DOES BILINGUAL EDUCATION AFFECT EDUCATIONAL ATTAINMENT AND LABOR MARKET OUTCOMES?

Evidence from the National Education Longitudinal Study of 1988 and High School and Beyond

Mark Hugo Lopez
School of Public Affairs
University of Maryland

November 2002

I thank Orley Ashenfelter, David Card, Linda Chavez, William Evans, Joseph Guzman, Dean Hyslop, Stephen Krashen, Robert McMillan, Marie Mora and Cecilia Rouse for useful comments on previous drafts of this paper. I further thank four anonymous referees, members of the Princeton Labor Lunch, Princeton Industrial Relations Seminar Series, the College of William and Mary Labor Seminar Series, the New Mexico State University Economics series, Southwestern Economic Association Participants, BALAS Conference Participants, Western Economic Association Participants, Participants at the Virginia Commonwealth University Labor Economics Seminar Series, the University of Maryland Labor Economics Seminar series, and the School of Public Affairs at the University of Maryland for helpful comments. Support from the Industrial Relations Section of Princeton University and the School of Public Affairs at the University of Maryland is gratefully acknowledged. All errors in fact or interpretation are my own.

DOES BILINGUAL EDUCATION AFFECT EDUCATIONAL ATTAINMENT AND LABOR MARKET OUTCOMES?

Evidence from the National Longitudinal Study of 1988 and High School and Beyond

ABSTRACT

Over the past 20 years, the U.S. school system has witnessed a large influx of students who are identified as limited English proficient (LEP). Schools have responded to this increase by instituting a variety of English programs for non-native English speakers with the primary objective of improving the English language skills of these students. This paper examines the relationship between participation in an English language assistance program and several educational and labor market outcomes. Using matched propensity score analysis samples and an instrumental variables approach to address negative selection into these programs, I find little evidence that these programs improve students' educational attainment and labor market earnings.

JEL Codes: I21, I28, J15

I. INTRODUCTION

Economists, educators and policy analysts have long recognized that English proficiency is a critical determinant of success in the United States' labor market.[1] For example, Hispanic workers earn 20-30% less than white-non-Hispanic workers (Chavez 1991, Chapa 1989). Two explanations are generally proposed to explain this stylized fact. First, Hispanics generally acquire less human capital than whites. This manifests itself in the form of fewer years of education completed and lower levels of English proficiency. Neal and Johnson (1996) present evidence that the wage gap for Hispanic men and women is almost entirely due to lower levels of skill acquisition before students enter the job market. Similarly, Trejo (1997) argues that, for third generation Mexican Americans, the wage gap is driven by low-levels of educational and English capital acquisition rather than unequal payoffs to that human capital. Second, the low earnings of Hispanics are compounded by the recentness of their immigration to the U.S. Chavez (1991) and Rodriguez (1997) argue that subsequent generations of Hispanics are finding economic success as previous waves of non-Hispanic immigrants have, requiring no special assistance to ease their transition into the U.S. economy. Others, however, argue that policies are necessary to ease and accelerate the transition of limited English proficient populations into the economic mainstream (Krashen 1996).

One policy that attempts to achieve that goal is bilingual education, a program specifically designed to improve the English language skills of limited English proficient (LEP) students.[2] The marginal cost of primary and secondary school-based bilingual programs

---

[1] See, for example, Chiswick and Miller (1995), Mora (1996), Bloom and Grenier (1993), Smith (1990), Kossoudji (1988), McManus (1985), Grenier (1984), McManus et al (1983), Reimers (1983), and Carliner (1976).

[2] As I explain in detail later in the paper, the rubric of "bilingual education" encompasses a wide variety of programs, including English as a Second Language (ESL) programs, native language instruction programs, immersion programs, and Bilingual-Bicultural programs.

amounts to 2 to 3 billion dollars for the 3.2 million LEP students in U.S. schools.[3] However, despite recognition that improving educational attainment and educational opportunity for LEP students can theoretically lead to better academic outcomes for these students, considerable disagreement over whether bilingual education programs as implemented help achieve that goal continues despite over 25 years of school districts providing bilingual program services.[4] Recently, voters in California passed Proposition 227, which reduced the prevalence of bilingual education programs in the State of California, beginning in the academic year 1998-1999 (Lopez 2000).

Studies evaluating the effectiveness of bilingual education programs have centered mainly on the effects of bilingual education programs on English proficiency and reading achievement outcomes. Using a meta-analysis, Greene (1998) finds that on average, scientifically acceptable research on bilingual education programs suggests modest improvements in student English academic achievement. Greene further suggests that in the few, small randomized studies that do exist; there are modest gains in English achievement. However, measuring the success of bilingual education is not limited solely to English proficiency and reading achievement; there are other outcomes on which bilingual education could have an impact for limited English proficient participants. These include the effects of bilingual programs on retention rates, college attendance rates, educational attainment and labor market earnings. Little work assesses the impact of bilingual education programs on these outcomes.[5]

---

[3] Figures on the cost of bilingual education programs in the U.S. are difficult to find. The ALEC Foundation provides one of the few estimates available (at 2 to 3 billion), and Youngblood (1995) and the City of New York (1994) both estimate the cost at somewhere between an additional $900 to $1,000 dollars per student in New York City. Part of the difficulty in estimating the cost of these programs is the way in which they are classified, and the fact that the federal government does not require reporting on the part of state or local education agencies on the costs of programs.

[4] Crawford(1995) provides an excellent review of the history and politics of bilingual education.

[5] Exceptions include Guzman (2002), Lopez and Mora (1998), Mora(1998), Cheng (1996) and Curiel, Rosenthal, and Richeck (1986).

In this paper I examine whether participation in an English language assistance program, such as bilingual education, affects educational attainment and labor market performance. The educational attainment measures I use are whether or not a student has completed a high school diploma, excluding the GED, right out of high school, enrolled in a four year college after high school, having attained a high school diploma or a bachelor's degree by age 28, and years of education completed all by the age of 28. I also consider the impact of English language assistance programs on earnings in the labor market at age 28.

Using a propensity score analysis to identify an appropriate comparison group against which to compare bilingual education participants and attempting an instrumental variables analysis, I find no evidence of positive effects for bilingual education in this data. In fact most effects are negative. This is relative to a group of students who would have qualified for participation in a bilingual education program, but were not assigned to a program. The paper proceeds as follows. In section II I discuss bilingual education and its structure in the U.S.; in section III I discuss the data I use in this paper. In section IV I model the effect of bilingual education participation on outcomes. I present my results in section V, and conclude with section VI.

II. BACKGROUND ON BILINGUAL EDUCATION

English language assistance programs are not characterized by any strict federal policy. However, the passage of the Bilingual Education Act of 1967 and the subsequent authorization of funds for bilingual education programs under Title VII of the Elementary and Secondary Education Act of 1968 led to the development of experimental bilingual education programs without mandating the existence of such programs. Currently the federal government spends approximately $150 million a year on a variety of national bilingual education activities ranging from program development to the gathering of statistics on the number of LEP students in the U.S. and the services they receive.

Following the Federal government's lead, by 1981 several had states passed laws that either allowed for non-language subjects to be taught in languages other than English or mandated the establishment of bilingual education programs. As of 2001, California and Arizona voters had passed laws that significantly reduced the prevalence of bilingual education in public schools. Massachusetts and Colorado are currently considering similar laws in 2002.

In the U.S. there exist three forms of English language assistance programs, each with the stated objective of teaching English to non-English speaking children:

- English as a Second Language (ESL): Limited English proficient students spend most of the day in English-only classes, but are removed part of the day for English instruction.

- Transitional Bilingual Education (TBE): Limited English proficient students are taught non-language subjects in their native language until their English is sufficient enough to learn wholly in English.

- Developmental Bilingual Education: Native-English speaking and LEP students are taught in both English and the native language of the LEP students. The focus of the program is two-way bilingualism.[6]

Most schools offer services extensive enough to fit under the definition of transitional bilingual education, but depending on the resources of the school and the needs of the students, programs other than transitional bilingual education may be offered. Evidence from the Department of Education on the number of LEP students and the services they receive suggests that over 80 percent of services offered by school districts are of a form similar to transitional bilingual education (Henderson et al. 1994).[7]

---

[6] These definitions come from the ALEC Report, 1994.

[7] For the purposes of evaluation, I do not identify English immersion as a bilingual education program since schools spend no extra money on teachers or books when a student is immersed into English-only environments.

III. DATA AND SAMPLE

Data

The data I use are drawn from the National Center for Education Statistics' National Education Longitudinal Study of 1988 (NELS:88) and High School and Beyond (HSB), both of which offer extensive information on the language use of students and their participation in bilingual education programs. NELS:88 is a stratified sample of public and private school students in the eighth grade during 1988. In the initial base year survey, there were approximately 25,000 students; follow-up surveys were conducted in 1990, 1992, 1994, and 1999. HSB, also a stratified sample, contains data from interviews with high school sophomores in 1980, and again in 1982, 1984, 1986 and 1992, gathering information on a number of background characteristics, academic and high school characteristics, aspirations, and work experience. I employ the sophomore cross-section of background, school, and language characteristics from the base year survey in 1980, and information on long-run educational attainment and labor market performance from the fourth follow-up in 1992. Further, I use only public school students who participated in all waves of the survey through 1992 for NELS:88, and base year, first follow-up, and fourth follow-up (1992) participants for HSB.

One advantage of the NELS:88 is that it began in 1988 when survey participants were in the eighth grade. Thus students had the opportunity to participate in well-established bilingual programs. Further, this survey's design captures those students most at risk of dropping out, i.e. those who plan to leave school immediately upon reaching their sixteenth or seventeenth birthdays, or those who had been held back in earlier grades.[8]

[8] Boozer (1992) reports that "early" dropout status appears particularly sensitive to background factors and schools characteristics. Furthermore, Fernandez and Nielsen (1986) find that among Latino students, early dropouts (i.e. before or during the tenth grade) account for a sizable fraction of Hispanic dropouts.

In contrast to NELS:88, HSB provides data on the types of programs attended, allowing me to distinguish between English language assistance program types. Further, HSB provides important post-secondary information on students ten years after high school.

Sample Construction

To evaluate the effectiveness of English language assistance programs, ideally I would conduct an experiment in which some students who are identified as eligible for a program are randomly assigned to a treatment group (program participation) and the remainder is assigned to a control group (no program participation). After some period of time, I could simply compare the treatment and control groups on a range of outcomes to determine the effects, positive or negative, of programs like bilingual education. Since I am unable to conduct such an experiment, my first objective is to construct a comparison group by identifying a group of students who did not participate in an English language assistance program, but who were likely to have been eligible. Since these programs are offered only to LEP students, I define an "at risk" group of potentially eligible participants using guidelines based on the actual LEP identification practices of school districts. I then compare the outcomes of those students who received bilingual education with those who qualified for the program but were immersed in an English-only environment. [9]

Typically, the procedure most school districts follow for LEP identification begins by asking a student (sometimes with the parent) to complete a Home Language Survey.[10] This survey asks students and parents what the home and first language of a student is, in addition to a few other questions on language usage. If a student answers "a language other than English" to

[9] See Willig (1985) for a discussion of issues directly related to bilingual education program evaluation. Another standard for evaluation of bilingual education programs would be to compare bilingual education participants to native English speaking students. However, this is not a well-defined experiment since the appropriate experiment is to evaluate the effect of bilingual education programs on a group of randomly assigned students who would qualify for the program.

[10] See Fleishman and Hopstock (1993) for a discussion of procedures school districts use. Approximately 89% of school districts use a home language survey in their LEP identification strategy.

questions on home language and first language status, a student's proficiency in the native language and English is assessed. Then, if a student's measured level of English proficiency is sufficiently low, the student is identified as LEP. Note that cut-off and test requirements vary by state and district (Lopez 2000, Fleishman and Hopstock 1993, Board of Education, City of New York 1994).[11] Thus, depending on available resources and state requirements, the student is recommended for an English language assistance program. Parental consent is usually obtained before a student is enrolled in any program, though not always.

