1-4-2012

Value Added of Teachers in High-Poverty Schools and Lower Poverty Schools

Tim R. Sass
Georgia State University

Jane Hannaway
American Institutes for Research

David Figlio
Northwestern University

Li Feng
Texas State University

W.J. Usery Workplace Research Group Paper Series

Working Paper 2012-1-4
January 2012

Value Added of Teachers in High Poverty Schools and Lower Poverty Schools

Tim R. Sass
Georgia State University

Jane Hannaway
American Institutes for Research

Zeyu Xu
American Institutes for Research

David Figlio
Northwestern University

Li Feng
Texas State University

Georgia State University
Andrew Young School of Policy Studies

Value Added of Teachers in High-Poverty Schools and Lower Poverty Schools

Tim R. Sass
Georgia State University, CALDER

Jane Hannaway
Zeyu Xu
American Institutes for Research, CALDER

David Figlio
Northwestern University, CALDER

Li Feng
Texas State University

January 19, 2012

Abstract

Using student-level microdata from 2000-2001 to 2004-2005 from Florida and North Carolina, we compare the effectiveness of teachers in schools serving primarily students from low-income families (>70% free-and-reduced-price-lunch students) with teachers in schools serving more advantaged students. The results show that the average effectiveness of teachers in high poverty schools is in general less than teachers in other schools and there is significantly greater variation in teacher quality among high poverty schools. These differences are largely driven by less productive teachers at the bottom of the teacher effectiveness distribution in high-poverty schools. The bulk of the quality differential is due to differences in the unmeasured characteristics of teachers. We find that the gain in productivity to more experienced teachers from additional experience is much stronger in lower-poverty schools. The lower return to experience in high-poverty schools does not appear to be a result of differences in the quality of teachers who leave teaching or who switch schools, however. Our findings suggest that measures that induce highly effective teachers to move to high-poverty schools and which promote an environment in which teachers' skills will improve over time are more likely to be successful.

We gratefully acknowledge support from the U. S. Department of Education, Contract No. ED-04-CO-0036/0002 to RTI, UI Subcontract No. 15-312-0209234. We also greatly benefitted from comments on an earlier draft made by U. S. Department of Education staff, including Andy Abrams, David Goodwin and Stephanie Stullich. We would also like to thank the North Carolina Education Research Data Center (NCERDC) at Duke University and the Florida Department of Education for their assistance in providing the data. The authors are fully responsible for the opinions expressed herein as well as any errors or omissions.

I. INTRODUCTION

Whether measured by student achievement or educational attainment, the divergence in student performance between schools serving primarily low-income populations and those enrolling students from more affluent families is stark. In 2009, only 14 percent of 4th grade students from high-poverty schools scored at or above the “proficient” level in reading on the National Assessment of Educational Progress (NAEP) whereas half the students in low-poverty schools met or exceeded the threshold for proficiency. Similarly, in math just 17 percent of students from high-poverty schools scored at the proficient level or above on the NAEP while 60 percent of students from low-poverty schools performed at the proficient level or better. The differences are even larger for students in large cities, where the proportion of students who are eligible for free and reduced-price lunch (FRL) earning a proficient score in either math or reading is lower than for FRL students in the rest of the nation (Uzzell, et al., 2010). Twelfth-grade students in high-poverty schools are also less likely to earn a high school diploma than their counterparts from low-poverty schools (68 versus 91 percent) and are less likely to attend a four-year college (28 percent versus 52 percent) (Aud, et al., 2010). Correspondingly, nearly 90 percent of so-called “dropout factories,” where fewer than 60 percent of high-school freshman are still attending the same school in grade 12, are schools serving large numbers of low-income students. The majority of these schools with low graduation rates are located in the nation’s cites (Balfanz and Letgers, 2005). In half of the 100 largest cities in the United States 50 percent or more of high school students attend “dropout factories” (Balfanz and Letgers, 2004).

While student, parental and neighborhood factors undoubtedly contribute to the observed performance differential between students in low and high-poverty schools, it is hard to deny that systematic differences in school quality are partly to blame for the observed gaps in achievement and educational attainment. The source of quality differentials between schools serving primarily low-income students and those serving more affluent students, and hence the appropriate policies to ameliorate differences in school quality, are less clear, however.

Differences in teacher quality would appear to be the most likely reason for disparities in the quality of high-poverty and lower-poverty schools. Recent empirical evidence finds teachers to be the most important schooling factor affecting student achievement. Rockoff (2004), Rivkin, Hanushek and Kain (2005) and Aaronson, Barrow, and Sander (2007) show that persistent measures of teachers' contributions to student achievement or "value added" vary tremendously across teachers, and that the within-school variation in value added is at least as large as between-school variation. Similarly, Figlio and Lucas (2004) demonstrate that within-school variation in grading standards (and presumably other teacher behaviors as well) is nearly the same as the overall variation in grading standards.

Previous research has also highlighted disparities in the qualifications of teachers in schools serving primarily disadvantaged and minority students versus teachers in schools with more advantaged student bodies (Clotfelter, et al. (2005), Goldhaber, Choi and Cramer (2007), Lankford, et al. (2002)). However, while observed teacher characteristics (e.g. educational attainment, certification status, years of experience beyond the first few years, etc.) vary across schools, these differences are only weakly related to teacher performance (Harris and Sass, 2011; Clotfelter, Ladd and Vigdor, 2007).

Several recent studies also clearly establish that schools serving disadvantaged students have more difficulty hiring and retaining teachers. Teachers in general appear to prefer schools that serve students from higher income families and who are higher-achieving, and teachers working in schools with more highly disadvantaged students are more likely to leave their school district or transfer to a lower-needs school within their district (Lankford, Loeb and Wyckoff, 2002; Hanushek, Kain and Rivkin, 2004; Boyd, Lankford, Loeb and Wyckoff, 2005; Clotfelter, Ladd and Vigdor, 2005; Imazeki, 2005; Scafidi, Sjoquist and Stinebrickner, 2007; Feng, 2009). Accountability pressures can exacerbate the problems that schools serving low-achieving student populations face in retaining high-quality teachers (Feng, Figlio and Sass, 2010). Teachers with better earning opportunities are also more likely to leave teaching (Dolton and van der Klaauw, 1999), which might make it even harder for schools in economically disadvantaged area to hold on to good teachers. It may also be harder for these schools to recruit teachers to begin with, as potential teachers tend to prefer to work in schools near where they grew up (Boyd, Lankford, Loeb and Wyckoff, 2005).

The combination of evidence on the importance of teacher quality, differences in observable qualifications of teachers across schools, and the mobility patterns of teachers has led many observers to conclude that the quality of teachers in high-poverty schools is generally inferior to that of teachers in lower-poverty schools. This view has fueled policy initiatives designed to encourage promising new teachers to teach in high-poverty schools ("Teach for America") and provide incentives for existing teachers to move from lower-poverty schools to high-poverty schools. For example, the recent "Race-to-the-Top" competition graded applicants in part on whether they had plans to provide financial incentives to teach in high-poverty schools.[1] Similarly, the U.S. Department of Education has funded a set of experiments in seven school districts throughout the country, known as the "Talent Transfer Initiative" that provide

[1] The federal Race to the Top guidelines required states to include plans for ensuring an "equitable distribution of effective teachers and principals." Applicants were graded in part on "The extent to which the State, in collaboration with its participating LEAs … has a high-quality plan … to ensure that students in high-poverty and/or high minority schools … have equitable access to highly effective teachers … Plans … may include, but are not limited to, the implementation of incentives and strategies in such areas as recruitment [and] compensation."

differential compensation to highly effective teachers who agree to teach in high-poverty schools.

Despite the circumstantial evidence, in fact little is known about the relative productivity of teachers in schools serving economically disadvantaged student populations and those enrolling students from more affluent families. In this paper we seek to fill this void and inform the debate on teacher labor market policies by addressing four related research questions:

1. How does the average contribution of teachers in high poverty elementary schools compare to that of teachers in lower poverty elementary schools in terms of student achievement gains in mathematics and in reading/language arts?

2. Are there differences in the variation of teacher effectiveness within schools serving largely students from low-income families vis-à-vis the set of schools serving more affluent populations?

3. To what extent do observed teacher characteristics (e.g., certification, experience, education) in high poverty elementary schools and in schools with lower poverty levels explain differences in teacher contributions to student learning?

4. To what extent does teacher mobility contribute to differences in teacher value-added across high and lower-poverty schools?

We find that the average effectiveness of teachers in high poverty schools is in general less than teachers in other schools and there is significantly greater variation in teacher quality among high poverty schools. These differences are largely driven by less productive teachers at the bottom of the teacher effectiveness distribution in high-poverty schools. Teachers at the top of the effectiveness distribution are generally similar across school settings. Differences in observable characteristics of teachers, such as experience, certification status and educational attainment, explain at most one-fourth of the difference in teacher quality across high and lower-poverty schools; most of the quality differential stems from differences in the unmeasured characteristics of teachers. A number of factors appear to contribute to the difference in unmeasured teacher quality between high and lower poverty schools. In Florida, the lowest quality early career (0-2 years of experience) teachers in high poverty schools are much worse than the lowest quality early career teachers in lower poverty schools. However, in North Carolina the differences in productivity of early career teachers are not significantly different across school types throughout the teacher quality distribution. In both states the gain in productivity to more experienced teachers from additional experience is much stronger in lower-poverty schools. The lower return to experience in high-poverty schools does not appear to be a result of differences in the quality of teachers who leave teaching or who switch schools, however. Rather, it may be the case that the effect of experience on teacher productivity may depend on the setting in which the experience is acquired. If there are positive spillovers among teachers that depend on teacher quality (ie. teacher "peer effects") or if exposure to challenging student populations lessens the future productivity of teachers (i.e. leads to "burn out"), teachers in schools serving large proportions of low-income students may simply not improve much as time goes by.

Our analysis has important implications for public policy. Given the large variation in teacher effectiveness, major improvements in student outcomes could be realized if schools were to identify the most successful teachers and deploy them in the settings where they could make the most difference. Having improved information about the likely contribution of teachers in high poverty schools and in other schools, and the degree to which these differences can be explained by the types of factors that are observable ex ante, may help frame the magnitude of the potential problem of staffing schools serving disadvantaged students. And understanding the variation of measured teacher effectiveness within and across high poverty schools and other schools can offer insight into the potential scope for and design of teacher compensation policies aimed at attracting and retaining highly effective teachers in the most challenging schools. It may also have implications for the design of performance accountability systems, in particular the balance targeted to the school level and individual teacher level.

II. Data and Sample

In this study we use student-level microdata from two states, Florida and North Carolina. Currently, these are the only states in which teachers and students can be linked to specific classrooms across all schools in the state for several years. The two states' data systems have some differences (e.g., different tests), but are generally comparable in terms of characteristics of students and teachers. We coordinate the analysis to ensure reasonable comparability in results across the two states.[2]

A. Sources

The primary source of data for Florida is the Florida Department of Education’s K-20 Education Data Warehouse (FL-EDW). The FL-EDW is a longitudinal data system that includes individual records for students and school personnel, as well as information on courses and schools. Both students and teachers can be linked to specific classrooms with a unique course offering identification variable. Available student-level information includes basic demographics (age, race/ethnicity, gender), program participation (special education, limited English proficiency, gifted, free/reduced-price lunch) and test scores. Teacher data includes demographics, experience, educational attainment, certification status and certification exam scores. Basic demographic and experience information is also available for principals.

[2] Specific variable definitions for the two states are provided in an Appendix.

North Carolina data used in this analysis were collected by the North Carolina Department of Instruction (NCDPI) and contain detailed administrative records on students, teachers, and schools. Each student enrolled in North Carolina public schools is assigned a unique randomized identifier which allows us to track individual students over time. Student level information includes the school attended, ethnicity, sex, free/reduced price lunch status, English proficiency, special education status, level of parental education, and state assessment scores (end-of-grade (EOG) test scores in reading and math from grades 3 through 8). School level data include information on student membership, locale, grade span, and AYP status. Data are compiled each year by the North Carolina Education Research Data Center (NCERDC) at Duke University. We formed a student level longitudinal file by merging annual datasets. School and teacher information was then linked to the student-level longitudinal data file.

B. Achievement Measures

Florida

Annual testing in all grades 3-10 began in Florida in the 1999-00 school year with the administration of the Florida Comprehensive Achievement Test – Norm Reference Test (FCAT-NRT). The FCAT-NRT is a version of the Stanford Achievement Test and, as the name implies, is tied to national norms. The FCAT-NRT is essentially a "no-stakes" test in Florida; it is not used for accountability purposes. In contrast, the FCAT Sunshine State Standards exam (FCAT-SSS) is used to assign school grades and determines individual student retention and secondary school graduation. The FCAT-SSS is a criterion-reference test based on the curriculum standards established in Florida. The FCAT-SSS was first administered in all grades 3-10 in the 2000-01 school year. We use the FCAT-SSS in this analysis for grades 3 through 5. Scores for the FCAT-SSS are normalized (with a mean of zero and a standard deviation of one) by grade, year and subject in order to control for any variation in the test over time. Differences in student scores from year to year are used to show student progress in terms of their performance relative to their peers' performance.

