

DRAFT FOR COMMENTS

DO HIGHER STATE TEST SCORES IN TEXAS MAKE FOR BETTER HIGH SCHOOL OUTCOMES?

Martin Carnoy, Susanna Loeb, and Tiffany L. Smith

School of Education
Stanford University
Stanford, CA 94305-3096

Carnoy@leland.stanford.edu

Sloeb@leland.stanford.edu

Tlsmith@leland.stanford.edu

December, 2000

Paper presented at the American Educational Research Association Annual Meeting, New Orleans, April 24-28. We are grateful to the U.S. Department of Education, OERI, for sponsoring this research, under a grant to the Consortium for Policy Analysis in Education. We are also grateful to Sarah Turner for her helpful comments.

Making schools accountable through state testing was the pre-eminent educational reform of the 1990s. Thirty-nine states now administer some form of performance-based assessment; 24 states attach stakes to their tests; and 40 states use tests scores for school accountability purposes (Stecher and Barron, 1999). Proponents argue that using student scores on curriculum-based tests as a measure of school effectiveness encourages teachers to teach the curriculum. It sets a minimum standard on which schools can be judged; and it quantifies school “quality” in a way that parents and politicians can easily understand. By setting student improvement goals for schools, the state can motivate school personnel to reach continuously higher, while also identifying those schools unwilling or unable to meet the prescribed goals.

Critics argue that such testing does not promote “real” improvements in student learning. Rather, teachers and principals are motivated to meet “standards” by teaching the test. Instead of creating an improved learning environment, these crude forms of assessment may reduce opportunities to learn higher-order skills particularly for low-income students (McNeil and Valenzuela, 1999). Critics also claim that state testing increases disadvantaged students' probability of dropping out by forcing students to repeat grades (Haney, 1999; 2000; Shrag, 2000).

The most visible state-testing program is in Texas. The Texas Assessment of Academic Skills (TAAS) is a battery of state tests given every spring to all students in public schools in grades three to eight and again in grade ten, where passing it serves as a requirement for high school graduation. Schools are evaluated both on the percentage of all of their students passing the TAAS and on the percentage of their low-income and minority students passing. Rewards for doing well and sanctions for doing poorly are both implicit and explicit. Schools that perform well relative to state norms are given an “exemplary” designation and financial bonuses to spend on pet projects. Schools that do poorly are given an “inadequate” designation. "Inadequate"

schools get new management if they do not improve by the following year. Designations are widely publicized, so parents know how their children's school rates. Since designations take into account the proportion of disadvantaged students and the proportion of African-American and Latino students in a school, being exemplary in a poor or largely-minority school may mean a lower pass rate than in an all-white or high-income school.

The reason so many educational policymakers and politicians nationwide are looking to Texas is simple: the state has apparently achieved great success in raising average test scores and in closing the gap between disadvantaged and advantaged students, at least in the lower grades. Students in Texas have made substantial gains on the TAAS in all grades. In addition, they have made gains on an independent measure of achievement gains, the National Assessment of Educational Progress (NAEP), though these gains are not as great as those on the TAAS (Klein, Hamilton, McCaffrey, and Stecher, 2000). The NAEP results show Texas and North Carolina making the largest average gains among all states from 1990 to 1997 (Grissmer and Flanagan, 1998). When states' fourth grade students are ranked by their 1996 NAEP mathematics scores, only five states come out ahead of Texas. Significantly, all five have much lower minority populations than Texas. The gains have been smaller on the fourth grade NAEP reading test but are still higher than gains nationally.

Because of Texas' large Latino and African-American student population, educational gains in the state depend heavily on how well these minority groups do in school. The 1996 NAEP results for Texas' eighth graders and high school seniors are not as positive as those for fourth grade, mainly because the gap between minorities and whites has not decreased (and may have widened) in the higher grades (Fisher, 2000; Haney, 1999). This may be a question of time. When the 2000 NAEP results come out this year, we will gain insight into whether the fourth graders of 1996 (now the eighth graders of 2000) presage a more learned high school class of

2004 or whether Texas's accountability system is mainly good at raising minority students' elementary school NAEP scores, but not scores in middle and high school.

The effect of TAAS-type accountability on student performance in the higher grades is important. It does not make much sense to claim that student outcomes are improving if the criterion used to measure academic achievement does not result in "outcomes that count" for students' life success, such as increased school attainment or increased achievement that sustains over time. In today's world, we measure better education by problem-solving competencies, high school completion and college attendance and completion. These are the "signals" that society values.

Even if the implementation of the TASS has not lead to higher achievement or attainment for students, it may be beneficial in providing parents with information about their children's achievement and about the relative performance of the schools in their neighborhoods. However, if the TAAS leads to poorer student outcomes, then critics would have a strong case for seeking alternatives to state testing. Haney (2000) argues just that. He finds that high school completion rates have faltered as a result of the TAAS, especially among Latino and African-American students.

The object of this paper is to review Haney's findings and examine, to the extent the data allow, the impact that TAAS has had students' educational attainment. We do this in two ways. First, we assess trends over time in statewide measures of test scores, progression through high school, high school completion, and high school seniors' college plans. Second, we analyze data on high schools to estimate whether rising test scores are coincident with rising dropout rates. We find strikingly high retention in the ninth grade that has increased over time. However, this trend began well before the implementation of the TAAS. If there is a link between retention and state policies it is likely to date back to the 1984 reform. Across high schools, we find little

relationship between the TAAS and dropout rates. To the extent that this relationship does exist, it appears that higher TAAS scores in schools are associated with reduced, and not increased, dropout behavior.

