TRANSFORMING TENURE:
Using Value-Added Modeling to Identify Ineffective Teachers

Marcus A. Winters
Senior Fellow, Manhattan Institute

Public school teachers in the United States are famously difficult to dismiss. The reason is simple: after three years on the job, most receive tenure—after a brief and subjective evaluation process (typically, a classroom visit or two by an administrator or another teacher) in which few receive negative ratings. Once tenured, teachers are armored against efforts to remove them, and most do not face any serious reevaluation to ensure that their skills stay up to standard. With this traditional approach, tenured teachers sometimes lose their positions for insubordination, criminal conduct, gross neglect, or other reasons—but almost never for simply being bad at the job.

This state of affairs protects teachers (both good and bad) quite well but is clearly harmful to students. The effects of a poor teacher, research has shown, haunt pupils for years afterward. Being assigned to such a teacher reduces the amount that a student learns in school and is associated with lower earnings in adulthood (in part because having an inadequate teacher makes a child more likely to have an early pregnancy and less likely to go to college). An education system that protects bad teachers does a grave disservice to the children in its care.

In recent years, some school districts have experimented with changes in tenure rules. They seek the power to remove ineffective teachers and, in some jurisdictions, to reevaluate teachers throughout their careers.

A keystone of this reform movement is the replacement of subjective evaluation with quantifiable measures of each teacher's effectiveness. The quantitative method is known as value-added modeling (VAM), a statistical analysis of student scores that seeks to identify how much an individual teacher contributes to a pupil's progress over the years. The use of VAM in teacher evaluations is growing, but the method remains extremely controversial. Critics often claim that it does not and cannot measure actual teacher quality.

This paper addresses that claim. Part I analyzes data from Florida public schools to show that a VAM score in a teacher's third year is a good predictor of that teacher's success in his or her fifth year. Having established that VAM is a useful predictive tool, Part II of the paper addresses the most effective ways that VAM can be used in tenure reform.

VAM is not a perfect measure of teacher quality because, like any statistical test, it is subject to random measurement errors. So it should not be regarded as the "magic bullet" solution to the problem of evaluating teacher performance. However, the method is reliable enough to be part of a sensible policy of tenure reform—one that replaces "automatic" tenure with rigorous evaluation of new candidates and periodic reexamination of those who have already received tenure.

About the Author

Marcus A. Winters is a senior fellow at the Manhattan Institute and an assistant professor at the University of Colorado Colorado Springs. He conducts research and writes extensively on education policy, including topics such as school choice, high school graduation rates, accountability, and special education. Winters has performed several studies on a variety of education policy issues including high-stakes testing, performance-pay for teachers, and the effects of vouchers on the public school system. His research has been published in the journals *Educational Evaluation and Policy Analysis*, *Education Finance and Policy*, *Economics of Education Review*, *Teachers College Record*, and *Education Next*. His op-ed articles have appeared in numerous newspapers and magazines, including *The Wall Street Journal*, *The Washington Post*, *USA Today*, the *New York Post*, the New York *Daily News*, the *Weekly Standard*, and *National Affairs*. He is often quoted in the media on education issues. Winters received a B.A. in political science from Ohio University in 2002, and a Ph.D. in economics from the University of Arkansas in 2008.

CONTENTS

Transforming Tenure: Using Value-Added Modeling to Identify Ineffective Teachers

Marcus A. Winters

INTRODUCTION

Tenure and the Problem of Teacher Quality

Bad teachers substantially harm a child's prospects. Studies have found that an ineffective teacher can cost pupils as much as a grade level's worth of learning during a single school year.[1] Further, bad teachers—those who do not make any measurable contribution to their students' advancement—make students more likely to have an early pregnancy, reduce the chances that they will go to college, and have a negative impact years later on their pupils' earnings as adults.[2] A wide body of research has shown that even as teacher quality is a school's most important driver of achievement, teacher quality varies a great deal from classroom to classroom in public schools.[3]

Since 2009, a few school districts around the nation have been experimenting with changes to business as usual, seeking ways to improve the quality of their teachers. Though these districts remain a small minority, the reform effort has gathered steam, especially in the past year. One of its most controversial suggestions is the redefinition—or even elimination—of tenure for public school teachers.