North Carolina

The state of North Carolina has required schools to administer math and reading end-of-grade exams for 3rd-8th graders since 1994 and subject-specific end-of-course exams in secondary school since 1996. These tests are typically administered during the last two weeks of the school year. At the elementary school level, math and reading scores are normalized by year and grade (with a mean of zero and standard deviation of one) so that these test scores are comparable across year and grade. Using these normalized scores, we create gain scores by taking the difference between the current year's score and the score from the previous year, and use these as our dependent variable in our elementary school analysis. We focus on grades 3 through 5 in elementary schools.

C. Samples

The study period covers school years 2000-01 through 2004-05 in both Florida and North Carolina. In both states we focus on elementary schools, grades 3-5. In order to ensure that we correctly attribute student achievement gains to individual teachers, we include only students in self-contained classrooms.

It should be noted that although teachers are directly linked to students in Florida, this is not the case for North Carolina data. In North Carolina, individual student test scores are linked to test proctors, who may not necessarily be the students' class instructors. Therefore, in North Carolina we take additional steps to verify if a test proctor is the actual teacher in a classroom. This is done by aggregating the test-score data (which contain unique proctor IDs) to the classroom level and linking test classes with classroom level data on instruction classes (which contain unique teacher IDs) and then comparing the aggregate student demographic characteristics (class size, number of White students, and number of male students) of a test classroom with those of an instruction classroom. If the test classroom sufficiently resembles the instruction classroom, we have reasonable confidence that the test proctor is also the instructor and therefore establish links between that teacher and his students. This strategy has been successfully applied to North Carolina data in earlier research,[3] and in this study we keep only those teachers who can be verified to be instructors in classrooms under study.

Second, we restrict the samples to reduce heterogeneity at the school level by excluding charter schools. The restriction has a relatively small effect on the number of schools included. In our first study year (2000/01) there were no charter schools in North Carolina and only 30 in Florida.

Third, in order to reduce noise and generate more reliable teacher value-added estimates, we limit analysis to classes with 10-40 students. Classes with more than 40 students are more likely to be test classes rather than instruction classes. Therefore including these large classes increases the probability of inappropriately attributing student performance to test proctors instead of instructors. On the other hand, small classes offer us too few student observations for each teacher, reducing the precision of teacher performance estimates for those teachers

Fourth, we eliminate schools that switched poverty status (i.e. moved between the 0-70 and 70-100 percent FRL categories). This facilitates our comparision of different types of schools and eliminates schools on the “borderline” of our 70 percent FRL poverty designation.

Table 1 summarizes the size of our analysis sample relative to the population of teachers in each state. We compare schools with poverty levels (defined by the percentage of students

[3] See Xu, Hannaway and Taylor, 2011, for details.

who are eligible for free/reduced-price lunch) above 70 percent with schools below 70 percent. We also compare schools at greater extremes: those above 70 percent poverty with those less than 30 percent poverty.[4] After applying the sample restrictions we are able to produce value-added estimates for over 9,000 unique elementary school teachers in Florida and nearly 8,000 teachers in North Carolina. The distribution of teachers across schools is somewhat different across the two states. In Florida, 35 percent of teachers with value-added scores teach in high-poverty (70-100% FRL) schools whereas only 19 percent of North Carolina teachers with value-added estimates teach is such schools. Likewise, 20 percent of the Florida teachers in our sample are in the most affluent schools (0-30 percent FRL) while 27 percent of North Carolina teachers are in the least impoverished schools.

Table 2 presents descriptive statistics for students, teachers and principals in our sample, broken down by school poverty level. Test of mean differences across school poverty categories are calculated based on standard errors clustered at the school level for students and teachers. As expected, higher and lower poverty schools serve different student populations. Compared to lower poverty (<70% FRL and <30% FRL) schools, high poverty (>70% FRL) schools in both states have a larger share of African-American and Hispanic students. Indeed, the average percent of Black students in high poverty schools is more than twice the percent in lower (<70% FRL) poverty schools and more than four times the percent in the lowest (<30% FRL) poverty schools. And in both states student performance levels[5] are considerably lower in low poverty schools than in high poverty schools.[6]

---

[4] To some extent these breaks are arbitrary, but have some rationale. Seventy percent is about the average poverty level in Title I schools in both states (NC=63 percent; FL=72 percent). And 30 percent is about the average poverty level in non-Title I schools in both states (NC=31 percent; FL=30 percent).
[5] Note these are normalized performance measures, as discussed earlier.
[6] Indeed, in North Carolina students in schools with 0-70 percent FRL and in schools with 0-30 percent FRL outperform their counterparts in schools with more than 70% FRL students, on average, by about 0.5 and 0.7 standard deviations respectively in both math and reading. These findings are comparable to the Florida results.

In Florida the average performance *gains* in math are significantly higher in high poverty (>70%) schools than other schools. This is different from North Carolina, where the average performance gains in math in high poverty schools are not different than gains in lower poverty schools. Likewise, in Florida average reading gains are lower in the most affluent schools (0-30 percent FRL) than in the highest poverty schools (70-100 percent FRL) while in North Carolina the opposite is true; reading gains are higher in the least impoverished schools. In all cases the mean differences in achievement gains are quite small, ranging from two to four percent of a standard deviation.

Teacher qualifications are also different in high poverty and lower poverty schools in both states. Although Florida tends to have less experienced teachers in general, in both Florida and North Carolina high poverty schools have a larger percentage of first-year teachers than lower poverty schools.[7] In Florida, the rate of inexperienced teachers (those teachers with two or fewer years of experience) in high poverty schools is one-and-a-half times that in lower-poverty schools while in North Carolina the differences are much more modest. Florida teachers in high poverty schools are also less likely to have a graduate degree and teachers in high poverty schools are less likely to hold a regular license or be National Board certified than teachers in lower-poverty schools in both states. In Florida, principals in high poverty schools are more likely to be new to the school than principals in lower poverty schools. In North Carolina, the PRAXIS scores of teachers in high poverty schools are also significantly lower (by 0.03 to 0.04 standard deviations) than the Praxis scores of teachers in lower poverty schools.

[7] The higher proportion of rookie teachers in Florida is likely due to the fact that Florida was experiencing rapid population growth during the sample period and class sizes were restricted due to a constitutional class-size limit.

III. Analytic Strategy

A. General Approach

Our basic strategy involves four steps. First, we estimate the determinants of individual student achievement in a "value-added" framework that takes into account prior schooling inputs (captured by the prior-year test score), student and family characteristics, peer influences, teacher characteristics and school characteristics.[8] A relatively generic version of the value-added model can be specified as:

$$A_{ijkmt} = \lambda A_{it-1} + \boldsymbol{\beta}_1 \mathbf{X}_{it} + \boldsymbol{\beta}_2 \mathbf{P}_{-ijmt} + \boldsymbol{\beta}_3 \mathbf{S}_{mt} + \delta_k + \nu_{it}\,, \qquad (1)$$

Where A represents student achievement, **X** is a vector of student/family characteristics, **P** is a vector of classroom peer characteristics and **S** is a vector of school-level characteristics. The subscripts i, j, k and m denote individual students, classrooms, teachers and schools respectively. The variable δ_k is an indicator for teacher k and ν_{it} is a random error term. The estimated value of δ_k represents the effect of an individual teacher on her students' average performance, holding constant other factors that influence student achievement, or simply the "teacher effect." Commonly estimated models differ along three dimensions: the assumed persistence of prior educational inputs, how student characteristics are taken into account and whether teacher effects are weighted based on their precision. As specified, equation (1) allows for partial decay over time of the impact of schooling inputs on achievement. If the effect of past school-based inputs does not decay, then λ=1 and the dependent variable can be expressed as the gain in student achievement, ΔA. The vector **X** can be limited to observable time-varying and time-invariant student and family characteristics. Alternatively, a student fixed effect can be included in the

[8] For a detailed discussion of value added models and the assumptions underlying them, see Boardman and Murnane (1979), Todd and Wolpin (2003), Rivkin (2007) and Harris, Sass and Semykina (2011).

model that captures both observed and unobserved time-constant student/family characteristics. Finally, the precision of teacher effects will vary with the number of tested students per teacher. The varying degree of precision can be taken into account by "shrinking" less precise estimates toward the population mean (which equals zero in the case of teacher fixed effects).

In our empirical analysis we initially estimate eight different variants of the value-added model (partial/complete persistence x with/without student fixed effects x non-shrunken/shrunken). However, for the sake of parsimony we focus on "shrunken" estimates from the model with partial persistence and no student fixed effects, as recent experimental and simulation-based evidence suggests this model is likely to produce relatively unbiased estimates of teacher effects under a range of conditions (Kane and Staiger (2008), Guarino, Reckase and Wooldridge (2011)). We distinguish between schools serving primarily students from low-income families (70 percent or more FRL) and those serving fewer students in poverty (less than 70 percent FRL).[9] Separate teacher-effect estimates are produced for each teacher unique teacher/school-type combination.

Despite estimating a variety of value-added models, there is no guarantee that our estimates of teacher quality are unbiased. As Rothstein (2010) notes, if students are dynamically assigned to teachers on the basis of prior unobserved shocks to student achievement and these shocks are serially correlated, then controlling for observable student characteristics or even adjusting for unobserved time-invariant student heterogeneity via student fixed effects, will not be sufficient to produce unbiased teacher effects. Using data from a single cohort of students in North Carolina, Rothstein uncovers evidence of future teacher "effects" on current achievement,

[9] We only present estimates broken down by high-poverty (70+% FRL) and lower-poverty (<70% FRL) school types. However, we also estimated teacher/school-type effects for high-poverty (70+% FRL) and low poverty (<30% FRL) schools. Results using this breakdown of school types are similar to those presented in the paper available upon request.

suggesting value-added measures of teacher performance are indeed biased. However, Koedel and Betts (2011) find evidence that dynamic sorting of student and teachers to classrooms is transitory and that observing teachers over multiple time periods mitigates the dynamic sorting bias envisioned by Rothstein. Further, Kane and Staiger (2008), compare experimental evidence on the impacts of randomly assigned pairs of teachers to their pre-experimental value-added scores and fail to reject the equivalence of the two measures for a number of value-added models. Likewise, based on simulation evidence, Guarino, Reckase and Wooldridge (2011) find that teacher rankings generated from partial-persistence value-added models can correlate relatively well with true rankings under a variety of plausible student-teacher assignment mechanisms. For the present analysis, we are concerned with comparing average teacher quality across different types of schools, rather than distinguishing the quality of individual teachers. So long as any potential bias associated with non-random sorting of students and teachers to classrooms is not systematically related to school poverty status, then our estimates of teacher quality differences across schools will be valid.

In the second step of our analysis, we separate the estimated teacher-school type effects into sub-samples of teachers in higher and lower poverty schools. We determine how the mean teacher effect and the dispersion of teacher effects vary across schools serving different proportions of students in poverty. For each sub-sample we then determine the relationship between teacher credentials and effectiveness by regressing the teacher effect estimates on a set of observable teacher characteristics, including experience indicators, and indicator for post-baccalaureate degrees, and indicators for certification status.

Third, we decompose the difference in teacher effectiveness across schools serving students from different poverty levels into three components: differences due to differences in observable teacher characteristics or endowments, differences due to differences in the marginal effect of teacher characteristics on student achievement and differences due to the interaction of the differences in endowments and marginal effects. Thus, for example, we can determine if differences in average teacher quality are due to differences in observable qualifications or to differences in the "payoff" from qualifications (e.g. the impact of additional experience on student achievement) across school types.

Finally, we explore how differences in teacher mobility may explain differences in the experience-effectiveness profile of teachers in high-poverty and lower-poverty schools. We compare the effectiveness or value-added of "stayers" and "leavers" for each school type.

B. Methodological Challenges

Non-random sorting of teachers and students

As described above, one of the key challenges in estimating a teacher's contribution to student performance is the non-random sorting between teachers and students both across and within schools. To mitigate bias resulting from non-random matching of teachers to students based on time-invariant student characteristics such as innate ability (i.e. "static tracking"), we estimate models with student fixed effects in addition to models with controls for observed student covariates. [10] If selection is based solely on fixed characteristics (whether observed or not), the student fixed effects would eliminate any selection bias. Further, our value-added estimates are based on multiple observations of teachers over time. Thus even if teacher assignments are based on time-varying student performance ("dynamic tracking"), but the

[10] With student fixed effects and the assumption of non-decay of prior inputs the value-added model becomes: $A_{it} = \lambda A_{it-1} + \boldsymbol{\beta}_1 \mathbf{X}_{it} + \boldsymbol{\beta}_2 \mathbf{P}_{-ijmt} + \boldsymbol{\beta}_3 \mathbf{S}_{mt} + \gamma_i + \delta_k + \nu_{it}$, where γ_i represents the individual student-specific effect. The student covariate model replaces student fixed effects with student time-invariant (and quasi-time-invariant) characteristics.

assignment strategies are transitory (e.g. because principal turnover leads to varying assignment rules), are estimates should not be seriously biased.

