# TEACHER TRAINING, TEACHER QUALITY AND STUDENT ACHIEVEMENT*

by

Douglas N. Harris
Department of Educational Policy Studies
University of Wisconsin – Madison
213 Education Building
Madison, WI 53706
Email: dnharris3@wisc.edu

Tim R. Sass
Department of Economics
Florida State University
288 Bellamy Building
Tallahassee, FL 32306
Email: tsass@fsu.edu

Original version: March 9, 2006
This version: February 5, 2007

*Currently under peer review*.

**Abstract**

We study the effects of various types of education and training on the ability of teachers to promote student achievement. Previous studies on the subject have been hampered by inadequate measures of teacher training and difficulties addressing the non-random selection of teachers to students and of teachers to training. We address all of these limitations by estimating models with student, teacher, and school fixed effects using an extensive database from the state of Florida. Our results suggest that teacher training generally has little influence on productivity. One exception is that content-focused teacher professional development is positively associated with productivity in middle and high school math. In addition, more experienced teachers appear more effective in teaching elementary and middle school reading. There is no evidence that either pre-service (undergraduate) training or the scholastic aptitude of teachers influences their ability to increase student achievement. These results call into question previous findings based on models that do not adequately control for the various forms of selection bias.

*We wish to thank the staff of the Florida Department of Education's K-20 Education Data Warehouse for their assistance in obtaining and interpreting the data used in this study. The views expressed is this paper are solely our own and do not necessarily reflect the opinions of the Florida Department of Education. This research is part of a larger project assessing teacher quality being funded by grant R305M040121 from the U.S. Department of Education.

# I. Introduction

It is generally acknowledged that promoting teacher quality is a key element in improving primary and secondary education in the United States. Indeed, one of the primary goals of the *No Child Left Behind* law is to have a "highly qualified teacher" in every classroom. Despite decades of research, however, there is no consensus on what factors enhance teacher quality.

We focus here on the relationship between teacher productivity and teacher training, including formal pre-service university education, in-service professional development, and informal training acquired through on-the-job experience. Previous research on teacher training has yielded highly inconsistent results and has fueled a wide range of policy prescriptions. Some studies find that formal education is important and these have been interpreted as support for strengthening existing teacher preparation programs in universities[1] and increased expenditures on post-college training. Equally common, however, is the finding that formal education is irrelevant, leading others to argue for the elimination of colleges of education.[2]

One reason for the uncertainty regarding the effects of teacher training is that all past studies have suffered from one of three methodological problems. First, it is difficult to measure productivity, especially in teaching where a student's own ability, the influences of a student's peers, and other characteristics of schools also affect measured outcomes. The problem is exacerbated by the fact that assignment of students and teachers to classrooms is usually not

---

[1] See, for example, the American Federation of Teachers' publication, "Where we Stand: Teacher Quality" at http://www.aft.org/pubs-reports/downloads/teachers/TQres.pdf

[2] See George Will's commentary, "Ed Schools vs. Education," *Newsweek*, January 16, 2006. Will argues "The surest, quickest way to add quality to primary and secondary education would be addition by subtraction: Close all the schools of education."

random, leading to possible correlations between observed teacher attributes and unobserved student characteristics. Second, it is difficult to obtain data that link the education and training of teachers to the achievement of the students they teach. Third, like in other occupations, there is an inherent selection problem in evaluating the effects of education and training on teacher productivity. Unobserved characteristics, such as "innate" ability, may affect the amount and types of education and training as well as subsequent performance of teachers in the classroom. Addressing all of these issues in a single study presents significant data and estimation challenges.

In this paper we present new evidence on the effects of teacher pre-service formal education and in-service professional development training on teacher productivity using a unique statewide administrative database from Florida. The Florida data allow us to tie student performance to the identity of their classroom teacher and in turn link teachers to their in-service training, their college coursework and their pre-college entrance exam scores. These extremely rich data provide a unique opportunity to address the selection problem associated with the acquisition of education and training by teachers and generate reliable estimates of the influence of different types of teacher education and training on teacher productivity in the classroom.

Our analysis proceeds in two steps. First, we estimate student achievement models that include a rich set of covariates that measure time-varying student and classroom peer characteristics as well as fixed effects that control for unmeasured time-invariant student, teacher and school characteristics. This first-stage model includes detailed data on the quantity and characteristics of education and training teachers receive after they have entered the classroom, including both graduate education and workshops sponsored by schools and school districts (called "in-service" or professional development training). We also include measures of teacher experience, which represents informal on-the-job training. This first step yields estimates of the fixed effect for each teacher, which represents the teacher's contribution to student achievement or "value added" that does not vary over her career.[3] In the second step we take the estimated teacher fixed effect and regress it on characteristics of the teacher's undergraduate coursework. We control for teacher pre-college cognitive/verbal ability with college entrance exam scores.

We begin in section II by describing past literature on teacher training, including both studies that analyze student achievement gains over two periods and studies using panel data or data from various types of experiments. Our methodology and data are discussed in sections III and IV, respectively. Our results, presented in section V, suggest that previous research that does not account for unmeasured student and teacher characteristics has produced invalid results and conclusions regarding teacher training. The implications of our findings are discussed in section VI.

## II. Previous Literature on the Effects of Teacher Training

In early work on teacher productivity, researchers estimated education production functions by regressing aggregate student achievement levels on measures of teacher training and various other controls using cross-sectional data (see review by Hanushek (1986)). A subsequent generation of studies used student-level two-year test-score gains and richer sets of teacher training variables. In the last several years this approach has been taken a step further by

[3] The term "value-added" has two rather different meanings in the education literature. Sometimes if refers to education production function models where the dependent variable is the gain in student achievement or student learning. The second meaning, which we use here, is simply the teacher's marginal product with respect to student achievement.

estimating achievement gain models with student-level panel data that use student fixed effects or take advantage of random assignment to control for unobserved student heterogeneity.

### *A. Gain Score Studies*

Over a dozen studies have estimated models of the relationship between teacher education/training and student achievement using data on the change in student test scores between two points in time.[4] The studies are summarized in Table 1 and most have been extensively reviewed in Wayne and Youngs (2003) and elsewhere.[5] The primary disadvantage of the gain score studies is their inability to fully control for the characteristics of students who are assigned to a given teacher. As noted in column (c) of Table 1, these studies rely on observed student characteristics or "covariates" to account for student heterogeneity. However, they cannot control for unobserved characteristics like innate ability and motivation. There is evidence that better trained and more experienced teachers tend to get students of greater ability and with fewer discipline problems (e.g., Clotfelter et al. (2005), Feng (2005)). Given this positive matching between student quality and teacher training, the gain score studies' failure to control for unobserved student characteristics would tend to upwardly bias estimates of teacher value-added associated with education and training.

The gain score studies include a variety of covariates that measure potentially important teacher characteristics. All studies include basic information about teacher race and gender and

---

[4] With the exception of Angrist and Lavy (2001) below, all of the studies reviewed are based on U.S. data.

[5] See other reviews by Rice (2003), Wilson and Floden (2003), and Wilson, Floden and Ferrini-Mundy (2001). There is substantial overlap in the studies reviewed in these relative to Wayne and Youngs (2003).

several include teachers' scores on certification tests and college entrance exams.[6] The teacher characteristic that is arguably most frequently associated with teacher effectiveness in these studies is the score on tests related to verbal and cognitive skills, such as some college entrance exams. As shown in Table 1, five of the early studies include such a variable (Ehrenberg and Brewer (1995), Ferguson and Ladd (1996), Hanushek (1992), Murnane and Phillips (1981), Rowan, et al. (1997)) and only one of these has such a score for teachers taken before they entered college (Ferguson and Ladd (1996)).[7] In this last study, the authors measure the pre-college ability of teachers by the school-level average composite score on the American College Test (ACT). Holding constant average teacher experience and the proportion of teachers with a master's degree, they find that average ACT scores of teachers are positively correlated with achievement score gains of fourth graders. The failure of most existing studies to include measures of pre-college ability is problematic since higher-ability prospective teachers are likely to attend more prestigious universities and may complete more challenging coursework, thereby biasing upward the estimated impact of college preparation on teacher productivity.

While the gain-score studies lack controls for unobserved student heterogeneity and frequently do not account for teachers' pre-college ability, they often possess good measures of teacher training (see column (a) of Table 1). Nearly all include teacher experience and degree

[6] The licensure score in this case may reflect either stronger cognitive and verbal ability or better teaching skills. One contribution of the present study is to include separate measures of these two factors which therefore aids in the interpretation.

[7] Note that Ehrenberg and Brewer (1995) do not use gains of individual students, but rather rely on changes in scores across different groups of students, often called "synthetic" gains. Also, other studies have focused on teacher scores on certification tests taken after college, such as the National Teacher Examination (NTE), but these scores conflate student cognitive/verbal ability with learning that may take place while in college. In addition to Coltfelter et al. (2005) and Summers and Wolfe (1977) shown in Table 1, Ferguson (1991), Sheehan and Marcus (1978), and Strauss and Sawyer (1986) have studied the relationship between teacher tests and teacher productivity. This latter group of studies is excluded from the table and the remainder of the review because they do not consider any type of teacher training.

level. In addition, many include the undergraduate college major and/or a measure of the selectivity/prestige of the college attended. Eberts and Stone (1984), while excluding college selectivity, is the only early study to include information about in-service training.

## *B. Studies with Panel Data or Random Assignment of Students*

There is evidence that the use of panel data with student fixed effects yields substantially different and more valid estimates of teacher productivity than do gain-score and related models (e.g., Harris and Sass (2005a), Rivkin, et al. (2005)). This calls into question the validity of the results obtained in the earlier generation of studies described in the previous section. We identified four studies of teacher training that use student fixed effects. In addition, we found five studies that address non-random selection through various types of experimental design.[8] These are summarized in the bottom section of Table 1.

Rivkin, et al. were arguably the first to study teacher training while controlling for unmeasured student ability (1998, 2005). While much richer than the two-period analyses of the gain-score studies, their data only identify the grade level and school each student attended and thus can only measure the impact of the average characteristics of teachers across all classrooms in each grade and school. An additional limitation of the Rivkin et al. study is that it includes only two rough measures of teacher training, experience and attainment of advanced degrees, excluding other measures such as undergraduate courses, college major, grades, or other information.

---

[8] We exclude from the discussion Angrist and Lavy (2001) because the results are based on international data. We also do not discuss Hanushek et al. (2005), which addresses somewhat similar questions and uses the same data source as Rivkin et al., but which does not focus on teacher training.

The lack of good measures of teacher training and pre-service education is a common shortcoming of the recent literature. Only two of the recent studies include a measure of pre-service training other than highest degree earned (Betts et al. (2003), Clotfelter et al. (2005)).[9] Further, only two studies include a measure of in-service professional development training (Angrist and Lavy (2001), Jacob and Lefgren (2004)). However these two studies exclude any information on teacher pre-service education.

