

## **Bilingual education and English proficiency**

**Christopher Jepsen**  
University of Kentucky  
Department of Economics  
Center for Business and Economic Research

**January 2009**

***Preferred citation***

Jepsen, C. (2009, January). Bilingual education and English proficiency. *University of Kentucky Center for Poverty Research Discussion Paper Series*, DP2009-01. Retrieved [Date] from <http://www.ukcpr.org/Publications/DP2009-01.pdf>.

Author correspondence: Christopher Jepsen, University of Kentucky, Department of Economics, 335BA Gatton Building, Lexington, KY 40506-0034. E-mail: [jepsen@uky.edu](mailto:jepsen@uky.edu).

University of Kentucky Center for Poverty Research, 302D Mathews Building, Lexington, KY, 40506-0047  
Phone: 859-257-7641; Fax: 859-257-6959; E-mail: [jspra2@uky.edu](mailto:jspra2@uky.edu)

# Bilingual Education and English Proficiency

Christopher Jepsen\*  
University of Kentucky  
335BA Gatton Business and Economics Building  
Lexington, KY 40506-0034  
jepsen@uky.edu

January 2009

## Abstract:

English Learners, students who are not proficient in English and speak a non-English language at home, make up more than 10 percent of the nation's K-12 student body. Achieving proficiency in English for these students is a major goal of both state and federal education policy, motivating the provision of bilingual education policies. Using data for nearly 500,000 English Learners from California, I show that students in bilingual education have substantially lower English proficiency than other English Learners in first and second grades. In contrast, there is little difference between bilingual education and other programs for students in grades three through five. These results hold across fixed effects, propensity score, and instrumental variables models.

---

\* I thank Laura Brien, David Figlio, Arturo Gonzalez, Scott Hankins, Brian Jacob, Lisa Jepsen, Robert McMillan, Abigail Payne, Heather Rose, Jon Sonstelie, Christopher Taber, Kenneth Troske, and seminar participants at the Bureau of Labor Statistics, the Public Policy Institute of California, the University of California – Santa Barbara, and the University of Kentucky for useful comments. Thanks also to Mark Fetler and Michelle Center for providing access to – and assistance in using – restricted-use data from the California Department of Education, as well as to Eric Brunner for sharing voting data and programs for California. Shelley de Alth and Ria Sengupta provided outstanding research assistance. All opinions and mistakes are mine alone.

## **Introduction**

Over 10 percent of the nation's K-12 students are not proficient in English and speak a language other than English at home (Padolsky, 2005). In California, these students – called English Learners – make up more than 25 percent of the K-12 student body. Although California has by far the most English Learner (EL) students of any state in the country, many other states are experiencing rapid growth in the population of EL students. For example, the EL population has more than doubled in Colorado, Nevada, and Oregon over the last ten years, resulting in a student population that is more than ten percent EL in each state. Therefore, an analysis of the experience of EL programs in California can provide guidance for other states.

The academic and labor-market success of EL students depends in large part on their ability to become proficient in English. Standardized tests are usually given in English, and EL students have consistently lower test scores than students whose native language is English. English language proficiency is critical for overall academic success. Furthermore, several authors have demonstrated the positive link between English proficiency and success in the labor market (Gonzalez, 2000; Trejo, 2003; Bleakley and Chin, 2004).

As most educators and researchers would agree, instruction targeting English proficiency is necessary for students with little English proficiency. For these students, the key research question concerns the most effective type of instructional services that they receive rather than whether or not they should receive services at all. For each student, the California data contain four categories for English language instruction (collectively referred to as EL services): (1) English Language Development (ELD), (2) ELD and Specially Designed Academic Instruction Delivered in English (SDAIE), (3) ELD and SDAIE with primary language support, and (4) ELD

and academic subjects through primary language.<sup>1</sup> The last category is defined as “bilingual education” due to the use of primary language in academic subjects.<sup>2</sup> For all these programs, academic instruction focuses on teaching the academic material, with the acknowledgment that the students have limited English proficiency. Bilingual programs address this issue by providing part or all of the instruction in the primary language. Other EL programs address this issue by providing additional explanation of terms, using pictures to add context, and in some cases providing small amounts of clarification in the primary language. In principle, the overall techniques are similar across programs. There are two important differences between bilingual programs and other EL programs. The first is the obvious difference in the language of instruction: bilingual programs teach academic subjects in the primary language, whereas other EL programs teach these subjects in English. The second difference is that bilingual programs attempt to improve primary language skills in addition to English language skills, whereas other EL programs focus only on improving English language skills. There is much debate over the role of native language instruction, yet much of this debate is based on belief rather than on empirical evidence (Greene, 1998).

Whereas previous work on bilingual education has focused on academic achievement (Lopez, 2003; Gordon and Hoxby, 2002; Greene, 1998), the goal of this paper is to look at the effect of bilingual education on English proficiency.<sup>3</sup> To that end, I use administrative data on almost 500,000 Spanish-speaking EL students in California to compare English Learners who receive bilingual education (academic instruction in the primary language) with English Learners who receive EL services that provide academic instruction in English. In California, students are

---

<sup>1</sup> There is a fifth category for “other” EL services, treated as non-bilingual education, which contains few students.

<sup>2</sup> The bilingual education category includes all types of bilingual programs such as early-exit programs as well as dual-immersion programs.

<sup>3</sup> Matsudaira (2005) looks at the effects of receiving any instructional services, regardless of the language of instruction, on academic achievement.

not randomly assigned to bilingual programs. Schools choose whether to offer bilingual education, and students need to obtain parental permission in order to participate in bilingual programs, both consequences of 1998's Proposition 227. I use multiple techniques to control for this non-random selection. In particular, I use propensity score analysis to match bilingual students with students receiving other EL services, based on the likelihood of participating in bilingual education. I also estimate instrumental variables (IV) models, using changes in bilingual availability around Proposition 227 as instruments for bilingual education.<sup>4</sup>

The results of this study show that bilingual education has a sizable negative effect on English proficiency for students in first and second grades, with magnitudes often in excess of 0.3 standard deviations. This pattern is consistent across three different methods: OLS, matching estimators, and IV. In later elementary grades (3-5), the effect of bilingual education is usually less than 0.1 standard deviations, and it is sometimes not significantly different from zero. This negative effect is not surprising. Non-bilingual programs are narrowly focused on English proficiency, whereas bilingual education programs encourage proficiency in the student's native language and in English. The pattern of results is also consistent with common approaches to dual-immersion bilingual education, where the percentage of instruction in English is as low as ten percent in kindergarten compared to nearly 50 percent by fifth grade.

### **Relation to Previous Work**

Many researchers have written about bilingual education, but most of these studies are not able to identify the causal effect of bilingual education as opposed to other EL programs on student outcomes (Greene, 1998). Some of the studies lack a comparison group (i.e. students not in bilingual programs), whereas others do not control for observable (or unobservable)

---

<sup>4</sup> This approach is similar in spirit to the IV model used in Gordon and Hoxby (2002) in their analysis of bilingual education and academic achievement.

differences between students in bilingual programs and other students. In this paper, the discussion is limited to rigorous analyses that attempt to control for the non-random selection of students into bilingual education.

Among the studies with comparison groups and controls for selection, several authors focus on the effects of primary language instruction, usually in academic subjects, on individual students. This restrictive definition is quite similar to the definition of bilingual education that I employ. Federman (2000), Lopez (2003), Lopez and Mora (1998), and Mora (2000) find insignificant or negative effects of bilingual programs on labor-market outcomes such as educational attainment and earnings (using a combination of methods, including propensity score and IV estimation).<sup>5</sup> In contrast, Cheng (1996) finds positive IV effects of bilingual education on both these outcomes, although the exogeneity of his instruments is questionable. A limitation of these studies is that their comparison groups include all students who speak a language other than English at home.<sup>6</sup> The comparison groups include students who are proficient in English as well as students who are not proficient in English.

California's mandated study of Proposition 227 finds insignificant effects of bilingual education on student-level mathematics and reading test scores (Parrish et al., 2006). Their preferred model is a hierarchical linear model for a large panel of student-level data from the Los Angeles school district. However, they acknowledge that this model may not fully control for the non-random selection of students into bilingual education programs.

Gordon and Hoxby (2002) look at the effect of bilingual education (again defined as academic instruction in the primary language) on grade-level academic achievement in California. In their preferred IV model, they find a positive effect of bilingual education on

---

<sup>5</sup> These studies use National Educational Longitudinal Survey and High School and Beyond data.

<sup>6</sup> The comparison group is chosen because the authors cannot identify English Learner status but they can identify language minority status.

achievement in multiple subjects, although these effects are concentrated in the earlier grades. They look separately at the effects of bilingual education on the performance of EL students as well as on the performance of English-proficient students.

Greene (1998) conducts a meta-analysis of 11 rigorous studies on bilingual education. He finds a positive effect of bilingual education on academic achievement (primarily reading). In other words, students in bilingual education programs outperform other EL students, students who receive EL services in English or who receive no EL services. August and Shanahan (2006) provide a more recent meta-analysis and reach similar conclusions. Genesee et al. (2006) summarize the education literature and find positive associations between native-language instruction and various measures of English proficiency and academic achievement.

Rather than looking at bilingual education, Matsudaira (2005) includes all English Language services in a single category. He uses a regression-discontinuity design and finds negligible differences in mathematics and reading achievement between students who receive English Language services and students who receive no such services.

In sum, these studies find mixed effects of bilingual education on academic achievement and labor-market outcomes. They use a host of techniques to control for the nonrandom placement of students into bilingual education. Their data sets are either aggregated to the grade level or contain English proficient students along with English Learners. However, none of them considers the effect of bilingual education on English proficiency (as distinct from academic achievement in English, reading, or writing).

## **Data**

In 2001, California implemented a statewide assessment, the California English Language Development Test (CELDT), to measure proficiency in English. The CELDT is

administered to *all* EL students in the fall of each school year. The CELDT consists of three portions: listening and speaking, reading, and writing. The listening and speaking portion is administered individually to students in all grades, where the tester asks each student a set of questions to measure both specific and general skills (Legislative Analyst's Office, 2004). The reading and writing portions of the test, administered to students in grades 2 through 12, are standardized tests with multiple-choice and short-answer sections. An overall scale score is calculated based on the scale scores for each individual portion.

The CELDT data contain extensive student level test score data on English proficiency, but they contain no measures of academic achievement. Scale scores are available for each portion of the English proficiency test (listening and speaking, reading, and writing), as well as for the overall score. These scores are available for the current school year as well as the previous school year. Comparisons of CELDT scores are only valid within each grade span (K-2, 3-5, 6-8, and 9-12). In other words, a scale score of 500 in grade one or two may or may not represent the same level of English proficiency as a scale score of 500 in grade four or five. Consequently, separate regressions are estimated for each grade and year.

As mentioned above, the CELDT data identify the type of EL services provided. In addition, CELDT demographic data contain information on primary language, gender, age, mobility, and program participation (for programs other than EL). The mobility information includes the number of years in the current school and the number of years in U.S. schools.<sup>7</sup> The non-EL programs include special education receipt and Title I receipt (i.e. supplemental services, typically targeting low-income students).

---

<sup>7</sup> Data prior to 2004 also contain information on the years in the current district. The data on years in U.S. schools are quite noisy, so I do not use them in the analysis.



Student-level CELDT data are supplemented with school-level data from the California Basic Education Data System (CBEDS). These data include school-level information on teachers (because researchers cannot link EL students to their individual teachers), school demographics, and academic achievement (for non-EL students) in mathematics.

The CELDT data are a repeated cross section rather than a panel. They contain the previous year's CELDT score (in addition to the current year's CELDT score), but all other data is from the current school year. The bilingual education variable is from the current year, which I treat as a proxy for bilingual education in the previous year. For a subset of students, I was able to merge CELDT data across years using the school identifier, birth date, grade, and gender. For these students, most had the same EL program for both years. Furthermore, the regression results were qualitatively similar when using the previous year's EL program rather than current year's EL program. I use data from the current year because the sample is much larger and more representative, and the results are not sensitive to this choice.

