

Gershenson, Seth

Working Paper

Should Value-Added Models Control for Student Absences?

IZA Discussion Papers, No. 9978

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Gershenson, Seth (2016) : Should Value-Added Models Control for Student Absences?, IZA Discussion Papers, No. 9978, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/142417>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 9978

Should Value-Added Models Control for Student Absences?

Seth Gershenson

May 2016

Should Value-Added Models Control for Student Absences?

Seth Gershenson

*American University
and IZA*

Discussion Paper No. 9978
May 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Should Value-Added Models Control for Student Absences?*

Whether or not value-added models should control for contemporaneous student absences is theoretically ambiguous, as such absences are only partly outside of teachers' control. Teachers often feel strongly that value-added models should account for student attendance, and many districts' value-added models condition on lagged student absences as a result. Using matched teacher-student administrative data from a state-wide longitudinal data system, this note investigates the practical importance of this modeling decision for value-added measures of teacher effectiveness (VAMs). This is done by comparing VAM-based rankings of teacher effectiveness generated by value-added models that either control for current absences, control for lagged absences, or exclude student absences altogether. Regardless of how between-school differences are accounted for, VAM-based rankings of teacher effectiveness are insensitive to how, and whether, student absences enter the value-added model's conditioning set. Spearman Rank Correlations are always larger than 0.99 for both math and reading VAMs, suggesting that whether or not value-added models control for annual student absences is a relatively unimportant modeling decision, at least in the context of self-contained primary school classroom teachers. These results are consistent with recent research suggesting that simply conditioning on lagged achievement yields approximately unbiased VAMs. Moreover, these findings suggest that controlling for student absences in teacher evaluation systems' value-added models is a relatively inexpensive way to increase teacher buy-in.

JEL Classification: I2

Keywords: teacher effectiveness, value added models, student absences,
teacher evaluation

Corresponding author:

Seth Gershenson
School of Public Affairs
American University
4400 Massachusetts Avenue, NW
Washington DC, 20016-8070
USA
E-mail: gershens@american.edu

* The author is grateful for financial support from the Spencer Foundation. Opinions reflect those of the author and do not necessarily reflect those of the granting agency. Cory Koedel provided helpful comments on an earlier draft. Andrew Brannegan and Michael Hayes provided helpful research assistance. Any remaining errors are my own.

1. Introduction

Teachers are one of the most important school-provided educational inputs (Hanushek & Rivkin 2010; Harris 2011). Providing effective teachers to all students, particularly to students from traditionally disadvantaged backgrounds, may help to close achievement gaps and improve the quality of public education more generally (Harris 2011). However, identifying effective teachers is difficult in practice (Harris 2011). Value-added measures (VAMs) of teacher effectiveness are gaining popularity, but not without some controversy (e.g., Chetty, Friedman, & Rockoff 2014; Harris 2011; Sass, Semykina, & Harris 2014). Intuitively, VAMs attempt to identify individual teachers' contributions to student achievement gains by conditioning on previous achievement (i.e., lagged test scores) in the hope that previous achievement acts as a sufficient statistic, or proxy, for the unobserved history of familial, educational, and individual inputs received to that point by students (Sass et al., 2014). Indeed, whether value-added models adequately control for unobserved historical inputs is one of the fundamental concerns regarding the validity of VAMs (Chetty et al. 2014; Sass et al. 2014).

It is reassuring that recent research on VAMs generally finds that simply controlling for lagged achievement yields reasonably good estimates of teacher effects (e.g., Chetty et al. 2014; Guarino, Reckase, & Wooldridge 2015; Kane & Staiger 2008). This result is consistent with the finding that most sorting of students to classrooms is based on lagged achievement (Chetty et al. 2014; Dieterle et al. 2015). However, teacher and student sorting is only one potential source of bias. A less studied potential source of bias is unobserved contemporaneous student effort, which is likely to be at least partially outside teachers' control. Indeed, in a survey of Baltimore City Public School (BCPS) teachers, Balch and Koedel (2014) find that one of teachers' largest

concerns regarding VAMs is whether they will hold teachers accountable for the performance of students who are frequently absent.

