Putting School Surveys to the Test*

Joshua Angrist† Peter Hull‡ Russell Legate-Yang§

Parag A. Pathak¶ Christopher R. Walters‖

October 1, 2024

Abstract

School districts increasingly gauge school quality by student surveys of school climate and student engagement. We use data from New York City's middle and high schools to compare the long-run predictive validity of survey-based value-added models (VAMs) and more conventional VAMs based on student test scores. Our analysis leverages the New York school match, which includes an element of random assignment, to validate a wide range of value-added estimates. We then contrast predictions computed using survey- and test-based measures of school quality for school effects on consequential outcomes related to high school graduation and college enrollment. For high schools, both survey and test score VAMs predict school effects on graduation, but survey value-added is nearly uninformative for middle schools and for effects on advanced diplomas. Moreover, conditional on test score value-added, survey value-added fails to predict school effects on college enrollment and persistence. While surveys aren't worthless, our findings suggest test score information dominates survey information as a guide to school quality.

*Financial support from Gates Foundation is gratefully acknowledged. We also thank Eryn Heying, Talia Gerstle, and Niamh McLoughlin for invaluable administrative support. Fatima Djalalova provided expert research assistance. The research described here was carried out under data-use agreements between MIT and the New York City school district. The IRB of record was MIT COUHES, under protocol 1904808765.

†MIT and NBER. Email: angrist@mit.edu

‡Brown University and NBER. Email: peter_hull@brown.edu

§MIT. Email: rcly@mit.edu

¶MIT and NBER. Email: ppathak@mit.edu

‖University of California, Berkeley and NBER. Email: crwalters@econ.berkeley.edu

1 Introduction

For decades, school districts and states have relied on standardized tests to measure school performance—but this has started to change. The 2015 *Every Student Succeeds Act* requires states to include non-academic factors for purposes of federally-mandated school accountability reports (Jordan and Hamilton, 2020). At the same time, social scientists and educators have argued for the importance of non-cognitive development for students' later life outcomes (see, e.g., Dweck, 2006; Duckworth et al., 2007; Cunha and Heckman, 2007; Jackson et al., 2020; Deming, 2023). Responding to such currents, school districts increasingly disseminate and highlight survey-based measures of school climate and student engagement.

The emerging emphasis on survey-based measures has often and intentionally come at the expense of test-based accountability schemes. New York City is a leading example of this trend: in 2013, New York abandoned a largely test-based accountability regime established under outgoing Mayor Michael Bloomberg. In the Bloomberg era, many schools—especially those serving low income students—struggled to earn high grades on the city's "school report cards."[1] In 2014, the mayoral administration of Bill de Blasio introduced an accountability scheme that prioritizes survey data over standardized tests. Based in part on a redesigned school climate survey, the new measures are meant to provide a more complete picture of school performance (Merrill et al., 2018). In practice, the new ratings feature variables such as the proportion of students indicating that they "learn a lot from feedback on their work." Critics of the new approach argue that information of this sort is inherently vague and likely to be inconsistent (Wall, 2014).

This paper evaluates and contrasts the predictive validity of survey-based and test-based measures of school performance. Our analysis builds on Angrist et al. (2021) by leveraging the random assignment embedded in New York's middle and high school matches to validate causal school effects in a value added modeling (VAM) framework. We use this framework

[1]Bloomberg-era report cards were 85 percent test-score based (Hernández, 2013). Corradini (2024) examines the impact of school report cards and other performance measures on student application patterns.

to estimate and validate effects of individual school attendance on achievement as measured by test scores, a summary of survey-based performance indicators, and consequential longer-term outcomes related to high school graduation and college attendance. Our VAMs for test scores and longer-term outcomes correct the substantial selection bias in school average achievement levels. For surveys, average response levels accurately measure school causal effects. Given this equivalence, we then ask whether students who attend schools with high average survey responses (*survey levels*) and high test score value added (*test VAM*) are more likely to graduate high school and to attend and persist in college. We're especially interested in the relative merits of survey levels and test VAM as predictors of school effects on high school graduation and college attendance. We assess the value of these alternative metrics in a decision-theoretic model in which applicants aim to pick a school so as to maximize expected college attendance and persistence given all available information (survey levels, test score value-added, or both).

We apply this approach to New York, where the school survey program elicits information related to diverse aspects of the school environment. With over a million potential respondents, the New York survey is one of the most ambitious efforts of its kind. The New York school survey design draws on research identifying aspects of the school environment that are widely seen as related to school improvement (Merrill et al., 2018). The survey has intellectual roots in a similar effort based in the Chicago public schools (Bryk et al., 2010). As in related work by Jackson et al. (2020), our survey measures use student responses only. Much of the information on the New York City Department of Education school ratings website, presented to families in the form of "school quality snapshots," originates in these surveys. School quality snapshots present test-based measures as well, though the details are found on a separate and relatively dense "school performance dashboard" that most parents are less likely to encounter.[2]

[2]See, for example, the quality snapshot for Beacon HS and the corresponding performance dashboard. Response rates for the New York parents' survey are low while teacher survey results hard to interpret since teachers are not randomly assigned to schools.

Our findings broadly suggest that test score information dominates survey information as a guide to school quality. A high school that increases test scores by one standard deviations is found to increase on-time college enrollment by 18 percentage points, on average. The corresponding survey levels association is much smaller, at 4 percentage points; moreover, it shrinks to a statistically insignificant 2.6 percentage points once test VAM is controlled for. A similar pattern arises for persistence in college through two years. For middle schools, test VAM is similarly predictive of college enrollment while survey levels is close to uninformative with or without the test VAM control. Only for relatively lower-stakes high school diploma outcomes do survey and test VAM display similar predictiveness, while test VAM continues to dominate for advanced high school diplomas. For a parent choosing schools to maximize the odds of college enrollment or persistence, simulations suggest that test VAM information achieves 30% to 37% of the payoff of an oracle who knows college effects with certainty. On top of test VAM, survey information does essentially nothing to close the gap between feasible and oracle payoffs.

The next section describes the New York City schools setting as well as the survey and test score data underpinning our analysis. Section 3 details the econometric framework used to identify, estimate, and validate VAMs for school effects on surveys, test scores, and longer-run outcomes related to graduation and college enrollment. This section also explains how we gauge the predictive power of survey and test VAM for school effects on longer-run outcomes. Our analysis is predicated on the notion that parents and educators aim to use published school quality information to gauge a schools' impact on important longer-run outcomes. Section 5 formalizes this view by modeling an optimal school choice problem in which parents use estimated VAMs and survey averages to maximize true college value-added for their students.

