
education policy analysis archives

A peer-reviewed, independent,
open access, multilingual journal



Arizona State University

Volume 22 Number 97

October 20th, 2014

ISSN 1068-2341

Anticipating and Incorporating Stakeholder Feedback when Developing Value-Added Models

Ryan Balch

Baltimore City Schools

&

Cory Koedel

University of Missouri

United States

Citation: Balch, R., & Koedel, C. (2014). Anticipating and incorporating stakeholder feedback when developing value-added models. *Education Policy Analysis Archives*, 22(97).

<http://dx.doi.org/10.14507/epaa.v22.1701>

Anticipating and incorporating stakeholder feedback when developing value-added models.

Abstract: State and local education agencies across the United States are increasingly adopting rigorous teacher evaluation systems. Most systems formally incorporate teacher performance as measured by student test-score growth, sometimes by state mandate. An important consideration that will influence the long-term persistence and efficacy of these systems is stakeholder buy-in, including buy-in from teachers. In this study we document common questions from teachers about value-added measures and provide research-based responses to these questions. The questions come from teachers in Baltimore City Public Schools, who are evaluated using a combined measure of which value-added is one component. We focus on teacher questions about value-added because value-added generates the most concern from teachers. We also connect teacher concerns about value-added to other components of the evaluation system, such as classroom observations, although at present these other components

Journal website: <http://epaa.asu.edu/ojs/>

Facebook: /EPAAA

Twitter: @epaa_aape

Manuscript received: 3/25/2014

Revisions received: 06/06/2014

Accepted: 06/10/2014

have not garnered as much attention from teachers.

Keywords: teacher evaluation; value-added models; stakeholder engagement

Anticipando e incorporando la información de los interesados en el desarrollo de modelos de valor agregado.

Resumen: Agencias de educación estatales y locales en los Estados Unidos están adoptando cada vez más sistemas rigurosos de evaluación docente. La mayoría de los sistemas incorporan formalmente el desempeño docente, medida por el crecimiento de los resultados de prueba estandarizadas de los estudiantes, a veces por mandato estatal. Una consideración importante que influirá en la persistencia a largo plazo y la eficacia de estos sistemas es la aceptación de los interesados, incluida la aceptación por parte de los profesores. En este estudio documentamos preguntas comunes de los profesores acerca de las medidas de valor añadido y proporcionamos respuestas basadas en la investigación a estas preguntas. Las preguntas vienen de los maestros en Baltimore City Public Schools, que se evalúan mediante una medida combinada de los cuales un componente es el valor agregado. Nos centramos en las preguntas de los docentes sobre el valor añadido debido que es la que genera la mayor preocupación de los docentes. También relacionamos preocupaciones de los maestros sobre el valor añadido a otros componentes del sistema de evaluación, tales como observaciones en el aula, aunque en la actualidad estos otros componentes no han generado tanta atención.

Palabras clave: evaluación docente; modelos de valor agregado; participación de los interesados.

Antecipando e incorporando a informação dos interessados no desenvolvimento de modelos de valor agregado.

Resumo: As agências educacionais estatais e locais nos Estados Unidos estão adotando cada vez mais sistemas de avaliação mais rigorosos dos docentes. A maioria dos sistemas incorporam formalmente o desempenho dos docentes medidas pela evolução dos resultados dos alunos em testes padronizados, às vezes por mandato estadual. Uma consideração importante que irá influenciar a persistência a longo prazo a eficácia destes sistemas é a aceitação das partes interessadas, incluindo a aceitação por parte dos professores. Neste estudo documentamos perguntas comuns de professores sobre medidas de valor agregado e fornecemos respostas baseadas em pesquisas sobre essas perguntas. As perguntas vêm de professores das Escolas Públicas da cidade de Baltimore, que são avaliados por uma medida combinada onde um dos componentes e uma medida de valor agregado. Nós nos concentramos sobre as questões dos professores sobre o valor agregado, porque geraram a maior preocupação entre os docentes. Preocupações também relacionavam o valor agregado dos professores com outros componentes do sistema de avaliação, tais como observações em sala de aula, embora estes outros componentes não têm gerado muita atenção.

Palavras-chave: avaliação de professores; modelos de valor acrescentado; participação das partes interessadas.

Introduction

School districts and state education agencies across the United States are developing rigorous performance evaluation systems for teachers. These systems typically aim to produce “combined measures” for teachers that measure performance along multiple dimensions (Amrein-Beardsley & Barnett, 2012; Kane & Staiger, 2012; Dee & Wyckoff, 2013; Mihaly et al., 2013; Strunk, Weinstein, &

Makkonnen, forthcoming). The move to performance-based evaluations for teachers marks a fundamental shift away from the long-standing, qualifications-based approach. This is appealing because research has consistently documented large differences in teacher performance as measured by outputs (Chetty, Freedman, & Rockoff, 2014a/b; Hanushek & Rivkin, 2010; Rockoff, 2004) but has failed to link these differences to observable teacher qualifications (Kane, Rockoff, & Staiger, 2008; Koedel & Betts, 2007; Nye et al., 2004; Rivkin et al., 2005).

A vast literature on value-added has developed over the past 10-15 years (for a recent overview see Hanushek and Rivkin, 2010). This literature has greatly informed the current policy application of value-added models (VAMs) in teacher evaluation systems. The increasing prevalence of VAMs in this role has been supported and criticized by different segments of the scholarly community. On the one hand, proponents of using value-added point to the persistent information contained in value-added measures and argue that the benefits of using value-added likely outweigh the costs associated with its limitations (e.g., see Glazerman et al., 2010; Chetty, Friedman, & Rockoff, 2014 a/b). Critics caution against the over-reliance on value-added for evaluative purposes, raising concerns about the accuracy and stability of the measures and the potential for VAM-based incentives to narrow and over-simplify schooling curricula (e.g., see Amrein-Beardsley, 2014; Baker et al., 2010).

