Teacher Quality and Dropout Outcomes in a Large, Urban School District

Cory Koedel*
University of Missouri

June 2008

Recent research shows that variation in teacher quality has large effects on student performance. However, this research is based entirely on student test scores. Focusing on high-school math teachers, this paper evaluates teacher quality in terms of another educational outcome of great interest – graduation. I use a unique instrumental variables approach to identify teacher effects and find that differences in teacher quality have large effects on graduation outcomes. Because teacher effects on graduation outcomes will be more pronounced for students who are on the graduation margin, the results imply an avenue through which high-quality teachers are more productive with disadvantaged students.

*I would like to thank Andrew Zau and many administrators at San Diego Unified School District, in particular Karen Bachofer and Peter Bell, for assistance with data issues. I also thank Julian Betts, Julie Cullen, Yixiao Sun, Nora Gordon, Daniel Millimet, participants at the UCSD applied seminar and two anonymous referees for useful comments and suggestions and the Spencer Foundation for research support. The underlying project that provided the data for this study has been funded by the Public Policy Institute of California and directed by Julian Betts.

Education Secretary Margaret Spellings recently referred to a small group of largely urban schools as "dropout factories” and did so with good reason – these schools are graduating less than 50 percent of their students. In fact, Orfield et al. (2004) show that in almost half of the high schools in the 100 largest urban school districts in the country, 12th-grade classes are less than 50 percent of the size of 9th-grade classes four years earlier. This “graduation rate crisis,” as it has been dubbed by these authors, is of great economic significance. For example, Ashenfelter and Krueger (1994) estimate that an extra year of schooling corresponds to a 12 to 16 percent increase in wages and Barrow and Rouse (2004) estimate that in 2003, high school graduates earned approximately 75 percent more than high school dropouts annually.[1] Furthermore, in addition to the costs of dropping out borne by individuals, high dropout rates are also associated with negative externalities. Lochner and Moretti (2004) estimate that a 1-percentage-point increase in high school completion among men ages 20 to 60 would save the United States $1.4 billion per year by reducing costs associated with crime.

The recent teacher-quality literature overwhelmingly indicates that differences in teacher quality have large effects on student performance measured by test scores.[2] Given this result, and that graduation outcomes are of such great economic importance, it seems natural to ask whether teacher quality affects graduation outcomes.[3] Econometrically, analyzing teacher quality in terms of graduation outcomes is complicated by the non-random assignment of students to

[1] Rouse (1999) does a follow-up study to Ashenfelter and Krueger and finds that these authors may have overstated the returns to schooling, which are actually closer to 10 percent per year. Estimates from Barrow and Rouse incorporate the facts that high school graduates earn higher wages and work more hours.
[2] See, for example, Aaronson et al. (2007), Hanushek et al. (2005), Koedel (2007), Nye et al. (2004), Rivkin et al. (2005), Rockoff (2004).
[3] Loeb and Page (2000) find that teacher salary increases have a positive effect on graduation rates ten years later. However, they do not evaluate the extent to which variation in teacher quality, measured at the micro level, affects graduation outcomes.

teachers. In the test-score literature, panel datasets have been exploited to remove bias generated by this non-random student-teacher matching. Specifically, test-score studies have relied on lagged measures of performance and student fixed effects to remove sorting bias.[4] However, in the analysis of graduation outcomes these methods cannot be implemented because graduation outcomes cannot be tracked over time as can test-score outcomes. At a given point in time, a student simply drops out of school or does not.

I use a unique instrumental variables approach based on variation from yearly school-level staffing changes to identify teacher effects on graduation outcomes. Differences in teacher quality are shown to have non-negligible effects, even within schools. This implies that improvements in teacher quality can help mitigate the graduation-rate crisis faced by many urban school districts across the country. More generally, this finding is also relevant in the debate over which types of students are more responsive to changes in teacher quality. Whereas recent research by Clotfelter et al. (2006), based on student test-score performance, indicates that the returns to teacher quality may be higher for advantaged students, the evidence here shows that the weakest students also have much to gain from improvements in teacher quality.

I. Empirical Strategy

Students' teachers throughout high school are likely to be endogenous to their graduation outcomes. This endogeneity may manifest itself either through direct teacher selection by students within subjects, or through subject selection (e.g., choosing to take calculus) that affects teacher selection. In addition, graduation outcomes may be correlated with the assignment of students to teachers by administrators. To identify teacher effects, I exploit exogenous variation

[4] Recent work by Rothstein (2008) suggests that these methods may be insufficient for removing sorting bias.

in the exposure of students to teachers over time. This variation comes from yearly school-level staffing changes and is highest among math teachers, who are the focus of this study. I jointly model students' graduation outcomes with teacher selection and estimate teacher effects via maximum likelihood.

Consider the following empirical model of the student dropout decision. Let D_i^* denote the net benefit to student *i* of dropping out of high school where:

$$D_i^* = \beta_0 + X_i\beta_1 + T_i\theta_J + \varepsilon_i \qquad \textbf{(1)}$$

In equation (1), the vector X_i includes controls for demographics, socioeconomic status, English-learner status, whether or not the student switched schools during high school and the initial math class taken in ninth grade for each student. The J-dimensional vector T_i indicates which teachers taught student *i* during high school and is likely to be endogenous to the dropout outcome. Student *i*'s decision to drop out, D_i, is a zero-one indicator that is equal to one if $D_i^* \geq 0$ and equal to zero otherwise.

The initial-math-course controls in equation (1) provide a measure of pre-high school performance as pre-high school performance determines initial math-course placement in high school.[5] Additionally, they capture the effects of tracking within schools. For example, if students who start high school in remedial math classes are more likely to attend remedial

[5] I also considered including pre-high school test scores in the dropout specification but there was a substantial portion of the student sample that did not have test-score records for the 8th grade. This is because I focus on underperforming schools in San Diego which tend to have the most transient student populations (see Section II).

English classes, remedial history classes, and so on, these controls will capture the effects of this tracking. Not surprisingly, the initial-math-course controls are strong predictors of graduation.

Teacher selection throughout high school for each student, T_i, can be similarly modeled. The underlying teacher-selection function, T_i*, depends on the same explanatory variables as in equation (1) and an additional set of controls that capture student-teacher exposure. For all j in which the corresponding entry in the vector $T_i^* \geq 0$, student i is taught by teacher j in high school.[6,7]

$$\begin{aligned} T_{i1}^* &= \alpha_0 + X_i\alpha_1 + Z_{i1}\alpha_2 + u_{i1} \\ T_{i2}^* &= \gamma_0 + X_i\gamma_1 + Z_{i2}\gamma_2 + u_{i2} \\ &\vdots \\ T_{iJ}^* &= \delta_0 + X_i\delta_1 + Z_{iJ}\delta_2 + u_{iJ} \end{aligned} \qquad \textbf{(2)}$$

Each observed T_{ij} is a dichotomous outcome equal to one if student i had teacher j at any point in high school and equal to zero otherwise. In equation set (2), the vector Z_{ij} is excluded from the dropout equation and used to identify the effect of teacher j.

[6] The teacher selection equations are not mutually exclusive. Theoretically, a by-year multinomial model of teacher selection may be more complete in that, in most cases, the choice of one teacher may exclude choosing others within years. However, the parameter space for such a model would be so large that it would be infeasible to estimate.

[7] The vast majority of student-teacher matches are for one year although some students and teachers were matched for up to three years in the data. The model does not allow the effect of multiple years with a given teacher to differ from the effect of a single semester with that teacher. However, students that are taught by the same teacher in multiple years are also likely to be in more exposed to that teacher based on the year and math-course level in which they enter high school. Differences in this student-teacher exposure will be used to identify the teacher effects as discussed below.

The error terms ε, $u_1,\ldots,$ u_J are assumed to be joint-normally distributed with mean zero and variance-covariance matrix Ω:

$$\begin{matrix} \varepsilon_i \\ u_{i1} \\ \vdots \\ u_{iJ} \end{matrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \Omega\right) \quad \text{Where} \quad \Omega = \begin{bmatrix} \sigma^2_{\varepsilon\varepsilon} & \sigma_{\varepsilon u_1} & \cdots & \sigma_{\varepsilon u_J} \\ \sigma_{\varepsilon u_1} & \sigma^2_{u_1 u_1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \sigma_{\varepsilon u_J} & \cdots & \cdots & \sigma^2_{u_J u_J} \end{bmatrix} \tag{3}$$

In (3), I assume the standard normalization $\sigma^2_{\varepsilon\varepsilon} = \sigma^2_{u_1u_1} = \ldots = \sigma^2_{u_Ju_J} = 1$. In the absence of this assumption, the coefficients and error variances in (1) and (2) are only identified up to their proportions. The distributional assumption in (3) allows for correlation between the error terms in the teacher-selection and dropout equations. Given this framework, the dropout decision is specified as a probit and teacher selection as a binary, endogenous determinant of graduation.

The probit-based estimation strategy is preferred to its linear analog. One reason is that a linear model lacks flexibility. For example, if the dropout outcome were specified as a linear function of the explanatory variables it would restrict teacher effects to be the same across students regardless of the X_i's; despite the fact that students' probabilities of dropping out vary widely across the X_i's (see Table 2 and Appendix Tables A.1 – A.3). This inflexibility in the linear model will result in predicted probabilities of the dependent variables that are not between zero and one.[8]

[8] Additionally, Monte Carlo exercises performed by Bhattacharya et al. (2006) strongly support the probit-based approach over the linear-probability-model analog, particularly where $P(D_i = 1)$ is relatively small as is the case in my models. In fact, even when the multivariate probit is misspecified (i.e., the underlying data generating process is

My identification strategy relies on variation in the classes taught by teachers over time to identify teacher effects. For example, consider a math teacher who teaches four classes of algebra and one class of geometry in one year. In the next year, this teacher might teach two classes of algebra and three classes of geometry. Furthermore, some teachers move in and out schools over time. Figure 1 shows four examples of variation in the proportion of the total number of student semesters taught in different subjects over time for four different teachers used in this analysis. Given this variation in the classes taught by teachers over time, differences in the classes taken and the timing of classes taken by students create variation in student-teacher exposure. For example, depending on what year a given student happens to take geometry, the probability of that student being taught by teacher *j* may change simply because teacher *j* teaches more (or fewer) geometry classes in that year.