There are two additional studies using student fixed effects. Jepsen (2005) differs from Rivkin et al. in two ways. First, Jepsen is able to measure teacher characteristics at the teacher level rather than rely on grade-level aggregates. He also includes measures of teacher attitudes and behaviors, such as their enthusiasm and use of computers. While such measures add to the richness of the control variables, there is also a concern that these may absorb some of the training effects. For instance, having more education may make teachers more likely to use computers, which may in turn contribute to their productivity. In Jepsen's model, this would downwardly bias the measured effect of training.

Another study, Rockoff (2004), includes student fixed effects but focuses only on teacher training obtained via on-the-job experience. He points out, as have others (Murnane and Phillips (1981), Wayne and Youngs (2003)), that experience may be correlated with other unobserved teacher characteristics that affect student learning. For example, less effective teachers might be more likely to leave the profession and this may give the appearance that experience raises

[9] Betts et al. (2003) is also distinctive because it is the only study to simultaneously include both student fixed effects and student covariates to address student-to-teacher selection. This would not normally be possible because such covariates are usually either time invariant (e.g., race) or changes in covariates are not generally measured (e.g., parental education). In such cases, the covariates are perfectly collinear with the fixed effect. However, Betts et al. have some time-varying student covariates and therefore include these with the student fixed effects. We are able to use the same approach in our analysis in section IV.

teacher value-added when, in reality, less effective teachers are simply exiting the sample. Alternatively, selection could work in the opposite direction; more able teachers with higher opportunity costs may be more likely to leave the profession, leading to a spurious negative correlation between teacher experience and student achievement. Rockoff addresses this issue by including a teacher fixed effect to control for unmeasured time-invariant teacher ability, along with the experience measures. The teacher-specific effect should purge the influence of teacher ability on experience, yielding unbiased estimates of the marginal product of experience.[10] More generally, this means that teacher effects should be included in any study that attempts to estimate the effects of time-varying training, such as experience and professional development. Besides Rockoff (2004), the other recent studies of teacher training do not include a teacher effect and thus may suffer from omitted variable bias.

Four studies directly address the potential non-random assignment of students to teachers by using data where students are actually or apparently randomly assigned. In these cases, there is no need to include a student fixed effect because any differences across students will be random and therefore will not bias the estimates of the effects of teacher training. It is important to note, however, that random assignment does not eliminate the need for valid teacher controls because these are still necessary to account for non-random assignment of teachers to training.

---

[10] While the inclusion of teacher effects greatly reduces the potential bias associated with teacher attrition, it does not necessarily eliminate it for two reasons. First, since multiple observations are required to compute teacher effects, elementary school teachers who leave after one year are necessarily excluded. This is not a significant problem for middle and high-school teachers, however, since they teach multiple classes within a single period. Second, if there is an unobserved time-varying component of teacher productivity that is correlated with the likelihood of attrition, then this will not be fully captured by the teacher effect. For example, as noted by Murnane and Phillips (1981) and others, the presence of young children in the home may lower teacher productivity and also increase the likelihood of attrition. We test whether teacher-specific effects eliminate attrition bias in our empirical work below.

Three of the random assignment studies, Dee (2004), Ding and Lehrer (2005), and Nye et al. (2004), use data from the Tennessee class size reduction experiment that, while not focused on teacher training, did involve random assignment of student to teachers and collection of some data on teacher training. In the experiment, students and teachers in kindergarten through 3rd grade were randomly assigned to small and large classrooms. Clotfelter, et al. (2005) take a different approach to random assignment using data from North Carolina. They identify schools that appear to randomly assign students across classrooms within the school. This is done by comparing the actual distribution of students in each classroom with a uniform distribution based on six criteria (race, gender, participation in the subsidized school lunch program, student mobility, prior test score and teacher-reported parents' education).

Nearly all of the above studies potentially suffer from the non-random assignment of teachers to training, addressing the issue only by including covariates to control for important teacher characteristics. This is analogous to using student covariates to address the non-random assignment of students to teachers, which has already been shown to be problematic. Again, Rockoff (2004) goes beyond this by including the teacher fixed effect. Angrist and Lavy (2001) and Jacob and Lefgren (2004) also address this concern, but instead do so by making use of "natural experiments." Angrist and Lavy rely on a large and sudden infusion of funds to a group of schools in Israel that was devoted mainly to teacher professional development. Similarly, Jacob and Lefgren rely on the fact that the level of professional development in the Chicago public school district was based (exogenously) on the average level of student test scores. Throughout the remainder of the study we refer to the group of studies in this section as "panel/experiment" studies.

### *C. Results and Discussion*

The results of the above studies are shown in Table 2. We give greatest weight to those studies that have an array of teacher training measures and control for at least one of the forms of non-random selection. Regarding the gain score studies, these include the seven studies that identify the undergraduate college major and three others that exclude such measures, but include controls for teacher cognitive/verbal ability. Among the panel/experiment studies, we give greater weight to the Rockoff study because of his approach to studying teacher experience and to Betts et al. study because they include the college major combined with student fixed effects.

Starting with the effects of pre-service training, Betts et al. (2003) find mixed evidence on which college majors are most closely associated with teacher value-added. In elementary school, they find that "other" majors have higher value-added, but that science majors have lower value-added, compared with the education major.[11] In middle and high school, teachers with majors in the social sciences had higher value-added. One possibly surprising finding is that math majors are no different in affecting student math scores compared with education majors. This conflicts with gain-score studies finding that math majors are more effective (Aarsonson et al. (2003), Goldhaber and Brewer (1997), Monk (1994)) and suggests again the importance of controlling for non-random matching of students and teachers with student fixed effects.

Betts et al. find positive effects of master's degrees on students' scores in elementary mathematics and high school reading. None of the other panel/experiment studies finds any

[11] The specified majors for elementary teachers are English, science, social science, foreign language, math, and "other" major. They also include minor fields in the same areas. Education major is the excluded category. For middle and high school that use the same list, but add "minor in subject taught."

effect of master's degrees on reading. One possible explanation is that the other panel/experiment studies focus on reading at lower grade levels. Another is that the other three studies containing a graduate degree variable (Clotfelter et al. (2005), Jepsen (2005) and Rivkin et al. (2005)) have no measures of undergraduate education, which may be correlated with graduate degrees, creating an omitted variables bias. .

Teacher experience consistently has a positive, but often diminishing effect on teacher value-added according to the panel/experiment studies. Rockoff's results are especially convincing because he accounts for unmeasured teacher characteristics with teacher fixed effects. Moreover, his results suggest that the positive effect of experience found in the other studies is likely under-estimated. Overall, five of the six panel/experiment studies find a positive effect of experience (Betts et al. being the sole exception) and even these may suffer from a downward bias, for the reason described by Rockoff.

Only one of the panel/experiment studies includes a teacher test score. Clotfelter et al. include scores on a teacher licensure exam and find a positive and significant relationship with productivity in math. Five of the gain score models that include similar variables also find the same result. The only two that do not find such a result also include a variable describing university prestige (Murnane and Phillips (1981), Summers and Wolfe (1977)). It is possible that these two variables are capturing the same teacher characteristic—general verbal or quantitative ability—which would explain why the test score is a significant predictor only when the other is excluded from the estimation.

Overall, these studies using panel data or random assignment paint a somewhat different picture than the gain score studies. Undergraduate training in the subject area taught appeared to be important in gain score studies, but not in the two recent studies on the topic (Betts et al.

(2003), Clotfelter et al. (2005)). Experience, on the other hand, was often significant in gain score studies and this result is even more consistent in the panel/experiment studies. Every recent study that has considered experience finds that it is significant in at least one grade or subject; and the effects are somewhat stronger for reading compared with math.

But even the more recent panel/experiment studies all suffer from at least one significant methodological limitation. Most of these studies have extremely limited measures of training. Even those with richer data (Betts et al. (2003), Clotfelter et al. (2005)) are missing measures found to be significant in previous studies. Also, none of the studies reviewed is able to adequately address both of the forms of non-random selection that are likely to influence the results. We are able to address all of these issues in the analysis that follows.

## III. Econometric Model and Estimation Strategies

### *A. Measuring Teacher Productivity and Within-Career Education and Training*

While the issue of measuring a teacher's output is controversial, particularly outside the economics literature, we shall simply define the relevant product as student achievement measured by standardized tests. Consequently, we view a teacher's productivity as their contribution to student achievement, holding other inputs constant. To empirically measure the impact of education and training on teacher productivity it is therefore necessary to first develop a model of student achievement. We begin with a general specification of the standard "educational production function" that relates student achievement to vectors of time-varying student/family inputs (X), classroom-level inputs (C), school inputs (S) and time-invariant student/family characteristics ($\gamma$):

$$A_{it} - A_{it-1} = \Delta A_{it} = \boldsymbol{\alpha}_1 \mathbf{X}_{it} + \boldsymbol{\alpha}_2 \mathbf{C}_{ijmt} + \boldsymbol{\alpha}_3 \mathbf{S}_{mt} + \gamma_i + \varepsilon_{it} \quad (1)$$

The subscripts denote individuals (i), classrooms (j), schools (m) and time (t).

Equation (1) is a restricted form of the cumulative achievement function specified by Todd and Wolpin (2003) where the achievement level at time t depends on the individual's initial endowment (eg. innate ability) and their entire history of individual, family and schooling inputs.[12] Although often not stated, there are a number of implicit assumptions underlying the education production function specified in (1). First, it is assumed that the cumulative achievement function does not vary with age, is additively separable, and linear. Family inputs are assumed constant over time, and the impact of parental inputs on achievement, along with the impact of the initial individual endowment on achievement, induce a (student-specific) constant increment in achievement in each period. This allows the combination of these time-invariant inputs to individual achievement gains to be represented by the student-specific fixed component, $\gamma_i$. Third, the marginal impacts of all prior school inputs decay geometrically with the time between the application of the input and the measurement of achievement at the same rate. Thus lagged achievement serves as a sufficient statistic for all prior schooling inputs. Fourth, to remove individual lagged score from the right-hand side of the gain equation, it is further assumed that the decay rate is actually zero—that is, that school inputs applied at any point in time have an immediate and permanent impact on cumulative achievement.[13] A thorough discussion of these assumptions and the derivation of the linear education production function model can be found in Todd and Wolpin (2003) and Sass (2006).

---

[12] It is important to note that while the dependent variable is the change in student achievement, equation (1) is a model of student achievement levels, not achievement growth. The lagged value of achievement on the left hand side serves to represent the cumulative effect of all prior schooling inputs on current achievement.

[13] Thus, for example, the quality of a child's kindergarten must have the same impact on his cumulative achievement as of the end of the kindergarten year as it does on his achievement at age 18. While a strong assumption, this allows the impact of all prior schooling inputs to be captured by the lagged achievement score, $A_{t-1}$, on the left-hand side of the equation. Otherwise, equation (1) would contain a lagged dependent variable on the right hand side and thus could not be consistently estimated by ordinary least squares.