BACKGROUND

Texas' current educational reform has its roots in two distinct conflicts. The first was the challenge to the unequal distribution of resources among Texas school districts, a result of the need to supplement state funding of education with local property taxes. These additional taxes created large differences among districts, differences that were highly correlated with the ethnic composition of the district. Texas' largest minorities—Latinos and African-Americans—went to schools that received considerably fewer resources than schools that were predominantly Anglo. The second conflict arose in the 1970s from a new group of businesses based on high technology and services that challenged the hegemony of Texas' traditional agricultural and oil interests.

The present educational reform began in 1984 with a push by a group of businessmen, headed by H. Ross Perot, to bring Texas into the high-tech age and to resolve pressures by low-income minority groups for equalized school funding.¹ By 1991, the reform had gone through two rounds (1984 and 1987) and was institutionalized into Texas politics. Aside from increasing funds to low spending districts, it included a strategic plan that recommended “new learning standards for each grade, measuring learning by linking state wide assessments to those standards, holding schools accountable for results, but not dictating to teachers and principals how to achieve the results” (Grissmer and Flanagan, 1998, p. 28). Ann Richards, the newly-elected governor, implemented this “decentralization” reform, that gave control to the state over

¹ The state legislature had already passed the Equal Opportunities Act in 1979 to begin to address financial inequities, but the Reform Act of 1984 was the major educational reform of the period.

standards and testing, but to schools for choosing how to meet state goals. The reform was continued after 1995 by Richards' successor, George W. Bush.

In 1990-91, the TAAS was used for the first time in the tenth grade as a requirement for high school graduation. It was applied to the present complement of grades in 1994. The TAAS was still a basic-skills test but was somewhat more difficult and more comprehensive than the earlier TEAMS test used in the 1980s and early 1990s.

The Texas Education Agency sets state curriculum standards (requiring three math courses, including Algebra I in ninth grade, for high school graduation), administers the TAAS, and determines the levels of test performance considered unacceptable, acceptable, recognized, and exemplary. Local school districts can set higher standards and are responsible for implementing improvements to bring students in their schools to satisfactory performance levels on the TAAS.

With the increased publicity given to school-by-school success, adequacy, or failure on the TAAS and the increased national attention given to Texas' success in increasing test scores, the TAAS began to take on a life of its own. Local school districts, responsible for the testing standards, saw the simplest route to getting higher pass rates was to focus on the test itself. From our observations and interviews, it appears that teachers and principals in schools with a higher percentage of lower income, African-American and Latino pupils are more likely to focus on teaching the test than those in schools with higher-income pupils.²

Achieving exemplary status is also important for high-income schools. Even in high-income schools, however, it is the lower income, lower-achieving pupils who appear to get more test preparation. This is logical. High- or even middle-achieving pupils are likely to find the test

² These schools teach both the elements of the curriculum that will appear on the test and techniques for taking the test. Our observations suggest that the majority of the time is spent on the material and not on test-taking skills.

relatively easy and have a high probability of passing. Allocating much time to teaching these students test material is unlikely to increase school ranking as measured by pass rates.

Since the test is geared to assess student performance on a prescribed curriculum, it can be convincingly argued that teaching the test to lower-achieving students is not such a bad thing. Not only that, but because of the way the test results are measured, minority scores and scores of economically-disadvantaged students carry their own weight, forcing schools to focus on these students. All of this has had the effect, testing advocates claim, of improving the reading and math of lowest-income and minority groups the most. Thus, the original goals in the 1984 Act of pushing the system up from the bottom have been maintained. Joe Johnson of the University of Texas' Dana Center claims that:

... teachers were expected to get students to demonstrate proficiency on those objectives, and so, in essence, schools are doing what they have been asked to do. They are teaching students these objectives. So, when you look at those objectives, many of them are central to what we would want anybody's children to learn about reading, mathematics, and writing. What's absolutely clear, when you visit many of the schools and look at instruction in those schools, you see that students are in fact getting opportunities to learn challenging content. They're learning the objectives that are being presented, and they're able to demonstrate their learning on the TAAS test, as well as on other measures (Interview, June 4, 1999).

Critics, however, are not convinced. Linda McNeil and Angela Valenzuela argue that,

The pressure to raise TAAS scores leads teachers to substitute commercial TAAS-prep materials for the substance of the curriculum...Subjects tested by TAAS (reading, writing, and mathematics) are reduced, in the test and test-prep materials, to isolated facts and fragments of facts. This artificial treatment of these isolated components may enable children to recognize those components on a multiple-choice test, but does not necessarily enable them to use these components in other contexts (McNeil and Valenzuela, 1999, p. 5)

OVERVIEW OF ACADEMIC OUTCOMES

Test Scores: Passing rates on the math TAAS have improved substantially in every grade, particularly for disadvantaged students. Table 1 presents the results for all students who

took the test. When special education students who took the test are omitted, the results are slightly higher. It should be noted that these figures are the percentage of students who "passed"

Table 1. Percent Meeting Minimum Expectations on Mathematics TAAS, by Grade, 1994-1999 (percent)

	<i>1994</i>	<i>1995</i>	<i>1996</i>	<i>1997</i>	<i>1998</i>	<i>1999</i>
Grade 3						
Whites	72	81	82	86	86	90
African-Americans	41	53	57	64	62	65
Latinos	49	62	65	72	71	79
Disadvantaged	47	60	63	70	68	75
Grade 6						
Whites	70	75	84	87	89	93
African-Americans	37	40	57	62	70	75
Latinos	46	47	63	69	75	81
Disadvantaged	43	46	61	67	73	80
Grade 8						
Whites	70	70	78	83	88	92
African-Americans	32	30	44	55	66	74
Latinos	40	37	51	61	71	80
Disadvantaged	37	35	49	59	69	78
Grade 10						
Whites	68	71	75	81	85	89
African-Americans	32	35	43	51	58	66
Latinos	40	42	51	57	65	73
Disadvantaged	39	40	49	55	63	71

Source: TEA (www.tea.state.tx.us/student.assessment/results)

the TAAS, not the percentage score on the TAAS. Gains on the reading test are as large or larger.