Against the backdrop of this ongoing scholarly discourse, state and local education agencies across the United States continue to move toward an increased reliance on test-based performance measures to inform decision making (Winters & Cowen, 2013). The academic literature underlying much of the discussion about model choice has focused mostly on the challenging task of developing statistically informative and reliable value-added measures for teachers, and work in this area is ongoing. However, the process of selecting a model, and the success of the model in achieving its policy objectives, involves more than just statistical considerations (e.g., also see Lincove, Osborne, Dillon, & Mills, 2014). Stakeholder support is important, and teachers are among the biggest stakeholders in these developing systems.

The contribution of the present study is to document and critically examine teacher feedback about value-added measures within the larger framework of a “combined measure” evaluation system. Underlying our interest in teacher perceptions about value-added is the view that teachers are important stakeholders in their own evaluations (Freeman, 1984). Incorporating their concerns into the evaluation process is one way to improve teachers’ active participation in control, which Jones (1997) argues is in the best interest of improving workforce efficacy.¹ As discussed by Ehlert, Koedel, Parsons, and Podgursky (forthcoming, 2014), there is substantial informational value embodied in recently-developed measures of school and teacher effectiveness. Teachers may be more likely to leverage these measures to improve instruction if they are more comfortable with the process by which they are developed.

The context for our study is the teacher evaluation system in Baltimore City Public Schools (BCPS), where teacher value-added accounts for up to 35 percent of teachers’ total ratings. Teacher concerns about the evaluation system were commonly focused on issues related to value-added. We document and discuss the four most common questions raised by BCPS teachers about measuring student growth using value-added. Drawing on available research evidence, we provide responses to these concerns that are methodologically accurate and proved useful for BCPS staff in communicating with teachers. We additionally offer suggestions for simple adjustments that can be

¹ Jones (1997) also advocates for employee participation in the returns to productivity. Given that his work is in the context of a private-market firm, he describes this as participation in financial returns – e.g., profit sharing. Analogous concepts could be developed in the education context but this extends beyond the scope of the present study.

made to standard models to appease teacher concerns, even in cases where the statistical implications of doing so may seem small.

We also briefly examine whether teacher concerns about value-added are relevant for the other components of combined measures of teaching effectiveness. The most common alternative components of combined measures are teacher observations, student surveys, and professional-expectation measures. Teacher observations and professional expectations are currently part of the BCPS system and student surveys are being field tested for possible future inclusion. Although teachers in Baltimore appear to be more concerned with value-added than the other measures, we discuss how many of the issues that teachers raise about value-added are also relevant for other combined-measure components. It would be proactive for school districts, state education agencies and the research community to work together to develop a larger and more-rigorous evidence base on these non-growth-based performance measures and their statistical properties. This will allow for more effective responses to stakeholder concerns that are likely to arise as combined-measure teacher evaluation systems continue to mature.

Background

The Teacher Effectiveness Evaluation System in BCPS

Baltimore City Public Schools implemented the Teacher Effectiveness Evaluation system during the 2013-2014 school year. The system evaluates teachers based on professional practice and student growth as shown in Figure 1.

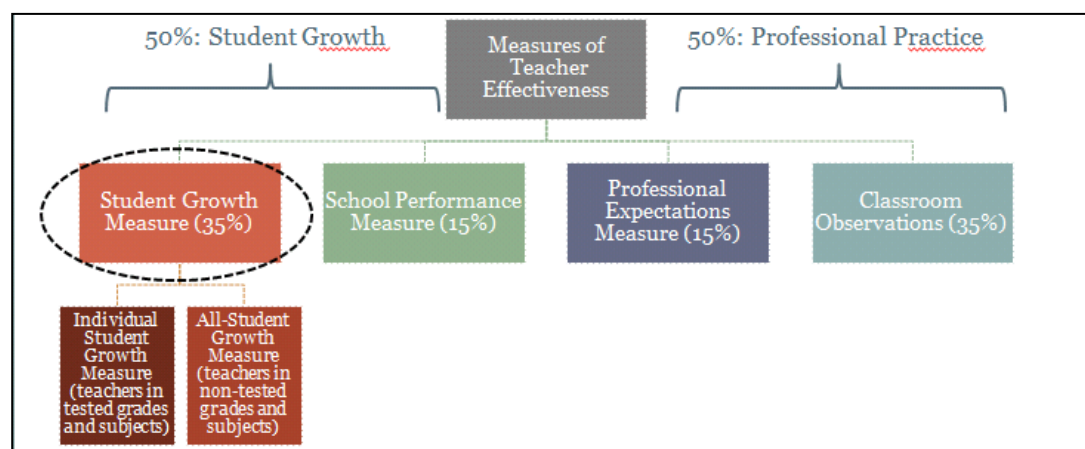


Figure 1. Teacher effectiveness evaluation in Baltimore City Schools

Fifty percent of teachers' evaluation scores derive from measures of student growth, which are subdivided into a school-wide component (15 percent) and a teacher-level component (35 percent). The school performance measure incorporates school-level measures of achievement, growth and school climate. In tested grades and subjects, the teacher-level component is based on value-added to test scores for the students assigned to individual teachers. Student learning objectives (SLOs) are being field-tested in BCPS this year for teachers in non-tested grades and subjects and will be used for the teacher-level component beginning in 2014-2015 (school-level value-added was used in 2013-2014 in place of the teacher-level component for these teachers given the absence of an alternative).