Since neither data set contains information on a student's LEP status at the time of initial enrollment in the U.S. school system, I cannot exactly replicate the selection process used by states.[12] However, I can replicate the Home Language Survey portion of the LEP identification process by using information from the language portions of each survey to define an individual as "at risk" for bilingual education if he/she responded that a language other than English is spoken at home. I further limit my sample to public school students since these programs are uniquely provided at the public school level only. I call this sample the language minority sample. This definition generates a sample with 2,339 observations in the NELS:88 and a sample of 1,823 in HSB. [13]

Unfortunately, the samples I identify from both data sets may not truly represent the pool of potential participants; there may be individuals who have answered that a language other than English is spoken at home, but whose first language is English. These students would never have qualified for participation in any English language program (So 1983). Therefore any estimate of the average effect of bilingual education on the outcomes of interest using the

[11] California recently adopted a single test for all districts to use in identifying limited English proficient students, or English learners. See the California Department of Education for more information.

[12] NELS:88 does contain information on LEP status in the eighth grade, but this information is unreliable since many students identified as LEP by a teacher reported speaking only English. (McArthur 1993)

[13] Appendix Figures 1 and 2 compare the NELS:88 and HSB skip patterns for identifying English language assistance program participation with the procedures most schools use. The survey skip patterns and questions were developed by the U.S. Department of Education's Office of Bilingual Education and Minority Language Affairs (NCES 1982, NCES 1990) and closely match the procedure schools used.

language minority samples may potentially be biased downward. At the same time, students who had been chosen as survey participants, but were designated by school officials as unable to answer the NELS:88 and HSB questionnaires because of language difficulties, were excluded from the data collection. Because of the design of both surveys, estimated program effects could be biased upward since these students are likely to have been participants in any English language assistance program. In the NELS:88 a special attempt, the new student supplement, was made in the first follow-up year of 1990 to sample these base year ineligibles with abbreviated base year and first follow-up questionnaires. Only those students whose English proficiency by 1990 was sufficient to complete an English questionnaire were included in the supplement. Data on these students became available with the release of the second follow-up. The omission of students with very low levels of English proficiency could potentially bias upward the estimate of the effect of bilingual education programs. In HSB, no special attempts were made to sample limited English proficient children.[14]

Defining Program Participation

Information in the NELS:88 makes it possible to identify bilingual education program participation in two ways. First, I define a student as having received treatment in a bilingual education program if he or she had "…ever been enrolled in an English language/language assistance program, that is, a program for students whose native language is not English" or had received instruction in a native language in mathematics, science or history during the student's first two years in the U.S. school system.[15]

[14] See Lopez (1996) for a more detailed discussion of the difficulties involved in identifying the appropriate comparison group. Although these data are not ideal for an evaluation of bilingual education, they are the best available with which to evaluate the long-run impacts of this increasingly important program. Further, the wealth of family background characteristics available in each data set helps to reduce potential statistical biases. I address the issue of alternative comparison groups later in the paper.

[15] This is Question 29 of the base year student survey. I expand on this definition of bilingual education program participation for the HSB sample below.

An overall measure of bilingual education may obscure differences in effects for younger and older students. On the one hand, there may not be a difference between the comparison and treatment groups among younger students since whether or not a student is in the program, young students learn a new language rather quickly. For older students, placement in a bilingual education program could drastically improve their ability to learn English, leading to a positive effect for older, but not younger, students. On the other hand, bilingual education teachers may be able to capitalize on younger students' ability to learn a language quickly, making the program more effective for younger than for older students. For example, programs for older students may not only attempt to teach English, but also other subjects, such as math. These additional factors may make the later programs both more extensive in treatment and more difficult to implement (Harley 1988, Long 1990). Second, it is possible to identify two program types, English as a second language (ESL) and Transitional Bilingual Education (TBE), although data on native language instruction is only for a student's first two years in the U.S., potentially limiting my ability to distinguish between program types.

The HSB survey also asks several questions about the nature of a student's exposure to English language assistance programs. Students were asked if they had taken an English class for non-English speakers; if they had taken other subjects (such as mathematics and/or science) in a language other than English; if they had taken reading and writing in another language; and if they had participated in ancestral history classes or bilingual-bicultural programs. Students were further asked for their participation in these programs during first to sixth, seventh to ninth, and tenth to twelfth grades. This level of detail on program types is HSB's main advantage over NELS:88. Using these questions on program participation, I construct a single measure of bilingual education participation that equals one if the student had been exposed to an English for non-English speakers class *or* a class in another subject that was taught in the student's native language.[16]

[16] I do not use the components for "reading and writing" in another language since it is not clear if the respondent is answering about foreign language classes or some aspect of a bilingual education program (So 1983). Also, I do not use "ancestral history" classes since I am interested only in the effect of learning English, not of ancestral history classes, on the outcomes of interest.

Second, I refine my definition of English language program participation to delineate two types of programs, just as in NELS:88: ESL and TBE. I classify students as ESL participants if they only participated in an English for non-English speakers class. I classify students as having participated in a TBE program if they received other classes, such as mathematics or science, in another language. Using these definitions, 69 percent of bilingual education students participated in TBE programs, and 31 percent in ESL programs in the HSB language minority sample. These percentages are very close to national estimates of the reported distribution of participation by program type. (Hopstock and Fleishman 1993)

Sample Descriptive Statistics

Tables 1a and 1b report the means of selected demographic variables for different populations. Table 1a contains information for the sample of individuals I use from the NELS:88 and Table 1b presents information for the sample I use from HSB. I use panel weights from both the NELS:88 and HSB to calculate these statistics.

In Tables 1a and 1b column (1), I present the means for all NELS:88 and HSB participants, respectively. The second and third columns respectively compare the means for individuals whose home language is English and those for whom a language other than English is spoken in the home. Column (3) is the "at-risk" sample I define for this paper. I call this "at-risk" sample the language minority sample. I further divide this sample into participants and non-participants. Column (4) exhibits the characteristics for non-participants while column (5) presents the means for bilingual program participants. Columns (6) and (7) do the same for the sample members who participated in an ESL program and a TBE program respectively.

First I compare the characteristics of members in the language minority sample to those students who are in the English-only sample. As expected, members of the language minority sample in both the NELS:88 and HSB are more likely to be Asian, Native American, and Latino. Language minority sample members are also more likely to be first or second generation

immigrants and to come from families in lower socio-economic (SES) quartiles than the English-only sample.[17]

The comparison between participants and non-participants indicates that, in both data sets, program students are more likely to be non-white, to be recent immigrants or from families of more recent immigration, to be of lower socioeconomic status, and to have lower test scores than non-participants. This fits squarely with the reported characteristics of bilingual education participant populations elsewhere (Board of Education, New York City, 1994). In addition, participants are less likely to attain a HS diploma (77.3 percent versus 82.4 percent for non-participants) and are less likely to have enrolled in a four-year college (37.3 percent versus 32.9 percent respectively) for the NELS:88 sample. Similarly, in the HSB sample, participants are less likely to have a high school diploma (94.2 percent vs. 87.4 percent respectively) and, consequently, a bachelor's degree (29.4 percent vs. 16.6 percent respectively). Also, the average number of years of education for participants is 12.9, while that for non-participants is 13.4 years.

Table 1a also presents evidence on the difference between programs. Among participants in the NELS:88, TBE students are more likely to be in lower test and SES quartiles than ESL students. On all other characteristics and background measures, the distribution of these variables across TBE and ESL students is similar except in the probability of attaining a HS diploma and attending a four-year college. Students who participated in TBE programs are more likely to complete a HS diploma. These same students though are less likely to attend a four-year college than their ESL counterparts (conditional on obtaining a high school diploma). For the HSB data, Table 1b also shows differences in outcomes across programs. In this case, ESL

[17] I define immigrant "generation" as follows: first generation students are those students who were born abroad and whose parents were also born abroad; second generation students are those students who were born in the U.S., but whose parents (both) were born abroad; second/third generation students are those students who were born in the U.S., but for whom one parent was born in the U.S. and one was born abroad; and third generation students are those students who were born in the U.S., and whose parents (both) were born in the U.S.

students are observed to be doing better on all outcomes than their TBE counterparts, somewhat in contrast to the results of the NELS:88.

This first review of the data shows that English language assistance program participants have lower levels of educational attainment. However, given the vastly different individual characteristics of students who receive language assistance compared with those who do not, these simple comparisons between the two groups cannot be interpreted as causal.

IV. ESTIMATION STRATEGIES

In order to incorporate covariates and estimate the causal effect of bilingual education on later outcomes, I model the effect of bilingual education on an outcome (y_i^*) as:

$$y_i^* = X_i\beta_1 + JS_i\beta_2 + BED_i\gamma_0 + e_i\,, \qquad (1)$$

where y_i^* is the outcome to be considered, such as years of completed education, the natural logarithm of earnings, the propensity to attain a bachelor's degree, or the propensity to stay in high school. For the discrete outcomes,

$$y_i = 1 \text{ if } y_i^* > 0 \qquad (2)$$
$$y_i = 0 \text{ otherwise}$$

where y_i equals one if the latent variable y_i^* is greater than zero, e.g. the student has attained a HS diploma.[18] y_i^* is assumed to be a function of several observed characteristics including personal characteristics X_i, which includes measures such as race, gender, age, region of residence, urban status, immigrant generation, years since arrival, whether or not the student has

[18] This measure of HS diploma attainment does not include students who achieved a GED. See Evans and Schwab (1995) for a discussion.

had any schooling outside of the U.S., and socio-economic status as measured by parental education, family size and family income level. y_i^* is also a function of general school characteristics, JS_i, such as the pupil-teacher ratio, percent of the student body that is Latino and African American, and whether or not the school offers bilingual education programs. Finally, BED_i is an indicator that equals one for those who participated in an English language assistance program, with γ_o representing the program treatment effect. Given the sample I have constructed, this treatment effect measures the impact of English language assistance program participants relative to English immersion for a limited-English-proficient population conditional on included covariates.

Difficulties in interpreting γ_o will arise if there are omitted factors that determine program participation and educational attainment. With most education programs one might assume positive self-selection bias, leading to an overestimate of the effect of the program on the outcome. However, here it is unclear whether selection is positive or negative. On the one hand, in many bilingual programs parents must decide whether or not to place their children in a program, and in fact in some school districts a program only begins through the parents' initiative (Crawford 1995). In this case, one might expect positive selection in that children whose parents are more involved with their children's education, and therefore enroll their children in bilingual education programs, also achieve higher levels of educational attainment. On the other hand, as shown in Tables 1a and 1b, over 40 percent of English language assistance program students come from poor families and about one-third fall into the lower test quartiles on the exams administered to all NELS:88 and HSB participants. These data suggest negative selection into programs on the basis of observable characteristics. If school administrators have access to a more complete set of variables representing the students' language ability, and if students with lower values on these characteristics are more likely to drop out, then selection will most likely be negative.

I attempt to address the issue of selection three ways. First, I control for an extensive list of observed personal and school characteristics. Using background measures from NELS:88 and

HSB as proxies for the characteristics of each student at the time of entry into the U.S. school system, I estimate simple probit and OLS models for the outcomes of interest, conditional on observed background characteristics.[19] If there were no other characteristics that were correlated with bilingual education participation and the outcome of interest, γ_o in equation (1) would represent an unbiased estimate of the treatment effect of program participation. But since there may be unobservable characteristics that are correlated with program participation, and the sample I have constructed may not truly represent the set of potential bilingual education participants, γ_o may not represent the true treatment effect of bilingual education programs. Thus γ_o maybe biased up or down since selection can be either positive or negative.

As a second approach, I explore different samples of analysis to test the robustness of the single equation estimation strategy described above. The language minority sample non-participants I employ above may include some students who would not have qualified for a program because they were already fluent in English at the time of their initial enrollment in the U.S. school system despite a non-English home or first language. To investigate this concern, I estimate the program treatment effect across different sub-samples. The first set of sub-samples are constructed around an observable characteristic, and include students whose first or home languages were not English; immigrants and the children of immigrants (the group most likely to participate in the program); U.S. born students born to U.S. born parents; Hispanics (the most represented group in programs and limited English proficient populations); non-Hispanics; students from families with parents who have completed less than a HS education; and students from families with parents who have completed a HS diploma or more.