The state-level NAEP scores show improvements, though not as consistent or striking, during this time period. For example, gains in scores on the fourth grade math NAEP in Texas far outpaced math NAEP gains nationally from 1992 to 1996. The NAEP reading scores in Texas also improved and kept pace with national gains between 1994 and 1998 (Fisher, 2000). Texas eighth grade math scores on the NAEP went up about one-third of a standard deviation between 1990 and 1996. The rise in the eighth grade math TAAS between 1994 and 1998 was about one-half a standard deviation (Grissmer and Flanagan, 1998; Fisher, 2000). Although NAEP reading gains are not available for 1994-98, 76 percent of Texas eighth graders in 1998 read at basic

levels or better compared with 74 percent nationally (Fisher, 2000). On the whole, the test scores of Texas students have improved over the 1990s, reflected by TAAS scores and other measures.

An important issue, however, is whether these much higher passing rates on the TAAS (and higher NAEP scores) are reflected in other measures of academic success. Although the TAAS is more difficult than was the TEAMS, it is still a test of very basic skills. Indeed, one analysis argues that the minimum competency test given in the tenth grade is essentially an eighth grade test, and maybe not even that (see Sandra Stosky, www.taxresearch.org, 1999). For example, 57 percent of blacks, 67 percent of Hispanics, and 83 percent of whites passed the 1996 fourth grade math TAAS test. On the fourth grade math NAEP test, 47 percent of blacks, 55 percent of Hispanics, and 85 percent of whites passed at a minimum competency level.

In addition, it is not clear that the convergence in scores across racial/ethnic groups seen with the TAAS, carries through to other outcomes. Fisher (2000) and Klein, Hamilton, McCaffrey, and Stecher (2000) argue that the TAAS scores show the black-white and Latino-white gap declining on the fourth grade math and reading tests while the NAEP results suggest that the gap is staying constant or increasing. Although the results for eighth grade math are less clear because the NAEP gains are measured for 1990-1996 and the TAAS for 1994-1999, Fisher claims that these suggest that the TAAS may sharply overstate minority gains.

Enrollment Trends: Two other academic outcomes worth considering are the dropout rate and the “finishing rate,” (i.e. the proportion of ninth or tenth graders who get to twelfth grade or to graduation). Dropout rates in middle and high school (seventh to twelfth grade) as reported by the Texas Education Agency (TEA) declined for all groups from 1994 to 1999, a continued decline from even higher dropout rates in the 1980s. The average reported dropout rate for all groups went from 6.1 percent in 1989-90 (Shrag, 1999) to 2.8 percent in 1994 to 1.6

percent in 1999.³ Latinos had the highest dropout rate in both years, but that declined as well, from 4.2 percent to 2.3 percent. In the population of Texas high schools (ninth to twelfth grade), dropout rates followed a similar pattern, declining from 2.6 percent in 1993-94 to 1.3 percent in 1998-99. Rates in low-income high schools fell even further, from almost six percent in 1993-94 to 1.8 percent in 1998-99 (see Table 2, below).

Critics have expressed serious doubts about the meaning of the dropout rate figures (Haney, 2000; Shrag, 1999). Many students disappear in the accounting process. Haney (2000) argues that the reported dropout rate is not consistent with the high school finishing rate. He finds that the ratio of high school graduates to enrollment in the ninth grade three years earlier declined suddenly for white and black and Latino students simultaneously with the introduction of the tenth grade TAAS in 1990-91. He argues that the finishing rate then recovered for white students the following year, 1991-92, and rose slowly and steadily for the rest of the decade. For blacks and Latinos, Haney argues, the finishing rate also rebounded but less (Haney, 2000, Figures 5.1 and 5.2). “This indicates that the TAAS exit test has been associated with a 50 percent increase in the gap in progression from grade nine to high school graduation for nonwhite students as compared to white students” (Haney, 2000, p. 68).

Haney is not the only one who finds a disparity between the official dropout rate and student progression through high school. Colvin (1999) writes, “Statewide, one out of three white students and one out of two African American and Latino students did not graduate on time with their class in 1998. In addition, of those who graduated, only about a third of the African-American and Latino students had taken a full complement of college prep courses.”

Ninth to twelfth grade progression: We plot the three year rolling average of the ratio of high school graduates to ninth grade enrollment three years earlier in Figure 1a and the ratio of

³ See www.tea.state.tx.us/perfreport/aeis/hist/state/html

twelfth grade enrollment to ninth grade enrollment three years earlier in Figure 1b. The ratios clearly fell between the early 1980s and the early 1990s. However, our estimates suggest that the downward trend ended shortly after the implementation of the tenth grade TAAS in the early 1990s. Tenth graders in the first year of the test would not have been in twelfth grade until 1992. By 1992, almost all of the decrease in high school finishing had already occurred. The introduction of TAAS clearly did not cause these declines. (Note that the 12th grade enrollment numbers in Figure 1b appear incorrect for black students. We discuss this further below).

Ninth grade retention: The decreases that we see in high school finishing were caused either by a decrease in the relative size of the twelfth grade or an increase in the relative size of the ninth grade. Haney points to the increased retention of Texas ninth graders—especially black and Latino ninth graders—after 1985 (Haney, 2000, Figure 5.3). He estimates that the ratio of ninth to eighth graders one year earlier rose steadily for minority groups from 1985-1988 and then again from 1990-1995. The total increase in retention between 1985 and 1995, as measured by this ratio, was about 20 percent, from 1.10 to 1.30. The rise for whites in the same period was much slower but positive, from 1.05 to 1.10.