For years, teachers' unions and their supporters have described tenure as a necessary bulwark against arbitrary or discriminatory termination, which was a common practice before the advent of modern employment law and labor standards. But the current tenure system protects bad teachers as well as good ones. (Very few tenured teachers are ever forced to leave the classroom.) We know that teachers vary in quality and that removing less competent teachers has the potential to improve students' education. Therefore, we can be sure that pupils are ill-served by a system that ensures that bad teachers cannot be fired. As its defenders like to point out, tenure ensures only that teachers receive due process before they are terminated. However, in most

school systems, the required due process is so burdensome—and has so small a chance of success—that in practice, poor performance is rarely a firing offense. To be rid of a teacher for poor performance, in most public school systems, an administrator must carefully document several proofs of incompetence over a sustained period of time—in the form of botched lesson plans, improper classroom development, and observed poor practice. These "proof points" are inherently subjective, and each is contestable in the hearing process. Meanwhile, measurements of actual outcomes—how much students have learned in the teacher's classroom—are rarely considered. This is why poor classroom performance is so rarely cited as a reason for dismissal. For instance, competence was mentioned in only eight of the 45 cases in which tenured teachers were terminated in New York City in 2008 and 2009. And six of those eight included other charges such as insubordination or misconduct.[4]

One might argue that worthy teachers with good records have earned some protection against the effects of a personal crisis or a rough year in the classroom. Tenure, though, is not reserved for proven educators. On the contrary, public school teachers are offered lifetime tenure very early in their careers—usually after three years—and the offer seldom has much to do with their performance. As of 2011, according to a review of tenure laws by the National Council on Teacher Quality (2011), only eight states require that performance of a teacher's students be central to deciding whether to award a teacher tenure. That actually represents considerable progress, since in 2009 the NCTQ found that not a single state awarded tenure primarily based on effectiveness. Moreover, in most American public schools, that early-career tenure decision is often the only systematic examination of a teacher's worth. Tenured teachers are rarely reexamined to ensure that their skills are maintained.

Why are measurements of effectiveness given so little weight in tenure processes? The simple answer is that, until recently, such measures did not exist. Tenure rules were written when performance was evaluated entirely on the basis of a classroom visit or two by an experienced observer. School systems simply lacked any objective measure of the teacher's contribution to student learning. Today, better measuring tools exist, but the rules remain as written. When tenure is decided, nearly all the teachers in a typical school system receive a satisfactory or higher rating.[5]

School systems need a better approach to tenure. Job protection, if it is to be offered at all, should be restricted to the best teachers. And policies should permit reevaluations, lest once-worthy teachers be protected long after their performance has faltered. Most important, tenure should be related to meaningful and objective measurements of teaching effectiveness.

On this last point, modern statistical tools present a promising avenue for reform. These measures, used in tandem with traditional subjective measures of teacher quality, could help administrators make better-informed decisions about which teachers should receive tenure and which should be denied it. Statistical evaluations can also be used to identify experienced teachers who are performing poorly, with an objectivity that reduces the risk of a teacher being persecuted by an administrator.

To those dissatisfied with the status quo, one technique in particular seems to offer a good basis for reform, and it has been implemented in many recent attempts to change tenure rules in order to improve teacher quality. It is the method known as value-added modeling (VAM). VAM uses a complex statistical procedure to determine each teacher's independent contribution to improvement in his or her students' test scores.

Many school systems across the nation have recently used, or are currently considering using, VAM assessments when making employment decisions. For instance, under new laws passed in Colorado in 2010, Tennessee in 2011, and just recently in New Jersey, teachers in those states will lose their tenure if they receive below-satisfactory performance ratings in two consecutive years. Those ratings are based, in part, on VAM.