As an alternative approach to identifying an appropriate sample of analysis, I use a propensity score analysis in which I select a sub-set of the language minority sample by matching each participant to a non-participant based on a propensity score. Only those comparison observations that are sufficiently matched to a participant are included in this

[19] Although the assumption that family background characteristics have remained constant since a student's entry into the U.S. school system is strong, family background characteristics tend to be highly correlated over time. (Mayer 1997)

matched sub-sample. I call this matched sub-sample the propensity score sample. Then within the propensity score samples for HSB and NELS:88, I compare participants with non-participants to evaluate the treatment effect from my sample. This method, first proposed by Rosenbaum and Rubin (1984, 1985), has been shown to be almost as good as a randomized experiment (Dehejia and Wahba, 1998a and 1998b).

To construct the propensity score sample, I first estimate a propensity score for the probability that an individual student participated in a bilingual education program. The propensity score models bilingual education program participation, BED_i^*, and is estimated for the language minority samples described above as determined by a continuous latent variable:

$$BED_i^* = X_i\delta_1 + Z_i\delta_2 + v_i \tag{3}$$

$$BED_i = 1 \text{ if } BED_i^* > 0 \qquad BED_i = 0 \text{ otherwise.} \tag{4}$$

Ideally, the latent variable BED_i^* would be a function of two sets of variables, each measured at the time of the student's initial enrollment into the U.S. school system: a vector of covariates X_i as in equation (1), and an additional set of covariates, Z_i, that represents language characteristics of the student. BED_i^* is unobserved; instead I only observe the realization of an indicator BED_i for whether the student has actually participated in a program; if the latent variable BED_i^* is greater than zero, the student participates, i.e., $BED_i = 1$. Once the model is estimated, the predicted probability of being in a program is calculated. This predicted probability is the conditional propensity score for each student.

The propensity score serves two purposes. First, it serves as a diagnostic to measure how well a sample from a non-experimental setting has been constructed. This is achieved by comparing the estimated kernel density of propensity scores for students who participated in the program, and those who did not. If a sample is well balanced, the kernel density estimate for participants and non-participants should look similar. Plots of the propensity score for

participants and a plot of the propensity score for non-participants from the language minority sample reveals that the propensity score distributions for the participants and non-participants are very different. See Figures 3 and 4 for NELS:88 and HSB respectively. This suggests that comparisons using the language minority sample may produce biased estimates of the program treatment effect since the comparisons and participants are different on observable dimensions.[20]

The second purpose for the propensity score is to provide a metric upon which to match participants to non-participants from the language minority sample. To accomplish this, I first rank all language minority sample members on their estimated propensity score. Then, I discard those language minority sample non-participants that have a propensity score greater than the maximum score for participants, or a propensity score lower than the minimum score for participants. Finally, each participant is matched to the nearest comparison sample member, based on the estimated propensity score. This matching is done with replacement.[21] Last, all language minority sample non-participants who are not matched to a participant are discarded. This matching procedure identifies 514 non-participants in the NELS:88 language minority sample, and 651 non-participants in the HSB language minority sample. Together with the participants, these samples comprise the propensity score samples.

Kernel density plots of the propensity score distributions for both the NELS:88 and HSB propensity score sample participants and matched non-participants are shown in the bottom panels of Figures 3 and 4 respectively. Once matching is complete, the distributions of propensity scores for participants and matched non-participants look very similar, in contrast to

---

[20] Dehejia and Wahba (1998a and 1998b) suggest an algorithm for estimating the propensity score. This includes defining a parsimonious probability model to describe participation in the program, estimating that model, and then checking the full sample for variables that are unbalanced across treatments and comparisons within a propensity score strata. If any unbalanced characteristics are found, Dehija and Wahba suggest that the unbalanced variable be interacted with other balanced variables. For the propensity score model I estimate here, all background characteristics are balanced across bilingual education participants and non-participants sufficiently. Thus, I do not estimate the propensity score model with any interactions.

[21] This method is the nearest-match method with replacement suggested by Dehejia and Wahba (1998a). Reported standard errors below have been adjusted for sampling with replacement.

the propensity score distributions for the language minority sample non-participants.[22] This suggests that the propensity score samples are fairly well balanced relative to the language minority sample. Further, examination of observable characteristics in appendix Table A1 suggests that both propensity score non-participant samples are well matched as they are both very similar to the participant samples from both data sets.

Third, to account for non-random selection I model bilingual education program participation, BED_i^*, as determined by a continuous latent variable:

$$BED_i^* = X_i *1 + Z_i *2 + v_i \quad (5)$$

$$BED_i = 1 \text{ if } BED_i^* > 0 \quad (6)$$
$$BED_i = 0 \text{ otherwise.}$$

The latent variable BED_i^* is a function of two sets of variables, each measured from the time of the student's initial enrollment into the U.S. school system: a vector of covariates X_i as in equation (1); and a vector of instruments, Z_i, which represents a set of factors which affect the chances of a student being assigned to a bilingual program, but do not affect the error terms for the outcomes of interest. One candidate for BED_i^* is a student's English proficiency score from a school assessment test administered after having been identified as a potential LEP student by the Home Language Survey. However, this score is unobserved; instead I only observe the realization of an indicator BED_i for whether the student has actually participated in a bilingual program; if the latent variable BED_i^* is greater than zero, the student participates in bilingual education, i.e. $BED_i = 1$.

Potential candidates for instruments are a set of language controls for the first language and home language of the student at the time of entry into the school system as reported by the student or the parent,

[22] For the propensity score samples I have constructed, the sample means for participants and non-participants are presented in Appendix Table A1. On observed characteristics, the propensity score sample non-participants look more like the participants than the full language non-participants do. For both the NELS:88 and HSB, the propensity score sample non-participants are more likely to be of first immigrant generation, more likely to be non-white, of lower socioeconomic background, and lower test quartile than language minority sample non-participants. Similarly, the predicted propensity score for the propensity score comparison groups in both cases is closer to that of the treatment groups, although they are statistically distinguishable from each other.

and a set of controls for the legislative status of bilingual education in the student's state of residence at the time of initial enrollment.[23] Since identification as a potential LEP student is based upon the home language and first language status of a student, if some language other than English is spoken in the home, then the student is more likely to be a program participant, even within the "at risk" sample. Particularly, Spanish speakers should be most likely to participate relative to all other language groups since programs exist most commonly for Spanish speakers (Fleishman and Hopstock 1993). Therefore I use measures of first language and home language as of the eighth grade, in the case of NELS:88, or as of the tenth grade, in the case of HSB, as instrumental variables.

Cheng (1996) has suggested another set of instruments based on the timing of a student's arrival in the U.S. relative to that student's language minority classmates. He suggests using a student's years since arrival rank at a student's school as an instrument. This is calculated simply by ranking survey participants within each school by the number of years they have been in the U.S. (a low number represents a recent arrival). Depending on a student's years since arrival rank, participation in an English language program will be more likely for relatively recent arrivals than for students who have been in the U.S. educational system longer. Thus a student's rank will be negatively correlated with program participation as relatively more recent arrivals are more likely to receive treatment because of their lower levels of English proficiency. Following Cheng, I construct the rank of a student among her school classmates in time spent in the U.S. and develop four dummies indicating the student's quartile rank at her school.

First-stage results are presented in Table 7. Clearly those students whose first language and home language is a language other than English are more likely to be in a bilingual program in both the NELS:88 and HSB. These results are as expected, and mimic closely what the home language survey would do, identifying students whose home language is something other than English as potential

[23] I have tried the bilingual education legislation status of a state as an instrument in the estimation that is to follow, but found the power of the instrument to be weak, suggesting that even though a state may have legislative requirements for bilingual education programs, most of the decisions on a bilingual program are made at the district level.

participants for bilingual programs. For the years since arrival rank variable, it appears that students who arrived most recently relative to their classmates are more likely to be in bilingual programs.

I estimate an IV model that is linear probability in both stages for those outcomes that are dichotomous in both the first and second stages. [24] Under the assumption that the excluded instruments have no direct effect on the outcomes of interest, these models will produce a consistent estimate of the English language program treatment effect.

V. ESTIMATION RESULTS

A. Basic Results Using the Language minority sample

Staying in School

Basic probit estimates of the effect of English language assistance programs on obtaining a HS diploma and enrolling in a four-year college using the language minority sample are reported in Table 2; marginal effects are presented in brackets. Columns (1) to (2) and (4) to (5) of Table 2 provide evidence on the overall effect of English language assistance programs and columns (3) and (6) compare ESL and TBE programs. Without controls for other factors, participants are less likely to get a HS diploma than non-participants by 4.6 percentage points. Controlling for personal, immigrant, and school characteristics substantially reduces the magnitude of the difference to a statistically insignificant 0.6 percentage points. Clearly, differences in the observable characteristics of students account for most of the raw gap in the educational attainment of participants and non-participants.

[24] In addition, I have attempted to address selection using a Heckman (1978) two-step selection model and a bivariate probit model. I estimated a bivariate probit model for those outcomes that are dichotomous in both the first and second stages. However, even though I employed many different sets of instruments in an attempt to address selection, each set of instruments either failed to produce a consistent story with existing data, or proved to be weak predictors in the first stage. In both cases, the estimated coefficients from each IV and bivariate probit estimate were either implausibly large, or relatively close to the single equation estimates I report. These results are available upon request from the author.

As discussed above, however, the overall effect may obscure important differences between English language assistance programs. In column (3), I examine the effect of ESL programs versus TBE. It appears that participation in TBE programs is associated with a 5.4 percentage point difference in attaining a HS diploma, although the effect is not statistically significant. However, students who participated in ESL are 1.4 percentage points less to get a HS diploma relative to non-participants, although again the difference is statistically insignificant.

For enrolling in a four-year college, I observe a slightly different pattern. Basic probit estimates of the effect of program participation on enrolling in a four-year college are contained in Table 2, columns (4) to (6). As before, columns (4) and (5) provide evidence on the overall effect of English language assistance program participation; column (6) separates compares programs. Column (4) shows that, without controls for other factors, participants are less likely to enroll in a four-year college than non-participants by 5.1 percentage points. However, controlling for a similar set of characteristics as in column (2), participants are no more or less likely to enroll in a four-year college than their immersed counterparts. Furthermore, there is no difference between program treatments, as evidenced by column (6).

Thus, using the NELS:88, it appears that English language assistance program students are no more or less likely than their immersed counterparts to obtain a HS diploma or enroll in a four-year college. Further, no observable differences are obvious in comparing programs.

Long Term Effects on Educational Attainment

Turning to the HSB, Table 3 reports estimates of the impact of English language assistance programs on long-term educational attainment and earnings using the language minority sample. Columns (1) to (3) present estimates for the probability that a student has attained a high school diploma or equivalent by 1992. Marginal effects are in brackets. Column (1) of Table 3 shows that, without controls for other factors, English language assistance students are less likely to have obtained a high school diploma by 1992. However, controlling for

personal, immigrant, and the general characteristics of the student's high school in 1980 reduces the size of the program coefficient by more than a third, although it remains statistically significant. This suggests that, controlling for a host of background characteristics, the conditional probability of an English language assistance program student attaining a high school diploma is 4.1 percentage points lower than that of a students who does not receive treatment. Note that these initial results contrast with those in Table 2 using the NELS:88 that indicate that participants are just as likely to complete high school as all non-participants.[25]

Probit estimates of the effect of English language assistance programs on attaining a bachelor's degree by 1992 are presented in columns (4)-(6) of Table 3. In column (4), participants are 14.5 percentage points less likely to have attained a bachelor's degree by 1992, and the difference is statistically significant. Adding controls for personal, immigrant, and high school characteristics in column (5) again reduces the magnitude of the effect. Conditional on these covariates, the probability that English language assistance students attain a bachelor's degree is a statistically significant 11.2 percentage points lower than for non-participants.