Distinguishing Teacher and School-Setting Effects

A second challenge to estimating teacher value-added is distinguishing teacher effects from school effects. Our model includes school-level characteristics, where available, to control for observed characteristics of the school. In particular, in Florida we account for observed traits of the principal, including administrative experience, the square of experience, and whether the principal is new to the school. However, to the extent that there are important unmeasured school characteristics that influence student achievement gains, they will become part of the estimated teacher effect. Often value-added models include school "fixed effects" in order to capture the impact of unobserved time-invariant school-level factors. However, when school fixed effects are employed, individual teacher effects are measured relative to other teachers at the same school. Such a strategy is not relevant in the present study since our goal is to make comparisons of teacher quality across schools.

In order to determine the extent to which teacher estimates are picking up unmeasured school-level influences we conduct an analysis with teacher-school type effects, where school-type refers to the school's poverty level. Each unique "teacher-school type" combination is therefore treated as a separate teacher effect. Thus teachers who teach in both high poverty (>70%) and lower poverty (<70%) schools would generate two teacher effects estimates, one for years in which they are teaching in a high poverty school and another for years in which they are teaching in a lower poverty school.[11] For those teachers who switch school types (and thus generate two teacher effect estimates) we compare their estimated effects across the two school

[11] We do not conduct an analysis with the <30% poverty because of an inadequate number of "switchers" in this smaller set of schools.

types. If the teacher effects for these “switchers” are not significantly different across school types, it suggests that our estimates of teacher quality are not significantly biased by unmeasured school-level characteristics. If, however, there are significant differences in the within-teacher effects across school types, there are two possible conclusions. It could be that there are significant school-quality differences between school types or that teachers perform better in some school environments than in others (i.e. a match-specific component to teacher quality).

Noise in Teacher Effect Estimates

As noted earlier, the estimated teacher effects are essentially the average gain of a teacher’s students, conditional on student, peer and school characteristics that are beyond the control of the teacher. Student test scores tend to be “noisy,” that is the same student will not achieve the same score on an exam each time they take it due to random factors like whether they got a good night’s sleep, whether they are feeling ill on exam day, and whether or not they get lucky when guessing between a couple of possible multiple-choice answers. Such random fluctuations tend to cancel out when averaging over a large number of students. We therefore impose a restriction of a minimum of 10 students per teacher when reporting estimated teacher effects.

In addition to placing a restriction on the minimum number of students per teacher, it is also possible to mitigate dispersion in estimated teacher effects through the use of “Empirical Bayes” or “shrinkage” estimators (see Morris (1983), Jacob and Lefgren (2005)). The shrunken estimates of teacher productivity are essentially a weighted average of the individual teacher effect estimates and the average teacher effect estimate, with greater weight given to the individual estimates the smaller is their standard error. As noted by McCaffrey et al. (2010), standard fixed-effects software routines compute fixed effects relative to some arbitrary hold-out unit (e.g. an omitted teacher), which can produce incorrect standard errors and thus inappropriate shrunken estimates. Therefore, to estimate the teacher effects and their standard errors we employ the Stata routine *felsdvregdm*, developed by Mihaly et al. (2010), which imposes a sum-to-zero constraint on the teacher estimated teacher effects and produces the appropriate standard errors for computing the shrunken estimates of teacher value added. For each state, our estimates of teacher effectiveness can therefore be interpreted as the impact of a teacher on student learning gains relative to the average teacher in our sample.

IV. RESULTS

A. Differences in Mean Teacher Value-Added Across High/Lower Poverty Schools

Table 3 compares value-added estimates of teachers in high and lower poverty elementary schools by state and subject. Eight sets of estimates and their standard errors are shown for each subject in the table, based on alternatives to three key specification choices: partial persistence (PP) versus complete persistence (CP) in prior schooling inputs, use of student covariates (SC) versus student fixed effects (SFE) to control for student heterogeneity and whether or not shrinkage techniques are employed. Standard errors for the mean value-added as well as for the difference in means are also provided. For North Carolina, the results are relatively robust across alternative value-added model specifications. In all sixteen subject/model combinations, the point estimate on the difference between average teacher quality in high-poverty and lower-poverty schools is positive and the difference is statistically significant in 13 of the 16 variants. For reading teachers in Florida the results are also relatively consistent across models, with positive point estimates in six of eight cases and in all four cases where the mean difference is significant, average value-added is lower in high-poverty schools.

For math in Florida the results are more varied, with two positive and significant mean differences and three negative and significant mean differences.

Recent evidence by Kane and Staiger (2008) and by Guarino, Reckase and Wooldridge (2011) suggests that models with partial persistence and without student fixed effects are likely to exhibit minimal bias under a variety of plausible teacher-student sorting scenarios. If we focus on this model then we obtain results that are consistent across both states and subjects. The non-shrunken estimates from this model indicate significantly higher average teacher quality in lower-poverty schools for both math and reading in both states. The shrunken estimates also indicate higher average teacher quality in lower-poverty schools, though the difference is not statistically significant for math teachers in Florida. The differences in average teacher quality across school types vary from one to seven percent of a standard deviation, roughly comparable to the difference in performance between a rookie teacher and a teacher with two years of experience (Harris and Sass, 2011).

In addition to differences in teacher quality across high-poverty and lower-poverty schools, we also estimate differences in average teacher quality across students within a given school type. Table 4A presents differences in average teacher quality between FRL and non-FRL students in high poverty and lower poverty schools in Florida. Corresponding estimates for North Carolina are presented in Table 4B. Interestingly, in Florida we find there is a tendency for FRL students in lower poverty schools to have lower quality teachers, while there is no consistent pattern in high poverty schools. Of the 16 comparisons (8 value-added models x 2 subjects), in 14 we find that average teacher quality is worse for FRL students than for non-FRL students in lower poverty schools. In contrast, in high poverty schools, results are mixed; in 8 cases FRL students have lower quality teachers than their classmates who come from more affluent families and in the other 8 cases they have higher quality teachers on average. If we focus on results from the preferred value-added model (Shrunken, SC, PP), in all four subject/school type combinations, students who are eligible for free/reduced-price lunch are taught by less effective teachers on average.

In North Carolina the results are much more uniform across school settings. In both high poverty and lower poverty schools in North Carolina FRL students tend to be taught by lower quality teachers relative to their schoolmates from more affluent families. For math, the difference is positive and statistically significant in all 16 cases (8 models x 2 school types). However, the differences in average teacher quality are generally small; in most cases they are less than one percent of a standard deviation. In the preferred specification (Shrunken, SC, PP) the differences are largest, at about three percent of a standard deviation in lower poverty schools and roughly one percent of a standard deviation in high poverty schools. In reading the results are somewhat more mixed; the difference in teacher quality between FRL and non-FRL students is statistically significant in only 10 of 16 cases. However, in the preferred specification the difference is statistically significant for both high and lower-poverty schools and is about the same as the estimated teacher quality differential in math (about one to two percent of a standard deviation).As discussed above, one challenge to our findings is that the teacher value-added estimates are absorbing the effects of unmeasured school characteristics. Thus unmeasured school conditions affecting student performance are attributed to teachers. Including a variety of school-level controls (e.g. percent of students FRL, percent of students with limited English proficiency, etc.) serves to minimize this problem. However, to further examine this issue we compare the effectiveness of teachers who taught in both high and lower-poverty (<70% FRL) schools. For these teachers we have two estimates of their effectiveness, one in a high poverty school and one in a lower poverty school. The results of this comparison, presented in Table 5, suggest there is no consistent bias resulting from omission of unobserved school characteristics. Among the 32 model/subject/state combinations, the within-teacher difference in performance across school types is significant in only six cases. In these six cases the difference is evenly split between positive and negative. If we focus on the preferred specification of student covariates and partial persistence, the difference is never significant in math, but is significant in reading for both states. However, the difference is negative in Florida and positive in North Carolina.

B. Differences in the Variability of Teacher Value-Added Across High/Lower Poverty Schools

In addition to mean value added, we also compare the within-school-type variation of teacher effectiveness. Results in Table 3 indicate that the standard deviation in teacher quality is almost always significantly higher in high-poverty schools. In Table 6 we compare teacher value-added across school types at various points of the quality distribution (based on the partial persistence, student covariate shrunken estimates). The overall variation in productivity among teachers is striking, as has been shown in earlier research (e.g., Rivkin, Hanushek and Kain, 2005). For both high-poverty and lower-poverty schools the difference between teachers' effectiveness at the 10th percentile and the 90th percentile in math is over one-third of a standard deviation in both states; in reading it is about one-fifth of a standard deviation.

The standard deviations associated with high poverty schools are significantly larger than those associated with lower poverty schools for each of the 16 model/subject combinations in Florida and for all eight of the non-shrunken estimates in North Carolina. By definition the variation in shrunken teacher effects is smaller and thus the differences across school types would be smaller as well. Thus it is not surprising that in North Carolina, only three of the eight shrunken estimate variances are significantly different across school types. The evidence thus suggests that while high-poverty schools tend to have slightly lower quality teachers on average, there is much greater heterogeneity in teacher quality within the group of schools serving primarily students from low-income households.

The difference in teacher quality across school types varies in a consistent way across teacher performance levels. For each of the four subject/state combinations, the divergence in teacher quality across school types is greatest at the low end of the teacher quality distribution and the teacher quality advantage of lower poverty schools diminishes as you move up the teacher quality distribution. For both reading and math the difference in quality between the lowest-performing teachers in high-poverty schools and those in lower-poverty schools in Florida is about 0.06 standard deviations while in North Carolina the difference is somewhat smaller, 0.03 to 0.04 standard deviations. The quality difference gradient is less in North Carolina than in Florida. In North Carolina, the teacher quality differential in math goes form 0.04 standard deviations at the bottom to being insignificantly different from zero among the best teachers in each school type. For reading teachers in North Carolina the quality differential diminishes only slightly, 0.025 standard deviations between the lowest quality teachers and 0.015 among the top decile of teachers. In Florida, the differential across school types in math drops from 0.06 at the low end to -0.02 between the best teachers. Among reading teachers it also drops considerably, though remains positive, going from 0.06 to 0.02. Put simply, the worst teachers in high poverty schools are less effective than the teachers at the bottom of the distribution in lower poverty schools. The gap narrows as one moves up the quality distribution, however. The best teachers in high-poverty schools are nearly as good or, in one case, slightly better than the best teachers in lower-poverty schools.

The distribution of teacher quality within school type is also presented graphically in Figures 1A and 1B for Florida and in Figures 2A and 2B for North Carolina. Not only are the means of the distributions different, the difference in the skew of the distributions is readily apparent, particularly the non-shrunken estimates of math teacher performance.

The differential at the bottom end of the teacher quality distribution between high and low-poverty schools poses important policy questions. Is the difference due to recruiting lower quality teachers in the first place in high-poverty schools, or is it due to having more inexperienced teachers in those schools? If inexperienced teachers in different school types are of equal quality, then policy efforts should focus on salary and other mechanisms for retaining teachers. If high-poverty schools get worse draws in the first place, then targeted hiring incentives may be optimal.

To explore this issue, we restrict our sample of teachers to those with two or fewer years of experience, and compare their effectiveness between high and low-poverty schools in Table 7. In North Carolina, the effectiveness of inexperienced teachers (both math and reading) in high-poverty schools is not significantly different from that in low-poverty schools. In contrast, in Florida the differences in teacher quality between school types across the teacher quality distribution for inexperienced teachers is very similar to that of all teachers combined. Thus it appears that the causes of teacher quality variation between high and low-poverty schools are different in the two states. In Florida, the differences are at least partly driven by differences in the quality distribution of new teachers. It appears that in Florida the worst recruits in high-poverty schools are of much lower quality than the worst recruits in more affluent schools. However, the differences in the quality of the best recruits is smaller, and in fact the very best recruits in high-poverty schools are somewhat better at teaching elementary math than are the top recruits in lower poverty schools. In North Carolina, high-poverty schools do not seem to be at a significant disadvantage in terms of recruiting new elementary school teachers, as the differences in average teacher quality are insignificant at each point in the teacher quality distribution. However, one must be cautious interpreting these results. Recall that there is a much smaller proportion of new teachers in North Carolina than in Florida during the sample period. Thus there may simply be a lack of statistical power to distinguish quality differences in new recruits in North Carolina.

The differences in the quality distribution of early-career teachers across school poverty types are also illustrated in Figures 3A and 3B (for math and reading teachers in Florida) and Figures 4A and 4B (for math and reading teachers in North Carolina). Differences are most readily apparent in the non-shrunken estimates and are more pronounced in reading than in math.

Our results thus far indicate that teacher quality is not uniformly lower in high-poverty schools. Rather, differences in average teacher quality appear to be driven in large measure by the relatively poor performance of the least effective teachers in high-poverty schools. It appears that the differential at the bottom end of the teacher quality distribution between high and low-poverty schools is driven in part by recruiting some very low-quality teachers by high-poverty schools, particularly in Florida.