2 Setting and Data

2.1 Data

Our sample includes fifth-grade applicants to New York City middle schools for the 2015-2016 school year and eighth-grade applicants to New York City high schools for 2012-2013 through 2016-2017 school years. The New York City Department of Education (DOE) provided data on each applicant's ranking over schools and priorities at each school program, and the offers made by the centralized assignment system. We also have data on school enrollment, student demographics, and achievement outcomes. Data on achievement outcomes include New York State standardized test scores in grades 3-8 and SAT scores. Our middle school test outcome is the average of grade 6 math and ELA scores; in high school, we use SAT scores, a test typically taken in grade 11. Test scores are standardized to be mean zero and standard deviation one among all test takers in the year. The NYC DOE also provided student-level survey responses, information on high school graduation, and information on college-going and persistence, from the National Student Clearinghouse.

New York uses the student-proposing deferred acceptance (DA) algorithm with a single lottery tie-breaker to assign students to schools at both middle and high school. Students rank programs within a school. At either middle or high school, New York applicants can rank at most 12 programs. The system uses a mix of lottery and non-lottery tie-breakers. At the city's "unscreened" schools, the tie-breaker is a random lottery number. At "screened" schools, the non-lottery tie-breakers can include past test scores, interview rankings, attendance, or grades.

Our empirical strategy uses the method in Abdulkadiroğlu et al. (2017, 2022) to compute each applicant's probability (or "risk") of assignment to each school. Our analysis aggregates the program-level assignment risk for multiple programs in a school to the school-level assignment risk. Since New York has non-lottery tie-breaking, the risk of assignment depends on bandwidths like in regression discontinuity designs. The upshot is that conditioning of

the DA propensity score generates a randomized trial, with strata defined by score values.

Table 1 describes the full sample of students and schools. The sample with risk, shown in columns (2) and (4), are those facing random assignment at least one school. About three-quarters of New York's high school students are Hispanic or Black, and nearly 80% qualify for a subsidized school meal, a traditional indicator of poverty. The demographic makeup of students, the distribution of baseline scores, and the enrollment patterns of students with risk are similar to the full sample. Middle school students are similar to high school students in most respects, but are somewhat less likely to be Hispanic or Black students.

The New York high school risk sample covers 429 out of 483 schools and the middle school risk sample covers 406 out of 616 schools. The vast majority of schools with risk have have lotteries. Middle schools are also less likely to be screened than high schools.

2.2 New York School Surveys

New York first fielded the NYC School Survey in 2006, and it has been deployed annually to students in grade 6-12, all families, and all teachers and staff ever since. Between February and April, respondents complete the survey through an online platform, which captures student and family ID information, and anonymized information from educators. The initial survey design was inspired by surveys developed by the University of Chicago Consortium of School Research (Bryk et al., 2010). The design changed in 2014 following the election of a new mayor (Merrill et al., 2018). Our data covers responses to this redesigned survey. During our time period, 76% of students take the survey in high school and 91% take the survey in middle school. Parent responses are lower (35% for HS and 57% for MS). We ignore educator responses because those may be confounded by sorting of teachers to schools.

Our empirical analysis constructs a survey index for high (middle) school students by computing the largest principal component of survey responses in the 3rd (1st) year of school.[3] Survey measurement timing corresponds to typical testing times for our test score

[3]We follow Jackson et al. (2020) who uses principal component analysis to construct a behavior index

score outcomes. Responses to each question are standardized to be mean zero and standard deviation one among all respondents in the year. For high school students, the index combines questions common across cohorts. Missing survey answers are imputed with the mean. Students who answer no questions are omitted from survey calculations. As with test scores, we standardize the survey index to have mean zero and standard deviation one among all respondents each year.

3 Econometric Framework

3.1 Estimating Causal Value-Added

A collection of outcomes Y_{ik}, $k \in \{1, \ldots, K\}$, delineate student i's educational trajectory. Long-run outcomes, like graduation and college enrollment, concisely summarize student achievement and orient education policy. Short-run outcomes include standardized test scores and survey responses recording a student's assessment of her school environment. These short run test and survey measures may support future outcomes or represent independently valuable skills like grit or critical thinking. A constant-effects causal model describes student i's potential values of each outcome k if she were to attend district school $j \in \{1, \ldots, J\}$:

$$Y_{ijk} = \beta_{jk} + \varepsilon_{ik}, \tag{1}$$

where β_{jk} is school j's value-added to outcome k. The collection of value-added effects $\{\beta_{jk}\}_{k=1}^{K}$ characterizes school j's contributions to student development on all K dimensions. The term ε_{ik} reflects other factors that influence student i's outcome on dimension k, such as family background, motivation, and ability. We normalize $E[\varepsilon_{ik}] = 0$ so that β_{jk} is the population mean of potential outcome k at school j.

from a combination of grades, on-time grade progression, absences, and suspensions; see Heckman et al. (2006) for a related approach.

School enrollment links potential and observed outcomes:

$$Y_{ik} = \sum_j D_{ij}\beta_{jk} + \varepsilon_{ik}, \tag{2}$$

where D_{ij} is an indicator for student i's enrollment at school j. Students do not enroll randomly, so the average outcome at school j may mistake student heterogeneity for value-added. In particular, $E[Y_{ik}|D_{ij} = 1] = \beta_{jk} + E[\varepsilon_{ik}|D_{ij} = 1]$, which differs from β_{jk} when enrollment choices correlate with ε_{ik}.

We address such selection bias by controlling for observed factors that generate correlation between school attendance and potential outcomes. Let X_i denote a vector of control variables (de-meaned so that $E[X_i] = 0$), and let γ_k denote the coefficient from a projection of ε_{ik} on X_i. Substituting this projection into equation (2) yields

$$Y_{ik} = \sum_j D_{ij}\beta_{jk} + X_i'\gamma_k + e_{ik}, \tag{3}$$

where $E[e_{ik}] = E[e_{ik}X_i] = 0$ by definition of γ_k. Our value-added analysis is based on the assumption that $E[D_{ij}e_{ik}] = 0$ for all j and k. This is a selection-on-observables restriction requiring additive control for X_i to eliminate all selection bias in comparisons of outcomes across schools. Selection-on-observables implies equation (3) coincides with an ordinary least squares (OLS) regression of Y_{ik} on school indicators with controls for X_i.

In our empirical analysis the control vector X_i includes student demographics, lagged achievement test scores, and information from New York's centralized match. Specifically, we control for the *assignment propensity score* for each school, defined as the conditional probability of assignment to the school given a student's preferences and priorities in the match. The inclusion of these assignment risk controls makes equation (3) a *risk-controlled value-added model* (RC VAM). In a study of value-added in New York middle and high schools and Denver middle schools, Angrist et al. (2021) show a close correspondence between RC VAM estimates and instrumental variables (IV) estimates based on random tie-breaking

within the match. This suggests risk controls are an effective means of limiting selection bias. Section 4.1 reports the results of similar tests for the NYC samples and outcomes used here, again finding little evidence of bias in RC VAM estimates.

