DO MORE EFFECTIVE TEACHERS EARN MORE OUTSIDE THE CLASSROOM?

Matthew M. Chingos

Brown Center on Education
    Policy
Brookings Institution
Washington, DC 20036
mchingos@brookings.edu

Martin R. West

(corresponding author)
Harvard Graduate School
    of Education
Cambridge, MA 02138
martin_west@gse.harvard.edu

Abstract

We examine earnings records for more than 130,000 classroom teachers employed by Florida public schools between the 2001–2 and 2006–7 school years, roughly 35,000 of whom left the classroom during that time. A majority of those leaving the classroom remained employed by public school districts. Among teachers in grades 4–8 leaving for other industries, a 1 standard deviation increase in estimated value added to student math and reading achievement is associated with 6–8 percent higher earnings outside teaching. The relationship between effectiveness and earnings is stronger in other industries than it is for the same groups of teachers while in the classroom, suggesting that current teacher compensation systems do not fully account for the higher opportunity wages of effective teachers.

1. INTRODUCTION

Every debate about teacher compensation policy—be it overall salary levels, merit pay, or bonuses for working in high-need schools—draws heavily on arguments about what current teachers could be making in other sectors of the economy (their opportunity wages).[1] Discussions about how to recruit and retain the best teachers similarly emphasize the need to ensure that teaching positions are attractive to the most effective current and potential members of the profession. Unfortunately, very little evidence is available regarding the nonteaching job opportunities available to teachers in general and to specific groups of teachers defined in terms of their effectiveness in promoting student achievement. This article seeks to fill that gap using a unique administrative data set that links the experiences of current and former teachers, including the academic performance of their students, to the earnings records of those same teachers.

Theory is ambiguous as to the expected relationship between value-added measures (VAMs) of teacher effectiveness and earnings outside education. On one hand, to the extent that such measures accurately capture teacher effectiveness and teacher effectiveness reflects general skills that enhance productivity in other occupations, high-value-added teachers who leave for other jobs should earn more than their low-value-added counterparts. On the other hand, VAMs are at best noisy indicators of teacher quality and may be biased by nonrandom sorting of students and teachers (Rothstein 2010). In addition, it is commonly argued that good teaching requires a highly idiosyncratic skill set that may not translate into success in other fields, a contention that is supported by the difficulty of predicting teacher effectiveness with such measures of general skills as SAT scores, college selectivity, and tests of cognitive ability (Kane, Rockoff, and Staiger 2007; Harris and Sass 2007; Rockoff et al. 2008) despite the strong power of such measures to predict wages in the labor market as a whole (see, e.g., Murnane et al. 2001 on the labor market returns of cognitive skills).[2] By examining the relationship between value added while teaching and subsequent earnings in other industries, we therefore provide an external (to public education) validation of VAMs and shed light on the competitive pressures facing school districts due to realities of the outside labor market.

---

1. See, for example, Ballou and Podgursky 1997; Boardman, Darling-Hammond, and Mullin 1982; Murnane et al. 1991; and Podgursky and Tongrut 2006.
2. However, Ferguson and Ladd (1996) do find that ACT exam scores predict teacher effectiveness. And even if teacher effectiveness is uncorrelated with measures such as SAT scores, value added could still be correlated with nonteaching earnings if employers observe general noncognitive skills that are useful in both industries (e.g., motivation or work ethic) but are not correlated with SAT scores.

Existing research supports the notion that alternative labor market opportunities affect the decisions of teachers to enter and remain in the classroom. Dolton and van der Klaauw (1995, 1999) show that teacher salaries and opportunity wages in the United Kingdom strongly influence the decisions of a national sample of early-career teachers to remain in the profession. Earlier research by Murnane and Olsen (1989, 1990) also found that opportunity wages affected teachers' career lengths in Michigan and North Carolina. More recently, Hoxby and Leigh (2004) provide evidence that the post-1960 decline in the aptitude of female teachers in the United States reflected union-induced compression of wages with respect to aptitude within teaching. However, each of these articles derives estimates of teachers' opportunity wages based on their observed characteristics and geographic location. Although this approach avoids concerns about selection bias that arise when using observed wages of teachers who leave the profession, it means that none of them has direct information on the earnings of leaving teachers or on those teachers' classroom effectiveness.

Two recent articles use state administrative data sets to examine the earnings of teachers leaving the classroom for other occupations. Podgursky, Monroe, and Watson (2004) merge employment data on Missouri teachers with earnings information from the state's unemployment insurance system, while Scafidi, Sjoquist, and Stinebrickner (2006) do the same for Georgia. Both sets of authors conclude that, contrary to common perceptions, very few exiting teachers take jobs that pay more than their prior salaries as teachers.[3] Neither analysis, however, contains information on the classroom effectiveness of leaving teachers.

We contribute to this literature by providing what is, to our knowledge, the first evidence on the relationship between the earnings of exiting teachers and their estimated effectiveness while teaching. More specifically, we follow the careers of more than 130,000 classroom teachers employed by Florida public schools between the 2001–2 and 2006–7 school years, roughly 35,000 of whom left the classroom during that time. Uniform quarterly earnings data are available from the state's unemployment insurance system for all current teachers and for those former teachers who worked elsewhere in the state. We are therefore able to observe the industries in which departing teachers worked and their average annual earnings in their new careers.

---

3. Stinebrickner (2002) performs a similar exercise for an oversample of teachers included in the National Longitudinal Study of the High School Class of 1972, but his ability to draw inferences about the current labor market for teachers is limited by the relatively small number of teachers in the data set and by the fact that most of the teachers in his data graduated from college in 1976.

For math and reading teachers in grades 4–8, we use information on their students' performance on state tests to estimate VAMs of classroom effectiveness. This information makes it possible to compare the relationships between classroom effectiveness and earnings inside and outside teaching for the common samples of teachers who left the classroom for other positions in public school districts or for another industry altogether.

Our analysis yields several noteworthy results. First, a majority of those leaving the classroom remained employed by public school districts, suggesting that a substantial amount of attrition from classroom teaching reflects movement into administrative or other nonteaching positions within the public school system. The growth in the number of administrative and support positions in American school systems in recent decades is well documented in the academic literature (Hanushek and Rivkin 1997). The latest data from the U.S. Department of Education (2008) indicate that the ratio of students to total full-time equivalent staff employed by public school systems fell by more than 41 percent (from 13.6 to 8.0) between 1970 and 2006, including a 13 percent reduction after 1990.[4] Our finding that many former teachers continue working in the public schools suggests that this pattern may well have led to higher rates of attrition from classroom teaching.

The median annual earnings of teachers moving into other sectors of the Florida economy fell by more than 20 percent upon leaving teaching, with the largest drops observed for females and teachers in grades K–8. These large overall declines, however, include a substantial share of former teachers who appear to have moved into part-time work. Among former teachers who were likely working full time, male teachers saw a slight increase in median earnings upon leaving for another industry, while female teachers experienced only a modest decline. However, the earnings distribution for former teachers is more dispersed than their earnings while teaching, suggesting substantial wage compression within the teaching profession.

Most important, we show that teachers who were more effective in the classroom (as measured by their students' test score gains) earn more in other industries than other teachers who left the profession. Among full-time workers, a 1 standard deviation increase in estimated value added is associated with 6–8 percent greater earnings outside teaching. The relationship between classroom effectiveness and earnings is consistently stronger outside teaching than for the same groups of teachers while in the classroom. The patterns strongly

---

4. The number of teachers hired for each student also rose during this period but at a slower rate than the number of total staff. The number of students per full-time equivalent teacher decreased by 23 percent between 1970 and 2006 and by 10 percent after 1990.

suggest that existing teacher compensation systems do not fully account for differences in short-term opportunity wages among current teachers.

The remainder of the article is organized as follows. The next two sections introduce our data and methods. Section 4 compares the earnings distribution of teachers who left the classroom for other jobs in public school districts and for jobs outside the public education system with their earnings while teaching. Section 5 compares the relationship between teacher value added and earnings inside and outside teaching for fourth- through eighth-grade teachers. Section 6 discusses the implications of our results for teacher compensation policy.

2. DATA

The central innovation in our study is the linking of information from the Florida Department of Education's K–20 Education Data Warehouse (EDW) with earnings records from the Florida Education and Training Placement Information Program (FETPIP). Our EDW data extract contains observations of every student in Florida who took the state assessment tests from 1998–99 to 2008–9, with each student linked to his or her courses (and corresponding teachers) for 2001–2 through 2008–9. The FETPIP data consist of state unemployment compensation records that include the earnings of current public schoolteachers and former teachers working in Florida from the first quarter of 2001 through the third quarter of 2008.

The EDW data include test score results from the Florida Comprehensive Assessment Test (FCAT), the state accountability system's "high-stakes" test, and the Stanford Achievement Test (SAT), a nationally norm-referenced test that was administered to students alongside the FCAT until 2008 but was not used for accountability purposes. Beginning in 2001, students in grades 3–10 took both tests each year in math and reading. Thus annual gain scores can be calculated for virtually all students in grades 4–10 beginning in 2002. The data also contain information on the demographic and educational characteristics of each student, including gender, race, free or reduced price lunch eligibility, limited English proficiency status, special education status, days in attendance, and age.

The EDW data also contain detailed information on individual teachers, including their demographic characteristics and experience. We construct an employment file based on course enrollment data (which matches students and teachers) in order to track whether and where teachers were employed in classroom teaching positions in a given year. A separate file from the EDW enables us to identify the specific jobs of former classroom teachers who remained employed in the public schools. The teacher experience variable we construct reflects all years the teacher has spent in the profession, including both public and private schools in Florida and other states. Given our interest

in the subsequent earnings of exiting teachers, we restrict our analysis to observations of teachers who were fifty-four years of age or younger in order to exclude likely retirees.[5]

The FETPIP data indicate, for each teacher, quarter, and employer, the amount earned, the randomly generated employer ID number, the number of workers at the firm, and the employer's North American Industry Classification System (NAICS) code. Only earnings received in Florida and reported to the state unemployment compensation system are included in these records.