The other 50 percent of teachers' evaluation scores come from classroom observations (35 percent) and professional expectation measures (15 percent), similarly to other "combined measure"

evaluation systems. We do not provide a lengthy discussion of the evaluation rubrics for these components of the evaluation system because the focus of the present study is on the value-added measures. Interested readers can find more information about the other components of the system through Baltimore City Schools’ website (see <http://www.baltimorecityschools.org/Page/23121>).

The value-added model. As noted above, the contribution of the present study is to document and address key teacher concerns about the value-added component of the evaluation system. Based on multiple information and training sessions with teachers, school principals and union representatives, and a districtwide survey administered by BCPS in the spring of 2013, teacher concerns about the evaluation system were commonly focused on value-added and, in particular, teacher-level value-added.

The BCPS value-added model was developed in collaboration with the American Institutes for Research (AIR). The model predicts students’ current-year test score using prior scores and information about the student and school. Table 1 lists the variables that are used as controls in the BCPS model. The goal of the model is to produce a conditional expected score for each student. The conditional expected score – or “predicted score” – depends on the student’s prior score history, individual characteristics, and his or her schooling environment (per the control variables listed in Table 1). Teacher performance is evaluated by looking for systematic deviations from expectation for students taught by a particular teacher. A teacher whose students perform exactly as expected is exactly average. A teacher whose students systematically exceed their expected scores is above average.

Table 1
Control Variables Included in the BCPS Value-Added Model

Student-level	School-level
Prior test-score performance*	School FARMS percentage
Prior-year attendance	School Special Education percentage
English-language learner status	Percent of students out of age level
Skipped-grade indicator	Average number of special education hours per student
Free and Reduced Meals (FARMS) status	
Least Restrictive Environment Status	
Student mobility	
Different from modal age in grade	
Repeater status	

* Prior test score controls depend on the grade and subject of the model. Value-added models are estimated for students in grades 2 through high school (in some courses).

The BCPS model is conceptually similar to models used in other locales including Washington DC (Isenberg & Hock, 2011), New York City (Value-Added Research Center & New York City Department of Education, 2010) and Pittsburgh (Johnson et al., 2012), among others; as well as models used to estimate teacher value-added in the academic literature (e.g., see Aaronson, Barrow, & Sander, 2007; Goldhaber & Hansen, 2013; Sass et al., 2012). While there are a number of technical features that differentiate the BCPS approach to measuring teacher value-added, we avoid a lengthy discussion of the technical aspects of the models here. The most notable difference is that teacher effects in BCPS are specified as random, and estimated within a hierarchical framework that also accounts for random school effects. The other models mentioned in the text estimate teacher effects as fixed (we refer the interested reader to Wooldridge, 2010, for a discussion of the tradeoffs associated with specifying teacher effects as random versus fixed). Of importance for the present

study is the qualitative similarity between the BCPS model and other available models. As we describe in the next section, the key distinguishing feature of the value-added approach relative to the common alternative – and the alternative available to BCPS in Maryland (see Section 2.3) – is that the value-added model explicitly conditions on student and schooling circumstance to construct predicted test scores for students.

Statewide context and the policy rationale for the BCPS VAM. While all districts in Maryland are required to evaluate teachers based in part on student performance, BCPS is the only district in Maryland that uses a value-added model for this purpose. The Maryland Department of Education allows districts to create local models for teacher evaluation provided they meet certain criteria. The criteria stipulate that (a) there is a 50/50 split between measures of professional practice and student growth, (b) at least 20 percent of the evaluation is based on state standardized tests, and (c) a maximum of 35 percent weight can be applied to any one measure of student growth.

If a district opts not to create a local model, it is obligated to use the default state model. The measure of student growth within the state's model is known as the Maryland Tiered Achievement Index (M-TAI), which is not VAM-based. M-TAI divides each of the three levels of achievement (basic, proficient, and advanced) into three performance levels (low-basic, basic, high-basic, low-proficient, etc.). Students accumulate points for moving up through the performance levels.

The fundamental issue with the Maryland measure that led to BCPS's decision to develop a value-added model for teacher evaluation is that it does not control for any student or school characteristics. Thus, it implicitly assumes that students at all levels of achievement, from all socioeconomic backgrounds, and in all schooling environments, can achieve similar gains. However, it is an empirical fact that students in different circumstances do not achieve similar gains (McCaffrey, 2013; Meyer et al., 2009). The failure of M-TAI to account for schooling context is particularly problematic for BCPS, which is a high-poverty district where more than 85 percent of students qualify for free or reduced price lunch. Although BCPS has high expectations for all students, there is danger in conflating expectations for students with expectations for personnel (Ehlert et al., forthcoming, 2014). As will become clear below, teacher feedback on the model suggests that teachers prefer the BCPS value-added approach to the alternative provided by the Maryland state department of education.

Addressing and Responding to Key Teacher Concerns about the Value Added Model

We now turn to the key concerns raised by teachers about the value-added model at BCPS. As noted above, teacher feedback about value-added was collected from several sources: (1) a district wide survey that was administered in the spring of 2013 yielding responses from 497 teachers, with 68 percent of teachers noting particular interest in value-added, (2) five information sessions that were conducted for both teachers and principals during the field test of the evaluation system, also in the spring of 2013, and (3) eight summer training sessions on student growth that were conducted prior to the implementation of the growth model in the summer preceding the 2013-2014 school year.