Finally, estimates of the difference in years of education are presented in columns (7) to (9) of Table 3. Before controlling for other covariates, I find that program participant students attain 0.66 of a year less education than their non-participant counterparts. Adding controls for other characteristics reduces the effect of being in a program by almost 50 percent: participant students attain fewer years of education than their immersed counterparts controlling for background characteristics, with the statistically significant difference of approximately 0.39 of a year.

Overall, these results suggests that by age 28 participants have less educational attainment than their immersed counterparts. Even if these coefficient estimates are slightly

[25] One needs to be cautious in comparing the NELS:88 and HSB in results. Survey participants from the NELS:88 are younger than their HSB counterparts, suggesting that when the NELS:88 sample is observed at age 28, there may be similar patters of educational attainment and labor market performance as the sample from HSB, though this may not necessarily be true. With the most recent wave of the NELS:88 available, this is now a testable hypothesis, and one that I have begun to explore.

inflated, due to unobserved heterogeneity, it is unlikely that these programs improve educational attainment in the long run.

The Effect of English Language Assistance Programs on Earnings

Table 3 also reports estimates of the effect of language programs on earnings. From columns (10) to (11), the effect of an English language assistance program on log earnings is positive but always statistically indistinguishable from zero. Thus, despite attaining fewer years of education overall, the education-adjusted earnings of bilingual education participants are about equivalent to those of their immersed counterparts. However, this overall earnings effect, may obscure variation in the experiences of different ethnicities and immigrant generations. Among Hispanics, Lopez and Mora (1998) observe large differences in wages for first and second-generation students who participated in bilingual education relative to non-participants. Guzman (2002) finds similar results for Hispanic students as well. This pattern may be due to the fact that Hispanic bilingual education students possess greater amounts of labor market experience than immersed students since they tend to leave school earlier. Among other ethnic groups, and across other generations, no statistically discernable differences were observed.

B. Refining the Sample of Analysis

Results using Alternative Samples of Analysis

Since the language minority sample contains individuals who never would have qualified for bilingual education programs, I perform the above analysis for different sub-groups of the language minority sample. Examining all outcomes for both samples, the coefficient estimates using different sub-samples identified on observable characteristics are very similar to that estimated using the language minority samples, reproduced in the first row, though there are some patterns. First, for the NELS:88, while there is no statistical difference in the probability of attaining a high school diploma, non-Hispanics who have participated in a English language

assistance program are less likely, by 10.7 percentage points, to have attained a high school diploma. However, for program participants whose parents have less than a high school education are more likely to attain a high school diploma, by 11.2 percentage points, than similar students who did not participate in a program. For the HSB sample, program participants whose first language is something other than English are just as likely as non-participants to complete a high school diploma by 1992, in contrast to the negative effects observed in other populations. A similar story for attaining a bachelor's degree by 1992 is reported for program participants who come from families where the parental education level is less than high school. In this case, English language assistance program participants are no more or less likely than non-participants to attain a bachelor's degree.

With regards to earnings, shown in column (6), there is only one statistically significant coefficient, that for non-Hispanics. If fluent English speakers had seriously contaminated the non-participant sample contained in the language minority samples, I would have expected to find that the more refined sub-samples would have yielded smaller coefficient estimates. Although this exercise cannot definitively prove that the language minority sample is appropriate, the similarity in the coefficient estimates suggests that the results are fairly robust to the choice of a sub-group based on observable characteristics.

Results using Propensity Score Samples (Matched Sub-Samples)

Results using the propensity score samples are shown in Table 5 for the NELS:88 sample and Tables 6 for the HSB sample. Concentrating on the NELS:88 results in Table 5 first, using the propensity score comparison sample, there are no statistically significant differences between participants and non-participants in the observed HS diploma attainment rates and four-year college matriculation rates. All the estimated coefficients are close to zero in magnitude, and suggest that participants are actually no more or less likely to attain a HS diploma or enroll in a four year college than their immersed, but similar, counterparts. With the propensity score sample for NELS:88, there appear to be little to no impacts of English language programs in

altering the dropout rate among students at risk for participating in the program both with regression controls and without regression controls, though differences may appear as young people age.

Tables 6 presents the results using the propensity score sample from HSB. Columns (1), (3), (5), and (7) reproduce results from Table (3). The top panel of Table 6 presents linear probability results for attaining a high school diploma and a bachelor's degree, both by 1992. The bottom panel presents OLS results for years of education attained by 1992, and the level of average earnings in 1990 and 1991. In both panels of Tables 6, a pattern similar to that of Table 3 is apparent. English language assistance program students are less likely to have completed their high school diploma by 1992, less likely to have completed a bachelor's degree by 1992, and have attained fewer years of education than propensity score sample non-participants. With regard to earnings, while the estimated effect of participation in language programs is similar when comparing the language minority sample and propensity score sample results, the observed differences are still not statistically significant, implying that there is no effect of English language programs on earnings.

While the samples I have used in my analysis of Tables 2 and 3 may not truly represent all potential participants in English language assistance programs, one thing is clear. When adjusting the comparison sample employed, using either predetermined characteristics to define a sample of analysis, or attempting to match a set of non-participants from the language minority sample to participants, the basic story of Tables 2 and 3 does not change: the estimated effect programs is non-positive.

C. Instrumental Variables Results

One concern about the previous analysis is that adding covariates lowered the estimated coefficient on program participation. In each case, these patterns lead to the question of whether the estimated regression effects might actually be overstated due to unobserved factors that are correlated

both with educational attainment and participation in a program. To address this, I estimate instrumental variable models for each outcome in this paper.

IV results of the effect of English language assistance programs on attaining a HS diploma and attending a four-year college are presented in Table 8a. Columns (1) to (3) present results for attaining a HS diploma. Columns (4) to (6) represent the results for enrolling in a four-year college.

These results, no matter which combination of instruments is used, suggest that English language assistance program students and no more or less likely to enroll in a four year college or attain a HS diploma. In several cases as well, the F-statistics on the first stage instruments are low, especially for the rank instrument, suggesting that the instruments are weak predictors in the first stage.

In the upper panel of Table 8b, I report IV results for high school diploma completion by 1992 and whether or not a student attains a BA by 1992. The main result of Table 4 does not change; program students are still less likely to obtain a high school diploma by 1992 than their immersed counterparts, and are also much less likely to get their bachelor's degree, although the estimated coefficients are all not statistically significant.

In the lower panel of Table 8b, the effect of English language assistance program participation on years of education completed is more negative that the results reported in Table 3. The IV point estimates are all larger in size, than in Table 4. I also report F-statistics for the instruments in the first stage. In each case, the F-statistics are small, and raise concerns that the reported IV results are biased towards the OLS results of Table 3 (Bound, Jaeger, Baker 1995).

Considering the results for annual earnings, the IV results essentially are indistinguishable from zero. All results suggest that no difference in earnings exist in the early labor market experiences of these students, despite observing differences in educational attainment, even though the estimated coefficients are wildly different in magnitude yet positive. The observed educational attainment differences may magnify themselves later, and lead to significant differences in earnings between participants and non-participants. By the age of 28, though, we are not seeing any real differences between groups.

I report a generalized-method-of-moments (GMM) error-orthogonality test statistic for each instrumental variables estimate contained in Table 8. This statistic is calculated by obtaining the R^2 from

a regression of the residuals from the second-stage equation on the first stage instruments, and multiplying this by the number of observations used. This statistic is distributed as a χ^2 with k-q degrees of freedom where k is the number of instruments and q is the number of endogenous variables. A weak association between the instruments and the second-stage residuals will make it less likely that one rejects the null hypothesis that the equations are properly specified (Newey 1985).

The GMM statistics I report in Table 8 suggest that I reject the null hypothesis (at the 5% level of significance) that the structural equations are properly specified in all specifications using language characteristics as instruments, except for years of education attained and finishing a HS diploma by 1992 for the HSB sample. This implies that the instruments used may not be completely valid, possibly reflecting differences in program identification procedures that go beyond the home language survey and that I do not observe. These may manifest themselves in differences in the quality and types of programs students have participated in and are offered, or alternatively may represent better abilities of students to report language characteristics than their parents. The GMM results here suggest that the instrumental variables estimates be interpreted with caution since the instruments may be correlated with other unobservable factors that, in turn, may be correlated with educational attainment.

VI. CONCLUSION

Using a propensity score and IV analysis to address negative selection into ELA programs, and using a broad array of short-run and long-run educational attainment measures, as well as labor market earnings at age 28, I consistently find that bilingual education fails to improve student outcomes. In all measures of educational attainment and earnings evaluated in this paper, ELA programs appears to inhibit educational attainment, although the estimated magnitudes are not as large as one might expect given the debate about one specific program, bilingual education occurring today (Krashen 1997, Glenn 1997, Porter 1996, Chavez 1991). In fact, it is socio-economic factors, and not necessarily participation in the program, that explains a large portion of the poor performance of ELA program students.

However, my measures of English language assistance program participation do not entirely distinguish between every possible program type currently offered in the U.S. Thus, I am very broadly measuring the impact of programs one might observe in U.S. schools and a refinement of this measure may point to programs that are successful in practice and those that are not.

Even with this caveat in mind though, the findings I report in this paper support the decision California voters made in June 1998 to pass Proposition 227, the Unz initiative. The passage of that initiative has significantly reduced the prevalence of transitional bilingual education programs in California public schools. However, despite developments in California, language programs in other states designed for LEP students continue to be increasingly important and more extensive as the population of LEP students grows. Further, Hispanics, earn less than non-Hispanic whites because Hispanics obtain lower levels of human capital.

Identifying the determinants of educational attainment and English skill acquisition has become more important as Hispanics become the largest racial/ethnic minority population. On the surface, bilingual education appears to be an appropriate policy to address this relatively low human capital investment. However, evidence presented here does not support bilingual education, or other English language assistance programs, as an appropriate policy for reducing the earnings and education gap in the long-run between Hispanic and non-Hispanics specifically, or LEP individuals and English speaking monolinguals generally. This suggests that policy makers may need to reconsider the impact and organization of current bilingual education programs, especially given that language minority populations are expected to grow in the coming years.

References

American Legislative Exchange Council (ALEC), "Bilingual Education in the United States 1991-1992," Special Supplement to *The Report Card on American Education*, Washington, D.C., 1994.

Board of Education of the City of New York, "Educational Progress of students in bilingual and ESL programs: A longitudinal study, 1990-1994," 1994.

Boozer, Michael A., "Segregation, Dropouts, and School Quality," unpublished paper, Princeton University, 1992.

Bloom, David, and Gilles Grenier, "Language, Employment, and earnings in the United States: Spanish-English Differentials from 1970 to 1990," *NBER Working Paper No. 4584*, December 1993.

Carliner, Geoffrey, "Returns to Education for Blacks, Anglos, and Five Spanish Groups," *The Journal of Human Resources*, Spring, 1976, 11, 172-184.

Chapa, Jorge, "The Myth of Hispanic Progress: Trends in the Educational and Economics Attainment of Mexican Americans," *Harvard Journal of Hispanic Policy*, volume 4, 1989-1990.

Chavez, Linda, *Out of The Barrio* , Basic Books, USA, 1991.

Cheng, Chang-kuo, "The Impacts of Bilingual Education on Educational Attainment," unpublished dissertation, University of Maryland College Park, 1996.

Chiswick, Barry, and Paul Miller, "The Endogeneity between Language and earnings: International Analyses," *Journal of Labor Economics*, volume 13, no. 2, April 1995.

Crawford, James, *Bilingual Education: History Politics Theory and Practice*. Crane Publishing Company, Inc., Trenton, New Jersey, 1995.

Curiel, Herman, James A. Rosenthal, and Herbert G. Richek, "Impacts of Bilingual Education on Secondary School Grades, Attendance, Retentions, and Drop-out," *Hispanic Journal of Behavioral Sciences*, vol. 8, No. 4, pp. 357-367, 1986.