We inflate all quarterly earnings to 2008 dollars using the average of the monthly Consumer Price Indexes (CPIs) for each quarter. We then calculate each current and former teacher's earnings in each school year from 2001–2 through 2007–8 by combining the earnings from the last quarter of the calendar year in which the school year began with the earnings from the first three quarters of the calendar year in which the school year ended. For example, earnings for 2001–2 are calculated as earnings from October 2001 through September 2002.[6] In order to exclude likely part-time teachers, we include only teachers who earned at least $20,000 from the public schools in at least one year that they were a classroom teacher.[7]

We divide this sample of classroom teachers into three groups: those who remained as teachers during our observation period, those who left teaching for other jobs within Florida public school districts, and those who left teaching for other nonteaching jobs in the state.[8] There are also a substantial number of teachers who left teaching but do not appear in the FETPIP wage data (because they withdrew from the labor force, became self-employed, or left Florida); these teachers are included in the descriptive statistics presented in appendix

---

5. Most Florida teachers become eligible for retirement with full pension benefits at age sixty-two or after thirty years of service. Assuming an entry age of twenty-five, the earliest possible retirement age with full benefits is therefore fifty-five.

6. The timing of the collection of the quarterly earnings data requires that we include September in the wrong school year, introducing a modest amount of measurement error.

7. Our measure of annual earnings combines income reported from all sources, including side jobs of teachers taking on additional employment. Fortunately such "moonlighting" activity appears to be relatively rare: among teachers remaining in the classroom throughout our sample period, mean annual earnings from sources other than public school districts were $657, or 1.4 percent of their mean total earnings of $46,014. Teachers leaving the classroom for other positions in public school districts had mean earnings from outside sources of $795, or 1.8 percent of their mean total earnings of $43,773. Teachers leaving the classroom for employment elsewhere earned somewhat more from outside sources while teaching, but the $2,267 in mean earnings they received from outside sources still only amounted to 5.8 percent of their mean total earnings of $39,161. Excluding earnings received from sources other than public school districts does not change the earnings-effectiveness patterns for teachers.

8. We identify whether a teacher was employed as a teacher in a given year using the student course enrollment files. We then use the FETPIP employer codes to identify teachers who left teaching but remained employed by a public school district. A former teacher is defined as working in public schools in a given academic year if he received the majority of his earnings from a public school district. Teachers who, during their nonteaching years, earned the majority of their earnings in public schools in some years and elsewhere in others are included in both groups of leavers (with only the relevant years of earnings considered for each group).
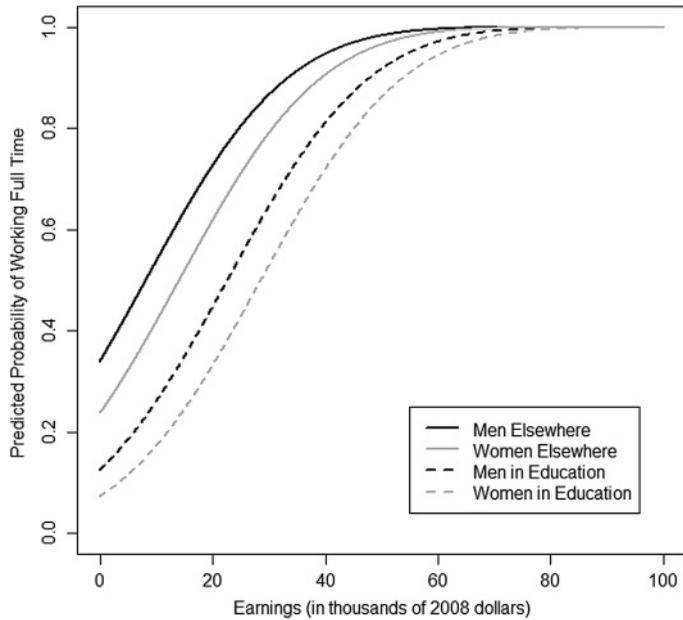
Figure 1. Predicted Probability of Working Full Time

table A.1 but are excluded from our main analyses. In order to allow for leaves of absence as well as transitions between teaching and nonteaching jobs (in both the public schools and elsewhere), we do not count either in-teaching or nonteaching earnings experienced during an individual's first year outside the classroom.[9] Teachers who are observed for only one year immediately after leaving the classroom are therefore excluded from the analysis. We also exclude all teachers who first appear in the last two years of our data (2006–7 or 2007–8) because such teachers could never be observed beyond their first year of nonteaching earnings.

Finally, in much of our analysis we weight former teachers by their estimated probability of working full time. These predicted probabilities were calculated by conducting a probit analysis of whether former teachers in the nationally representative 2004–5 Teacher Follow-up Survey (TFS) reported that they were working full time.[10] Predictor variables included in the regression were total annual earnings (inflated to 2008 dollars), gender, and whether they were working in the education sector or elsewhere. Figure 1 plots the predicted

---

9. Among teachers who were in the classroom in 2001–2 but not in 2002–3, 21 percent had returned to the classroom in 2003–4, suggesting that a sizable number of apparent leavers return after just one year of absence.

10. We excluded as potential retirees former teachers fifty-five years and older (as we do in our analysis of the Florida data), leaving us with a sample of 615 former teachers, 383 in education and 232 in other industries.

probabilities as a function of annual earnings, gender, and whether the teacher remained in the field of education.

3. ANALYTIC STRATEGY

Our primary aim is to examine how teachers' opportunity wages vary with their effectiveness. We first explain our approach to measuring teacher effectiveness and describe the models used to relate these measures to current and subsequent earnings. We then discuss the implications for the interpretation of our results of sample selection issues and the fact that we are able to track teachers for only a short period of time after they leave the classroom.

Measuring Teacher Effectiveness

To measure teacher effectiveness we use student test score data to construct VAMs for teachers of math and reading courses in grades 4–8.[11] Value-added measures attempt to isolate the causal effect that teachers have on their students' test scores.[12] Because the limited time period over which we observe teachers leads to considerable error in measuring their effectiveness, we focus our analysis on VAMs that average estimates based on the high-stakes FCAT and the low-stakes SAT.[13] However, all findings reported below concerning the relationship between this summary effectiveness measure and earnings are qualitatively similar to those obtained using measures based only on the FCAT or only on the SAT.[14]

We use the course files to match fourth- and fifth-grade students (most of whom are enrolled in self-contained classrooms) to their primary teacher and sixth- through eighth-grade students to their math and reading/English language arts teachers.[15] A very small number of fourth- and fifth-grade students

---

11. Although Florida also tests students annually in grades 9 and 10, the wide variety of math and reading course offerings in these grades makes it difficult to construct reliable VAMs.

12. Of course, teacher quality has other dimensions, most of which are difficult to measure and all of which are outside the scope of our analysis. The interpretation of our main findings hinges on the assumption that the test instruments used are on average a reasonable proxy for students' overall academic development.

13. We find that SAT value added in the current year is a significant predictor of FCAT value added in the following year conditional on FCAT value added in the current year, suggesting that averaging the two measures reduces measurement error.

14. The correlation coefficient between effectiveness of teachers as measured by their students' FCAT performance and effectiveness as measured by the SAT is 0.70 and 0.61 for fourth and fifth grade math and reading, respectively, and 0.66 and 0.53 for sixth through eighth grade math and reading. Separate estimates of the relationship between the SAT-based and FCAT-based teacher effectiveness measures and earnings are available from the authors upon request.

15. For fourth- and fifth-grade students, the course files do not always clearly identify the student's regular classroom teacher. In order to match the maximum number of students to their teachers, we examined students' general (e.g., self-contained classroom), math, and reading teachers and matched them to the one or two teachers with whom they spent at least 40 percent of their academic (general, math, and reading) time. We then dropped students who were matched to two teachers and students who were not matched to any teachers (32 percent of all students in these

who were in classrooms of fewer than five or more than forty students are dropped from the analysis.

To generate value-added estimates for each teacher, we regress students' math and reading test scores separately on their prior-year test scores (including squared and cubed terms); vectors of student, classroom, and school characteristics; dummy variables for teacher experience; and grade-by-year fixed effects.[16] Additional student-level control variables include the number of days absent the previous year and dummy variables for race, gender, limited English proficiency status, special education status, migrant status, free or reduced price lunch eligibility, whether the student was repeating the grade, and whether the student made a structural or nonstructural move to a new school that year.[17] Classroom- and school-level control variables include all the student-level characteristics (except prior-year test scores) aggregated to the appropriate level and class size.

The model, then, is

$$A_{it} = \omega A_{i,t-1} + \beta X_{it} + \gamma C_{it} + \phi S_{it} + \delta W_{it} + \pi_{it} + \varepsilon_{it}, \tag{1}$$

where A_{it} is the test score of student i in year t (standardized by grade and year to have a mean of zero and standard deviation of 1); $A_{i,t-1}$ includes the student's prior-year test scores in both subjects (and their squared and cubed terms); X, C, and S are student-, classroom-, and school-level characteristics; W is a vector of teacher experience dummy variables; π is a vector of grade-by-year fixed effects; and ε is a standard zero-mean error term. Although it is not common to include experience as a control variable when estimating teacher value added, we do so because considerable evidence indicates that teachers improve in their first few years in the classroom (see, e.g., Rockoff 2004), and we are interested in teachers' eventual ability in the profession rather than their effectiveness in a particular year.[18]

---

grades). A large and increasing number of fourth- and fifth-grade students in Florida appear to have more than one regular classroom teacher, perhaps reflecting an increase in team teaching. We match sixth- through eighth-grade students to their primary reading and math teachers in a similar fashion. For each subject, we linked each student to the teacher(s) with whom they spent at least 40 percent of their time in that subject and dropped students who were matched to two teachers in a given subject as well as those who were not matched to any teachers in that subject (16 percent of students in math and 38 percent of students in reading).

16. We control for teacher experience using a dummy variable for each of the first twenty years of experience, so the omitted category includes all teachers with more than twenty years of experience.