The district wide survey provides an indication of teachers' general concerns about value-added. Although it was not designed to elicit detailed questions from teachers about value-added in particular, 68 percent of teachers indicated on the survey that value-added was an aspect of the new evaluation system about which they wanted additional information. Teacher inquiries during the general information sessions and student-growth training sessions were collected to identify specific concerns about value-added. The general information sessions were held during the school year at locations throughout the city. The summer training sessions were held for principals, assistant

principals, and union representatives from each school. In-session questions, and questions submitted in writing afterward, were documented by staff members and analyzed for frequency and common themes. The objective of the data collection effort was to improve the quality of the sessions moving forward.

Based on the documentation by BCPS staff, we identify four key issues that consistently came up with regard to the use of value-added for teacher evaluations:

1. *Differentiated Students*. How can the model deal with a teacher who has students who are different for some reason (e.g., poverty, special education, etc.)? Will that teacher be treated unfairly by the model?
2. *Student Attendance*. Will teachers be held accountable for students who do not regularly attend class?
3. *Outside Events and Policies*. How can the model account for major events (e.g., school closings for snow) or initiatives (e.g., Common Core implementation) that impact achievement?
4. *Ex Ante Expectations*. Why can't teachers have their predicted scores – the target average performance levels for their students – in advance?

Below we elaborate on each of these questions and provide recommendations for how to respond to teachers drawn from the experiences of BCPS staff.

Differentiated Students. The unique features of teachers' students and classrooms are often of primary concern in discussions on measuring student growth. One of the most useful tools in allaying fears about value-added is to describe how value-added assigns predicted scores to students based on average growth for students sharing similar characteristics and in similar environments within the district. The control variables in the BCPS model allow for a response to teachers along the lines of: "We look at average growth in this district for students that look like your students and attend similar schools. Then we ask whether your students do better, the same, or worse than these other students." This response conveys two critical points. First, it highlights the fact that students in Baltimore City are compared to other students in Baltimore City. Second, it allows teachers to feel that it is more about how they teach than about who they teach. This sentiment is supported by research showing that the scope for bias in standard value-added models is small (e.g., see Chetty, Friedman, & Rockoff, 2014a; Kane et al., 2013). Note that such a response would not be possible if BCPS had adopted a sparser model along the lines of the default Maryland state model.

Teachers of special education students are particularly concerned about having a fair basis for comparison when measuring student growth. It is common in the development of value-added models to use a 0/1 indicator variable to account for student special-education status. However, based on teacher feedback, BCPS has improved on the specificity of coding for special-education students by creating a variable called the Least Restrictive Environment (LRE) for these students. The LRE variable reflects the percentage of time that students spend in a general education setting. Students coded as LRE-A spend 80 percent or more of the time in general education, LRE-B students spend between 40-80 percent of the time in a general education classroom, and LRE-C students spend less than 40 percent of the time in a general education classroom.

The LRE variable directly acknowledges the variability within special education in K-12 schools, which is an important concern for teachers of special education students. Although correlational analyses suggest that the model with the finer LRE controls produces estimated teacher effects that are highly correlated with estimates from the coarser model overall, the value of the LRE controls in terms of facilitating stakeholder participation in control (as discussed by Jones, 1997) is noteworthy. Furthermore, recent studies have pointed out that high correlations across models overall can mask important differences in output for some groups of teachers and schools (Goldhaber, Walch, & Gabele, 2013; Ehlert et al., 2013).

Student Attendance. Student attendance is a key concern for teachers, particularly at the middle and high school levels in an urban school system. Teachers are worried that they will be judged based on the performance of students to whom they are not consistently exposed. In response to teacher feedback, and as shown in Table 1, the BCPS model controls for students' prior-year attendance to address teacher concerns in this regard.

While teachers appreciate the model's explicit accounting of attendance, a common request is that *current year* attendance be including as a control variable. While a statistician would recognize that current year attendance is endogenous, teachers are unlikely to appreciate such an explanation. An alternative, more-effective strategy for BCPS has been to frame the issue in terms of the role played by teachers in determining current attendance. Teachers are typically receptive to the idea that they can have a positive effect on attendance, which is supported by research (e.g., see Duckworth & DeJung, 1989; Roderick et al., 1997). However, if the model controls for attendance directly then teachers would not receive credit for their influence in this way. BCPS tells teachers: "We want to make sure that if a teacher improves a student's attendance, and this helps improve achievement, we don't take away from the credit that the teacher receives." This is a layman's way of explaining the endogeneity problem. Coupled with the fact that the BCPS model controls for lagged attendance, which is a proxy for attendance that is not directly affected by the current teacher(s), this explanation has been a successful communication strategy for BCPS.

Outside Events and Policies. Every year there are major events (e.g., school closings for snow) and/or new policies and procedures are implemented (e.g., the Common Core) in the District. Sometimes, these events and policies are experienced by all schools in Maryland and, other times, the events are unique to BCPS. A common teacher concern is that these events will impact student achievement, and in turn, teachers' value-added scores.

BCPS has communicated to teachers that the model is constructed to examine how teachers perform compared to average growth in the district. If all students in Baltimore are affected by the event then everyone remains on a level playing field. As a specific example, it can be conveyed to teachers that, "If all students score lower because of excessive school closings due to snow, then average growth for students just like the ones in your classroom will also be lower." Teacher feedback during the informational and training sessions at BCPS indicates that they value the locality of the BCPS model. BCPS would have had difficulty in addressing this common teacher concern if it had opted to compare BCPS teachers to teachers statewide without accounting for variation across districts in the length of the school year, snow events, curriculum changes, etc. (as would have been the case with the Maryland default model).