If I were to assume that students' own math-course choices are exogenous to their dropout decisions, I could instrument for teacher selection using each student's own class schedule throughout high school. For example, the math-course path *algebra-geometry-intermediate algebra-precalculus*, followed over a four year sequence by some student, would imply exposure to a specific set of teachers. However, the course choices made by students during high school are unlikely to be exogenous to their dropout decisions. Therefore, rather than use each student's own class schedule to instrument for teacher selection, I instead use each student's entry-level math course in ninth grade to project her subsequent math-course path based on sample-wide averages (probabilistic). For example, for students who took algebra in the ninth grade at school

---

not multivariate normal), they show that it still produces treatment-effect estimates that are generally less biased than the analogous linear-based specification.

X, I can map out the proportion who took each type of math course in subsequent years and in this way create an average math-course path for these students at school X. Because high-school students generally take math courses in a particular sequence, students' ninth-grade math courses are strong predictors of their math-course paths.[9]

I create math-course paths at each school based on seven possible entry-level math courses.[10] For each student, I replace her own math-course choices with the average math-course path corresponding to her entry-level math course at her school, thereby removing the potentially endogenous decisions of individual students from the instrument sets. To avoid building the effects of the treatments (teachers) into the instruments, and to further alleviate endogeneity concerns, I exclude students' own year-cohorts when calculating the math course paths.

For teacher-selection equation *j*, I interact the students' projected math-course paths with indicator variables for the subject-years that teacher *j* teaches to complete the instrument sets. I define seven subject-types by which students and teachers can be matched in any given year: pre-algebra (that is, anything below algebra), algebra, geometry, advanced geometry, intermediate algebra, advanced intermediate algebra and advanced math (pre-calculus and calculus). The incorporation of teachers' teaching schedules into the instrument sets means that the teacher effects are identified from contemporaneous student-teacher exposure and will

[9] Individual students deviate from the standard math-course path given their entry-level math course for many reasons. Most commonly, a student may be required to re-take a course. Students' math-course paths are averaged within groups such that some fraction of each group is expected to deviate from the standard math-course path given their entry-level math course.

[10] The seven math-course paths are based on the following entry-level mathematics classifications: No math, pre-algebra, pre-algebra/algebra, algebra, algebra/geometry, geometry, advanced geometry. Pre-algebra/algebra and algebra/geometry indicate split years. None of the (few) students who entered high school at a level above advanced geometry failed to graduate. Therefore, these students were omitted from the analysis. Recall that students' entry-level math classes are included directly into the dropout equation in addition to the teacher selection equations. Therefore, the instruments are just the projected math-course paths.

improve instrument performance. To see this, note that all of the instruments used here are based on the same underlying information about students' class schedules. For example, conditional on knowing the probability of a student taking algebra in year X, knowing the probability of her taking geometry will do little to better predict the probability of her being taught by an algebra teacher in year X. Adding additional instruments that provide little in terms of extra explanatory power can increase estimation bias (Murray, 2006). The interaction of students' class schedules with teachers' teaching schedules provides an intuitive and systematic approach to limiting the presence of largely irrelevant instruments in the teacher selection equations.

Equations (4) and (5) detail the instrument construction, performed within schools. Let h index a student's year-cohort based on the year in which he or she was in the ninth grade. Let D_{ihskt} be an indicator for whether student i in year-cohort h started high school in math-course s and took math-course k in year t of high school. Let H represent the number of cohorts, N_h the number of students in cohort h, and N_{hs} the number of students in cohort h who started high-school in math-course s. Then C_{ihskt} is defined as the predicted probability that student i in year-cohort h who started high school in math-course s took math-course k in year t, estimated using the course selections of students in the other year-cohorts who took the same entry-level math course:

$$C_{ihskt} = \frac{1}{H-1}\sum_{l \neq h}[(\frac{1}{N_{ls}})\sum_{n}^{N_l} D_{nlskt}] \qquad \textbf{(4)}$$

Let C_i be a kt-dimensional row vector where the entries indicate the probabilities of student i taking each math-course k in each year t, as calculated by (4). Let G_j be a kt-by-kt diagonal matrix where the diagonal entries are set to one in any subject-year that teacher j teaches and zero otherwise. The *kt*-dimensional row vector of instruments for any given student-teacher match, Z_{ij}, can be written:

$$Z_{ij} = C_i * G_j \tag{5}$$

The instrument vector in (5) includes zero entries in subject-years where teacher *j* does not teach and/or where student *i* does not take a given math course. The column rank of the instrument matrix for teacher *j*, Z_j, is K_j, where K_j is the number of subject-years where teacher j teaches.

II. Data

I use administrative data linking students and teachers at the classroom level from the San Diego Unified School District (SDUSD). SDUSD is the second largest school district in California (enrolling approximately 141,000 students in 1999-2000) and the student population is approximately 27 percent white, 37 percent Hispanic, 18 percent Asian/Pacific Islander and 16 percent black. Twenty-eight percent of the students at SDUSD are English Learners, and 60 percent are eligible for meal assistance. Both of these shares are larger than those of the state of California as a whole. As far as standardized testing performance, students at SDUSD trailed very slightly behind the national average in reading in 1999-2000. On the contrary, SDUSD students narrowly exceeded national norms in math.[11] The California Department of Education reported the 4-year derived dropout rate at SDUSD to be approximately 13 percent in 1999-2000.[12]

---

[11] District characteristics summarized from Betts et al. (2003).

[12] The dropout outcome may be measured with error as a result of the data collection process. When a student leaves the district, the district relies on the student's new school to request a transcript to verify that the student did not drop out. If there is no such request, SDUSD uses the available contact information for the student to attempt to determine whether a dropout has occurred. When transcripts are not requested and the student cannot be reached, the student will generally be considered a dropout. Hausman et al. (1998) show that probit estimates may be inconsistent when there is measurement error in the dependent variable. Although I also consider linear SIV

Because dropouts occur in all years of high school, it is important to observe each student from the 9th through (potentially) the 12th grade. The student data consist of four successive year-cohorts beginning with students in the 9th grade in 1997-1998 and ending with students in the 9th grade in 2000-2001. This latter group entered the 12th grade in the final year of the data panel, 2003-2004. To be included in the analysis, students had to be enrolled at SDUSD in the 9th grade.

SDUSD is a geographically large district with 16 standard, full-enrollment high schools.[13] However, there is considerable variation in the dropout rate across schools. For example, 4-year derived dropout rates at the school level ranged from less than one percent to over 20 percent in 1999-2000. These across-school differences in the dropout rate make it difficult to argue that teachers at each school at SDUSD are equally concerned with dropouts. That is, teachers at low-dropout-rate schools may not view deterring dropouts as a significant part of their job whereas teachers at high-dropout-rate schools, some of whom may watch one in five students fail to graduate, are unlikely to feel the same.

Because of these differences in dropout environments, teacher quality measured by dropout outcomes is likely to be a more relevant measure at high-dropout-rate schools.[14] With this in mind, I focus my analysis on the four schools that account for the most dropouts at SDUSD. The

---

models, evidence from Bhattacharya et al. (2006) indicates that linear-based models will produce substantially biased estimates of treatment effects under conditions similar to those found here.

[13] There are also some atypical schools at SDUSD that focus on helping at-risk students. I do not include these schools in the analysis.

[14] There is also reason to expect teacher quality to play a differential role in affecting dropout outcomes across schools on the student side. Specifically, students dropping out from low-dropout-rate schools are more likely to be extreme outliers whereas students dropping out from high-dropout-rate schools may be closer to the margin such that teacher quality may be more likely to make a difference.

dropout rate across these four schools ranges from 10 to 20 percent and these schools account for almost two-thirds of all dropouts from the 16 standard high schools at SDUSD.[15] I identify a student as being a part of school X's population if at any time in her schooling career she attended school X. I do not include special-education students in the analysis because these students often do not take typical math courses and are not exposed to the typical sets of math teachers at their schools.

I was unable to estimate the effects of all of the math teachers in the data because there were numerous teachers who taught just a small portion of the student sample. When these teachers were included into the multivariate model the likelihood function did not converge. Instead, I estimate teacher effects for the ten teachers at each school who taught the largest shares of the student population. These teachers taught between 43 and 59 percent of the math-class semesters taken by students at their respective schools over the course of the data panel (these shares are largely reflective of school size).[16] Within each school, the teacher effects are estimated relative to the average effect of the omitted teachers.[17]

Table 1 summarizes the data from each of the four schools. Note that schools 1, 2 and 3 are particularly disadvantaged (for example, see the shares of students participating in free/reduced-

[15] Although none of the schools at SDUSD qualifies as a "dropout factory" in the sense that less than half of the students graduate, the dropout rates at these schools are certainly non-negligible.

[16] There were four teachers who were in the top ten in terms of the number of student semesters taught at their respective schools but for whom I was unable to estimate an effect. For these four teachers, there was insufficient variation in their classes taught to identify teacher selection. In place of these teachers at their respective schools, I added the teachers who taught the next most students.

[17] The omitted teachers will bias the coefficients for the remaining teachers to the extent that the variation in classes taught between the sets of included and omitted teachers are correlated. Because of this, some of the individually estimated teacher coefficients may not be unbiased estimates for the effects of their respective teachers. However, the primary motivation in this analysis is to determine whether differences in teacher quality influence dropout outcomes – that is, to determine whether a distribution of teacher quality exists in terms of dropouts (see Section III). Therefore, teacher effects that are biased only though the omission of other teacher effects will still provide valuable insight as long as they are not systematically biased towards zero.

price lunch programs and test-score performance). School 4 is not nearly as disadvantaged as the other three schools based on conventional measures. Nonetheless, the dropout rate at school 4 is non-negligible and because of its size, it still accounts for a large fraction of the total number of dropouts at SDUSD.