Attendance is one measure of student motivation that is typically available in administrative datasets (Isenberg & Walsh 2014), so there is no problem of missing data. Rather, there is a conceptual question as to whether and how value-added models should control for student absences when interest is in unbiased estimates of teacher effectiveness (Harris 2011; Noell et al. 2008). On the one hand, at least some student absences are completely outside teachers' control (e.g., absences due to household shocks, illness, or inclement weather) and such absences should unambiguously be controlled for (Harris 2011), as absences are known to reduce academic achievement (e.g., Aucejo & Romano 2014; Goodman 2014).¹ On the other hand, recent research suggests that teachers affect student absences (Author 2016; Ladd & Sorensen 2014), and absences caused by teachers are endogeneous "bad controls" that should be omitted from the econometric model (Angrist & Pischke 2009, p. 64).

In practice, of course, the share of a given student's absences attributable to his or her teacher is unknown, leaving analysts with three imperfect options. First, as is done in the Washington DC (Isenberg & Walsh, 2014) and BCPS (Balch & Koedel, 2014) school districts, the value-added model could condition on lagged (previous year) student absences. This is attractive because it avoids the endogeneity concerns regarding contemporaneous absences, though is limited to the extent that past absences do not perfectly predict current absences. The concern is that changes over time in individual student's absence rates may be driven by

¹ For example, recent studies of the harmful effects of student absences on academic achievement use flu outbreaks and inclement weather to instrument for student absences and have powerful first stages (Aucejo & Romano 2014; Goodman 2014).

unobserved changes in health or home environments that also affect achievement.² For example, Author (2016) shows that about 32% of variation in student absences is within-student variation, and that the within-student SD in annual absences is 1.75 absences. This suggests that lagged absences are a good, but imperfect, proxy for current absences. Teachers in BCPS recognized this, and still asked for current-year absences to be included in the value-added model. Doing so is the second option available to practitioners. However, as noted above, contemporaneous student absences are endogenous and including them in the model will penalize teachers to promote student attendance. Finally, student absences could be completely excluded from the econometric model, at the risk of creating omitted variables bias.

Intuitively, each of the three options described above is imperfect in some way. Given that teacher buy-in is critical to the success of teacher evaluation systems and many teachers feel strongly about controlling for student absences (Balch & Koedel 2014), adapting value-added models to include admittedly endogeneous contemporaneous absences might be a trade-off worth making if doing so does not substantively affect VAM-based rankings of teacher effectiveness. This note investigates the practical implications of conditioning on either current or lagged student absences by comparing VAM-based rankings of teacher effectiveness generated by three value-added model specifications: those that do not condition on student absences, those that control for lagged absences, and those that control for current absences.

I compare rankings rather than point estimates for two reasons: rankings are arguably more policy relevant and VAMs can yield valid rankings of teacher effectiveness even when the point estimates themselves are biased (Guarino, Reckase, Stacy, & Wooldridge 2015).

Comparisons are made for a variety of model specifications using statewide administrative data.

² This is why recent analyses of absences' effects on achievement rely on instrumental variable strategies (e.g., Aucejo & Romano 2014; Goodman 2014).

In all cases, VAM-based rankings of teacher effectiveness are quite robust to conditioning on student absences, suggesting that this is a relatively unimportant modeling decision when, as is often the case in teacher evaluation systems, rankings of teacher effectiveness are of interest.

2. Methods and Data

The data used to generate the VAMs analyzed in this note are thoroughly described in Author (2016). The longitudinal administrative dataset encompasses the population of third through fifth graders in North Carolina’s public school system between the 2005-06 and 2009-10 academic years. Importantly, these student-level data contain end-of-grade math and reading scores, annual absences, and student-teacher links. The analytic sample consists of fourth and fifth graders between 2007 and 2010 because the third-grade and 2006 data are used as lags in the value-added models. The data are maintained and provided by the North Carolina Education Research Data Center (NCERDC).³