Ex Ante Expectations. BCPS describes value-added scores to teachers as coming from the average difference between students' predicted and actual scores. Although the statistics underlying how the predictions are generated are difficult to convey, the BCPS experience has been that describing the concept of value-added in this way improves teacher understanding (also see this tutorial provided by the Value-Added Research Center at the University of Wisconsin-Madison: <http://varc.wceruw.org/tutorials/oak/index.htm>). However, explaining value-added as the difference between students' actual and predicted performance leads to a new challenge: teachers logically request that they receive their predictions in advance.

This is an obvious and sensible thing to ask for. If teachers are going to be held to certain performance expectations, then why can those expectations not be laid out clearly *ex ante*? Of course, the problem is that student predictions can only be generated after the end-of-year test – the estimation of predicted student scores and teacher value-added occurs simultaneously. This must be the case for a number of reasons, most notably to protect teachers and the district from unanticipated events that may occur over the course of the year. BCPS has explained the timing

problem to teachers by focusing on the possibility of these events and providing concrete examples, such as the aforementioned possibility that the end-of-year test will have surprising and/or undesirable properties. Without the ability to predict these events into the future it is not possible to provide accurate *ex ante* predictions for students. For BCPS staff, this response has proved to be more effective than simply telling teachers that scores are not available. Through use of examples, teachers recognize that it is in their best interest that predicted scores for students are not formulated until after the year is complete.

Extending Teacher Concerns to Other Evaluation Components

The primary purpose of this paper is to identify and discuss common questions raised by teachers about value-added. We focus on value-added because it has garnered the most attention from teachers in BCPS. However, we also note that some of the issues raised by teachers in the context of value-added are also likely to apply to other components in combined measures of teaching effectiveness.

Currently, the main concerns about measures of professional practice voiced by teachers in BCPS apply to the consistency of observational ratings among principals. BCPS has responded to these concerns by implementing an observer certification program. The program requires principals to watch three or more sample videos and rate teachers according to the evaluation rubric, known as the Instructional Framework. Principal ratings must meet minimum standards for alignment with pre-determined rankings for each video in order to be able to observe teachers.

While the observer certification program is designed to address teachers' current concerns, it is reasonable to expect that in the future teachers may wish to apply some of the principles embedded in the value-added measures to the observational measures, such as their conditional nature. To date, we are not aware of a single district or state education agency in the United States that has implemented a *conditional* measure of observational teaching performance. That is, teacher observation scores do not account for student or schooling circumstance. In this way, they are more like the uncontrolled Maryland achievement metric (M-TAI) than the BCPS value-added model.

It is beyond the scope of the present study to empirically examine the consequences of the unconditional nature of teachers' observational performance scores. However, recent evidence indicates that there are systematic differences in observation scores across teachers working in different environments (Whitehurst, Chingos, & Lindquist, 2014). Over time, these systematic differences will become more apparent and decisions about how to handle them will need to be made by administrators. Based on the questions and concerns raised by teachers in BCPS about value-added, we expect teachers to favor conditional measures of observational performance. The determination of whether and how to implement a conditional observation metric would benefit from additional research. Compared to the vast research literature on value-added, the empirical literature on observational and other, non-VAM-based measures of teaching performance is thin (with much of the evidence coming from the MET project – e.g., see Kane & Staiger, 2012; Mihaly et al., 2013).

Concluding Remarks

The current study examines feedback about value-added performance measures from teachers and principals in Baltimore City Public Schools. BCPS has adopted a local teacher evaluation model that deviates from Maryland's default state model in an effort to account for local district context and control for student characteristics when evaluating the performance of personnel. We document specific questions and concerns raised during the field testing and

implementation of the BCPS value-added model, which is used as one component in a larger “combined measure” teacher evaluation system.

Key teacher concerns in BCPS include accounting for student characteristics, addressing student attendance, controlling for local events that may impact achievement, and explaining the lack of availability of *ex ante* student-performance predictions. Policymakers and researchers working with district and state education agencies may benefit from incorporating this stakeholder feedback into value-added models and the discussions that surround them. This has the potential to increase teacher engagement and help promote the sustainability of evaluation systems that can be useful for improving instruction.

References

- Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, 25(1), 95-135.
<http://dx.doi.org/10.1086/508733>
- Amrein-Beardsley, A. (2014). *Rethinking value-added models in education: Critical perspectives on tests and assessment-based accountability*. New York, NY: Routledge.
- Amrein-Beardsley, A. & Barnett, J. H. (2012). Working with error and uncertainty to increase measurement validity. *Educational Assessment, Evaluation and Accountability*, 24(4), 369-379.
<http://dx.doi.org/10.1007/s11092-012-9146-6>
- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., ... Shepard, L. A. (2010). Problems with the use of student test scores to evaluate teachers. Economic Policy Institute Briefing Paper. Retrieved from <http://www.epi.org/publication/bp278/>
- Chetty, R., Friedman, J. N., & Rockoff, J.E. (2014a). Measuring the impact of teachers I: Evaluating bias in teacher value-Added estimates. *American Economic Review*, 104(9), 2593-2632.
<http://dx.doi.org/10.1257/aer.104.9.2593>
- Chetty, R., Friedman, J.N., & Rockoff, J.E. (2014b). Measuring the impact of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9), 2633-2679. <http://dx.doi.org/10.1257/aer.104.9.2633>
- Thomas, D., & Wyckoff, J. (2013). Incentives, selection and teacher performance. Evidence from IMPACT. NBER Working Paper No. 19529. Retrieved from <http://www.nber.org/papers/w19529>
- Duckworth, K., & DeJung, J. (1989). Inhibiting class cutting among high school students. *The High School Journal*, 72(4), 188-195.
- Ehlert, M., Koedel, C., Parsons, E. & Podgursky, M. (forthcoming). Selecting growth measures for use in school evaluation systems: Should proportionality matter? *Educational Policy*. Retrieved from http://economics.missouri.edu/working-papers/2012/WP1210_koedel.pdf
- Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2014). Choosing the right growth measure: Methods should compare similar schools and teachers. *Education Next*, 14(2), 66-71.
- Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2013). The sensitivity of value-added estimates to specification adjustments: Evidence from school- and teacher-level models in Missouri. *Statistics and Public Policy*, 1(1), 19-27.
- Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. New York, NY: Cambridge University Press.
- Glazerman, S., Goldhaber, D., Loeb, S., Raudenbush, S., Staiger, D., & Whitehurst, G.J. (2010). Evaluating teachers: The important role of value-added. Brown Center on Education Policy