Dehejia, Rajeev H., and Sadek Wahba, "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs," *NBER Working Paper No. 6586*, June 1998.

Dehejia, Rajeev H., and Sadek Wahba, "Propensity Score Matching Methods for Non-experimental Causal Studies," unpublished manuscript, University of Toronto, 1998.

Fernandez, Roberto M., and FranHois Nielsen, "Bilingualism and Hispanic Scholastic Achievement: Some Baseline Results," *Social Science Research*, 15, 43-70, 1986.

Fleischmann, Howard L., and Paul J. Hopstock, "Descriptive Study of Services to Limited English Proficient Students, Volume 1: Summary of Findings and Conclusions," Development Associates, Inc., Arlington, Virginia, 1993.

Glenn, Charles L., "What Does the National Research Council Study Tell Us About Educating Language Minority Children?," *READ Abstracts*, 1997. The Institute for Research in English Acquisition and Development.

Grenier, Gilles, "The Effects of Language Characteristics on the Wages of Hispanic-American Males," *The Journal of Human Resources*, Winter, 1984, 19, 25-52.

Greene, Jay, “A Meta-Analysis of the Effectiveness of Bilingual Education,” The Thomas Rivera Policy Institute, March 1998.

Guzman, Joseph, “Learning English: New Evidence on the Effectiveness of Bilingual Education,” *Education Next*, Fall 2002.

Harley, Bridgit, *Age in Second Language Acquisition*, Multilingual Matters, New York, Avon Publishing, 1986.

Heckman, James J., "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica* , 1978, 46, no. 6, 931-959.

Henderson, Allison, Brenda Donly, and William Strang, "Summary of the Bilingual Education State Educational Agency Program Survey of States' Limited English Proficient Persons and Available Educational Services, 1992-1993. Final Report," Special Issues Analysis Center, Development Associates, Inc., 1994.

Kossoudji, Sherrie A., "English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men, "*Journal of Labor Economics*, 1988, 6, no. 2, 205-228.

Krashen, Stephen D., *Under Attack: The Case Against Bilingual Education*, Language Education Associates, Culver City, CA, 1996.

Long, Michael, "Maturational Constraints on Language Development," *Studies in Second Language Acquisition*, 1990, 12, no. 3, pp. 251-285.

Lopez, Mark Hugo, 1996. “The Educational and Labor Market Impacts of Bilingual Education in the Short Run and Long Run: Evidence from the National Education Longitudinal Study of 1988 and High School and Beyond.” Unpublished Ph.D. dissertation, Princeton University.

Lopez, Mark Hugo, “Evaluating California’s Proposition 227,” unpublished paper, University of Maryland and the Public Policy Institute of California, 2000.

Lopez, Mark Hugo and Marie T. Mora, “Bilingual Education and Labor Market Earnings among Hispanics: Evidence using High School and Beyond,” University of Maryland mimeo, March 1998.

Mayer, Susan, *What Money Can’t Buy*, Harvard University Press, 1997.

McArthur, Edith K., "Language Characteristics and Schooling in the United States, A Changing Picture: 1979-1989," National Center for Education Statistics, NCES 93-699, 1993.

McManus, Walter S., "Labor Market Costs of Language Disparity: An Interpretation of Hispanic Earnings Differences," *American Economic Review*, volume 75, no. 4, 1985, pp. 818-827.

McManus, Walter S., "Labor Market Effects of Language Enclaves," *The Journal of Human Resources*, Spring, 1990, 25, pp. 228-252.

McManus, Walter S., William Gould, and Finis Welch, "Earnings of Hispanic Men: The Role of English Language Proficiency," *Journal of Labor Economics*, volume 1, no. 2, 1983, pp. 101-130.

Mora, Marie T., “The Effect of English Language Assistance Programs on the English Proficiency and Academic Achievement of Hispanic Students” unpublished paper, New Mexico State University, 1998.

Mora, Marie T., "English Proficiency, Bilingual Education, and the Earnings of Hispanics," unpublished Ph.D. dissertation, Texas A&M University, 1996.

National Center for Education Statistics, *High School and Beyond*, First Follow-Up Student, School Components User's Manuals, 1982. Public Release Files.

National Center for Education Statistics, *High School and Beyond*, Fifth Follow-Up Student, and School Components User's Manuals, 1992. Private Release Files.

National Center for Education Statistics, *National Education Longitudinal Study of 1988*, Base Year Student, Parent, Teacher, and School Component User's Manuals, 1990. Public Release Files.

National Center for Education Statistics, *National Education Longitudinal Study of 1988*, First Follow-Up Student, Teacher and School Components User's Manuals, 1992. Public Release Files.

National Center for Education Statistics, *National Education Longitudinal Study of 1988*, Second Follow-Up Student, Teacher and School Components User's Manuals, 1994. Public and Private Release Files.

Neal, Derek A., and William R. Johnson, "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy*, 1996, vol. 104, no. 5.

Porter, Rosalie Pedalino, "On the State of Bilingual Education 1990-1995: *Forked Tongue* continued," *Read Perspectives*, vol. III, no. 1, spring 1996.

Reimers, Cordelia W., "Labor Market Discrimination Against Hispanic And Black Men," *Review of Economics and Statistics*, November, 1983, 65, pp. 570-79.

Rodriguez, Gregory, "The Emerging Latino Middle Class," Institute for Public Policy Report, School of Public Policy, Pepperdine University, October 1997.

Rosenbaum, Paul R., and Donald Rubin, "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score," *The American Statistician*, January, 1985, 39, pp. 33-38.

Rosenbaum, Paul R., and Donald Rubin, "Reducing Bias in Observational Studies using Subclassification on the Propensity Score," *Journal of the American Statistical Association*, September, 1984, 79, pp. 516-524.

Smith, James P., "Hispanics and the American Dream: An Analysis of Hispanic Male Labor Market Wages 1940-1980," 1990, unpublished paper, RAND Corporation.

So, Alvin Y., "The High School and Beyond Data Set: Its relevance for Bilingual Education Research," Technical report, National Center for Bilingual research, Los Alamitos, California, NCBR-TN-5, 1983.

Thomas, Wayne P., and Virginia Collier, "School Effectiveness for Language Minority Students," unpublished manuscript, National Clearing House for Bilingual Education, 1997.

Trejo, Stephen J., "Why Do Mexican Americans Earn Low Wages?," *Journal of Political Economy*, 1997, vol. 105, no. 6.

White, Halbert, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and Direct Test of Heteroskedasticity," *Econometrica*, 48, 1980, 817-838.

Willig, Ann C., "A Meta Analysis of Selected Studies on the Effectiveness of Bilingual Education," *Review of Educational Research*, 1985, 55, no. 3, 269-317.

Youngblood, Marsha, "The Cost of Bilingual Education in the U.S.: A Review of the ALEC Report," *READ Perspectives*, Fall 1995, Vol. II, No. 2, pp. 37-50.

Table 1a

Demographics for Different Populations, NELS:88 Sample

	Population						
			Language Minority Sample				
	All of NELS 88	Not At-risk Sample	All	Non-Participants	English Program Participants	ESL	TBE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
English Program Participants		0.000	0.194 (0.400)	0.000	1.000	1.000	1.000
ESL		0.000	0.061 (0.240)	0.000	0.314 (0.465)	1.000	0.000
TBE		0.000	0.128 (0.334)	0.000	0.658 (0.475)	1.000	0.000
Asian	0.035	0.011	0.144 (0.351)	0.129 (0.335)	0.209 (0.407)	0.249 (0.433)	0.141 (0.349)
Native American	0.013	0.009	0.032 (0.175)	0.031 (0.174)	0.032 (0.177)	0.034 (0.182)	0.017 (0.129)
Latino	0.104	0.024	0.478 (0.449)	0.452 (0.498)	0.589 (0.492)	0.573 (0.495)	0.651 (0.478)
Black	0.132	0.149	0.050 (0.218)	0.050 (0.217)	0.051 (0.220)	0.062 (0.243)	0.009 (0.095)
White	0.712	0.805	0.291 (0.454)	0.336 (0.472)	0.104 (0.306)	0.070 (0.256)	0.159 (0.367)
Immigrant Generation							
First	0.041	0.011	0.183 (0.387)	0.127 (0.333)	0.415 (0.493)	0.423 (0.495)	0.413 (0.494)
Second	0.038	0.006	0.171 (0.377)	0.173 (0.379)	0.163 (0.370)	0.137 (0.344)	0.207 (0.407)
Second/Third	0.041	0.026	0.109 (0.312)	0.121 (0.327)	0.060 (0.237)	0.082 (0.275)	0.017 (0.130)
Third	0.786	0.878	0.405 (0.491)	0.452 (0.498)	0.210 (0.408)	0.204 (0.404)	0.215 (0.413)
Born in US	0.881	0.922	0.713 (0.453)	0.774 (0.418)	0.457 (0.499)	0.459 (0.499)	0.449 (0.499)
Schooling outside of U.S.	0.030	0.012	0.108 (0.311)	0.066 (0.249)	0.284 (0.451)	0.296 (0.497)	0.270 (0.446)
Lowest SES Quartile	0.248	0.218	0.411 (0.492)	0.366 (0.482)	0.600 (0.491)	0.545 (0.498)	0.692 (0.463)
Second	0.248	0.252	0.227 (0.419)	0.237 (0.426)	0.184 (0.388)	0.212 (0.409)	0.143 (0.317)
Third	0.242	0.258	0.173 (0.378)	0.189 (0.389)	0.121 (0.326)	0.133 (0.340)	0.084 (0.278)
Highest SES Quartile	0.255	0.265	0.189 (0.392)	0.211 (0.408)	0.099 (0.298)	0.111 (0.315)	0.081 (0.274)
Lowest Test Quartile	0.217	0.208	0.295 (0.456)	0.277 (0.447)	0.373 (0.484)	0.358 (0.480)	0.377 (0.406)
Second	0.227	0.224	0.265 (0.441)	0.268 (0.443)	0.249 (0.433)	0.249 (0.433)	0.267 (0.445)
Third	0.229	0.237	0.184 (0.387)	0.187 (0.390)	0.172 (0.377)	0.177 (0.382)	0.151 (0.359)
Highest Test Quartile	0.238	0.249	0.181 (0.385)	0.200 (0.400)	0.105 (0.302)	0.112 (0.315)	0.100 (0.301)
Attain HS Diploma*			0.814 (0.389)	0.824 (0.381)	0.773 (0.419)	0.750 (0.433)	0.829 (0.378)

Enroll in 4 year College*			0.365 (0.482) [1980]	0.373 (0.484) [1548]	0.329 (0.470) [432]	0.350 (0.478) [289]	0.266 (0.458) [136]
No. of Obs.	16,621	13,035	2339	1825	514	345	160

Source: Author's tabulations from the National Education Longitudinal Study of 1988 (NELS:88), base year (1988), first follow-up (1990), and second follow-up (1992) surveys. The NELS:88 full sample is composed of individuals who either participated in the base year, first follow-up, second follow-up, and third follow-up surveys or completed the new student supplement survey in 1990. See text for a discussion of members of the language minority sample (which is the sample used in this study). Some categories do not sum to one because of missing observations or rounding error. All means are weighted using the second follow-up panel weight.

* Note the smaller sample size. Students had to have completed high school to be eligible to enroll in a 4-year college, which reduces the sample size for analysis.