Some worry that because VAM is an imperfect measure of classroom effectiveness, it will incorrectly deny tenure protections to some effective teachers—or

even cause good teachers to lose their jobs. If so, VAM's negative impact might cancel out its benefits and result in no net improvement in the quality of a school district's teaching staff. After all, research shows that VAM is an imprecise measure of a teacher's true performance.[6]

For this report, I test the premise that a teacher's VAM score can help predict his or her future performance. I use data from Florida to replicate recent analyses by two scholars, Dan Goldhaber and Michael Hansen (2010), who used data from North Carolina. Consistent with their research, my results show that pre-tenure VAM scores are significantly related to student test-score performance in the teacher's classroom in later years. These results indicate that VAM often contains meaningful information about a teacher's future effectiveness, which can usefully inform employment decisions.

Obviously, the potential effects of any VAM-based tenure-reform policy would depend upon its design. Accordingly, the second part of this report looks at the number and type of teachers who would have been removed from the classroom ("deselected") rather than tenured under different sorts of VAM-based policies, had those policies been in place in Florida when the data were collected. These comparisons show that the effects of such policies on teacher quality will depend on the standard that a teacher must meet to receive a satisfactory rating and on whether a teacher can lose tenure after it has been granted. These design issues, though important, should not obscure the fundamental point: VAM-based tenure policies hold considerable promise for removing consistently ineffective teachers and thus improving teacher quality throughout the public school system.

Before considering the method and results from this report, it is worth emphasizing that though the analysis here focuses only on the influence of VAM on teacher tenure decisions, real-world policies will quite sensibly use VAM as only one measure of effectiveness when rating teachers. Therefore, this report has put VAM-based tenure policies to a hard test: by evaluating the effect of using VAM alone to identify and remove ineffective teachers, it has placed more reliance on VAM than a real district would. That the VAM approach passes this test is a striking indication of its usefulness.

It is important to recall that this analysis was created to test the ability of VAM to identify low-performing teachers under the structure of the current system. That is, the analysis assumes that teachers and school systems will not respond to the new rules by changing their other behaviors. This is unlikely to be the case in any real-world application of tenure reform. Instead, teachers could reasonably be expected to respond to a reformed tenure system in several ways. The reformed system might, for example, attract a different sort of candidate. Further, teachers could respond to the new possibilities—not receiving tenure or being removed from the classroom—in ways that are good for students (by increasing their effort level), or that have unpredictable effects (changing their teaching style), or that could have negative effects (emphasizing only testable material in the classroom).

Additional theoretical and empirical research is needed to map the real-world effects of incorporating VAM-based measures of teacher quality into employment decisions. However, understanding the ability of VAM to predict future performance and the type of teacher identified as ineffective by a VAM-based system is an essential first step.

Balancing the Needs of Teachers and Pupils

Though VAM is a powerful technique, it is undoubtedly an imperfect measure of a teacher's effectiveness. VAM is limited partly because it considers student performance only as measured by standardized tests, which are themselves imperfect measures of student achievement and account for only part of what school systems ask teachers to do. But even as a measure of the teacher's contribution to student test scores, VAM has potentially serious limitations.

Critics of VAM analysis rightly point out that, as a statistical tool, VAM must contend with measurement error—the inevitable fact that measurements of the same thing, taken at different times, will vary, and some of this variation will be essentially random.

VAM-based measures of teacher performance can be quite imprecise. When VAM is used to inform tenure decisions, it is likely that some average and even above-average teachers could be removed from the classroom because of a low VAM score caused by random variation in measurement over the years, rather than their own failures. The influence of measurement error can be mitigated by statistical adjustments and by incorporating multiple years of student performance when evaluating any particular teacher. But measurement error cannot be eliminated.