17. Students are identified as having made a structural move from the previous year if they are in a different school than in the previous year and are in the lowest grade offered in their new school, and as having made a nonstructural move if they are in a grade higher than the minimum grade of the new school.

18. Using a value-added model that does not control for experience actually yields slightly larger coefficients in the earnings regressions reported below. These results are included in the online appendix, which can be accessed at the *Education Finance and Policy* Web site (www.mitpressjournals.org/efp).

We estimate equation 1 separately by test (FCAT and SAT), subject (reading and math), and grade level (grades 4 and 5 and 6–8) and average the residuals by teacher and year to construct a VAM for each teacher in each year. Finally, we use the Bayesian (shrinkage) estimator described by Kane, Rockoff, and Staiger (2007) to isolate the persistent component of each teacher's value added.[19] This persistent component forms the effectiveness measure used throughout our analysis. Consistent with previous research, we find that teacher effectiveness varies more for math than for reading and elementary- than for middle-school teachers.[20]

Although widely used by researchers, the reliability of this kind of value-added model of teacher effectiveness using nonexperimental data continues to be debated (see, e.g., Rothstein 2010; Kane and Staiger 2008). The key potential confounding factor is the nonrandom matching of students and teachers both across and within schools, which would bias estimated teacher effects if there were unobserved differences across students that were not accounted for by the variables described above.[21]

It is unclear, however, whether or in what direction the nonrandom matching of students and teachers would bias our estimates of the relationship between effectiveness and earnings. Classical measurement error will of course attenuate that relationship—and should do so by similar amounts when comparing the wage structure inside and outside teaching for a common sample of teachers. It is not obvious a priori whether any additional error due to nonrandom matching would be systematically related to opportunity wages. As a robustness check, we therefore examine whether nonteaching earnings are systematically related to the average prior achievement of students assigned to each teacher (relative to other teachers within the same school) and whether our estimates of the effectiveness-earnings relationship are sensitive to controlling for this measure of student-teacher sorting.

19. One key difference is that for each year and teacher, Kane, Rockoff, and Staiger (2007) compute average residuals by class, whereas we compute average residuals by teacher (which is identical to class for fourth- and fifth-grade teachers because each teacher has only one class, but sixth- through eighth-grade teachers often teach multiple classes). We do this because the nature of the EDW course records makes it difficult to definitively assign middle school students to a specific math classroom even though we can confidently match them to a specific math teacher.

20. For all teachers for whom we are able to estimate effectiveness measures using FCAT scores, the standard deviations of these measures are (in standard deviations of student test scores) 0.11 and 0.05 for fourth and fifth grade math and reading, respectively, and 0.07 and 0.03 for sixth through eighth grade math and reading, respectively. The standard deviations of measures calculated using SAT scores are 0.08 and 0.05 for fourth and fifth grade math and reading, and 0.06 and 0.03 for sixth through eighth grade math and reading.

21. One strategy to reduce bias from nonrandom matching is to use school fixed effects to restrict the analysis to differences in teacher effectiveness within schools. We present models that control for school fixed effects and which also limit our analysis of the relationship between effectiveness and nonteaching earnings to teachers who left the same school as checks on the robustness of our preferred results.

Relating Effectiveness and Earnings

We use data at the teacher $*$ year level to estimate the relationship between classroom effectiveness and earnings both in and out of teaching. Our VAMs of teacher effectiveness are standardized separately by grade level (grades 4–5 and grades 6–8), test (FCAT and SAT), and subject area (math and reading) to have a mean of zero and a standard deviation of 1. Because most teachers in grades 4 and 5 teach both math and reading in self-contained classrooms, we average their VAM in both subjects. We also average the standardized VAMs from both the FCAT and SAT tests to reduce measurement error.

Combining elementary- and middle-school teachers, we estimate the following equation:

$$\log(earn_{it}) = \alpha + \lambda * VA_i + \phi_i + \varepsilon_{it}, \tag{2}$$

where $\log(earn_{it})$ is the natural logarithm of annual earnings for teacher i in academic year t, α is a constant, VA is the teacher's standardized VAM (which does not vary over time because it is calculated using all available data), ϕ is a set of dummy variables corresponding to the teacher's grade level and subject (middle school math or middle school English/reading, with elementary school teachers making up the omitted category), and ε is a zero-mean error term. Standard errors are adjusted for clustering at the teacher level to account for the fact that teachers appear multiple times in the data set for two reasons: (1) most teachers are observed in more than one year, and (2) some teachers are assigned to classrooms in more than one grade level and subject area.[22] We estimate equation 2 separately for each group of teachers (stayers, leavers who remain in public school districts, and leavers to other industries). For both groups of leavers, we estimate the model separately for the years in which they were teaching and those in which they were in other jobs.

We also estimate specifications of equation 2 that successively include district fixed effects and controls for teacher demographics. Because Florida's countywide school districts approximate local labor markets, district fixed effects should help eliminate any lingering bias in our measures of teacher

---

22. The standard errors we report do not account for the fact that teacher value added is an estimate. The appropriate solution to this issue would be to bootstrap the standard errors by repeatedly redrawing the coefficients from the value-added model in order to allow the effectiveness estimates to vary in each iteration (and also sampling with replacement from the set of teachers included in the model). However, it is not computationally feasible to calculate standard errors in this way for all our models. To confirm that this issue is not causing us to substantially understate our standard errors, we ran 250 replications for one subject-grade-test combination (FCAT math in grades 4 and 5). The uncorrected coefficient on value added in equation 2 for this subject-grade-test is 0.081 with a standard error of 0.024. The bootstrapped results yield a coefficient of 0.085 with a standard error of 0.026. Clearly the uncorrected standard errors reported in the article only slightly understate the standard errors that would be produced using the bootstrapping procedure.

effectiveness associated with geographic differences in job opportunities. The results of models including district fixed effects represent our preferred estimates, as they should provide policy makers with the best information on how opportunity wages vary with teacher effectiveness. Results including teacher demographic characteristics, in contrast, allow us to see whether any observed relationships between effectiveness and earnings reflect differences in job opportunities associated with those characteristics.

Sample Selection Issues

Ideally we would estimate equation 2 for a large sample of teachers who left teaching for reasons exogenous to their opportunity wages (e.g., seniority-based layoffs due to budget cuts). Unfortunately that is not possible in this study because Florida was rapidly expanding its teaching labor force during the period covered by our data due to enrollment growth and the implementation of a statewide class-size reduction policy (West and Chingos 2009). Our analysis is therefore based on the nonrandom sample of teachers who chose (or were forced) to leave the classroom during this period. Moreover, our reliance on the FETPIP data means that we are able to observe only the earnings of those exiting teachers who remained in Florida's formal labor market. Both these forms of sample selection could potentially influence our results on the relationship between classroom effectiveness and opportunity wages.

Selection in leaving teaching could influence our findings in two ways. First, to the extent that the relationship between effectiveness and opportunity wages differs for leavers, our findings may not generalize to the full population of teachers. Given the widespread interest among policy makers in improving teacher retention, however, those who elect to leave the profession are a highly relevant population to study. Moreover, we confirm below that the teachers we observe working in other industries are quite similar to the population of potential leavers in terms of their observable characteristics, including their estimated value added. We also show that the relationship between effectiveness and earnings while teaching is similar for stayers and leavers.

Selection in leaving teaching could also bias our estimates of the relationship between effectiveness and wages in other industries among leavers. In particular, one might worry that any observed relationship between effectiveness and nonteaching earnings is an artifact of higher-quality teachers leaving voluntarily if they find a better job opportunity and less effective teachers leaving involuntarily (and thus suffering a displacement loss). However, anecdotal evidence suggests that performance-based involuntary dismissals from public school teaching are exceedingly rare, and we show below that leavers are on average only slightly less effective than those who remain. We also address this concern empirically by looking separately at teachers who have tenure, for

whom virtually all leaving should be voluntary. Finally, we show that we obtain very similar results from a parallel analysis of teachers entering the classroom after having previously been employed in other industries. Although this is also a selected sample, the fact that we obtain similar findings for an entirely different set of teachers is reassuring.

It is theoretically unclear how selection in appearing in the FETPIP earnings data would influence our results, but it would introduce bias only if the relationship between effectiveness and opportunity wages differed for teachers whose earnings we do not observe. As documented below, nearly three-quarters of exiting teachers appear with positive earnings in at least one year, suggesting that the potential scope for bias is relatively small. This share far exceeds the 50 percent of exiting teachers in the nationwide TFS sample who report working either full time or part time in their first year after leaving, perhaps because it includes some teachers with very low earnings who would have reported to the TFS that they were not working. Interestingly, teachers leaving districts bordering other states are only 2 percentage points less likely to appear in the earnings data than teachers leaving other Florida districts. Based on these analyses, we conclude that few exiting teachers do not appear in the earnings data because they have left the state.

A final interpretive issue involves the relatively short length of time for which we are able to observe teachers' earnings after leaving the profession. Given our data, the longest we are able to observe the earnings of any exiting teacher is six years after leaving the classroom. It is of course likely that people are forward looking and would consider the entire future lifetime earnings in other industries when making occupational choices, not simply their earnings in the first few years (see, e.g., Willis and Rosen 1979). To the extent that earnings increase more over time in other industries than in public education, our evidence on the earnings of leaving teachers may understate their long-term opportunity wages. More importantly, our results concerning the relationship between effectiveness and earnings could be misleading to the extent that this relationship grows stronger or weaker over time. We therefore also discuss results of analyses in which we interact effectiveness with the length of time after which the teacher has left the classroom in order to discern whether that relationship could be expected to change over a longer period.

4. INDUSTRY CHOICE AND MEDIAN EARNINGS AMONG LEAVING TEACHERS

We begin our empirical analysis by examining the overall rate of attrition among Florida classroom teachers, the new industries chosen by those leaving the classroom, and their median earnings. After applying the exclusion restrictions discussed above, our analytic sample includes roughly 131,000

teachers who taught core academic subjects (defined as general, math, and reading in grades K–5 and general, math, reading, social studies, science, and foreign language in grades 6–12) in grades K–12.

