of prior achievement levels and produce results that can largely reflect differences in contextual characteristics such as students' socio-economic background. Contextualised attainment models try to address these problems by measuring the impact of contextual characteristics upon a specific performance measure but are less useful in disentangling school effects upon student progress from other contextual characteristics and are therefore less useful in measuring school performance. Value-added models attempt to overcome these problems by incorporating student prior attainment measures and, in some cases, contextual characteristics. This enables a more refined analysis of progress in student performance that is more effective in disentangling the effects of various factors that affect student progress. These advantages allow for greater accuracy in measuring performance which then creates greater confidence in the interpretation of school performance measures.

In summary, this report argues that value-added modelling contributes to system-wide learning by accurately measuring higher and lower performing aspects of the education system; to school improvement through improved identification and analysis of 'what works'; to improved and more equitable transparent systems of school accountability and school choice that can then create well-defined incentives for schools to improve their performance; to the development of information systems that allow schools to analyse and evaluate their performance and strengthen the overall system of school evaluation; to systems of education funding that more effectively direct resources to areas of need; and, to overcoming entrenched socioeconomic inequalities that exist in societies that might be masked at the school level by indiscriminate and inaccurate performance measures.

### *Value-added modelling: A definition*

Given the advantages of using value-added modelling, it is essential that this report distinguishes value-added modelling from other statistical approaches. Across participating countries there has been a large variation in the use of value-added modelling and statistical analyses to analyse school performance. Such variation increases the importance of defining both 'value-added' and 'value-added modelling' to clearly differentiate them

from other types of statistical analyses. In this report, the value-added contribution of a school is defined as:

the contribution of a school to students' progress towards stated or prescribed education objectives (e.g. cognitive achievement). The contribution is net of other factors that contribute to students' educational progress.

From this definition of value-added it was possible to define value-added modelling as:

a class of statistical models that estimate the contributions of schools to student progress in stated or prescribed education objectives (e.g. cognitive achievement) measured at at least two points in time.

Particular value-added models might utilise a narrower definition of the estimation of school performance but this general definition can be applied to a variety of value-added specifications while still clearly delineating value-added modelling from other types of statistical analyses. Statistical analyses that have been undertaken in a number of countries to monitor school performance would not be considered to be value-added modelling using these definitions. Such analyses often did not include at least two measures of student performance that can be considered to be the basis of value-added modelling. These analyses have been defined in this report as contextualised attainment models. It was considered appropriate not to try to expand the definition of value-added modelling to fit the performance measures used in each participating country as it would decrease the effectiveness of the analysis.

A distinguishing feature of value-added modelling is the inclusion of prior performance measures that allow a more accurate estimation of the contribution of the school to student progress. Doran & Izumi (2004) emphasised the advantages of value-added modelling in tracking students over time compared to cross-sectional (or contextualised attainment) models that provide a 'snapshot' picture of student performance. Value-added modelling facilitates more detailed analysis of school improvement by estimating the contribution of the school to improvements in student performance over a given time period. Additionally, value-added models are able to better account for unobserved factors contributing to the initial performance measure, such as student ability that are a systemic problem in much contextualised attainment modelling (Raudenbush, 2004).

The inclusion of a prior performance measure allows a school's value-added to be estimated. The value-added should be interpreted as the contribution of the school to student performance between the two performance measures. This is an important issue as it is possible to employ different student assessments at different time intervals. Such differences

need to be recognised in interpretation of the contribution of individual schools (*i.e.* a school's value-added score). A key distinction is the subject matter of the student assessments as the school's value-added is being estimated only on the subject matter included in the assessments (this is discussed further in Chapter one). A further consideration is the timing of the assessments. A number of value-added estimations estimate the contribution of the school in a given year. However, a number of education systems do not have annual assessments or a structure of assessments that would permit the estimation of a single year value-added score. This is not to say that value-added cannot be estimated over a multiple-year timeframe. On the contrary, such estimations are made in a number of education systems. But it is important to recognise that these differ from single year value-added scores so that in discussion of schools' value-added scores it is made clear the subject matter and the time-span in which value-added is measured.

The importance of multiple attainment measures raises the issue of what should be considered an appropriate prior performance measure upon which to measure progress. There is considerable debate about the comparability of test scores and the conversion of scores into meaningful and comparable scales (Braun, 2000; Dorans et al., 2007; Patz, 2007; Kolen and Brennan, 2004). Of course, many value-added models do not actually require that the test scores be vertically scaled. They simply require that scores in successive grades be approximately linearly related and, in most cases, that is a reasonable measure (Doran and Cohen, 2005). This report does not discuss the development of student assessment instruments themselves: a review of the considerable literature analysing assessment issues is outside the scope of this report. However, the definition of value-added used in this report focuses on *progress in stated or prescribed education objectives (e.g. cognitive achievement)*. This precludes some contextualised attainment models that include intelligence measures such as IQ scores that might be considered to be a measure of general ability but are less suitable as a measure of prior attainment upon which to measure progress. In discussion of schools' value-added scores it should always be clear what the prior and current attainment measures and test scores actually represent and how they should therefore affect policy actions and schools.