3.2 Bias-Corrected Covariance Estimation

Our analysis focuses on relationships between school VAM parameters across multiple student outcomes. These relationships are summarized by covariance parameters of the form:

$$\sigma_{k\ell} = \frac{1}{J}\sum_{j=1}^{J}(\beta_{jk} - \bar{\beta}_k)(\beta_{j\ell} - \bar{\beta}_\ell), \tag{4}$$

where $\bar{\beta}_k = J^{-1}\sum_j \beta_{jk}$. For a single outcome (setting $\ell = k$), this expression gives the variance of value-added in the finite population of schools in NYC, a measure of the variability of school quality on dimension k. For two different outcomes ($\ell \neq k$), $\sigma_{k\ell}$ is the covariance between value-added parameters across equations, revealing whether schools that boost outcome k also tend to boost outcome ℓ.

VAM variance and covariance parameters can be expressed compactly as quadratic forms in a vector of unknown value-added parameters. Let $\beta = (\beta_{11}, \beta_{21}, ..., \beta_{J1}, \beta_{12}, ..., \beta_{JK})'$ denote the $JK \times 1$ vector collecting value-added coefficients across all schools and outcomes. Then we can write

$$\sigma_{k\ell} = \beta' A_{k\ell} \beta, \tag{5}$$

where $A_{k\ell}$ is a $JK \times JK$ matrix composed of $J \times J$ blocks of zeros at all positions except its (k, ℓ)th block, which equals $J^{-1}(I_J - J^{-1}\iota_J\iota_J')$, with I_J the $J \times J$ identity matrix and ι_J a $J \times 1$ vector of 1's.

We construct estimates of these quadratic forms using RC VAM estimates and their sampling variances and covariances. Let $\hat{\beta}$ denote the $JK \times 1$ vector of value-added estimates generated from seemingly unrelated regression (SUR) estimation of system (3) for all K outcomes. These estimates are assumed to satisfy $E[\hat{\beta}] = \beta$ and have sampling variance

described by a $JK \times JK$ matrix $V = E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)']$, which can be well-approximated based on standard asymptotic theory.

A simple plug-in estimator of $\sigma_{k\ell}$ substitutes $\hat{\beta}$ for β in (5). This plug-in approach yields a biased estimate due to the non-linearity of the quadratic form. Intuitively, the variance of value-added coefficients for each outcome is inflated due to noise in the VAM estimates, while the covariance between value-added estimates on two dimensions is biased due to correlation in the underlying student-level outcomes. These plug-in biases are formalized in the expression

$$E[\hat{\beta}' A_{k\ell} \hat{\beta}] = \beta' A_{k\ell} \beta + tr\left(A_{k\ell} V\right). \tag{6}$$

This bias formula suggests the following bias-corrected estimators for the VAM covariance components:

$$\hat{\sigma}_{k\ell} = \hat{\beta}' A_{k\ell} \hat{\beta} - tr\left(A_{k\ell} V\right). \tag{7}$$

We use these bias-corrected covariance estimates to construct coefficients from bivariate and multivariate regressions relating VAM parameters across outcomes. For example, the slope coefficient from a bivariate regression of β_{jk} on $\beta_{j\ell}$ is estimated as $\hat{\sigma}_{k\ell}/\hat{\sigma}_{\ell\ell}$. Such regression coefficients allow us to compare the predictive value of test score and survey VAM in explaining VAM for longer-run outcomes like college attendance. Standard errors for all functions of the VAM parameter estimates are computed via the delta method.

4 Results

4.1 Validating VAMs

We assess the VAM validity by implementing lottery-based bias tests developed in Angrist et al. (2016) and Angrist et al. (2017). These tests are based on estimates of the equation

$$Y_{ik} = \alpha_k + \varphi_k \hat{\beta}_{d(i)k} + p_i' \delta_k + \eta_{ik}, \tag{8}$$

where $\hat{\beta}_{d(i)k} = \sum_j D_{ij}\hat{\beta}_{jk}$ is the estimate of value-added on outcome k for student i's enrolled school $d(i)$, and $p_i = (p_{i1}, ..., p_{iL})'$ is a vector of assignment propensity scores for a set of L schools. This set is limited to oversubscribed schools, defined as those with non-degenerate assignment risk so $p_{i\ell} \in (0, 1)$ for some students i. The sample for VAM validation includes students with non-degenerate risk for at least one school.

We estimate equation (8) by two-stage least squares (2SLS), instrumenting for enrolled value-added $\hat{\beta}_{d(i)}$ with a set of L offer indicators $Z_{i\ell}$ for oversubscribed schools. These offers are randomly assigned conditional on the risk controls p_i. In the absence of bias in a VAM, random offers that shift students across schools should cause outcomes to increase one-for-one with enrolled VAM, implying the 2SLS forecast coefficient φ should equal one. The deviation of the 2SLS estimate of φ from unity therefore provides a sense of average VAM bias, labeled "forecast bias" by Angrist et al. (2017). With multiple oversubscribed schools ($L > 1$) φ is overidentified, yielding a 2SLS overidentification test that checks whether the forecast coefficients generated by each instrument are equal. An omnibus test for VAM bias checks all implications of a valid VAM by combining the forecast and overidentification tests into a test with L degrees of freedom.

Table 2 displays results of these tests for survey responses, test scores, high school graduation, and college enrollment, with estimates for high schools in panel A and middle schools in panel B. VAM validity tests for survey responses, reported in column (1), show little bias in unadjusted survey levels. The 2SLS forecast coefficient estimates are close to one and the omnibus test fails to reject the null hypothesis of no bias in raw survey levels. In light of their apparent unconfoundedness, we use unadjusted survey levels as our measure of causal effects on survey responses in the following analyses.

While survey levels exhibit little bias, test results in columns (2), (3), and (4) identify substantial selection bias in test, high school graduation, and college enrollment measures based on average achievement levels. These uncontrolled VAM specifications yield forecast coefficients between 0.13 and 0.51 and omnibus tests that decisively reject the omnibus null of

no bias. In contrast, estimates in columns (5), (6), and (7) suggest that RC VAM estimates align with school causal effects for most outcomes studied here. 2SLS forecast coefficient estimates are statistically indistinguishable from one for test and college outcomes in both middle and high school and for the graduation outcome in middle school, with corresponding omnibus tests that generally fail to reject the null of RC VAM validity at conventional levels. The final row of each panel shows bias-corrected estimates of the standard deviation of value-added based on the VAM considered in each column, given by $\sqrt{\hat{\sigma}_{kk}}$. Consistent with the view that RC VAM eliminates the substantial selection bias in levels-based quality measures, RC VAM generates lower estimates for the standard deviations of school quality than do the levels measures. Nonetheless, RC VAM estimates still point to substantial variation in school quality across all outcomes for both middle and high schools. Across high schools, for example, a one-standard deviation increase in school quality boosts test scores by 0.12 standard deviations, survey responses by 0.49 standard deviations, and on-time enrollment at any college by 8 percentage points.