C. Sources of Variation in Teacher Quality Between High-Poverty and Lower-Poverty Schools

In order to better understand the factors contributing to the differences in the distribution of teacher quality across school types we break down the value-added differential using the Blinder-Oaxaca decomposition approach. In the labor economics literature the Blinder-Oaxaca decomposition is often used to divide the wage differential between groups (e.g. male/female or black/white) into three parts: (i) that due to differences in group characteristics, (ii) that due to differences in the marginal return to characteristics and (iii) interactions between differences in group characteristics and differences in returns. In the present context we can express the decomposition of mean differences in teacher value-added as follows:

$$\bar{\delta}_k^{HP} - \bar{\delta}_k^{LP} = \boldsymbol{\gamma}^{LP}(\bar{\boldsymbol{T}}^{HP} - \bar{\boldsymbol{T}}^{LP}) + (\boldsymbol{\gamma}^{HP} - \boldsymbol{\gamma}^{LP})\bar{\boldsymbol{T}}^{LP} + (\boldsymbol{\gamma}^{HP} - \boldsymbol{\gamma}^{LP})(\bar{\boldsymbol{T}}^{HP} - \bar{\boldsymbol{T}}^{LP}) \tag{2}$$

where **T** is a vector of teacher characteristics, γ is the associated coefficient vector and the superscripts refer to teachers in high poverty and lower poverty schools. The first term in (2) is the part of the teacher quality differential due to differences in the observable characteristics of teachers, the second term is the contribution of differences in the marginal effects of teacher characteristics to the difference in average value added and the third term accounts for the interaction between differences in characteristics and returns.

Table 8 presents results from the Blinder-Oaxaca decomposition of the variation in estimated teacher effectiveness.[12] Observable teacher characteristics include a vector of teacher experience categories and indicators for teacher educational attainment and licensure status. For both states and both subjects, differences in the observable characteristics of teachers account for only a modest proportion of the difference in teacher valued-added across school poverty categories; no more than 25 percent.[13] Among the observable characteristics, only differences in professional certification status (i.e. temporary versus full certification) are statistically significant. The majority of the difference in teacher quality across school types is explained by differences in the marginal effects of characteristics. By far, the largest estimated contribution is

[12] For regressions where the estimated teacher effect is the dependent variable (Tables 8 and 9) we employ the standard (non-shrunken) teacher effect estimates since the shrunken estimates would distort the estimated marginal effects of teacher characteristics on teacher productivity. Given that the dependent variable (the teacher effect) is an estimated value we employ feasible generalized least squares to account for estimation error.

[13] The reference group in the decomposition is teachers in high-poverty schools. Therefore the differences in teacher effectiveness due to differences in the marginal effects are evaluated at the mean characteristics of teachers in high-poverty schools.

from the coefficient on the constant term. This is equivalent to the "pure discrimination" term in wage differential analyses and signifies that most of the difference in teacher quality between high and lower poverty schools is due to differences in unmeasured teacher characteristics.

The differences in the marginal returns to specific characteristics are discernible in the regression results presented in Table 9.[14] In Florida, the productivity enhancement from experience appears to be larger and continues for a much longer period in lower poverty schools than in high poverty schools. For math, there is a two-percent-of-a-standard-deviation productivity advantage to experience through 20 years of experience. After that the difference in effectiveness between veterans and the least experienced (0-2 years of experience) is not significant. In high-poverty schools there is no significant difference in the productivity of experienced and inexperperienced teachers. In reading the payoff to experience is generally greater and more persistent than in math, though like in math, the experience premium diminishes more rapidly in high poverty schools than in lower poverty schools. For North Carolina we observe a similar pattern. In both math and reading the premium to experience in lower poverty schools is quantitatively substantial (three to six percent of a standard deviation) and remains statistically significant for even the most veteran teachers. In contrast, in high poverty schools teachers with 3-5 years of experience are more productive than early-career teachers (those with 0-2 years of experience), but the differential diminishes and becomes statistically insignificant after five years of experience.

To summarize, our findings thus far indicate that teachers in high poverty schools tend to be of lower quality, though the estimated differential is greater and more robust in North Carolina than in Florida and is larger in reading than in math. Across both states and both

[14] To account for the fact that the dependent variable (teacher value added) is an estimate, rather than a known value, we employ the method developed by Lewis and Linzer (2005) and embodied in the Stata routine *edvreg*.

subjects there is significantly greater variation in effectiveness among teachers working in high-poverty schools. The greater variability in quality among teachers working in high-poverty schools appears to be caused by lower quality of the least effective teachers. The best teachers in high-poverty schools are on par with the best teachers in lower-poverty schools but the least effective teachers in high-poverty schools are much less effective than their counterparts in lower-poverty schools. The observed differences in teacher quality across school types appear to be due primarily to differences in the unmeasured characteristics of teachers. We also find that the experience-productivity relationship appears to be much stronger in lower-poverty schools. We explore possible explanations for the differential returns to experience in the next section.

D. Why is the Return to Teacher Experience Different in High-Poverty Schools?

We posit four possible explanations for the observed differences in the relationship between teacher experience and teacher productivity in different school settings. The first two possibilities relate to differences in teacher career paths. Recent work on teacher labor market decisions suggests that the relationship between teacher quality and teacher attrition/mobility may vary across school settings (Boyd, et al., 2007; Goldhaber, Gross and Player, 2007). If relatively low-quality new teachers in high-poverty schools are less likely to leave their initial schools than are their counterparts in schools serving more affluent student populations, then the experience-quality profile would be flatter in high-poverty schools. Differential attrition of low-quality teachers could be due to lower opportunity costs or less effective monitoring of teacher performance in high-poverty schools. Alternatively, differences in teacher mobility across school types could be driving observed differences in teacher quality-experience patterns. There might be a "Dance of the Lemons" whereby low-productivity teachers with some experience in lower-poverty schools eventually migrate to high poverty schools. This would tend to lower the average quality of experienced teachers in high-poverty schools (and raise the average the quality of experienced teachers in lower-poverty schools) which would make the slope of the experience-quality relationship appear flatter in high-poverty schools.

In Table 10 we present the average quality of early-career (0-2 years of experience) teachers, broken down by school type and mobility. We do not find consistent support for the notion that lower-poverty schools do a better job at culling out low-performing early-career teachers. For math in Florida, we find no difference in the first-year performance of stayers and leavers in lower poverty schools, whereas in high poverty schools the teachers who stay past the first year are of much lower quality than those who leave their initial school placement. The reverse is true for second-year teachers, however. In low-poverty schools the teachers who stay beyond their second year are of lower quality than those who leave whereas in high poverty schools there is no significant difference between movers and stayers at the end of the second year. For reading in Florida, there are no differences in quality between early-career movers and stayers in lower poverty schools whereas in high-poverty schools first-year leavers are worse than stayers and second-year leavers are better than stayers. In North Carolina, there are no significant differences in the relative quality of early-career stayers and leavers across school types.

In addition to differences in attrition/mobility, observed differences in the quality-experience profile across school types could result from direct effects of the school environment on changes in teacher quality over time. Teachers who stay in high-poverty schools may simply "burn-out" faster, resulting in smaller increases in productivity over time compared to teachers working in less stressful environments. While we have no direct evidence on the burn-out hypotheses, it is consistent with evidence that regular education teachers are more likely to leave schools with challenging student populations (Lankford, Loeb and Wyckoff, 2002; Hanushek, Kain and Rivkin, 2004; Boyd, Lankford, Loeb and Wyckoff, 2005; Clotfelter, Ladd and Vigdor, 2005; Imazeki, 2005; Scafidi, Sjoquist and Stinebrickner, 2007; Feng, 2009) and that special education teachers often cite the stress of working with students with special needs and the lack of pupil progress relative to effort expended as reasons for switching from special to regular education (Billingsley and Cross (1991)). Further, peer effects may play a role in the apparent diminished effect of experience on productivity in high-poverty schools. Feng and Sass (2008) find that teachers are more likely to leave their initial placement the greater is the gap between their own productivity and the average quality of other teachers at their school. The most effective teachers who transfer tend to go to schools whose faculties are in the top quartile of teacher quality. Jackson and Bruegmann (2009) show that a teacher's students have higher achievement gains the higher the value-added of the teacher's colleagues.

In Table 11 we present evidence on the relationship between within-teacher mobility and teacher quality across school types.[15] Differences across states are readily apparent. In North Carolina the ratio of teachers migrating from high poverty schools to lower poverty schools relative to movement in the opposite direction is about two to one whereas in Florida it is nearly four to one. In Florida we find that teachers who switch from high-poverty to lower poverty schools are about three percent of a standard deviation more productive after the switch. This is at least consistent that lower poverty schools provide environments that enhance teacher productivity over time, either through positive teacher peer effects or reduced stress. The number of teachers moving in the opposite direction in Florida, as well as switchers in both directions in North Carolina are too small to discern and significant differences.

[15] The quality differentials for switchers in Tables 5 and 11 are not directly comparable because the samples are different. In Table 5, teachers who switch schools types more than once are included. Table 10 only includes one-time switchers.

V. Conclusions and Policy Implications

This study focuses on the effectiveness of teachers in high poverty and lower poverty schools. Large and persistent differences in student performance between schools serving large proportions of students from low-income households and those primarily students from more affluent households exist. Prior research indicates that teachers are the most important school factor affecting student achievement and thus policies designed to impact teacher quality are the logical starting point for reducing gaps in school performance. Prior federal efforts, primarily Title I, as well as state efforts target dollars to schools serving the most disadvantaged students, but they provide wide local latitude in determining what happens in these schools, including the assignment of teachers.

Our findings show that teachers in high poverty schools are generally less effective than teachers in lower poverty schools, though the magnitude of the difference varies across states. We do find consistent evidence, however, that the variation in effectiveness among teachers in high poverty schools is greater than the variation among teachers in lower poverty schools. Differences in the distribution of teacher quality appear to be driven by the relatively poor performance of the least effective teachers in high-poverty schools; the best teachers in high-poverty schools are on par with the best teachers in lower-poverty schools but the least effective teachers in high-poverty schools are much less effective than their counterparts in lower-poverty schools. The observed differences in teacher quality across school types are largely due to differences in the unmeasured traits of teachers, rather than by observable characteristics. For early career teachers, the least effective teachers in high poverty schools are much less productive than the lowest quality teachers in lower poverty schools. However, the difference is not statistically significant in North Carolina. For more experienced teachers in both states the experience-productivity relationship appears to be much stronger in lower-poverty schools. The lower return to experience in high-poverty schools does not seem to be a result of differences in the quality of teachers who leave teaching or who switch schools. Rather, it may be the case that the effect of experience on teacher productivity depends on the setting in which the experience is acquired. If there are positive peer effects among teachers that depend on teacher quality or if exposure to challenging student populations produces "burn out" and lessens the future productivity of teachers, teachers in schools serving large proportions of low-income students may simply not improve much as time goes by. Consistent with these hypotheses we find some evidence that teachers who switch from high-poverty to lower-poverty schools are more productive after the switch.

Our results have implications for both school accountability and teacher labor-market policies. Because high poverty schools also tend to have lower performance levels, they tend to be schools that are identified as not performing up to par and subject to accountability pressure. To the extent that school-level measures are used in accountability systems, they are likely to inadequately appreciate the contributions of the best teachers in these schools who are performing at least as well as the top teachers in more advantaged and higher performing schools. Our findings on the variability of teacher effectiveness call for accountability mechanisms that take into account not only school level performance measures, but also the individual level teacher contributions.

Our results suggest that solutions to the achievement gap between high and lower-poverty schools may be complex. Changing the quality of new recruits or importing teachers with good credentials into high poverty schools may not be sufficient. Rather, our findings indicate that measures that promote retention of the most-effective teachers already in high-poverty schools, that induce highly effective experienced teachers in lower-poverty schools to move to high-poverty schools and which promote an environment in which teachers' skills will improve over time are more likely to be successful.

REFERENCES

Aaronson, Daniel, Lisa Barrow, and William Sander (2007). "Teachers and Student Achievement in the Chicago Public High Schools," *Journal of Labor Economics* 25: 95–135.

Aud, Susan, William Hussar, Michael Planty, Thomas Snyder, Kevin Blanco, Mary Ann Fox, Lauren Frohlich, Jana Kemp, Lauren Drake (2010). *The Condition of Education 2010*. (NCES 2010-028). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington DC.

Balfanz, Robert and Nettie Legters (2004). "Locating the Dropout Crisis: Which High Schools Produce the Nation's Dropouts? Where are they Located? Who Attends Them?, Report No. 70, Center for Research on the Education of Students Placed at Risk (CRESPAR).

Balfanz, Robert and Nettie Legters (2005). "The Graduation Gap: Using Promoting Power to Examine the Number and Characteristics of High Schools with High and Low Graduation Rates in the Nation and Each State," CSOS Policy Brief, Johns Hopkins University.

Billingsley, Bonnie and Lawrence H. Cross. 1991. "Teachers Decisions to Transfer from Special to General Education," *Journal of Special Education* 24(4):496-511.

Boardman, Anthony E., and Richard J. Murnane. 1979. "Using Panel Data to Improve Estimates of the Determinants of Educational Achievement," *Sociology of Education* 52:113-121.

Boyd, Donald, Hamilton Lankford, Susanna Loeb and James Wyckoff (2005), "The Draw of Home: How Teachers' Preferences for Proximity Disadvantage Urban Schools," *Journal of Policy Analysis and Management*, 24(1):113–132.

Boyd, Don, Pam Grossman, Hamp Lankford, Susanna Loeb, and Jim Wyckoff (2008). "Who Leaves? Teacher Attrition and Student Achievement." *NBER Working Papers 14022*, Cambridge, MA: National Bureau of Economic Research.