There are a few sample restrictions in the data. Students in their first year in a California school have not been exposed to any EL program and therefore are excluded from the sample. Students who skipped a grade or repeated a grade are also excluded. Because 99 percent of students in bilingual programs in California are native Spanish speakers, I only include native Spanish speakers in the analysis sample. One or two percent of EL students attend charter schools, and they are excluded from the analysis. Similarly, students in nontraditional schools such as alternative schools, special education schools, and county offices of education also are excluded. Finally, less than one percent of the students have invalid values of either test scores or grade spans, so they are excluded from the sample.

## **Descriptive Characteristics**

Table 1 presents the descriptive statistics for the analysis sample. The statistics are separated by year (2003 and 2004), as well as by grade (one through five). The average test score is usually between 450 and 550 for the current year, with a standard deviation between 50 and 65. The percentage of students in bilingual education programs is highest for students in first and second grades, with a percentage of 13 to 14 percent. The percentage is around nine percent for fourth-grade students and six percent for fifth-grade students. These small percentages are to be expected given the provision of Proposition 227 that students be taught “overwhelmingly” in English. Roughly 80 percent are eligible for Title I services, and between four percent (first grade) and 11 percent (fifth grade) receive special education services. The average student attends a school that has little language diversity, and the majority of students in the school are English Learners.

### **Participation in Bilingual Education**

This section takes a closer look at the relationship between bilingual education participation and student and school characteristics. In California, the participation decision for bilingual education is a two-part process. The first step for a student to participate in bilingual education is for the student to attend a school that offers bilingual education, as schools are not required to offer bilingual education. If the student attends a school that offers bilingual education, the student may participate in bilingual education if at least one parent visits the school and signs a permission slip. Thus, the placement of students into bilingual programs is a subjective process by the schools – whether to offer bilingual education – and by the parents – whether to enroll their children in bilingual education. I focus on the outcome of this process: whether or not the student participates in bilingual education.

Table 2 contains the marginal effects from a probit on bilingual participation, presented in percentage terms. The marginal effects are calculated at the mean for continuous variables; for dichotomous variables, the marginal effect is the change in probability of bilingual participation from changing the dichotomous variable from zero to one.<sup>8</sup> Recall that bilingual participation is largest in the younger grades – around 14 percent in first grade – and smallest in the oldest grades – around six percent in fifth grade. Participation in bilingual education is negatively associated with previous CELDT scores in reading/listening and writing, but the association is positive for reading scores. For first-grade students, a one-standard deviation increase in the previous listening and speaking score (about 90 points – see Table 1) is associated with a decreased likelihood of participation in bilingual education of about four percentage points– approximately one-third of the mean participation rate. The effect is similar for second-grade students. When measured as a percentage of the bilingual participation rate, the effect of previous test scores is sizable, ranging from 27 to 57 percent.

Other student characteristics have significant differences in bilingual education participation. Females generally have slightly higher participation. Students receiving special education services correspond to noticeably lower participation in bilingual education. On the other hand, students who have spent more time in their current school correspond with somewhat higher participation.

Selected school characteristics have significant relationships with bilingual education. The percentage of EL students in the school has a large, positive relationship with bilingual education participation. Similarly, the language homogeneity index – higher for schools with little diversity in the home language of EL students – also has a large and positive relationship. Perhaps the positive coefficients on these two variables reflect economics of scale – more

---

<sup>8</sup> Although not reported, results from a linear probability model are quite similar.

bilingual education in schools with higher concentrations of students speaking the same non-English language. For grades one through three, schools with higher California Standards Test (CST) scores – a test of academic achievement in English and Language Arts – correspond with modestly lower participation.

### **OLS Model and Results**

The pattern of results from the participation probit suggests that students in bilingual programs differ from students in other instructional programs along many dimensions. In order to control for these differences, the starting point for the analysis is an ordinary least squares (OLS) regression, as shown in equation (1):

$$(1) \quad ENGPROF_i = BILINGUAL_i \alpha + X_i \beta + \varepsilon_i.$$

In this equation,  $ENGPROF_i$  is student  $i$ 's scale score from the current school year of the CELDT (2003 or 2004). Recall that I estimate separate models for each year and for each grade.

Because students in kindergarten and first grade only take the listening/speaking portion of the test, the only outcome variable for the grades one and two is the listening/speaking scale score.

For grades three through five, separate models are estimated for each of the three portions: listening/speaking, reading, and writing.

$BILINGUAL_i$  is a dummy variable for whether or not the student receives academic instruction in the student's native language (Spanish, as the sample is limited to Spanish-speakers). In other words, the comparison group contains EL students who do not receive academic instruction in their native language.

$X_i$  can be partitioned into three components (and a constant term),  $LAGPROF_i$  (lagged English proficiency),  $STUDENT_i$ , and  $SCHOOL_i$ .  $LAGPROF_i$  contains the scale scores from the previous year's CELDT. Again, for grades one and two,  $LAGPROF_i$  has only one test score, the

previous year's listening/speaking portion. For grades three through five,  $LAGPROF_i$  contains three test scores, one for each portion of the test. This structure allows for a more comprehensive effect of lagged English proficiency.

$STUDENT_i$  is a vector of student-level characteristics such as sex, special education receipt, and participation in the Title I program (a proxy for family income), as listed in Tables 1 and 2.  $SCHOOL_i$  is a vector of school-level characteristics from the previous school year.<sup>9</sup>

The OLS results are in Table 3. The top panel contains the results for the listening and speaking test. The middle panel contains the results for the reading test, and the bottom panel contains the results for the writing test. Each table contains a separate column for each grade and year. For each test and grade, the first column is from 2003 and the second column is from 2004. Appendix Tables 1a through 1c contain the results for the covariates other than bilingual education.

Bilingual education has a negative coefficient for the listening and speaking test, although the coefficient is much larger for grades one and two. Specifically, the effect is above 0.4 standard deviations for grade one and is above 0.2 standard deviations for grade two. In contrast, the coefficient is 0.05 standard deviations in grades three and four, and the coefficient for grade five is not statistically different from zero at the five-percent level. In other words, the negative association between bilingual education and listening/speaking proficiency is concentrated in first and second grades; by third grade, the effect is extremely small.

The coefficients for bilingual education in the reading proficiency regressions vary by grade. In third grade, bilingual education is associated with a small, positive effect of 0.06 to

---

<sup>9</sup> Because the CELDT is administered toward the beginning of the school year, most of the value added since the last administration of the CELDT occurred in the previous school year. For example, the 2004 CELDT contains school characteristics from the 2003-2004 school year. There is an implicit assumption that students attended their current school in the previous school year. The results are similar if I use current school year characteristics instead.

0.08 standard deviations. In fourth and five grades, the results vary by year. This pattern of results suggests that bilingual education does not have a strong association – either positive or negative – with reading test scores.

The bottom panel of Table illustrates that bilingual education has a modest, negative association with writing proficiency. The coefficient is negative and significant in all years and grades. In third grade, bilingual education is associated with a 0.10 or 0.11 decrease in writing proficiency. For fourth grade the effect size is a 0.06 to 0.08 standard-deviation decrease, and in fifth grade the effect size is a 0.06 standard-deviation decrease.

Bilingual education is not the only determinant of English proficiency, as illustrated by the results in Appendix Tables 1a through 1c. As expected, previous test scores have strong positive effects: Students do not lose English proficiency from one year to the next. Females and students who have spent more time in their current school have higher test scores, whereas disadvantaged students – Title I students and students receiving special education services – have lower test scores.

At the school level, average test scores (of non-EL students) in English/Language Arts also are positively associated with English proficiency. The school's percentage of EL students is associated with lower listening and speaking proficiency. The percent of teachers who are fully credentialed has a positive association with reading proficiency and writing proficiency. Other school characteristics, including the access to teachers specifically authorized to teach EL students, do not consistent effects on English proficiency.

Table 4 contains the results from a school fixed-effects estimator. This model controls for between-school differences in fixed (observable and unobservable) factors. Rossell (2005) illustrates how California schools had great variation in defining bilingual education and other

EL instructional programs. In other words, two schools may offer the same program, but one school calls it bilingual education and one school calls it something else. Rather than identifying the bilingual education effects off of such situations, the school fixed-effects model is identified only off of within-school differences in bilingual education. The results are similar to the OLS results in Table 3, although the magnitudes of the effect may vary slightly. These findings suggest that between-school variation in bilingual education programs is not driving the OLS results.

### **Propensity Score Models and Results**

In this section, propensity score analysis provides a robustness check on the OLS results in Table 3. For bilingual education, the propensity score equals the likelihood of participating in bilingual education, calculated as the predicted probability from the probit equation of bilingual education receipt reported in Table 2. The propensity scores obtained from a logit model are nearly identical. The independent variables are defined as in equation (1).<sup>10</sup>

Figure 1 illustrates the predicted propensity scores for Fall 2004, separately by grade. As expected, propensity scores are higher for students who participate in bilingual education than for those who do not. The area of common support is 100 percent for both cohorts (and for 2003, although not shown in the figure). Both of these factors are consistent with a valid propensity score estimator, but the validity of the estimator ultimately hinges on the assumption that observables capture all relevant determinants of bilingual education participation.

Propensity scores dominate the matching literature. However, there are many types of propensity score estimators, and no single estimator dominates the literature (see, for example,

---

<sup>10</sup> Researchers often use interaction terms or higher-ordered terms (for continuous independent variables) in order to improve the goodness-of-fit of the probit (or logit) generating the propensity score. However, there is no theoretical basis for including these extra terms. Their goal is to eliminate any statistically significant differences in variables between the treatment and control groups. However, given the size of my sample, it is extremely difficult to eliminate *all* statistically significant differences between students in bilingual education and other EL students.

Frolich, 2004). Because the focus of the paper is on the effectiveness of bilingual education, not propensity score estimators, I employ three common methods: kernel density, nearest neighbor, and caliper.<sup>11</sup> Essentially, the kernel density estimator creates a weighted average of the comparison group (in this case, EL students not in bilingual education) for each student in the treatment group (in this case, EL students in bilingual education).<sup>12</sup>

In the nearest-neighbor approach, each student in bilingual education is matched with the four students in other EL programs who have the most similar propensity scores. I choose four because simulation results in Abadie and Imbens (2006) worked best – in terms of mean-squared error – for four neighbors, although my results for four are quite similar to results with other numbers of matches (from 1 to 5).<sup>13</sup>

The third method is a caliper or radius method. I match each student in bilingual programs with all students in other EL programs who have a propensity score with a given amount, called the radius. I choose a radius of 0.000004 (which produces a similar average number of matches to the four nearest neighbors), although results with radii of 0.000001 to 0.000005 produce nearly identical results.

Table 5 contains the results from all three types of propensity score estimators. The reported results are simple differences in means between the two groups, the full set of students in bilingual education and the matched set of EL students in other programs. These differences are reported in terms of standard deviations of the dependent variable (current test score).

Regression-adjusted estimates are similar to the simple averages, and therefore are not reported.

---

<sup>11</sup> I use a matching program in Stata called `psmatch2` (see Leuven and Sianesi, 2003) to calculate the propensity score estimators. For a recent review of matching estimators, with an application to job training programs, see Mueser, Troske, and Gorislavsky (2007) and the references cited there.

<sup>12</sup> I use the default options in `psmatch2`: an Epanechnikov kernel (as recommended by Silverman, 1986) with a bandwidth of 0.08.

<sup>13</sup> I match with replacement, and I include ties (i.e. students with identical propensity scores), even if including them raises the number of closest neighbors above four. The results are not sensitive to either of these assumptions.



The results do not account for the variation resulting from the matching process. Given the similarity of the results to the OLS results for the full sample, it is extremely doubtful that controlling for that variation will substantially alter the findings. Furthermore, Abadie and Imbens (2006) show that bootstrapping the standard errors is not necessarily appropriate.