III. Results

The idea of an absolute "teacher effect" is inconsequential because every student has a teacher. The question of interest is whether *differences* in teacher quality influence student outcomes. To answer this question in terms of dropouts, I estimate teacher effects from the model in Section I for 40 teachers at the four schools described in Table 1. As indicated above, these teacher effects are estimated relative to the average effect of the omitted teachers at each school where the omitted-teacher groups are comprised of the teachers who teach the fewest students. The extent to which the 40 estimated teacher effects differ from the within-school, average-omitted-teacher effects, and each other, will determine the extent to which differences in teacher quality influence dropout outcomes. If teacher quality did *not* influence dropout outcomes, all of the teacher effects would be statistically indistinguishable from zero.

I estimate the dropout model separately for each school, which means that the analysis evaluates within-school variation in teacher quality. The by-school approach is appealing given the computational demands of the multivariate probit. Furthermore, I reject the null hypothesis that the coefficients on the non-teacher variables in the models are equal across the four schools.[18] The primary implication of focusing on within-school variation in teacher quality is that my

[18] I test for parameter equality using a multivariate-probit model where I substitute for individual teachers with teacher qualifications (see Section VI). Estimating the four-school model with the 40 individual teacher effects using the multivariate-probit approach, which requires the simultaneous estimation of 41 equations, was not feasible.

results will understate the potential effects of changes in teacher quality on dropout outcomes to the extent that quality varies across schools.

Teacher effects are estimated from two different specifications at each school. First, I use a basic probit that ignores any endogeneity between teacher selection and dropout outcomes. Next, I run the multivariate probit described above in which I instrument for teacher selection and estimate teacher effects via simulated maximum likelihood.[19,20] In each model at each school, I reject the null hypothesis that the included teacher effects are jointly insignificant at the 5-percent level of confidence or better. The null hypothesis that the error terms in the dropout and teacher-selection equations are uncorrelated is rejected at the 1-percent level of confidence at all four schools.

Table 2 shows the non-teacher results from the models for school 1.[21] The non-teacher results for schools 2 through 4 are available in Appendix Tables A.1 – A.3. Although there is heterogeneity across the four schools in the parameter estimates for the non-teacher components

[19] I use the mvprobit module in Stata by Cappellari and Jenkins (2003) to estimate the model. This module uses estimates from the individual univariate probit specifications as initial parameter values.

[20] The econometrics literature provides limited coverage on the topics of instrument validity and relevance for models such as the multivariate probit (nonlinear, with multiple endogenous variables). With a lack of available formal tests for instrument evaluation, I make the following two assertions about the instruments used here: First, with regard to instrument relevance, univariate probits suggest that the instruments are strong predictors of teacher selection. In 39 of the 40 univariate-probit teacher-selection equations, Wald tests reject instrument irrelevance at the 1-percent level of confidence (the 40th equation rejects irrelevance at the 10-percent level). Although these tests are informative, note that they are imperfect because the instruments are predicting multiple endogenous variables and the univariate probits do not take this into account. Second, with regard to instrument validity, Hansen-J over-identification tests performed on the linear analogs to the multivariate probits provide no suggestion that the instruments are invalid. P-values from these tests range from 0.28 to 0.70 across the four schools. However, these tests are again imperfect. Work by Bhattacharya et al. (2006) shows that the linear-probability-model versions of the multivariate probits will perform poorly, particularly in the application here where dropout outcomes are relatively low-probability events.

[21] I do not report marginal effects here. Unlike for the teachers, calculating the marginal effects on dropout outcomes for the non-teacher components is cumbersome because each non-teacher component affects dropouts directly and through teacher selection. The results in Table 2 and the corresponding appendix tables are meant only to provide a basic overview of the models.

of the models, there are a few general patterns worth mentioning. First, as expected, the entry-level math courses are strong predictors of dropout outcomes with higher entry-level courses being associated with lower dropout rates. Second, conditional on these entry-level math courses, typically important demographic characteristics do not significantly predict dropouts. For example, parental education, race and gender indicators are generally not significant predictors of dropout outcomes once students' entry-level math courses are controlled for in the models. Finally, in addition to entry-level math courses and some teacher effects (as shown in Tables 3 – 6), the only other consistently significant predictor of students' dropout outcomes is English-learner status.

Tables 3 through 6 detail the estimated teacher effects at the four schools. Columns 1 and 2 report coefficient estimates from the basic models and columns 3 and 4 report estimates from the multivariate probits.[22] Teachers are ordered by their estimated marginal effects in the multivariate probits. At schools 1 through 4 respectively, 20, 14.8, 13.5 and 9.9 percent of the student samples ultimately drop out of school.

Across the four schools, 13 out of the 40 estimated math-teacher coefficients (or 33 percent) are statistically different from the average effect of the omitted teachers indicating that differences in teacher quality can indeed affect dropout outcomes. Furthermore, the magnitudes of the (marginal) teacher effects imply that they are economically meaningful, ranging from 4.2 to 14.1 percentage points. For example, at school 3, where 13.5 percent of the student sample ultimately

[22] Teachers' marginal effects are straightforward to calculate because teachers enter into the model as independent variables only in the dropout equation. Note that the reported marginal effects are calculated as the average of the marginal effects across students. Standard errors are approximated using the delta method where the explanatory variables are evaluated at their sample averages within each school.

drops out of school, five teachers have marginal effects that, relative to the average effect of the omitted teachers, are of a magnitude greater than 5 percentage points. These estimates imply a significant margin by which teacher quality can affect dropout outcomes.

The point estimates for the teacher effects are predominantly negative. Of the 13 statistically significant teacher effects, 12 are negative. For the remaining 27 teacher effects that are statistically insignificant, 21 of 27 have a negative point estimate. If these insignificant teacher effects were all true zeros, the probability of observing 21 of 27 negative point estimates by chance is less than one third of one percent. Overall, the results imply that the teachers who teach the most students at these schools are generally more effective at reducing dropout rates than those that teach the least (recall that the average-omitted-teacher effects are based on the teachers who teach the fewest students). There are numerous potential explanations for this. One possibility is that the results reflect selection. On the one hand, teachers who teach the fewest students at these low-performing schools may do so because they are not offered more classes than absolutely necessary by administrators because administrators knows they are of low quality. It may also be that the teachers who teach the most students select into teaching at these schools precisely because they are effective at deterring dropout outcomes. If this were the case, some of these teachers might actually choose to work with the most disadvantaged students at their schools. The empirical results are consistent with this hypothesis. The changes in the teacher-effect estimates when moving from the endogenous specification to the instrumental variables specification (columns 2 and 4 in the tables above) suggest that some of the teachers who are best at deterring dropout outcomes may be matched with students who are more likely to drop out (for example, see teachers 1 and 2 at school 2, teachers 1, 2, 3 and 4 at school 3, etc.).

Because the multivariate-probit estimates are much noisier than the basic-probit estimates, these changes in the point estimates are merely suggestive. Nonetheless, they could reflect a concerted effort by these teachers (and administrators) to deter dropouts.

A second possibility is that the results reflect the fact that the included teachers are more qualified, and more experienced, than the omitted teachers. I investigate this explanation in detail in Section VI.

IV. Potential Sources of Estimation Bias

Here I consider two potential sources of bias in the teacher-effect estimates from the dropout models. First, my empirical approach necessitates a focus on math teachers because variation in the classes taught by teachers over time, which is used to identify the teacher effects, is largest among math teachers. School administrators and popular media have argued that math is a decisive subject for students' graduation outcomes (Helfand, 2006). For example, speaking about algebra in 2006, Los Angeles Unified School District Superintendent Roy Romer was quoted as saying "It triggers dropouts more than any single subject." However, one concern with my approach is that exposure to an above-average math teacher may also imply exposure to an above-average English teacher, an above-average history teacher, and so on, which will influence dropout outcomes. On the one hand, even if the estimated math-teacher effects are biased by other teacher effects, the primary question of whether differences in teacher quality affect dropout outcomes can still be answered. Nonetheless, clearer inference can be made if bias from non-math teachers is minimized. There are three factors limiting, if not removing, this bias in my estimates. First, I include controls for students' entry-level math courses in the dropout equation (1), which will capture any across-subject tracking effects at the group level (note that teacher effects are identified from group-level student exposure). Second, because the teacher effects are identified from group-level student exposure, they should not be confounded by the potentially high correlation of teacher quality across subjects within students. Third, there is only very weak student tracking across subjects in San Diego secondary schools (Koedel, 2007). Given my identification strategy, if teacher quality in non-math subjects were to influence my results it would have to be the case that staffing changes in non-math subjects happened to occur as staffing changes occurred in math *and* students' group-level class-taking behaviors were highly correlated across subjects. For example, even with contemporaneous staffing changes across subjects, pre-algebra students would have to predominantly take the same English class(es) for the pre-algebra instrument to predict English-teacher selection with any relevance. This does not appear to be the case at SDUSD. Koedel (2007) provides empirical evidence indicating that math students are highly dispersed into the classrooms of other teachers in different subjects. One explanation for the wide dispersion is structural: math in secondary school is not a grade-level specific subject whereas most other subjects are.

A second concern with my empirical approach is that some teachers might only teach students at certain grade levels, which could create a mechanical relationship between teacher selection and dropouts. This concern is somewhat alleviated because math is not a grade-level-specific subject in secondary school, meaning that math teachers teach students in multiple grades, and because students drop out in all grades in high school.[23] More importantly, the exposure from which the teacher effects are identified is not grade-level specific. To see this, consider a teacher, teacher

[23] In fact, 95 percent of the math teachers evaluated in this study taught students in at least three out of the four grade levels in high school and 90 percent taught students in all four grade levels.

A, who taught algebra at school X in year Y. Even if all of the students taught by teacher A happened to be in the ninth grade, the instruments will identify teacher A's effect based on *all* students who took algebra in year Y, regardless of grade-level. That is, the instruments do not distinguish students' grade levels so teacher A's effect is identified based on all students whose cohort-specific math-course paths include some positive probability of taking algebra in year Y. Empirically, I can verify that teachers' grade-level shares do not predict their estimated dropout effects. After estimating the teacher effects from the multivariate probits above, I run a second-stage weighted regression where I regress the estimated teacher effects on school indicator variables and the shares of each teacher's students who are taught in each subject and at each grade level. For example, the grade-level share for the 11^{th} grade for teacher X is calculated as the total number of 11^{th}-grade students taught by teacher X divided by the total number of students taught by teacher X. The grade-level shares are all insignificant in this regression with t-statistics ranging from 0.5 to 0.7 in absolute value.