4.2 Contrasting Test VAM and Survey Levels

Comparisons of the predictive value of test VAM and survey levels reveal that test score value-added is a more reliable predictor of impacts on college attendance and persistence. These results appear in Table 3, which uses bias-corrected variance and covariance estimates $\hat{\sigma}_{k\ell}$ derived from VAMs to summarize relationships between causal effects across dimensions. Column (1) of panel A shows that among NYC high schools, schools that improve test scores by a standard deviation have on average 18 percentage points higher college value-added, an estimate that is highly statistically significant. The corresponding estimate for survey levels in column (2) is much smaller. Moreover, column (3) demonstrates that including both test VAM and survey levels in a multivariate regression for on-time college VAM renders the survey coefficient closer to zero and not statistically significant, while the test score coefficient and the r-squared are essentially unchanged. This indicates that the modest predictive value

of the survey measure in the bivariate model is explained by its correlation with test score VAM.

This pattern is remarkably similar for other college outcomes and at the middle school level. Results for high school value-added on two-year college persistence, reported in columns (4)-(6) of panel A, again show a stronger bivariate relationship for test score VAM than survey levels, and including both measures simultaneously results in a survey coefficient close to zero and almost no improvement in r-squared. Panel B shows corresponding results for on-time college enrollment value-added among NYC middle schools (the sample window here does not allow for an analysis of persistence). The relationship between college value-added and test score value-added is similar in middle school and high school: an increase in test VAM by one standard deviation of test scores is associated with a 17 percentage point causal increase in college attendance. In contrast, the correlation between college VAM and survey levels is close to zero and insignificant with or without test score VAM included. While these results demonstrate that the predictive value of test score VAM exceeds that of survey levels for college outcomes, it is worth noting that much of the variation in college value-added is unexplained by either measure, as evidenced by the low bias-corrected r-squared's reported in the bottom row of each panel.

Results for high school graduation outcomes are broadly similar, though we find a somewhat stronger role of survey levels for Regents Diplomas in high school. Table 4 repeats the analysis of Table 3 using high school graduation value-added rather than college value-added as the dependent variable. For on-time receipt of the typical Regents Diploma among NYC high schools, test VAM and survey levels are roughly equally predictive, as shown in columns (1)-(2) of panel A. Note, however, that the VAM bias tests Table 2 decisively reject RC VAM validity for this outcome, so this pattern may be partly driven by selection bias. Results for more demanding high school graduation outcomes are more in line with the college results from Table 3. Columns (4)-(6) show that test score VAM is more predictive of value-added for Advanced Diploma receipt among high schools. Panel B shows that survey levels is a

weak predictor of value-added on either type of diploma among middle schools, while test score VAM has modest positive predictive value for Advanced Diploma receipt.

5 VAM Without Regrets

Here, we embed alternative measures of causal school quality in a decision-theoretic framework that quantifies the value of different types of school information. Suppose a parent seeks to promote college attendance for his or her child, and has access to a list of estimates that may include survey levels, test score VAM, or both. This parent chooses a school to maximize expected college value-added given the available information. We quantify the value of an information portfolio by computing the *regret* of the resulting decision rule, defined as the difference in expected college value-added relative to an *oracle* parent who knows college value-added for all schools with certainty (the true β_{jk}'s for the college outcome). Such an oracle's decisions provide an upper bound on how well a parent could possibly do in this decision problem, so estimates of regret relative to the oracle allow for a comparison of the value of different information sets.

Table 5 reports estimates of regret based on simulations of decisions targeting college value-added for high schools. We compute these metrics by assuming the β_{jk}'s are drawn from a multivariate normal distribution with variances and covariances described by the bias-corrected VAM covariance parameters $\hat{\sigma}_{k\ell}$. The simulation draws a vector of true VAM parameters, then provides a hypothetical parent with either true college value-added (the oracle) or noisy test VAM and survey levels estimates, with mean-zero normally-distributed noise drawn based on the sampling variance matrix V. Parents then choose a most-preferred school based on posterior mean predictions of college value-added given the observed noisy estimates. To increase the simulation's realism we assume parents only consider schools listed on the rank-ordered preference lists submitted to the match in addition to the 10 most popular schools among other students at their child's middle school, rather than choosing

between all schools in the district.

Estimates of regret demonstrate that test score VAM is a better guide to school effects on college attendance than is survey levels. A parent with no information (who therefore ranks schools randomly) experiences regret of 12 percentage points for college enrollment and 8 percentage points of regret for for two-year college persistence. Adding survey levels to the information set provides only modest reductions in regret compared to a random-choice benchmark. In contrast, using test value-added for decision-making reduces regret by 30 to 35 percent. Moreover, adding survey levels to a parent's information set provides essentially no reduction in regret once the parent knows test score value-added. These results demonstrate that survey metrics provide little useful information about which schools promote college attendance after accounting for impacts on test scores.

6 Summary and Conclusions

US public school districts increasingly report school quality metrics based on measures beyond academic achievement. Such measures often include surveys assessing school climate and student satisfaction. This paper investigates the reliability of school surveys in predicting causal impacts on longer-term outcomes like college attendance, and contrasts forecasts based on surveys and achievement-based measures of test score value-added. Our key findings demonstrate that test score value-added predicts long-term impacts more reliably than school surveys, and surveys provide little incremental value beyond test score metrics when both are available.

References

Abdulkadiroğlu, A., J. D. Angrist, Y. Narita, and P. A. Pathak (2017): "Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation," *Econometrica*, 85, 1373–1432.

——— (2022): "Breaking Ties: Regression Discontinuity Design Meets Market Design," *Econometrica*, 90, 117–151.

Angrist, J. D., P. D. Hull, P. A. Pathak, and C. R. Walters (2016): "Interpreting Tests of School VAM Validity," *American Economic Review: Papers & Proceedings*, 106, 388–392.

——— (2017): "Leveraging Lotteries for School Value-Added: Testing and Estimation," *Quarterly Journal of Economics*, 132, 871–919.

——— (2021): "Credible School Value-Added with Undersubscribed School Lotteries," MIT Blueprint Labs Working Paper. Forthcoming, *Review of Economics and Statistics*.

Bryk, A., P. B. Sebring, E. Allensworth, S. Luppescu, and J. Q. Easton (2010): *Organizing Schools for Improvement: Lessons from Chicago*, University of Chicago Press.

Corradini, V. (2024): "Information and Access in School Choice Systems: Evidence from New York City," Working paper, MIT Economics.