From the perspective of teachers (and their unions), the collateral damage of even a single teacher losing tenure from an inaccurately low VAM score is unacceptable. However, the issue is not as cut-and-dried from the perspective of the student. A tenure-reform policy based on VAM will be an improvement for students if it removes enough low-performing teachers to improve overall teacher quality in a school district. If student achievement is our most pressing concern, we need to consider the possible consequences of VAM-based policies on whole districts, even as we acknowledge the potential for error in individual cases.

No evaluation system creates a perfect measure of an employee's productivity. VAM, then, should not be judged against a nonexistent ideal but rather evaluated for its potential to improve on the current system's ability to predict future performance. In the analyses that follow, this was my goal: to assess whether a tenure policy based on VAM would tend to improve a school district's overall teacher quality.

PART I: VAM IS A RELIABLE PREDICTOR OF FUTURE PERFORMANCE

Following Goldhaber and Hansen's work from North Carolina, my primary analysis uses a simple value-added model to estimate a teacher's contribution to student test scores during the first two years in the classroom. I then evaluate the relationship between this measure and the achievement of students in the teacher's classroom during his or her fifth year. If the previous VAM measure of teacher quality is a significant predictor of the teacher's later achievement, we can conclude that

VAM provides reliable information about a teacher's future performance.

The analyses use detailed data about Florida students' performance on the state's annual high-stakes math and reading exams, the Florida Comprehensive Assessment Test (FCAT) in the spring semesters from 2002 through 2009.[7] Though individuals are not identified by name, the data set permits the analyst to follow the performance of each student over time. It also includes identifying variables for each teacher and a variable used to match students to teachers in classrooms.

My analyses only include students in the fourth and fifth grades. In later grades, students change teachers for each subject, making the assessment of teacher impact far more difficult. Further, testing in Florida begins in the third grade, and the analysis requires a baseline achievement score for the year before the study period. Therefore, grades before fourth are not available for this method.

I used student reading scores to create a simple value-added model by grade and year (a later check showed that results would be similar had I used math scores). The model accounted for the impact on test scores of such observed student characteristics as race/ethnicity, gender, and socioeconomic status (as measured by whether the children were eligible for free or reduced-priced lunches). After controlling for these and other variables, I was able to arrive at the estimated contribution of individual teachers to their students' test scores.[8]

With a measure of teacher impact in place for each student, I could then look at the data at the teacher level to develop a rolling measure of each teacher's quality over the years. As we have mentioned, most school systems offer tenure after three years in the classroom. Therefore, I calculated each teacher's average VAM score during his or her first three years in the classroom.

Finally, I took the measure of each teacher's average value-added score during his or her first three years back to the student-year data set. I used the VAM from those first three years to help predict each

teacher's students' achievement in the teacher's fifth year (the 2007–08 school year). What I was looking for was a significant and meaningful relationship between pre-tenure VAM score and the performance of students in the teacher's future classroom years later.

Relationship Between Pre-Tenure VAM and Later Student Performance

The results of the analysis are reported in Table 1. The first column reports the results from a regression analysis (a statistical method for showing the relationship among several variables) in which I mapped the relationship between student achievement and a teacher's having a master's degree. (The master's is often used as a proxy for skill and commitment in current evaluation systems.) Consistent with previous research, I find no relationship between a teacher having a master's degree and student outcomes.

The second column reports the results of a regression analyzing the relationship between the teacher's average VAM score during the first three years in the classroom and the performance of that teacher's students during his or her fifth year in the classroom. The result shows a statistically significant and substantial relationship between the teacher's pre-tenure average VAM score and achievement in that teacher's classroom several years later. The third column shows

Table I: Results

	Model 1	Model 2	Model 3
Masters Degree	0.136		0.115
	[0.103]		[0.0906]
Average Pre-Tenure VAM		0.628***	0.626***
		[0.0569]	[0.0565]

*** Significant at 1% level

Each column in the table represents the results of an independent regression. Three models are presented in order to illustrate the effect of controlling for possession of a master's degree or pre-tenure value added on the relationships. Dependent variable in all models is the student's reading test score in the spring of 2008-09. Model 1 considers the relationship between student achievement and whether a teacher has a Master's degree; Model 2 considers the relationship between student achievement and the teacher's average pre-tenure value-added score; and Model 3 includes both Master's degree and average pre-tenure score. Independent variables of interest listed by row. Models also control for observed student characteristics and prior student test score. Standard errors clustered by school reported in the brackets.

that a control for whether the teacher has a master's degree has no meaningful influence on the finding.