The overall attrition rate among Florida teachers in these subjects who are not yet at retirement age is relatively modest. Among teachers who were in the classroom in the 2001–2 school year, for example, 91 percent were teaching in 2003, 83 percent in 2004, 78 percent in 2005, 73 percent in 2006, 70 percent in 2007, and 66 percent in 2008 (hereafter we often refer to school years using the calendar year of the spring semester). The analogous numbers for those employed in 2003 are similar: 90 percent in 2004, 83 percent in 2005, 77 percent in 2006, 74 percent in 2007, and 68 percent in 2008. The share of teachers remaining in the classroom in a given year includes both individuals who taught continuously and those returning from temporary leaves of absence. The steadily declining percentages for both cohorts therefore indicate that the number of leaving teachers each year exceeds the number returning from a leave.

Table 1 reports the destination industries of teachers leaving the classroom for more than one consecutive year between 2002 and 2008 using a set of categories based on the North American Industry Classification System (NAICS). Exiting teachers with multiple employers are assigned to the industry in which they received the majority of their earnings.[23] The analysis is necessarily limited to former teachers who had reported earnings in Florida and therefore excludes those withdrawing from the labor market and those moving to another state. The percentage of all exiting teachers for whom earnings data are available is 74 percent, as shown in the last column of table 1. It is higher for males than females at the elementary school level. Appendix table A.1, which compares the observed characteristics of exiting teachers who did not have reported earnings in Florida with those who did, indicates that these teachers are less likely to be male, black, or Hispanic and are modestly younger on average.

Perhaps surprisingly, a majority of leavers across all grade levels for whom earnings data are available continue to draw most of their earnings from public school districts. The modal such teacher (41 percent) moved to a position as a specialist or aide assigned to specific student populations (e.g., in special education or Title I). Substantial proportions also became principals or assistant principals (16 percent) or entered positions supporting classroom teachers, such as curriculum or instruction specialists and subject coaches

---

23. The tiny fraction (less than 0.5 percent) of teachers that did not earn the majority of earnings from a single industry are included in the "Other" category.

Table 1. Destination Industries of Former Florida Teachers, 2002–8

	Public Elem/Sec Schools	Private Elem/Sec Schools	Other Education	Professional Services	Health Care	Other Services	Public Admin.	Other or Multiple	N	Workers as % of Leavers
All K–8 teachers	63%	4%	6%	10%	6%	3%	3%	6%	18,774	74%
All male K–8 teachers	52%	3%	6%	14%	5%	4%	6%	10%	3,266	78%
All female K–8 teachers	65%	5%	5%	9%	6%	3%	2%	5%	15,508	73%
All 9–12 teachers	50%	4%	10%	13%	6%	4%	5%	8%	7,164	74%
All male 9–12 teachers	44%	4%	10%	16%	5%	4%	6%	11%	2,696	74%
All female 9–12 teachers	54%	4%	10%	11%	7%	4%	4%	7%	4,468	74%

Notes: Categories are based on North American Industry Classification System (NAICS) codes. Other Education includes all Educational Services other than Elementary and Secondary Schools. Professional Services includes Information; Finance and Insurance; Real Estate and Rental and Leasing; Professional, Scientific, and Technical Services; Management of Companies and Enterprises; and Administrative Support and Waste Management and Remediation Services. Health Care includes Health Care and Social Assistance. Other Services includes Arts, Entertainment, and Recreation; Accommodation and Food Services; and Other Services (except Public Administration). "Other" includes Agriculture, Forestry, Fishing, and Hunting; Mining, Quarrying, and Oil and Gas Extraction; Utilities; Construction; Manufacturing; Wholesale Trade; Retail Trade; and Transportation and Warehousing.

(14 percent). The share of former classroom teachers remaining employed by districts is highest for elementary and middle grade teachers (63 percent) and lowest among high school teachers (50 percent). Females who remain in the state's workforce are between 10 and 13 percentage points more likely to remain employed by districts than males, depending on the grade level. Given the large share of teachers who remained employed by public school districts, the remainder of our analysis looks separately at the experiences of those who did and did not stay in public schools.

Only 4 percent of exiting teachers moved into teaching positions in private elementary and secondary schools. A slightly larger share left for other jobs in educational services outside elementary and secondary schools. Among those leaving the public school system, most therefore left for jobs in noneducation industries. The most common category for former teachers at all grade levels was Professional Services, which includes Information; Finance and Insurance; Real Estate and Rental and Leasing; Professional, Scientific, and Technical Services; Management of Companies and Enterprises; and Administrative and Support and Waste Management and Remediation Services.[24]

Table 2a reports median average annual earnings for all leaving teachers for the same sample of teachers for whom outside earnings data are available. Median earnings for former teachers remaining in public school districts were higher than for the same group while teaching, but the differences were modest. Former teachers working elsewhere, meanwhile, experienced a substantial decline in median earnings. The decline was largest for female teachers, who saw median earnings drop by one-fifth to one-third, but males also saw earnings declines of about one-tenth.

Figure 2a compares the distribution of average annual earnings for exiting teachers who remained employed by public school districts before and after leaving the classroom; figure 2b does the same for leavers to other industries. The plots indicate that a substantial share of both groups of leavers had very low levels of earnings after leaving, suggesting that they moved into part-time employment. Among teachers not working in public school districts, the distribution peaks at a level below $10,000. Although movement into part-time employment is clearly an important phenomenon in teacher labor markets, our primary aim in this article is to compare the structure of compensation for full-time employment inside and outside teaching. We therefore also look at results that weight individuals based on the predicted probability that they

24. Roughly similar shares of leavers entered jobs classified as Health Care, Other Services (including Arts, Entertainment, and Recreation; Accommodation and Food Services; and Other Services), Public Administration, or Other (including Agriculture, Forestry, Fishing, and Hunting; Mining, Quarrying, and Oil and Gas Extraction; Utilities; Construction; Manufacturing; Wholesale Trade; Retail Trade; and Transportation and Warehousing).

Table 2a. Median Earnings of Florida Teachers, 2002–8

	Stayers	Leavers Working in Public School Districts			Leavers Working Elsewhere in Florida		
			Not			Not in Public	
	Teaching	Teaching	Teaching	Difference	Teaching	Schools	Difference
All K–8 teachers	$42,804	$41,870	$44,570	$2,701	$37,736	$28,152	−$9,584
	79,163	12,691	12,691	6%	6,908	6,908	−25%
All male K–8 teachers	$43,865	$43,358	$49,510	$6,152	$38,701	$34,509	−$4,192
	11,061	1,870	1,870	14%	1,550	1,550	−11%
All female K–8 teachers	$42,644	$41,599	$43,854	$2,255	$37,503	$26,254	−$11,249
	68,102	10,821	10,821	5%	5,358	5,358	−30%
All 9–12 teachers	$44,948	$43,409	$47,543	$4,133	$38,928	$33,101	−$5,827
	28,442	4,007	4,007	10%	3,486	3,486	−15%
All male 9–12 teachers	$46,246	$44,724	$50,432	$5,708	$39,988	$36,915	−$3,074
	10,240	1,347	1,347	13%	1,483	1,483	−8%
All female 9–12 teachers	$44,212	$42,755	$45,769	$3,014	$38,046	$29,878	−$8,168
	18,202	2,660	2,660	7%	2,003	2,003	−21%

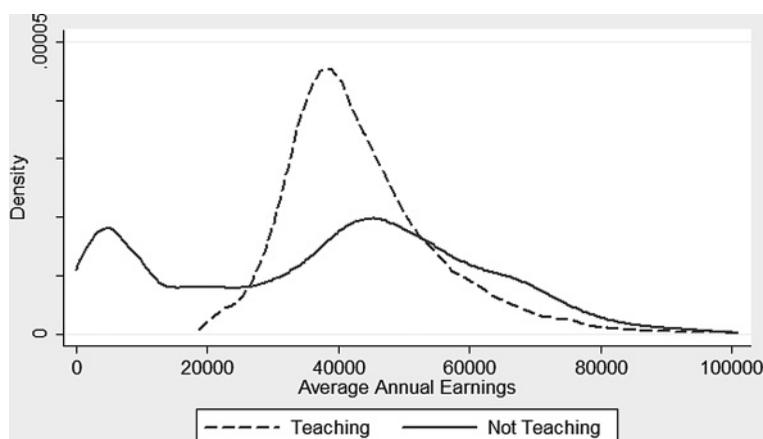*Note:* The numbers of teachers are in italics.

Figure 2a. Average Annual Earnings in and out of Teaching: Leavers in Public Schools (*N* = 14,812), Unweighted. *Note:* Teachers with average annual earnings over $100,000 are excluded from figure.

were working full time (we refer to the sample represented by these weighted results as "likely full-time workers").[25]

---

25. In the regression analysis, each teacher ∗ year observation is weighted by the probability that the teacher was working full time that year. Table 2b and figures 3a and 3b weight each teacher by the average predicted probability of working full time over all years that she worked after leaving the classroom. Appendix table A.1 shows that the weighted samples are disproportionately male
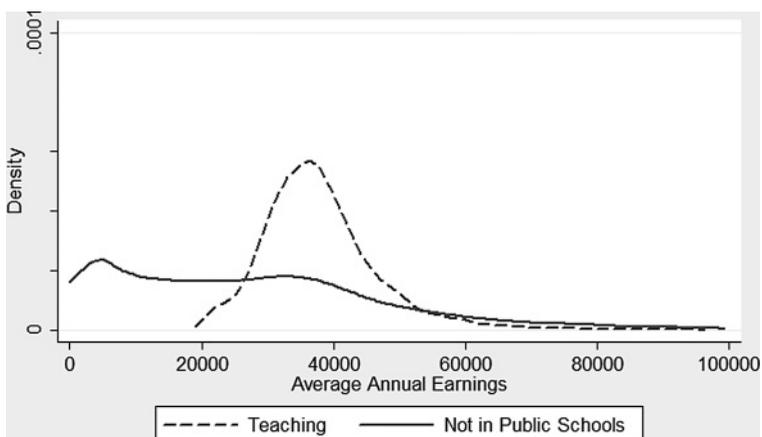
Figure 2b. Average Annual Earnings in and out of Public Schools: Leavers Not in Public Schools (*N* = 9,163), Unweighted. *Note:* Teachers with average annual earnings over $100,000 are excluded from figure.