We plot three year rolling averages of the ratio of ninth grade enrollment to eighth grade enrollment one year earlier (Figure 2). Our results confirm that retention of black and Latino students has been rising rapidly since the early 1980s. The similar trends for these two groups suggest that the increased retention of Latino students is not being driven solely by in-migration during the past decade. Retention rates for whites have also increased, but at a much slower pace.

Haney, Shrag, and others postulate that higher ninth grade retention rates are closely tied to the tenth grade TAAS and result in higher high school dropout rates and lower graduation rates. There is no doubt that retention rates for minority students have risen greatly during the

reform period and have only tailed off in the past three to four years. However, there is no compelling evidence that the implementation of the tenth grade TAAS in 1990-91 was responsible for rising retention rates. Increasing standards and assessment, which began in 1984 in Texas, are associated with increased retention, particularly of blacks and Latinos, but this increase does not appear to be linked specifically with the introduction of the tenth grade TAAS.

Rising retention rates should be associated with declining finishing rates if students, retained in ninth grade, then drop out of school. This was the case in the late 1980s, for both whites and minorities, suggesting that dropout rates increased during this time. However, increased retention does not necessarily result in decreased finishing. Haney implies that the reported decline in dropout rates in the 1990s is inconsistent with rising retention rates in the early 1990s. But except for the sudden drop in finishing rates in 1990-91, there appears to be little relation between Haney's estimates of changes in ninth to eighth grade ratio and his estimates of twelfth to ninth grade ratios during this period. Retention rates rise for black, Latino, and whites students increased but finishing rates did not fall.

Other high school finishing measures: A third measure of attrition, twelfth grade enrollment compared to eighth grade enrollment four years earlier, avoids much of the problem of rising enrollment in the ninth grade year. The rolling average ratio remains virtually constant for whites at 0.8 from 1987-1998, and for Latinos at 0.7. Attrition only increases substantially for African-Americans, and that occurred in the 1980s, not in the 1990s (Figure 3).⁴ In Figure 4, a plot of the ratio of twelfth grade enrollment to tenth grade enrollment two years earlier, shows a

⁴ The rapid decline in twelfth grade enrollment relative to eighth and ninth grade enrollment four and three years earlier for African-Americans is probably an artifact of errors in enrollment statistics for that group in the 1980s. Considering the data on graduates relative to ninth and eighth grade enrollment, where the ratio is only somewhat higher for blacks than for Latinos in the 1980s, the error is probably in the twelfth grade enrollment for blacks.

decline for all groups but especially for blacks and Latinos, in the 1980s. For blacks this ratio has increased since 1993.

The bottom line is that about 30 percent of black and Latino eighth graders and about 20 percent of white eighth graders statewide do not finish with their cohort. About 15 percent of white tenth graders, about 20 percent of black tenth graders, and about 25 percent of Latino tenth graders do not reach the twelfth grade with their cohort. However, it is difficult to argue that the tenth grade TAAS, in and of itself, is to blame for either higher retention rates in ninth grade or for higher attrition rates for eighth and tenth graders.

College plans: The intention to go to college is another important measure of student success. The proportion of high school seniors who take the SAT or ACT, their scores on the SAT/ACT, and the rates at which Texas high school graduates go to college all rose sharply in the 1980s, but have slowed down in recent years. Figure 5 shows the proportion of Texas high school graduates with college plans. This percentage fell for whites and Latinos in the 1990s, but rose slightly for African-Americans. For high school graduates the gap between black/Latino and white college plans has halved, from 17 percent in 1985 to about eight percent in 1997. Overall enrollment in Texas colleges rose six percent in 1990-1996. White enrollment fell in these years, so the entire increase was due to black and Latino college enrollment increases (NCES, 1999).

Progression through high school: We summarize the enrollment, graduation, and college plans trends for the white, black and Latino students in Figures 6a, 6b and 6c. For whites the difference between eighth and ninth grade enrollment increased in the 1990s, but this did not seem to result in a greater gap between eighth and tenth grade enrollment during this time, nor in an increase in the difference between tenth and eleventh grade enrollment. The difference between twelfth grade enrollment and graduation actually decreased in the 1990s.

For African-Americans, we observe some important anomalies in the data. For example, the twelfth grade numbers appear to be wrong for 1988 and 1989. However, the data do show a clear increase in the number of black students being retained in the ninth grade through out the time period. The mid-to-late 80s and the mid-to-late 90s saw the steepest increase in the retention rate. The gap between eighth and tenth grade enrollment widened in the early to mid-1990s, indicating that many of these retained students may be dropping out of school. The difference between tenth and eleventh grade enrollment also increased. But the difference between eighth grade and graduation has remained fairly constant, in part because the gap between the graduation rate and 11th grade enrollment decreased in the 1990s. This suggests that retained students who dropped out would have been the ones most likely to drop out in the later grades if they had not been retained.

For Latinos the trends are very similar to those for blacks, except that for Latinos there is an increase in the gap between eighth grade enrollment and graduation, occurring between 1988 and 1994. The percent of high school graduates who have college plans has increased substantially for both blacks and Latinos.⁵

Summary of enrollment trends: The statewide enrollment data shows large changes in enrollment patterns over the past 15 years. In particular, ninth grade retention has increased dramatically. Moreover, the proportion of eighth graders who make it to twelfth grade four years later decreased until the early 1990s and appears to have increased somewhat in recent years. However, no obvious relationship emerges between the implementation of the high stakes TAAS and these statewide enrollment trends.

⁵ The irony, then, is that when blacks and Latinos graduate high school, they are much more likely than in the past to go to college, either two- or four-year, yet their likelihood of graduating high school with their eighth grade class or even getting a regular high school degree remains about 15 percent less than for whites—at a low 0.65, compared with almost 0.8 for whites.