Cunha, F. and J. Heckman (2007): "The Technology of Skill Formation," 97(2), 31–47.

Deming, D. (2023): "Multidimensional Human Capital and the Wage Structure," Chapter prepared for the Handbook of the Economics of Education.

Duckworth, A. L., C. Peterson, M. D. Matthews, and D. R. Kelly (2007): "Grit: Perservance and Passion for Long-Term Goals," .

Dweck, C. (2006): "Mindset: The New Psychology of Success," Random House.

Heckman, J. J., J. Stixrud, and S. Urzua (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 26, 411–482.

Hernández, J. C. (2013): "Bloomberg Issues Final Letter Grades for New York Schools," Chalkbeat.

Jackson, K. C., S. C. Porter, J. Q. Easton, A. Blanchard, and S. Kiguel (2020): "School Effects on Socioemotional Development, School-Based Arrests, and Educational Attainment," *American Economic Review: Insights*, 2, 491–508.

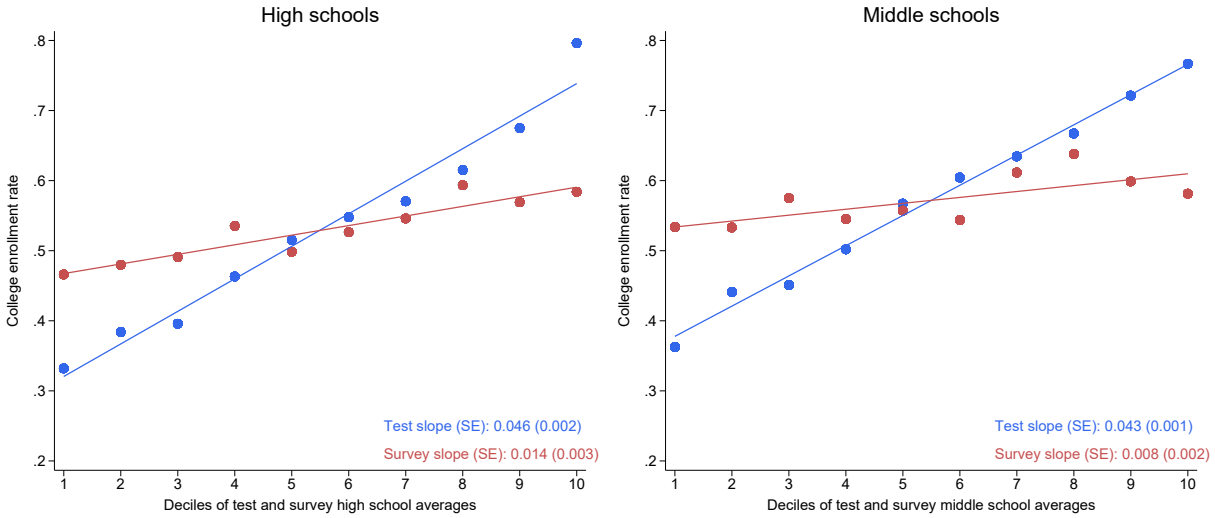
Jordan, P. W. and L. S. Hamilton (2020): "Walking a Fine Line: School Climate Surveys in State ESSA Plans," FutureEd.

Merrill, L., C. Lafayette, and S. Goldberg (2018): "Redesigning the Annual NYC School Survey: Lessons from a Research-Practice Partnership," Research Alliance for NYC Schools, NYU Steinhardt, Available at: https://steinhardt.nyu.edu/sites/default/files/2021-01/Lessons_from_a_Research-Practice_Partnership.pdf.

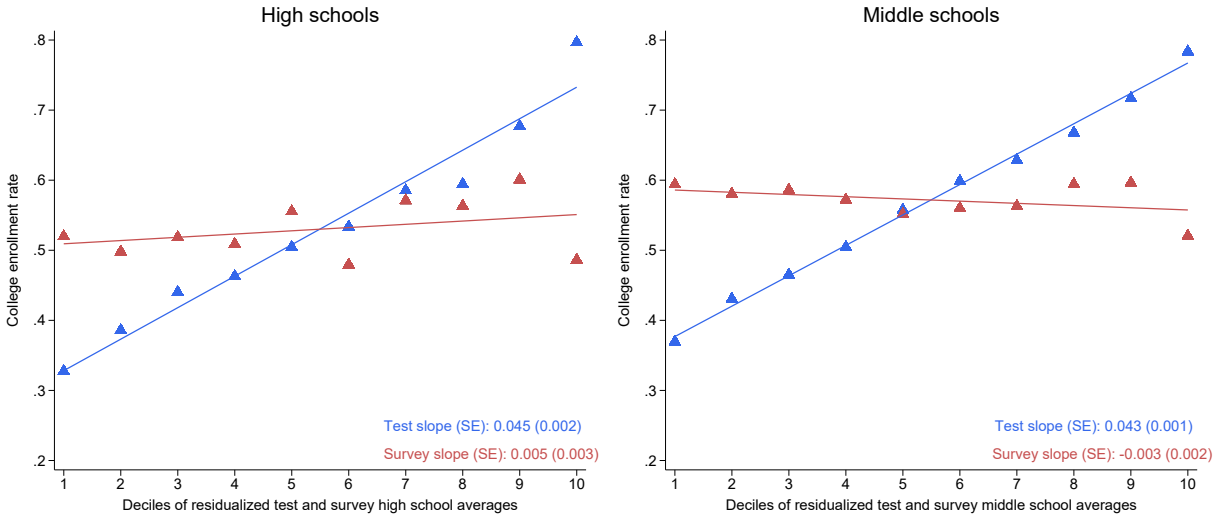
Wall, P. (2014): "Under Gentler Rating System, Schools Will No Longer Be Ranked or Graded," *ChalkBeat*.

Figure 1. College enrollment by math and survey school averages

A. School averages



B. Residualized school averages



Notes: These binned scatterplots depict average college enrollment rates conditional on test and survey school averages. Bins in Panel A are deciles of test and survey school averages. Bins in Panel B are deciles of residualized test and survey school averages. Residualized test averages are residuals from regressing test school averages on survey school averages, and vice versa for residualized survey averages. The test is SAT in high school and the New York State sixth-grade state assessment in middle school. The survey is an index constructed from the NYC School Survey; see the data appendix for details.