The estimators in Table 5 are quite similar to the OLS estimators (measured in standard deviation effects) for the full sample (Table 3). With one exception, the full sample OLS standard-deviation effect lies between the three matching estimators. For example, the full sample OLS coefficient is -0.48 standard deviations for first-grade students (Table 3a, column 1). This coefficient falls between the kernel density estimate of -0.51 and the radius/caliper estimate of -0.43 (both from Table 5, column 1). There are several potential explanations for the similarity between matching estimators and OLS coefficients. The area of common support is 100 percent for the kernel and nearest neighbor matching, but falls to roughly 60 to 80 percent for radius matching. This high percentage suggests that propensity score samples may be similar, but not identical, to the overall sample (controlling for observables). A related explanation is that selection on observables is largely linear. In this case, OLS models control for selection and therefore the propensity score – a nonlinear function of observables – provides little additional information.

A third explanation is that the matching procedure based on the propensity score is not appropriate. Propensity scores are convenient because they reduce the matching problem to one variable, the propensity score. Matching across multiple variables – i.e. directly on the observables (or a subset of them) rather than on a nonlinear function of them – is difficult to implement successfully (Smith and Todd, 2005). For example, Mueser, Troske, and Gorislavsky

(2007) and Frolich (2004) find that Mahalanobis distance methods perform worse than propensity score methods.<sup>14</sup> Therefore, I do not pursue multi-dimensional matching estimators.

### **Instrumental Variables Models and Results**

One concern with the OLS and matching estimators is that the bilingual education coefficient may be picking up the effect of other variables that cannot be measured. For example, suppose that bilingual education students receive much more parental participation (such as parent-teacher conferences, help with homework, etc.) than students in other EL programs. Because the data do not contain measures of parental involvement, OLS and matching estimators may be confounding the effect of bilingual education and parental involvement.

The results in this section use an instrumental variables (IV) approach to address this concern. I consider two sets of instruments. The first set of instruments are measures of changes in bilingual education participation as a result of the 1998 state proposition severely restricting access to bilingual education. The second and alternate set of instruments is the percentage of voters who voted for the proposition limiting bilingual education. Passed in the summer of 1998, Proposition 227 required that EL students be taught “overwhelmingly” in English, and it created barriers to providing bilingual education (most notably the requirement that each student in bilingual education obtain parental permission). As expected, the percentage of EL students in bilingual education dropped sharply after 1998.<sup>15</sup>

---

<sup>14</sup> Similarly, Frolich (2004) finds that using the odds ratio of the propensity score, as suggested by Imbens (2000), performed more poorly than directly using propensity score methods.

<sup>15</sup> In order to make sure that the instrument is not identified from schools that dropped bilingual programs after Proposition 227, I estimate additional models that exclude schools where no EL students are in bilingual education programs (as well as schools where all EL students are in bilingual programs). The results from these models are nearly identical to the results from the models reported in this section.

Gordon and Hoxby (2002) also use an instrument based on Proposition 227 in their analysis of bilingual education and academic achievement. Their instrument is the change in the percent of students eligible for bilingual education services if schools strictly adopted Proposition 227. In other words, they set the percentage of students eligible for bilingual education to zero for the post-Proposition 227 period. For the pre-Proposition 227 period, the percentage is the percentage of EL students (calculated separately for each grade and school). They also assume that bilingual education is not offered for grades with fewer than 10 EL students because expired federal and California laws set the threshold for offering bilingual education at 10 students.<sup>16</sup>

Part of the appeal of this instrument is that it relies solely on the exogenous “shock” to bilingual education resulting from Proposition 227. However, this instrument is highly correlated with the percentage of students classified as EL students, one of the school level variables in equation (1). In fact, using Gordon and Hoxby’s (2002) instrument always produces insignificant first-stage coefficients; therefore, it is not used.

Instead, I use two related, school-level instruments for the first set of instruments. The first is the mechanical change in the percentage of EL students participating in bilingual education as a result of Proposition 227. As in Gordon and Hoxby (2002), the mechanical percentage is set to zero for the post-Proposition period. The pre-Proposition measure is the school-level percentage of students enrolled in bilingual education in spring 1998 (i.e. right before Proposition 227).<sup>17</sup> Because the post-Proposition percentage is set to zero, the

---

<sup>16</sup> This threshold has no effect on bilingual education participation in the 2003 and 2004 CELDT data. Other thresholds, such as 5 or 15, also have no effect. Therefore, thresholds are not used in the definition of the instruments in this paper.

<sup>17</sup> The school is chosen as the level of analysis in order to reduce the influence of measurement error based on within-school changes in either bilingual enrollment or total enrollment. When the instrument is defined at the grade level (as in Gordon and Hoxby, 2002), the instrument is insignificant in the first-stage regression (predicting bilingual participation).

mechanical change (i.e. post-Proposition minus pre-Proposition) is equal to the opposite of the pre-Proposition percentage. This percentage has a strong, negative effect on the first-stage regressions, i.e. linear probability models on the likelihood of participating in a bilingual program. First-stage results are contained in Appendix Table 2.

In order for this instrument to be valid, it must not be directly correlated with student-level English proficiency. Because the post-Proposition 227 percentage is set to zero, the instrument depends entirely on the percentage of students enrolled in bilingual programs in spring 1998. The outcome is a student-level measure of English proficiency. Very few if any of the English Learners in grades K-5 in 2003 or 2004 were in school before Proposition 227, therefore minimizing the possibility that these students were directly affected by Proposition 227. Still, it is possible that this percentage is related to unobserved school characteristics that affect English proficiency.

I also use a second instrument that uses changes over time in the school-level percent bilingual. Rather than measuring the mechanical change in percent bilingual, this additional instrument measures the actual change in percent bilingual between spring 1998 and spring 1999 (at the school level). Any school-specific attributes that are constant between 1998 and 1999 would not be captured by this instrument. However, changes in school attributes as a result of Proposition 227 would be captured by the instrument. This instrument is positive and highly significant in the first-stage regression (participation in bilingual education).

The IV results for these two instruments are in Table 6. The table reports the full model, where each cell is from a separate IV model. The layout is the same as Table 4 (the results for the school fixed effect model). The first two columns are for grade one, the second two for grade two, and so on. Within each grade, the first column is for 2003 and the second is for 2004. The

dependent variable is the English listening/speaking proficiency in the top panel, the English reading proficiency in the middle panel, and the English writing proficiency in the bottom panel.

Bilingual education has a large, negative association with listening and speaking proficiency in first and second grade. The IV coefficients show no evidence of a negative effect for listening and speaking proficiency in grades three through five, as the coefficients are always positive and in some cases statistically significant. For reading proficiency, bilingual education has a positive association with bilingual education for students in grades three and four, and the coefficients for grade five are not statistically different from zero. For writing proficiency, bilingual education has a modest, negative association for grades three and four, and the coefficient for grade five varies across years.

As mentioned previously, I also consider a second IV strategy that uses the voting behavior on Proposition 227. Specifically, I match each Census block to the nearest elementary school, and then calculate the percentage of “yes” votes on Proposition 227.<sup>18</sup> Presumably, schools with more support for Proposition 227 are less likely to offer bilingual education than schools with less support for the Proposition. Indeed, the first stage results in Appendix Table 3 are consistent with this hypothesis. For this instrument to be valid, however, I assume that voting behavior in 1998 is not directly correlated with student-level English proficiency in 2003 and 2004. However, I have no way to test this assumption directly.

The results from this IV analysis are in Table 7. In other words, the instrument in Table 7 is the percentage of voters who voted “yes” for Proposition 227. The other two instruments mentioned previously are excluded in order to provide an alternate identification strategy based on voting data rather than on education data. The IV results from the voting data (Table 7) are similar in spirit to earlier tables, but they are less precisely estimated. The large, negative

---

<sup>18</sup> I thank Eric Brunner for sharing data and programs allowing me to link voting data with school data.

coefficients in grades 1 and 2 are still evident, but the results for later grades are very noisy. For example, the bilingual coefficients on the fifth-grade listening and speaking test are above 50, but the standard errors are in excess of 20. The imprecise estimates in Table 7 are not surprising because the instrument is not perfect: school attendance boundaries and Census block boundaries overlap. Still, the pattern of results is consistent with the other IV estimates.

In sum, the results are generally consistent across all three methods (OLS, propensity score, and IV). Compared to other EL programs, bilingual education has a sizable, negative association with English listening/speaking proficiency for EL students in first and second grade. For EL students in grades three through five, bilingual education has a small and often insignificant association with English proficiency.

### **Bilingual Education by Previous Proficiency Score**

As seen above, EL students in bilingual education have dramatically lower English listening and speaking proficiency than EL students in other EL programs in first and second grades. Bilingual education has small and often insignificant associations with English proficiency for students in grades three through five. These findings are robust across years and across estimation methods. This section contains an exploration of whether bilingual education has differential effects based on students' prior English proficiency levels.

Specifically, I estimate an OLS model similar to equation (1), except for the addition of variables interacting the bilingual education variable with dummy variables for each quartile of the previous year's listening and speaking score. The results for the listening and speaking test are in Table 8, and results for reading and writing are in Appendix Table 4.<sup>19</sup>

---

<sup>19</sup> The results are similar when the quartile interactions terms are replaced with a single interaction term between bilingual education and the previous year's listening and speaking test score.

The results for all three English proficiency test show the same pattern: the bilingual education coefficient improves dramatically with the quartile of lagged listening and speaking proficiency. In other words, the students who appear to benefit the most from bilingual education relative to other EL programs are the students who had strong English listening/speaking proficiency in the previous year. For students in the lowest quartile, bilingual education is associated with lower listening and speaking test scores of more than 40 points (approximately 0.6 standard deviations) among first-grade students. For students in the third-lowest quartile, the bilingual coefficient is only -11 points in 2003 and -7 points in 2004.). Other grade levels also have a positive relationship between lagged listening and speaking proficiency and the bilingual coefficient, although the range of effects is much smaller.

#### *Attrition*

One concern with the results in all the tables is the relationship between student attrition and bilingual education. By using lagged CELDT scores, the sample is restricted to students who took the CELDT in the current year and in the previous year. If attrition from one year to the next is systematically related to bilingual education status, then the estimates in the table are not valid. There are two sources of attrition in the CELDT data. First, attrition occurs if a student leaves the California public school system (CELDT test score data are available for students who switch public schools within California). Second, attrition occurs if a student is no longer considered an English Learner. Once a student is considered proficient in English, that student no longer is considered an English Learner and stops taking the CELDT.

Because the data are a repeated cross-section, I cannot directly measure attrition. Instead, I calculate attrition indirectly by attempting to merge the fall 2003 cross-section with the fall 2004 cross-section. Student identifiers are not available, so I match students based on school,

grade, gender, and birth date. Therefore, I limit the attrition analysis to students in the fall 2003 cross-section who can be uniquely identified based on the matching variables (school, grade, gender, and birth date). I focus on kindergarten through fourth grade because these students would be in grades one through five in the fall 2004 cohort. Table 9 contains the percentage of students who were successfully matched in the fall 2004 cohort. The top panel combines all students, whereas the bottom two panels provide separate information by bilingual status.

The table shows that substantial attrition occurs: 65 percent of students in the fall 2003 cohort are in the fall 2004 cohort. Students in bilingual programs have lower attrition than students in other EL programs, although the difference declines as the grade increases. The results are similar for each quartile of the 2003 listening and speaking test.<sup>20</sup> This pattern of results is not consistent with the hypothesis that the large, negative bilingual education coefficients in grades K-2 and small, often insignificant, bilingual coefficients in grades 3-5 occurs because the high-achieving, non-bilingual students in grades 3-5 were more likely to become proficient in English. On the other hand, the lower attrition for students in bilingual programs is not consistent with the hypothesis that the bilingual education coefficients are biased down because bilingual education students are economically disadvantaged and therefore more likely to leave California schools. Thus, the table suggests that the results in this paper are not being driven by systematic attrition of EL students.

## **Discussion**

This paper studies the causal effect of bilingual education versus other EL programs on English proficiency. The comparison group is the set of students receiving other EL programs because most educators and policymakers agree that these EL students need some sort of EL

---

<sup>20</sup> Note that the quartile is calculated for all students in the appropriate grade level. Separate quartiles are not calculated for students in bilingual programs and for students in other EL programs.



program. Not offering them any EL program is no more plausible an option than not offering them a mathematics program. The study uses a unique administrative data set of California EL students in elementary grades. Students are not randomly assigned to bilingual education: In California, Proposition 227 (1998) requires parental permission slips before students are allowed in bilingual education classrooms. I use two approaches to control for this selection: propensity score analysis, which controls for selection on observables, and instrumental variables, where the instruments are based on the abrupt change in bilingual education availability caused by Proposition 227.