Results reported in Table 1 demonstrate that the value-added assessment of the teacher's effectiveness prior to the tenure decision is a significant predictor of the teacher's later effectiveness. Thus, VAM measures early in a teacher's career appear to be good predictors of how well a teacher will perform in the future. As mentioned, this result is consistent with the previous findings of Goldhaber and Hansen, who used data from North Carolina; it is important to note that data from another state's school system, based on data from a different standardized test, show the same relationship between early-career VAM scores and later student success.

PART II: COMPARING THE EFFECTS OF DIFFERENT VAM-BASED POLICIES

Accepting that VAM can help predict future success for teachers, I turn to the next practical question for school districts: How should VAM be incorporated into tenure policy?

Policymakers must first consider the level of performance that a teacher has to meet to avoid an ineffective rating. This bar must not be set too low, or the VAM will have little impact on quality. For instance, a VAM-based policy that removes a large school district's single worst teacher might have a substantial effect for the few students who would have been assigned to that teacher's classroom but would have an infinitesimal effect on overall teacher quality throughout the school system.

A second issue to consider is whether a teacher who receives tenure under a reformed system would keep it going forward (as is currently the case) or whether teachers could be continually reviewed. If tenure continues to be decided in teachers' third year on the job and they experience no further significant reviews, the impact of any quality-improvement effort will be limited to teachers at the start of their careers. This means that the policy might affect too few teachers and do nothing about older teachers whose effectiveness is fading.

Finally, policymakers must consider how to use multiple years of VAM scores to assign tenure or identify teachers for removal. The measurement error inherent in VAM analysis, along with other administrative issues, should lead school systems to use multiyear measures when making employment decisions. Policymakers could respond to this need by comparing teacher performance using the average VAM score over a multiyear period or, as districts in Colorado and Tennessee have already done, by removing teachers after they receive consecutive poor ratings.

Table 2 reports the number of students in Florida who were attached to teachers who would have been fired according to different versions of a VAM-based policy: first, one that removes a teacher who has received a poor rating based on the previous three years' performance; second, a policy that removes teachers only after they have demonstrated below-standard performance during their first three years in the classroom; and third, a policy that removes teachers who perform below a particular standard during consecutive years.

The table shows that different versions of a tenure-reform policy would benefit different numbers of students. As would be expected, policies that simply raise the VAM score considered acceptable will affect a greater number of teachers, and thus students. Similarly, the most impactful policy is one that affects all teachers, regardless of whether they have previously been granted tenure.

The table also shows that the most conservative policy—that is, the policy that leads to the fewest teacher removals—removes teachers based on consecutive bad ratings rather than their average rating relative to other teachers during a multiyear period. That result occurs because under a system based on

consecutive poor ratings, teachers who earned a single low rating—perhaps because of random error—have the opportunity to "correct" the result by meeting the standard the next year. On the other hand, a policy that removes all teachers whose average score is below a particular percentile will always remove that percentage of teachers. By definition, a policy that removes teachers whose average VAM is below the fifth percentile of all average VAM scores during that period will remove 5 percent of the teachers, while a policy that removes teachers if they consecutively score below the fifth percentile will keep a teacher who scores in the third percentile during one year and the seventh percentile the next.

The effect of a tenure-reform policy on overall teacher quality in the school system depends both on the number and quality of teachers denied tenure under such a policy. Figures 1 through 9 compare the distribution of the 2008–09 VAM scores of teachers who would have been deselected at the end of the 2007–08 school year, according to these different systems, with those of teachers who would have avoided removal.