Table 1b

Demographics for Different Populations, HSB Sample

	Population						
	All of HSB	Not At-risk Sample	Language Minority Sample				
			All	Non-Participants	English Program Participants	ESL	TBE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
English Program Participants			0.319 (0.467)	0.000	1.000	1.000	1.000
TBE			0.219 (0.414)	0.000	0.687 (0.464)	0.000	1.000
ESL			0.100 (0.300)	0.000	0.313 (0.463)	1.000	0.000
Asian Am.	0.011	0.003	0.062 (0.241)	0.052 (0.222)	0.082 (0.276)	0.103 (0.305)	0.074 (0.262)
Native Am.	0.009	0.007	0.021 (0.144)	0.020 (0.141)	0.023 (0.151)	0.013 (0.106)	0.029 (0.167)
Latino	0.125	0.091	0.371 (0.483)	0.315 (0.465)	0.491 (0.500)	0.446 (0.498)	0.511 (0.500)
Black	0.108	0.121	0.039 (0.193)	0.037 (0.188)	0.042 (0.202)	0.013 (0.115)	0.056 (0.229)
White	0.747	0.786	0.506 (0.500)	0.576 (0.494)	0.356 (0.479)	0.426 (0.496)	0.324 (0.469)
Immigrant Generation							
First	0.019	0.005	0.102 (0.302)	0.077 (0.267)	0.153 (0.360)	0.150 (0.358)	0.154 (0.261)
Second	0.041	0.024	0.127 (0.333)	0.136 (0.342)	0.109 (0.312)	0.096 (0.296)	0.115 (0.380)
Second/third	0.068	0.054	0.127 (0.333)	0.134 (0.341)	0.111 (0.315)	0.086 (0.282)	0.122 (0.321)
Third	0.854	0.903	0.510 (0.500)	0.535 (0.499)	0.456 (0.498)	0.536 (0.500)	0.420 (0.494)
Lowest SES Quartile	0.233	0.223	0.322 (0.467)	0.271 (0.446)	0.430 (0.495)	0.355 (0.480)	0.464 (0.500)
Second	0.239	0.243	0.215 (0.411)	0.226 (0.418)	0.190 (0.392)	0.186 (0.390)	0.192 (0.394)
Third	0.245	0.248	0.226 (0.418)	0.252 (0.434)	0.170 (0.376)	0.210 (0.408)	0.152 (0.360)
Highest SES Quartile	0.245	0.247	0.202 (0.402)	0.223 (0.416)	0.159 (0.366)	0.205 (0.404)	0.138 (0.345)
Lowest Test Quartile	0.207	0.204	0.243 (0.429)	0.180 (0.384)	0.379 (0.486)	0.252 (0.435)	0.437 (0.497)
Second	0.221	0.223	0.215 (0.411)	0.195 (0.396)	0.257 (0.438)	0.303 (0.461)	0.237 (0.425)
Third	0.236	0.242	0.189 (0.391)	0.215 (0.411)	0.132 (0.338)	0.157 (0.364)	0.120 (0.326)
Highest Test Quartile	0.257	0.256	0.241 (0.428)	0.306 (0.461)	0.100 (0.300)	0.176 (0.382)	0.064 (0.246)
HS Diploma by '92	0.909	0.906	0.920 (0.271)	0.942 (0.234)	0.874 (0.332)	0.924 (0.265)	0.851 (0.356)
BA by '92	0.238	0.233	0.255 (0.436)	0.294 (0.456)	0.166 (0.372)	0.237 (0.426)	0.131 (0.338)

Years of Education (st. dev.)	13.211 (1.869)	13.183 (1.858)	13.241 (1.845)	13.421 (1.903)	12.854 (1.648)	13.310 (1.833)	12.647 (1.514)
Log of Earnings (st. dev.) [sample size]	9.673 (0.853) [10,422]	9.673 (0.847) [8,147]	9.647 (0.907) [1603]	9.638 (0.933) [1051]	9.665 (0.849 [552]	9.755 (0.804) [187]	9.623 (0.867) [365]
No. of Obs.	10,862	8,278	1823	1174	649	218	431

Source: Author's tabulations from the High School and Beyond (HSB), base year (1980), first follow-up (1982), and fourth follow-up (1992) surveys. The HSB full sample is composed of sophomores who participated in all waves of the survey. See text for a discussion of members of the language minority sample (which is the sample used in this study). Some categories do not sum to one because of missing observations or rounding error. Earnings are calculated as the average of an individual's real earnings in 1990 and 1991; see text for a discussion of the construction of this variable. Log earnings are reported in this table, along with the standard deviation and sample size for the relevant population. All means are weighted using the fourth follow-up cross-sectional weight.

Table 2

The Effect of Bilingual Education on Educational Attainment, NELS:88

Probit Results Using the Language Minority Sample

Dependent Variable:	Attain High School Diploma			Enroll in a 4-year College		
	(1)	(2)	(3)	(4)	(5)	(6)
English Program Participant	-0.166 (0.126) [-0.046]	0.028 (0.124) [0.006]		-0.137 (0.104) [-0.051]	-0.075 (0.099) [-0.027]	
TBE			0.028 (0.158) [0.054]			-0.062 (0.173) [-0.022]
ESL			-0.066 (0.157) [-0.014]			-0.063 (0.108) [-0.023]
Other Controls:						
Personal Characteristics		Yes	Yes		Yes	Yes
Jr. High School Characteristics		Yes	Yes		Yes	Yes
P-value for F-test on ESL = TBE"						
Log Likelihood	-1118.2461	-920.7412	-918.4228	-1296.647	-1112.041	-1112.144

The above are probit models with standard errors in parentheses and marginal effects in brackets. All standard errors have been corrected for heteroscedasticity and sample clustering. All results are weighted using third follow-up panel weights from NELS:88. Data comes from NELS:88; there are 2,339 observations; see text for a discussion of sample members. A dummy was included if information on a student's bilingual program status was missing or indeterminate. Other Controls include controls for personal characteristics[race (Asian, Black, Latino, Native American, and other), region and urban status of the student's Jr. High School, parental education level, and family income level, immigrant characteristics (dummies for immigrant generation, whether or not the student had any schooling outside of the U.S., and years since arrival)], Jr. High School general characteristics [(ethnic group percentages, pupil/teacher ratio, percent of school which receives free lunch, percent of class which dropped out before 10th grade, and the starting salary for a new teacher with a BA].

Table 3

The Long Run Effect of Bilingual Education, HSB
Using the Language Minority Sample

Dependent Variable	Attained High School Diploma by 1992			Attained a Bachelor's Degree by 1992			Years of Education Attained by 1992			Log of Average Earnings 1990 and 1991		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
English Language Participant	-0.461	-0.363		-0.494	-0.434		-0.656	-0.394		0.024	0.093	
	(0.118)	(0.118)		(0.112)	(0.123)		(0.132)	(0.130)		(0.071)	(0.058)	
	[-0.074]	[-0.041]		[-0.145]	[-0.112]							
ESL			-0.158			-0.259			-0.126			0.157
			(0.192)			(0.178)			(0.219)			(0.075)
			[-0.017]			[-0.065]						
TBE			-0.453			-0.560			-0.534			0.057
			(0.130)			(0.144)			(0.129)			(0.070)
			[-0.056]			[-0.131]						
P-Value on: ESL = TBE												
Other Controls												
Personal Characteristics		Yes	Yes		Yes	Yes		Yes	Yes		Yes	Yes
Sr. High School general Characteristics		Yes	Yes		Yes	Yes		Yes	Yes		Yes	Yes
Log Likelihood	-493.662	-424.013	-422.529	-901.811	-738.048	-736.271						
R^2							0.0314	0.2281	0.2312	0.0002	0.3949	0.3957

The results of columns (1) to (6) are probit results with standard errors in parentheses and marginal effects in brackets; dependent variables are indicated at the top of each column. All standard errors have been corrected for heteroscedasticity and sample clustering. All results are weighted using the fourth-follow-up cross-sectional weight in HSB. See text for a discussion of the construction of the measures of bilingual education participation. A dummy has been included for students missing information on bilingual education status. Columns (7) to (12) are OLS results; dependent variables are indicated at the top of each column Other covariates include controls for personal characteristics (race/ethnicity, region and urban status of high school, family income level as eight dummy variables, parental education level, and whether or not the family owns a home in 1980, dummies for immigrant generation and years since arrival), Sr. High school general characteristics (racial composition of the high school, starting teacher's salary, per pupil district expenditures, the pupil/teacher ratio.) In the log earnings equation, additional controls include work experience (measured in months), work experience squared, marital status, most recent job tenure, and years of education completed.

Table 4

The Effect of Bilingual Education Using different Samples of Analysis

	NELS:88 Sample		HSB Sample			
Dependent Variable:	Attained High School Diploma	Enrolled in 4-year College	Attained High School Diploma by 1992	Attained a Bachelor's Degree by 1992	Years of Education Attained by 1992	Log of Average Earnings from 1990-1991
	Probit Results	Probit Results	Probit Results	Probit Results	OLS Results	OLS Results
Sample of Analysis	(1)	(2)	(3)	(4)	(5)	(6)
Full Sample	0.028 (0.124) [0.006]	-0.075 (0.069) [-0.027]	-0.363 (0.118) [-0.041]	-0.434 (0.123) [-0.112]	-0.394 (0.138)	0.093 (0.058)
First Language Something other than English	-0.063 (0.133) [-0.014]	-0.030 (0.101) [-0.011]	-0.255 (0.180) [-0.021]	-0.465 (0.177) [-0.062]	-0.344 (0.152)	0.018 (0.075)
Home Language Something other than English	-0.062 (0.137) [-0.014]	-0.007 (0.112) [-0.022]	-0.116 (0.166) [-0.012]	-0.424 (0.191) (-0.074)	-0.290 (0.164)	-0.053 (0.092)
Immigrants and Children of Immigrants	0.258 (0.161) [0.043]	-0.004 (0.119) [-0.001]	-0.515 (0.231) [-0.017]	-0.281 (0.169) [-0.084]	-0.298 (0.193)	-0.091 (0.087)
Non-Immigrants	0.028 (0.124) [0.006]	-0.079 (0.029) [-0.026]	-0.363 (0.117) [-0.091]	-0.434 (0.128) [-0.112]	-0.394 (0.122)	0.093 (0.058)
Hispanic	0.098 (0.132) [0.080]	-0.059 (0.130) [-0.019]	-0.367 (0.146) [-0.074]	-0.341 (0.167) [-0.051]	-0.261 (0.151)	0.007 (0.075)
Non-Hispanic	-0.563 (0.235) [-0.107]	-0.056 (0.151) [-0.022]	-0.325 (0.192) [-0.017]	-0.576 (0.163) [-0.165]	-0.475 (0.179)	-0.157 (0.076)
Parent Education Level less than HS	0.398 (0.149) [0.112]	-0.261 (0.176) [-0.062]	0.005 (0.249) [0.0002]	0.007 (0.294) [0.0006]	0.270 (0.205)	0.128 (0.128)
Parent Education Level HS or above	-0.249 (0.167) [-0.047]	0.037 (0.136) [0.014]	-0.514 (0.151) [-0.041]	-0.459 (0.146) [-0.135]	-0.494 (0.165)	0.121 (0.070)

All coefficients are reported English language assistance program coefficients for each sub-sample. All results are weighted using the NELS:88 third follow-up panel weight in columns (1) and (2), and the HSB fourth follow-up cross-sectional weight for all other estimates. See the text for a full list of included covariates in all of the above specifications. Results in columns (1) to (4) are probit results; marginal effects are reported in brackets. Standard errors have been corrected for heteroscedasticity.

Table 5

The Effect of English Language Programs on Educational Attainment, NELS:88

Linear Probability Results with Propensity score samples

Dependent Variable:	Attained a High School Diploma				Enrolled in a 4-year College			
	Language Minority Sample	Propensity score Sample	Language Minority Sample	Propensity score Sample	Language Minority Sample	Propensity score Sample	Language Minority Sample	Propensity score Sample
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
English Program Participant	-0.047 (0.037)	0.056 (0.048)	0.007 (0.032)	0.053 (0.032)	-0.051 (0.038)	-0.021 (0.044)	-0.023 (0.031)	0.019 (0.039)
Other Controls:								
Personal Characteristics			Yes	Yes			Yes	Yes
Jr. High School Characteristics			Yes	Yes			Yes	Yes
Sample Size	2339	1028	2339	1028	1980	841	1980	841
Log Likelihood	0.0038	0.0064	0.1671	0.2968	0.0026	0.0022	0.1750	0.2671

All models are linear probability models. Standard errors are reported in parentheses and marginal coefficients are reported in brackets. All results are weighted using the NELS:88 second follow-up panel weight. For additional covariates, see footnote to Table 2 and text. For discussion of the construction of the propensity score sample, see text.