The results from both selection techniques are similar to the OLS results. In grades one and two, students in bilingual education program have lower listening and speaking proficiency scores than other EL students. The size of this effect is large – over 0.3 standard deviations in many cases. Furthermore, the effect of bilingual education is inversely related to students' previous listening and speaking proficiency. Students with low English listening and speaking proficiency in the previous year have particularly large, negative effects of bilingual education, whereas students with higher levels of English proficiency suffer little if any negative consequences of bilingual education. Policymakers should consider this adverse effect on English listening and speaking when evaluating bilingual education programs, as they represent a cost to the youngest and weakest students.

For students in grades three through five, the effects of bilingual education versus other EL programs are much smaller. In several cases, the effects on English proficiency are insignificant or even positive. In terms of standard deviations, nearly all of the effects are under 0.1 standard deviations, and many are under 0.05 standard deviations. Although these effects are

sometimes statistically significant, their small size suggests fewer concerns about the negative consequences of bilingual education on English proficiency for older students.

Why would bilingual education have a large, negative effect of English listening and speaking proficiency for students in early elementary grades but smaller (and sometimes insignificant) effects in later grades? Bilingual programs in kindergarten and first grade use native language much more extensively than English; some sources suggest only 10 percent of the instruction provided by dual-immersion programs is in English (Ovando and Collier, 1998). By fourth or fifth grade, the language split is roughly equal between English and Spanish. If this 90/10 model (as it is known) is common in California, then lower English proficiency gains might be expected for bilingual education students in first and second grades, as compared to other EL programs where a higher proportion of overall instruction takes place in English.

Because student-level academic achievement data are not available, this paper looks at only one outcome: English proficiency as measured by the CELDT data. Bilingual education has many goals, including overall academic proficiency. Still, English proficiency in the younger grades is essential for overall academic achievement. Although bilingual education is not as successful as other EL programs for early elementary school students, it may have other benefits that are not captured in the administrative data used here. Previous work finds mixed effects for academic achievement (for example, Lopez, 2003; Gordon and Hoxby, 2002). Further research on all relevant outcomes (not just English proficiency and academic achievement) is needed.

## References

- Abadie, Alberto, and Guido W. Imbens (2002), "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica*, 74(1): 235-268.
- August, Diane, and Timothy Shanahan, editors (2006). *Developing Literacy in Second Language Learners: Report of the National Literacy Panel on Language Minority Children and Youth*. New York: Lawrence Erlbaum Associates.
- Bleakley, Hoyt, and Aimee Chin (2004), "Language Skills and Earnings: Evidence from Childhood Immigrants," *The Review of Economics and Statistics*, 86(2): 481-496.
- Cheng, Chang-Kuo (1996), *Impact of Bilingual Education on Educational Attainment*. Ph.D. Dissertation, University of Maryland.
- Federman, Maya (2000), *Bilingual Education, College Major Choice, and College Remediation*. Ph.D Dissertation, Harvard University.
- Frolich, Markus (2004), "Finite-Sample Properties of Propensity-Score Matching and Weighting Estimators," *The Review of Economics and Statistics*, 86(1): 77-90.
- Genesee, Fred, Kathryn Lindholm-Leary, William M. Saunders, and Donna Christian (2006). *Education English Language Learners: A Synthesis of Research Evidence*. New York: Cambridge University Press.
- Gonzalez, Arturo (2000), "The Acquisition and Labor Market Value of Four English Skills: New Evidence from the NALS," *Contemporary Economic Policy*, 18(3): 259-269.
- Gordon, Nora, and Caroline Hoxby (2002), "Achievement Effects of Bilingual Education versus English Immersion: Evidence from California's Proposition 227," unpublished manuscript, University of California-San Diego.
- Greene, Jay (1998), "A Meta-Analysis of the Effectiveness of Bilingual Education," The Tomas Rivera Policy Institute.
- Guzman, Joseph M. (2002), "Learning English," *Education Next*, 2(3): 58-65.
- Imbens, Guido (2000), "The Role of the Propensity Score in Estimating Dose-Response Functions," *Biometrica*, 87: 706-710.
- Legislative Analyst's Office (2004), *A Look at the Progress of English Learner Students*, Sacramento, California.
- Leuven, Edwin, and Barbara Sianesi (2003), "Psmatch2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing," <http://ideas.repec.org/c/boc/bocode/s432001.html>.

- Lopez, Mark Hugo (2003), "Do English Language Assistance Programs Affect Educational Attainment and Labor Market Outcomes? Evidence from the National Education Longitudinal Study of 1988 and High School and Beyond," unpublished manuscript, Maryland University.
- Lopez, Mark Hugo, and Marie T. Mora (1998), "The Labor-Market Effects of Bilingual Education among Hispanic Workers," *READ Perspectives* 5: 33-51.
- Matsudaira, Jordan (2005), "Sinking or Swimming? Evaluating the Impact of English Immersion versus Bilingual Education on Student Achievement," unpublished manuscript, University of California.
- Mora, Marie T. (2000), "English-Language Assistance Programs, English Skill Acquisition, and the Academic Progress of High School Language Minority Students," *Policy Studies Journal* 28(4): 721-738.
- Mueser, Peter R., Kenneth R. Troske, and Alexey Goriaslavsky (2007), "Using State Administrative Data to Measure Program Performance," *Review of Economics and Statistics* 89(4): 761-783.
- Ovando, Carlos J., and Virginia P. Collier (1998). *Bilingual and ESL Classrooms: Teaching in Multicultural Contexts*. San Francisco: McGraw Hill.
- Padolsky, Daniel (2005), "National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs FAQ No. 1," accessed on line October 14, 2005 (<http://www.ncele.gwu.edu/expert/faq/01leps.htm>).
- Parrish, Thomas B., et al. (2006), "Effects of the Implementation of Proposition 227 on the Education of English Learners, K-12: Findings from a Five-Year Evaluation," Prepared for the California Department of Education.
- Rossell, Christine H. (2005), "The Educational Effectiveness of Bilingual Education," *Education Next* 3(4): 44-52.
- Silverman, Bernard W. (1986). *Density Estimation for Statistics and Data Analysis*. New York. Chapman and Hall.
- Smith, Jeffrey A., and Petra E. Todd (2005), "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125(1-2): 305-353.
- Trejo, Stephen J. (2003), "Intergenerational Progress of Mexican-Origin Workers in the U.S. Labor Market," *Journal of Human Resources* 38(3): 467-489.

Table 1 - Descriptive Statistics by Grade and Year

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
<i>Current CELDT Scale Scores</i>										
Listening/Speaking	500.7 (59.0)	497.3 (58.4)	537.0 (58.8)	537.2 (56.9)	505.1 (57.3)	504.5 (54.9)	530.4 (61.3)	533.1 (59.6)	546.1 (64.0)	551.2 (62.7)
Reading					469.7 (49.2)	471.6 (49.5)	497.1 (49.4)	497.3 (49.7)	514.8 (50.4)	516.5 (50.7)
Writing					491.9 (53.2)	497.2 (54.2)	514.3 (49.8)	517.4 (51.1)	526.4 (48.9)	530.3 (50.2)
<i>Lagged CELDT Scale Scores</i>										
Listening/Speaking	413.7 (88.4)	409.2 (93.6)	482.8 (65.7)	493.7 (70.2)	511.9 (68.2)	529.5 (71.5)	490.5 (69.6)	494.7 (68.5)	504.4 (75.2)	515.8 (74.8)
Reading					443.2 (47.3)	445.7 (46.6)	461.8 (48.3)	463.3 (50.1)	485.3 (50.2)	486.5 (51.5)
Writing					460.1 (59.6)	465.0 (59.4)	474.2 (59.0)	483.4 (59.9)	494.6 (58.6)	502.2 (59.8)
<i>Bilingual Education</i>	0.135 (0.341)	0.141 (0.348)	0.134 (0.341)	0.138 (0.345)	0.122 (0.327)	0.122 (0.328)	0.091 (0.287)	0.092 (0.289)	0.060 (0.238)	0.061 (0.239)
<i>Student Demographics</i>										
Female	0.490 (0.500)	0.488 (0.500)	0.492 (0.500)	0.492 (0.500)	0.492 (0.500)	0.492 (0.500)	0.490 (0.500)	0.488 (0.500)	0.484 (0.500)	0.481 (0.500)
Title I	0.817 (0.386)	0.776 (0.417)	0.830 (0.376)	0.792 (0.406)	0.843 (0.364)	0.804 (0.397)	0.850 (0.357)	0.824 (0.381)	0.861 (0.346)	0.836 (0.371)
Special education	0.041 (0.199)	0.047 (0.211)	0.056 (0.230)	0.065 (0.246)	0.074 (0.262)	0.079 (0.270)	0.088 (0.283)	0.098 (0.298)	0.103 (0.304)	0.113 (0.316)
Years in current school	1.872 (0.361)	1.883 (0.342)	2.629 (0.701)	2.641 (0.657)	3.290 (1.070)	3.288 (1.045)	3.803 (1.464)	3.808 (1.460)	4.234 (1.862)	4.227 (1.841)

Table 1 (Continued) - Descriptive Statistics by Grade and Year

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
<i>School Demographics</i>										
Language homogeneity index	0.861 (0.179)	0.860 (0.179)	0.863 (0.177)	0.866 (0.176)	0.865 (0.176)	0.865 (0.176)	0.865 (0.177)	0.866 (0.176)	0.862 (0.179)	0.868 (0.176)
Class size	19.15 (1.33)	19.53 (1.87)	19.14 (1.34)	19.48 (1.92)	19.14 (1.45)	19.52 (1.93)	18.79 (2.93)	19.17 (3.21)	18.33 (4.05)	18.72 (4.29)
Percent EL	57.44 (22.60)	55.13 (21.86)	57.45 (22.25)	56.21 (21.85)	57.98 (22.20)	55.89 (21.76)	57.52 (22.22)	55.97 (21.72)	56.97 (22.15)	55.82 (21.56)
Average teacher experience	10.78 (2.81)	11.60 (2.98)	10.79 (2.78)	11.58 (2.93)	10.79 (2.78)	11.61 (2.93)	10.78 (2.78)	11.55 (2.91)	10.80 (2.81)	11.55 (2.94)
Percent BA or less	27.31 (19.13)	23.54 (18.52)	27.29 (19.11)	24.22 (18.36)	27.68 (19.04)	24.44 (18.48)	27.96 (19.07)	25.03 (18.44)	27.51 (19.00)	25.31 (18.57)
Percent BA + 30	45.89 (20.73)	48.78 (21.19)	45.99 (20.97)	47.76 (21.02)	45.69 (20.52)	47.33 (21.02)	45.35 (20.37)	46.99 (20.59)	45.63 (20.49)	46.72 (20.50)
Percent fully credentialled	87.22 (12.40)	91.55 (10.12)	87.35 (12.06)	91.34 (9.90)	87.13 (12.14)	91.27 (9.95)	86.96 (12.15)	91.00 (10.03)	87.10 (12.01)	90.77 (10.30)
EL teacher access	0.09 (0.10)	0.09 (0.10)	0.09 (0.11)	0.09 (0.10)	0.08 (0.09)	0.09 (0.12)	0.08 (0.09)	0.09 (0.10)	0.08 (0.08)	0.09 (0.09)
Percent new students	16.80 (8.29)	17.82 (9.38)	17.05 (8.39)	18.00 (9.43)	17.00 (8.41)	18.31 (10.26)	16.85 (8.08)	18.27 (10.34)	17.24 (8.69)	18.30 (10.39)
Average CST score	327.3 (35.6)	320.4 (62.2)	328.5 (26.3)	322.6 (56.6)	328.9 (23.6)	323.5 (54.1)	329.3 (23.1)	324.7 (51.3)	329.2 (22.9)	325.3 (49.5)
Observations	114,259	110,349	121,637	119,368	115,188	120,396	112,624	109,983	100,010	101,376

Note: Standard deviations are in parentheses.