The vector of classroom inputs can be divided into four components: peer characteristics, $\mathbf{P}_{\text{-ijmt}}$ (where the subscript –i students other than individual i in the classroom), time-varying teacher characteristics (eg. experience and in-service training), $\mathbf{T}_{\text{kt}}$ (where k indexes teachers), time-invariant teacher characteristics (eg. innate ability and pre-service education), $\delta_k$, and non-teacher classroom-level inputs (such as books, computers, etc.), $\mathbf{Z}_j$. If we assume that, except for teacher quality, there is no variation in education inputs across classrooms within a school, the effect of $\mathbf{Z}$ becomes part of the school-level input vector, $\mathbf{S}_m$. If we further assume that school-level inputs are constant over the time span of analysis, they can be captured by a school fixed component, $\phi_m$. The education production function can then be expressed as:

$$\Delta \mathrm{A}_{\mathrm{it}} = \boldsymbol{\beta}_1 \mathbf{X}_{\mathrm{it}} + \boldsymbol{\beta}_2 \mathbf{P}_{-\mathrm{ijmt}} + \boldsymbol{\beta}_3 \mathbf{T}_{\mathrm{kt}} + \gamma_{\mathrm{i}} + \delta_{\mathrm{k}} + \phi_{\mathrm{m}} + \nu_{\mathrm{it}} \tag{2}$$

Where $\nu_{it}$ is a normally distributed, mean zero error.

We include three measures of teacher education and training in the vector of time-varying teacher characteristics, $\mathbf{T}_{\text{kt}}$. Experience, representing on-the-job training, is captured by a truncated quadratic function, experience and experience squared up to ten years of experience, and a dummy variable indicating teachers with more than 10 years of experience. The omitted category is teachers with zero experience. This specification allows for non-linear effects of teacher experience on student achievement and avoids perfect collinearity between experience and time that would result from a continuous linear measure of teacher experience.[14] In-service training is measured by a vector of variables representing the number of hours spent in various types of professional development courses. Both current-year hours of training as well as the amount of training in each of the three prior years are measured separately to allow for delayed implementation of new teaching strategies, human capital depreciation and possible negative

[14] Rockoff (2004) pursues a similar strategy and finds that the marginal effect of experience generally goes to zero around five years of experience.

impacts of contemporaneous training on student achievement associated with absences from the classroom. Finally, attainment of post-baccalaureate degrees is included to capture the effects of additional formal education obtained after entering the teaching profession. The vector of coefficients on these time-varying teacher characteristics, $\beta_3$, thus represents the impact of within-career education and training on teacher productivity.

## *B. Computational Issues*

Estimation of (2) is computationally challenging since it includes three levels of fixed effects: individual students ($\gamma_i$), teachers ($\delta_k$) and schools ($\phi_m$). Standard fixed effects methods eliminate one effect by demeaning the data with respect to the variable of interest (eg. deviations from student means). Additional effects must then be explicitly modeled through the inclusion of dummy variable regressors. Given our data includes tens of thousands of teachers and thousands of schools, such standard methods are infeasible.

We combine two different approaches to solve the computational problem associated with estimating a three-level fixed effects model. First, we utilize the "spell fixed effects" method proposed by Andrews, Schank and Upward (2004) and combine the teacher and school fixed effects into a single effect, $\eta_{km} = \delta_k + \phi_m$. This combined effect represents each unique teacher/school combination or "spell." The education production function thus becomes:

$$\Delta A_{it} = \boldsymbol{\beta}_1 \mathbf{X}_{it} + \boldsymbol{\beta}_2 \mathbf{P}_{-ijmt} + \boldsymbol{\beta}_3 \mathbf{T}_{kt} + \gamma_i + \eta_{km} + \nu_{it} \tag{3}$$

The second approach is an extension of the iterative fixed effects estimator recently proposed by Arcidiacono, et al. (2005).[15] The essence of the Arcidiacono et al. method is to estimate the fixed effect for each individual by calculating each individual's error in each time

[15] Arcidiacono, et al derive their estimator in the context of a model with only fixed effects and no other covariates. However, it is straightforward to extend their approach to models with covariates. Details of the derivation are available upon request.

period (ie. actual outcome minus the individual's predicted outcome) and then compute the mean of these errors for each individual over time. With each estimate the individual fixed effects are recomputed and the process is iterated until the coefficient estimates converge.

Taking deviations from the spell means, the achievement equation becomes:

$$(\Delta A_{it} - \overline{\Delta A}_{km}) = \boldsymbol{\beta}_1(\mathbf{X}_{it} - \overline{\mathbf{X}}_{km}) + \boldsymbol{\beta}_2(\mathbf{P}_{-ijmt} - \overline{\mathbf{P}}_{km}) + \boldsymbol{\beta}_3(\mathbf{T}_{kt} - \overline{\mathbf{T}}_{km}) + (\gamma_i - \overline{\gamma}_{km}) + \nu_{it} \tag{4}$$

where the overbar and km subscript denote the mean of the relevant variable over all students and all time periods covered by teacher k at school m. Subtracting the de-meaned individual effect from both sides yields:

$$(\Delta A_{it} - \overline{\Delta A}_{km}) - (\gamma_i - \overline{\gamma}_{km}) = \boldsymbol{\beta}_1(\mathbf{X}_{it} - \overline{\mathbf{X}}_{km}) + \boldsymbol{\beta}_2(\mathbf{P}_{-ijmt} - \overline{\mathbf{P}}_{km}) + \boldsymbol{\beta}_3(\mathbf{T}_{kt} - \overline{\mathbf{T}}_{km}) + \nu_{it} \tag{5}$$

Equation (5) is estimated by ordinary least squares (OLS), using initial guesses for the individual effects. This produces estimates of $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_3$ which are then used to calculate predicted outcomes for each individual and in turn update the estimated individual effects. The process is iterated until the coefficient estimates converge. Standard errors are obtained by bootstrapping.[16]

[16] The standard errors from the bootstrap procedure do not account for clustering of students within a classroom or classrooms within a school. This is partly compensated for by the fact that we include classroom peer measures and teacher fixed effects (which correspond to a common average error for all students a teacher ever teaches). Unfortunately, a procedure to account for clustering of errors at multiple levels within a bootstrap framework is not currently available. Cameron, Gelbach and Miller (2006a) derive corrected standard errors for the case where clustering occurs at multiple levels (eg. students in a class, classes within a teacher, teachers within a school), but there is not complete nesting (over time students switch teachers and teachers switch schools). However, this is method is based on OLS regressions and is not applicable to our iterative model. Cameron, Gelbach and Miller (2006b) derive a bootstrap procedure for determining the corrected standard errors when there is clustering at a single level, but the authors have not yet worked out a bootstrap procedure for the case of multi-way clustering.

*C. Measuring the Effects of Pre-Service Education on Teacher Productivity*

In order to gauge the effects of teacher ability and college preparation on future productivity we follow a two-step estimation procedure first proposed by Dickens and Ross (1984). In the first step we calculate the estimated teacher-school effects from the estimation of equation (5). The teacher-school spell fixed effects can be expressed as the difference between the average achievement gain for all students in group km minus the product of the estimated coefficients and the group averages of the explanatory variables:

$$\eta_{km} = \overline{\Delta A}_{km} - \bar{\gamma}_{km} - \boldsymbol{\beta}_1 \overline{\mathbf{X}}_{km} - \boldsymbol{\beta}_2 \overline{\mathbf{P}}_{km} - \boldsymbol{\beta}_3 \overline{\mathbf{T}}_{km} \qquad (6)$$

These teacher-school effects can be decomposed into three components: the school effect, the portion of the teacher effect due to pre-college ability, and the part of teacher effect due to the education they receive in college. We measure pre-college ability by a teacher's college entrance exam score. In the second step we gauge the impact of pre-service education on later teacher productivity by regressing the estimated teacher-school effects on a vector of pre-service education variables for teacher k, $U_k$, their entrance exam scores, $E_k$, a set of school dummies, $\phi_m$, and a random error.

$$\eta_{km} = \boldsymbol{\omega}_1 \mathbf{U}_k + \boldsymbol{\omega}_2 \mathbf{E}_k + \phi_m + \xi_{km} \qquad (7)$$

The estimates of the coefficient vector $\omega_1$ indicate the marginal effect of changes in the characteristics of a teacher's pre-service education on their future productivity. Following Dickens and Katz (1986), equation (7) is estimated by weighted least squares, with the square root of the numbers of students per teacher/school spell as weights.

# IV. Data

We make use of a unique panel data set of school administrative records from Florida[17] that allows us to overcome many of the challenges associated with measuring the impact of education and training on teacher productivity. The data cover nine school years, 1995-1996 through 2003-2004, and include all public-school students and their teachers throughout Florida. Achievement test data are available for both math and reading in each of grades 3-10 for the years 1999-2000 through 2003-2004.[18] Summary statistics are available in Table 3.

Like statewide databases in North Carolina and Texas, the Florida data track individual students over time and thus allow us to control for unobserved student characteristics with student-specific fixed effects. However, unlike other statewide databases, we can precisely match both students and teachers to specific classrooms at all grade levels.[19] This classroom-level data, with consistent teacher identification over time, allows us to control for time-invariant teacher characteristics via fixed effects. Another advantage of the data is that we can determine the specific classroom assignments of middle-school and high-school students, who typically

---

[17] A more detailed description of the data is provided in Sass (2006).

[18] The state of Florida currently administers two sets of reading and math tests to all 3rd through 10th graders in Florida. The "Sunshine State Standards" Florida Comprehensive Achievement Test (FCAT-SSS) is a criterion-based exam designed to test for the skills that students are expected to master at each grade level. The second test is the FCAT Norm-Referenced Test (FCAT-NRT), a version of the Stanford-9 achievement test used throughout the country. The scores on the Stanford-9 are scaled so that a one-point increase in the score at one place on the scale is equivalent to a one-point increase anywhere else on the scale. The Stanford-9 is a vertically scaled exam, thus scale scores typically increase with the grade level. We use FCAT-NRT scale scores in all of the analysis. The use of vertically scaled scores to evaluate student achievement is important since a one-unit change has the same meaning for low- and high-achieving students. Other types of measures, such as standard deviations from the mean score are potentially problematic; it is not clear that a 0.1 standard deviation increase in a test score, starting one standard deviation from the mean, is the same as a 0.1 standard deviation increase for someone with an initial score equal to the sample mean.

[19] Currently, the Texas data do not provide a way to link teachers and students to specific classrooms. For North Carolina, one can only (imperfectly) match specific teachers and students to classrooms at the elementary school level. Matching is done by identifying the person who administers each student the annual standardized test, which at the elementary school level is typically the classroom teacher.

rotate through classrooms during the day for different subjects. This allows us to better separate the effects of teachers from students and their peers. It also affords us the opportunity to see how the impacts of teacher education and training vary at different grade levels. For example, one might expect that a teacher's content knowledge of mathematics might be more important in a high school trigonometry class than in a fourth grade class learning arithmetic.