The result that the teacher-effect estimates are not influenced by students' grade levels is partly attributable to my focus on math teachers at disadvantaged schools. All of the math teachers at these schools teach some low-level classes (i.e., geometry and below) because there are not enough students taking high-level classes to allow teachers to specialize in these classes. Also, the low-level classes contain more students in higher grade levels because many students never take high-level math. Therefore, students at these disadvantaged schools are not heavily stratified across math teachers by grade level. If an analysis were based on more advantaged schools, where there were large populations of students taking advanced math classes, some teachers might only teach advanced classes with older students. Furthermore, low-level math classes at advantaged schools will likely contain fewer older students, implying that students will be more stratified by grade level across teachers. In this case, students' grade levels might be expected to have more influence on the estimated teacher effects and on the estimated variance of teacher quality (for example, differences in teacher quality among 12th-grade calculus teachers would be unlikely to affect dropout outcomes).

V. Mechanisms

Here I consider *how* variation in teacher quality might affect dropout outcomes. Specifically, I look to see if the dropout effects estimated in Section III can be linked to teachers' effects on student performance measured by grade reports or standardized test scores.

I estimate teacher effects on students' grade reports using the same instrumental-variables approach as the dropout models where I instrument for teacher selection using students' projected math-course paths. To maintain consistency with the dropout analysis, I evaluate teachers based on students' grade point averages (GPAs) for their entire high-school careers. Because GPAs are not binary outcomes, I revert to a linear specification and estimate teacher effects via GMM. I consider three models at each school where the dependent variables for each student are overall GPA in high school, overall math GPA in high school and overall non-math GPA in high school. This latter model is of particular interest because it excludes math teachers' direct effects on GPAs by excluding math grades. GPA data were not available for all students – I use just the fraction of the student sample for which I have grade-report data at each school.[24]

[24] The grade-report data are imperfect. In addition to some students not having grades, others only have grades for some classes. I aggregate students GPAs for their entire high-school careers at SDUSD which may somewhat offset the missing-data problems.

Figure 2 documents the teacher-effect results from the grade-report models for the four schools based on students' overall GPAs. Teachers are ordered as in Tables 3 through 6 above. Appendix B provides data tables corresponding to Figure 2. These tables also show each teacher's separate effects on math and non-math GPA outcomes.

Because I non-randomly chose the teachers for whom I estimate dropout effects, these teachers are likely to differ more from the group of omitted teachers than from each other. However, even among the group of 40 teachers that I evaluate here there is a clear negative relationship between teachers' dropout effects and GPA effects, implying a positive quality correlation. To evaluate this relationship, I divide each teacher effect in each model by its own standard error and estimate the correlations between the vectors of weighted GPA effects and weighted dropout effects. These correlations between the total-GPA/math-GPA/non-math-GPA effects and the dropout effects are -0.24/-0.08/-0.26. Furthermore, note that for the 13 teachers who have statistically significant dropout effects across the four schools, none have statistically significant GPA effects that would imply a negative quality correlation (i.e., same-signed dropout and GPA effects). In fact, 9 of the 13 have opposite-signed total-GPA effects and 11 of 13 have opposite-signed non-math-GPA effects (just 4 of which are statistically significant in each case). The finding that effective math teachers, measured by dropout rates, improve students' grades in non-math subjects suggests that these teachers are either improving students' cognitive skills in these subjects or generating an effort response that is reflected in the grade reports. The latter explanation is consistent with evidence from the education literature, which suggests that grade reports for disadvantaged students are strongly affected by student effort (more so than for their advantaged peers whose grade reports depend more on academic performance - see, for example, Howley et al., 2000, or Stiggins et al., 1989).

The GPA results also alleviate one possible concern with this analysis: that the teachers who are the most "effective" at reducing dropout rates are simply easy teachers who allow students to pass classes that they should perhaps not pass. Given that the correlation between teachers' dropout effects and non-math GPA effects is larger than the correlation between teachers' dropout effects and math-GPA effects, this explanation for the dropout effect seems unlikely. Furthermore, the difference between these correlations is even more pronounced among the subset of teachers who have statistically significant dropout effects. The correlation between these teachers' non-math-GPA effects and their dropout effects is -0.47. For their math-GPA effects, it is just -0.04. In conjunction with the non-math-GPA effects, the negligible math-GPA effects support work by Figlio and Lucas (2004), who show that teachers with higher grading standards produce better results for students. That is, if the non-math-GPA effects reflect increased student effort, the fact that this increased effort has no effect on students' math GPAs suggests that the most effective math teachers are not easy graders.

I also consider the extent to which teachers' dropout effects can be linked to their effects on standardized test scores. I use results from Koedel (2007) for comparison, who evaluates teacher value-added to test scores for secondary-school teachers at SDUSD. I compare the effects of the math teachers evaluated here, measured by dropout outcomes, to the effects of these same math teachers measured by students' test scores. The estimates from Koedel (2007) are based on math test scores from the ninth through eleventh grades on the Stanford 9 exam (students are not tested in the twelfth grade at SDUSD), which is not a high stakes test. The value-added comparison does not indicate any link between teacher effectiveness measured by test scores and teacher effectiveness measured by dropout outcomes. This might be expected if dropout effects reflect teachers' abilities to improve student effort but not cognitive performance, or if standardized tests are inflexible enough that they cannot detect improvements in schooling performance that might accompany a student's decision to complete high school.

Finally, note that in analyzing students' GPAs and test scores, teacher effects may be driven by the performance of a different population of students than in the case of dropouts. For example, if a given teacher is particularly effective with disadvantaged students but particularly ineffective with advantaged students, GPA and test-score analyses might indicate that this teacher, overall, is essentially average. However, in the case of dropouts, the advantaged students are not on the dropout margin to begin with so this teacher's dropout effect will be predominantly driven by her effect on disadvantaged students. Put differently, this teacher's advantaged students may underperform on standardized tests but are unlikely to drop out.

VI. Teacher Effectiveness and Resume Characteristics

The analysis thus far shows that individual teachers can have large effects on dropout outcomes. On the one hand, this information will be useful to policymakers in itself because it presents evidence of an additional dimension along which teacher quality affects student outcomes. However, it provides little direct guidance to policymakers hoping to identify effective teachers. In this section, I attempt to link teachers' dropout effects to their observable qualifications, which would allow administrators to more easily identify the teachers who will be effective in deterring dropouts.

Table 7 details the differences between the samples of included and omitted teachers in terms of observable qualifications. Given that the included teachers were chosen because they taught the most students, it is not surprising that they are more experienced. Also, a much larger share of the included teachers have the basic qualifications that are commonly associated with effective teaching in math.

To estimate the causal effects of teacher qualifications on dropout outcomes, I again use students' class schedules as in the dropout and grade-report analyses. However, here I predict student exposure to *teacher qualifications* rather than to specific teachers. I consider four teacher qualifications that are commonly associated with teacher quality. First, I look at exposure to math teachers who are inexperienced, where I define inexperienced teachers as those with three years of teaching experience or less. Second, I evaluate the effect of teachers' education levels as measured by whether teachers have master's degrees. Third, I consider two measures that might be associated with math-teacher quality specifically: whether teachers have bachelors' degrees in math and whether teachers are fully authorized to teach math. The latter qualification, the math authorization, requires that teachers complete a set of university courses prescribed by the California Commission on Teacher Credentialing. I do not look at the effect of exposure to fully credentialed teachers because over 94% of teachers at SDUSD are fully credentialed, leaving little variation to identify the effect of exposure to this qualification.

Unlike the teacher-specific analysis, it is unintuitive to model the exposure of students to teacher qualifications as binary. For example, more than half of the teachers across the four schools have math authorizations and almost half have master's degrees. Instead, I consider the effects of the *total numbers* of "qualified" teachers who teach each student as measured by the four criteria described above (or, in the case of teacher experience, the total number of "unqualified" teachers). The outcome of interest is still the binary dropout outcome; however, the teacher-qualification measures are not binary. Although a simple solution would be to specify the teacher-qualification and dropout equations as linear and run models that are otherwise analogous to the multivariate probit, Bhattacharya et al. (2006) indicates that such an approach will produce substantially biased estimates of treatment effects. As an alternative, I preserve the probit-based approach developed in Section I by re-classifying the teacher-qualification measures as binary. Specifically, at each school, I define the number of qualified teachers that teach each student as either "above-average" or "below-average" within that school.[25] I then use students' class schedules to instrument for exposure to an above-average number of qualified teachers in a 5-equation multivariate probit.[26] The equation of interest is again the dropout equation and there are four teacher-qualification equations. Table 8 details the estimated marginal effects of having been taught by an above-average number of qualified teachers as measured by each of the observable qualifications described above.

[25] I also consider fractional-response models where each teacher-qualification measure is coded as the share of the semesters of high-school in which a student is taught by a teacher with that qualification. These models produce results that are qualitatively similar to those presented below.

[26] Unlike in the teacher-specific multivariate probits, there is no obvious and consistent way to pare down the instrument sets for the teacher-qualification equations. Therefore, in the models for each school, I select four different but potentially overlapping subsets of the class-schedule instruments and run each qualifications model four times. The results from the four models within each school are virtually identical regardless of which subset of the instruments I use to identify the qualifications effects. This should be expected if the instruments are truly valid given that they are identifying teacher effects from the same underlying source of variation. Because the results are so similar across the four models at each school, there is no clear justification for presenting the results based on one subset of instruments over another so I simply chose one set of results to show for each school in Table 8.