Table 1. Samples

	NYC high schools		NYC middle schools	
	All	With risk	All	With risk
	(1)	(2)	(3)	(4)
Demographics				
Hispanic	0.433	0.467	0.403	0.406
Black	0.289	0.275	0.271	0.224
Asian	0.143	0.146	0.164	0.199
White	0.126	0.104	0.148	0.159
Female	0.494	0.504	0.494	0.486
Free/reduced price lunch	0.786	0.794	0.738	0.733
Special education	0.189	0.072	0.207	0.191
English language learner	0.114	0.099	0.112	0.114
Baselines (std.)				
Math	-0.090	-0.026	0.032	0.100
ELA	-0.071	0.008	0.016	0.065
Enrollment				
Screened	0.131	0.074	0.048	0.027
Lottery	0.869	0.926	0.952	0.973
Share non-compliant	0.340	0.294	0.366	0.365
Share not offered	0.222	0.106	0.213	0.150
Students	269,492	77,057	68,528	24,612
Schools	483	483	616	584
Lotteries (schools with risk)		429		406

Notes: This table describes the middle and high school student samples used to estimates school value-added. Column 1 describes high school students enrolled in 9th grade in the 2012-2013 through 2016-2017 school years. Column 3 describes middle school students enrolled in 6th grade in the 2015-2016 school year. Columns 2 and 4 describe the corresponding samples of applicants with assignment risk at at least one school. Baseline characteristics and lagged scores are from 5th grade for middle school students and 7th and 8th grade for high school students. Baseline scores are standardized to be mean zero and standard deviation one in the student-level test score distribution, separately by year. Screened schools are defined as schools without any lottery programs. The share non-compliant is defined as the proportion of students who enroll other than where offered a seat; this includes students receiving no offers.

Table 2. Tests for bias in value-added models

	Levels				VAMs		
	Survey	Test	Regents diploma on-time	College on-time	Test	Regents diploma on-time	College on-time
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A: High schools				
Forecast coefficient	0.975	0.126	0.132	0.296	1.01	0.796	0.936
	(0.041)	(0.017)	(0.028)	(0.029)	(0.056)	(0.047)	(0.051)
Bias tests							
Forecast bias	0.381	2511	956	591	0.026	18.7	1.59
	[0.537]	[0.000]	[0.000]	[0.000]	[0.871]	[0.000]	[0.208]
Omnibus (20 d.f.)	17.5	2569	1010	740	22.9	51.7	27.2
	[0.621]	[0.000]	[0.000]	[0.000]	[0.294]	[0.000]	[0.129]
First-stage F	1045	1224	1644	1690	1023	1507	1547
Risk sample	49036	49614	63556	65369	50415	63866	66958
Lottery schools	429	429	429	429	429	429	429
VAM std. dev.	0.275	0.493	0.158	0.160	0.119	0.079	0.076
			Panel B: Middle schools				
Forecast coefficient	0.939	0.294	0.510	0.352	1.06	0.878	0.820
	(0.060)	(0.043)	(0.129)	(0.133)	(0.065)	(0.120)	(0.144)
Bias tests							
Forecast bias	1.02	272	14.4	23.7	0.762	1.03	1.56
	[0.313]	[0.000]	[0.000]	[0.000]	[0.383]	[0.31]	[0.211]
Omnibus (20 d.f.)	26.4	314	31.5	38.2	22.6	20.2	30.5
	[0.152]	[0.000]	[0.049]	[0.008]	[0.308]	[0.446]	[0.062]
First-stage F	157	66.6	99.4	84.5	159	128	122
Risk sample	22324	22875	18119	18738	23415	18068	18821
Lottery schools	403	405	404	405	406	406	406
VAM std. dev.	0.397	0.552	0.110	0.142	0.193	0.104	0.110

Notes: This table reports tests for bias in value-added models (VAMs) and levels. VAM includes cubic functions of baseline math and ELA scores and indicators for sex, race, subsidized lunch, special education, limited English proficiency, each interacted with application year. VAMs also include assignment propensity scores and running variable controls. High school VAMs add cubic functions of baseline attendance, disciplinary incidents, GPA, and credit accumulation. The first row in each panel reports the VAM standard deviation, calculated by subtracting a bias correction from the plug-in VAM variance. Forecast coefficients are from instrumental variables regressions of student outcomes on VAM fitted values, instrumenting fitted values with binned assignment indicators. Assignments are binned by ventile of the estimated VAM. IV models control for propensity scores, running variable controls, and baseline demographics and achievement. Test and survey outcomes are standardized to be mean zero and standard deviation one in the student-level distribution, separately by year. College and graduation outcomes are binary. The forecast bias test checks whether the forecast coefficient equals 1. The omnibus test combines the forecast test with an overidentification test that checks overidentifying restrictions implicit in the procedure used to estimate the forecast coefficient. Standard errors are reported in parentheses; test p-values are reported in brackets.

Table 3. Value-added effects on college enrollment and persistence

	On-time enrollment			2 year persistence		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: High schools			
Regressors						
Test VAM	0.186		0.172	0.147		0.147
	(0.037)		(0.038)	(0.027)		(0.028)
Survey levels		0.042	0.026		0.014	0.000
		(0.014)	(0.015)		(0.010)	(0.011)
R-squared	0.104	0.031	0.115	0.160	0.008	0.160
Schools		462			462	
Outcome mean		0.590			0.361	
			Panel B: Middle schools			
Regressors						
Test VAM	0.165		0.182			
	(0.025)		(0.026)			
Survey levels		0.004	-0.023			
		(0.012)	(0.012)			
R-squared	0.146	0.000	0.159			
Schools		594				
Outcome mean		0.611				

Notes: This table reports regressions of college value-added on test value-added and survey levels. Regression coefficients and r-squared's are derived from an estimate of the VAM covariance matrix. Entries in this matrix are covariances between VAMs for different outcomes with a correction to remove quadratic form plug-in bias. A single, stacked regression estimates VAM coefficients and sampling errors for all outcomes. Delta method standard errors are also corrected for quadratic form plug-in bias and are reported in parentheses. Standard errors are adjusted to account for a finite sample of schools.

Table 4. Value-added effects on high school graduation

	Regents diploma on-time			Advanced diploma on-time		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: High schools			
Regressors						
Test VAM	0.067		0.012	0.165		0.164
	(0.040)		(0.040)	(0.042)		(0.043)
Survey levels		0.103	0.101		0.017	0.002
		(0.015)	(0.016)		(0.015)	(0.015)
R-squared	0.010	0.139	0.139	0.052	0.003	0.052
Schools		458			458	
Outcome mean		0.725			0.195	
			Panel B: Middle schools			
Regressors						
Test VAM	0.019		-0.001	0.080		0.064
	(0.029)		(0.030)	(0.027)		(0.028)
Survey levels		0.026	0.027		0.031	0.021
		(0.012)	(0.012)		(0.012)	(0.012)
R-squared	0.002	0.019	0.019	0.032	0.023	0.042
Schools		591			591	
Outcome mean		0.830			0.361	

Notes: This table reports regressions of graduation value-added on test and survey value-added. Regression coefficients and r-squared's are derived from an estimate of the VAM covariance matrix. Entries in this matrix are covariances between VAMs for different outcomes with a correction to remove quadratic form plug-in bias. A single, stacked regression estimates VAM coefficients and sampling errors for all outcomes. Delta method standard errors are also corrected for quadratic form plug-in bias and are reported in parentheses. Standard errors are adjusted to account for a finite sample of schools.