Not only does our data directly link students and teachers to specific classrooms, it also provides information on the proportion of time spent in each class. This is potentially important for correctly matching teachers and their students at the elementary school level. While primary school students typically receive all of their academic instruction from a single teacher in a single "self-contained" classroom, this is far from universal. In Florida, five percent of elementary school students enrolled in self-contained classrooms are also enrolled in a separate math course, four percent in a separate reading course and four percent in a separate language arts course. In addition, nearly 13 percent of elementary students enrolled in self-contained elementary classes are also enrolled in some type of exceptional student education course apart from their regular classroom, either special-education or gifted courses.[20]

We restrict our analysis of student achievement to students who receive instruction in the relevant subject area in only one classroom. Only elementary school students in "self-contained" classrooms are included. Elementary students spending less than one hour per day in the class are not considered as a member of the classroom peer group. At the middle and high-school level students who are enrolled in more than course in the relevant subject area (mathematics and reading/language arts) are dropped, though all students enrolled in a course are included in the

[20] Since previous studies lack data on students' complete course enrollments, they either ignore the fact that students may receive instruction outside their primary classroom or deal with the issue in an ad-hoc fashion.

measurement of peer-group characteristics. To avoid atypical classroom settings and jointly taught classes we consider only courses in which 10-50 students are enrolled and there is only one “primary instructor” of record for the class. Finally, we eliminate charter schools from the analysis since they may have differing curricular emphases and student-peer and student-teacher interactions may differ in fundamental ways from traditional public schools.

Another unique aspect of the Florida data is the inclusion of information on post-secondary students. For relatively young teachers (those who attended a Florida public university or community college since 1995) our data include complete college transcript information, including entrance exam scores, courses taken and degrees received. Because Florida has a uniform course numbering system, we are able to create variables that describe each course according to its focus on teacher content knowledge, pedagogical knowledge, and classroom observation/practice in teaching.[21] We then aggregate these measures for each teacher to capture the relevant characteristics of each teacher’s entire undergraduate training. We are also know the major associated with each college degree and can thus distinguish future teachers who graduated with an education major from those who earned a degree in various non-education disciplines like mathematics and English literature.

[21] Courses were coded using the course descriptions in the State of Florida State Course Numbering System. The following categories are used: *Education theory/foundations* includes courses that cover general education theory or general issues in education. *Pedagogical-instructional* includes general instructional methods and theories to instruction. *Pedagogical-management* includes classroom management issues in general or for different groups of students. *Pedagogical-content* includes combinations of subject and pedagogy. *Other development* includes issues such as ethics, professionalism or administration. *Classroom observation* includes observation in the classroom. *Classroom practice* includes courses that require field experience. *Subject content* includes subject content (e.g. math). Each course was assigned a total value of one which was, in some cases, distributed over several types of training. An example may help to illustrate: SCE 4361 Introduction to Middle School Science Teaching was coded as *pedagogical-content* (0.3) and *classroom practice* ( 0.7). This is the based on the course description: “Introduction to the roles and responsibilities of science teachers with an emphasis on middle school students. Extensive fieldwork required.”

# V. Results

## A. Effects of Experience and Professional Development Training

Estimates of the student achievement model, equation (5), without and with teacher fixed effects are presented in Tables 4 and 5, respectively. Following Rockoff (2004), experience effects are indicated by three measures: experience and experience squared (for the first ten years of experience), and experience beyond ten years. In-service education is measured by the total hours of professional development training, independent of course content. In Table 6, we estimate a similar model (including teacher fixed effects) except that content-oriented professional development is separated from "total" professional development. One general conclusion that emerges is that the effects of these different types of training vary by subject, grade level and the timing of the training. We discuss this below in further detail.

Estimates of the model without teacher fixed effects, presented in Table 4, indicate that experience enhances teacher productivity at all grade levels in math and in elementary school reading. There is evidence that veteran teachers with ten or more years of experience outperform less experienced teachers in middle school reading as well. However, we find no impact of experience on teacher productivity for high-school reading teachers. This later finding may be due to the fact that the curricular focus of high-school language arts classes is on literature and less on the mechanics of reading. Also, reading scores are more likely to be influenced by non-school factors, such as how many books students read in their free time.

Without teacher fixed effects we also find that there is no net impact of contemporaneous professional development on teacher productivity. However, prior-year professional development coursework is positively correlated with current productivity. In particular, there are positive effects of prior professional development for math teachers at all grade levels and for elementary school teachers in reading.

When the models presented in Table 4 are re-estimated with teacher fixed effects most of the coefficients on the experience variables become insignificant, as shown in Table 5. With teacher fixed effects included, teacher experience only has a positive effect on student achievement in middle school math. The coefficients on the three experience measures imply that the marginal effect of moving from zero to one year of experience is roughly 0.37 scale score points on the student achievement test. This translates to 0.015 of a standard deviation in student achievement gains or 0.010 of a standard deviation in the achievement level. The marginal effect diminishes and eventually becomes negative at roughly the eighth year of experience and there are no positive effects of experience beyond 10 years in this or any other subject-grade combination. Our finding that the only apparent effects are in math contrasts with other studies which have more frequently found positive experience effects in reading.

These weak effects of experience are perhaps not surprising given the evidence of other recent panel studies summarized in Table 2. In particular, the only other study to employ teacher fixed effects, Rockoff (2004) also obtains relatively weak experience effects that vary across subject levels. This suggests three possible explanations. First, it could be that despite the non-linear specification of experience there is still some multicollinearity among the experience variables, teacher fixed effects, and time/grade dummies. Second, there may be insufficient within-teacher variation in experience over the span of our relatively short panel to get precise estimates of experience effects. Alternatively, it could be that past results, which excluded teacher fixed effects and generally found strong impacts of experience, were subject to a form of selection bias. If less effective teachers are more likely to leave the profession then comparisons across teachers with different experience could be reflecting both selection on unmeasured teacher characteristics as well as true experience effects. The evidence in Tables 4 and 5 provides some supports for all of these explanations. In some cases (e.g., elementary math), the point estimates are similar after the teacher effects are added, but the precision is reduced, suggesting the multicollinearity explanation. In other cases, such as high school math, the point estimates change significantly, suggesting that the exclusion of the teacher effects introduces bias.

As noted in our literature review, even models with teacher fixed-effects may yield biased estimates of the impact of experience if unobserved time-varying teacher characteristics (e.g. pregnancy and the presence of young children) affect both teacher attrition and teacher productivity. Following the suggestion of Semykina and Wooldridge (2005), we test for attrition bias by adding a variable indicating attrition in the next period. The coefficient on the additional variable is always insignificant, indicating our teacher fixed effects sufficiently account for unobserved teacher heterogeneity.

The results for PD with teacher effects in the model are in many ways similar to those of experience. As with teacher experience, the precision of the estimated effects of in-service training change significantly if we include teacher fixed effects. With teacher effects in the model there is little evidence of PD effects, except again in middle school math. Interestingly, in this subject and grade, the positive effects do not occur in the period in which the PD occurs, but PD does appear to improve teaching in each of the three subsequent years. A similar, but less consistent effect of PD is found in high school math, but this effect is apparent in only one of the lag terms. There is also one instance (elementary reading) where PD appears to have a negative effect

The possibility of a lagged PD effect has been observed elsewhere (Goldhaber and Anthony (2004), Harris and Sass (2005b)). One possible explanation is that PD reduces the amount of time teachers have to devote to their students when the PD is taking place. In addition, if substitute teachers are hired so that the PD can take place during school hours, and if the substitutes are less effective or unable to maintain the continuity of instruction in the teacher's absence, then this too may reduce measured teacher value-added. Therefore, at time t in our results, any positive effect may be more than offset by a negative effect. At time t-1, one would expect the teacher to be able to incorporate the training into teaching lessons and effects should begin to appear, if any exist. The fact that the effects are largest three years after PD has taken place suggests that the effects of training can take many years to show up in student achievement scores.

In Table 6, we identify the effects of different types of PD and find that the effects of total PD on middle school math, discussed above, are attributable mainly to content-oriented training. The lagged coefficients for this PD type remain consistently positive and statistically significant. The lagged effect of other in-service hours on middle school math is also positive and significant, but this is true in only one case and this effect is much smaller in magnitude than the content PD effects. The only other substantive change in the results, compared with Table 4, is that there now appear to be some (lagged) effects of PD on high school reading.

These weak and inconsistent findings regarding PD are consistent with the only two studies that consider the effects of PD (Angrist and Lavy, 2001; Jacob and Lefgren, 2004)), which found no effects. The results in the present study are somewhat more positive (again, especially in middle school math), perhaps because the previous studies estimate only the short-term effects of PD which, as we have shown, are smaller than the longer-term effects.

We consider the impact of advanced degrees in Table 7. Since our model includes teacher fixed effects, post-baccalaureate degrees earned prior to the period of analysis wash out when we demean the data. Thus our approach measures the impact of *changes* in the possession of an advanced degree (for a given teacher) during the period of study.[22] Our results indicate that obtaining an advanced degree during one's teaching career does not enhance productivity and may actually reduce productivity in high school math and middle school reading. This may be because the graduate degrees include a combination of pedagogy and content and our other evidence suggests that only the latter has a positive influence on teacher productivity.

Other explanations for the graduate degree results arise from issues of methodology. Most previous studies suffer from selection bias, as noted earlier, and our solution is to study the effects of the graduate degree attainment within teachers using teacher fixed effects. However, this approach imposes the implicit assumption that the receipt of the graduate degree reflects a sudden infusion of new preparation. In reality, the receipt of the degree is the culmination of several years of graduate courses whose influence may already be reflected in the teacher effects, especially for those teachers who take graduate courses over many years before receiving a graduate degree. Another scenario is that teachers load up on courses in the academic year preceding the receipt of the degree and therefore have less attention to devote to their students as noted above. We found evidence above of such a contemporaneous decline in productivity when we considered the effects of other forms of professional development.

[22] The estimated coefficient on the advanced-degree variable measures the average productivity differential between the time before and the time after receipt of the degree. Before the degree is received some knowledge may have already been acquired through coursework already completed, thus biasing the estimated effect toward zero. However, work toward an advanced degree may take away from time available for class preparation and other teaching-related activities, which would tend to lower productivity before receipt of the degree and upwardly bias the estimated impact of the degree.

## *B. Pre-Service Training Effects*

The results in Tables 5, 6 and 7 are based on models that include teacher effects. In this section we use these teacher effects as the dependent variable in order to analyze the effects of pre-service training, as shown in equation (7). Table 8 displays the results for the effects of teachers of various undergraduate majors, including various types of education degrees, math and English. Note that the sample size drops significantly because our data only have information on teachers' college majors dating back to 1995 and, even for those with college transcripts in the database, there are a substantial number of missing observations on the college entrance exam.[23] Our results exclude observations with missing data on these key variables. Nonetheless, our remaining sample size is still larger than many previous studies on the subject.