Table 2b. Median Earnings of Florida Teachers, Weighted by Probability of Working Full Time

	Stayers	Leavers Working in Public School Districts			Leavers Working Elsewhere in Florida		
	Teaching	Teaching	Not Teaching	Difference	Teaching	Not in Public Schools	Difference
All K–8 teachers	$42,804	$41,870	$51,571	$9,702	$37,736	$35,445	−$2,291
	79,163	*12,691*	*12,691*	*23%*	*6,908*	*6,908*	*−6%*
All male K–8 teachers	$43,865	$43,358	$54,622	$11,264	$38,701	$38,568	−$133
	11,061	*1,870*	*1,870*	*26%*	*1,550*	*1,550*	*0%*
All female K–8 teachers	$42,644	$41,599	$50,984	$9,385	$37,503	$34,385	−$3,118
	68,102	*10,821*	*10,821*	*23%*	*5,358*	*5,358*	*−8%*
All 9–12 teachers	$44,948	$43,409	$54,891	$11,482	$38,928	$39,196	$267
	28,442	*4,007*	*4,007*	*26%*	*3,486*	*3,486*	*1%*
All male 9–12 teachers	$46,246	$44,724	$56,891	$12,167	$39,988	$41,181	$1,193
	10,240	*1,347*	*1,347*	*27%*	*1,483*	*1,483*	*3%*
All female 9–12 teachers	$44,212	$42,755	$53,492	$10,737	$38,046	$37,586	−$460
	18,202	*2,660*	*2,660*	*25%*	*2,003*	*2,003*	*−1%*

Note: The numbers of teachers appear in italics.

Table 2b replicates the analysis of changes in median earnings among likely full-time workers. As expected, median earnings for former teachers increase by a greater amount among those remaining in public school districts when emphasizing full-time employment. Among teachers leaving for

relative to the unweighted samples of exiting teachers but are otherwise similar in terms of the demographic characteristics that we observe.
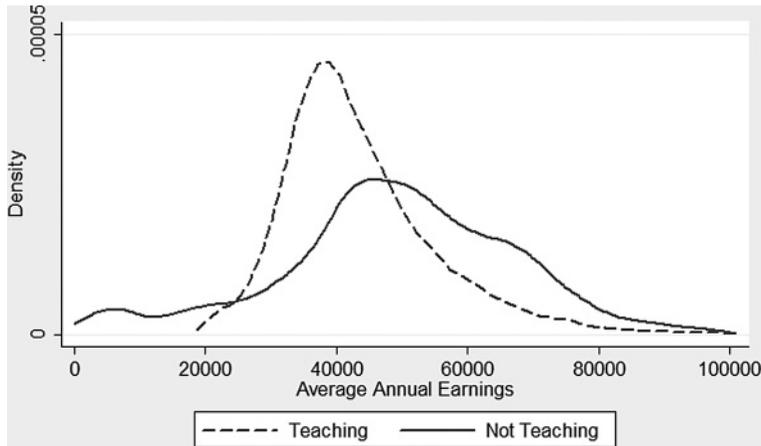
Figure 3a. Average Annual Earnings in and out of Teaching: Leavers in Public Schools ($N = 14{,}812$), Weighted. *Note:* Teachers with average annual earnings over $100,000 are excluded from figure.

other industries, median earnings declined by just 6 percent for elementary and middle school teachers and were essentially unchanged for high school teachers. Female teachers in the younger grades experienced a modest decline in wages of 8 percent, while men in the older grades saw a small increase of 3 percent. On average, the earnings of men in the younger grades and women in the older grades barely changed at all.

These results indicate that Florida teachers leaving for full-time employment in other industries are not taking jobs that are on average better compensated than their positions as teachers. Taking into account fringe benefits, which are typically more generous for public school teachers than for workers in other sectors of the economy (Costrell and Podgursky 2009), would presumably strengthen this conclusion.

Figures 3a and 3b, which display the distribution of average annual earnings for the likely full-time workers, confirm that the distribution of earnings among former teachers is considerably more dispersed than the earnings of the same groups while teaching. This is especially true of those leaving the classroom for other industries. Among leavers working within public school districts, the standard deviation of annual earnings among likely full-time workers increased from $11,652 to $18,064 as they left classroom teaching. The 90th/50th percentile earnings ratio for this group remained at 1.4, however, suggesting that the additional variation in earnings was primarily at the bottom of the earnings distribution. Among leavers working elsewhere in Florida, the standard deviation of annual earnings increased from $8,867 to $20,050 and the 90/50 earnings ratio jumped from 1.3 to 1.8.
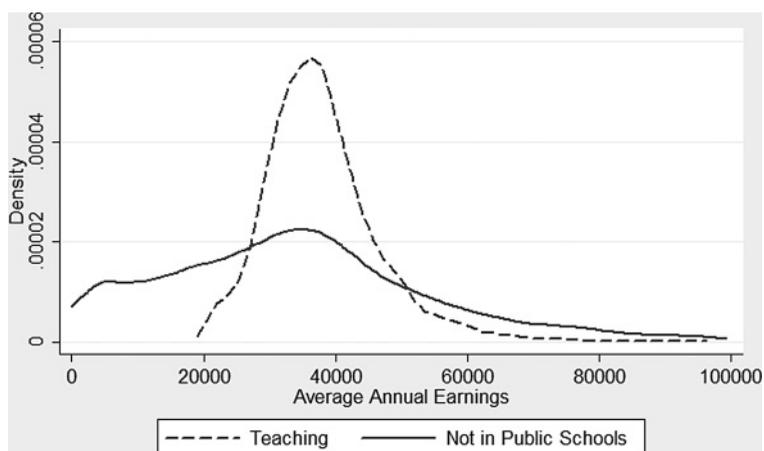
Figure 3b. Average Annual Earnings in and out of Public Schools: Leavers Not in Public Schools (*N* = 9,163), Weighted. *Note:* Teachers with average annual earnings over $100,000 are excluded from figure.

These patterns suggest substantial wage compression (relative to individual opportunity wages) among classroom teachers, but they also likely reflect the greater heterogeneity in the working conditions of jobs taken by the former teachers (as compared with when they were all in the same profession). The relatively low variability in wages among teachers could be attractive to some individuals, such as those expecting to take temporary breaks from the labor force. The next section, however, considers whether the additional variation in the nonteaching earnings of former teachers is related to their relative effectiveness while teaching.

5. TEACHER EFFECTIVENESS AND EARNINGS

To examine the relationship between classroom effectiveness and earnings, we use the sample of teachers in grades 4–8 for whom we are able to calculate VAMs of their contribution to student learning based on equation 1. We first examine the raw (unadjusted) relationship between estimated value added and (log) earnings while teaching and while not teaching. Then we examine the robustness of these results to the inclusion of district fixed effects and standard demographic controls. We focus throughout on the sample of likely full-time workers (with each teacher * year observation after the teacher left the classroom weighted by the probability that the worker was employed full time). Parallel unweighted results are qualitatively similar and available from the authors upon request.

Appendix table A.1 provides summary statistics for the three groups of teachers in our main analysis: teachers who remained in the classroom

throughout the period covered by our data (stayers), teachers who left for non-teaching positions in public school districts (leavers working in public school districts), and teachers who left for positions with other employers (leavers working elsewhere in Florida).[26] Among likely full-time workers, stayers and leavers working in public school districts had slightly higher average value added (0.02–0.03 standard deviations above the mean) than leavers elsewhere (−0.05).[27] In addition to being a little less effective on average, those leaving for full-time employment outside public school districts had modestly lower earnings while teaching and were more likely than stayers to be male, black, and Hispanic and less likely to have a master's degree in any field. Teachers with master's degrees were overrepresented among leavers for other positions in public school districts, suggesting that districts use this credential to screen internal candidates for nonteaching jobs.[28]

Table 3 presents estimates of the raw relationship between individual teacher effectiveness and the log of annual earnings. For both groups of leavers, the correlation is presented separately using annual earnings while teaching and while in other positions. Comparing the results for either group while teaching and while not teaching indicates differences in the structure of compensation inside and outside teaching for a common sample of teachers. Comparing the results for stayers with those for both groups of leavers while teaching provides suggestive evidence on the extent to which findings concerning the structure of compensation while teaching for leavers are generalizable to the full population of teachers in these grades.

The first row reveals a modest but statistically significant correlation between estimated value added and earnings while teaching for all three groups, with point estimates suggesting that a 1 standard deviation increase in value added is associated with an increase in earnings of 0.5 percent (for stayers), 1.2 percent (for leavers remaining employed by public school districts), and 1.3 percent (for leavers working elsewhere). Because our VAMs control for teacher experience, this relationship should not be driven by salary schedules offering higher pay for more senior teachers. It may instead reflect Florida's bonus program for schools improving their performance on the state's accountability system, pilot district-level merit pay schemes, or simply

---

26. Appendix table A.1 also provides information on leavers who do not appear in the Florida earnings data, who we observe only while teaching. Compared with the other leavers in the value-added sample, these teachers are more likely to be white and female and are on average somewhat younger.

27. The finding that exiting teachers are slightly less effective than those who remain is consistent with other research in Florida (Feng and Sass 2011; West and Chingos 2009) and with a growing body of evidence from other settings (Hanushek et al. 2005; Krieg 2006; Goldhaber, Gross, and Player 2010).

28. Florida principals are in fact required to hold a professional certificate obtained through a master's degree in educational leadership, administration, or administration and supervision.