Table 2 - Probit Marginal Effects for Bilingual Education Participation by Grade and Year, As Percentages

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
Lagged listen/ speak	-0.040 (0.003)	-0.046 (0.003)	-0.073 (0.004)	-0.084 (0.004)	-0.012 (0.002)	-0.023 (0.003)	-0.020 (0.002)	-0.009 (0.003)	-0.010 (0.002)	-0.006 (0.002)
Lagged reading					-0.007 (0.004)	-0.018 (0.004)	0.029 (0.004)	0.016 (0.003)	0.016 (0.004)	0.004 (0.003)
Lagged writing					-0.080 (0.005)	-0.076 (0.005)	-0.062 (0.005)	-0.066 (0.005)	-0.033 (0.004)	-0.033 (0.004)
Female	0.07 (0.19)	0.25 (0.19)	-0.07 (0.18)	0.45 (0.19)	0.72 (0.17)	0.89 (0.17)	0.41 (0.15)	0.56 (0.15)	0.15 (0.13)	0.45 (0.13)
Title I	0.63 (1.26)	-0.29 (1.31)	0.57 (1.19)	0.05 (1.21)	-0.46 (1.27)	-0.31 (1.14)	-0.05 (1.01)	2.13 (0.77)	0.23 (0.82)	1.88 (0.60)
Special education	-2.35 (0.55)	-1.93 (0.57)	-3.09 (0.48)	-3.41 (0.46)	-5.25 (0.39)	-4.83 (0.36)	-3.41 (0.32)	-4.03 (0.34)	-2.31 (0.29)	-2.30 (0.27)
Years in current school	1.66 (1.00)	3.77 (0.76)	2.54 (0.51)	4.46 (0.47)	2.30 (0.31)	3.28 (0.25)	1.50 (0.20)	2.00 (0.18)	0.48 (0.11)	0.69 (0.11)
Language index	15.76 (3.88)	12.59 (3.83)	14.78 (3.51)	11.68 (3.62)	14.48 (3.32)	12.39 (3.14)	13.51 (2.78)	8.76 (2.78)	7.65 (2.13)	4.35 (2.37)
Class size	-1.66 (0.45)	-1.13 (0.35)	-1.79 (0.40)	-0.91 (0.35)	-1.59 (0.36)	-1.13 (0.28)	-0.32 (0.16)	-0.47 (0.12)	-0.37 (0.09)	-0.40 (0.08)
Percent EL	0.23 (0.03)	0.28 (0.04)	0.22 (0.03)	0.23 (0.03)	0.20 (0.03)	0.18 (0.03)	0.13 (0.02)	0.10 (0.02)	0.06 (0.02)	0.06 (0.02)
Average teacher experience	-0.66 (0.93)	2.17 (1.02)	-0.40 (0.89)	2.35 (1.00)	0.12 (0.89)	1.71 (0.89)	0.12 (0.81)	3.22 (0.85)	-0.38 (0.70)	1.86 (0.63)
Percent BA or less	0.02 (0.05)	0.09 (0.05)	-0.04 (0.05)	0.14 (0.05)	-0.06 (0.05)	0.09 (0.04)	-0.03 (0.04)	0.10 (0.04)	-0.07 (0.03)	0.10 (0.03)
Percent BA + 30	0.08 (0.03)	0.03 (0.04)	0.02 (0.03)	0.10 (0.04)	0.01 (0.03)	0.05 (0.03)	0.01 (0.03)	0.03 (0.03)	-0.02 (0.02)	0.03 (0.02)
Percent fully credentialed	0.12 (0.06)	0.05 (0.08)	0.12 (0.05)	0.04 (0.07)	0.11 (0.05)	0.09 (0.06)	0.15 (0.04)	0.08 (0.05)	0.11 (0.03)	0.06 (0.04)
EL teacher access	-10.85 (9.55)	-8.09 (8.17)	-12.18 (9.85)	1.75 (4.23)	-5.94 (8.34)	1.08 (2.52)	2.19 (3.76)	4.22 (1.85)	2.74 (3.17)	3.39 (2.20)
Percent new students	-0.03 (0.06)	0.03 (0.05)	-0.13 (0.06)	0.02 (0.05)	-0.06 (0.06)	0.02 (0.04)	0.07 (0.05)	0.02 (0.04)	0.05 (0.04)	0.04 (0.03)
Average CST score	-0.13 (0.05)	-0.06 (0.05)	-0.16 (0.04)	-0.06 (0.04)	-0.11 (0.04)	-0.07 (0.03)	-0.12 (0.03)	-0.06 (0.03)	-0.09 (0.03)	-0.05 (0.03)
Observations	114,259	110,349	121,637	119,368	115,188	120,396	112,624	109,983	100,010	101,376

Notes: Standard errors clustered by school are in parentheses. Each column contains results from a separate model. Each model also contains controls for age, grade, squared teacher experience, and dummy variables for missing variables.

Table 3 - OLS Results, by Grade and Year

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
<i>Dependent Variable - Listening/Speaking Proficiency</i>										
Bilingual education	-28.60	-25.35	-18.15	-12.80	-2.94	-1.61	-2.27	-2.35	-2.74	-2.74
	(1.28)	(1.42)	(0.87)	(0.90)	(0.93)	(0.82)	(1.07)	(0.92)	(1.41)	(1.51)
	[-0.48]	[-0.43]	[-0.31]	[-0.22]	[-0.05]	[-0.03]	[-0.04]	[-0.04]	[-0.04]	[-0.04]
<i>Dependent Variable - Reading Proficiency</i>										
Bilingual education					3.15	4.07	-0.30	1.25	-2.45	-1.06
					(0.60)	(0.56)	(0.60)	(0.55)	(0.83)	(0.75)
					[0.06]	[0.08]	[-0.01]	[0.03]	[-0.05]	[-0.02]
<i>Dependent Variables - Writing Proficiency</i>										
Bilingual education					-5.59	-5.18	-4.14	-3.04	-2.69	-3.17
					(0.65)	(0.65)	(0.66)	(0.62)	(0.79)	(0.85)
					[-0.11]	[-0.10]	[-0.08]	[-0.06]	[-0.06]	[-0.06]
Observations	114,259	110,349	121,637	119,368	115,188	120,396	112,624	109,983	100,010	101,376

Notes: Standard errors clustered by school are in parentheses. The brackets measure the bilingual effect in standard deviations of the dependent variable. Unlike previous tables, each cell (coefficient, standard error, and standard-deviation effect) is from a separate regression. See the variables and notes to Appendix Tables 1a through 1c for a more complete list of additional control variables.



Table 4 - School Fixed Effect Models of English Proficiency, by Test, Grade, and Year

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
<i>Dependent Variable - Listening/Speaking Proficiency</i>										
Bilingual education	-26.10	-25.59	-20.89	-13.63	-4.41	-4.01	-5.90	-3.43	-6.56	-5.89
	(0.59)	(0.59)	(0.61)	(0.57)	(0.62)	(0.55)	(0.74)	(0.69)	(0.97)	(0.90)
	[-0.44]	[-0.44]	[-0.36]	[-0.24]	[-0.08]	[-0.07]	[-0.10]	[-0.06]	[-0.10]	[-0.09]
<i>Dependent Variable - Reading Proficiency</i>										
Bilingual education					2.22	3.20	-1.46	0.59	-3.77	-3.86
					(0.50)	(0.48)	(0.53)	(0.52)	(0.67)	(0.64)
					[0.05]	[0.06]	[-0.03]	[0.01]	[-0.07]	[-0.08]
<i>Dependent Variables - Writing Proficiency</i>										
Bilingual education					-5.75	-4.78	-3.99	-3.31	-3.34	-5.34
					(0.51)	(0.49)	(0.53)	(0.54)	(0.65)	(0.64)
					[-0.11]	[-0.09]	[-0.08]	[-0.06]	[-0.07]	[-0.11]
Observations	114,259	110,349	121,637	119,368	115,188	120,396	112,624	109,983	100,010	101,376

Notes: Standard errors clustered by school are in parentheses. The brackets measure the bilingual effect in standard deviations of the dependent variable. Each cell (coefficient, standard error, and standard-deviation effect) is from a separate regression. See the variables and notes to Appendix Tables 1a through 1c for a more complete list of additional control variables.

Table 5 - Propensity Score Estimators for Bilingual Education Effect, by Test, Grade, and Year

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
<i>Dependent Variable - Listening/Speaking Proficiency</i>										
Kernel density	[-0.51]**	[-0.48]**	[-0.33]**	[-0.27]**	[-0.06]**	[-0.06]**	[-0.03]**	[-0.07]**	[-0.08]**	[-0.09]**
Four nearest neighbors	[-0.49]**	[-0.47]**	[-0.31]**	[-0.26]**	[-0.03]**	[-0.03]**	[-0.01]	[-0.04]**	[-0.05]**	[-0.03]
Radius <= 0.00004	[-0.43]**	[-0.39]**	[-0.31]**	[-0.18]**	[-0.04]**	[-0.02]	[-0.04]**	[-0.02]	[-0.06]**	[-0.08]**
<i>Dependent Variable - Reading Proficiency</i>										
Kernel density					[0.04]**	[0.04]**	[-0.02]*	[-0.01]	[-0.10]**	[-0.08]**
Four nearest neighbors					[0.07]**	[0.06]**	[0.01]	[0.02]	[-0.06]**	[-0.04]**
Radius <= 0.00004					[0.06]**	[0.10]**	[0.002]	[0.10]**	[-0.04]**	[-0.01]
<i>Dependent Variables - Writing Proficiency</i>										
Kernel density					[-0.12]**	[-0.13]**	[-0.09]**	[-0.10]**	[-0.10]**	[-0.12]**
Four nearest neighbors					[-0.09]**	[-0.10]**	[-0.05]**	[-0.07]**	[-0.07]**	[-0.07]**
Radius <= 0.00004					[-0.12]**	[-0.10]**	[-0.09]**	[-0.11]**	[-0.06]**	[-0.09]**

Notes: The brackets measure the bilingual effect in standard deviations of the dependent variable. \* denotes statistical significance at the five-percent level (two-sided test) and \*\* denotes statistical significance at the one-percent level (two-sided test). Simple standard errors (i.e. not bootstrapped) are calculated for calculating statistical significance. Each cell is from a separate propensity score estimator.

Table 6 - Instrumental Variables Models for English Proficiency, by Test, Grade, and Year  
 Instruments are Measures of Proposition-227 Induced Changes in School-Level Bilingual Education Availability

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
<i>Dependent Variable - Listening/Speaking Proficiency</i>										
Bilingual education	-43.99	-32.75	-23.01	-15.25	2.19	6.46	2.24	3.53	2.12	7.90
	(3.26)	(2.83)	(2.70)	(2.04)	(2.43)	(2.33)	(3.26)	(3.13)	(5.53)	(5.30)
	[-0.75]	[-0.56]	[-0.39]	[-0.27]	[0.04]	[0.12]	[0.04]	[0.06]	[0.03]	[0.13]
<i>Dependent Variable - Reading Proficiency</i>										
Bilingual education					9.23	9.94	3.21	4.55	1.71	0.91
					(1.51)	(1.43)	(1.61)	(1.54)	(2.34)	(2.36)
					[0.19]	[0.20]	[0.07]	[0.09]	[0.03]	[0.02]
<i>Dependent Variables - Writing Proficiency</i>										
Bilingual education					-6.69	-4.29	-6.34	-3.08	-0.71	5.59
					(1.60)	(1.47)	(1.66)	(1.60)	(2.34)	(2.58)
					[-0.13]	[-0.08]	[-0.13]	[-0.06]	[-0.01]	[0.11]
Observations	114,259	110,349	121,637	119,368	115,188	120,396	112,624	109,983	100,010	101,376

Notes: Standard errors clustered by school are in parentheses. The brackets measure the bilingual effect in standard deviations of the dependent variable. Each cell (coefficient, standard error, and standard-deviation effect) is from a separate regression. See the variables and notes to Appendix Tables 1a through 1c for a more complete list of additional control variables.