We find that teacher productivity is not significantly related to the choice of college major. This is consistent with Betts, Zau, and Rice (2003), the only other panel or random assignment study to include data on college major (see Table 2).[24] We also find, contrary to several early studies using two-year gain scores, that teacher test scores (in this case, SAT) are not associated with teacher value-added and may be negatively associated in high school math.

Table 9 also displays the estimated effects of pre-service training but, in this case, we focus on the specific content of education courses. Again, few of the coefficients are significant, but there are some notable exceptions. Pedagogical-content knowledge appears to contribute positively to teacher effectiveness in elementary and middle-school math, although not in

---

[23] To maximize the amount of college entrance exam information available we include data from the state university system, the community college system and a database on applicants for a state merit-based scholarship known as "bright futures." ACT scores as well as community college placement exam scores were converted to SAT scores using concordance tables.

[24] Estimates of the high school math equation produce the anomalous finding that teacher productivity is inversely related to scores on the quantitative portion of the SAT

reading or in the other grade levels. Also, the quantity of English literature credits is positively correlated with teacher value-added at the high school level. In contrast, mathematics and statistics coursework outside the college of education does not appear to increase teacher productivity and in fact is negatively correlated with teacher value-added in some instances. The SAT score continues to be largely insignificant, confirming the results from Table 8.

Overall, while our results are inconsistent across grades and subjects, it appears that colleges of education might improve the performance of their graduates, and schools might improve their existing teachers, by placing somewhat greater emphasis on content knowledge, including that which is pedagogically oriented. We find no evidence that pedagogical training itself contributes positively to teacher value-added and there is some evidence graduates of colleges of education are, on the average, less effective than others. This conclusion is suggested by both the apparently positive effects of content-oriented courses in teacher preparation programs and the effects of content-oriented in-service professional development in middle and high school math.

## VI. Summary and Conclusions

This study provides a substantially different view of the determinants of teacher quality compared with much of the past literature. We find no evidence that teachers with higher college entrance exam scores or who receive undergraduate degrees in the subject they teach are more effective than other teachers. Content-oriented professional development appears to positively influence teacher productivity in both pre-service and in-service forms, but only in middle and school math. In other respects, our results reinforce previous findings.

Panel/experiment studies find that more experienced teachers are more effective, although we find this only holds in elementary and middle school reading.

Our literature review and analysis also highlight some of the critical methodological issues that arise in studying teacher productivity. No previous study has been able to account for the non-random assignment of students to teachers, and of teachers to training, while also including an extensive set of teacher training measures. Our analysis here is the first to do so and we find evidence that failure to account for these methodological issues have in fact led to systematic biases in past studies. We also find that addressing all of these issues creates new methodological challenges, such as potential multicollinearity between experience and teacher effects. The use of teacher fixed effects also changes the interpretation of estimated coefficients in important ways, as highlighted by the discussion of graduate degrees.

These results can help policymakers to move closer toward more justifiable policies regarding teacher training, although the implications are not yet entirely clear. On the one hand, the lack of influence of undergraduate degrees suggests that reducing restrictions based on college majors is an appropriate policy direction. However, the importance of experience, as least at some grades, also means there are significant benefits to retaining teachers over and above the avoidance of the cost of hiring new teachers. What is clearer is that policies, rather than being broad and rigid, will need to account for differences in training effects that vary across grades and subjects, as well as perhaps teachers and schools.

## References


Andrews, Martyn, Thorsten Schank and Richard Upward. 2004. "Practical Estimation Methods for Linked Employer-Employee Data." Unpublished. Manchester: University of Manchester.

Angrist, Joshua D. and Victor Lavy. 2001. "Does Teacher Training Affect Pupil Learning? Evidence from Matched Pairings in Jerusalem Public Schools." *Journal of Labor Economics* 19(2):343-69.

Aaronson, Daniel, Lisa Barrow and William Sander. 2003. "Teachers and Student Achievement in the Chicago Public High Schools." Unpublished. Chicago: Federal Reserve Bank of Chicago.

Arcidiacono, Peter, Gigi Foster, Natalie Goodpaster and Josh Kinsler. 2005. "Estimating Spillovers in the Classroom with Panel Data." Unpublished. Durham: Duke University.

Betts, Julian R., Andrew C. Zau and Lorien A. Rice. 2003. Determinants of Student Achievement: New Evidence from San Diego. San Diego: Public Policy Institute of California.

Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller. 2006a. , "Robust Inference with Multi-way Clustering." Unpublished. Davis: University of California at Davis.

Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller. 2006b. , " Bootstrap-Based Improvements for Inference with Clustered Errors." Unpublished. Davis: University of California at Davis.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2005. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." Unpublished. Durham: Duke University.

Dee, Thomas S. 2004. "Teachers, Race and Student Achievement in a Randomized Experiment." *Review of Economics and Statistics*, 86(1):195-210.

Dickens, William T. and Brian A. Ross. 1984. "Consistent Estimation Using Data From More Than One Sample." Technical Working Paper #33. Cambridge, MA: National Bureau of Economic Research.

Dickens, William T. and Katz, Lawrence F. 1986. "Interindustry Wage Diffferences and Industry Characteristics." Working Paper #2014. Cambridge, MA: National Bureau of Economic Research.

Ding, Weili and Steven F. Lehrer. 2005. "Accounting for Unobserved Ability Heterogeneity Within Education Production Functions." Unpublished. Kingston: Queen's University.

Eberts, Randall W. and Joe A. Stone. 1984. Unions and Public Schools: The Effect of Collective Bargaining on American Education. Lexington: Heath.

Ehrenberg, Ronald G. and Dominic J. Brewer. 1994. “Do School and Teacher Characteristics Matter? Evidence from High School and Beyond.” *Economics of Education Review* 13(1):1-17.

Ehrenberg, Ronald G. and Dominic J. Brewer. 1995. “Did Teachers' Verbal Ability and Race Matter in the 1960s? Coleman revisited.” *Economics of Education Review* 14(1):1-21.

Feng, Li. 2005. “Hire Today, Gone Tomorrow: The Determinants of Attrition Among Public School Teachers.” Unpublished. Tallahassee: Florida State University.

Ferguson, Ronald. F. 1991. “Paying for Public Education: New evidence on How and Why Money Matters.” *Harvard Journal on Legislation* 28(2):465-98.

Ferguson, Ronald F. and Ladd, Helen F. 1996. “How and Why Money Matters: An Analysis of Alabama Schools” in Holding Schools Accountable: Performance Based Reform in Education, ed. Helen F. Ladd. Washington: Brookings Institution.

Goldhaber, Dan and Emily Anthony. 2004. “Can Teacher Quality Be Effectively Assessed?” Unpublished. Seattle: University of Washington.

Goldhaber, Dan D. and Dominic J. Brewer. 1997. “Why Don’t Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity.” Journal of Human Resources 32(3):505-23.

Goldhaber, Dan D. and Dominic J. Brewer. 2001. “Evaluating the Evidence on Teacher Certification: A Rejoinder.” *Educational Evaluation and Policy Analysis* 23(1):79-86.

Hanushek, Eric A. 1986. “The Economics of Schooling: Production and Efficiency in Public Schools.” *Journal of Economic Literature* 24(3): 1141-77.

Hanushek, Eric A. 1992. “The Trade-off Between Child Quantity and Quality.” *Journal of Political Economy* 100(1):84-117.

Hanushek, Eric A., John F. Kain, Daniel M. O’Brien, and Steven G. Rivkin. 2005. “The Market for Teacher Quality.” Working Paper #11154. Cambridge, MA: National Bureau of Economic Research.

Harnisch, Delwyn L. 1987. “Characteristics Associated with Effective Public High Schools.” *Journal of Educational Research* 80(4):233-241.

Harris, Douglas and Tim R. Sass. 2005a. “Advances in Value-Added: Theory and Applications Related to Teacher Quality.” Unpublished. Tallahassee: Florida State University.

Harris, Douglas and Tim R. Sass. 2005b. "The Effects of NBPTS Teachers on Student Achievement." Unpublished. Tallahassee: Florida State University.

Hill, Heather, Brian Rowan, and Deborah Loewenberg Ball. 2005. "Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement." *American Educational Research Journal* 42(2):371-406.

Jacob, Brian A. and Lars Lefgren. 2004. "The Impact of Teacher Training on Student Achievement: Quasi-Experimental Evidence from School Reform Efforts in Chicago." *Journal of Human Resources* 39(1):50-79.

Jepsen, Christopher. 2005. "Teacher Characteristics and Student Achievement: Evidence from Teacher Surveys." *Journal of Urban Economics* 57(2):302-19.

Kiesling, Herbert J. 1984. "Assignment Practices and the Relationship of Instructional Time to the Reading Performance of Elementary School Children." *Economics of Education Review* 3(4):341-50.

Link, Charles R. and Edward C. Ratledge. 1979. "Student Perceptions, I.Q. and Achievement," *Journal of Human Resources* 14(1) 98-111.

Monk, David H. 1994. "Subject Area Preparation of Secondary Math and Science Teachers and Student Achievement." *Economics of Education Review* 13(2):125-45.

Murnane, Richard J. 1975. *The Impact of School Resources on the Learning of Inner City Children*. Cambridge, MA: Ballinger.

Murnane, Richard J. and Barbara R. Phillips. 1981. What do Effective Teachers of Inner-City Children Have in Common?" *Social Science Research* 10(1):83-100.

Nye, Barbara, Spyros Konstantopoulos and Larry V. Hedges. 2004. "How Large are Teacher Effects?" *Educational Evaluation and Policy Analysis* 26(3):237-257.

Rice, Jennifer K. 2003. *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*. Washington: Economic Policy Institute.

Rivkin, Steven G., Eric A. Hanushek and John F. Kain. 1998. "Teachers, Schools and Academic Achievement," Working Paper #6691. Cambridge, MA: National Bureau of Economic Research.

Rivkin, Steven G., Eric A. Hanushek and John F. Kain. 2005. "Teachers, Schools and Academic Achievement." *Econometrica* 73(2):417-58.

Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review* 94(2): 247-52.

Rowan, Brian, Fang-Shen Chiang and Robert J. Miller. 1997. “Using Research on Employees' Performance to Study the Effects of Teachers on Students' Achievement.” *Sociology of Education* 70(4):256-84.

Sass, Tim R. 2006. “Charter Schools and Student Achievement in Florida.” *Education Finance and Policy* 1(1):91-122.

Semykina, Anastasia and Jeffrey Wooldridge. 2005. “Estimating Panel Data Models in the Presence of Endogeneity and Selection: Theory and Application.” Unpublished. East Lansing: Michigan State University.