Table 3. Relationship between Log(Earnings) and Effectiveness: Teachers in Value-Added Sample

	Stayers	Leavers Working in Public School Districts		Leavers Working Elsewhere in Florida	
	Teaching	Teaching	Not Teaching	Teaching	Not in Public Schools
All teachers	0.005 [0.001]*** 36,737	0.012 [0.004]*** 5,352	0.025 [0.007]***	0.013 [0.006]** 2,710	0.075 [0.020]***
All 4–5 teachers, math/reading average	0.009 [0.002]*** 20,921	0.014 [0.005]*** 3,042	0.033 [0.009]***	0.020 [0.008]*** 1,240	0.093 [0.027]***
All 4–5 teachers, math	0.010 [0.002]*** 20,916	0.015 [0.005]*** 3,041	0.031 [0.008]***	0.023 [0.007]*** 1,240	0.099 [0.027]***
All 4–5 teachers, reading	0.006 [0.002]*** 20,919	0.009 [0.005]* 3,042	0.027 [0.009]***	0.004 [0.008] 1,240	0.055 [0.029]*
All 6–8 math teachers	0.004 [0.002] 9,969	0.009 [0.008] 1,408	0.018 [0.014]	0.010 [0.009] 878	0.065 [0.035]*
All 6–8 reading teachers	0.000 [0.002] 11,716	0.010 [0.007] 1,713	0.014 [0.013]	0.008 [0.011] 973	0.053 [0.038]

Notes: Robust standard errors adjusted for clustering at the teacher level are in brackets; numbers of teachers are in italics (the number of teacher ∗ year observations is greater). Each value-added measure is standardized to have a mean of zero and standard deviation of 1. In the All Teachers results (row 1), dummies for grades 6–8 math and grades 6–8 reading are included in all regressions. *significant at 10%; **significant at 5%; ***significant at 1%.

differences in average compensation levels across school districts (West and Chingos 2009).

Among leavers for other positions in public school districts, the point estimate for the relationship between value added and post-teaching earnings is only modestly larger than the comparable estimate for earnings while teaching (0.025 vs. 0.012). Among leavers for jobs in other industries, however, the relationship between value added and subsequent earnings increases in magnitude and remains statistically significant, with a 1 standard deviation increase in value added associated with an increase in earnings of 7.5 percent. The estimate, which is statistically significantly different at the 1 percent level from the analogous estimate for the same group while teaching, provides evidence that the returns to skills associated with classroom effectiveness are greater outside the public school system.[29] The analogous coefficient on a

29. We test the significance of this difference by running a pooled model (including earnings while both teaching and not teaching) that includes a dummy variable that identifies the nonteaching

VAM based only on FCAT scores is 0.063; for a measure based only on SAT scores it is 0.061. That the coefficient on the average of the two measures is modestly larger is consistent with the hypothesis that averaging the two measures reduces measurement error.

The remaining rows of table 3 provide separate estimates of the same relationship for elementary school teachers using only their math or reading VAM and for middle school math and reading teachers. None of the estimates for these groups differs by statistically significant amounts from the analogous estimates for the pooled sample. In order to maximize statistical power, we therefore focus on the pooled sample as we consider the robustness of these results to the inclusion of district fixed effects and demographic control variables. We focus on the sample of leavers that worked outside the public schools.[30]

Table 4 estimates the relationship between value added and earnings conditional on district fixed effects and demographic variables including gender, race and ethnicity, age, and education credentials. As noted above, because Florida school districts are coterminous with counties, district fixed effects are useful to account for differences in regional labor markets. They may also eliminate any lingering bias in our estimates of teacher value added that are correlated with differences across districts. The additional control variables are included to see whether the relationship between value added and earnings reflects differences in earnings opportunities across demographic groups.

The coefficients on the control variables yield several interesting patterns. Males earned 19 percent more than females while not teaching, as compared with 2 percent more while teaching. Although this pattern might reflect gender differences in the earnings potential of leaving teachers, it could also be the case that male teachers are more likely than females to prioritize salary when considering other job opportunities.[31] A master's degree is associated with earnings that are 18 percent higher than those of other former teachers now working in other industries. The earnings premium for a doctoral degree (a credential held by only 3 percent of all leavers) is far greater outside teaching, perhaps suggesting that teachers who have invested in this credential are likely to leave education only for highly paid positions.

More important, the results concerning the relationship between effectiveness and earnings are only modestly changed from the analysis in table 3.

---

observations and an interaction between that dummy and our measure of value added. The coefficient on the interaction term and its standard error allow us to test the significance of the difference between the value-added coefficient while teaching compared with while not teaching. These models also include interactions between the nonteaching dummy and any control variables.

30. Analogous results for leavers who remained in the public schools are available from the authors.

31. However, the relationship between effectiveness and nonteaching earnings does not differ significantly for males and females.

Table 4. Relationship between Log(Earnings) and Effectiveness with Controls: Leavers in Value-Added Sample Weighted by Probability of Working Full Time

	Leavers Working Elsewhere in Florida			
	Teaching	Not in Public Schools	Teaching	Not in Public Schools
Value added in math and/ or reading	0.010 [0.005]*	0.061 [0.019]**	0.012 [0.005]*	0.062 [0.019]**
Male			0.020 [0.009]*	0.188 [0.039]**
Black			−0.010 [0.011]	−0.087 [0.044]*
Hispanic			−0.020 [0.014]	0.070 [0.065]
Age in 2007			0.004 [0.001]**	0.000 [0.002]
Master's degree			0.086 [0.009]**	0.177 [0.037]**
Doctoral degree			0.161 [0.028]**	0.464 [0.084]**
District fixed effects?	Yes	Yes	Yes	Yes
Teacher ∗ year observations	6,998	6,112	6,969	6,057
Number of teachers	2,710		2,690	
R^2	0.15	0.03	0.20	0.05

Notes: Robust standard errors adjusted for clustering at the teacher level are in brackets. The value-added measures are all standardized (separately for grades 4–5, 6–8 math, and 6–8 reading) to have a mean of zero and standard deviation of 1. Dummies for grades 6–8 math and grades 6–8 reading are included in all regressions.
*significant at 5%; **significant at 1%.

Among leavers for other industries, controlling for district fixed effects reduces the coefficient from 0.075 to 0.061; adding the other control variables leaves it essentially unchanged at 0.062. However, the relative magnitudes of the teaching and nonteaching coefficients remain largely unchanged: in the uncontrolled results, the nonteaching coefficient is 5.8 times as large as the teaching coefficient; in the results with controls, the nonteaching coefficient is 5.2 times as large as the teaching coefficient. The difference in the coefficients across models is statistically significant at the 1 percent level when only district fixed effects are added and at the 5 percent level when the full set of controls is added.

Table 5 presents the results of models that replace the continuous measure of teacher value added with dummy variables identifying teachers in each quartile of the effectiveness distribution. This nonlinear specification indicates that teachers in the 3rd and 4th effectiveness quartiles earn 9 percent and 19 percent more, respectively, than teachers in the bottom quartile. The

Table 5. Relationship between Log(Earnings) and Effectiveness Quartile: Leavers in Value-Added Sample Weighted by Probability of Working Full Time

	Leavers Working Elsewhere in Florida			
	Teaching	Not in Public Schools	Teaching	Not in Public Schools
Second value-added quartile	−0.012 [0.010]	0.058 [0.043]	−0.007 [0.010]	0.075 [0.043]*
Third value-added quartile	−0.007 [0.010]	0.094 [0.045]**	0.001 [0.010]	0.099 [0.045]**
Top value-added quartile	0.033 [0.013]**	0.186 [0.051]***	0.034 [0.012]***	0.186 [0.051]***
District fixed effects?	Yes	Yes	Yes	Yes
Controls?	No	No	Yes	Yes
Teacher ∗ year observations	6,998	6,112	6,969	6,057
Number of teachers	2,710		2,690	
R^2	0.15	0.03	0.20	0.05

Notes: Robust standard errors adjusted for clustering at the teacher level are in brackets. The omitted value-added quartile dummy is for the bottom quartile. Dummies for grades 6–8 math and grades 6–8 reading are included in all regressions. Controls include teacher gender, race/ethnicity, age in 2007, and education.
*significant at 10%; **significant at 5%; ***significant at 1%.

difference between the coefficients on the variables identifying the top two quartiles of teachers is statistically significant at the 10 percent level. Although the magnitude of the top-quartile coefficient suggests that highly effective teachers may have particularly strong earnings opportunities in other industries, the results in table 5 do not allow us to reject a linear relationship between effectiveness and earnings.

As noted above, although the results in tables 3–5 suggest that more effective teachers have better opportunity wages outside teaching, the observed relationship between effectiveness and nonteaching earnings could be misleading if higher-quality teachers leave only when they find a better job opportunity and less effective teachers leave involuntarily and thus suffer a displacement loss. To address this concern, we estimated the effectiveness-earnings relationship separately for teachers who had at least three years of experience in the Florida public schools and thus were tenured, effectively making performance-based dismissal impossible. These estimates are less precise due to the reduced sample size (the number of teacher ∗ year observations drops by more than half), but they are qualitatively similar to and never statistically significantly different from the estimates for the full sample.[32]

---

32. The uncontrolled coefficient on value added is 0.071, which is statistically significant at the 5 percent level. The coefficient conditional on district fixed effects is 0.039, which is not statistically significant

A remaining concern is that our results reflect bias in the value-added estimates due to the nonrandom matching of students and teachers. For example, this might be the case if principals assign unobservably better students to teachers who have higher opportunity wages in an effort to keep them from leaving the profession. As a partial test of this proposition, we calculate a measure of student-teaching sorting on prior-year test scores and examine whether it is related to nonteaching wages. We calculate this measure by first regressing students' prior-year test scores (from the FCAT and SAT in both reading and math) on school ∗ grade ∗ year fixed effects and then aggregating the residuals to the teacher ∗ year level. We then aggregate the teacher * year averages to the teacher level using the shrinkage estimator described earlier. Finally, we standardize the result to have a mean of zero and standard deviation of 1 across all teachers so that it is directly comparable to our measure of teacher value added. This sorting measure indicates to what extent each teacher is systematically assigned students with higher or lower initial test scores than other students in the same school, grade, and year.