Though each figure represents a different policy, all show that teachers who would have been fired in 2008–09 were less effective than teachers who would have survived review. However, the figures illustrate that some teachers who were observed to be performing at or above the mean in 2008–09 would have been fired according to any version of tenure reform. The risk—of firing teachers whose later performance is above average—increases as the standard for failure is set higher. For example, a policy that removes teachers performing below the 25th percentile sets a higher standard than a policy that removes those scoring below the fifth percentile. But that policy is more likely to remove teachers whose later effectiveness would prove to be well above average.

Table 2: VAM-Based Teacher Firing			
Percentile	5th	10th	25th
Remove if Average 3 Year Below Xth Percentile	3,651	7,804	20,580
Only Fourth Year Teachers Subjected to Potential Deselection	2,539	4,999	12,635
Remove if Receive 2 Consecutive VAM Scores Below Xth Percentile	580	1,685	9,457

The figures also enable us to compare the later performance of teachers who would have been deselected according to different policy styles. As was done in Table 2, we consider the quality of teachers deselected according to a policy that: a) removes any teacher whose average VAM score over a three-year period was below the Xth percentile; b) removes only entering fourth-year teachers whose average VAM score over their first three years was below the Xth percentile among all teachers; or c) removes any teacher with a VAM score below the Xth percentile among all teachers for consecutive years.

The figures illustrate that the most conservative policy design—that is, the policy least likely to remove teachers who later perform well in the classroom—removes those teachers who score below the Xth percentile during consecutive years. As Table 2 illustrates, this is the policy design that removes the smallest number of teachers. On the other hand, a policy that removes any teacher whose average VAM score over a three-year period is below the Xth percentile will tend to remove more teachers who would later demonstrate themselves to be effective, though even this policy will tend to remove more ineffective teachers than effective ones.

CONCLUSION

Like previous research found in North Carolina, my analysis of Florida data found that pre-tenure VAM scores often provide information about a teacher's future quality. Thus, VAM analysis can help replace "automatic" tenure with employment decisions based on reliable evaluations. It can be part of tenure reform and thus can contribute to improving public education in the United States.

But which tenure-reform policies would make best use of this technique? I addressed this question by pinpointing the teachers in the Florida data who would have been removed from the classroom according to several different types of policies and performance standards. I found that any VAM-based policy would have removed teachers who, on average, performed worse than their peers later in their careers.

However, different versions of VAM-based policies proved to have different consequences. Specifically, certain versions increased the risk that effective teachers (as measured by VAM) would be removed. For example, a policy could target teachers for removal if they have two or more periods of consecutive poor performance. Alternately, the policy could simply score teachers on an average of their performance ratings for a given number of years. I found that the latter policy was more likely than the former to result in the removal of effective teachers (teachers who, despite a "bad patch" in the records, would prove to be effective later). Another way to increase this risk of "false positives," I found, was to set the performance bar high. Such policies, applied to the Florida data, would also have resulted in the removal of teachers who would later demonstrate effective performance.

These results tell tenure reformers that they should consider the number and type of teachers likely to be denied tenure or removed from the classroom under their proposed policies. This will help them design policies that balance the interests of students in need of great teachers and the legitimate interests of teachers concerned that they will be inappropriately removed from the classroom because of a randomly low VAM score.

The need for well-designed policies should not obscure the finding that public schools can indeed use VAM to help identify teachers for tenure or removal. Instead, these results underscore the importance of blending VAM with sound policies. This report does not argue that VAM should be used in isolation to evaluate teachers for tenure or to make any other employment decisions. VAM, as we have seen, is subject to random measurement errors, and so must be combined with other methods of teacher evaluation.

The lesson of this report and of other research is that VAM can be a useful piece of a comprehensive evaluation system. Claims that it is unreliable should be rejected. VAM, when combined with other evaluation methods and well-designed policies, can and should be part of a reformed system that improves teacher quality and thus gives America's public school pupils a better start in life.