Table 7 - Instrumental Variables Models for English Proficiency, by Test, Grade, and Year  
Instrument is the Percentage of "Yes" Votes on Proposition 227

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
<i>Dependent Variable - Listening/Speaking Proficiency</i>										
Bilingual education	-53.75	-33.66	-25.37	-16.65	-12.34	12.11	0.36	9.56	51.96	59.95
	(9.93)	(8.81)	(8.27)	(6.82)	(7.60)	(6.85)	(12.38)	(10.07)	(33.28)	(23.92)
	[-0.91]	[-0.58]	[-0.43]	[-0.29]	[-0.22]	[0.22]	[0.01]	[0.16]	[0.81]	[0.96]
<i>Dependent Variable - Reading Proficiency</i>										
Bilingual education					5.54	8.64	-7.67	2.24	-22.11	0.01
					(4.59)	(3.92)	(5.73)	(5.10)	(13.04)	(9.46)
					[0.11]	[0.17]	[-0.16]	[0.05]	[-0.44]	[<0.01]
<i>Dependent Variables - Writing Proficiency</i>										
Bilingual education					-12.37	-7.88	-20.18	-8.97	-38.27	-0.76
					(4.64)	(4.12)	(6.13)	(5.61)	(15.55)	(10.70)
					[-0.23]	[-0.15]	[-0.41]	[-0.18]	[-0.78]	[-0.02]
Observations	111,167	107,579	118,427	116,101	112,159	117,152	109,498	106,841	95,178	96,528

Notes: Standard errors clustered by school are in parentheses. The brackets measure the bilingual effect in standard deviations of the dependent variable. Each cell (coefficient, standard error, and standard-deviation effect) is from a separate regression. See the variables and notes to Appendix Tables 1a through 1c for a more complete list of additional control variables.

Table 8 - Bilingual Education Interacted with Previous Listening/Speaking Quartile, by Grade and Year  
 Dependent Variable - Listening and Speaking Proficiency

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
Bilingual * 1st quartile	-42.78 (2.01)	-43.72 (2.19)	-28.55 (1.11)	-24.89 (1.11)	-15.19 (1.19)	-9.72 (1.00)	-11.75 (1.27)	-8.49 (1.17)	-11.07 (1.62)	-8.88 (1.65)
Bilingual * 2nd quartile	-34.31 (1.46)	-23.97 (1.37)	-13.63 (0.94)	-5.04 (0.97)	2.88 (1.01)	5.37 (0.98)	2.13 (1.31)	-0.62 (1.13)	2.41 (1.77)	1.24 (1.93)
Bilingual * 3rd quartile	-11.41 (1.21)	-6.79 (1.30)	-6.52 (1.17)	0.47 (1.09)	11.24 (1.33)	10.33 (1.00)	8.24 (1.64)	5.23 (1.31)	4.97 (2.22)	3.57 (2.23)
Bilingual * 4th quartile	1.99 (1.47)	4.36 (1.70)	-3.70 (1.66)	1.91 (1.70)	10.92 (1.61)	4.64 (1.76)	9.63 (1.98)	3.08 (1.71)	3.79 (3.16)	-3.08 (3.19)
Observations	114,259	110,349	121,637	119,368	115,188	120,396	112,624	109,983	100,010	101,376

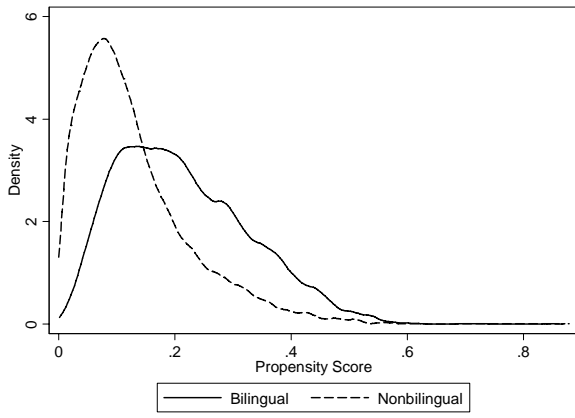
Notes: Standard errors clustered by school are in parentheses. Each column is from a separate regression. See the variables and notes to Appendix Tables 1a through 1c for a more complete list of additional control variables.

Table 9 - Percentage of Students in Fall 2003 Cohort Who Are in Fall 2004 Cohort  
By EL Program, Grade, and Fall 2003 Listening/Speaking Quartile

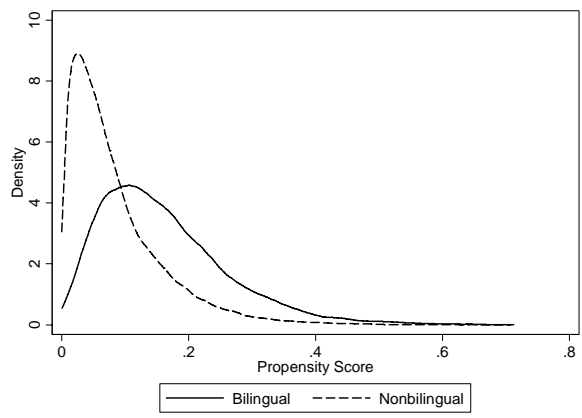
	Fall 2003 Grade					
	KG to 4	KG	1	2	3	4
<b>All Students</b>						
All Quartiles	65.1%	52.0%	69.9%	70.5%	66.9%	64.6%
1st Quartile	55.2%	41.6%	64.2%	67.3%	64.5%	65.6%
2nd Quartile	68.5%	53.2%	72.5%	73.7%	72.3%	71.7%
3rd Quartile	71.7%	55.8%	72.8%	73.3%	71.8%	66.7%
4th Quartile	65.0%	56.5%	70.3%	69.1%	61.8%	57.1%
<b>Not Bilingual</b>						
All Quartiles	64.5%	51.3%	69.2%	70.0%	66.5%	64.7%
1st Quartile	53.4%	40.0%	61.5%	66.0%	63.8%	66.0%
2nd Quartile	67.7%	52.3%	71.9%	73.2%	72.1%	72.0%
3rd Quartile	71.5%	55.2%	72.3%	72.8%	71.7%	66.8%
4th Quartile	64.8%	56.1%	69.9%	68.7%	61.4%	57.2%
<b>Bilingual</b>						
All Quartiles	70.1%	60.6%	74.9%	74.4%	70.3%	63.5%
1st Quartile	65.0%	54.4%	71.9%	71.6%	68.1%	63.1%
2nd Quartile	73.9%	61.9%	77.4%	76.3%	74.5%	69.3%
3rd Quartile	74.5%	65.5%	78.9%	78.1%	73.3%	65.8%
4th Quartile	69.1%	65.3%	76.5%	76.1%	67.1%	55.6%

Figure 1 – Fall 2004 Propensity Scores by Grade and Bilingual Education Participation

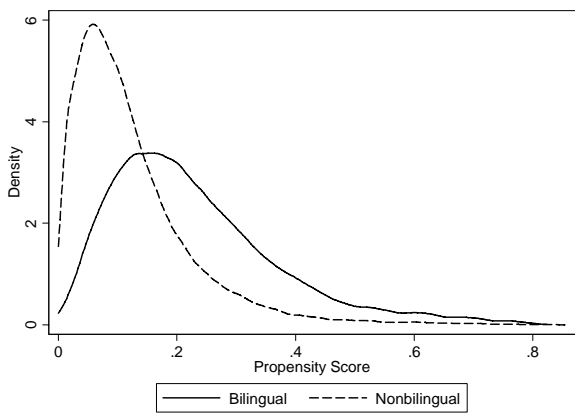
Grade 1



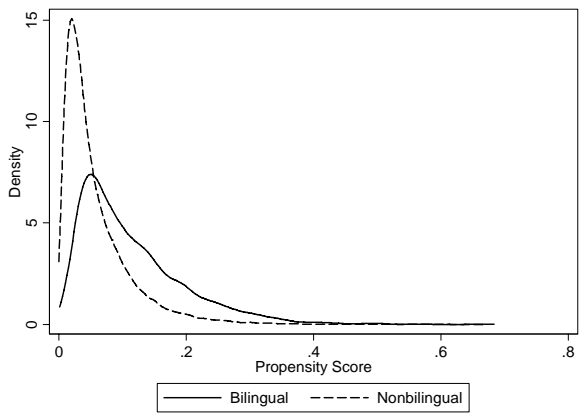
Grade 4



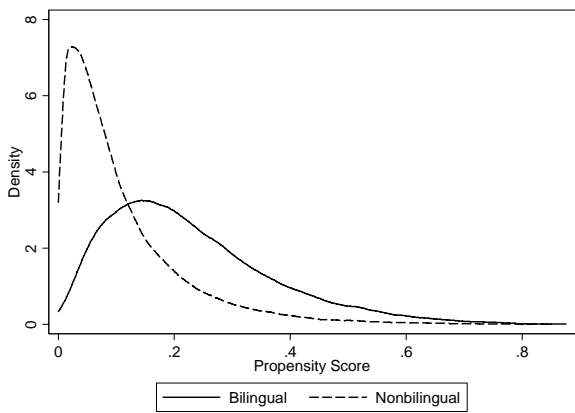
Grade 2



Grade 5



Grade 3



Appendix Table 1a - OLS Results for Listening/Speaking Proficiency, by Grade and Year

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
Bilingual education	-28.60 (1.28)	-25.35 (1.42)	-18.15 (0.87)	-12.80 (0.90)	-2.94 (0.93)	-1.61 (0.82)	-2.27 (1.07)	-2.35 (0.92)	-2.74 (1.41)	-2.74 (1.51)
Lagged listening/speaking	0.300 (0.004)	0.309 (0.004)	0.395 (0.004)	0.420 (0.005)	0.220 (0.006)	0.260 (0.004)	0.224 (0.006)	0.359 (0.005)	0.189 (0.006)	0.309 (0.004)
Lagged reading					0.227 (0.006)	0.224 (0.005)	0.272 (0.006)	0.169 (0.005)	0.268 (0.007)	0.177 (0.006)
Lagged writing					0.181 (0.006)	0.123 (0.005)	0.159 (0.006)	0.104 (0.005)	0.186 (0.007)	0.118 (0.006)
Female	0.738 (0.302)	2.234 (0.294)	1.745 (0.299)	2.344 (0.274)	1.583 (0.296)	0.955 (0.259)	3.265 (0.318)	0.061 (0.284)	3.161 (0.360)	-0.753 (0.317)
Title I	-1.218 (0.883)	-1.571 (0.823)	-1.541 (0.861)	-1.697 (0.750)	-3.325 (0.778)	-0.686 (0.788)	-1.961 (0.904)	-0.572 (0.813)	-1.836 (0.967)	-1.726 (0.858)
Special education	-24.13 (0.956)	-26.63 (0.951)	-18.70 (0.726)	-20.58 (0.706)	-17.86 (0.607)	-17.50 (0.577)	-11.79 (0.615)	-14.26 (0.588)	-9.64 (0.691)	-10.51 (0.595)
Years in current school	4.872 (0.833)	4.638 (0.753)	2.016 (0.368)	1.191 (0.296)	0.851 (0.215)	0.639 (0.182)	0.983 (0.174)	0.267 (0.146)	0.913 (0.146)	0.276 (0.119)
Language index	5.431 (2.106)	0.561 (1.811)	1.800 (1.941)	0.044 (1.797)	2.958 (1.953)	1.486 (1.668)	3.922 (2.184)	0.887 (1.807)	2.136 (2.309)	-0.948 (2.174)
Class size	0.265 (0.349)	0.347 (0.236)	0.374 (0.286)	-0.156 (0.267)	-0.107 (0.223)	-0.388 (0.177)	-0.180 (0.161)	0.014 (0.131)	-0.395 (0.122)	0.074 (0.088)
Percent EL	-0.194 (0.022)	-0.207 (0.022)	-0.128 (0.020)	-0.134 (0.019)	-0.130 (0.018)	-0.126 (0.017)	-0.110 (0.021)	-0.108 (0.020)	-0.064 (0.023)	-0.129 (0.020)
Avg teacher experience	0.193 (0.930)	-0.079 (0.593)	-0.408 (0.727)	-0.081 (0.509)	0.311 (0.627)	-0.330 (0.523)	0.694 (0.718)	0.432 (0.586)	0.275 (0.772)	-0.528 (0.633)