Sheehan, Daniel S. and Mary Marcus. 1978. “Teacher Performance on the National Teacher Examinations and Student Mathematics and Vocabulary Achievement.” *Journal of Education Research* 71(3):134-36.

Strauss, Robert P. and Elizabeth A. Sawyer. 1986. “Some New Evidence on Teacher and Student Competencies.” *Economics of Education Review* 5(1):41-48.

Summers, Anita A. and Barbara L. Wolfe. 1977. “Do Schools Make a Difference?” *American Economic Review* 67(4):639-52.

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.

Wayne, Andrew J. and Peter Youngs. 2003. “Teacher Characteristics and Student Achievement Gains.” *Review of Educational Research* 73(1):89-122.

Wilson, Suzanne, Robert E. Floden and Joan Ferrini-Mundy. 2001. *Teacher Preparation Research: Current Knowledge, Gaps, and Recommendations*. Seattle, WA: Center for the Study of Teaching and Policy.

Wilson, Suzanne and Robert Floden. 2003. *Creating Effective Teachers: Concise Answers for Hard Questions*. New York: AACTE Publications.

## Table 1: Summary of Past Research Methods

| *Studies* | *Teacher training measures* (a) | *Method addressing non-random assignment of teachers to training* (b) | *Method addressing non-random assignment of teachers to students* (c) |
|---|---|---|---|
| **Gain Score Studies** | | | |
| Aaronson, et al. (2003) | College selectivity, major | Covariates; certification | Covariates |
| Eberts and Stone (1984) | College courses, in-service | Covariates; time use | Covariates |
| Ehrenberg and Brewer (1994) | College selectivity † | Covariates | Covariates |
| Ehrenberg and Brewer (1995) | n.a. | Covariates; verbal test | Covariates |
| Ferguson and Ladd (1996) | n.a.† | Covariates; ACT | Covariates |
| Goldhaber and Brewer (1997) | Degrees by subject | Covariates; certified | Covariates |
| Goldhaber and Brewer (2001) | College major in subject | Covariates; certification | Covariates |
| Hanushek (1992) | n.a. | Covariates; word test | Covariates |
| Harnisch (1987) | n.a.* | Covariates | Covariates |
| Hill, et al. (2005) | n.a.* | Covariates; content knowledge | Covariates |
| Kiesling (1984) | n.a.* | Covariates; instruct., prep. time | Covariates |
| Link and Ratledge (1979) | n.a. | Covariates | Covariates |
| Monk (1994) | College courses (detail) | Covariates | Covariates |
| Murnane (1975) | College major | Covariates; college GPA | Covariates |
| Murnane and Phillips (1981) | College prestige | Covariates; verbal test | Covariates |
| Rowan, et al. (1997) | College major * | Covariates; math test | Covariates |
| Summers and Wolfe (1977) | College selectivity | Covariates; NTE | Covariates |
| **Panel/Random Assignment Studies** | | | |
| Angrist and Lavy | Professional develop. * † | Covariates; matched sample | Fixed effect |
| Betts et al. (2003) | College major | Covariates; certif., college major | Fixed effect; covariate |
| Clotfelter, et al. (2005) | College selectivity | Covariates; licensure exam | Random assign. |
| Dee (2004) | n.a. | Covariates | Random assign. |
| Ding and Lehrer (2005) | n.a. | Covariates | Random assign. |
| Jacob and Lefgren (2004) | Professional develop. * | Quasi-experiment | Covariates |
| Jepsen (2005) | n.a. | Covariates; practices, certif.. | Fixed effect |
| Nye, et al (2004) | n.a. | Covariates | Random assign. |
| Rivkin, et al. (2005) | n.a. | Covariates; certification | Fixed effect |
| Rockoff (2004) | n.a. | Covariates; teacher fixed effect | Fixed effect |

As indicated in the text, all of these studies use longitudinal student test scores. Experience and highest degree level are included in all studies and therefore not listed in column (a), except where indicated by an asterisk (*). In column (b), certification is listed as a means of addressing non-random assignment of teachers to training, although this is really only true for in-service (certification occurs after pre-service training and therefore cannot cause pre-service training). Teacher race and gender are included as controls in all studies, except where noted. Also, the symbol (†) indicates that teacher quality variables are averaged across teachers, either within grades or schools, rather than for individual teachers. Note that only the Betts et al. study uses both a student fixed effect and covariates. The main reason is that, in most data sets, student covariates do not vary over time. However, in the Betts et al. study, some student covariates change over time, allowing for the simultaneous inclusion of fixed effects. This is noteworthy because the present study uses the same approach.

**Table 2: Results of Past Studies, By Teacher Training Type**
[Training Variable (Subject, Grade, Result)]

| Studies | *Undergraduate training* *(a)* | *Graduate training and professional development* *(b)* | *Experience* *(c)* | *Test score* *(d)* |
|---|---|---|---|---|
| **Gain Score Studies** | | | | |
| Aaronson, et al. (2003) | Degree educ, math/sci (MH ++, RH ++) | n.a. | (MH 0, RH 0) | n.a. |
| Eberts and Stone (1984) | Math courses (ME 0) | MA (ME --) PD (ME 0) | (0) | |
| Ehrenberg and Brewer (1994) | n.a. | MA (CH -) | (CH 0) | n.a. |
| Ehrenberg and Brewer (1995) | n.a. | MA (CE 0, CH 0) | (CE +, CH 0) | (CE ++, CH ++) |
| Ferguson and Ladd (1996) | n.a. | MA (ME ++, RE 0, MM ++, RM ++) | (ALL 0) | (ME 0, RE ++, MM ++) |
| Goldhaber and Brewer (1997) | Degree math (MH ++) | MA in subject (MH ++) | (MH 0) | n.a. |
| Goldhaber and Brewer (2000) | Degree in subject (MH 0, SH -) or educ (MH 0, SH -) | MA in subject (MH 0, SH 0) or educ (MH 0, SH 0) | (MH 0, SH ++) | n.a. |
| Hanushek (1992) | n.a. | MA (RE 0) | (RE ++) | (RE +) |
| Harnisch (1987) | n.a. | MA (ALL 0) | n.a. | n.a. |
| Hill et al. (2005) | n.a. | n.a. | (ME 0) | (ME ++) |
| Kiesling (1984) | n.a. | MA (RE -) | (RE ++) | n.a. |
| Link and Ratledge (1979) | n.a. | MA (RE 0) | (RE 0) | n.a. |
| Monk (1994) | Subject-specific content (MH ++, SH Mix) and educ. courses (MH +, SH Mix) | Subject-specific content (MH +, SH Mix), and educ. courses (MH +, SH Mix) | (MH +, SH 0) | n.a. |
| Murnane (1975) | Degree educ (ME 0, RE 0) | MA (ME -, RE 0) | (ME ++, RE ++) | n.a. |
| Murnane and Phillips (1981) | Univ. prestige (RE 0) | MA (RE 0) | (RE +) | (RE -) |
| Rowan, et al. (1997) | Degree math (MH +) | n.a. | n.a. | (MH ++) |
| Summers and Wolfe (1977) | Univ. prestige (CE +) | n.a. | (CE 0) | (CE -) |
| **Panel, Random Assign. Studies** | | | | |
| Angrist and Lavy (2001) | n.a. | PD (ME +, RE +) | n.a. | n.a. |
| Betts et al. (2003) | College major (ALL Mix) | MA (ME +, RE 0, MM 0, RM 0, MH 0, RH ++) | (ME 0, RE 0, MM +, RM 0, MH 0, RH 0) | n.a. |
| Clotfelter, et al. (2005) | Univ. prestige (ME 0, RE 0) | MA (ME --, RE --) | (ME ++, RE ++) | (ME ++, RE 0) |
| Dee (2004) | n.a. | MA (ME +, RE 0) | (ME 0, RE ++) | n.a. |
| Ding and Lehrer (2005) | n.a. | MA (ME 0, RE 0) | (ME 0, RE +) | |
| Jacob and Lefgren (2004) | n.a. | PD (ME 0, RE 0) | n.a. | n.a. |
| Jepsen (2005) | n.a. | >BA (ME 0, RE 0) | (ME +, RE +) | n.a. |
| Nye, et al. (2004) | n.a. | MA (ME +, RE 0) | (ME +, RE +) | n.a. |
| Rivkin, et al. (2005) | n.a. | MA (MM 0, RM 0) | (MM +, RM 0) | n.a. |
| Rockoff (2004) | n.a. | MA(ME 0, RE -) | (ME 0, RE ++) | n.a. |

Each cell starts by listing the specific variable under consideration, except in column (c) where experience is defined the same way across studies and (d) because the teacher tests are defined in table 1. These variables include: MA = master's degree; PD = professional development; BA = bachelor's degree. The first letter in parentheses indicates the subject area: M = math, R = reading, S = science, and C = composite of math and reading. The second letter indicates the grade level: E = elementary, M = middle school, and H = high school. This is followed by information regarding the effects of the specified variable on student achievement scores in the previously specified subject and grade in the preferred specifications: ++ = positive and significant in nearly all preferred specifications; + = often positive and significant; 0 = insignificant; - = often negative and significant; -- = negative and significant in nearly all preferred specifications; and Mix = mix of positive/significant and negative/significant.