Table 8 suggests that students who are exposed to sets of more-educated teachers are more likely to graduate. The teacher-education effect is particularly pronounced at schools 3 and 4. However, with regard to the other teacher qualifications, no clear patterns emerge. Note that exposure to more inexperienced teachers seems to *reduce* dropout rates at school 3, which contradicts both intuition and a substantial body of empirical evidence showing that inexperienced teachers are generally less effective (for example, see Rivkin, Hanushek and Kain, 2005). One explanation for this finding is that the results in Table 8 are biased by negative selection. Recall that the four schools evaluated here are four of the most disadvantaged schools in the district. Like most school districts, teachers with better qualifications are generally favored in the hiring process at SDUSD. If teaching positions at these disadvantaged schools are generally undesirable, highly-qualified teachers who teach at these schools may be the least able as evidenced by their inability to obtain employment at better schools.[27] Given the possibility of negative-selection bias, it is impossible to rule out causal teacher-qualification effects. Furthermore, to the extent such bias does exist, it strengthens the evidence that teacher education may be a particularly effective tool in reducing dropouts.

VII. Concluding Remarks

This paper uses a unique instrumental-variables approach to identify teacher effects on students' dropout outcomes. Differences in teacher quality are shown to play an important role in determining these outcomes, implying that improvements in teacher quality can help mitigate the graduation-rate crisis faced by many urban school districts across the country (as dubbed by Orfield et al., 2004). However, the current structure of teacher recruitment and compensation in

[27] As discussed in Section III, it is also possible that some teachers prefer teaching at these schools.

most school districts, which generally provides little incentive for teachers to teach at disadvantaged schools, has resulted in the consistent failure of urban schools to attract and retain high-quality teachers (Lankford et al., 2002). This will limit the extent to which teacher-quality improvements can be used to combat high dropout rates. Ultimately, the evidence here suggests that increasing the ability of urban schools to recruit and retain high-quality teachers has the potential to reduce student dropout rates significantly.

Figure 1. Examples of Variation in the Classes Taught Over Time, by Subject, for Four Teachers at School 1

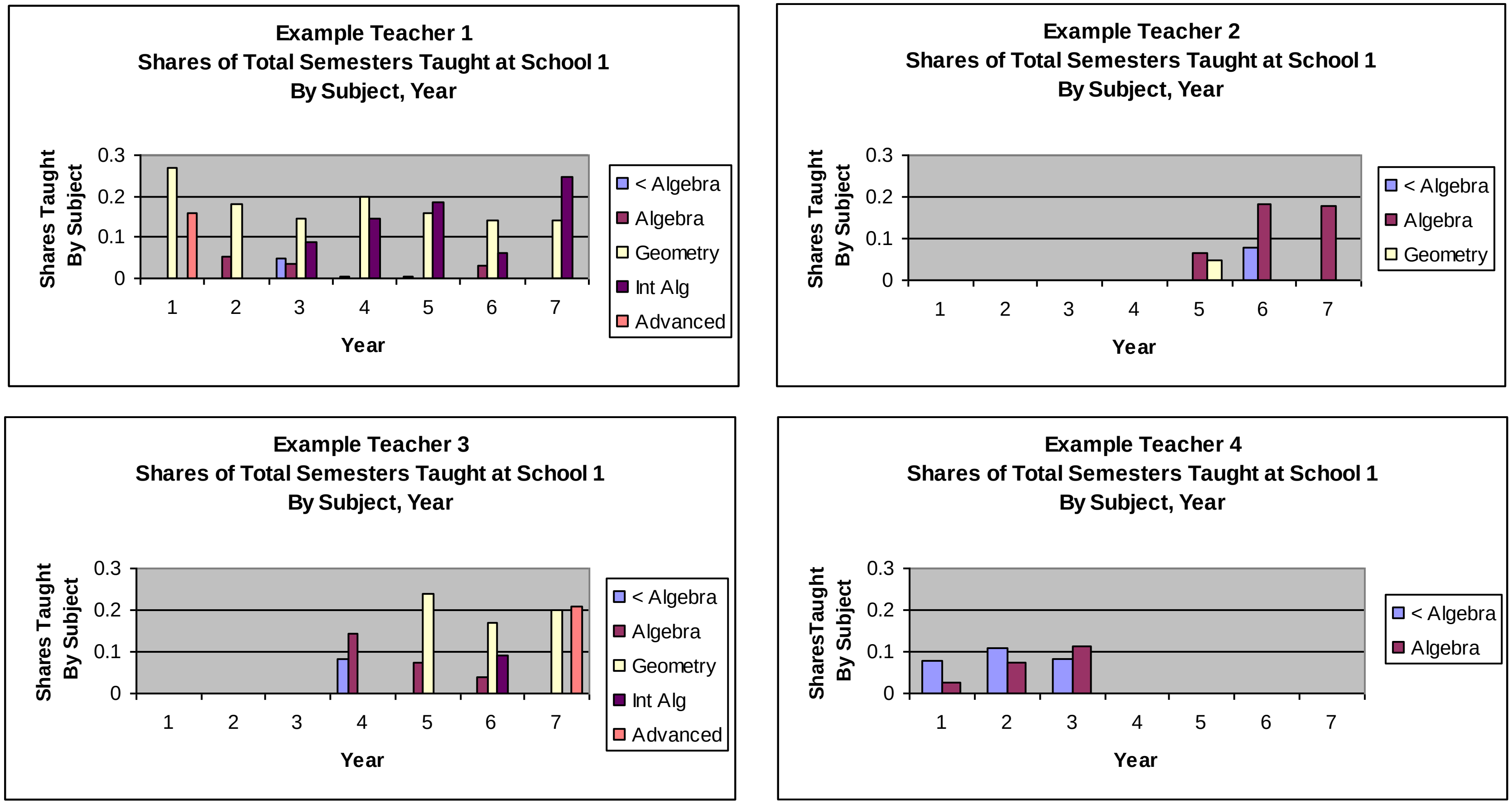
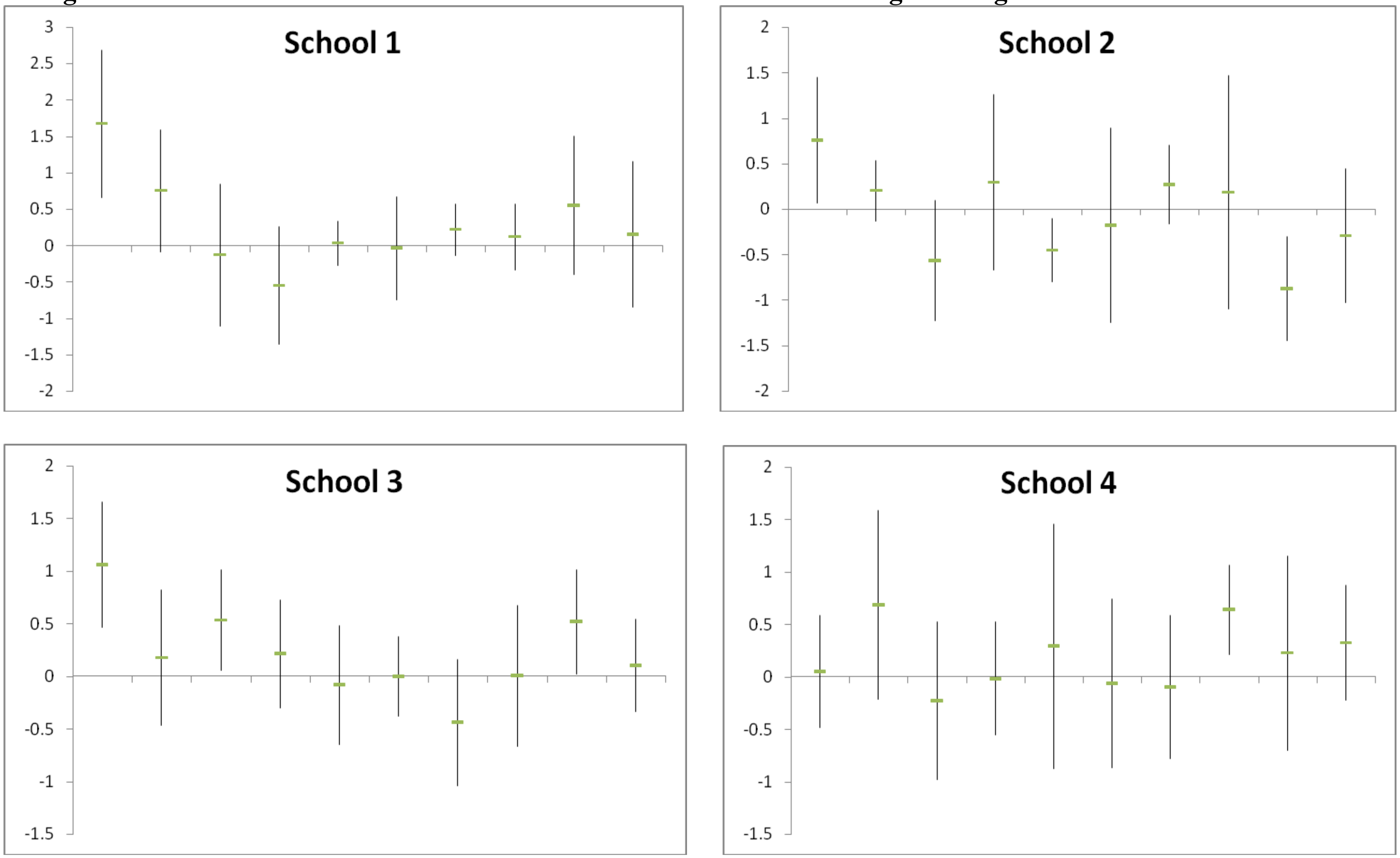


Figure 2. Teacher Effects on Students' Cumulative Grade Point Averages in High School for all Schools



Note: Figure shows point estimates and corresponding 95-percent confidence intervals.