Overall, the TAAS has had relatively little effect improving other high school outcomes, or, for that matter, making them worse. Yet, the TAAS may have had differential effects across schools. To explore this further we use school level data to test whether there is a relationship across schools between gains in the TAAS passing rate and other student outcomes. We now turn to these estimates.

MODEL AND DATA FOR SCHOOL-LEVEL ANALYSIS

At the individual level, educational achievement has a strong positive effect on educational attainment (Carnoy and DeAngelis, 2000). However, at the high school level, when compulsory education laws do not compel students to complete, the relationship is more complicated. Increasing average achievement in high school (as measured by TAAS passing rates) can produce increased average student attainment. But, at the same time, increased average achievement can be the result of higher dropout rates among lower achieving students. So rising average achievement in high school may result in higher completion rates, or may be the result of lower completion rates. In the first case, rising achievement is positively related to rising completion rates; in the second, it is negatively related. In the case where neither effect dominates, they may offset each other.⁶

We are interested in the relationship between TAAS scores and other student outcomes. Imagine two possible scenarios. In the first, an emphasis on increasing TAAS scores increases the overall quality of schooling, leading to gains in student learning on multiple dimensions and decreases the dropout rate. In an alternative scenario, however, increased emphasis on TAAS comes at the expense of other learning or leads to efforts to screen students *before* they take the

⁶ The same complexity pervades the relationship between rising achievement and the proportion of students taking the SAT/ACT.

TAAS. This may lead to increases in the dropout rate either as low performing students are forced out of schools in order to increase school average TAAS scores or as students choose to leave. Students may choose to leave because they fear that they will do poorly on TAAS, because they are retained in earlier grades, or because the learning environment has deteriorated.

We would like to assess whether and to what extent the TAAS has affected students' educational attainment. We are particularly interested in whether schools that have been able to increase their TAAS score have done so in conjunction with increased student attainment or have paid the price of decreased student attainment. In keeping with this goal, we will look across schools and see the relationship between the change in the TAAS score and the change in student attainment. Have schools that have experienced an increase in average score seen an increase in student attainment as well?

Unfortunately, a good measure of student attainment is not available so we have to use proxies. We considered three measures: the school reported dropout rate, the tenth to twelfth grade retention rate, and the ninth to twelfth grade retention rate. We define the tenth to twelfth grade retention rate as the number of students in the twelfth grade in one year divided by the number of students in the tenth grade two years earlier. The ninth to twelfth grade retention rate is defined similarly. In order to determine which measure to use, we ran some simple correlations. The correlation between dropout rate and tenth to twelfth grade retention is -0.23 for 1998-99 indicating that the two measures may be picking up similar trends but that, at least one of these measures, has large error. Consistent with measurement problems, differencing makes the correlations even lower (likely a result of substantial measurement error) so that the correlation between the change in tenth to twelfth grade retention and the change in the dropout rate is -0.065. However, the ninth to twelfth grade measure appears to have even greater flaws.

Surprisingly, the tenth to twelfth grade retention rate is not correlated with the ninth to twelfth grade retention rate ($r=-0.033$ for 1998-99) and the dropout rate is, in fact, positively correlated with ninth to twelfth grade retention ($r=0.18$).

One check on the usefulness of our attainment measures is to look at the correlation between the measures and the socioeconomic status of the school. In general we would expect a positive correlation between attainment and status. We find that the percent of students who are enrolled in the free lunch program is essentially equally correlated with both the dropout rate and the tenth to twelfth grade retention rate (0.16 and -0.17 respectively) while it is uncorrelated with the ninth to twelfth grade retention rate (-0.014).⁷ This is a further indication that the ninth to twelfth grade measure may not be a good measure, and so we eliminated it from our choice of attainment measures. While we look at both remaining measures, we concentrate on the dropout rate because we have it for the longest time period, 1994 to 1999. We can construct the tenth to twelfth grade retention rate only from 1996 to 1999.

Table 2 describes our sample; providing means, standard deviations, and sample sizes for the main variables that we use. It presents these descriptives for the full sample of high schools, as well as for urban, suburban, non-metro, and rural schools. In addition, we separate urban schools into those in which less than 40 percent of students enroll in free lunch programs and those in which greater than or equal to 40 percent of students enroll in free lunch programs. We see that urban schools perform considerably lower on the TAAS than other schools and that poor urban schools have the lowest average performance.

In keeping with the state trend, average TAAS scores have improved over time in all groups. We see that dropout rates are particularly high in poor urban districts though they have

⁷ The results are similar, though slightly stronger, if we look only at urban schools: the tenth to twelfth grade retention rate is not correlated with the ninth to twelfth grade retention rate ($r=-0.016$ for 1998-99); the correlation between dropout rate and tenth to twelfth grade retention is -0.28; the percent of students who are enrolled in the