ENDNOTES

[1] E.g., Hanushek (1992) finds that students assigned to a teacher whose students have results in the 75th percentile (i.e., whose scores are better than three-quarters of their fellow pupils') will test one year and a half ahead of where they started when the school year is over. Students with teachers in the 25th percentile, on the other hand, end up with scores that are only a half-year better than their starting point.

[2] Chetty, Friedman, and Rockoff (2011).

[3] See Hanushek and Rivkin (2010).

[4] E-mail correspondence with the Department of Education.

[5] Weisberg, Sexton, Mulhern, and Keeling (2009).

[6] See, e.g., McCaffrey, Sass, Lockwood, and Mihaly (2009).

[7] The analyses use a rich student-level panel data set acquired from the Florida K-20 data warehouse.

[8] Consistent with previous research, I adjust the teacher effects according to the empirical Bayes estimator.

REFERENCES

Chetty, R., J. N. Friedman, and J. E. Rockoff (2011). "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood." NBER Working Paper 17699.

Goldhaber, Dan, and Michael Hansen (2010). "Using Performance on the Job to Inform Teacher Tenure Decisions." *American Economic Review: Papers and Proceedings* 100: 250–55.

Hanushek, Eric A. (1992). "The Trade-Off Between Child Quantity and Quality." *Journal of Political Economy* 100, no. 1: 84–117.

———, and Steven G. Rivkin (2010). "Using Value-Added Measures of Teacher Quality." Urban Institute.

McCaffrey, Daniel F., Tim R. Sass, J. R. Lockwood, and Kata Mihaly (2009). "The Intertemporal Variability of Teacher Effect Estimates." *Education Finance and Policy* 4, no. 4: 572–606.

National Council on Teacher Quality (2011). 2011 State Teacher Policy Yearbook.

Weisberg, D., S. Sexton, J. Mulhern, and D. Keeling (2009). "The Widget Effect." New Teacher Project.
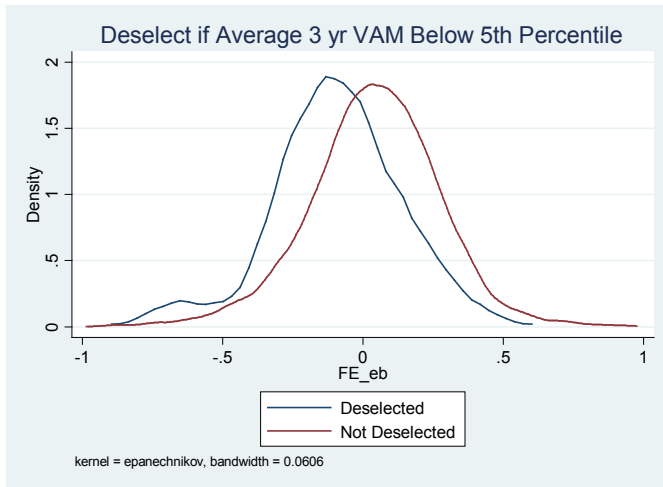
Figure I