Table 1. Basic Data Summary

	School 1	School 2	School 3	School 4
Initial Year of Data Panel	1997-1998	1997-1998	1997-1998	1997-1998
Final Year of Data Panel	2003-2004	2003-2004	2003-2004	2003-2004
Total Number of Students in Data Panel	3072	2518	2217	3779
% Female	48	51	48	49
% English Learner	49	42	49	15
% Black	22	13	29	21
% Hispanic	57	68	36	22
% Asian	15	3	27	50
% White	6	15	8	7
% Parental Ed = College or more	7	9	7	26
% Free/Reduced-Price Lunch*	100	63	83	44
Average Test-Scores (Math)**	-0.46	-0.14	-0.46	0.04
% of Students Who Ultimately Drop Out	0.200	0.148	0.135	0.099
Share of Student Semesters Taught by the Top 10 Teachers	0.514	0.547	0.586	0.428

*Total share of students on free or reduced-price lunch at school level in the 1999-2000 school year
** Average score for 9^{th}-grade test takers at each school over the course of data panel relative to the 9^{th}-grade district-wide average, reported in standard deviations of the test. Based on all students who took the test at the relevant school in the 9^{th} grade.

Table 2. Non-Teacher Results from School 1 - Dependent Variable: Indicator for Whether a Dropout Occurred

Variable	Basic Probit	Multivariate IV Probit
English Learner (EL)	0.186	0.206
	(0.063)***	(0.071)**
Re-designated from EL to non-EL	-0.606	-0.594
During High School	(0.118)***	(0.121)***
Female	-0.093	-0.083
	(0.054)*	(0.054)
Asian	0.016	0.01
	(0.149)	(0.143)
Black	-0.024	-0.038
	(0.136)	(0.136)
Hispanic	0.216	0.176
	(0.133)	(0.133)
Max Parental Ed = High School	-0.060	-0.061
	(0.126)	(0.124)
Max Parental Ed = Some College	-0.241	-0.202
	(0.156)	(0.157)
Max Parental Ed = College Graduate	-0.101	-0.132
	(0.165)	(0.165)
Max Parental Ed = Graduate School	-0.853	-0.825
	(0.464)*	(0.456)*
Max Parental Ed = Unknown	0.026	0.075
	(0.088)	(0.096)
Student changed schools mid-year at	0.166	0.162
some point during high school	(0.142)	(0.149)
9^{th} Grade Math = No Math	-0.156	-0.148
	(0.121)	(0.123)
9^{th} Grade Math = Part Algebra, Part	-0.159	-0.158
Pre-Algebra	(0.081)**	(0.081)*
9^{th} Grade Math = Algebra	-0.183	-0.168
	(0.068)***	(0.073)**
9^{th} Grade Math = Part Algebra, Part	-0.301	-0.327
Geometry	(0.178)*	(0.191)*
9^{th} Grade Math = Geometry	-0.489	-0.385
	(0.198)**	(0.205)*
9^{th} Grade Math = Advanced Geometry	-0.420	-0.238
	(0.210)**	(0.227)
Constant	-0.771	-0.813
	(0.154)***	(0.174)***
Number of student observations	3072	3072

Notes: Standard errors in parentheses. All students who took intermediate algebra (> advanced geometry) in 9^{th} grade graduated high school. Omitted variables are: Indicator variables for non-EL, non re-designated non-EL, white, male, parental education is high school dropout, 9^{th} grade math class is pre-algebra and all teachers other than those listed in Table 2.
***Significant at 1% level of confidence.
**Significant at 5% level of confidence.
*Significant at 10% level of confidence.

Table 3. Results from School 1 - Dependent Variable: Indicator for Whether a Dropout Occurred

Teacher	Basic Probit	Basic Probit (Marginal Effects)	Multivariate IV Probit	Multivariate IV Probit (Marginal Effects)
Teacher 1	-0.395	-0.090	-0.674	-0.141
	(0.105)***	(0.020)***	(0.179)***	(0.057)**
Teacher 2	0.109	0.030	-0.114	-0.029
	(0.078)	(0.022)	(0.163)	(0.042)
Teacher 3	-0.045	-0.012	-0.112	-0.029
	(0.076)	(0.019)	(0.162)	(0.042)
Teacher 4	-0.189	-0.047	-0.108	-0.028
	(0.077)**	(0.018)**	(0.150)	(0.039)
Teacher 5	-0.021	-0.005	-0.103	-0.026
	(0.073)	(0.019)	(0.147)	(0.038)
Teacher 6	0.013	0.003	-0.025	-0.007
	(0.082)	(0.022)	(0.167)	(0.043)
Teacher 7	-0.213	-0.052	0.062	0.017
	(0.097)**	(0.022)**	(0.172)	(0.045)
Teacher 8	-0.127	-0.032	0.108	0.029
	(.072)*	(0.017)*	(0.154)	(0.041)
Teacher 9	-0.007	-0.002	0.144	0.039
	(0.082)	(0.022)	(0.176)	(0.047)
Teacher 10	0.210	0.059	0.359	0.104
	(0.078)***	(0.023)**	(0.168)**	(0.050)**
Number of student observations	3072	3072	3072	3072

Notes: Estimates are relative to the omitted group of teachers as described in the text. Standard errors are in parentheses.

***Significant at 1% level of confidence
**Significant at 5% level of confidence
*Significant at 10% level of confidence

Table 4. Results from School 2 - Dependent Variable: Indicator for Whether a Dropout Occurred

Teacher	Basic Probit	Basic Probit (Marginal Effects)	Multivariate IV Probit	Multivariate IV Probit (Marginal Effects)
Teacher 1	-0.479	-0.079	-0.587	-0.102
	(0.112)***	(0.015)***	(0.174)***	(0.042)**
Teacher 2	-0.268	-0.048	-0.355	-0.067
	(0.106)**	(0.017)**	(0.163)**	(0.035)*
Teacher 3	-0.107	-0.020	-0.228	-0.045
	(0.099)	(0.018)	(0.167)	(0.035)
Teacher 4	0.116	0.024	-0.216	-0.042
	(0.112)	(0.025)	(0.202)	(0.042)
Teacher 5	-0.087	-0.017	-0.179	-0.035
	(0.108)	(0.020)	(0.184)	(0.038)
Teacher 6	-0.154	-0.029	-0.168	-0.034
	(0.082)*	(0.015)*	(0.139)	(0.028)
Teacher 7	0.004	0.001	-0.123	-0.025
	(0.094)	(0.019)	(0.171)	(0.035)
Teacher 8	-0.173	-0.032	-0.118	-0.024
	(0.111)	(0.019)	(0.207)	(0.042)
Teacher 9	-0.214	-0.039	-0.067	-0.014
	(0.087)**	(0.015)**	(0.160)	(0.032)
Teacher 10	0.164	0.034	0.151	0.033
	(0.101)	(0.022)	(0.171)	(0.037)
Number of student observations	2518	2518	2518	2518

Notes: Estimates are relative to the omitted group of teachers as described in the text. Standard errors are in parentheses.

***Significant at 1% level of confidence
**Significant at 5% level of confidence
*Significant at 10% level of confidence

Table 5. Results from School 3 - Dependent Variable: Indicator for Whether a Dropout Occurred

Teacher	Basic Probit	Basic Probit (Marginal Effects)	Multivariate IV Probit	Multivariate IV Probit (Marginal Effects)
Teacher 1	-0.480	-0.073	-0.573	-0.097
	(0.104)***	(0.013)***	(0.159)***	(0.037)**
Teacher 2	-0.415	-0.060	-0.565	-0.089
	(0.140)***	(0.016)***	(0.185)***	(0.042)**
Teacher 3	-0.266	-0.042	-0.453	-0.077
	(0.115)**	(0.016)**	(0.183)**	(0.038)**
Teacher 4	-0.236	-0.039	-0.343	-0.061
	(0.103)**	(0.015)**	(0.165)**	(0.033)*
Teacher 5	-0.360	-0.055	-0.319	-0.058
	(0.107)***	(0.014)***	(0.182)*	(0.035)*
Teacher 6	-0.044	-0.007	-0.239	-0.044
	(0.094)	(0.016)	(0.166)	(0.032)
Teacher 7	-0.060	-0.011	-0.041	-0.008
	(0.0106)	(0.018)	(0.178)	(0.033)
Teacher 8	-0.218	-0.034	-0.023	-0.005
	(0.138)	(0.020)	(0.209)	(0.039)
Teacher 9	-0.012	-0.003	0.058	0.012
	(0.122)	(0.021)	(0.176)	(0.033)
Teacher 10	0.023	0.004	0.273	0.059
	(0.107)	(0.019)	(0.182)	(0.037)
Number of student observations	2217	2217	2217	2217

Notes: Estimates are relative to the omitted group of teachers as described in the text. Standard errors are in parentheses.

***Significant at 1% level of confidence
**Significant at 5% level of confidence
*Significant at 10% level of confidence

Table 6. Results from School 4 - Dependent Variable: Indicator for Whether a Dropout Occurred

Teacher	Basic Probit	Basic Probit (Marginal Effects)	Multivariate IV Probit	Multivariate IV Probit (Marginal Effects)
Teacher 1	-0.496	-0.052	-0.549	-0.064
	(0.117)***	(0.009)***	(0.161)***	(0.026)**
Teacher 2	-0.475	-0.050	-0.416	-0.052
	(0.118)***	(0.009)***	(0.157)***	(0.024)**
Teacher 3	-0.323	-0.037	-0.313	-0.042
	(0.096)***	(0.009)***	(0.156)*	(0.022)*
Teacher 4	-0.382	-0.044	-0.292	-0.040
	(0.087)***	(0.008)***	(0.140)**	(0.020)**
Teacher 5	-0.395	-0.042	-0.289	-0.038
	(0.134)***	(0.011)***	(0.226)	(0.031)
Teacher 6	-0.202	-0.025	-0.170	-0.024
	(0.095)**	(0.010)**	(0.154)	(0.021)
Teacher 7	-0.194	-0.024	-0.105	-0.015
	(0.118)*	(0.013)*	(0.191)	(0.026)
Teacher 8	-0.082	-0.011	-0.090	-0.013
	(0.089)	(0.011)	(0.148)	(0.020)
Teacher 9	-0.064	-0.008	-0.079	-0.012
	(0.088)	(0.011)	(0.146)	(0.020)
Teacher 10	-0.162	-0.019	-0.032	-0.005
	(0.100)	(0.011)	(0.168)	(0.023)
Number of student observations	3779	3779	3779	3779

Notes: Estimates are relative to the omitted group of teachers as described in the text. Standard errors are in parentheses.