Clotfelter, Charles T. Helen F. Ladd and Jacob Vigdor (2005). "Who Teaches Whom? Race and the Distribution of Novice Teachers," *Economics of Education Review* 24(4):377-392.

Clotfelter, Charles T., Helen F. Ladd and Jacob L. Vigdor (2007), "How and Why Do Teacher Credentials Matter for Student Achievement?" *NBER Working Papers* 12828, Cambridge, MA: National Bureau of Economic Research.

Dolton, P. and W. Van der Klaauw (1999), "The Turnover of Teachers: A Competing Risks Explanation," *Review of Economics and Statistics*, 81(3):543-552.

Feng, Li (2009), "Opportunity Wages, Classroom Characteristics, and Teacher Mobility," *Southern Economic Journal*, vol. 75(4):1165-1190.

Feng, Li, David Figlio and Tim R. Sass (2010). "School Accountability and Teacher Mobility," *NBER Working Papers* 16070, Cambridge, MA: National Bureau of Economic Research, Inc.

Feng, Li and Tim R. Sass (2008). "Teacher Quality and Teacher Mobility," unpublished manuscript, Florida State Unversity.

Figlio, D. and M. Lucas (2004), "Do High Grading Standards Affect Student Performance?" *Journal of Public Economics*, 88(9-10):1815-34.

Fowler, R.C. 2003, "The Massachusetts Signing Bonus Program for New Teachers: A Model of Teacher Preparation Worth Copying?" *Education Policy Analysis Archives* 11.

Goldhaber, Dan, Hyung-Jai Choi, and Lauren Cramer (2007). "A Descriptive Analysis of the Distribution of NBPTS-Certified Teachers in North Carolina," *Economics of Education Review* 26(2):160-172.

Goldhaber, Dan, Bethany Gross, and Daniel Player (2011). "Teacher Career Paths, Teacher Quality, and Persistence in the Classroom: Are Public Schools Keeping their Best?" *Journal of Policy Analysis and Management* 30(1): 57-87.

Guarino, Cassandra M., Mark Reckase and Jeffrey Wooldridge (2011). "Can Value-Added Measures of Teacher Performance be Trusted?" Unpublished manuscript.

Harris, Douglas, Tim R. Sass and Anastasia Semykina (2011). "Value-Added Models and the Measurement of Teacher Productivity." Unpublished manuscript.

Imazeki, Jennifer (2005). "Teacher Salaries and Teacher Attrition," *Economics of Education Review* 24(4):431-449.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin (2004), "Why Public Schools Lose Teachers," *Journal of Human Resources* 39(2): 326-354.

Hanushek, E.A., J.F. Kain, D. O'Brien and S.G. Rivkin (2005), "The Market for Teacher Quality," *NBER Working Papers* 11154, National Bureau of Economic Research, Inc.

Harris, Douglas N. and Tim R. Sass (2010). "Teacher Training, Teacher Quality and Student Achievement. *Journal of Public Economics* 95(7-8): 798:812.

Jackson, C. Kirabo and Elias Bruegmann (2009). "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers." *American Economic Journal: Applied Economics,* 1(4): 85-108.

Jacob, Brian A. and Lars Lefgren (2005). "Principals as Agents: Subjective Performance Measurement in Education." Working Paper #11463. Cambridge, MA: National Bureau of Economic Research.

Kane, Thomas J. and Douglas O. Staiger (2008). Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation." Working Paper #14607. Cambridge, MA: National Bureau of Economic Research.

Lewis, Jeffrey B. and Drew A. Linzer (2005). "Estimating Regression Models in Which the Dependent Variable Is Based on Estimates." *Political Analysis*, 13(4): 345–364.

Lankford, Hamilton, Loeb, Susanna, and Wyckoff, James (2002). "Teacher Sorting and the Plight of Urban Schools. A Descriptive Analysis," *Educational Evaluation and Policy Analysis*, 24(1), 37–62.

McCaffrey, Daniel F., J. R. Lockwood, Kata Mihaly and Tim R. Sass (2010). "A Review of Stata Routines for Fixed Effects Estimation in Normal Linear Models." Unpublished manuscript.

Mihaly, Kata, Daniel F. McCaffrey, J.R. Lockwood and Tim R. Sass (2010). "Centering and Reference Groups for Estimates of Fixed Effects: Modifications to felsdvreg." *Stata Journal*, 10(1), 82-103.

Morris, Carl N. (1983). "Practical Empirical Bayes Inference: Theory and Applications." *Journal of the American Statistical Association* 78(381): 47-55.

Rivkin, Steven G., Eric A. Hanushek and John F. Kain (2005), "Teachers, Schools, and Academic Achievement," *Econometrica* 73(2): 417–458.

Rivkin, Steven G. (2007), "Value-Added Analysis and Education Policy," *CALDER Briefs* 1, National Center for Analysis of Longitudinal Data in Education Research, November, 2007.

Rockoff, Jonah E. (2004), "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," *American Economic Review*, Vol. 94, No.2, 247-52.

Scafidi, Benjamin, David L. Sjoquist and Todd R. Stinebrickner (2007). "Race, Poverty, and Teacher Mobility," *Economics of Education Review* 26(2):145-159.

Todd, Petra E. and Kenneth I. Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement," *Economic Journal* 113(485):F3-F33.

Uzzell, Renata, Candace Simon, Amanda Horwitz, Anne Hyslop, Sharon Lewis and Michael Casserly (2010). "Beating the Odds: Analysis of Student Performance on State Assessments and NAEP." Washington, DC: Council of the Great City Schools.

Xu, Zeyu, Jane Hannaway and Colin Taylor (2011). "Making a Difference? The Effects of Teach for America in High School." *Journal of Policy Analysis and Management* 30(3): 447-469.

Table 1A
Number of teachers in state and sub-samples, by state and school poverty category - Math

	Florida				North Carolina			
	0-30% FRL	30-70% FRL	70-100% FRL	Total	0-30% FRL	30-70% FRL	70-100% FRL	Total
Teachers of Relevant Classes in State	2,929	6,588	5,781	15,298	3,574	9,645	4,149	15,487
Teachers Linked to Students	2,929	6,588	5,781	15,298	2,468	6,606	2,625	10,383
Eliminate Charter School Students	2,644	6,433	5,614	14,691	2,399	6,591	2,610	10,273
Eliminate Classes with <10 or >40 Students	2,570	6,210	5,315	14,095	2,260	6,124	2,386	9,496
Eliminate Students with Invalid Test Grade Level	2,570	6,207	5,308	14,085	2,260	6,124	2,386	9,496
Eliminate Students with Missing Gain Scores	2,190	5,388	4,727	12,305	2,241	6,056	2,351	9,388
Eliminate Students in Schools that Switch Poverty Status	2,131	4,771	3,796	10,698	2,204	5,154	1,557	8,115
Eliminate Obs. with Missing Data on Characteristics of Students & Teachers	1,837	4,121	3,212	9,170	2,170	5,060	1,529	7,965

[1] Classes include self-contained classes in grades 4 and 5. Since teachers can switch between schools in different poverty categories, and since in some schools the percentage of FRL students is not available in all years, the sum of the teachers in school poverty categories do not add up to the total number of unique teachers independent of school poverty category.

[2] In North Carolina individual student test scores are linked to test proctors, who are not necessarily classroom instructors. We compare the student demographic characteristics of the test classrooms and the instructional classrooms and keep only those teachers who can be verified to be instructors. In Florida, there exist course offering codes that identify each unique classroom. Both students and teachers can be linked to specific course offerings.

Table 1B
Number of teachers in state and sub-samples, by state and school poverty category - Reading

	Florida				North Carolina			
	0-30% FRL	30-70% FRL	70-100% FRL	Total	0-30% FRL	30-70% FRL	70-100% FRL	Total
Teachers of Relevant Classes in State	2,934	6,592	5,788	15,314	3,574	9,645	4,149	15,487
Teachers Linked to Students	2,934	6,592	5,788	15,314	2,468	6,606	2,625	10,383
Eliminate Charter School Students	2,649	6,437	5,621	14,707	2,399	6,591	2,610	10,273
Eliminate Classes with <10 or >40 Students	2,575	6,212	5,323	14,110	2,260	6,124	2,386	9,496
Eliminate Students with Invalid Test Grade Level	2,575	6,209	5,315	14,099	2,260	6,124	2,386	9,496
Eliminate Students with Missing Gain Scores	2,273	5,579	4,822	12,674	2,238	6,042	2,347	9,367
Eliminate Students in Schools that Switch Poverty Status	2,209	4,937	3,869	11,006	2,201	5,142	1,554	8,098
Eliminate Obs. with Missing Data on Characteristics of Students &Teachers	1,879	4,260	3,257	9,396	2,167	5,058	1,526	7,957

[1] Classes include self-contained classes in grades 4 and 5. Since teachers can switch between schools in different poverty categories, and since in some schools the percentage of FRL students is not available in all years, the sum of the teachers in school poverty categories do not add up to the total number of unique teachers independent of school poverty category.

[2] In North Carolina individual student test scores are linked to test proctors, who are not necessarily classroom instructors. We compare the student demographic characteristics of the test classrooms and the instructional classrooms and keep only those teachers who can be verified to be instructors. In Florida, there exist course offering codes that identify each unique classroom. Both students and teachers can be linked to specific course offerings.

Table 2
Student, teacher and school characteristics in analytic sample, by state and school poverty category

Student, teacher and school characteristic	Florida 0-30% FRL	Florida 0-70% FRL	Florida 70-100% FRL	North Carolina 0-30% FRL	North Carolina 0-70% FRL	North Carolina 70-100% FRL
Students						
Male (%)	50.16 **	50.00 **	48.97	50.75	50.67	50.89
White (%)	74.29 **	64.15 **	16.85	79.02 **	70.22 **	18.81
Black (%)	9.30 **	14.08 **	47.13	13.20 **	20.49 **	64.22
Hispanic (%)	10.60 **	16.39 **	33.37	2.88 **	4.62 **	8.43
Other minorities (%)	5.81 **	5.37 **	2.65	4.90 **	4.67 **	8.54
Student performance (level scores)						
Math	0.44 **	0.26 **	-0.23	0.44 **	0.18 **	-0.39
Reading	0.45 **	0.28 **	-0.26	0.39 **	0.16 **	-0.41
Student performance (gain scores)						
Math	-0.04 **	-0.03 **	0.02	0.02	0.01	0.02
Reading	-0.02 **	-0.03	-0.03	0.02 **	0.01 **	-0.01
Teachers/Principals						
Experience (% of teachers)						
0 years	8.42 **	9.73 **	12.91	2.59 **	3.22 **	6.10
1-2 years	15.30 **	17.04 **	25.71	14.08 *	15.96	16.86
3-5 years	16.72	16.32	16.45	17.04	16.30	16.15
6-12 years	24.62 **	23.01 **	18.88	25.99	24.45	24.58
13-20 years	19.33 **	17.42 **	11.48	17.76 **	16.48 **	12.45
21-27 years	9.65 **	9.47 **	8.10	14.08	14.08	12.26
28 or more years	5.96	7.00	6.47	8.45 **	9.51 *	11.61
Graduate degree (%)	33.03 **	31.38 **	29.04	31.98	30.15	29.54
Regular license (%)	94.27 **	92.90 **	88.51	97.92 **	97.07 **	93.37
NBPTS certified (%)	3.99 **	3.38 **	1.67	13.29 **	11.17 **	6.18
Praxis score	—	—	—	0.21 **	0.15 **	-0.30
Principal's experience	11.43	11.31	11.06	—	—	—
New principal indicator (%)	12.22 **	14.82 **	18.94	—	—	—
Schools						
Poverty level (% FRPL)	19.71 **	40.35 **	85.41	20.41 **	41.10 **	86.01
LEP students (%)	2.48 **	3.18 **	7.83	—	—	—
Special education students (%)	14.51 *	14.15	12.93	—	—	—

— Not available

Note: All statistics are based on the math analysis sample, except for reading achievement test scores (which are based on the reading analysis sample). * denotes that the estimate is statistically different from the corresponding estimate for schools with 70-100% FRL students at the 5% level, and ** denotes significance at the 1% level.