Table 5. Regret relative to oracle school choice

	College on-time	2 year persistence at college	Regents HS diploma ontime	Advanced HS diploma ontime
	(1)	(2)	(3)	(4)
	Panel A: High school			
Applicant's information set:				
Nothing (random choice)	0.121	0.078	0.139	0.151
Regret avoided	[-0%]	[-0%]	[-0%]	[-0%]
Everything (oracle choice)	0	0	0	0
	[-100%]	[-100%]	[-100%]	[-100%]
Test VA	0.084	0.048	0.125	0.118
	[-30.5%]	[-37.6%]	[-10.0%]	[-22.0%]
Survey levels	0.100	0.071	0.088	0.144
	[-17.3%]	[-8.8%]	[-36.6%]	[-4.6%]
Test VA, survey levels	0.082	0.048	0.087	0.118
	[-32.3%]	[-37.6%]	[-37.8%]	[-21.8%]

Notes: This table reports average regret relative to oracle school choice for applicants who seek to maximize college enrollment value-added. Among schools she ranked in the match, an applicant picks a school to maximize expected college value-added given information about estimated test score VAM and survey levels. An applicant with no information picks a random school. Regret is the ex-post college value-added loss compared to the optimal choice. Simulations draw school value-added from the joint distribution of college, math, and survey effects in Tables 3 and 4. VAM estimation errors drawn from a mean-zero normal distribution with a school-specific variance from the joint VAM regression. High school applicants' choice sets contain 16 schools on average.

Putting School Surveys to the Test: Supplemental Appendix

Joshua Angrist, MIT & NBER
Peter Hull, Brown University & NBER
Russell Legate-Yang, MIT
Parag Pathak, MIT & NBER
Christopher Walters, UC Berkeley & NBER

September 2024

Appendix Figures and Tables

Table A1. Tests for statistical balance

A. High schools

	Test		Survey		Regents diploma on-time		College on-time	
	Uncontrolled (1)	Controlled (2)	Uncontrolled (3)	Controlled (4)	Uncontrolled (5)	Controlled (6)	Uncontrolled (7)	Controlled (8)
Demographics								
Hispanic	-0.033	-0.004	0.019	0.001	0.057	-0.001	0.007	-0.007
	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)	(0.004)	(0.001)	(0.003)
Black	-0.067	-0.002	-0.056	-0.003	-0.012	-0.007	-0.017	0.002
	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
Asian	0.057	0.002	0.035	-0.002	-0.005	0.005	0.020	0.003
	(0.001)	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
White	0.042	0.003	0.001	0.004	-0.041	0.003	-0.010	0.002
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
Female	0.027	-0.001	0.020	0.002	-0.016	-0.002	0.036	-0.002
	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)	(0.004)	(0.001)	(0.003)
Free/reduced price lunch	-0.049	-0.002	-0.018	-0.005	0.014	-0.003	-0.015	-0.004
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
Special education	-0.015	0.000	-0.010	0.004	-0.002	0.001	-0.002	-0.002
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
English language learner	-0.027	-0.003	0.009	-0.002	0.024	-0.002	-0.002	-0.003
	(0.001)	(0.003)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.002)
Baselines (std.)								
Math	0.296	0.011	0.139	-0.003	-0.043	0.013	0.088	0.012
	(0.002)	(0.006)	(0.002)	(0.005)	(0.003)	(0.006)	(0.002)	(0.005)
ELA	0.267	0.000	0.108	-0.005	-0.052	0.002	0.085	0.012
	(0.002)	(0.006)	(0.002)	(0.005)	(0.003)	(0.006)	(0.002)	(0.005)
Joint F test p-value	0.000	0.596	0.000	0.094	0.000	0.121	0.000	0.128
Students	269,546	77,066	269,546	77,066	269,546	77,066	269,546	77,066

B. Middle schools

	Test		Survey		Regents diploma on-time		College on-time	
	Uncontrolled (1)	Controlled (2)	Uncontrolled (3)	Controlled (4)	Uncontrolled (5)	Controlled (6)	Uncontrolled (7)	Controlled (8)
Demographics								
Hispanic	-0.032	0.006	-0.032	0.009	-0.049	-0.014	-0.022	-0.007
	(0.003)	(0.008)	(0.003)	(0.007)	(0.003)	(0.011)	(0.003)	(0.010)
Black	-0.065	-0.010	-0.020	-0.016	0.012	0.001	-0.038	-0.004
	(0.002)	(0.008)	(0.003)	(0.007)	(0.003)	(0.010)	(0.003)	(0.009)
Asian	0.038	0.008	0.006	0.003	0.001	0.011	0.029	0.009
	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.006)	(0.002)	(0.006)
White	0.058	-0.007	0.043	0.002	0.034	-0.001	0.029	-0.000
	(0.002)	(0.005)	(0.002)	(0.004)	(0.003)	(0.006)	(0.002)	(0.006)
Female	0.007	0.001	0.008	-0.011	-0.006	-0.020	0.013	-0.009
	(0.003)	(0.008)	(0.003)	(0.007)	(0.003)	(0.011)	(0.003)	(0.010)
Free/reduced price lunch	-0.048	-0.009	-0.041	-0.011	-0.056	0.009	-0.058	-0.004
	(0.002)	(0.007)	(0.002)	(0.006)	(0.003)	(0.008)	(0.003)	(0.008)
Special education	-0.020	-0.005	-0.011	0.002	-0.020	0.003	-0.023	0.003
	(0.002)	(0.007)	(0.002)	(0.006)	(0.003)	(0.010)	(0.002)	(0.008)
English language learner	-0.009	-0.006	-0.017	0.009	-0.022	-0.011	-0.004	0.001
	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.007)	(0.002)	(0.007)
Baselines (std.)								
Math	0.175	0.031	0.118	0.009	0.102	0.015	0.148	0.027
	(0.005)	(0.016)	(0.005)	(0.014)	(0.007)	(0.021)	(0.005)	(0.019)
ELA	0.149	0.017	0.115	0.003	0.121	0.013	0.143	0.026
	(0.005)	(0.016)	(0.005)	(0.015)	(0.007)	(0.021)	(0.005)	(0.019)
Joint F test p-value	0.000	0.131	0.000	0.135	0.000	0.152	0.000	0.731
Students	68,550	24,615	68,550	24,615	68,550	24,615	68,550	24,615

Notes: This table reports balance statistics, estimated by regressing baseline covariates on the VAM of the offered school and an indicator for any offer. Rows report the estimated coefficient on offered VAM. Columns 2, 4, 6, and 8 control for expected VAM, any offer risk, and running variable controls. Robust standard errors are reported in parentheses. The joint F test row reports p-values from the test of no imbalance among all baseline covariates.