Table 2: Sample Statistics for Key Analysis Variables

	Overall	Urban	Suburban	Nonmetro	Rural	Urban <40% Free Lunch	Urban >40% Free Lunch
TAAS 10th grade pass rate 1998-99	78.89 (12.02)	70.12 (12.84)	78.45 (9.82)	80.38 (10.26)	81.90 (12.70)	76.96 (10.69)	64.08 (11.52)
	<i>1134</i>	<i>175</i>	<i>208</i>	<i>416</i>	<i>335</i>	<i>82</i>	<i>93</i>
TAAS 10th grade pass rate 1993-94	55.57 (16.37)	44.26 (16.95)	57.43 (14.66)	56.90 (13.76)	58.66 (17.60)	56.42 (14.53)	33.53 (10.45)
	<i>1134</i>	<i>175</i>	<i>208</i>	<i>416</i>	<i>335</i>	<i>82</i>	<i>93</i>
Dropout rate 1998-99	1.33 (1.32)	1.66 (1.03)	1.40 (0.95)	1.39 (1.52)	1.03 (1.32)	1.50 (0.89)	1.80 (1.12)
	<i>1134</i>	<i>175</i>	<i>208</i>	<i>416</i>	<i>335</i>	<i>82</i>	<i>93</i>
Dropout rate 1993-94	2.55 (2.32)	4.61 (3.28)	2.84 (1.83)	2.53 (1.88)	1.34 (1.57)	3.25 (2.51)	5.82 (3.42)
	<i>1132</i>	<i>174</i>	<i>207</i>	<i>416</i>	<i>335</i>	<i>82</i>	<i>92</i>
Ratio 12th grade enrollment 98-99 to 10th grade 96-97	0.83 (0.12)	0.74 (0.11)	0.80 (0.10)	0.84 (0.09)	0.88 (0.14)	0.78 (0.09)	0.70 (0.11)
	<i>1134</i>	<i>175</i>	<i>208</i>	<i>416</i>	<i>335</i>	<i>82</i>	<i>93</i>
Ratio 12th grade enrollment 95-96 to 10th grade 93-94	0.83 (0.13)	0.72 (0.10)	0.81 (0.11)	0.83 (0.10)	0.89 (0.16)	0.77 (0.10)	0.68 (0.09)
	<i>1129</i>	<i>174</i>	<i>208</i>	<i>416</i>	<i>331</i>	<i>81</i>	<i>93</i>
% black enrollment 1998-99	10.68 (16.50)	22.49 (26.51)	11.46 (16.14)	8.79 (11.90)	6.39 (11.12)	18.44 (21.00)	26.06 (30.23)
	<i>1134</i>	<i>175</i>	<i>208</i>	<i>416</i>	<i>335</i>	<i>82</i>	<i>93</i>
% Hispanic enrollment 1998-99	28.75 (28.15)	47.61 (31.31)	28.37 (30.44)	24.52 (25.79)	24.40 (23.45)	29.40 (19.32)	63.66 (31.13)
	<i>1134</i>	<i>175</i>	<i>208</i>	<i>416</i>	<i>335</i>	<i>82</i>	<i>93</i>
% lunch program enrollment 1998-99	37.46 (20.78)	44.39 (22.17)	29.03 (24.52)	33.09 (17.49)	44.49 (17.70)	24.76 (9.85)	61.69 (14.07)
	<i>1134</i>	<i>175</i>	<i>208</i>	<i>416</i>	<i>335</i>	<i>82</i>	<i>93</i>
Note: Standard deviations are in parentheses and sample sizes are in italics							

lunch program is essentially equally correlated with both the dropout rate and the tenth to twelfth grade retention rate (0.28 and -0.35 respectively) and uncorrelated with the ninth to twelfth grade retention rate (0.06).

improved dramatically over this time period. As a comparison, the progression rates from 10th to 12th grade are lower and have improved the most in poor urban schools. However, the extent of improvement is small and much less than the decline in dropout rate.

REGRESSION RESULTS

Table 3 presents bivariate regression results for the relationship between the change in the TAAS score from 1994 to 1999 and the change in the dropout rate from 1994 to 1999. The coefficient measures the percentage change in the dropout rate associated with a one percent increase in the TAAS score. We find that, on average, schools in which TAAS scores increased by ten percent, dropout rates fell by 0.24 percent. This relationship is driven by poor urban schools. In urban schools in which greater than 40 percent of the students are eligible for the free or reduced price lunch program, a 10 percent increase in the TAAS is associated with an almost 12 percent decrease in the dropout rate. These results suggest that schools that saw increases in the TAAS, also saw increases in student attainment.

Table 3: Bivariate Regression Results of the Change in Log Dropout Rates on the Change in Log TAAS Score.

	Overall	Urban	Suburban	NonMetro	Rural	Urban < 40% Lunch	Urban > 40% Lunch
Coefficient	-0.024	-0.104	-0.030	-0.014	-0.010	0.437	-1.178
Std Error	0.005	0.019	0.011	0.008	0.005	0.565	0.275
R-Square	0.024	0.151	0.035	0.008	0.011	0.009	0.162

There are a number of possible explanations for these findings. It may be the case that schools with increasing TAAS scores have actually improved and thus the dropout rate has fallen. However, there are a number of other hypotheses worth checking. First, it is possible that schools that see a rise in the economic situation of their students may see a gain in both TAAS scores and other outcomes even if an effort to increase the TAAS score hurt student attainment. In other words, the fact that changing economic situations would tend to have a similar effect on both test scores and attainment may overshadow a possible negative effect that increased emphasis on the TAAS may have on student attainment. In order to take this possibility into account, we reran the above analysis but this time included controls for changes in the percent of students in the free lunch program both at the campus and district level, as well as controls for changes in the percent of black students and in the percent of Hispanic students. These results are given in Table 4a. The same trends are evident. Increases in the TAAS score are associated with decreases in the dropout rate and this is particularly true for poor urban schools. In poor urban schools a ten percent increase in the TAAS is associated with an 11 percent decrease in the dropout rate.

Second, it may be the case that students dropout before the tenth grade because they know they will need to pass the TAAS in order to graduate. In order to check whether this phenomenon is driving our results, we added a control for the change in the ninth to tenth grade retention rate. This measure is calculated as the ratio of tenth graders in 1998-99 to ninth graders in 1997-98 minus the ratio of tenth graders in 1994-95 to ninth graders in 1993-94. Table 4b gives these results. This addition makes little difference. In poor urban schools a ten percent increase in the TAAS again is associated with an 11 percent decrease in the dropout rate.