Appendix Table 1a (Continued) - OLS Results for Listening/Speaking Proficiency, by Grade and Year

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
Percent BA or less	-0.066 (0.039)	-0.057 (0.034)	-0.071 (0.037)	0.069 (0.028)	-0.039 (0.031)	-0.017 (0.029)	0.011 (0.036)	0.029 (0.030)	-0.049 (0.040)	0.019 (0.032)
Percent BA +30	-0.074 (0.027)	-0.072 (0.024)	-0.047 (0.025)	-0.027 (0.019)	0.019 (0.022)	-0.019 (0.020)	0.038 (0.026)	0.014 (0.022)	-0.013 (0.027)	-0.018 (0.025)
Percent fully credentialed	-0.002 (0.047)	-0.027 (0.058)	-0.031 (0.036)	0.055 (0.049)	-0.035 (0.037)	-0.032 (0.044)	-0.050 (0.048)	0.006 (0.051)	-0.018 (0.049)	0.042 (0.052)
EL teacher access	5.424 (2.920)	-0.259 (2.701)	5.125 (3.215)	-1.155 (2.453)	0.818 (3.209)	2.619 (2.050)	4.710 (3.577)	8.828 (2.379)	10.338 (3.988)	3.841 (3.079)
Percent new students	-0.024 (0.052)	-0.169 (0.059)	0.017 (0.042)	-0.032 (0.030)	-0.047 (0.043)	-0.042 (0.029)	-0.065 (0.051)	-0.025 (0.029)	0.050 (0.055)	-0.031 (0.031)
Average CST score	0.137 (0.032)	0.071 (0.030)	0.139 (0.025)	0.113 (0.022)	0.113 (0.024)	0.092 (0.022)	0.099 (0.029)	0.087 (0.024)	0.150 (0.032)	0.073 (0.028)
Observations	114,259	110,349	121,637	119,368	115,188	120,396	112,624	109,983	100,010	101,376

Notes: Standard errors clustered by school are in parentheses. The brackets measure the bilingual effect in standard deviations of the dependent variable. R-squareds range from 0.29 to 0.39. All models also contain controls for squared teacher experience, age, grade, and dummy variables for missing variables (except bilingual education).

Appendix Table 1b - OLS Results for Reading Proficiency, by Grade and Year

	Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004
Bilingual education	3.15 (0.60)	4.07 (0.56)	-0.30 (0.60)	1.25 (0.55)	-2.45 (0.83)	-1.06 (0.75)
Lagged listen/speak	0.025 (0.003)	0.019 (0.002)	0.021 (0.003)	0.031 (0.003)	0.026 (0.003)	0.026 (0.003)
Lagged reading	0.406 (0.005)	0.427 (0.004)	0.457 (0.004)	0.431 (0.004)	0.470 (0.006)	0.466 (0.005)
Lagged writing	0.214 (0.005)	0.220 (0.005)	0.194 (0.004)	0.212 (0.004)	0.160 (0.006)	0.173 (0.005)
Female	3.319 (0.227)	2.930 (0.217)	5.600 (0.214)	3.092 (0.220)	5.865 (0.236)	2.770 (0.231)
Title I	-1.621 (0.517)	-1.737 (0.453)	-1.835 (0.496)	-1.421 (0.435)	-1.230 (0.508)	-0.777 (0.451)
Special education	-17.13 (0.615)	-16.22 (0.526)	-17.98 (0.507)	-17.65 (0.479)	-19.68 (0.518)	-18.24 (0.472)
Years in current schl	-0.316 (0.157)	-0.219 (0.142)	-0.259 (0.096)	-0.428 (0.090)	-0.172 (0.087)	-0.313 (0.076)
Language index	-0.972 (1.210)	-2.269 (1.088)	-1.071 (1.071)	-2.896 (0.984)	-2.594 (1.010)	-2.219 (1.040)
Class size	-0.036 (0.138)	-0.259 (0.113)	0.036 (0.063)	0.018 (0.058)	0.096 (0.049)	0.011 (0.047)
Percent EL	-0.019 (0.011)	0.009 (0.010)	0.012 (0.010)	0.012 (0.010)	0.021 (0.010)	0.006 (0.010)
Avg teacher exp	-0.009 (0.348)	0.189 (0.326)	0.823 (0.354)	0.419 (0.320)	-0.257 (0.372)	-0.239 (0.290)
Pct BA or less	0.005 (0.020)	-0.047 (0.017)	0.051 (0.019)	-0.023 (0.016)	0.003 (0.019)	-0.030 (0.017)
Pct BA +30	0.002 (0.013)	-0.034 (0.013)	0.017 (0.012)	-0.013 (0.012)	-0.005 (0.012)	-0.028 (0.012)
Pct fully credentialed	0.076 (0.026)	0.056 (0.028)	0.080 (0.022)	0.061 (0.026)	0.099 (0.021)	0.086 (0.026)
EL teacher access	-1.503 (1.646)	-2.317 (1.085)	1.193 (2.161)	-1.784 (1.310)	-2.142 (2.010)	0.899 (1.764)
Percent new students	-0.022 (0.024)	0.007 (0.017)	0.010 (0.023)	0.005 (0.017)	0.048 (0.022)	0.006 (0.016)
Average CST score	0.108 (0.016)	0.109 (0.013)	0.119 (0.014)	0.090 (0.012)	0.153 (0.013)	0.092 (0.012)
Observations	115,188	120,396	112,624	109,983	100,010	101,376

Notes: Standard errors clustered by school are in parentheses. The brackets measure the bilingual effect in standard deviations of the dependent variable. R-squareds range from 0.44 to 0.52. All models also contain controls for squared teacher experience, age, grade, and dummy variables for missing variables (except bilingual education).

Appendix Table 1c - OLS Results for Writing Proficiency, by Grade and Year

	Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004
Bilingual education	-5.59 (0.65)	-5.18 (0.65)	-4.14 (0.66)	-3.04 (0.62)	-2.69 (0.79)	-3.17 (0.85)
Lagged listen/speak	0.048 (0.003)	0.039 (0.003)	0.034 (0.003)	0.022 (0.003)	0.034 (0.003)	0.015 (0.003)
Lagged reading	0.309 (0.005)	0.328 (0.005)	0.271 (0.004)	0.253 (0.004)	0.250 (0.005)	0.237 (0.004)
Lagged writing	0.339 (0.006)	0.352 (0.005)	0.341 (0.005)	0.377 (0.005)	0.326 (0.006)	0.369 (0.005)
Female	4.239 (0.238)	4.022 (0.222)	5.150 (0.216)	4.388 (0.224)	5.264 (0.228)	3.904 (0.230)
Title I	-1.193 (0.511)	-0.923 (0.472)	-0.309 (0.461)	-0.446 (0.436)	-0.271 (0.481)	-0.697 (0.475)
Special education	-25.13 (0.715)	-24.71 (0.601)	-18.97 (0.551)	-19.33 (0.519)	-17.79 (0.511)	-16.22 (0.496)
Years in current schl	0.212 (0.175)	0.324 (0.144)	0.357 (0.100)	0.080 (0.098)	0.295 (0.085)	0.021 (0.083)
Language index	2.070 (1.278)	1.776 (1.173)	1.736 (1.101)	-0.118 (1.043)	-0.608 (1.035)	0.933 (1.174)
Class size	0.119 (0.136)	-0.432 (0.109)	0.020 (0.066)	-0.030 (0.059)	0.013 (0.044)	-0.007 (0.058)
Percent EL	-0.016 (0.012)	0.001 (0.011)	0.008 (0.010)	0.005 (0.010)	0.023 (0.010)	0.006 (0.011)
Avg teacher exp	0.972 (0.367)	-0.134 (0.352)	0.995 (0.356)	0.309 (0.331)	-0.014 (0.367)	0.025 (0.329)
Pct BA or less	0.002 (0.022)	-0.052 (0.019)	0.022 (0.017)	-0.012 (0.017)	-0.020 (0.020)	-0.028 (0.019)
Pct BA +30	0.009 (0.014)	-0.010 (0.013)	0.028 (0.012)	0.016 (0.013)	0.005 (0.013)	0.010 (0.013)
Pct fully credentialed	0.126 (0.024)	0.126 (0.028)	0.106 (0.020)	0.114 (0.027)	0.110 (0.022)	0.119 (0.028)
EL teacher access	-1.181 (1.867)	-1.495 (1.178)	-0.947 (2.251)	-0.834 (1.524)	-2.194 (2.046)	3.149 (1.953)
Percent new students	0.027 (0.027)	0.033 (0.019)	-0.009 (0.025)	0.016 (0.021)	0.042 (0.022)	0.029 (0.017)
Average CST score	0.111 (0.017)	0.106 (0.013)	0.112 (0.015)	0.096 (0.012)	0.135 (0.013)	0.078 (0.013)
Observations	115,188	120,396	112,624	109,983	100,010	101,376

Notes: Standard errors clustered by school are in parentheses. R-squareds range from 0.49 to 0.53. All models also contain controls for squared teacher experience, age, grade, and dummy variables for missing variables (except bilingual education).

Appendix Table 2 - Instrumental Variables First Stage Results  
 Linear Probability Model for Participation in Bilingual Education  
 Instruments are Measures of Proposition-227 Induced Changes in School-Level Bilingual Education Availability

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
Actual change in pct bilingual	0.59238 (0.02958)	0.64013 (0.03020)	0.57623 (0.02929)	0.61179 (0.03046)	0.54749 (0.03009)	0.56398 (0.02934)	0.42764 (0.02985)	0.46106 (0.02998)	0.42764 (0.02985)	0.30427 (0.02910)
Mechanical change in pct bilingual	-0.66134 (0.02726)	-0.69743 (0.02629)	-0.65361 (0.02748)	-0.69755 (0.02651)	-0.64044 (0.02774)	-0.65087 (0.02592)	-0.53905 (0.02828)	-0.54340 (0.02741)	-0.53905 (0.02828)	-0.36746 (0.02695)
Lagged listening/speaking	-0.00039 (0.00004)	-0.00045 (0.00003)	-0.00075 (0.00005)	-0.00079 (0.00004)	-0.00017 (0.00003)	-0.00027 (0.00003)	-0.00021 (0.00003)	-0.00013 (0.00003)	-0.00021 (0.00003)	-0.00011 (0.00002)
Lagged reading					-0.00008 (0.00005)	-0.00010 (0.00004)	0.00014 (0.00004)	0.00014 (0.00003)	0.00014 (0.00004)	0.00004 (0.00004)
Lagged writing					-0.00066 (0.00005)	-0.00071 (0.00005)	-0.00048 (0.00005)	-0.00058 (0.00005)	-0.00048 (0.00005)	-0.00035 (0.00004)
Female	-0.00072 (0.00180)	0.00205 (0.00179)	0.00161 (0.00172)	0.00323 (0.00168)	0.00648 (0.00175)	0.00930 (0.00166)	0.00426 (0.00162)	0.00453 (0.00156)	0.00426 (0.00162)	0.00500 (0.00146)
Title I	-0.00187 (0.00910)	-0.01087 (0.00933)	-0.00533 (0.00874)	-0.01441 (0.00919)	-0.01508 (0.00932)	-0.01777 (0.00935)	-0.01228 (0.00831)	-0.00048 (0.00869)	-0.01228 (0.00831)	0.00471 (0.00755)
Special education	-0.03446 (0.00527)	-0.03198 (0.00532)	-0.04090 (0.00471)	-0.04099 (0.00468)	-0.06241 (0.00478)	-0.05902 (0.00480)	-0.04162 (0.00409)	-0.04703 (0.00430)	-0.04162 (0.00409)	-0.02920 (0.00335)
Years in current school	0.02714 (0.01057)	0.03506 (0.00769)	0.02890 (0.00485)	0.03428 (0.00397)	0.02163 (0.00279)	0.02411 (0.00210)	0.01197 (0.00173)	0.01510 (0.00149)	0.01197 (0.00173)	0.00448 (0.00122)
Language index	-0.02175 (0.02223)	-0.06645 (0.02264)	-0.02752 (0.02187)	-0.07005 (0.02285)	-0.02202 (0.02224)	-0.04504 (0.02188)	-0.02536 (0.01971)	-0.04973 (0.02074)	-0.02536 (0.01971)	-0.05611 (0.02004)
Class size	-0.00688 (0.00345)	-0.00857 (0.00242)	-0.00780 (0.00310)	-0.00626 (0.00323)	-0.00765 (0.00298)	-0.00935 (0.00200)	-0.00460 (0.00121)	-0.00691 (0.00131)	-0.00460 (0.00121)	-0.00749 (0.00155)
Percent EL	0.00080 (0.00027)	0.00073 (0.00029)	0.00065 (0.00025)	0.00031 (0.00029)	0.00052 (0.00026)	-0.00002 (0.00025)	0.00013 (0.00023)	-0.00055 (0.00022)	0.00013 (0.00023)	-0.00041 (0.00019)