Table A2. Tests for attrition

	High school		Middle school	
	College	Graduation	College	Graduation
	(1)	(2)	(3)	(4)
Offered value-added				
College on-time	0.001	-0.000	-0.005	-0.006
	(0.002)	(0.002)	(0.008)	(0.008)
Regents diploma on-time	0.003	0.004	0.019	0.012
	(0.003)	(0.003)	(0.009)	(0.010)
Test	0.002	0.006	-0.003	-0.002
	(0.003)	(0.003)	(0.007)	(0.007)
Survey	0.007	0.012	0.002	0.001
	(0.002)	(0.003)	(0.006)	(0.006)
Mean follow-up rate	0.831	0.785	0.768	0.713
N	328,054		71,815	

Notes: This table reports differential attrition estimates statistics, estimated by regressing an indicator for non-missing follow-up outcome on the VAM of the offered school and an indicator for any offer. All models control for expected VAM, any offer risk, and running variable controls. Robust standard errors are reported in parentheses.

Table A3. Tests for bias in value-added models

	Levels				VAMs			
	Advanced diploma on-time	4 year college enrollment	2 year persistence at college	2 year persistence at 4 year college	Advanced diploma on-time	4 year college enrollment	2 year persistence at college	2 year persistence at 4 year college
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: High schools								
Forecast coefficient	0.453	0.203	0.125	0.080	1.28	1.16	0.871	1.15
	(0.018)	(0.027)	(0.031)	(0.029)	(0.037)	(0.060)	(0.078)	(0.085)
Bias tests								
Forecast bias	928	854	811	1039	58.8	6.95	2.76	2.95
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.008]	[0.097]	[0.086]
Omnibus (20 d.f.)	1075	945	871	1099	75.0	23.8	20.7	29.8
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.252]	[0.417]	[0.073]
First-stage F	2014	1600	1766	1785	1912	1255	1550	1494
Risk sample	62319	65952	65632	64955	63940	66077	66485	66958
Lottery schools	429	429	429	429	429	429	429	429
VAM std. dev.	0.191	0.183	0.164	0.164	0.086	0.07	0.051	0.044
Panel B: Middle schools								
Forecast coefficient	0.559	0.203			0.894	0.891		
	(0.123)	(0.136)			(0.120)	(0.157)		
Bias tests								
Forecast bias	12.8	34.5			0.777	0.482		
	[0.000]	[0.000]			[0.378]	[0.488]		
Omnibus (20 d.f.)	35.5	60.1			18.2	14.9		
	[0.018]	[0.000]			[0.576]	[0.784]		
First-stage F	53.6	63.3			154	102		
Risk sample	17792	18559			17833	19041		
Lottery schools	403	403			405	406		
VAM std. dev.	0.216	0.164			0.117	0.105		

Notes: This table reports tests for bias in value-added models and levels for an extended set of outcomes. These estimates follow the procedures described in the note to Table 2.

Table A4. Value-added effects on 4 year college enrollment and persistence

	On-time enrollment			2 year persistence		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: High schools			
Regressors						
Test VAM	0.171		0.157	0.122		0.120
	(0.031)		(0.033)	(0.022)		(0.023)
Survey levels		0.041	0.026		0.016	0.005
		(0.013)	(0.013)		(0.009)	(0.009)
R-squared	0.109	0.036	0.123	0.156	0.016	0.158
Schools		462			462	
Outcome mean		0.436			0.294	
			Panel B: Middle schools			
Regressors						
Test VAM	0.160		0.171			
	(0.024)		(0.025)			
Survey levels		0.010	-0.016			
		(0.012)	(0.011)			
R-squared	0.173	0.003	0.181			
Schools		594				
Outcome mean		0.500				

Notes: This table reports regressions of 4 year college enrollment and persistence value-added on math value-added and survey levels. College persistence is available only for the high school cohorts. The models and derivation procedure used to compute these estimates are described in the note in Table 3.

Table A5. Instrumental variables estimates of value-added effects

	College enrollment on-time			2 year persistence at college			Regents HS diploma on-time			Advanced HS diploma on-time		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: High schools												
Regressors												
Test VAM	0.160		0.161	0.103		0.099	-0.021		-0.064	0.192		0.187
	(0.054)		(0.056)	(0.048)		(0.050)	(0.045)		(0.046)	(0.044)		(0.046)
Survey levels		0.014	-0.002		0.016	0.007		0.073	0.079		0.027	0.009
		(0.021)	(0.022)		(0.018)	(0.019)		(0.017)	(0.018)		(0.017)	(0.018)
Lottery schools			431						431			
Panel B: Middle schools												
Regressors												
Test VAM	0.115		0.130				-0.008		-0.034	0.089		0.100
	(0.061)		(0.065)				(0.054)		(0.056)	(0.055)		(0.059)
Survey levels		-0.002	-0.022					0.030	0.035		0.001	-0.014
		(0.027)	(0.029)					(0.023)	(0.024)		(0.025)	(0.026)
Lottery schools		424							423			

Notes: This table reports instrumental variable value-added model (IV VAM) estimates of test and survey value-added effects on high school graduation and college achievement. The estimates in column (1) report regressions of college enrollment value-added on estimated test VAM. These estimates are derived via the omitted variables bias (OVB) formula from student-level IV estimates of the regression of college enrollment outcome on estimated test score VAM. OVB calculations use VAM covariance estimates and standard errors as in Tables 3 and 4.

Table A6. Instrumental variables estimates of regret relative to oracle school choice

	College on-time	2 year persistence at college	Regents HS diploma ontime	Advanced HS diploma ontime
	(1)	(2)	(3)	(4)
		Panel A: High school		
Applicant's information set:				
Nothing (random choice)	0.134	0.145	0.127	0.132
Regret avoided	[-0%]	[-0%]	[-0%]	[-0%]
Everything (oracle choice)	0	0	0	0
	[-100%]	[-100%]	[-100%]	[-100%]
Test VA	0.102	0.120	0.124	0.093
	[-23.9%]	[-17.5%]	[-2.4%]	[-29.3%]
Survey levels	0.128	0.140	0.090	0.119
	[-4.3%]	[-3.9%]	[-29.1%]	[-9.4%]
Test VA, survey levels	0.102	0.120	0.090	0.093
	[-23.7%]	[-17.5%]	[-29.5%]	[-29.4%]

Notes: This table reports IV VAM estimates of the regret figures reported in Table 5. These estimates are computed from simulations that calibrate the joint distribution of school quality using the IV VAM estimates in Appendix Table A5.