Table 3
Means and standard deviations of teacher value-added, by state, school poverty category, subject and model

Subject and model	Florida 0-70% FRL	Florida 70-100% FRL	Florida Difference		North Carolina 0-70% FRL	North Carolina 70-100% FRL	North Carolina Difference	
Means								
Math								
Non-shrunken, SC, CP	0.0053	0.0186	-0.0133	**	0.0052	-0.0228	0.0280	**
	(0.0019)	(0.0035)	(0.0040)		(0.0028)	(0.0064)	(0.0069)	
Non-shrunken, SFE, CP	0.0089	0.0175	-0.0087		0.0105	-0.0462	0.0567	**
	(0.0041)	(0.0076)	(0.0086)		(0.0071)	(0.0165)	(0.0179)	
Non-shrunken, SC, PP	0.0214	-0.0038	0.0252	**	0.0099	-0.0438	0.0537	**
	(0.0019)	(0.0035)	(0.0040)		(0.0028)	(0.0063)	(0.0069)	
Non-shrunken, SFE, PP	0.0010	0.0189	-0.0179	**	0.0028	-0.0122	0.0150	
	(0.0022)	(0.0042)	(0.0047)		(0.004)	(0.0092)	(0.0100)	
Shrunken, SC, CP	0.0067	0.0185	-0.0119	**	0.0057	-0.0154	0.0211	**
	(0.0015)	(0.0026)	(0.0030)		(0.0023)	(0.005)	(0.0055)	
Shrunken, SFE, CP	0.0054	0.0099	-0.0045		0.0024	-0.0096	0.0120	*
	(0.0015)	(0.0027)	(0.0031)		(0.0027)	(0.0055)	(0.0061)	
Shrunken, SC, PP	0.0205	0.0018	0.0186	**	0.0099	-0.0331	0.0430	**
	(0.0015)	(0.0027)	(0.0031)		(0.0024)	(0.0051)	(0.0057)	
Shrunken, SFE, PP	0.0012	0.0127	-0.0115	**	0.0006	-0.0013	0.0019	
	(0.0011)	(0.0020)	(0.0023)		(0.0017)	(0.0037)	(0.0040)	
Reading								
Non-shrunken, SC, CP	0.0101	-0.0073	0.0174	**	0.0058	-0.0257	0.0315	**
	(0.0014)	(0.0027)	(0.0031)		(0.0020)	(0.0053)	(0.0056)	
Non-shrunken, SFE, CP	0.0014	0.0106	-0.0092		0.0097	-0.0431	0.0528	**
	(0.0041)	(0.0067)	(0.0078)		(0.0072)	(0.0172)	(0.0187)	
Non-shrunken, SC, PP	0.0288	-0.0353	0.0641	**	0.0134	-0.0593	0.0727	**
	(0.0015)	(0.0027)	(0.0031)		(0.0020)	(0.0051)	(0.0055)	
Non-shrunken, SFE, PP	0.0076	-0.0058	0.0134	*	0.0027	-0.0118	0.0145	
	(0.0023)	(0.0036)	(0.0043)		(0.0042)	(0.0107)	(0.0115)	
Shrunken, SC, CP	0.0080	-0.0012	0.0092	**	0.0054	-0.0120	0.0174	**
	(0.0008)	(0.0014)	(0.0016)		(0.0012)	(0.0027)	(0.0029)	
Shrunken, SFE, CP	-0.0000	0.0059	-0.0059	**	0.0021	-0.0066	0.0087	**
	(0.0010)	(0.0017)	(0.0019)		(0.0010)	(0.0022)	(0.0025)	
Shrunken, SC, PP	0.0225	-0.0184	0.0409	**	0.0111	-0.0333	0.0444	**
	(0.0009)	(0.0017)	(0.0019)		(0.0013)	(0.0030)	(0.0033)	
Shrunken, SFE, PP	0.0032	-0.0002	0.0034	**	0.0031	-0.0062	0.0093	**
	(0.0005)	(0.0008)	(0.0009)		(0.0012)	(0.0027)	(0.0030)	
Standard deviations								
Math								
Non-shrunken, SC, CP	0.1812	0.2366	-0.0554	**	0.2026	0.2221	-0.0195	**
Non-shrunken, SFE, CP	0.3913	0.5067	-0.1154	**	0.5196	0.5762	-0.0566	**
Non-shrunken, SC, PP	0.1807	0.2346	-0.0539	**	0.2029	0.2195	-0.0166	**
Non-shrunken, SFE, PP	0.2134	0.2789	-0.0655	**	0.2949	0.3212	-0.0263	**
Shrunken, SC, CP	0.1386	0.1732	-0.0346	**	0.1683	0.1760	-0.0077	*
Shrunken, SFE, CP	0.1430	0.1785	-0.0355	**	0.1966	0.1911	0.0055	
Shrunken, SC, PP	0.1456	0.1829	-0.0373	**	0.1735	0.1797	-0.0062	
Shrunken, SFE, PP	0.1005	0.1327	-0.0322	**	0.1259	0.1283	-0.0024	
Reading								
Non-shrunken, SC, CP	0.1380	0.1819	-0.0439	**	0.1438	0.1836	-0.0398	**
Non-shrunken, SFE, CP	0.3902	0.4494	-0.0592	**	0.5289	0.6008	-0.0719	**
Non-shrunken, SC, PP	0.1405	0.1842	-0.0437	**	0.1455	0.1785	-0.0330	**
Non-shrunken, SFE, PP	0.2171	0.2426	-0.0255	**	0.3077	0.3747	-0.0670	**
Shrunken, SC, CP	0.0748	0.0909	-0.0161	**	0.0848	0.0929	-0.0081	**
Shrunken, SFE, CP	0.0937	0.1127	-0.0190	**	0.0769	0.0783	-0.0014	
Shrunken, SC, PP	0.0911	0.1118	-0.0207	**	0.0969	0.1039	-0.0070	**
Shrunken, SFE, PP	0.0449	0.0530	-0.0081	**	0.0918	0.0938	-0.0020	

Note: * denotes that the difference is statistically significant at the 5% level, and ** denotes significance at the 1% level.

Table 4A—Florida
Difference in Mean Value Added for FRL and Non-FRL Students, by state, school poverty category, subject and model

	0-70% FRL Schools				70-100% FRL Schools			
	Non-FRL Students	FRL Students	Difference		Non-FRL Students	FRL Students	Difference	
Math								
Non-shrunken, SC, CP	0.0172	0.0158	0.0013	**	0.0387	0.0436	-0.0049	**
	(0.1656)	(0.1746)	(0.0005)		(0.2101)	(0.2204)	(0.0013)	
Non-shrunken, SFE, CP	0.0261	0.0229	0.0031	**	0.0383	0.0471	-0.0088	**
	(0.3421)	(0.3517)	(0.0010)		(0.4238)	(0.4487)	(0.0026)	
Non-shrunken, SC, PP	0.0439	0.0242	0.0197	**	0.0349	0.0240	0.0109	**
	(0.1648)	(0.1746)	(0.0005)		(0.2119)	(0.2235)	(0.0013)	
Non-shrunken, SFE, PP	0.0118	0.0085	0.0032	**	0.0346	0.0460	-0.0113	**
	(0.1878)	(0.1951)	(0.0005)		(0.2385)	(0.2523)	(0.0014)	
Shrunken, SC, CP	0.0172	0.0158	0.0013	**	0.0387	0.0436	-0.0048	**
	(0.1656)	(0.1746)	(0.0005)		(0.2101)	(0.2203)	(0.0013)	
Shrunken, SFE, CP	0.0260	0.0229	0.0031	**	0.0383	0.0471	-0.0088	**
	(0.3421)	(0.3517)	(0.0010)		(0.4238)	(0.4487)	(0.0026)	
Shrunken, SC, PP	0.0439	0.0242	0.0197	**	0.0349	0.0240	0.0109	**
	(0.1648)	(0.1746)	(0.0005)		(0.2119)	(0.2235)	(0.0013)	
Shrunken, SFE, PP	0.0118	0.0085	0.0032	**	0..0346	0.0460	-0.0113	**
	(0.1878)	(0.1951)	(0.0005)		(0.2385)	(0.2523)	(0.0014)	
Reading								
Non-shrunken, SC, CP	0.0202	0.0148	0.0053	**	0.0045	0.0035	0.0010	
	(0.1219)	(0.1242)	(0.0003)		(0.1559)	(0.1661)	(0.0009)	
Non-shrunken, SFE, CP	0.0053	0.0143	-0.0090	**	0.0201	0.0338	-0.0137	**
	(0.3377)	(0.3396)	(0.0009)		(0.3813)	(0.3959)	(0.0023)	
Non-shrunken, SC, PP	0.0482	0.0268	0.0213	**	-0.0018	0.-0.0229	0.0210	**
	(0.1250)	(0.1281)	(0.0003)		(0.1586)	(0.1723)	(0.0009)	
Non-shrunken, SFE, PP	0.0139	0.0099	0.0039	**	0.0156	0.0054	0.0102	**
	(0.1880)	(0.1971)	(0.0005)		(0.1994)	(0.2115)	(0.0012)	
Shrunken, SC, CP	0.0202	0.0148	0.0053	**	0.0045	0.0035	0.0010	
	(0.1219)	(0.1242)	(0.0003)		(0.1559)	(0.1661)	(0.0009)	
Shrunken, SFE, CP	0.0053	0.0143	-0.0090	**	0.0201	0.0338	-0.0137	**
	(0.3377)	(0.3396)	(0.0009)		(0.3813)	(0.3959)	(0.0023)	
Shrunken, SC, PP	0.0482	0.0268	0.0213	**	-0.0018	-0.0229	0.0210	**
	(0.1250)	(0.1281)	(0.0003)		(0.1586)	(0.1723)	(0.0009)	
Shrunken, SFE, PP	0.0139	0.0099	0.0039	**	0.0156	0.0054	102.0000	**
	(0.1880)	(0.1971)	(0.0005)		(0.1994)	(0.2115)	(0.0012)	

Note: * denotes that the difference is statistically significant at the 5% level, and ** denotes significance at the 1% level.

Table 4B—North Carolina
Difference in Mean Value Added for FRL and Non-FRL Students,
by state, school poverty category, subject and model

	0-70% FRL Schools				70-100% FRL Schools			
	Non-FRL Students	FRL Students	Difference		Non-FRL Students	FRL Students	Difference	
Math								
Non-shrunken, SC, CP	0.0221	0.0049	0.0172	**	0.0044	-0.0036	0.0080	**
	(0.1886)	(0.1951)	(0.0006)		(0.2085)	(0.2184)	(0.0021)	
Non-shrunken, SFE, CP	0.0296	0.0220	0.0076	**	-0.0088	-0.0253	0.0165	**
	(0.4784)	(0.4815)	(0.0016)		(0.5665)	(0.5599)	(0.0056)	
Non-shrunken, SC, PP	0.0330	0.0026	0.0304	**	-0.0100	-0.0216	0.0116	**
	(0.1905)	(0.1943)	(0.0006)		(0.2081)	(0.2171)	(0.0021)	
Non-shrunken, SFE, PP	0.0169	0.0081	0.0088	**	0.0127	0.0035	0.0092	**
	(0.2719)	(0.2673)	(0.0009)		(0.3142)	(0.3158)	(0.0031)	
Shrunken, SC, CP	0.0201	0.0052	0.0149	**	0.0060	-0.0007	0.0067	**
	(0.1626)	(0.1670)	(0.0005)		(0.1758)	(0.1821)	(0.0018)	
Shrunken, SFE, CP	0.0130	0.0083	0.0047	**	0.0014	-0.0020	0.0034	
	(0.1967)	(0.1964)	(0.0007)		(0.1834)	(0.1916)	(0.0018)	
Shrunken, SC, PP	0.0302	0.0032	0.0270	**	-0.0064	-0.0162	0.0098	**
	(0.1683)	(0.1705)	(0.0006)		(0.1799)	(0.1858)	(0.0018)	
Shrunken, SFE, PP	0.0079	0.0040	0.0039	**	0.0070	0.0040	0.0030	*
	(0.1269)	(0.1255)	(0.0004)		(0.1316)	(0.1323)	(0.0013)	
Reading								
Non-shrunken, SC, CP	0.0169	0.0059	0.0110	**	-0.0119	-0.0163	0.0044	**
	(0.1307)	(0.1361)	(0.0004)		(0.1609)	(0.1721)	(0.0016)	
Non-shrunken, SFE, CP	0.0246	0.0235	0.0011		-0.0244	-0.0192	-0.0052	
	(0.4837)	(0.4930)	(0.0016)		(0.5778)	(0.5729)	(0.0057)	
Non-shrunken, SC, PP	0.0307	0.0070	0.0237	**	-0.0335	-0.0450	0.0115	**
	(0.1351)	(0.1370)	(0.0005)		(0.1623)	(0.1712)	(0.0016)	
Non-shrunken, SFE, PP	0.0124	0.0147	-0.0023	*	0.0106	0.0132	-0.0026	
	(0.2842)	(0.2869)	(0.0009)		(0.3687)	(0.3748)	(0.0037)	
Shrunken, SC, CP	0.0123	0.0051	0.0072	**	-0.0048	-0.0075	0.0027	**
	(0.0844)	(0.0859)	(0.0003)		(0.0929)	(0.0969)	(0.0009)	
Shrunken, SFE, CP	0.0055	0.0036	0.0019	**	-0.0032	-0.0037	0.0005	
	(0.0795)	(0.0811)	(0.0003)		(0.0754)	(0.0782)	(0.0008)	
Shrunken, SC, PP	0.0235	0.0063	0.0172	**	-0.0197	-0.0271	0.0074	**
	(0.0968)	(0.0964)	(0.0003)		(0.1060)	(0.1094)	(0.0011)	
Shrunken, SFE, PP	0.0064	0.0065	-0.0001		-0.0010	-0.0005	-0.0005	
	(0.0926)	(0.0956)	(0.0003)		(0.0933)	(0.0970)	(0.0009)	

Note: * denotes that the difference is statistically significant at the 5% level, and ** denotes significance at the 1% level.