- at Brookings Policy Report. Retrieved from <http://www.brookings.edu/research/reports/2010/11/17-evaluating-teachers>
- Goldhaber, D., & Hansen, M. (2013). Is it just a bad class? Assessing the long-term stability of estimated teacher performance. *Economica*, 80(319), 589-612. <http://dx.doi.org/10.1111/ecca.12002>
- Goldhaber, D., Walch, J., & Gabele, B. (2013). Does the model matter? Exploring the relationship between different student achievement-based teacher assessments. *Statistics and Public Policy*, 1(1), 28-39.
- Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review*, 100(2), 267-271. <http://dx.doi.org/10.1257/aer.100.2.267>
- Isenberg, E., & Hock, H. (2010). Measuring school and teacher value added for IMPACT and TEAM in DC Public Schools: Final report. Unpublished report, Mathematica Policy Research. Retrieved from http://communications.mathematica-mpr.com/publications/pdfs/education/valueadded_tchrprt.pdf
- Johnson, M., Lipscomb, S., Gill, B., Booker, K., & Bruch, J. (2012). Value-added model for Pittsburgh Public Schools. Unpublished report, Mathematica Policy Research. Retrieved from <http://files.eric.ed.gov/fulltext/ED531790.pdf>
- Jones, D. (1997). Employees as stakeholders. *Business Strategy Review*, 8(2), 21-24. <http://dx.doi.org/10.1111/1467-8616.00016>
- Kane, T. J., & Staiger, D. O. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. Seattle, WA: Bill & Melinda Gates Foundation. Retrieved from http://www.metproject.org/downloads/MET_Gathering_Feedback_Practioner_Brief.pdf
- Kane, T. J., McCaffrey, D.F., Miller, T., & Staiger, D. O. (2013). *Have we identified effective teachers? Validating measures of effective teaching using random assignment*. Seattle, WA: Bill & Melinda Gates Foundation.
- Kane, T. J., Rockoff, J.E., & Staiger, D.O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, 27(6), 615-631. <http://dx.doi.org/10.1016/j.econedurev.2007.05.005>
- Koedel, C., & Betts, J.R. (2007). Re-examining the role of teacher quality in the educational production function. University of Missouri Working Paper No. 07-08.
- Lincove, J. A., Osborne, C., Dillon, A., & Mills, N. (2014). The politics and statistics of value-added modeling for accountability of teacher preparation Programs. *Journal of Teacher Education*, 65(1), 24-38. <http://dx.doi.org/10.1177/0022487113504108>
- McCaffrey, D. F. (2012). Do value-added methods level the playing field for teachers? Policy Report, Carnegie Foundation for the Advancement of Teaching. Retrieved from http://www.carnegieknowledgegenetwork.org/wp-content/uploads/2013/06/CKN_2012-10_McCaffrey.pdf
- Meyer, R. H., Dokumaci, E., Geraghty, E.L., McLaughlin, S., & Morgan, E.A. (2009). Wisconsin value-added model: A demonstration project. Final Report. Value-Added Research Center. Retrieved from <http://oea.dpi.wi.gov/files/oea/pdf/vademons.pdf>
- Mihaly, K., McCaffrey, D.F., Staiger, D.O., & Lockwood, J.R. (2013). A composite estimator of effective teaching. *RAND External Publication*, EP-50155.
- Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237-257. <http://dx.doi.org/10.3102/01623737026003237>
- Rivkin, S. G., Hanushek, E.A. & Kain, J.F. (2005). Teachers, schools and academic achievement. *Econometrica*, 73(2), 417-58. <http://dx.doi.org/10.1111/j.1468-0262.2005.00584.x>

- Rockoff, J. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review (P&P)*, 94(2), 247-252. <http://dx.doi.org/10.1257/0002828041302244>
- Roderick, M., Arney, M., Axelman, M., DaCosta, K., Steiger, C., Stone, S., ... Waxman, E. (1997). Habits hard to break: A new look at truancy in Chicago's public high schools. Research Brief, School of Social Service Administration at the University of Chicago. Retrieved from <https://ccsr.uchicago.edu/sites/default/files/publications/p0a09.pdf>
- Sass, T. R., Hannaway, J., Xu, Z., Figlio, D.N., & Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics*, 72(2-3), 104-122. <http://dx.doi.org/10.1016/j.jue.2012.04.004>
- Strunk, K. O., Weinstein, T.L., & Makkonnen, R. (forthcoming). Sorting out the signal: Do multiple measures of teachers' effectiveness provide consistent information to teachers and principals? *Education Policy Analysis Archives*.
- The New Teacher Project. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. TNTP Policy Report.
- Value-Added Research Center. (2010). NYC teacher data initiative: Technical report on the NYC value-added model. Unpublished report, Wisconsin Center for Education Research. University of Wisconsin-Madison. Retrieved from <http://schools.nyc.gov/NR/rdonlyres/A62750A4-B5F5-43C7-B9A3-F2B55CDF8949/87046/TDINYCTechnicalReportFinal072010.pdf>
- Whitehurst, G. J., Chingos, M.M., & Lindquist, K.M. (2014). Evaluating teachers with classroom observations: Lessons learned in four districts. Brown Center on Education Policy at Brookings Policy Report. Retrieved from <http://www.brookings.edu/research/reports/2014/05/13-teacher-evaluation-whitehurst-chingos>
- Winters, M. A., & Cowen, J.M. (2013). Would a value-added system of retention improve the distribution of teacher quality? A simulation of alternative policies. *Journal of Policy Analysis and Management*, 32(3), 634-654. <http://dx.doi.org/10.1002/pam.21705>
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (Second Edition). Cambridge, MA: MIT Press.