The following discussion of VAM estimation draws heavily from Author (2016), to which the interested reader is referred for further details. Teacher effects, or VAMs, are the estimated teacher fixed effects (FE) in the following value added model

$$y_{ijgst} = \alpha y_{i,t-1} + \beta \mathbf{x}_{it} + \gamma \mathbf{c}_{-i,jgst} + \theta_j + \pi_g + \omega_{st} + u_{ijgst}, \quad (1)$$

where $i, j, g, s,$ and t index students, teachers, grades, schools, and years, respectively; y is standardized student test scores; \mathbf{x} is a vector of observed student demographic and socioeconomic characteristics that sometimes includes either current or lagged annual absences; \mathbf{c} is a vector of classroom characteristics; $\theta, \pi,$ and $\omega,$ are teacher, grade, and school-by-year FE, respectively; and u is an idiosyncratic error term. Because within- or between-school

³ See http://www.childandfamilypolicy.duke.edu/project_detail.php?id=35.

comparisons of teachers are sometimes of policy interest, versions of equation (1) are also estimated that either remove ω from the model or replace it with a school FE (e.g., Goldhaber & Theobald 2012). However, equation (1) is the preferred specification because the school-by-year FE control for variation across both schools and time in the length of academic calendars, weather and health shocks, and policies that either directly influence student attendance or the way that student absences are administratively reported (Author 2016).

For both math and reading achievement, equation (1) will be estimated by OLS three times, for three possible specifications of the vector x : excluding absences, including lagged absences, and including current absences. The resulting VAM-based rankings are then compared in three ways. First, comparisons are made via simple Spearman rank correlations. Second, I compute the percentage of teachers who are above average in both rankings, and similarly for various quantiles of interest sometimes used by teacher evaluation regimes. Finally, because correlations and simple descriptives might mask large swings in rankings (Goldhaber & Theobald 2012), I report transition matrixes that provide more nuanced views of the sensitivity of teacher rankings to controlling for student absences.

3. Results

Table 1 summarizes the sensitivity of VAM-based rankings of teacher effectiveness to conditioning on annual student absences. Rankings of the baseline VAMs generated by equation (1) are quite stable for both math and reading, regardless of whether the value-added model controls for current or lagged absences, and school-by-year FE, school FE, or no school FE. The Spearman Rank Correlations are all close to one and teachers in the top (bottom) quartile or decile in the with-absence ranking are almost always in the top (bottom) quartile or decile in the

no-absence ranking. For example, 96 to 97% of teachers who are in the top decile of the math ranking generated by value-added models that do not control for student absences are similarly in the top decile of the math ranking generated by value-added models that do control for contemporaneous student absences in each of the different school-FE specifications. This high stability rate is the same for the bottom decile of the math distribution, and is even higher for models that instead condition on lagged student absences. A similar pattern is observed in the reading results, which overall are a bit more stable than the math results.

Together, the results presented in table 1 suggest that across a variety of specifications, VAM-based rankings are robust to whether or not value-added models control for student absences. However, as noted in section 2, the descriptive statistics presented in table 1 potentially mask practically important changes in rankings (Goldhaber & Theobald 2012). For this reason quintile transition matrixes are reported in table 2, where each matrix compares movement between rankings that do and do not adjust for annual student absences. The left column reports transition matrixes for math rankings and the right column does so for reading.

The first pair of matrixes compares current-absence to no-absence VAMs. Consistent with the summary statistics reported in table 1, the reading rankings are slightly more stable, though neither subject contains large swings. Each element on the diagonal is larger than 90% and nearly all inter-quintile movement is to an adjacent quintile. The upper-left and bottom-right elements of these matrixes are even higher, between 96 and 98%, indicating that the classification of neither the least nor most effective teachers changes for the vast majority of such teachers as a result of controlling for current student absences.

The remaining pairs of transition matrixes shown in table 2 provide similar results that are again consistent with those discussed in table 1. The lagged-absence VAMs appear to be ever

so slightly more stable than current-absence VAMs, there are no large swings in rankings, and 98 to 99% of teachers who fall in the bottom (top) quintile of one VAM-based ranking remain in the same quintile regardless of how, or whether, the model conditions on student absences.