Table 5
Differences in estimated teacher value-added for teachers who taught in both 0-70% FRL schools and 70-100% FRL schools, by state, subject and model

	Florida			North Carolina		
Subject and model	*Mean differences (high poverty - not high poverty)*	*Standard error*		*Mean differences (high poverty - not high poverty)*	*Standard error*	
<u>*Math*</u>						
Non-shrunken, SC, CP	-0.0056	0.0179		-0.0262	0.0273	
Non-shrunken, SFE, CP	0.0367	0.0436		-0.0228	0.0765	
Non-shrunken, SC, PP	-0.0249	0.0168		0.0008	0.0261	
Non-shrunken, SFE, PP	0.0607	0.0283	*	-0.0212	0.0485	
Shrunken, SC, CP	-0.0034	0.0123		-0.0225	0.0204	
Shrunken, SFE, CP	0.0063	0.0135		-0.0200	0.0309	
Shrunken, SC, PP	-0.0180	0.0123		-0.0012	0.0206	
Shrunken, SFE, PP	0.0159	0.0097		-0.0111	0.0217	
<u>*Reading*</u>						
Non-shrunken, SC, CP	-0.0320	0.0138	*	0.0209	0.0256	
Non-shrunken, SFE, CP	-0.0143	0.0410		0.1268	0.0878	
Non-shrunken, SC, PP	-0.0501	0.0133	**	0.0530	0.0254	*
Non-shrunken, SFE, PP	0.0128	0.0231		0.0558	0.0475	
Shrunken, SC, CP	-0.0111	0.0061		0.0111	0.0121	
Shrunken, SFE, CP	0.0019	0.0091		0.0193	0.0116	
Shrunken, SC, PP	-0.0251	0.0073	**	0.0312	0.0144	*
Shrunken, SFE, PP	-0.0006	0.0042		0.0214	0.0165	

Note: Differences between high and low-poverty teacher effects (t-test) are * significant at 5%, ** significant at 1%

Table 6
Teacher value-added at various percentiles,
by state, subject and school poverty category
(Shrunken Partial Persistence, Student Covariates Model)

	Florida				North Carolina			
Subject and teacher performance percentile	0-70% FRL	70-100% FRL	Difference		0-70% FRL	70-100% FRL	Difference	
Math								
10	-0.1582	-0.2225	0.0643	**	-0.2001	-0.2365	0.0364	**
25	-0.0748	-0.1163	0.0415	**	-0.1066	-0.1265	0.0199	**
50	0.0195	-0.0033	0.0228	**	0.0000	-0.0261	0.0261	**
75	0.1136	0.1176	-0.0040		0.1146	0.0991	0.0155	
90	0.2087	0.2299	-0.0212	**	0.2185	0.2045	0.0140	
Reading								
10	-0.0899	-0.1516	0.0617	**	-0.1022	-0.1277	0.0254	**
25	-0.0377	-0.0886	0.0509	**	-0.0510	-0.0699	0.0190	**
50	0.0230	-0.0191	0.0421	**	0.0058	-0.0122	0.0180	**
75	0.0820	0.0466	0.0354	**	0.0612	0.0446	0.0166	**
90	0.1352	0.1161	0.0191	**	0.1124	0.0975	0.0149	**

Note: * denotes statistical significance at the 5% level, and ** denotes statistical significance at the 1% level

Figure 1A
Distribution of teacher value-added for high and lower-poverty schools, math teachers in Florida

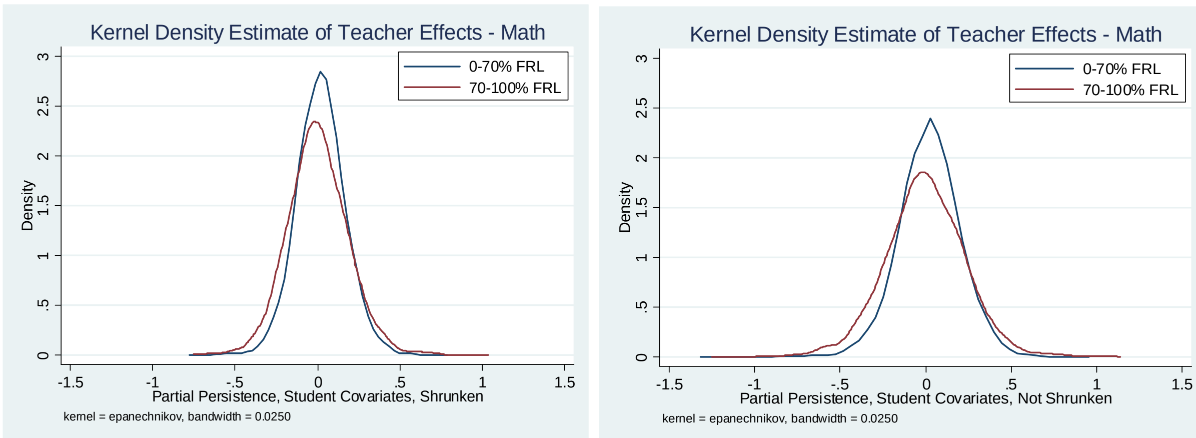

Figure 1B
Distribution of Teacher value-added for high and lower-poverty schools, reading teachers in Florida

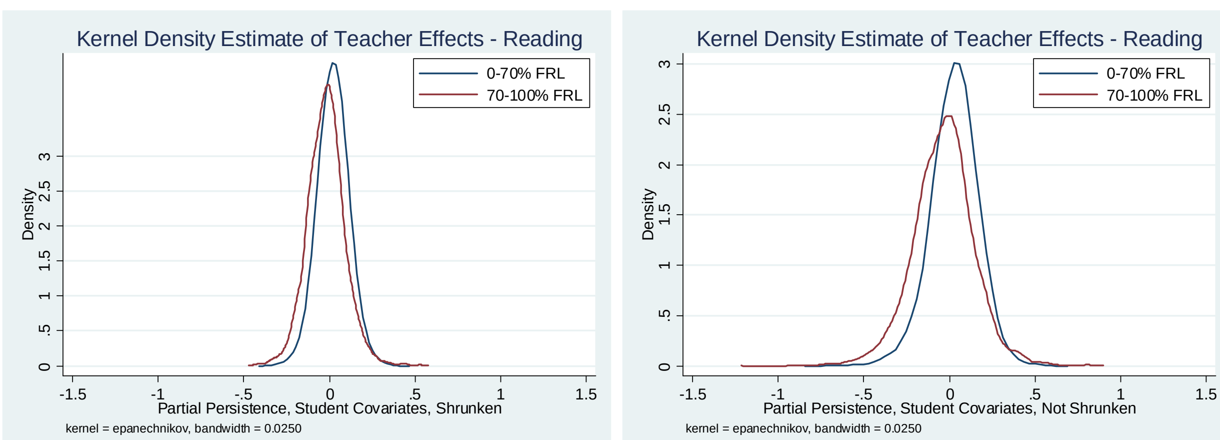

Figure 2A
Distribution of teacher value-added for high and lower-poverty schools, math teachers in North Carolina

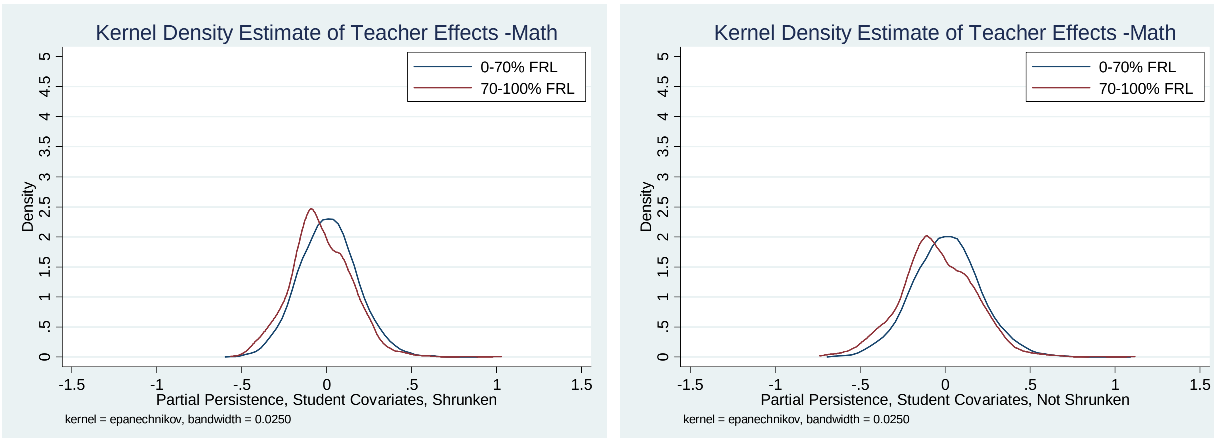

Figure 2B
Distribution of teacher value-added high and lower-poverty schools, reading teachers in North Carolina

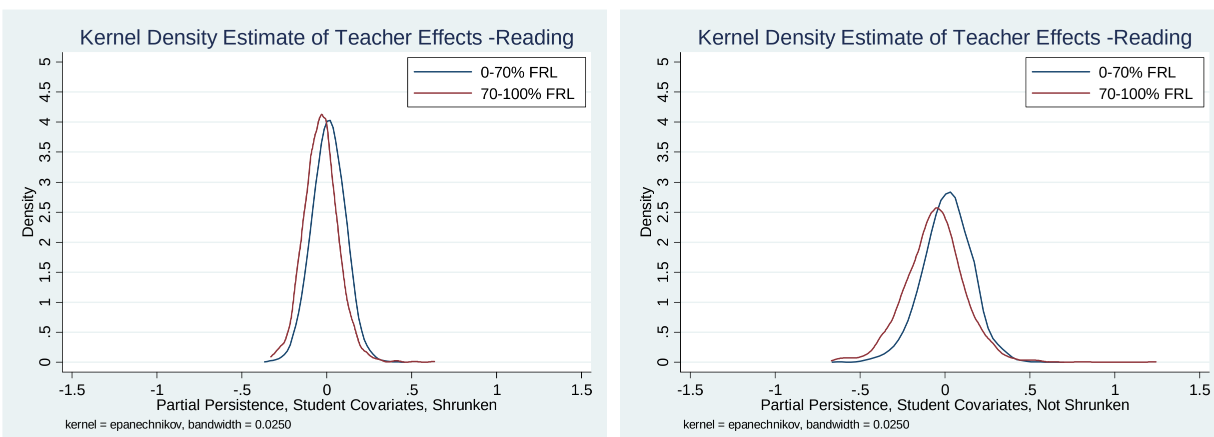

Table 7
Teacher value-added at various percentiles, by state, subject and school poverty category (teachers with 2 or fewer years of experience)
(Shrunken Partial Persistence, Student Covariates Model)

Subject and teacher performance percentile	Florida				North Carolina		
	0-70% FRL	*70-100% FRL*	Difference		*0-70% FRL*	*70-100% FRL*	Difference
Math							
10	-0.1625	-0.2213	0.0588	**	-0.2557	-0.3108	0.0551
25	-0.0821	-0.1200	0.0379	**	-0.1506	-0.2257	0.0752
50	0.0095	-0.0130	0.0225	**	-0.0160	-0.0735	0.0575
75	0.1037	0.1096	-0.0059		0.1893	0.1284	0.0609
90	0.1937	0.2136	-0.0199	**	0.3394	0.3553	-0.0159
Reading							
10	-0.0896	-0.1584	0.0688	**	-0.1243	-0.1536	0.0293
25	-0.0391	-0.0943	0.0552	**	-0.0717	-0.0991	0.0274
50	0.0159	-0.0278	0.0437	**	0.0032	-0.0266	0.0298
75	0.0725	0.0322	0.0403	**	0.0591	0.0720	-0.0128
90	0.1260	0.1042	0.0218	**	0.1469	0.1529	-0.0060

Note: * denotes statistical significance at the 5% level, ** denotes statistical significance at the 1% level

Figure 3A
Distribution of teacher value-added for high and lower-poverty schools, math teachers in Florida with 0-2 years of experience

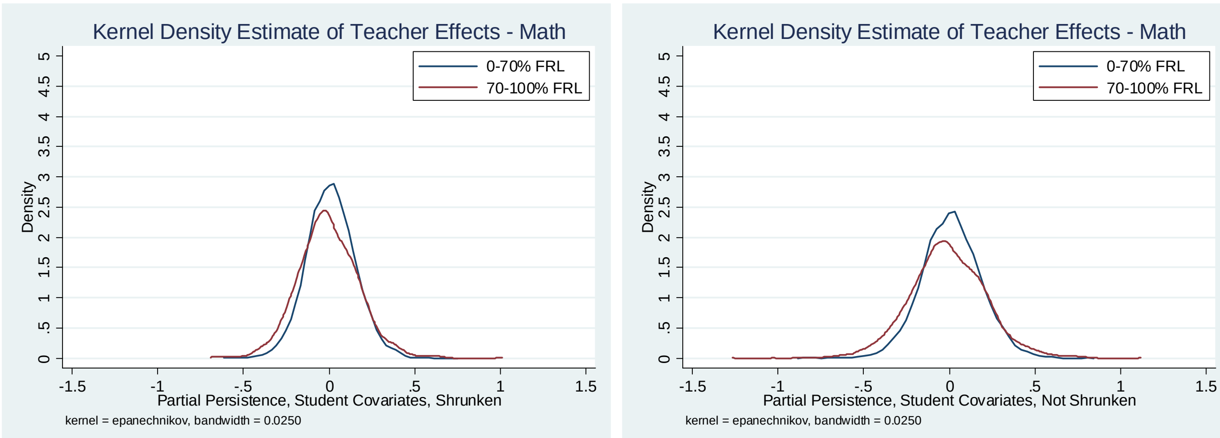

Figure 3B
Distribution of Teacher value-added for high and lower-poverty schools, reading teachers in Florida with 0-2 years of experience