***Significant at 1% level of confidence

**Significant at 5% level of confidence

*Significant at 10% level of confidence

Table 7. Differences in Observable Qualifications Between the Included and Omitted Teachers – All Schools

	Included Teachers	Omitted Teachers
Average Experience	13.30	10.90
Share With > 3 Years Average Experience	0.87	0.68
Share with Math Authorization	0.84	0.56
Share with BA in Math	0.69	0.38
Share with Master's Degree	0.48	0.43
N	40	115

Notes: Qualifications are averaged within teachers where relevant. For example, if a teacher taught for two years in the panel then that teacher's experience level would be averaged across the two years. To be included in the "omitted teachers" group in the table, teachers had to teach at least 50 student semesters in math over the course of the data panel. The "math authorization" requires that teachers complete a set of university course requirements as prescribed by the California Commission on Teacher Credentialing.

Table 8. The Effects of Exposure to an Above-Average Number of Qualified Teachers on Students' Dropout Outcomes

	School 1	School 2	School 3	School 4
Qualifications				
Master's Degree	-0.054	-0.039	-0.100	-0.050
	(0.039)	(0.026)	(0.034)***	(0.018)***
Bachelor's Degree in Math	-0.040	-0.048	-0.030	0.019
	(0.032)	(0.028)*	(0.027)	(0.019)
Full Math Authorization	0.000	-0.027	-0.034	-0.012
	(0.045)	(0.031)	(0.029)	(0.018)
Inexperienced Teacher	-0.012	0.017	-0.080	-0.016
(3 years or less experience)	(0.033)	(0.026)	(0.028)***	(0.014)
N	3072	2518	2217	3779

Notes: Marginal effects are reported and estimated as the average of the marginal effects across the student sample.
Standard errors in parenthesis.
***Significant at 1% level of confidence
**Significant at 5% level of confidence
*Significant at 10% level of confidence

References

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

[2] Ashenfelter, Orley and Alan Krueger, Estimates of the Economic Return to Schooling from a New Sample of Twins, *American Economic Review* 84 (1994) 1157-73.

[3] Barrow, Lisa and Cecilia Elena Rouse, The Economic Value of Education by Race and Ethnicity, *Economic Perspectives* 30 (2006) 14-27.

[4] Betts, Julian R., Andrew Zau and Lorien Rice, *Determinants of Student Achievement, New Evidence from San Diego*, Public Policy Institute of California (2003).

[5] Bhattacharya, Jay, Dana P. Goldman, and Daniel McCaffrey, Estimating Probit Models with Self-Selected Treatments, *Statistics in Medicine* 25 (2006) 389-413.

[6] Cappellari, Lorenzo and Stephen P. Jenkins, Multivariate Probit Regression Using Simulated Maximum Likelihood, *The Stata Journal* 3 (2003) 278-294.

[7] Clotfelter, Charles T., Helen F. Ladd and Jacob L. Vigdor, Teacher-Student Matching and the Assessment of Teacher Effectiveness, *Journal of Human Resources* 41 (2006) 778 - 820.

[8] Figlio, David and Maurice Lucas, Do High Grading Standards Affect Student Performance? *Journal of Public Economics* 88 (2004) 1815-34.

[9] Hanushek, Eric, John Kain, Daniel O'Brien and Steven Rivkin, The Market for Teacher Quality, NBER WP 11154 (2005).

[10] Hausman, Jerry A., Jason Abrevaya and Fiona M. Scott-Morton, Misclassification in the Dependent Variable in a Discrete Response Setting, *Journal of Econometrics* 87 (1998) 239-269.

[11] Helfand, Duke, A Formula for Failure in LA Schools, *Los Angeles Times* (1-30-2006).

[12] Howley, Aimee, Patricia S. Kusimo and Laurel Parrott, Grading and the Ethos of Effort, *Learning Environments Research* 3 (2000) 229-246.

[13] Koedel, Cory, Teacher Quality and Joint Production in Secondary School, University of Missouri Working Paper (2007).

[14] Lankford, Hamilton, Susanna Loeb and James Wyckoff, Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis, *Education Evaluation and Policy Analysis* 24 (2002) 37-62.

[15] Lochner, Lance and Enrico Moretti, The Effect of Education on Crime: Evidence from Prison Inmates, Arrests and Self-Reports, *American Economic Review* 94 (2004) 155-89.

[16] Loeb, Susanna and Marianne E. Page, Examining the Link between Teacher Wages and Student Outcomes: The Importance of Alternative Labor Market Opportunities and Non-Pecuniary Variation, *The Review of Economics and Statistics* 82 (2000) 393-408.

[17] Murray, Michael P., Avoiding Invalid Instruments and Coping with Weak Instruments, *Journal of Economic Perspectives* 20 (2006) 111-132.

[18] Nye, Barbara, Spyros Konstantopoulos and Larry V. Hedges, How large are teacher effects? *Educational Evaluation and Policy Analysis* 26 (2004) 237-257.

[19] Orfield, Gary, Daniel Losen, Johanna Wald and Christopher B. Swanson, *Losing our Future: How Minority Youth Are Being Left Behind by the Graduation Rate Crisis*, Cambridge, MA: The Civil Rights Project at Harvard University. Contributors: Advocates for Children of New York, The Civil Society Institute (2004).

[20] Rivkin, Steven, Eric Hanushek and John Kain, Teachers, Schools and Academic Achievement, *Econometrica* 79 (2005) 417-58.

[21] Rockoff, Jonah, The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data, *American Economic Review*, Papers and Proceedings (2004).

[22] Rothstein, Jesse, Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement, Manuscript, Princeton University (2008).

[23] Rouse, Cecilia Elena, Further Estimates of the Economic Return to Schooling from a New Sample of Twins, *Economics of Education Review* 18 (1999) 149-157.

[24] Stiggins, R. J., Frisbie, D.A. and Griswold, P.A., Inside High School Grading Practices: Building a Research Agenda, *Educational Measurement: Issues and Practice* 8 (1989) 5-14.

Appendix A
Non-Teacher Coefficient Estimates from Models for Schools 2, 3 and 4

Table A.1. Non-Teacher Results from School 2 - Dependent Variable: Indicator for Whether a Dropout Occurred

Variable	Basic Probit	Multivariate IV Probit
English Learner (EL)	0.339	0.334
	(0.080)***	(0.083)***
Re-designated from EL to non-EL	-0.466	-0.438
During High School	(0.099)***	(0.101)***
Female	-0.037	-0.027
	(0.067)	(0.066)
Asian	-0.423	-0.453
	(0.320)	(0.319)
Black	-0.025	-0.024
	(0.148)	(0.150)
Hispanic	0.202	0.217
	(0.130)	(0.129)
Max Parental Ed = High School	0.120	0.112
	(0.153)	(0.153)
Max Parental Ed = Some College	0.088	0.047
	(0.182)	(0.181)
Max Parental Ed = College Graduate	-0.362	-0.405
	(0.224)	(0.224)*
Max Parental Ed = Graduate School	-0.007	-0.011
	(0.33)	(0.366)
Max Parental Ed = Unknown	0.037	-0.023
	(0.114)	(0.115)
Student changed schools mid-year at	0.002	-0.103
some point during high school	(0.205)	(0.215)
9th Grade Math = No Math	-0.546	0.04
	(0.330)*	(0.25)
9th Grade Math = Part Algebra, Part	-0.529	-0.507
Pre-Algebra	(0.163)***	(0.165)***
9th Grade Math = Algebra	-0.322	-0.247
	(0.082)***	(0.093)***
9th Grade Math = Geometry	-0.558	-0.496
	(0.239)**	(0.247)**
9th Grade Math = Advanced Geometry	-0.805	-0.713
	(0.149)***	(0.165)***
Constant	-0.773	-0.716
	(0.162)***	(0.179)***
Number of student observations	2518	2518

Notes: Standard errors in parentheses. All students who took intermediate algebra (> advanced geometry) in 9th grade graduated high school. In addition, only six students took the algebra-geometry split at school 2 so that control is omitted from the model. Other omitted variables are: Indicator variables for non-EL, non re-designated non-EL, white, male, parental education is high school dropout, 9th grade math class is pre-algebra and all teachers other than those listed in Table 3.

***Significant at 1% level of confidence.

**Significant at 5% level of confidence.

*Significant at 10% level of confidence.

Table A.2. Non-Teacher Results from School 3 - Dependent Variable: Indicator for Whether a Dropout Occurred

Variable	Basic Probit	Multivariate IV Probit
English Learner (EL)	0.113	0.054
	(0.085)	(0.087)
Re-designated from EL to non-EL	-0.367	-0.321
During High School	(0.120)***	(0.123)**
Female	-0.113	-0.115
	(0.073)	(0.073)
Asian	-0.036	-0.022
	(0.165)	(0.169)
Black	-0.127	-0.145
	(0.155)	(0.154)
Hispanic	0.354	0.344
	(0.155)**	(0.154)**
Max Parental Ed = High School	-0.256	-0.206
	(0.172)	(0.170)
Max Parental Ed = Some College	-0.240	-0.194
	(0.187)	(0.187)
Max Parental Ed = College Graduate	-0.056	-0.046
	(0.205)	(0.209)
Max Parental Ed = Graduate School	-0.478	-0.393
	(0.570)	(0.561)
Max Parental Ed = Unknown	-0.047	-0.035
	(0.118)	(0.125)
Student changed schools mid-year at	0.397	0.374
some point during high school	(0.195)**	(0.205)*
9^{th} Grade Math = No Math	-0.134	-0.198
	(0.147)	(0.157)
9^{th} Grade Math = Part Algebra, Part	-0.0878	-0.111
Pre-Algebra	(0.101)	(0.106)
9^{th} Grade Math = Algebra	-0.292	-0.376
	(0.117)**	(0.130)***
9^{th} Grade Math = Part Algebra, Part	-0.380	-0.435
Geometry	(0.388)	(0.378)
9^{th} Grade Math = Geometry	-0.865	-0.968
	(0.356)**	(0.360)***
9^{th} Grade Math = Advanced Geometry	-0.981	-1.02
	(0.299)***	(0.305)***
Constant	-0.713	-0.606
	(0.187)	(0.212)
Number of student observations	2217	2217

Notes: Standard errors in parentheses. All students who took intermediate algebra (> advanced geometry) in 9^{th} grade graduated high school. Omitted variables are: Indicator variables for non-EL, non re-designated non-EL, white, male, parental education is high school dropout, 9^{th} grade math class is pre-algebra and all teachers other than those listed in Table 4.