4. Conclusion

This note fills a gap in the value-added literature by empirically examining the practical importance of controlling for student absences in value-added models when interest is creating rankings of teacher effectiveness. Administrative data from the large, diverse state of North Carolina conclusively shows that in several specifications, and for both math and reading achievement, controlling for student absences does not appreciably change VAM-based rankings of teacher effectiveness. Generally, these results suggest that this is a relatively unimportant modeling decision, at least in the primary school context in North Carolina. Administrators, practitioners, and policymakers might leverage this finding by working and communicating with teachers to ensure that teachers' concerns regarding the ways in which they are held accountable for the achievement of frequently absence students. Future work might conduct similar analyses in other states and districts, and in the secondary school context, where the causes and consequences of student absences might be different from those in primary school settings. Similarly, future work might examine the sensitivity of VAMs to controlling for other measures and types of student effort, engagement, and motivation, or non-cognitive skills more broadly, that are influenced by teachers (e.g., Jackson 2013).

References

- Angrist, J., & Pischke, S. (2009). *Mostly harmless econometrics: An empiricists' companion*. Princeton, NJ: Princeton Univ. Press.
- Aucejo, E. M., & Romano, T. F. (2014). Assessing the effect of school days and absences on test score performance. CEP Discussion Paper No. 1302.
- Balch, R., & Koedel, C. (2014). Anticipating and incorporating stakeholder feedback when developing value-added models. *Education Policy Analysis Archives*, 22(97).
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9), 2593-2632.
- Dieterle, S. G., Guarino, C. M., Reckase, M. D., & Wooldridge, J. M. (2015). How do principals assign students to teachers? Finding evidence in administrative data and the implications for value-added. *Journal of Policy Analysis & Management*, 34(1), 32-58.
- Goldhaber, D., and R. Theobald. (2012). *Do different value-added models tell us the same things?* Available <http://www.carnegieknowledge.org/briefs/value-added/different-growth-models/>. Accessed 11 November 2014.
- Goodman, J. (2014). Flaking out: Student absences and snow days as disruptions of instructional time (No. w20221). National Bureau of Economic Research.
- Guarino, C.M., Reckase, M.D., & Wooldridge, J.M. (2015). Can value-added measures of teacher performance be trusted? *Education Finance and Policy*, 10(1), 117-156.
- Guarino, C. M., Reckase, F., Stacy, B. W., & Wooldridge, J. M. (2015). Evaluating specification tests in the context of value-added estimation. *Journal of Research on Educational Effectiveness*, 8(1), 35-59.
- Hanushek, E. A., & Rivkin, S.G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review* 100(2), 267-271.
- Harris, D. N. (2011). *Value-added measures in education*. Cambridge, MA: Harvard Education Press.
- Isenberg, E., & Walsh, E. (2014). Measuring teacher value added in DC, 2012–2013 school year. Washington, DC: Mathematica Policy Research.
- Jackson, C. K., (2013). Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina. NBER Working Paper No. w18624.

Kane, T. J., and D.O. Staiger. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. NBER Working Paper No. w14607.

Ladd, H. F., and Sorensen, L. C. (2014). Returns to teacher experience: Student achievement and motivation in middle school. Calder Working Paper No. 112.

Noell, G. H., Porter, B. A., Patt, R. M., & Dahir, A. (2008). *Value added assessment of teacher preparation in Louisiana: 2004-2005 to 2006-2007*. Unpublished manuscript, Louisiana State University.

Sass, T. R., Semykina, A., & Harris, D. N. (2014). Value-added models and the measurement of teacher productivity. *Economics of Education Review*, 38, 9-23.