Table 4: Specification Checks

4a: Controls for Change in Free Lunch (Campus and District), in % Black, and in % Hispanic

	Overall	Urban	Suburban	Non-Metro	Rural	Urban < 40% Lunch	Urban > 40% Lunch
Coefficient	-0.398	-0.890	-0.423	-0.208	-0.237	0.699	-1.145
Std Error	0.119	0.232	0.310	0.221	0.217	0.560	0.306
R-Square	0.026	0.164	0.037	0.009	0.058	0.117	0.183

4b: Controls for Change in Free Lunch (Campus and District), in % Black, and in % Hispanic plus Change in the Retention Rate between ninth and tenth grade.

	Overall	Urban	Suburban	NonMetro	Rural	Urban < 40% Lunch	Urban > 40% Lunch
Coefficient	-0.406	-0.930	-0.449	-0.188	-0.248	0.783	-1.153
Std Error	0.120	0.234	0.324	0.223	0.217	0.538	0.307
R-Square	0.030	0.194	0.038	0.009	0.081	0.188	0.197

4c: Controls for Change in Free Lunch (Campus and District), in % Black, and in % Hispanic plus instrument for 1994-99 changes in TAAS with 1995-98 changes in TAAS

	Overall	Urban	Suburban	NonMetro	Rural	Urban < 40% Lunch	Urban > 40% Lunch
Coefficient	-0.714	-0.804	-0.312	0.041	-0.062	0.302	-0.677
Std Error	0.243	0.338	0.606	0.467	0.761	0.737	0.569
R-Square	0.022	0.193	0.037	0.005	0.077	0.176	0.175

There are two more factors worth considering. The first is effects of "regression to the mean." If we have a school that just happened to have a high performing class in 1993-94, they would have a low dropout rate and a high score on TAAS in that year. If we compare this with a normal class in 1998-99, then we would see that the TAAS scores and the dropout rate have moved together. That is, in this school the TAAS has gone down and the dropout rate has increased. Usually we count on regression to get rid of this problem because some school is likely to have had a particularly low performing class in 1993-94. However, in this case, this low performing class in another school only worsens the bias. In this second school we would see a gain over time in TAAS and a decrease in the dropout rate. The TAAS and student

attainment would again be moving together and we would overestimate the relationship between attainment and test score. Regression to the mean, then, leads to an upward bias in our estimates. Luckily, it is not difficult to adjust for this problem. We instrument for the change in the TAAS score from 1993-94 to 1998-99 with the change in the TAAS score from 1994-95 to 1997-98. This process essentially, uses only the variation across schools in the change in test score that is consistent between the two periods, and thus gets rid of the change due to particularly good or bad cohorts.

The same technique also helps alleviate our second concern. This second concern is measurement error. The test score is likely to be an imperfect measure of any class's ability on the test. This imperfection will lead to a bias toward zero in the estimated relationship between the dropout rate and the test score. Instrumenting reduces this bias by using variation in the test score change that is consistent over the two time periods. Table 4c presents the instrumented results that adjust for regression to the mean and measurement error. We find a slightly smaller effect in poor urban areas, but still, a ten percent increase in TAAS is associated with a seven percent decrease in the dropout rate both overall and in poor urban schools.

While we believe these results are informative, it is also useful to look at alternative outcome measures. Tables 5 and 6 gives the results using changes in the tenth to twelfth grade retention rate, changes in the ninth to tenth grade retention rate, changes in the percent of students taking the SATs, and changes in mean SAT scores (controlling for changes in the percent taking). The results for the alternative outcomes do not show a consistent relationship between TAAS scores and other student outcome measures. We do find a positive and statistically significant relationship between change in TAAS score and change in the tenth to twelfth grade retention rate for the full sample in the bivariate regression. The point estimate for poor urban schools indicates a stronger effect, consistent with the trends for the dropout rate

observed above; however this results is not statistically significant at conventional levels. We found no other statistically significant effects in the simple regressions. We also ran full regressions which instrumented for the change in TAAS score from 1994 to 1999 with the change from 1995 to 1998 and included controls for changes in district free lunch enrollment, campus free lunch enrollment, percent of Hispanic students, and percent of black students. These results are summarized by Table 6. None of the estimates is significant at conventional levels.

Table 5: Bivariate Regression Results for Alternative Outcome Measures

Full Sample

	tenth to twelfth retention	ninth to tenth retention	Percent taking SAT*	Mean SAT*
Coefficient	0.00078	-0.0056	0.029	-0.213
Std Error	0.00036	0.0215	0.036	0.165
R-Square	0.0042	0.0001	0.0054	0.0064

Urban Schools with > 40% of Students enrolled in Free Lunch Program

	tenth to twelfth retention	ninth to tenth retention	Percent taking SAT	Mean SAT
Coefficient	0.00171	0.00087	-0.00141	-0.1126
Std Error	0.00161	0.00233	0.1464	0.3043
R-Square	0.0112	0.0014	0.0222	0.0310

Percent taking the SAT includes additional control for the change in the dropout rate and Mean SAT includes additional controls for percent change in the dropout rate and change in percent taking the SAT.

Table 6: Full Regression Results for Alternative Outcome Measures

Full Sample

	tenth to twelfth retention	ninth to tenth retention	Percent taking SAT*	Mean SAT*
Coefficient	.0010535	.0127075	.0501501	-.2442977
Std Error	.0010098	.0607834	.0975375	.164205
R-Square	0.0100	0.0038	0.0059	0.0289

Urban Schools with > 40% of Students enrolled in Free Lunch Program

	tenth to twelfth retention	ninth to tenth retention	Percent taking SAT	Mean SAT
Coefficient	0.00096	-0.0060	-0.0143	-0.178
Std Error	0.00191	0.0045	0.3091	0.912
R-Square	0.018	0.137	0.084	0.115

Percent taking the SAT includes additional control for the change in the dropout rate and Mean SAT includes additional controls for both percent change in the dropout rate and change in percent taking the SAT. Mean SAT analysis for the first sample and tenth to twelfth retention for the urban sample include controls but are not IV, because of convergence difficulties.