About the Authors

Ryan Balch

Baltimore City Schools

ryanbalch@mystudentsurvey.com

Ryan Balch was the director of teacher, principal, and school evaluation for Baltimore City schools from 2012-2014. During this time, he oversaw the creation and implementation of the district's new systems of evaluation. Currently, Dr. Balch is the CEO of My Student Survey, a company that supports districts and states in the administration and development of student surveys of teacher practice. Ryan Balch completed his PhD in Education Policy at Vanderbilt University, where his dissertation focused on the development and validation of student surveys on teacher practice. He was the principal investigator for the student survey pilot of more than 15,000 students in seven districts as part of Race to the Top. Previously, Ryan worked as a science teacher and department chair for seven years in Georgia at Riverwood High School. He has a master's degree in Science Education from Georgia State University and a B.A. in Psychology from Duke University.

Cory Koedel

University of Missouri

koedelc@missouri.edu

Cory Koedel is an associate professor of economics and public policy at the University of Missouri, Columbia. His research is in the areas of teacher quality and compensation, curriculum evaluation, school choice and the efficacy of higher education institutions. His work has been widely cited in top academic journals in the fields of economics, education and public policy, and he serves on several technical advisory panels related to school and teacher evaluations for school districts, state education agencies and non-profit organizations. Dr. Koedel was awarded the Outstanding Dissertation Award from the American Educational Research Association (Division L) in 2008, and in 2012 he received the Junior Scholar Award from the same group. He received his PhD in economics from the University of California, San Diego in 2007.

education policy analysis archives

Volume 22 Number 97 October 20th, 2014 ISSN 1068-2341



Readers are free to copy, display, and distribute this article, as long as the work is attributed to the author(s) and **Education Policy Analysis Archives**, it is distributed for non-commercial purposes only, and no alteration or transformation is made in the work. More details of this Creative Commons license are available at <http://creativecommons.org/licenses/by-nc-sa/3.0/>. All other uses must be approved by the author(s) or **EPAA**. **EPAA** is published by the Mary Lou Fulton Institute and Graduate School of Education at Arizona State University. Articles are indexed in CIRC (Clasificación Integrada de Revistas Científicas, Spain), DIALNET (Spain), [Directory of Open Access Journals](#), EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), QUALIS A2 (Brazil), SCImago Journal Rank; SCOPUS, Socolar (China).

Please contribute commentaries at <http://epaa.info/wordpress/> and send errata notes to Gustavo E. Fischman fischman@asu.edu

Join EPAA's Facebook community at <https://www.facebook.com/EPAAAPE> and **Twitter feed** @epaa_aape.

education policy analysis archives
editorial board

Editor **Gustavo E. Fischman** (Arizona State University)

Associate Editors: **Audrey Amrein-Beardsley** (Arizona State University), **Rick Mintrop**, (University of California, Berkeley)
Jeanne M. Powers (Arizona State University)