Sorting on prior-year test scores should not bias our value-added estimates because we control for them, but this sorting measure may nonetheless serve as a proxy for sorting on unobservable characteristics related to student achievement. It is first worth noting that the sorting and VAMs are only modestly correlated ($r = .22$), which implies that there is a great deal of variation in our teacher VAM that is uncorrelated with student sorting. Table 6 shows that the sorting measure is correlated with nonteaching earnings, a pattern that could result from principals assigning higher-achieving classrooms to teachers who they expect have better outside earnings opportunities. However, the relationship is less than half as strong as for the VAM (compare column 2 of table 6 with the final column of table 3). More importantly, when the sorting and VAMs are examined together, the coefficient on the VAM is always larger than the coefficient on the sorting measure and the latter coefficient is always insignificant (although the difference between the two coefficients is never statistically significant). We interpret these results as evidence that the estimated relationship between value added and nonteaching earnings is not simply the result of bias in the value-added estimates.

As a second check of the robustness of our results to nonrandom matching of teachers and students, we examined the relationship between earnings and VAMs that condition on school fixed effects. These value-added estimates are no longer comparable across schools but have the advantage of controlling

---

but imprecisely estimated (its standard error is 0.028). For this group of teachers while teaching, the coefficient on value added is 0.016 (0.014 with district fixed effects). Complete results for teachers with at least three years of experience are included in the online appendix.

Table 6. Relationship between Log(Earnings), Sorting, and Effectiveness: Leavers in Value-Added Sample (Weighted)

			Leavers Working Elsewhere in Florida					
	Teaching	Not in Public Schools	Teaching	Not in Public Schools	Teaching	Not in Public Schools	Teaching	Not in Public Schools
Measure of sorting on prior-year test scores	−0.002 [0.005]	0.041 [0.018]**	−0.004 [0.005]	0.031 [0.018]*	−0.001 [0.004]	0.031 [0.018]*	0.004 [0.004]	0.034 [0.018]*
Value-added in math and/ or reading			0.014 [0.006]**	0.070 [0.020]***	0.011 [0.005]**	0.055 [0.019]***	0.011 [0.005]***	0.056 [0.019]***
Male							0.020 [0.009]**	0.188 [0.039]***
Black							−0.009 [0.011]	−0.082 [0.044]*
Hispanic							−0.019 [0.014]	0.075 [0.065]
Age in 2007							0.004 [0.001]***	0.000 [0.002]
Master's degree							0.087 [0.009]***	0.178 [0.037]***
Doctoral degree							0.161 [0.027]***	0.466 [0.084]***
District fixed effects?	No	No	No	No	Yes	Yes	Yes	Yes
Teacher * year observations	7,232	6,373	6,998	6,112	6,998	6,112	6,969	6,057
Number of teachers	2,834		2,710		2,710		2,690	
R²	0.00	0.00	0.00	0.01	0.15	0.03	0.20	0.05

Notes: Robust standard errors adjusted for clustering at the teacher level are in brackets. The value-added and sorting measures are all standardized (separately for grades 4–5, 6–8 math, and 6–8 reading) to have a mean of zero and standard deviation of 1. Dummies for grades 6–8 math and grades 6–8 reading are included in all regressions. *significant at 10%; **significant at 5%; ***significant at 1%.

for unobserved school characteristics that are constant over time. As shown in appendix table A.2, the wage regression with district effects yields an effectiveness-earnings relationship that is modestly smaller than our preferred estimate (0.049 instead of 0.061) but is still statistically significant at the 1 percent level. Adding school fixed effects to the wage equation, which is very demanding of the data as it compares only leaving teachers that left the same school, yields a coefficient estimate that is smaller (0.037) but still statistically significant (at the 10 percent level).

We also tested the robustness of our results to changing the value-added specification in two additional ways. First, we confirmed that excluding teacher experience from the value-added models produced similar results. Second, we checked whether our results might be influenced by transient changes in teachers' performance prior to their leaving the classroom (i.e., Ashenfelter dips) that are correlated with value added. We performed this latter test in two ways: (1) excluding the last year of data for all teachers for whom we are able to generate value-added estimates for multiple years and (2) excluding data from a teacher's final year *and* dropping all teachers who only appeared in the first year of our data. Results of these alternative specifications, both of which produce qualitatively similar results to our preferred estimates, are included in the online appendix.

As noted above, because we only observe teachers for at most six years after leaving the classroom, our results may overstate the long-term relationship between value added and earnings if this relationship changes over time. We address this issue by interacting value added with number of years not teaching.[33] We obtained coefficients of 0.053 on value added and 0.018 on the interaction (although the interaction is not statistically significant). In other words, the effectiveness-earnings relationship is positive and statistically significant in the first years of earnings that we examine and increases each year, but not by a statistically significant amount. With district fixed effects, we get 0.044 on value added and 0.013 on the interaction. We interpret these results as suggestive evidence that our findings concerning the relationship between effectiveness and earnings in other industries would be stronger if we had a longer panel.

Finally, we conducted a complementary analysis of the pre-teaching earnings of new teachers entering the profession from other industries (those with zero years experience in Florida public schools in their first year of teaching).[34]

---

33. We rescale the year not teaching variable so that first year not teaching is zero, when in fact it is the second year not teaching because we exclude data from the first year not teaching.
34. We restrict the sample to observations when the worker was at least twenty-five years old, excluding the first pre-teaching year as a potential transition year (so the analysis is limited to workers observed for at least two years before entering teaching).

We weight these observations by the predicted probability that the worker was employed full time, using data from the Current Population Survey.[35] Although the number of entering teachers is somewhat larger than the number of leaving teachers, selection issues make it more difficult to interpret the results for entrants. Both entering and leaving teachers are self-selected, but for leaving teachers we have data on the entire population of potential leavers and thus can examine selection on observables (most important, value added). For entering teachers we cannot identify or describe what the full population of potential teachers (or even applicants for teaching positions) looks like. Thus we cannot confirm that the entering teachers are similar to the pool of potential teachers, as we did in appendix table A.1 by comparing leaving teachers with the pool of potential leavers.

The results for the entering teachers, which are presented in appendix table A.3, are not statistically significantly different from the results for the leaving teachers. The point estimate of the relationship between effectiveness and earnings for entering teachers that conditions only on district fixed effects is weaker but still statistically significant. The results with control variables suggest that, among entering teachers, a 1 standard deviation increase in value added was associated with pre-teaching earnings that were 4–6 percent higher.

6. CONCLUSIONS

The results presented above represent the first evidence on the extent to which opportunity wages outside teaching are correlated with classroom effectiveness, a question of considerable interest to policy makers. Moreover, we are able to compare the relationship between these characteristics and earnings inside and outside the teaching profession for a common sample of leaving teachers.

There are several important limitations on this evidence. A relatively small number of the teachers in our data leave the classroom for other jobs—and many of those who leave teaching remain employed by public school districts. Those leaving teaching prior to retirement age are not a randomly selected group, and a substantial share of exiting teachers does not appear in our earnings data at all. Finally, our evidence comes from a single state. Though teacher compensation policies in Florida school districts closely resemble those in use elsewhere, it is possible that our findings are driven by peculiarities of the state's labor market or other education policies. Ongoing layoffs of teachers in response to budget cuts in multiple states may provide future

---

35. Specifically, we use data on Florida workers aged 25–55 with a BA or MA in the 2002, 2003, and 2004 March Current Population Survey to run a probit regression of full-time status on earnings and gender.

researchers with an opportunity to address these limitations, especially if it were to become possible to link the teacher identifiers in state education warehouses to administrative data from federal tax returns.[36]

Even so, our results have clear implications for teacher recruitment and compensation policies in Florida and likely beyond. We find greater dispersion in the earnings of leaving teachers after moving into other industries than for the same individuals while in the classroom, which suggests that teacher salaries are compressed relative to opportunity wages. This is hardly surprising, given the role that collective bargaining plays in the determination of teacher salaries, and it likely also reflects greater heterogeneity in working conditions across multiple industries. Yet it confirms that many teachers have better earnings opportunities outside the profession, while others likely earn more in teaching than would be the case elsewhere.

More importantly, we find that teachers who are most successful in raising student achievement earn more in other industries. Although teaching is surely a unique endeavor requiring specialized skills, at least some of the same attributes that make for effective teachers appear to be rewarded in the broader labor market. This suggests that alternative certification programs that successfully target professionals who have been successful in other industries may well provide a source of new teachers who are likely to be effective in the classroom.

The specific design of policies to offset differences in opportunity wages between more and less effective teachers is beyond the scope of our analysis, but it seems safe to conclude that experimentation with merit pay and other incentive schemes should continue. The success of financial incentives in retaining effective teachers will depend on the wage elasticity of decisions to remain in the classroom, which may be low if many exiting teachers leave for reasons other than the opportunity to earn more elsewhere. By ignoring the realities of the outside labor market, however, the dominant teacher compensation systems in American public school districts are ill designed to recruit and retain their most valuable employees.

36. Chetty et al. (forthcoming) link data from the Project STAR class-size experiment to administrative data from tax returns in order to be able to track students' earnings across multiple states.

REFERENCES

Ballou, Dale, and Michael Podgursky. 1997. *Teacher pay and teacher quality*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Boardman, Anthony E., Linda Darling-Hammond, and Stephen P. Mullin. 1982. A framework for the analysis of teachers' demand and supply. *Economics of Education Review* 2(2): 127–55.

Chetty, Raj, John Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Schanzenbach, and Danny Yagan. Forthcoming. How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Quarterly Journal of Economics*. In press.

Costrell, Robert M., and Michael Podgursky. 2009. Teacher retirement benefits. *Education Next* 9(2): 58–63.

Dolton, Peter, and Wilbert van der Klaauw. 1995. Leaving teaching in the UK: A duration analysis. *Economic Journal* 105(429): 431–44.