Table 6

The Effect of Bilingual Education on Educational Attainment and Labor Market Earnings, HSB

OLS Estimates with Propensity score samples

Dependent Variable:	Attained High School Diploma by 1992				Attained a Bachelor's Degree by 1992			
	Language Minority Sample OLS	Propensity score Sample OLS	Language Minority Sample OLS	Propensity score Sample OLS	Language Minority Sample OLS	Propensity score Sample OLS	Language Minority Sample OLS	Propensity score Sample OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
English Program Participant	-0.072	-0.072	-0.052	-0.059	-0.151	-0.062	-0.101	-0.062
	(0.021)	(0.022)	(0.021)	(0.022)	(0.0031)	(0.030)	(0.020)	(0.033)
Other Controls:								
Personal Characteristics			Yes	Yes			Yes	Yes
High School Characteristics			Yes	Yes			Yes	Yes
Sample Size	1823	1300	1823	1300	1637	1138	1637	1138
R-square	0.0147	0.0151	0.0908	0.1389	0.0321	0.0197	0.2075	0.2442
Dependent Variable:	Years of Education Attained by 1992				Log of Average Earnings in 1990 and 1991			
	Language Minority Sample OLS	Propensity score Sample OLS	Language Minority Sample OLS	Propensity score Sample OLS	Language Minority Sample OLS	Propensity score Sample OLS	Language Minority Sample OLS	Propensity score Sample OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
English Program Participant	-0.656	-0.399	-0.394	-0.381	0.024	-0.044	0.093	0.029
	(0.131)	(0.162)	(0.130)	(0.129)	(0.072)	(0.081)	(0.058)	(0.069)
Other Controls:								
Personal Characteristics			Yes	Yes			Yes	Yes
High School Characteristics			Yes	Yes			Yes	Yes
Sample Size	1823	1300	1823	1300	1603	1129	1603	1129
R-square	0.0314	0.0337	0.2281	0.2312	0.0002	0.0007	0.3949	0.4312

All models are OLS models. Standard errors are in parentheses. All standard errors have been corrected for heteroscedasticity and sample clustering. See text for a discussion of the construction of the HSB propensity score sample. All results are weighted using the HSB fourth follow-up cross-sectional weight. See footnote to Table 3 for a full listing of all included covariates in the above specifications.

Table 7
Determining Bilingual Education Program Participation
First Stage Linear Probability Results (weighted)

	NELS:88 Language Minority Sample		HSB Language Minority Sample	
	(1)	(2)	(3)	(4)
Student Reported:				
First Language:				
Spanish	0.124		0.070	
	(0.035)		(0.043)	
Other	0.111		0.032	
	(0.040)		(0.043)	
Home Language				
Spanish	0.048		0.079	
	(0.026)		(0.044)	
Other	0.053		0.150	
	(0.028)		(0.045)	
Student Reported Years Since Arrival Rank:				
Second Quartile		-0.049		0.087
		(0.048)		(0.033)
Third Quartile		-0.092		0.005
		(0.054)		(0.031)
Fourth Quartile (longest time since arrival)		-0.081		0.214
		(0.052)		(0.048)
R-square	0.2374	0.2215	0.1431	0.1422

The above models are OLS models. Marginal effects are reported in brackets.

Table 8a

The Effect of Bilingual Education in the Long Run, NELS:88

Instrumental Variables Second Stage Results(Weighted)

Dependent Variable:	Attain HS Diploma			Attend a 4 year College		
	OLS Results	Second Stage Results		OLS Results	Second Stage Results	
	(1)	(2)	(3)	(4)	(5)	(6)
English Program Participant	0.007 (0.032)	0.197 (0.191)	0.566 (0.766)	0.019 (0.039)	0.154 (0.332)	-0.489 (1.101)
P^2 (dof)		10.665(4)	2.175 (2)		11.781 (4)	2.356 (2)
Instruments						
Years Since Arrival Rank			yes			yes
Language Characteristics		yes			Yes	
F-Stat on Instruments in First Stage		6.67	1.01		6.67	1.00

Above models are linear probability models in the first stage. See the text for a full listing of included covariates in all of the above specifications. For the instruments student language characteristics are given by five dummy variables describing the first and home language of the student. The critical values for a chi-square statistic at the 5% level are: 5.99 for two degrees of freedom, and 9.49 for four degrees of freedom.

Table 8b

The Effect of Bilingual Education in the Long Run, HSB

Instrumental Variables Second Stage Results(Weighted)

Dependent Variable:	HS Diploma Attained by 1992			Attained BA by 1992		
	OLS Results	Second Stage Results		OLS Results	Second Stage Results	
	(1)	(2)	(3)	(4)	(5)	(6)
English Program Participant	-0.052 (0.021)	-0.660 (0.259)	-0.178 (0.190)	-0.062 (0.033)	-0.873 (0.422)	-0.248 (0.256)
P^2 (dof)		1.034 (2)	2.586 (4)		0.62 (2)	55.8 (4)
Instruments						
Years Since Arrival Rank		yes			yes	
Language Characteristics			yes			yes
F-Stat on Instruments in First Stage		3.46	2.93		1.85	3.21

	Years of Education Attained by 1992			Log of Average Earnings from 1990-1991		
	OLS Results	Second Stage Results		OLS Results	Second Stage Results	
	(7)	(8)	(9)	(10)	(11)	(12)
English Program Participant	-0.399 (0.162)	-3.327 (1.237)	-1.248 (1.102)	0.093 (0.058)	0.841 (0.562)	0.031 (0.485)
P^2 (dof)		0.172 (2)	6.896 (4)		0.4542 (2)	15.294 (4)
Instruments						
Years Since Arrival Rank		yes			yes	
Language Characteristics			yes			yes
F-Stat on Instruments in First Stage		3.46	2.93		2.96	2.00

Above models are linear probability models in the first stage. OLS results are taken from Table 4, columns (7) and (10). See the text for a full listing of included covariates in all of the above specifications. The data set includes 2,478 observations for columns (1) to (4), and 2,195 observations for columns (5) to (8). For the instruments student language characteristics are given by five dummy variables describing the first and home language of the student. The critical values for a chi-square statistic at the 5% level are: 5.99 for two degrees of freedom, and 9.49 for four degrees of freedom.

Appendix Table A1

Demographics for Propensity Score Comparison Groups from the NELS:88 and HSB Samples

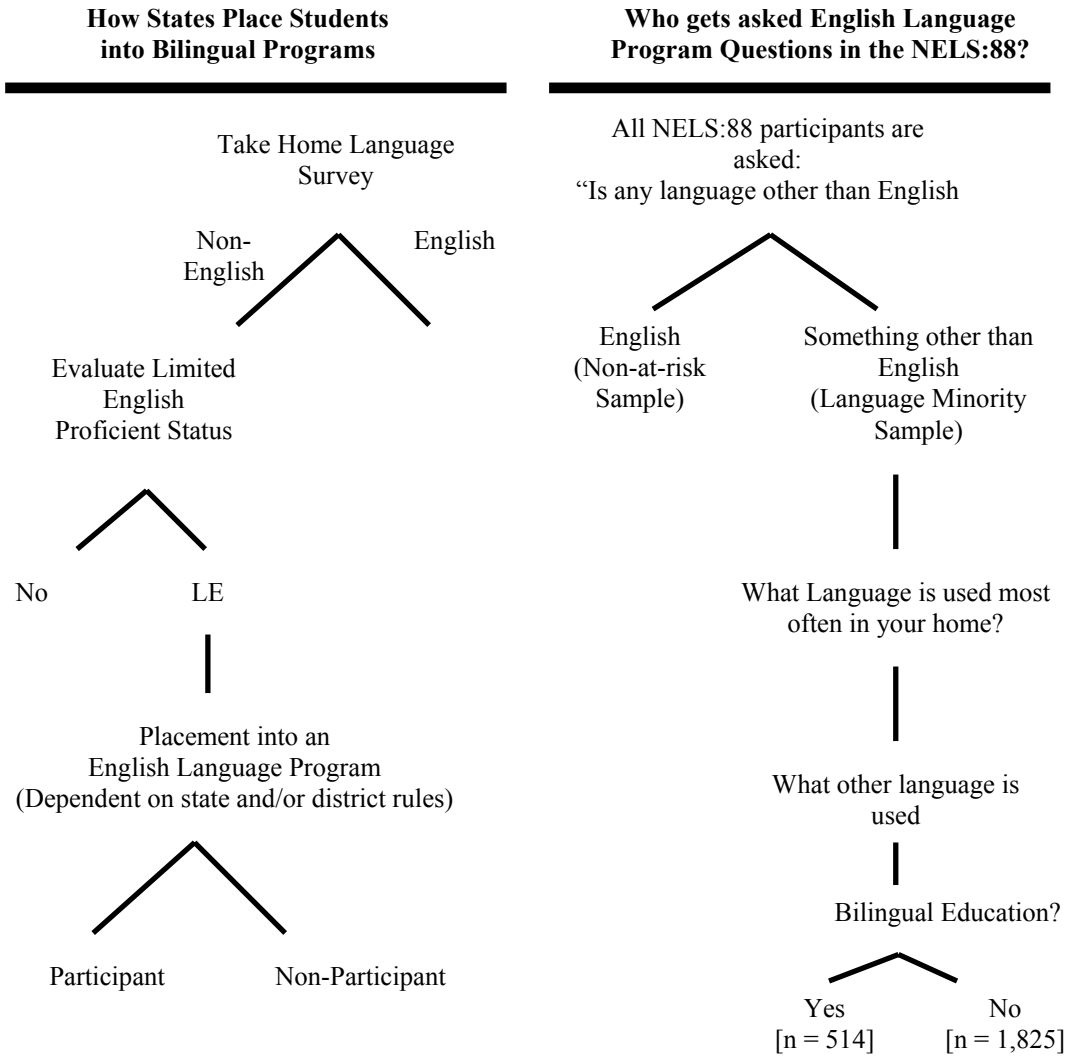
	NELS:88 Sample			HSB Sample		
	English Program Participants	Full Language Sample Non-participants	Propensity score sample Non-participants	English Program Participants	Full Language Sample Non-participants	Propensity score sample Non-participants
	(1)	(2)	(3)	(4)	(5)	(6)
English Program Participant	1.000	0.000	0.000	1.000	0.000	0.000
Asian Am.	0.209	0.129	0.225	0.082	0.052	0.082
	(0.407)	(0.335)	(0.418)	(0.276)	(0.222)	(0.274)
Native Am.	0.032	0.031	0.024	0.023	0.020	0.012
	(0.177)	(0.174)	(0.452)	(0.151)	(0.141)	(0.108)
Latino	0.589	0.452	0.647	0.491	0.315	0.430
	(0.492)	(0.498)	(0.478)	(0.500)	(0.465)	(0.496)
Black	0.051	0.050	0.020	0.042	0.037	0.041
	(0.220)	(0.217)	(0.142)	(0.202)	(0.188)	(0.199)
White	0.104	0.336	0.084	0.356	0.576	0.434
	(0.306)	(0.472)	(0.278)	(0.479)	(0.494)	(0.496)
Immigrant Generation						
First	0.415	0.127	0.353	0.153	0.077	0.095
	(0.493)	(0.333)	(0.478)	(0.360)	(0.267)	(0.293)
Second	0.163	0.173	0.194	0.109	0.136	0.137
	(0.370)	(0.379)	(0.396)	(0.312)	(0.342)	(0.344)
Second/third	0.060	0.121	0.140	0.111	0.134	0.118
	(0.237)	(0.327)	(0.318)	(0.315)	(0.341)	(0.324)
Third	0.210	0.452	0.182	0.456	0.535	0.499
	(0.408)	(0.498)	(0.386)	(0.498)	(0.499)	(0.500)
Lowest SES Quartile	0.600	0.366	0.567	0.430	0.271	0.380
	(0.491)	(0.482)	(0.496)	(0.495)	(0.446)	(0.486)
Second	0.184	0.237	0.183	0.190	0.226	0.141
	(0.388)	(0.426)	(0.387)	(0.392)	(0.418)	(0.348)
Third	0.121	0.189	0.124	0.170	0.252	0.213
	(0.326)	(0.389)	(0.329)	(0.376)	(0.434)	(0.410)
Highest SES Quartile	0.099	0.211	0.127	0.159	0.223	0.225
	(0.298)	(0.408)	(0.332)	(0.366)	(0.416)	(0.418)
Lowest Test Quartile	0.373	0.277	0.335	0.379	0.180	0.294
	(0.484)	(0.447)	(0.423)	(0.486)	(0.384)	(0.456)
Second	0.249	0.268	0.298	0.257	0.195	0.236
	(0.433)	(0.443)	(0.458)	(0.438)	(0.396)	(0.425)
Third	0.172	0.187	0.142	0.132	0.215	0.220
	(0.377)	(0.390)	(0.354)	(0.338)	(0.411)	(0.415)
Highest Test Quartile	0.105	0.200	0.114	0.100	0.306	0.109
	(0.302)	(0.400)	(0.328)	(0.300)	(0.461)	(0.312)
Attain HS Diploma	0.773	0.824	0.730	0.874	0.942	***
	(0.419)	(0.381)	(0.443)	(0.332)	(0.234)	
Attend 4 year College	0.773	0.373	0.331	0.166	0.294	***
	(0.419)	(0.484)	(0.474)	(0.372)	(0.456)	
[sample size]	[432]	[1548]	[409]			
HS Diploma by '92	***	***	***	12.854	13.421	0.932
				(1.648)	(1.903)	(0.251)