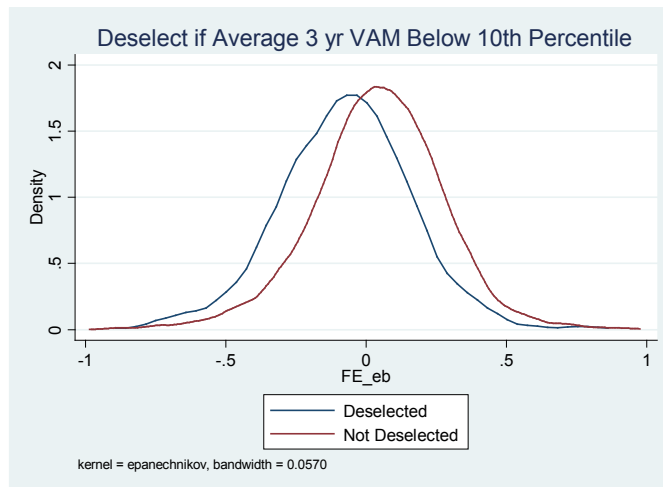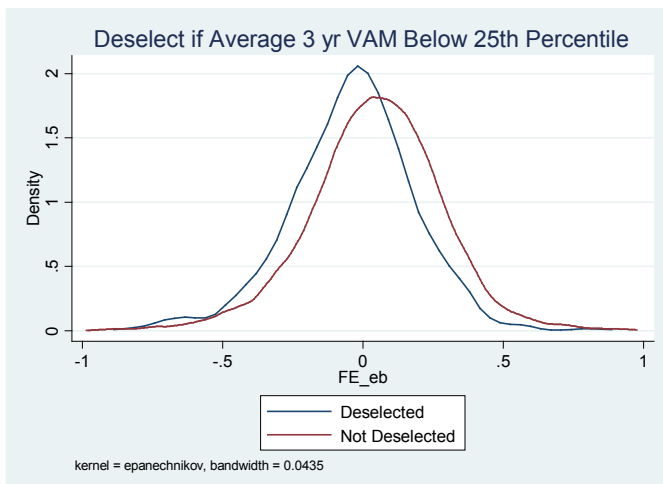Deselect if Average 3 yr VAM Below 5th Percentile

kernel = epanechnikov, bandwidth = 0.0606

Figure 2

Deselect if Average 3 yr VAM Below 10th Percentile

kernel = epanechnikov, bandwidth = 0.0570

Figure 3

Deselect if Average 3 yr VAM Below 25th Percentile

kernel = epanechnikov, bandwidth = 0.0435

Figure 4

Deselect if Consecutive VAM Below 5th Percentile

kernel = epanechnikov, bandwidth = 0.0645

Figure 5

Deselect if Consecutive VAM Below 5th Percentile

kernel = epanechnikov, bandwidth = 0.0713

Figure 6

Deselect if Consecutive VAM Below 5th Percentile

kernel = epanechnikov, bandwidth = 0.0523

Figure 7

Deselect if Average 3 yr VAM Below 5th Percentile - Only 4th Year

kernel = epanechnikov, bandwidth = 0.0775

Figure 8

Deselect if Consecutive VAM Below 10th Percentile - Only 4th Year

kernel = epanechnikov, bandwidth = 0.0609

Figure 9

Deselect if Consecutive VAM Below 25th Percentile - Only 4th Year

kernel = epanechnikov, bandwidth = 0.0480

The Manhattan Institute's **Center for State and Local Leadership (CSLL)** promotes promising new approaches to reform of state and local government. CSLL works on a broad range of issues, including public sector reform (specifically of pensions and health benefits), education reform, prisoner reentry, policing, public housing, infrastructure, immigration, and public service delivery. By spotlighting new ideas, and providing the research and proposals to inform creative new policies, the CSLL hopes to lay the groundwork for an environment in which commerce, employment, and a rich civic life can flourish.

The CSLL operates across the country, working in states such as California, Illinois, and Rhode Island, and cities such as Newark, New Jersey and Detroit, Michigan. The CSLL's tools include regular writing and research reports by affiliated Manhattan Institute scholars and senior fellows, along with public events and media appearances. The CSLL operates www.PublicSectorInc.org, a website devoted to analysis of the implications—financial and political—of the power wielded by public sector unions and allied elected officials. The CSLL also annually selects and showcases the Manhattan Institute's Urban Innovator Award, which recognizes a state or local leader whose combination of policy creativity and skill at implementation has led to groundbreaking improvements in public service, and the Manhattan Institute's Social Entrepreneurship Awards, which recognize those who identify social needs and take it upon themselves to address them privately.

www.manhattan-institute.org/csll