***Significant at 1% level of confidence.

**Significant at 5% level of confidence.

*Significant at 10% level of confidence.

Table A.3. Non-Teacher Results from School 4 - Dependent Variable: Indicator for Whether a Dropout Occurred

Variable	Basic Probit	Multivariate IV Probit
English Learner (EL)	0.616	0.622
	(0.088)***	(0.089)***
Re-designated from EL to non-EL	-0.486	-0.487
During High School	(0.131)***	(0.132)***
Female	-0.066	-0.070
	(0.061)	(0.061)
Asian	-0.188	-0.197
	(0.129)	(0.130)
Black	0.059	0.053
	(0.133)	(0.133)
Hispanic	-0.047	-0.052
	(0.136)	(0.137)
Max Parental Ed = High School	0.065	0.075
	(0.142)	(0.142)
Max Parental Ed = Some College	-0.109	-0.110
	(0.135)	(0.136)
Max Parental Ed = College Graduate	-0.090	-0.089
	(0.133)	(0.133)
Max Parental Ed = Graduate School	0.088	0.088
	(0.285)	(0.284)
Max Parental Ed = Unknown	-0.265	-0.245
	(0.129)**	(0.130)*
Student changed schools mid-year at	-0.005	0.044
some point during high school	(0.201)	(0.200)
9^{th} Grade Math = No Math	0.051	0.057
	(0.235)	(0.238)
9^{th} Grade Math = Part Algebra, Part	0.149	0.134
Pre-Algebra	(0.165)	(0.165)
9^{th} Grade Math = Algebra	-0.345	-0.352
	(0.072)***	(0.076)***
9^{th} Grade Math = Part Algebra, Part	-0.025	-0.071
Geometry	(0.701)	(0.710)
9^{th} Grade Math = Geometry	-0.296	-0.321
	(0.186)	(0.190)*
9^{th} Grade Math = Advanced Geometry	-1.078	-1.073
	(0.180)***	(0.193)***
Constant	-0.588	-0.638
	(0.163)***	(0.169)***
Number of student observations	3779	3779

Notes: Standard errors in parentheses. All students who took intermediate algebra (> advanced geometry) in 9^{th} grade graduated high school. Omitted variables are: Indicator variables for non-EL, non re-designated non-EL, white, male, parental education is high school dropout, 9^{th} grade math class is pre-algebra and all teachers other than those listed in Table 5.
***Significant at 1% level of confidence.
**Significant at 5% level of confidence.
*Significant at 10% level of confidence.

Appendix B
Teacher-Effect Estimates from the GPA Analysis

Table B.1. Comparison of Teachers' Dropout Effects and GPA Effects at School 1

Teacher	Multivariate IV Probit - Dropout (Marginal Effects)	Total GPA Effect	Math GPA Effect	Non-Math GPA Effect
Teacher 1	-0.141	1.674	1.417	1.790
	(0.057)**	(0.516)***	(0.609)**	(0.542)***
Teacher 2	-0.029	0.755	0.705	0.785
	(0.042)	(0.429)*	(0.470)	(0.448)*
Teacher 3	-0.029	-0.130	-0.572	-0.028
	(0.042)	(0.498)	(0.552)	(0.521)
Teacher 4	-0.028	-0.549	-0.368	-0.674
	(0.039)	(0.412)	(0.427)	(0.427)
Teacher 5	-0.026	0.032	0.071	0.043
	(0.038)	(0.155)	(0.179)	(0.161)
Teacher 6	-0.007	-0.037	-0.297	0.134
	(0.043)	(0.361)	(0.412)	(0.375)
Teacher 7	0.017	0.220	0.084	0.265
	(0.045)	(0.180)	(0.196)	(0.191)
Teacher 8	0.029	0.119	-0.169	0.219
	(0.041)	(0.231)	(0.266)	(0.241)
Teacher 9	0.039	0.553	1.030	0.433
	(0.047)	(0.488)	(0.533)**	(0.500)
Teacher 10	0.104	0.151	0.830	-0.062
	(0.050)**	(0.510)	(0.559)	(0.528)
Number of student observations	3072	2938	2938	2938

Notes: Estimates are relative to the omitted group of teachers as described in the text. Standard errors are in parentheses.

***Significant at 1% level of confidence
**Significant at 5% level of confidence
*Significant at 10% level of confidence

Table B.2. Comparison of Teachers' Dropout Effects and GPA Effects at School 2

Teacher	Multivariate IV Probit - Dropout (Marginal Effects)	Total GPA Effect	Math GPA Effect	Non-Math GPA Effect
Teacher 1	-0.102	0.758	0.121	0.821
	(0.042)**	(0.351)**	(0.351)	(0.352)**
Teacher 2	-0.067	0.206	0.267	0.192
	(0.035)**	(0.170)	(0.199)	(0.170)
Teacher 3	-0.045	-0.565	-1.063	-0.486
	(0.035)	(0.337)*	(0.382)***	(0.336)
Teacher 4	-0.042	0.294	0.528	0.225
	(0.042)	(0.491)	(0.559)	(0.494)
Teacher 5	-0.035	-0.451	-0.498	-0.424
	(0.038)	(0.177)**	(0.224)**	(0.178)**
Teacher 6	-0.034	-0.178	0.197	-0.274
	(0.028)	(0.547)	(0.653)	(0.544)
Teacher 7	-0.025	0.271	0.426	0.246
	(0.035)	(0.221)	(0.271)	(0.222)
Teacher 8	-0.024	0.183	-0.323	0.333
	(0.042)	(0.655)	(0.781)	(0.659)
Teacher 9	-0.014	-0.875	-0.121	-0.813
	(0.032)	(0.291)***	(0.351)	(0.289)***
Teacher 10	0.033	-0.293	0.004	-0.366
	(0.037)	(0.375)	(0.454)	(0.374)
Number of student observations	2518	2513	2513	2513

Notes: Estimates are relative to the omitted group of teachers as described in the text. Standard errors are in parentheses.

***Significant at 1% level of confidence
**Significant at 5% level of confidence
*Significant at 10% level of confidence

Table B.3. Comparison of Teachers' Dropout Effects and GPA Effects at School 3

Teacher	Multivariate IV Probit - Dropout (Marginal Effects)	Total GPA Effect	Math GPA Effect	Non-Math GPA Effect
Teacher 1	-0.097	1.06	1.180	0.949
	(0.037)**	(0.303)***	(0.397)***	(0.291)***
Teacher 2	-0.089	0.177	-0.067	0.171
	(0.042)**	(0.329)	(0.491)	(0.320)
Teacher 3	-0.077	0.533	0.519	0.490
	(0.038)*	(0.245)**	(0.403)	(0.226)**
Teacher 4	-0.061	0.215	0.074	0.219
	(0.033)*	(0.261)	(0.351)	(0.256)
Teacher 5	-0.058	-0.080	-0.579	0.080
	(0.035)*	(0.289)	(0.370)	(0.297)
Teacher 6	-0.044	-0.003	-0.038	-0.067
	(0.032)	(0.193)	(0.356)	(0.178)
Teacher 7	-0.008	-0.440	-0.924	-0.313
	(0.033)	(0.305)	(0.402)**	(0.297)
Teacher 8	-0.005	0.004	-0.227	0.048
	(0.039)	(0.343)	(0.415)	(0.345)
Teacher 9	0.012	0.518	0.841	0.523
	(0.033)	(0.254)**	(0.315)***	(0.253)**
Teacher 10	0.059	0.103	-0.243	0.185
	(0.037)	(0.224)	(0.297)	(0.221)
Number of student observations	2217	2156	2132	2156

Notes: Estimates are relative to the omitted group of teachers as described in the text. Standard errors are in parentheses.

***Significant at 1% level of confidence
**Significant at 5% level of confidence
*Significant at 10% level of confidence

Table B.4. Comparison of Teachers' Dropout Effects and GPA Effects at School 4

Teacher	Multivariate IV Probit - Dropout (Marginal Effects)	Total GPA Effect	Math GPA Effect	Non-Math GPA Effect
Teacher 1	-0.064	0.050	-0.134	0.120
	(0.026)**	(0.273)	(0.394)	(0.270)
Teacher 2	-0.052	0.686	1.519	0.622
	(0.024)**	(0.461)	(0.703)	(0.445)
Teacher 3	-0.042	-0.227	-0.636	-0.127
	(0.022)*	(0.382)	(0.562)	(0.371)
Teacher 4	-0.040	-0.015	-0.086	-0.057
	(0.020)**	(0.275)	(0.406)	(0.272)
Teacher 5	-0.038	0.293	0.664	0.358
	(0.031)	(0.596)	(0.908)	(0.572)
Teacher 6	-0.024	-0.06	0.099	-0.117
	(0.021)	(0.412)	(0.611)	(0.409)
Teacher 7	-0.015	-0.096	0.076	-0.179
	(0.026)	(0.348)	(0.524)	(0.343)
Teacher 8	-0.013	0.640	0.643	0.652
	(0.020)	(0.218)***	(0.327)**	(0.211)***
Teacher 9	-0.012	0.228	-0.519	0.312
	(0.020)	(0.472)	(0.710)	(0.461)
Teacher 10	-0.005	0.326	0.468	0.346
	(0.023)	(0.279)	(0.370)	(0.273)
Number of student observations	3779	3772	3770	3772

Notes: Estimates are relative to the omitted group of teachers as described in the text. Standard errors are in parentheses.

***Significant at 1% level of confidence
**Significant at 5% level of confidence
*Significant at 10% level of confidence