Innovative Applications of O.R.

Value added, educational accountability approaches and their effects on schools' rankings: Evidence from Chile[☆]

Claudio Thieme[a], Diego Prior[b,*], Emili Tortosa-Ausina[c], René Gempp[a]

[a] *Universidad Diego Portales, Chile*
[b] *Departament of Business, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain*
[c] *Universitat Jaume I and Ivie, Spain*

ARTICLE INFO

ABSTRACT

Value added models have been proposed to analyze different aspects related to school effectiveness on the basis of student growth. There is consensus in the literature about the need to control for socioeconomic status and other contextual variables at student and school level in the estimation of value added, for which the methodologies employed have largely relied on hierarchical linear models. However, this approach is problematic because results are based on comparisons to the school's average—implying no real incentive for performance excellence. Meanwhile, activity analysis models to estimate school value added have been unable to control for contextual variables at both the student and school levels. In this study we propose a robust frontier model to estimate contextual value added which merges relevant branches of the activity analysis literature, namely, metafrontiers and partial frontier methods. We provide an application to a large sample of Chilean schools, a relevant country to study due to the reforms made to its educational system that point out to the need of accountability measures. Results indicate not only the general relevance of including contextual variables but also how they contribute to explaining the performance differentials found for the three types of schools—public, privately-owned subsidized, and privately-owned fee-paying. Also, the results indicate that contextual value added models generate school rankings more consistent with the evaluation models currently used in Chile than any other type of evaluation models.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The development of indicators to evaluate the quality of education is a core element of countries' efforts to implement improvements in their education systems (Battauz, Bellio, & Gori, 2011). In many countries, this concern has motivated the adoption of accountability systems (Kane & Staiger, 2002), whose main objective is to evaluate school quality and report these results to parents, principals, teachers, or policy makers, who will use them to make choices about schools, to improve their professional practice or to develop educational policies.[1] The available empirical evidence in this regard has contributed to strengthen this tendency, showing that well designed accountability systems (i.e. those which find the responsibility attributable to each of the participants in the educational system) enable organizational improvement inside each school (Rouse, Hannaway, Goldhaber, & Figlio, 2007), as well as optimizing the educational outcomes (Carnoy & Loeb, 2002; Hanushek & Raymond, 2005). An underlying requisite of any accountability system is to use a robust methodology to disentangle what share of the students' achievement can be attributed to the school, and what share is simply the result of other variables beyond the school's control.

In terms of methodology, the general consensus is that students' educational achievement depends both on their personal characteristics as well as those of their school and context. In order to analyze these scenarios, the most common and ac-

[1] By way of example, see, for instance, the detailed information regarding school performance in the UK disclosed in http://www.education.gov.uk/schools/performance/.

cepted methodology is multilevel regression models (McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004), also known as hierarchical linear models, or regression models with random effects (Goldstein, 2003; Raudenbush & Bryk, 2002). The key characteristic of these methods is their capacity to disentangle what proportion of variance in student achievement can be explained by student variables (level 1), and what share can be explained by aggregate, or school, contextual variables (level 2). When multiple levels are considered, such as hierarchical systems of students nested in schools, it is possible to obtain a better understanding and measurement of the causes that explain students' learning processes (Aitkin & Longford, 1986). The multilevel approach is highly relevant when attempting to make decisions, specific to each student, school, or context, that contribute useful information to develop new improvement processes in schools, discourage managers' opportunistic behavior, signal a correct resource endowments policy (by establishing rewards and penalties), and make decisions on public policies.

The initial stages of research on school accountability were characterized by the use of cross sectional measures to estimate school performance (e.g. the mean annual results of standardized tests), but the current practice is to rely on panel data methods to evaluate student performance in order to estimate the academic growth of students throughout their school life—ideally also controlling for other relevant variables (Goldstein et al., 1993; Goldstein & Thomas, 1996; Gray, Jesson, Goldstein, Hedger, & Rasbash, 1995; Mortimore, Sammons, & Thomas, 1994; Sammons, 1995). In this context of school accountability, the value added (VA) of a given school can be broadly defined as the contribution that it makes to students' net progress (i.e. to the learning objectives) after the effects of other variables, external to the school, have been removed (Meyer, 1997). The basic value added model compares schools' performance controlling for students' previous achievement.