Dolton, Peter, and Wilbert van der Klaauw. 1999. The turnover of teachers: A competing risks explanation. *Review of Economics and Statistics* 81(3): 543–52.

Feng, Li, and Tim Sass. 2011. Teacher quality and teacher mobility. CALDER Working Paper No. 57, Urban Institute.

Ferguson, Ronald F., and Helen F. Ladd. 1996. How and why money matters: An analysis of Alabama schools. In *Holding schools accountable: Performance-based reform in education*, edited by Helen F. Ladd, pp. 265–98. Washington, DC: Brookings Institution Press.

Goldhaber, Dan, Betheny Gross, and Daniel Player. 2010. Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best? *Journal of Policy Analysis and Management* 30(1): 57–87.

Hanushek, Eric A., John F. Kain, Daniel M. O'Brien, and Steven G. Rivkin. 2005. The market for teacher quality. NBER Working Paper No. 11154.

Hanushek, Eric A., and Steven G. Rivkin. 1997. Understanding the twentieth-century growth in U.S. school spending. *Journal of Human Resources* 32(1): 35–68.

Harris, Douglas, and Tim Sass. 2007. Teacher training, teacher quality and student achievement. CALDER Working Paper No. 3, Urban Institute.

Hoxby, Caroline M., and Andrew Leigh. 2004. Pulled away or pushed out? Explaining the decline of teacher aptitude in the United States. *American Economic Review* 94(2): 236–40.

Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2007. What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review* 27(6): 615–31.

Kane, Thomas J., and Douglas O. Staiger. 2008. Estimating teacher impacts on student achievement: An experimental evaluation. NBER Working Paper No. 14607.

Krieg, John M. 2006. Teacher quality and attrition. *Economics of Education Review* 25(1): 13–27.

Murnane, Richard J., and Randall Olsen. 1989. The effects of salaries and opportunity costs on length of stay in teaching: Evidence from Michigan. *Review of Economics and Statistics* 71(2): 347–52.

Murnane, Richard J., and Randall Olsen. 1990. The effects of salaries and opportunity costs on length of stay in teaching: Evidence from North Carolina. *Journal of Human Resources* 25(1): 106–24.

Murnane, Richard J., Judith D. Singer, John B. Willett, James J. Kemple, and Randall J. Olsen. 1991. *Who will teach? Policies that matter.* Cambridge, MA: Harvard University Press.

Murnane, Richard J., John B. Willet, M. Jay Braatz, and Yves Duhaldeborde. 2001. Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. *Economics of Education Review* 20(4): 311–20.

Podgursky, Michael, Ryan Monroe, and Donald Watson. 2004. The academic quality of public school teachers: An analysis of entry and exit behavior. *Economics of Education Review* 23(5): 507–18.

Podgursky, Michael, and Ruttaya Tongrut. 2006. (Mis-)measuring the relative pay of public school teachers. *Education Finance and Policy* 1(4): 425–40.

Rockoff, Jonah E. 2004. The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review* 94(2): 247–52.

Rockoff, Jonah E., Brian A. Jacob, Thomas J. Kane, and Douglas O. Staiger. 2008. Can you recognize an effective teacher when you recruit one? *Education Finance and Policy* 6(1): 43–74.

Rothstein, Jesse. 2010. Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics* 94(2): 247–52.

Scafidi, Benjamin, David L. Sjoquist, and Todd R. Stinebrickner. 2006. Do teachers really leave for higher paying jobs in alternative occupations? *Advances in Economic Analysis and Policy* 6(1): 1–42.

Stinebrickner, Todd R. 2002. An analysis of occupational change and departures from the labor force: Evidence of the reasons that teachers quit. *Journal of Human Resources* 37(1): 192–216.

U.S. Department of Education. 2008. *Table 80: Staff employed in public elementary and secondary school systems by functional area: Selected years 1949–50 through fall 2006.* Digest of Education Statistics. Washington, DC: National Center for Education Statistics. Available http://nces.ed.gov/pubs2009/2009020_2a.pdf. Accessed 17 July 2011.

West, Martin R., and Matthew M. Chingos. 2009. Teacher effectiveness, mobility, and attrition in Florida. In *Performance incentives: Their growing impact on American K–12 education,* edited by Matthew G. Springer, pp. 251–71. Washington, DC: Brookings Institution Press.

Willis, Robert, and Sherwin Rosen. 1979. Education and self-selection. *Journal of Political Economy* 87(5): S7–S36.

APPENDIX

Table A.1. Summary Statistics, 4th- to 8th-Grade Teachers in Value-Added Sample

	Stayers	Leavers Working in Public School Districts		Leavers Working Elsewhere in Florida		Leavers Not Working or Not in Florida
	Unweighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
Value-added (standardized)	0.03	0.00	0.02	−0.05	−0.05	−0.03
Mean earnings while teaching	$45,354	$43,750	$43,750	$39,281	$39,281	$38,088
Mean earnings while not teaching		$43,903	$53,507	$31,720	$39,311	
Median earnings while teaching	$42,924	$42,016	$42,016	$38,364	$38,364	$37,025
Median earnings while not teaching		$46,109	$52,503	$29,315	$36,413	
Percent male	16%	17%	20%	25%	30%	16%
Percent black	17%	21%	23%	19%	19%	17%
Percent Hispanic	9%	12%	12%	12%	13%	5%
Age in 2007	42.7	40.5	41.3	37.9	38.0	37.1
Percent with master's	38%	54%	62%	30%	31%	29%
Percent with doctorate	2%	3%	3%	3%	3%	2%
Number	35,568	5,133	5,133	2,638	2,638	3,367

Note: "Weighted" indicates that the sample is weighted by the probability that each teacher was working full time.

Table A.2. Relationship between Log(Earnings) and Effectiveness with Controls: Leavers in Value-Added Sample Weighted by Probability of Working Full Time, School Fixed Effects (SFE) Models

	Leavers Working Elsewhere in Florida									
	Teaching	Not in Public Schools	Teaching	Not in Public Schools	Teaching	Not in Public Schools	Teaching	Not in Public Schools	Teaching	Not in Public Public Schools
Value-added in math and/or reading controlling for SFE	0.015 [0.007]**	0.059 [0.017]***	0.015 [0.007]**	0.049 [0.016]***	0.014 [0.006]**	0.047 [0.016]***	0.010 [0.006]*	0.037 [0.019]*	0.008 [0.006]	0.037 [0.019]*
Male					0.050 [0.013]***	0.197 [0.036]***			0.025 [0.015]*	0.202 [0.043]***
Black					−0.009 [0.014]	−0.085 [0.039]**			−0.016 [0.016]	−0.012 [0.051]
Hispanic					−0.003 [0.019]	0.113 [0.054]**			−0.026 [0.020]	0.120 [0.067]*
Age in 2007					0.003 [0.001]***	0.000 [0.002]			0.003 [0.001]***	0.003 [0.002]
Master's degree					0.105 [0.012]***	0.193 [0.032]***			0.102 [0.013]***	0.230 [0.040]***
Doctoral degree					0.175 [0.037]***	0.322 [0.092]***			0.170 [0.039]***	0.192 [0.123]
District fixed effects?	No	No	Yes	Yes	Yes	Yes	No	No	No	No
School fixed effects?	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Teacher * year observations	14,465	8,754	14,465	8,754	14,430	8,695	14,465	8,754	14,430	8,695
Number of teachers	3,477		3,477		3,456		3,477		3,456	
R²	0.00	0.00	0.06	0.03	0.08	0.05	0.21	0.34	0.22	0.35

Notes: Robust standard errors adjusted for clustering at the teacher level are in brackets. The value-added measures are all standardized (separately for grades 4–5, 6–8 math, and 6–8 reading) to have a mean of zero and standard deviation of 1. Dummies for grades 6–8 math and grades 6–8 reading are included in all regressions. *significant at 10%; **significant at 5%; ***significant at 1%.

Table A.3. Relationship between Log(Earnings) and Effectiveness: Entering Teachers in Value-Added Sample Weighted by Probability of Working Full Time

	Teaching	Not in Public Schools	Teaching	Not in Public Schools	Entering Teachers that Worked Elsewhere in Florida			
					Teaching	Not in Public Schools	Teaching	Not in Public Schools
Value-added in math and/or reading	0.003 [0.004]	0.040 [0.021]*	0.005 [0.004]	0.057 [0.021]**				
Second value-added quartile					−0.003 [0.009]	−0.031 [0.050]	−0.001 [0.009]	−0.025 [0.049]
Third value-added quartile					0.001 [0.009]	0.003 [0.050]	0.005 [0.009]	0.020 [0.049]
Top value-added quartile					0.013 [0.011]	0.105 [0.055]*	0.015 [0.011]	0.144 [0.055]**
Male			0.025 [0.008]**	0.061 [0.042]			0.024 [0.008]**	0.058 [0.042]
Black			0.019 [0.010]*	−0.017 [0.051]			0.019 [0.010]*	−0.023 [0.051]
Hispanic			−0.012 [0.014]	0.025 [0.059]			−0.012 [0.014]	0.021 [0.059]
Age in 2007			0.002 [0.000]**	0.017 [0.002]**			0.002 [0.000]**	0.017 [0.002]**

Master's degree	0.061 [0.008]**		0.061 [0.042]	0.061 [0.008]**		0.059 [0.042]
Doctoral degree	0.119 [0.029]**		0.070 [0.175]	0.118 [0.029]**		0.050 [0.174]
District fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
Teacher × year observations	9,617	7,816	7,789	9,597	7,816	7,789
Number of teachers	3,773		3,759	3,773		3,759
R^2	0.18	0.03	0.05	0.20	0.03	0.05

Notes: Robust standard errors adjusted for clustering at the teacher level are in brackets. The value-added measures are all standardized (separately for grades 4–5, 6–8 math, and 6–8 reading) to have a mean of zero and standard deviation of 1. Dummies for grades 6–8 math and grades 6–8 reading are included in all regressions. *significant at 10%; **significant at 1%.