Jessica Allen University of Colorado, Boulder

Gary Anderson New York University

Michael W. Apple University of Wisconsin, Madison

Angela Arzubiaga Arizona State University

David C. Berliner Arizona State University

Robert Bickel Marshall University

Henry Braun Boston College

Eric Camburn University of Wisconsin, Madison

Wendy C. Chi* University of Colorado, Boulder

Casey Cobb University of Connecticut

Arnold Danzig Arizona State University

Antonia Darder University of Illinois, Urbana-Champaign

Linda Darling-Hammond Stanford University

Chad d'Entremont Strategies for Children

John Diamond Harvard University

Tara Donahue Learning Point Associates

Sherman Dorn University of South Florida

Christopher Joseph Frey Bowling Green State University

Melissa Lynn Freeman* Adams State College

Amy Garrett Dikkers University of Minnesota

Gene V Glass Arizona State University

Ronald Glass University of California, Santa Cruz

Harvey Goldstein Bristol University

Jacob P. K. Gross Indiana University

Eric M. Haas WestEd

Kimberly Joy Howard* University of Southern California

Aimee Howley Ohio University

Craig Howley Ohio University

Steve Klees University of Maryland

Jackyung Lee SUNY Buffalo

Christopher Lubienski University of Illinois, Urbana-Champaign

Sarah Lubienski University of Illinois, Urbana-Champaign

Samuel R. Lucas University of California, Berkeley

Maria Martinez-Coslo University of Texas, Arlington

William Mathis University of Colorado, Boulder

Tristan McCowan Institute of Education, London

Heinrich Mintrop University of California, Berkeley

Michele S. Moses University of Colorado, Boulder

Julianne Moss University of Melbourne

Sharon Nichols University of Texas, San Antonio

Noga O'Connor University of Iowa

João Paraskveva University of Massachusetts, Dartmouth

Laurence Parker University of Illinois, Urbana-Champaign

Susan L. Robertson Bristol University

John Rogers University of California, Los Angeles

A. G. Rud Purdue University

Felicia C. Sanders The Pennsylvania State University

Janelle Scott University of California, Berkeley

Kimberly Scott Arizona State University

Dorothy Shipps Baruch College/CUNY

Maria Teresa Tatto Michigan State University

Larisa Warhol University of Connecticut

Cally Waite Social Science Research Council

John Weathers University of Colorado, Colorado Springs

Kevin Welner University of Colorado, Boulder

Ed Wiley University of Colorado, Boulder

Terrence G. Wiley Arizona State University

John Willinsky Stanford University

Kyo Yamashiro University of California, Los Angeles

* Members of the New Scholars Board

archivos analíticos de políticas educativas
consejo editorial

Editor: **Gustavo E. Fischman** (Arizona State University)

Editores. Asociados **Alejandro Canales** (UNAM) y **Jesús Romero Morante** (Universidad de Cantabria)

Armando Alcántara Santuario Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México

Claudio Almonacid Universidad Metropolitana de Ciencias de la Educación, Chile

Pilar Arnaiz Sánchez Universidad de Murcia, España

Xavier Besalú Costa Universitat de Girona, España

Jose Joaquín Brunner Universidad Diego Portales, Chile

Damián Canales Sánchez Instituto Nacional para la Evaluación de la Educación, México

María Caridad García Universidad Católica del Norte, Chile

Raimundo Cuesta Fernández IES Fray Luis de León, España

Marco Antonio Delgado Fuentes Universidad Iberoamericana, México

Inés Dussel FLACSO, Argentina

Rafael Feito Alonso Universidad Complutense de Madrid, España

Pedro Flores Crespo Universidad Iberoamericana, México

Verónica García Martínez Universidad Juárez Autónoma de Tabasco, México

Francisco F. García Pérez Universidad de Sevilla, España

Edna Luna Serrano Universidad Autónoma de Baja California, México

Alma Maldonado Departamento de Investigaciones Educativas, Centro de Investigación y de Estudios Avanzados, México

Alejandro Márquez Jiménez Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México

José Felipe Martínez Fernández University of California Los Angeles, USA

Fanni Muñoz Pontificia Universidad Católica de Perú

Imanol Ordorika Instituto de Investigaciones Económicas – UNAM, México

María Cristina Parra Sandoval Universidad de Zulia, Venezuela

Miguel A. Pereyra Universidad de Granada, España

Monica Pini Universidad Nacional de San Martín, Argentina

Paula Razquin UNESCO, Francia

Ignacio Rivas Flores Universidad de Málaga, España

Daniel Schugurensky Universidad de Toronto-Ontario Institute of Studies in Education, Canadá

Orlando Pulido Chaves Universidad Pedagógica Nacional, Colombia

José Gregorio Rodríguez Universidad Nacional de Colombia

Miriam Rodríguez Vargas Universidad Autónoma de Tamaulipas, México

Mario Rueda Beltrán Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México

José Luis San Fabián Maroto Universidad de Oviedo, España

Yengny Marisol Silva Laya Universidad Iberoamericana, México

Aida Terrón Bañuelos Universidad de Oviedo, España

Jurjo Torres Santomé Universidad de la Coruña, España

Antoni Verger Planells University of Amsterdam, Holanda

Mario Yapu Universidad Para la Investigación Estratégica, Bolivia

arquivos analíticos de políticas educativas
conselho editorial

Editor: **Gustavo E. Fischman** (Arizona State University)
Editores Associados: **Rosa Maria Bueno Fisher** e **Luis A. Gandin**
(Universidade Federal do Rio Grande do Sul)

Dalila Andrade de Oliveira Universidade Federal de Minas Gerais, Brasil
Paulo Carrano Universidade Federal Fluminense, Brasil
Alicia Maria Catalano de Bonamino Pontifícia Universidade Católica-Rio, Brasil
Fabiana de Amorim Marcello Universidade Luterana do Brasil, Canoas, Brasil
Alexandre Fernandez Vaz Universidade Federal de Santa Catarina, Brasil
Gaudêncio Frigotto Universidade do Estado do Rio de Janeiro, Brasil
Alfredo M Gomes Universidade Federal de Pernambuco, Brasil
Petronilha Beatriz Gonçalves e Silva Universidade Federal de São Carlos, Brasil
Nadja Herman Pontifícia Universidade Católica –Rio Grande do Sul, Brasil
José Machado Pais Instituto de Ciências Sociais da Universidade de Lisboa, Portugal
Wenceslao Machado de Oliveira Jr. Universidade Estadual de Campinas, Brasil

Jefferson Mainardes Universidade Estadual de Ponta Grossa, Brasil
Luciano Mendes de Faria Filho Universidade Federal de Minas Gerais, Brasil
Lia Raquel Moreira Oliveira Universidade do Minho, Portugal
Belmira Oliveira Bueno Universidade de São Paulo, Brasil
Antônio Teodoro Universidade Lusófona, Portugal
Pia L. Wong California State University Sacramento, U.S.A
Sandra Regina Sales Universidade Federal Rural do Rio de Janeiro, Brasil
Elba Siqueira Sá Barreto Fundação Carlos Chagas, Brasil
Manuela Terrasêca Universidade do Porto, Portugal
Robert Verhine Universidade Federal da Bahia, Brasil
Antônio A. S. Zuin Universidade Federal de São Carlos, Brasil