**Table 3**
**Summary Statistics**
**Florida Public School Students and Teachers, 1999/2000-2003/2004**

| | Math | | | Reading | | |
|---|---|---|---|---|---|---|
| | Elementary (Grades 3-5) | Middle (Grades 6-8) | High School (Grades 9-10) | Elementary (Grades 3-5) | Middle (Grades 6-8) | High School (Grades9-10) |
| *In-Service (Student-Level)Variables* | | | | | | |
| Achievement Gain | 20.358 | 14.529 | 11.847 | 16.835 | 16.053 | -2.639 |
| Std. Dev. of Achiev. Gain | 25.584 | 23.895 | 25.696 | 26.366 | 25.568 | 25.340 |
| Achievement Level | 647.164 | 682.616 | 718.448 | 652.405 | 692.181 | 701.054 |
| Std. Dev. of Achiev. Level | 37.841 | 38.942 | 39.505 | 39.818 | 40.485 | 34.860 |
| Number of Schools Attended | 1.040 | 1.036 | 1.024 | 1.040 | 1.033 | 1.026 |
| "Structural" Mover | 0.010 | 0.229 | 0.315 | 0.010 | 0.194 | 0.405 |
| "Non-Structural" Mover | 0.116 | 0.153 | 0.160 | 0.116 | 0.138 | 0.188 |
| Fraction Female Peers | 0.502 | 0.511 | 0.521 | 0.502 | 0.516 | 0.518 |
| Fraction Black Peers | 0.241 | 0.208 | 0.186 | 0.239 | 0.199 | 0.193 |
| Fraction Mover Peers | 0.126 | 0.381 | 0.474 | 0.126 | 0.331 | 0.593 |
| Fraction "Strc.-Mover" Peers | 0.010 | 0.229 | 0.315 | 0.010 | 0.194 | 0.405 |
| Average Age of Peers (Mo.) | 121.821 | 151.521 | 177.894 | 121.892 | 153.043 | 180.560 |
| Average Class Size | 25.850 | 27.444 | 27.981 | 25.854 | 26.837 | 27.858 |
| Experience (Up to 10 Yrs.) | 2.152 | 2.167 | 1.945 | 2.154 | 2.053 | 1.835 |
| Experience$^2$ (Up to 10 Yrs.) | 13.439 | 13.118 | 11.841 | 13.456 | 12.097 | 10.695 |
| 10+ Years of Experience | 0.422 | 0.385 | 0.459 | 0.422 | 0.388 | 0.431 |
| Total In-service Hours | 52.571 | 44.368 | 35.555 | 52.534 | 49.661 | 41.707 |
| Content In-service Hours | 19.760 | 13.460 | 12.590 | 19.686 | 16.419 | 14.460 |
| Other In-service Hours | 32.811 | 30.908 | 22.965 | 32.848 | 33.265 | 27.247 |
| Advanced Degree | 0.306 | 0.318 | 0.381 | 0.306 | 0.327 | 0.372 |
| *Pre-Service (Teacher/School Spell-Level) Variables* | | | | | | |
| Education Major | 0.949 | 0.793 | 0.622 | 0.950 | 0.587 | 0.451 |
| Math Ed. Major | 0.000 | 0.163 | 0.371 | | | |
| English Ed. Major | | | | 0.000 | 0.237 | 0.267 |
| Math Major | 0.000 | 0.040 | 0.122 | | | |
| English Major | | | | 0.000 | 0.281 | 0.430 |
| SAT Quantitative Score | 468.596 | 511.253 | 536.3012 | | | |
| SAT Verbal Score | | | | 482.622 | 508.376 | 523.922 |
| No. of Obs. (In-service) | 785,611 | 783,903 | 667,389 | 797,525 | 925,268 | 504,036 |
| No. of Obs. (Pre-service) | 2,116 | 854 | 561 | 2,140 | 1,090 | 677 |

**Table 4**
**Iterated OLS Estimates of the Effects of Teacher Experience and In-Service Training on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004**

| | Math | | | Reading | | |
|---|---|---|---|---|---|---|
| | Elementary (Grades 4-5) | Middle (Grades 6-8) | High School (Grades 9-10) | Elementary (Grades 4-5) | Middle (Grades 6-8) | High School (Grades9-10) |
| Experience (Up to 10 Yrs.) | 0.4985*** | 0.2126** | 0.3102*** | 0.4003*** | -0.0619 | 0.0233 |
| | (4.23) | (1.98) | (3.21) | (3.54) | (0.80) | (0.18) |
| Experience$^2$ (Up to 10 Yrs.) | -0.0401*** | -0.0148** | -0.0251*** | -0.0330*** | 0.0110 | -0.0015 |
| | (3.51) | (2.17) | (2.61) | (2.86) | (1.40) | (0.12) |
| Over 10 Yrs. of Experience | 1.6766*** | 1.1320*** | 0.6216*** | 1.7507*** | 0.6141*** | 0.2042 |
| | (6.81) | (5.43) | (3.11) | (10.52) | (3.66) | (0.78) |
| Total In-service Hours$_t$ | -0.0017 | 0.0021 | 0.0002 | 0.0003 | 0.0005 | -0.0025 |
| | (1.32) | (1.63) | (0.14) | (0.25) | (0.47) | (1.64) |
| Total In-service Hours$_{t-1}$ | 0.0034*** | 0.0048*** | -0.0011 | 0.0024** | -0.0006 | 0.0013 |
| | (3.34) | (4.19) | (0.61) | (1.98) | (0.65) | (0.67) |
| Total In-service Hours$_{t-2}$ | -0.0002 | 0.0026* | 0.0014 | -0.0005 | 0.0001 | -0.0029* |
| | (0.22) | (1.82) | (0.76) | (0.36) | (0.10) | (1.64) |
| Total In-service Hours$_{t-3}$ | 0.0009 | 0.0054*** | 0.0032** | 0.0020* | 0.0011 | -0.0019 |
| | (0.73) | (4.13) | (2.29) | (1.70) | (1.11) | (1.04) |
| Student Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher Fixed Effects | No | No | No | No | No | No |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of Students | 514,620 | 542,289 | 426,474 | 525,837 | 594,877 | 348,713 |
| Number of Observations | 785,780 | 784,423 | 667,698 | 797,694 | 925,900 | 504,231 |

Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores. Models include the following time varying student/class characteristics: number of schools attended by the student in the current year, “structural” move by student, “non-structural move” by student, indicator of a student repeating a grade, class size, fraction of classroom peers who are female, fraction of classroom peers who are black, average age (in months) of classroom peers, fraction of classroom peers who changed schools, fraction of classroom peers who made a “structural move.” All models also include year, grade level, and repeater-by-grade dummies. Absolute values of bootstrapped t-statistics, based on 50 repetitions, appear in parentheses. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test.

**Table 5**
**Iterated OLS Estimates of the Effects of Teacher Experience and In-Service Training on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004**

| | Math | | | Reading | | |
|---|---|---|---|---|---|---|
| | Elementary (Grades 4-5) | Middle (Grades 6-8) | High School (Grades 9-10) | Elementary (Grades 4-5) | Middle (Grades 6-8) | High School (Grades9-10) |
| Experience (Up to 10 Yrs.) | 0.4692 | 0.4247 | -0.0012 | 0.4728 | 0.0280 | -0.1002 |
| | (1.37) | (1.72)* | (0.00) | (1.33) | (0.12) | (0.31) |
| $Experience^2$ (Up to 10 Yrs.) | -0.0433 | -0.0523** | 0.0017 | -0.0467 | 0.0050 | 0.0280 |
| | (1.48) | (2.17) | (0.07) | (1.33) | (0.27) | (0.85) |
| Over 10 Yrs. of Experience | 0.8358 | -1.7980 | -0.5147 | 0.7951 | 0.7855 | 2.0019 |
| | (0.50) | (1.44) | (0.38) | (0.45) | (0.60) | (1.01) |
| Total In-service $Hours_t$ | -0.0008 | 0.0035 | -0.0031 | -0.0004 | -0.0005 | -0.0013 |
| | (0.35) | (1.24) | (1.15) | (0.23) | (0.27) | (0.41) |
| Total In-service $Hours_{t-1}$ | 0.0018 | 0.0069*** | -0.0008 | -0.0016 | -0.0011 | 0.0048 |
| | (0.65) | (2.95) | (0.31) | (0.62) | (0.61) | (1.45) |
| Total In-service $Hours_{t-2}$ | -0.0038 | 0.0039* | 0.0042 | -0.0040* | 0.0010 | 0.0011 |
| | (1.32) | (1.70) | (1.49) | (1.72) | (0.53) | (0.35) |
| Total In-service $Hours_{t-3}$ | -0.0013 | 0.0033* | 0.0051** | -0.0002 | 0.0016 | -0.0021 |
| | (0.56) | (1.67) | (2.31) | (0.07) | (0.90) | (0.74) |
| Student Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of Students | 514,620 | 542,289 | 426,474 | 525,837 | 594,877 | 348,713 |
| Number of Observations | 785,780 | 784,423 | 667,698 | 797,694 | 925,900 | 504,231 |

Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores. Models include the following time varying student/class characteristics: number of schools attended by the student in the current year, "structural" move by student, "non-structural move" by student, indicator of a student repeating a grade, class size, fraction of classroom peers who are female, fraction of classroom peers who are black, average age (in months) of classroom peers, fraction of classroom peers who changed schools, fraction of classroom peers who made a "structural move." All models also include year, grade level, and repeater-by-grade dummies. Absolute values of bootstrapped t-statistics, based on 50 repetitions, appear in parentheses. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test.

**Table 6**
**Iterated OLS Estimates of the Effects of Teacher Experience and In-Service Training on Student Math and Reading Achievement in Florida, 1999/2000-2003/2004**

| | Math | | | Reading | | |
|---|---|---|---|---|---|---|
| | Elementary (Grades 4-5) | Middle (Grades 6-8) | High School (Grades 9-10) | Elementary (Grades 4-5) | Middle (Grades 6-8) | High School (Grades9-10) |
| Experience (Up to 10 Yrs.) | 0.4601<br>(1.35) | 0.3864<br>(1.58) | -0.0175<br>(0.06) | 0.4873<br>(1.37) | 0.0300<br>(0.13) | -0.0907<br>(0.28) |
| Experience$^2$ (Up to 10 Yrs.) | -0.0418<br>(1.42) | -0.0480**<br>(2.02) | 0.0039<br>(0.17) | -0.0475<br>(1.35) | 0.0048<br>(0.25) | 0.0243<br>(0.74) |
| Over 10 Yrs. of Experience | 0.8950<br>(0.55) | -1.8130<br>(1.46) | -0.4013<br>(0.30) | 0.8512<br>(0.48) | 0.7969<br>(0.61) | 1.8698<br>(0.94) |
| Content In-service Hours$_t$ | -0.0005<br>(0.11) | 0.0019<br>(0.40) | 0.0051<br>(1.07) | -0.0043<br>(1.04) | 0.0010<br>(0.32) | 0.0040<br>(0.71) |
| Content In-service Hours$_{t-1}$ | 0.0013<br>(0.26) | 0.0123**<br>(2.54) | 0.0066<br>(1.36) | -0.0048<br>(1.14) | -0.0007<br>(0.20) | 0.0181***<br>(2.93) |
| Content In-service Hours$_{t-2}$ | -0. 0099*<br>(1.91) | 0.0248***<br>(5.07) | 0.0087*<br>(1.68) | -0.0070<br>(1.50) | -0.0047<br>(1.08) | -0.0045<br>(0.63) |
| Content In-service Hours$_{t-3}$ | -0.0003<br>(0.07) | 0.0123**<br>(2.25) | 0.0050<br>(1.05) | 0.0022<br>(0.49) | 0.0019<br>(0.38) | -0.0001<br>(0.02) |
| Other In-service Hours$_t$ | -0.0011<br>(0.44) | 0.0045<br>(1.31) | -0.0068**<br>(2.09) | 0.0020<br>(0.70) | -0.0016<br>(0.70) | -0.0062<br>(1.63) |
| Other In-service Hours$_{t-1}$ | 0.0018<br>(0.54) | 0.0056*<br>(1.85) | -0.0043<br>(1.19) | 0.0000<br>(0.01) | -0.0015<br>(0.73) | -0.0024<br>(0.62) |
| Other In-service Hours$_{t-2}$ | -0.0012<br>(0.38) | -0.0025<br>(1.00) | 0.0024<br>(0.75) | -0.0027<br>(0.91) | 0.0020<br>(1.08) | 0.0021<br>(0.67) |
| Other In-service Hours$_{t-3}$ | -0.0015<br>(0.55) | 0.0010<br>(0.44) | 0.0050<br>(1.72) | -0.0005<br>(0.23) | 0.0013<br>(0.67) | -0.0031<br>(0.97) |
| Student Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of Students | 514,620 | 542,289 | 426,474 | 525,837 | 594,877 | 348,713 |
| Number of Observations | 785,780 | 784,423 | 667,698 | 797,694 | 925,900 | 504,231 |

Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores. Models include the following time varying student/class characteristics: number of schools attended by the student in the current year, “structural” move by student, “non-structural move” by student, indicator of a student repeating a grade, class size, fraction of classroom peers who are female, fraction of classroom peers who are black, average age (in months) of classroom peers, fraction of classroom peers who changed schools, fraction of classroom peers who made a "structural move." All models also include year, grade level, and repeater-by-grade dummies. Absolute values of bootstrapped t-statistics, based on 50 repetitions, appear in parentheses. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test.