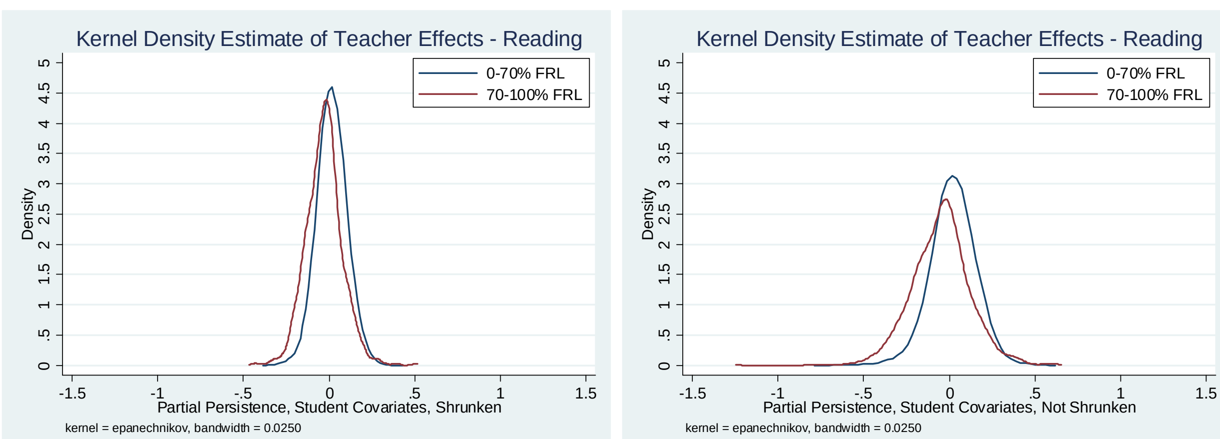

Figure 4A
Distribution of teacher value-added for high and lower-poverty schools, math teachers in North Carolina with 0-2 years of experience

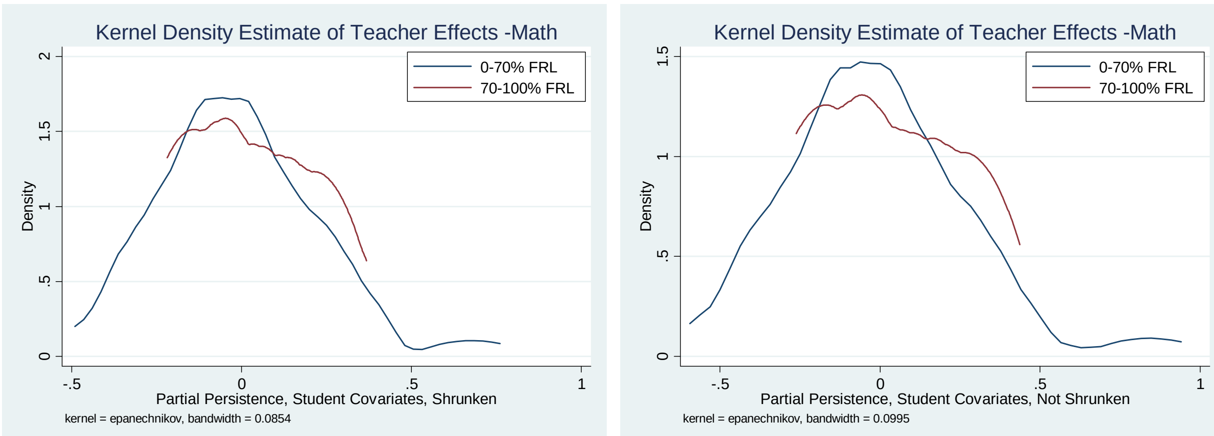

Figure 4B
Distribution of Teacher value-added for high and lower-poverty schools, reading teachers in North Carolina with 0-2 years of experience

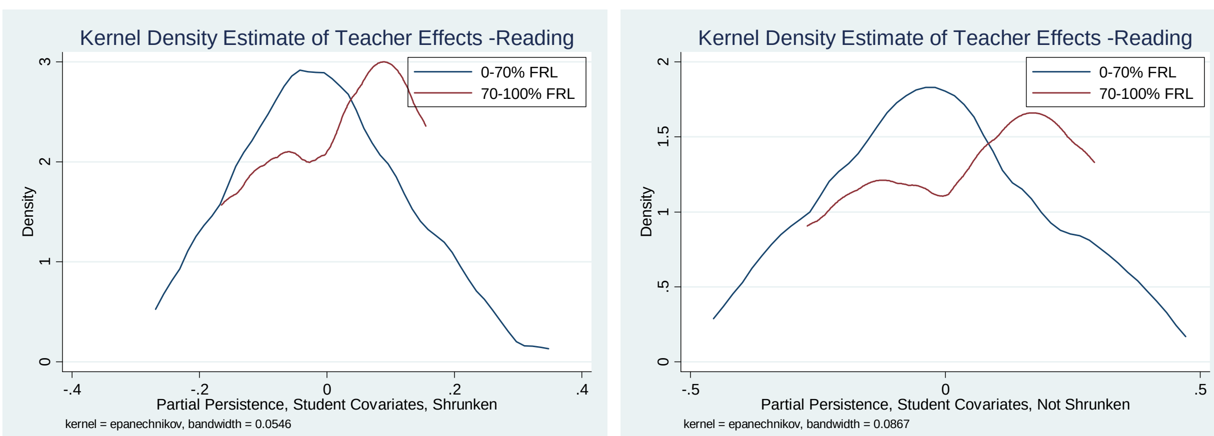

Table 8
Blinder-Oaxaca Decomposition of Differences in Estimated Teacher Value-Added by State, Subject and School Poverty Category
(Partial Persistence Model with Student Covariates -- not shrunken, GLS)

		Florida		North Carolina	
Subject, predicted value-added, overall difference and sources of difference		0-70% FRL	70-100% FRL	0-70% FRL	70-100% FRL
Math					
Mean predicted value-added		0.0283	0.0063	0.0113	-0.0410
Overall Difference		0.0220**		0.0518**	
Sources of difference					
Characteristics	- total	0.0056**		0.0036*	
	- experience	0.0022		0.0005	
	- advanced degree	0.0000		0.0000	
	- professional cert.	0.0033**		0.0031**	
Marginal effect of characteristics	- total	0.0165***		0.0469**	
	- experience	0.0015		0.0115	
	- advanced degree	0.0044		0.0010	
	- professional cert.	-0.0125		-0.0220	
	- constant	0.0231		0.0565	
Interaction of characteristics and marginal effects		-0.0001		0.0013	
Reading					
Mean predicted value-added		0.0345	-0.0284	0.0156	-0.0550
Overall Difference		0.0629**		0.0707**	
Sources of difference					
Characteristics - total		0.0055**		0.0026	
	- experience	0.0026		0.0011	
	- advanced degree	0.0001		0.0000	
	- professional cert.	0.0029**		0.0015	
Marginal effect of characteristics	- total	0.0557**		0.0659**	
	- experience	-0.0041		0.0077	
	- advanced degree	0.0008		0.0008	
	- professional cert.	-0.0130		0.0157	
	- constant	0.0719**		0.0416	
Interaction of characteristics and marginal effects		0.0017		0.0022	

Note: Difference is * significant at 5%, ** significant at 1%

Table 9
FGLS estimates of the determinants of teacher value-added by state, subject and school poverty category

Teacher Experience and Certification	Florida				North Carolina			
	0-70% FRL		70-100% FRL		0-70% FRL		70-100% FRL	
Math								
3-5 Years	0.0066		0.0156		0.0377	**	0.0445	*
	(0.0071)		(0.0126)		(0.0095)		(0.0204)	
6-12 Years	0.0234	**	0.0113		0.0571	**	0.0311	
	(0.0061)		(0.0113)		(0.0085)		(0.0186)	
13-20 Years	0.0221	**	0.0198		0.0513	**	0.0225	
	(0.0064)		(0.0135)		(0.0093)		(0.0224)	
21-27 Years	-0.0012		0.0100		0.0576	**	0.0334	
	(0.0078)		(0.0150)		(0.0097)		(0.0222)	
28-Plus Years	-0.0012		0.0007		0.0462	**	0.0488	*
	(0.0086)		(0.0158)		(0.0112)		(0.0231)	*
Advanced Degree	0.0142	**	-0.0002		-0.0060		-0.0094	
	(0.0041)		(0.0083)		(0.0061)		(0.0140)	
Certification	0.0449	**	0.0590	**	0.0675	**	0.0908	**
	(0.0098)		(0.0134)		(0.0210)		(0.0292)	
Reading								
3-5 Years	0.0154	**	0.0351	**	0.0326	**	0.0368	*
	(0.0054)		(0.0099)		(0.0067)		(0.0162)	
6-12 Years	0.0287	**	0.0198	*	0.0540	**	0.0212	
	(0.0046)		(0.0088)		(0.0060)		(0.0148)	
13-20 Years	0.0393	**	0.0162		0.0558	**	0.0332	
	(0.0048)		(0.0105)		(0.0065)		(0.0177)	
21-27 Years	0.0369	**	0.0206		0.0581	**	0.0603	**
	(0.0058)		(0.0116)		(0.0068)		(0.0175)	
28-Plus Years	0.0335	**	0.0177		0.0518	**	0.0689	**
	(0.0064)		(0.0123)		(0.0078)		(0.0183)	
Advanced Degree	0.0053		0.0025		-0.0109	*	-0.0135	
	(0.0031)		(0.0064)		(0.0043)		(0.0111)	
Certification	0.0374	**	0.0521	**	0.0596	**	0.0430	
	(0.0076)		(0.0107)		(0.0156)		(0.0237)	

Table 10
Mean teacher value-added for teachers with 0-2 years of experience by teacher mobility, state, subject and school poverty category

Subject and teacher mobility	Florida *0-70% FRL*	Florida *70-100% FRL*	North Carolina *0-70% FRL*	North Carolina *70-100% FRL*
Math				
Leaves at the end of 1st year	-0.0798	-0.0270	-0.1194	-0.1190
Stays beyond 1st year	-0.0971	-0.1219	-0.0862	-0.0377
Difference	**0.0172**	**0.0949** **	**-0.0332**	**-0.0814**
Leaves at the end of 2nd yr. (year 1 efficiency)	0.1269	-0.0437	-0.0799	-0.1007
Stays beyond 2nd year (year 1 efficiency)	-0.1382	0.0610	-0.0783	-0.0450
Difference	**0.2651** **	**-0.1047**	**-0.0016**	**-0.0557**
Leaves at the end of 2nd yr. (year 2 efficiency)	0.0935	0.0937	-0.0658	-0.1063
Stays beyond 2nd year (year 2 efficiency)	-0.0287	0.0892	-0.0250	-0.0420
Difference	**0.1222** *	**0.0045**	**-0.0408**	**-0.0643**
Reading				
Leaves at the end of 1st year	0.0852	-0.1041	-0.0489	-0.0632
Stays beyond 1st year	0.1300	-0.0387	-0.0362	-0.0175
Difference	**-0.0447**	**-0.0876** **	**-0.0127**	**-0.0456**
Leaves at the end of 2nd yr. (year 1 efficiency)	0.0790	-0.0647	-0.0669	-0.0309
Stays beyond 2nd year (year 1 efficiency)	0.0476	-0.1699	-0.0302	-0.0322
Difference	**0.0314**	**0.1053** *	**-0.0367**	**0.0013**
Leaves at the end of 2nd yr. (year 2 efficiency)	0.1246	-0.0628	-0.0156	-0.0059
Stays beyond 2nd year (year 2 efficiency)	0.1004	-0.0405	-0.0137	-0.0293
Difference	**0.0242**	**-0.0224**	**-0.0018**	**0.0234**

Note: * denotes statistical significance at the 5% level, ** denotes statistical significance at the 1% level.

Table 11
Mean teacher value-added for those who taught in both high and low-poverty settings, by state, subject and the direction of switch

Subject and direction of school change	Florida					North Carolina			
	N	Pre-switch performance	Post-switch performance	Difference		N	Pre-switch performance	Post-switch performance	Difference
Math									
Switch from high to low-poverty schools	149	0.0026	0.0367	0.0341	*	38	0.0022	-0.0270	-0.0292
Switch from low to high-poverty schools	40	-0.0222	-0.0067	0.0155		20	-0.0395	-0.0363	0.0032
Reading									
Switch from high to low-poverty schools	156	-0.0081	0.0218	0.0298	**	38	-0.0167	0.0023	0.0190
Switch from low to high-poverty schools	41	-0.0233	-0.0193	-0.0040		20	-0.0291	-0.0342	-0.0051

Note: * denotes statistical significance at the 5% level, ** denotes statistical significance at the 1% level