More complex value added models are also available, and over the last few years there has been a growing tendency to use *contextual* VA models, which allow researchers to control for socioeconomic status (SES), ethnic background, gender, and other variables that are not under the school's control or responsibility. Thus, contextual VA models provide an estimation of the net performance of schools by removing the effect of previous achievement and other preexisting differences among students (Ballou, Sanders, & Wright, 2004). It is generally agreed that contextual variables should be used to estimate VA models, especially when setting some form of accountability, or when disseminating the results, since results might be questionable if they do not take into account contextual characteristics of both students and schools. Although there is no consensus as to what specific contextual variables should be included in the model (Tekwe et al., 2004), socioeconomic status (SES) is usually one of them.

According to their characteristics, VA indicators emerge as an attractive methodology for several actors interested in measuring or improving school performance, including: (i) governments (which need to rely on objective accountability measures); (ii) politicians (who want to guarantee that the assessment of schools considers their ethnic and socioeconomic diversity); (iii) researchers (who need to study those factors contributing to school effectiveness using net indicators, which are not spuriously contaminated by the characteristics of students); (iv) teachers and school managers (who want objective measures of their performance, tuned for their specific student populations); (v) parents (who need to choose schools for their children according to their real capacity to add value to their students); and (vi) society as a whole, since this entails a more accurate and fair evaluation of the schools in the country (Drury & Doran, 2003; McCaffrey et al., 2004).

It is also crucial to understand that school effectiveness studies—including VA analysis—require using some kind of methodology to compare the schools being evaluated with a benchmark. In the case of VA research, the most popular methodology is multilevel regression models (see, for instance Goldstein et al., 1993; Gray et al., 1995; Cervini, 2009; Goddard & Goddard, 2001). An implicit assumption of this approach is to use the average school as a benchmark. However, this approach is not free from criticisms such as, for instance, that using the average as a benchmark is not an incentive for excellence (Bock, Wolfe, & Fisher, 1996; Kupermintz, 2003; McCaffrey, Lockwood, Koretz, & Hamilton, 2003), that according to traditional VA models test scores must be vertically scaled, and that the appropriate functional form for the model is not granted in advanced (Murphy, 2012; Ray, Evans, & McCormack, 2009).

An attractive approach to overcome this criticism is to consider the models derived from the activity analysis literature, which evaluate school performance by comparing any given school with the best observed performance. Instead of using a regression line as a benchmark, these methodologies consider a nonparametric frontier built either using Data Envelopment Analysis (DEA), or its nonconvex variant, namely, Free Disposal Hull (FDH).[2] In addition to explicitly defining an optimal benchmark, frontier models also allow several outputs to be used simultaneously (i.e. several concurrent measures of student and school performance), offering greater flexibility to estimate VA.

In this line of research, there has been a growing interest in developing approaches to estimate school effectiveness. For instance, Silva Portela and Thanassoulis (2001), De Witte, Thanassoulis, Simpson, Battisti, and Charlesworth-May (2010) and Portela, Camanho, and Keshvari, (2013) have developed methodologies to estimate basic VA models, whereas Thieme, Prior, and Tortosa-Ausina (2013) have proposed a model to analyse contextual effects in multilevel settings with cross-sectional data. However, the existing methodologies have not been able to estimate contextual VA, namely, to develop a frontier model able to estimate school VA effects controlling for students' previous achievement, and also for contextual variables at student and school levels. This development is crucial to further explore the use of frontier models to estimate contextual VA models in real world applications.

For this reason, the aim of this paper is both empirical and, to a lesser degree, methodological. Regarding the latter (at the methodological level) we propose a *robust* frontier model to estimate contextual value added (CVA) which combines both methodological contributions from multilevel modeling to school VA, as well as relevant proposals in the field of activity analysis methods—namely, the so-called *metafrontiers* (Battese, Rao, & O'Donnell, 2004) as well as the partial frontier methods (Cazals, Florens, & Simar, 2002).

Regarding the former (at the empirical level), we use this novel approach to analyze school effectiveness in Chile. This application is especially relevant for this country which, since the 1980s, has been implementing a series of reforms to its educational system (see Mizala & Romaguera, 2000), with strong emphasis on accountability measures. Among other reforms, the government transferred the management of public schools from the Ministry of Education to city councils, and allowed for the participation of private schools in the public system through a voucher system. Simultaneously, an accountability system was created, consisting

---

[2] We can also find parametric variants to this literature, among which SFA (Stochastic Frontier Analysis) is the most popular. Parametric and nonparametric methods have both advantages and disadvantages, some of which have been recently outlined by Badunenko, Henderson, and Kumbhakar (2012).