Even with the greater accuracy obtained with the use of value-added modelling, there remain some difficulties in measuring school performance. The interpretation of schools' value-added scores should include various caveats and cautions for correct interpretation. These issues are discussed in Part II of this report. While this discussion seeks to illustrate the various measurement issues in designing and utilising value-added modelling, it is not the intention to negate their considerable potential. To the contrary, accurate value-added estimations have great potential for use in policy development and school improvement initiatives and are a substantial

improvement on alternative measures. For example, Chapter Six discusses the statistical and methodological issues that must be addressed in the development and use of value-added modelling. These issues are highlighted not to deter the use of value-added modelling in education systems but to encourage their effective development in advancing specified policy objectives. In fact, a key reason why the use of value-added modelling is encouraged is that these statistical and methodological issues often create far greater problems of misspecification with other statistical approaches and school performance measures. These alternative approaches normally provide less accurate measures of school performance and are therefore less useful for effective system and school development. The attention given in this report to statistical and methodological issues is thus done to emphasise the need to develop and provide accurate value-added measures of school performance to both inform policy development and school improvement initiatives and to gain the confidence of stakeholders.

### *Format of this report*

This report is divided into three parts that might be suitable to slightly different audiences. Part I discusses the objectives and use of value-added modelling. This includes a discussion of the policy objectives (discussed in Chapter One) that can be advanced with value-added modelling. Linked to this issue is a discussion of how value-added information and school scores can be presented to different stakeholders, distinguishing between the presentation of value-added information for internal purposes, for public consumption, and presentation in the media. A number of examples are provided of effective presentation methods in countries in Chapter Two. The discussion of the presentation of value-added information for internal purposes focuses upon the application of value-added for modelling for school improvement purposes in Chapter Three. Central to this discussion is how the information can play a key role in fostering data-based decision-making in schools that utilise accurate performance measures to develop and monitor school improvement initiatives. This discussion views schools as learning organisations that undertake and benefit from analysis of different aspects of school and student performance. Focus is given to the targeted use of value-added modelling for: specific sub-groups of the student population and specific aspects of schools; setting performance targets and performance projections; identifying students in need of special assistance and early interventions; and, improving the overall system of school evaluations.

Part II discusses the design of value-added models and focuses upon the technical aspects of value-added modelling. Chapter Four discusses key design considerations in developing a system of value-added modelling and identifies the key issues that need to be addressed. Examples of the main types of value-added models are presented in Chapter Five to provide some tangible examples and to illustrate their various requirements, and how they

might be adapted to particular settings. Chapter Six discusses the key statistical and methodological considerations in the development of value-added modelling. These are emphasised in order to assist in the identification of the key criteria with which to choose a preferred value-added model(s) in an education system. A number of issues are presented with supporting analysis from participating countries discussed to highlight the steps that can be taken in choosing the appropriate value-added model. The point is made that a key aspect of this issue for administrators is to decide upon what is the most appropriate model to meet the objectives and planned use of value-added modelling.

Part III discusses the implementation of systems of value-added modelling in education systems. This discussion provides policy makers and administrators with guidance on how to implement a system that best meets their needs. Again, the experiences from participating countries are drawn upon to illustrate the key issues and potential strategies that can be employed. Chapter Seven focuses upon the initial steps that need to be taken in the development of the system leading up to, and including, the pilot phase of implementation. Chapter Eight discusses the ongoing development, with considerable attention given to the development of a communication and stakeholder engagement policy. This engagement policy should accompany the introduction of a system of value-added modelling and include training for pertinent users. The actions and consequences for school principals, teachers and other stakeholders will need to be clearly articulated to not only build confidence in a new system but also to assuage fears of the introduction of a system that can be perceived as potentially lacking in fairness and transparency. Specific strategies will need to be developed that explain the system and educate stakeholders in how value-added scores are calculated and how they will be used. As is illustrated in Part III, successful strategies have been developed that highlight the benefits of value-added modelling compared with other performance measures. In a number of countries, stakeholders have welcomed the development and use of value-added modelling: its greater accuracy provides a fairer measure of school performance that creates more equitable systems of school accountability and school choice and fosters more accurate and therefore effective school improvement initiatives.

Also included is a discussion of the main steps that need to be undertaken in the implementation of a system of value-added modelling. The discussion of these steps is not meant to provide an exhaustive list of all activities that need to be undertaken but should assist policy makers and administrators who hope to gain a quick understanding of the process required in the implementation of a system of value-added modelling. This is presented as a small separate section at the end of Part I to emphasise the importance of implementation issues and their connection to specific policy objectives and uses of value-added modelling.