BA by '92	***	***	***	9.665 (0.849	9.638 (0.933)	0.197 (0.398)
Years of Education	***	***	***	12.854 (1.648)	13.421 (1.903)	13.066 (1.735)
Log of Earnings [sample size]	***	***	***	9.665 (0.849 [552]	9.638 (0.933) [1051]	9.400 (0.895) [577]
Estimated Propensity Score	0.396 (0.235)	0.157 (0.157)	0.369 (0.222)	0.476 (0.449)	0.249 (0.183)	0.416 (0.203)
No. of Obs.		1698	514	649	1165	651

Source: For the NELS:88 sample, all tabulations are from base year (1988), first follow-up (1990), and second follow-up (1992) surveys. See text for a discussion of the construction of the language minority sample and the propensity score comparison sample. Some categories do not sum to one because of missing observations or rounding error. All means are weighted using the second follow-up panel weight.

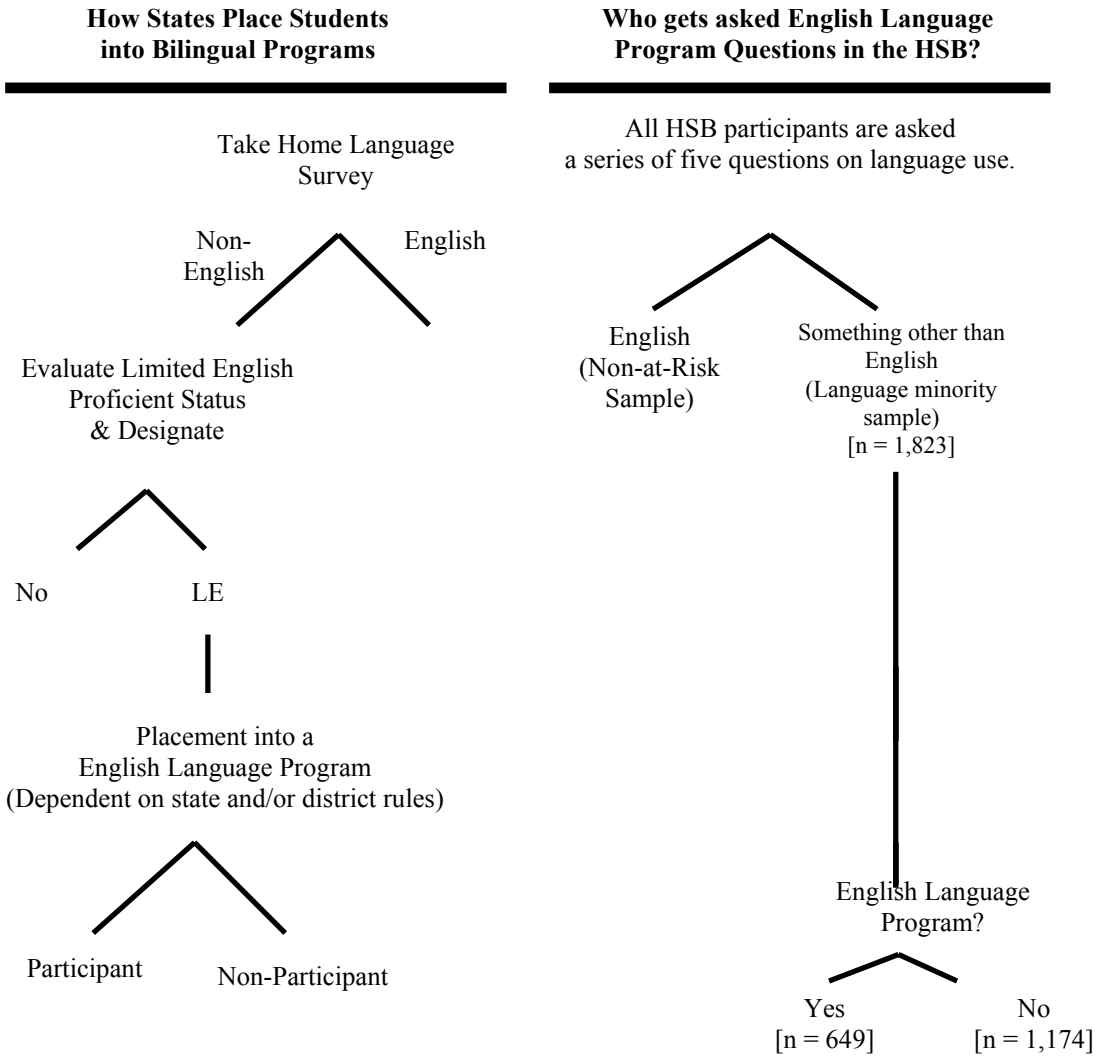
For the HSB sample, all tabulations are from the base year (1980), first follow-up (1982), and fourth follow-up (1992) surveys. The HSB full sample is composed of sophomores who participated in all waves of the survey. See text for a discussion of the construction of the language minority sample and the propensity score comparison sample. Some categories do not sum to one because of missing observations or rounding error. Earnings are calculated as the average of an individual's real earnings in 1990 and 1991. Log earnings are reported in this table, along with the standard deviation and sample size for the relevant population. All means are weighted using the fourth follow-up cross-sectional weight.

Figure 1: **NELS:88 Sample Selection**



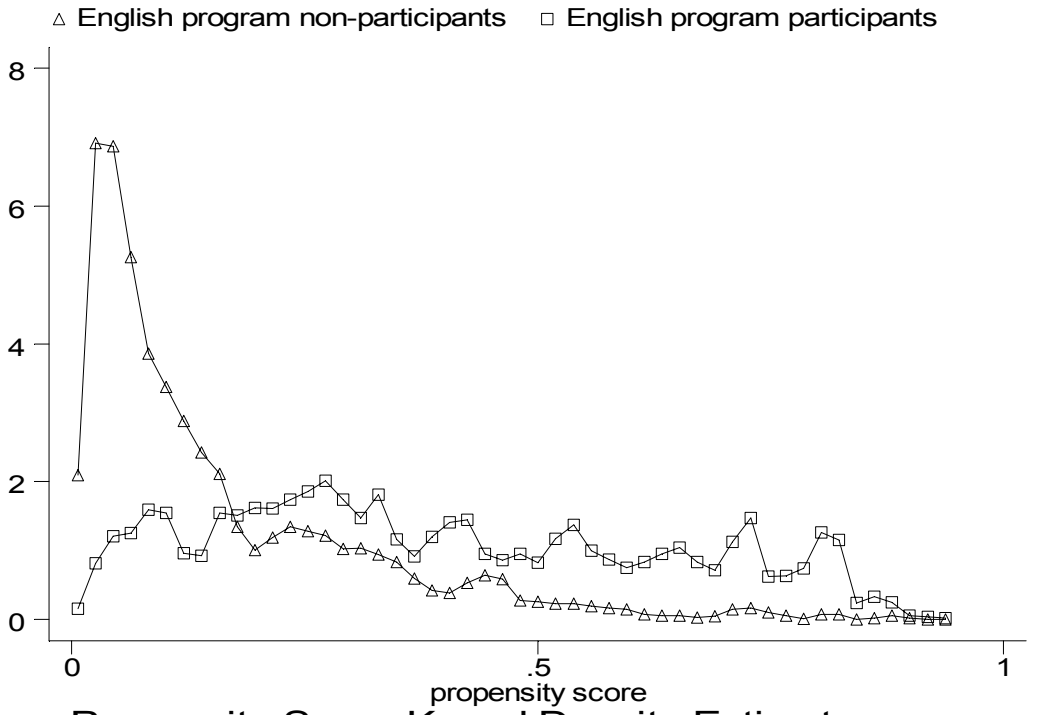
Notes: "n" refers to the sample size for each sub-population of NELS:88's base year through third follow-up participants.

Figure 2: HSB Sample Selection



Notes: "n" refers to the sample size for each sub-population of HSB's fourth follow-up

Figure 3: NELS:88 Kernel Density Estimates



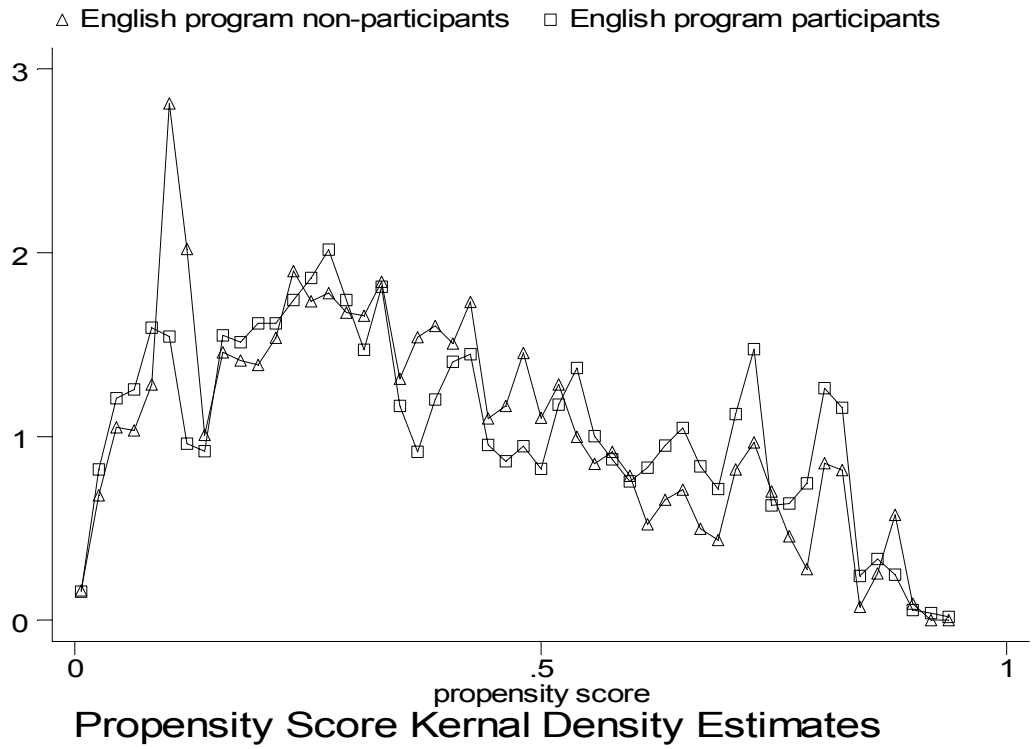


Figure 4: HSB Kernel Density Estimates