Appendix Table 2 (Continued) - Instrumental Variables First Stage Results  
 Linear Probability Model for Participation in Bilingual Education  
 Instruments are Measures of Proposition-227 Induced Changes in School-Level Bilingual Education Availability

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
Avg teacher exp	-0.00586 (0.00829)	0.01886 (0.00942)	-0.00270 (0.00872)	0.01822 (0.00884)	0.00132 (0.00826)	0.01155 (0.00859)	-0.00459 (0.00734)	0.02478 (0.00702)	-0.00459 (0.00734)	0.01443 (0.00534)
Pct BA or less	-0.00121 (0.00043)	-0.00053 (0.00043)	-0.00192 (0.00042)	-0.00009 (0.00044)	-0.00204 (0.00044)	-0.00047 (0.00041)	-0.00158 (0.00041)	0.00026 (0.00040)	-0.00158 (0.00041)	0.00069 (0.00034)
Pct BA +30	-0.00031 (0.00028)	-0.00084 (0.00028)	-0.00089 (0.00029)	-0.00028 (0.00028)	-0.00087 (0.00028)	-0.00055 (0.00026)	-0.00066 (0.00026)	-0.00057 (0.00025)	-0.00066 (0.00026)	-0.00036 (0.00022)
Pct fully credentialed	-0.00012 (0.00049)	0.00060 (0.00074)	-0.00001 (0.00047)	0.00060 (0.00068)	0.00012 (0.00045)	0.00101 (0.00058)	0.00077 (0.00037)	0.00115 (0.00053)	0.00077 (0.00037)	0.00075 (0.00045)
EL teacher access	0.03574 (0.02357)	0.02251 (0.02115)	0.02082 (0.01700)	0.05023 (0.02436)	0.04316 (0.02617)	0.01973 (0.01356)	0.03972 (0.02286)	0.02267 (0.01887)	0.03972 (0.02286)	0.01287 (0.02196)
Percent new students	0.00138 (0.00069)	0.00023 (0.00047)	0.00050 (0.00062)	0.00027 (0.00042)	0.00095 (0.00068)	0.00028 (0.00039)	0.00210 (0.00069)	0.00039 (0.00044)	0.00210 (0.00069)	0.00068 (0.00040)
Average CST score	-0.00026 (0.00039)	-0.00046 (0.00034)	-0.00054 (0.00034)	-0.00058 (0.00033)	-0.00036 (0.00036)	-0.00056 (0.00028)	-0.00058 (0.00032)	-0.00051 (0.00027)	-0.00058 (0.00032)	-0.00048 (0.00025)
Observations	114,259	110,349	121,637	119,368	115,188	120,396	112,624	109,983	100,010	101,376

Notes: Standard errors clustered by school are in parentheses. The results from each column are from a separate regression. See the variables and notes to Table 3 for a more complete list of additional control variables.

Appendix Table 3 - Instrumental Variables First Stage Results  
 Linear Probability Model for Participation in Bilingual Education  
 Instrument is the Percentage of "Yes" Votes on Proposition 227

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
Percentage voted "Yes" on Proposition 227	-0.00411 (0.00052)	-0.00440 (0.00055)	-0.00385 (0.00048)	-0.00401 (0.00050)	-0.00382 (0.00048)	-0.00423 (0.00046)	-0.00286 (0.00043)	-0.00285 (0.00042)	-0.00124 (0.00032)	-0.00159 (0.00035)
Lagged listening/speaking	-0.00049 (0.00004)	-0.00054 (0.00003)	-0.00094 (0.00005)	-0.00102 (0.00005)	-0.00015 (0.00003)	-0.00029 (0.00003)	-0.00024 (0.00003)	-0.00012 (0.00003)	-0.00012 (0.00003)	-0.00011 (0.00003)
Lagged reading					0.00004 (0.00005)	-0.00006 (0.00005)	0.00034 (0.00005)	0.00024 (0.00004)	0.00018 (0.00005)	0.00009 (0.00004)
Lagged writing					-0.00109 (0.00006)	-0.00105 (0.00006)	-0.00078 (0.00006)	-0.00084 (0.00006)	-0.00039 (0.00004)	-0.00043 (0.00005)
Female	0.00065 (0.00202)	0.00153 (0.00204)	0.00138 (0.00192)	0.00588 (0.00196)	0.00878 (0.00194)	0.01135 (0.00180)	0.00523 (0.00177)	0.00635 (0.00176)	0.00177 (0.00153)	0.00576 (0.00155)
Title I	-0.00447 (0.01253)	-0.01503 (0.01331)	-0.00612 (0.01185)	-0.01223 (0.01278)	-0.01088 (0.01244)	-0.01742 (0.01265)	-0.00565 (0.01032)	0.01511 (0.00934)	-0.00535 (0.00916)	0.01240 (0.00772)
Special education	-0.02591 (0.00596)	-0.02505 (0.00606)	-0.03856 (0.00547)	-0.04224 (0.00537)	-0.07272 (0.00533)	-0.07183 (0.00515)	-0.04609 (0.00448)	-0.05413 (0.00471)	-0.03002 (0.00373)	-0.03159 (0.00365)
Years in current school	0.01592 (0.01017)	0.03172 (0.00685)	0.02571 (0.00481)	0.03967 (0.00386)	0.02335 (0.00301)	0.03032 (0.00218)	0.01616 (0.00203)	0.01927 (0.00171)	0.00537 (0.00121)	0.00781 (0.00113)
Language index	0.13426 (0.02854)	0.09828 (0.02964)	0.13345 (0.02673)	0.09630 (0.02990)	0.13684 (0.02666)	0.10580 (0.02743)	0.11518 (0.02306)	0.07448 (0.02545)	0.05924 (0.01881)	0.02416 (0.02364)
Class size	-0.01501 (0.00434)	-0.00969 (0.00343)	-0.01679 (0.00381)	-0.00842 (0.00336)	-0.01544 (0.00343)	-0.01003 (0.00258)	-0.00421 (0.00185)	-0.00571 (0.00173)	-0.00241 (0.00150)	-0.00308 (0.00148)
Percent EL	0.00172 (0.00034)	0.00215 (0.00037)	0.00171 (0.00031)	0.00168 (0.00036)	0.00151 (0.00032)	0.00107 (0.00031)	0.00102 (0.00028)	0.00047 (0.00027)	0.00055 (0.00024)	0.00028 (0.00021)

Appendix Table 3 (Continued) - Instrumental Variables First Stage Results  
 Linear Probability Model for Participation in Bilingual Education  
 Instrument is the Percentage of "Yes" Votes on Proposition 227

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004	2003	2004	2003	2004
Avg teacher exp	-0.00957 (0.01080)	0.02283 (0.00945)	-0.00738 (0.01027)	0.02452 (0.00910)	-0.00081 (0.01043)	0.01853 (0.00890)	0.00089 (0.00925)	0.03147 (0.00669)	-0.00607 (0.00782)	0.01885 (0.00491)
Pct BA or less	-0.00034 (0.00056)	0.00062 (0.00056)	-0.00102 (0.00055)	0.00119 (0.00055)	-0.00124 (0.00054)	0.00068 (0.00052)	-0.00084 (0.00049)	0.00115 (0.00048)	-0.00109 (0.00040)	0.00128 (0.00038)
Pct BA +30	0.00061 (0.00035)	0.00007 (0.00036)	0.00004 (0.00035)	0.00081 (0.00036)	-0.00006 (0.00035)	0.00035 (0.00033)	-0.00001 (0.00031)	0.00021 (0.00030)	-0.00028 (0.00027)	0.00036 (0.00025)
Pct fully credentialed	0.00230 (0.00059)	0.00226 (0.00091)	0.00223 (0.00057)	0.00195 (0.00087)	0.00226 (0.00054)	0.00267 (0.00075)	0.00246 (0.00044)	0.00213 (0.00067)	0.00156 (0.00034)	0.00169 (0.00051)
EL teacher access	0.01970 (0.02977)	0.04313 (0.02731)	0.01158 (0.01903)	0.07424 (0.02867)	0.02761 (0.03279)	0.03403 (0.01737)	0.04508 (0.02787)	0.05493 (0.02233)	0.04529 (0.02767)	0.05385 (0.02411)
Percent new students	0.00072 (0.00071)	0.00075 (0.00064)	-0.00021 (0.00063)	0.00040 (0.00056)	0.00051 (0.00070)	0.00027 (0.00051)	0.00162 (0.00072)	0.00036 (0.00054)	0.00139 (0.00059)	0.00106 (0.00049)
Average CST score	-0.00019 (0.00050)	0.00024 (0.00047)	-0.00043 (0.00043)	0.00013 (0.00044)	-0.00003 (0.00045)	0.00015 (0.00035)	-0.00043 (0.00038)	-0.00019 (0.00034)	-0.00052 (0.00030)	-0.00018 (0.00028)
Observations	111,167	107,579	118,427	116,101	112,159	117,152	109,498	106,841	95,178	96,528

Notes: Standard errors clustered by school are in parentheses. The results from each column are from a separate regression. See the variables and notes to Table 3 for a more complete list of additional control variables.

Appendix Table 4 - Bilingual Education Interacted with Previous Listening/Speaking Quartile  
 Dependent Variable is Reading or Writing Proficiency for Each Grade and Year

	Grade 3		Grade 4		Grade 5	
	2003	2004	2003	2004	2003	2004
<i>Dependent Variable - Reading Proficiency</i>						
Bilingual * 1st quartile	-1.59 (0.76)	-0.13 (0.75)	-3.51 (0.78)	-1.36 (0.81)	-5.00 (1.10)	-4.71 (1.20)
Bilingual * 2nd quartile	3.48 (0.75)	6.74 (0.78)	-0.28 (0.80)	-0.30 (0.66)	-3.33 (0.97)	-1.32 (0.92)
Bilingual * 3rd quartile	8.55 (0.95)	9.53 (0.78)	3.69 (0.98)	4.53 (0.80)	-0.36 (1.21)	2.02 (0.97)
Bilingual * 4th quartile	12.46 (0.95)	11.22 (1.07)	6.58 (1.19)	7.78 (0.84)	4.14 (1.92)	5.49 (1.48)
<i>Dependent Variable - Writing Proficiency</i>						
Bilingual * 1st quartile	-10.85 (0.87)	-8.84 (0.84)	-8.69 (0.90)	-6.52 (0.91)	-6.39 (1.17)	-6.61 (1.56)
Bilingual * 2nd quartile	-3.75 (0.82)	-1.81 (0.82)	-1.81 (0.85)	-1.98 (0.80)	-1.83 (1.03)	-2.41 (0.96)
Bilingual * 3rd quartile	0.28 (0.99)	-0.36 (0.80)	1.11 (1.00)	-0.06 (0.87)	0.90 (1.09)	-0.04 (1.00)
Bilingual * 4th quartile	1.85 (0.93)	-1.50 (1.18)	0.82 (1.21)	1.47 (0.80)	2.49 (1.35)	0.48 (1.61)
Observations	115,188	120,396	112,624	109,983	100,010	101,376

Notes: Standard errors are in parentheses. The results from each column and subject are from a separate regression (there are a total of 12 regressions). See the variables and notes to Table 3 for a more complete list of additional control variables.