Table 1: Sensitivity of VAM Rankings to Controlling for Student Absences

	Spearman Correlation Coefficient	Both above 90th percentile	Both above 75th percentile	Both above mean	Both above median	Both below 25th percentile	Both below 10th percentile
Math							
<i>Current versus None</i>							
Baseline	0.99***	9.6%	24.2%	49.2%	48.9%	24.1%	9.6%
School FE	0.99***	9.6%	24.1%	49.6%	49.0%	24.2%	9.6%
No FE	0.99***	9.7%	24.3%	51.1%	49.2%	24.4%	9.7%
<i>Lagged versus Current</i>							
Baseline	0.99***	9.8%	24.5%	49.6%	49.3%	24.4%	9.7%
School FE	0.99***	9.7%	24.3%	49.8%	49.2%	24.4%	9.7%
No FE	0.99***	9.8%	24.5%	51.3%	49.5%	24.6%	9.8%
<i>Lagged versus None</i>							
Baseline	0.99***	9.7%	24.4%	49.7%	49.2%	24.4%	9.7%
School FE	0.99***	9.8%	24.5%	50.1%	49.4%	24.6%	9.8%
No FE	0.99***	9.8%	24.6%	51.5%	49.6%	24.7%	9.9%
Reading							
<i>Current versus None</i>							
Baseline	0.99***	9.8%	24.5%	47.8%	49.4%	24.5%	9.8%
School FE	0.99***	9.7%	24.6%	49.6%	49.5%	24.5%	9.8%
No FE	0.99***	9.8%	24.5%	51.9%	49.5%	24.6%	9.8%
<i>Lagged versus Current</i>							
Baseline	0.99***	9.9%	24.7%	48.0%	49.6%	24.7%	9.9%
School FE	0.99***	9.8%	24.7%	49.7%	49.6%	24.6%	9.9%
No FE	0.99***	9.8%	24.7%	52.1%	49.7%	24.7%	9.8%
<i>Lagged versus None</i>							
Baseline	0.99***	9.9%	24.7%	48.1%	49.6%	24.7%	9.9%
School FE	0.99***	9.9%	24.8%	49.9%	49.7%	24.8%	9.9%
No FE	0.99***	9.9%	24.8%	52.3%	49.7%	24.8%	9.9%

Notes: The analytic sample contains 13,391 unique teachers. Comparisons are made between versions of equation (1) that either exclude student absences, control for lagged student absences, or control for current student absences. The baseline model conditions on school-by-year fixed effects (FE). *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Sensitivity of VAM Rankings to Controlling for Student Absences – Transition Matrixes

Current	Math					Reading				
	None									
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Quintile 1	96.4%	3.6%	0.0%	0.0%	0.0%	98.2%	1.8%	0.0%	0.0%	0.0%
Quintile 2	3.5%	91.9%	4.6%	0.0%	0.0%	1.8%	95.3%	2.9%	0.0%	0.0%
Quintile 3	0.04%	4.5%	90.4%	5.0%	0.0%	0.0%	2.9%	94.2%	2.9%	0.0%
Quintile 4	0.0%	0.0%	5.0%	91.9%	3.1%	0.0%	0.0%	2.9%	95.2%	1.9%
Quintile 5	0.0%	0.0%	0.0%	3.1%	96.9%	0.0%	0.0%	0.0%	1.9%	98.1%
Lagged	Current					Current				
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Quintile 1	97.7%	2.3%	0.0%	0.0%	0.0%	98.8%	1.2%	0.0%	0.0%	0.0%
Quintile 2	2.2%	95.2%	2.7%	0.0%	0.0%	1.2%	97.2%	1.7%	0.0%	0.0%
Quintile 3	0.0%	2.7%	94.3%	3.1%	0.0%	0.0%	1.7%	96.4%	1.9%	0.0%
Quintile 4	0.0%	0.0%	3.0%	94.8%	2.1%	0.0%	0.0%	1.9%	97.1%	1.0%
Quintile 5	0.0%	0.0%	0.0%	2.1%	97.9%	0.0%	0.0%	0.0%	1.1%	98.9%
Lagged	None					None				
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Quintile 1	97.3%	2.7%	0.0%	0.0%	0.0%	98.7%	1.3%	0.0%	0.0%	0.0%
Quintile 2	2.6%	94.1%	3.3%	0.0%	0.0%	2.2%	95.1%	2.7%	0.0%	0.0%
Quintile 3	0.0%	3.3%	92.9%	3.8%	0.0%	0.0%	2.7%	94.3%	3.1%	0.0%
Quintile 4	0.0%	0.0%	3.8%	93.7%	2.6%	0.0%	0.0%	3.0%	94.8%	2.1%
Quintile 5	0.0%	0.0%	0.0%	2.5%	97.5%	0.0%	0.0%	0.0%	2.1%	97.9%

Notes: The North Carolina data contain 13,391 unique teachers. These models condition on school-by-year fixed effects.