**Table 7**
**Iterated OLS Estimates of the Effects of Teacher Experience,**
**In-Service Training and Advanced Degrees on Student Math and Reading**
**Achievement in Florida, 1999/2000-2003/2004**

| | Math | | | Reading | | |
|---|---|---|---|---|---|---|
| | Elementary (Grades 3-5) | Middle (Grades 6-8) | High School (Grades 9-10) | Elementary (Grades 3-5) | Middle (Grades 6-8) | High School (Grades9-10) |
| Experience (Up to 10 Yrs.) | 0.4730 | 0.3949 | -0.0252 | 0.5028* | 0.0196 | -0.0808 |
| | (1.46) | (1.59) | (0.10) | (1.70) | (0.09) | (0.27) |
| Experience$^2$ (Up to 10 Yrs.) | -0.0433 | -0.0483** | 0.0044 | -0.0494** | 0.0053 | 0.0231 |
| | (1.44) | (2.04) | (0.17) | (1.96) | (0.30) | (0.84) |
| Over 10 Yrs. of Experience | 0.8589 | -1.7296 | -0.4299 | 0.8220 | 0.6930 | 1.8175 |
| | (0.57) | (1.39) | (0.28) | (0.50) | (0.58) | (0.89) |
| Content In-service Hours$_t$ | -0.0006 | 0.0020 | 0.0047 | -0.0043 | 0.0010 | 0.0041 |
| | (0.18) | (0.49) | (1.07) | (1.01) | (0.46) | (0.75) |
| Content In-service Hours$_{t-1}$ | 0.0011 | 0.0123*** | 0.0064 | -0.0047 | -0.0007 | 0.0179*** |
| | (0.26) | (3.06) | (1.34) | (1.11) | (0.19) | (3.16) |
| Content In-service Hours$_{t-2}$ | -0.0099** | 0.0247*** | 0.0086 | -0.0070 | -0.0047 | -0.0044 |
| | (2.28) | (4.96) | (1.44) | (1.60) | (1.12) | (0.56) |
| Content In-service Hours$_{t-3}$ | -0.0003 | 0.0121** | 0.0050 | 0.0022 | 0.0019 | -0.0001 |
| | (0.06) | (2.34) | (1.21) | (0.44) | (0.38) | (0.01) |
| Other In-service Hours$_t$ | -0.0014 | 0.0047** | -0.0071** | 0.0020 | -0.0015 | -0.0063* |
| | (0.47) | (2.02) | (2.23) | (0.82) | (0.58) | (1.77) |
| Other In-service Hours$_{t-1}$ | 0.0015 | 0.0057* | -0.0043 | 0.0002 | -0.0017 | -0.0025 |
| | (0.55) | (1.95) | (1.30) | (0.09) | (0.78) | (0.63) |
| Other In-service Hours$_{t-2}$ | -0.0015 | -0.0025 | 0.0024 | -0.0027 | 0.0020 | 0.0020 |
| | (0.58) | (0.77) | (0.69) | (1.09) | (0.86) | (0.69) |
| Other In-service Hours$_{t-3}$ | -0.0018 | 0.0011 | 0.0045 | -0.0007 | 0.0011 | -0.0031 |
| | (0.81) | (0.41) | (1.47) | (0.31) | (0.56) | (0.90) |
| Advanced Degree | -0.7088 | 0.8510 | -1.9786*** | -0.5295 | -1.1143** | -0.8008 |
| | (1.54) | (1.48) | (3.68) | (0.95) | (2.20) | (1.00) |
| Student Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of Students | 514,593 | 542,043 | 426,382 | 525,809 | 594,679 | 348,595 |
| Number of Observations | 785,611 | 783,903 | 667,389 | 797,525 | 925,268 | 504,036 |

Since testing begins in grade 3, the grade-4 sample includes only those students who repeated grade 3 or grade 4 and thus have 3 annual test scores. Models include the following time varying student/class characteristics: number of schools attended by the student in the current year, “structural” move by student, “non-structural move” by student, indicator of a student repeating a grade, class size, fraction of classroom peers who are female, fraction of classroom peers who are black, average age (in months) of classroom peers, fraction of classroom peers who changed schools, fraction of classroom peers who made a “structural move.” All models also include year, grade level, and repeater-by-grade dummies. Absolute values of bootstrapped t-statistics, based on 50 repetitions, appear in parentheses. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test.

**Table 8**
**WLS Estimates of the Effects of College Major and College Entrance Exam Scores on a Teacher's "Value-Added" to Student Math and Reading Achievement in Florida, 1999/2000-2003/2004**

| | Math | | | Reading | | |
|---|---|---|---|---|---|---|
| | Elementary (Grades 3-5) | Middle (Grades 6-8) | High School (Grades 9-10) | Elementary (Grades 3-5) | Middle (Grades 6-8) | High School (Grades9-10) |
| Education Major | 0.6179 | -3.4137* | 0.8026 | 2.4382 | -0.1535 | 0.6135 |
| | (0.24) | (1.68) | (0.14) | (0.95) | (0.07) | (0.17) |
| Math Ed./English Ed. Major | | 2.5157 | -0.8950 | 14.0018 | 0.6627 | -1.1140 |
| | | (0.96) | (0.26) | (0.32) | (0.33) | (0.37) |
| Math/English Major | | -2.1675 | 3.2757 | -14.7469 | 0.7581 | 0.5436 |
| | | (0.39) | (0.78) | (0.71) | (0.37) | (0.19) |
| SAT Quant./Verbal Score | 0.0012 | 0.0109 | -0.0288** | -0.0017 | 0.0057 | 0.0004 |
| | (0.18) | (1.11) | (2.24) | (0.26) | (0.75) | (0.04) |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| R-squared | 0.542 | 0.519 | 0.674 | 0.502 | 0.435 | 0.557 |
| Number of Observations | 2,134 | 943 | 639 | 2,158 | 1,252 | 763 |

The dependent variable is the teacher-school spell fixed effect estimated from a model of student achievement using all Florida public school students in the relevant grades. Since testing begins in grade 3, the elementary school samples only include third graders who repeated a grade and thus have an achievement gain score for third grade. Observations are weighted by the square root of the number of students per teacher/school spell. Absolute values of t-statistics appear in parentheses. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test.

**Table 9**
**WLS Estimates of the Effects of College Course Work and College Entrance Exam Scores on a Teacher's "Value-Added" to Student Math and Reading Achievement in Florida, 1999/2000-2003/2004**

| | Math | | | Reading | | |
|---|---|---|---|---|---|---|
| | Elementary (Grades 3-5) | Middle (Grades 6-8) | High School (Grades 9-10) | Elementary (Grades 3-5) | Middle (Grades 6-8) | High School (Grades9-10) |
| Gen. Educ. Theory Credits | -0.3642<br>(0.33) | -1.4788<br>(1.03) | 0.1450<br>(0.06) | -1.0579<br>(0.95) | -3.1265**<br>(2.35) | -0.9648<br>(0.67) |
| Pedagogical - Instructional Credits | -1.1297<br>(1.23) | 0.6340<br>(0.53) | 2.1730<br>(1.19) | -1.6526*<br>(1.77) | 1.9494*<br>(1.84) | 1.9512<br>(1.23) |
| Pedagogical - Management Credits | -2.5321<br>(0.26) | 8.8258<br>(0.88) | | -5.7047<br>(0.44) | -2.8698*<br>(1.84) | |
| Pedagogical - Content Credits | 1.4643*<br>(1.95) | 1.7628*<br>(1.87) | -0.7593<br>(0.38) | 0.1434<br>(0.15) | 0.3295<br>(0.29) | -1.9271<br>(0.80) |
| Professional Development Credits | -1.4492<br>(0.90) | 2.2154<br>(0.83) | 2.4478<br>(1.49) | 1.0245<br>(0.66) | -0.8533<br>(0.49) | 1.0988<br>(0.56) |
| Classroom Practice Credits | -1.0555<br>(1.12) | -3.4134**<br>(2.48) | -0.6008<br>(0.25) | 0.4059<br>(0.43) | -1.2762<br>(0.90) | -0.0122<br>(0.01) |
| Subject Content Credits | 0.6501<br>(0.27) | 5.2721<br>(1.50) | 2.3557<br>(0.39) | 2.3081<br>(0.97) | 2.4178<br>(0.62) | -12.1705<br>(1.56) |
| Mathematics Credits | -0.0927<br>(0.68) | 0.4298<br>(1.45) | -0.8822**<br>(2.21) | | | |
| Statistics Credits | -4.1344*<br>(1.66) | -3.3730<br>(1.31) | 1.3495<br>(0.49) | | | |
| English Literature Credits | | | | -0.1049<br>(0.41) | 0.3235<br>(1.29) | 0.8873**<br>(2.38) |
| Math Education Credits | 0.4657<br>(0.32) | -0.9482<br>(0.72) | 0.1384<br>(0.09) | | | |
| Language Arts Educ. Credits | | | | 0.4745<br>(0.30) | 0.6993<br>(0.69) | 0.8109<br>(0.54) |
| SAT Quant./Verbal Score | -0.0056<br>(0.96) | 0.0034<br>(0.41) | -0.0190*<br>(1.73) | 0.0044<br>(0.80) | 0.0085<br>(1.21) | -0.0009<br>(0.11) |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| R-squared | 0.493 | 0.468 | 0.631 | 0.452 | 0.426 | 0.535 |
| Number of Observations | 2,684 | 1,124 | 795 | 2,732 | 1,355 | 833 |

The dependent variable is the teacher-school spell fixed effect estimated from a model of student achievement using all Florida public school students in the relevant grades. Since testing begins in grade 3, the elementary school samples only include third graders who repeated a grade and thus have an achievement gain score for third grade. Observations are weighted by the square root of the number of students per teacher/school spell. Absolute values of t-statistics appear in parentheses. * indicates statistical significance at the .10 level and ** indicates significance at the .05 level and *** indicates significance at the .01 level in a two-tailed test.