DISCUSSION

The explicit objective of Texas' school accountability system is to improve educational outcomes among the state's four million public elementary and secondary students. Rising pass rates on the TAAS suggest that this objective is being met. Texas' students have also made substantial gains on some of the NAEP tests, specifically in fourth and eighth grade math. This would appear to validate claims that rising pass rates on the TAAS may, indeed, indicate real learning gains.

Nevertheless, as we have shown, other important indicators of educational success, namely high school progression and graduation rates at best rose slightly in the 1990s, and then only in the past few years. It appears that rising TAAS scores on the tenth-grade high-stakes test have had at best a small impact on educational outcomes that count, namely high school completion and the likelihood of attending college. This is particularly troubling because high school graduation rates are relatively low in Texas. Only about 65 percent of African-American and Latino eighth graders and about 78 percent of white eighth graders graduate four years later.

Our results suggest that claims of rapidly increased retention of blacks and Latinos in the ninth grade since the implementation of the Texas education reform in 1984 are absolutely true. However, claims that increasing retention is directly attributable to the new tenth grade TAAS exam that students must pass to graduate is not supported by the data. Statewide enrollment grade-by-grade also do not support claims that statewide dropout rates increased after 1990.

We have shown, using school level data, that there is a positive relation between lower dropout rates and higher TAAS scores on the tenth grade TAAS in the 1990s. That is, those high schools that have had larger increases in their students' pass rates on the tenth grade TAAS also had have larger declines in dropout rates. The relationship is largest among urban high schools serving low-income students, and continues to be significant and large even when we correct for possible sources of bias in the coefficient of test score increase on dropout rate. However, our results also show that there is a much weaker relationship between increases in the tenth grade TAAS pass rate and the proportion of tenth graders reaching twelfth grade two years later. The corresponding relationship for urban high schools serving low- income students is positive but not significant. We also do not find a relationship between increasing TAAS score on the tenth grade test and either the proportion of SAT/ACT test takers or the high school's average SAT score.

There are a number of possible explanations for the weakness of the relationship we find between TAAS scores and other outcomes. First, the data simply may be too inaccurate to capture any effects that exist. The low correlation among our various high school progression and completion measures raises serious concern about the estimation of dropout rates by high schools. Much of our analysis depends on the reliability of our outcome measures. The data on high school enrollment comes directly from the Texas Education Agency, but there are evident anomalies in both the aggregate data and the data on individual schools.

Yet, let us assume that these data anomalies are such that our results remain valid: rising TAAS pass rates have not had much impact on educational attainment. The difficulty of the tenth grade TAAS may simply have been set at a level that would not substantially impact the prevailing dropout rate. The proportion of students statewide who cannot pass it (after repeated tries) may be synonymous with the proportion that would have dropped out anyway. This is not

an unlikely scenario, given the political sensitivity of high school completion rates. Another possibility is that not enough time has passed to observe consequences of the TAAS. If we are currently viewing the full effect the test, however, it appears that it has not been successful at improving educational outcomes with high social value, such as high school completion or college attendance.

Enrollment data for recent years hint at a possible turning point in high school progression trends. Ninth grade retention has leveled off in the past three years and graduation rates relative to ninth and eighth grade enrollment have increased. That could mean that high retention rates in ninth grade and high attrition rates in eighth to twelfth grade are relatively “short term” investments in raising the “quality” of Texas’ high school graduates. However, retention and attrition are still very high, particularly for minority groups. If they remain so, critics may be right that Texas accountability system is not an effective way to improve minority students’ opportunities to learn.

REFERENCES

- Stecher, Brian and Sheila Barron. 1999. Test-Based Accountability: The Perverse Consequences of Milepost Testing. CRESST and RAND Education. Paper presented at the annual meeting of the American Educational Research Association, April 21, 1999, Montreal, Canada.
- Carnoy, Martin and Karen DeAngelis. 2000. "Does "Ability Influence Individual Earnings, and If So, by How Much?" Stanford University School of Education (mimeo).
- Colvin, Richard Lee. 1999. "Texas Schools Gain Notice and Skepticism." *Los Angeles Times*, Part A Section, July 6, Home Edition.
- Fisher, Flo. 2000. "Tall Tales? Texas Testing Moves from the Pecos to Wobegon." No location (mimeo).
- Grissmer, David and Ann Flanagan. 1998. "Exploring Rapid Achievement Gains in North Carolina and Texas." Washington, D.C.: National Education Goals Panel.
- Haney, Walter. 1999. "Report for Testimony in *GI forum v. Texas Education Agency*." Boston College, School of Education (mimeo).
- Klein, Stephen P., Laura S. Hamilton, Daniel F. McCaffrey, Brian M. Stecher 2000. " What Do Test Scores in Texas Tell Us?" RAND Issue Paper.
- McNeil, Linda and Angela Valenzuela. 1999. "The Harmful Impact of the TAAS System of Testing in Texas: Beneath the Accountability Rhetoric." Houston, Texas: Rice University, Department of Education (mimeo).
- Shrag, Peter. 2000. "Too Good to Be True." *The American Prospect*, Vol. 11, no. 4 (January 3).

Figure 1a. Texas: High School Graduates Compared to 9th Grade Three Years Earlier, Three-Year Rolling Averages, by Race/Ethnicity, 1983-1997

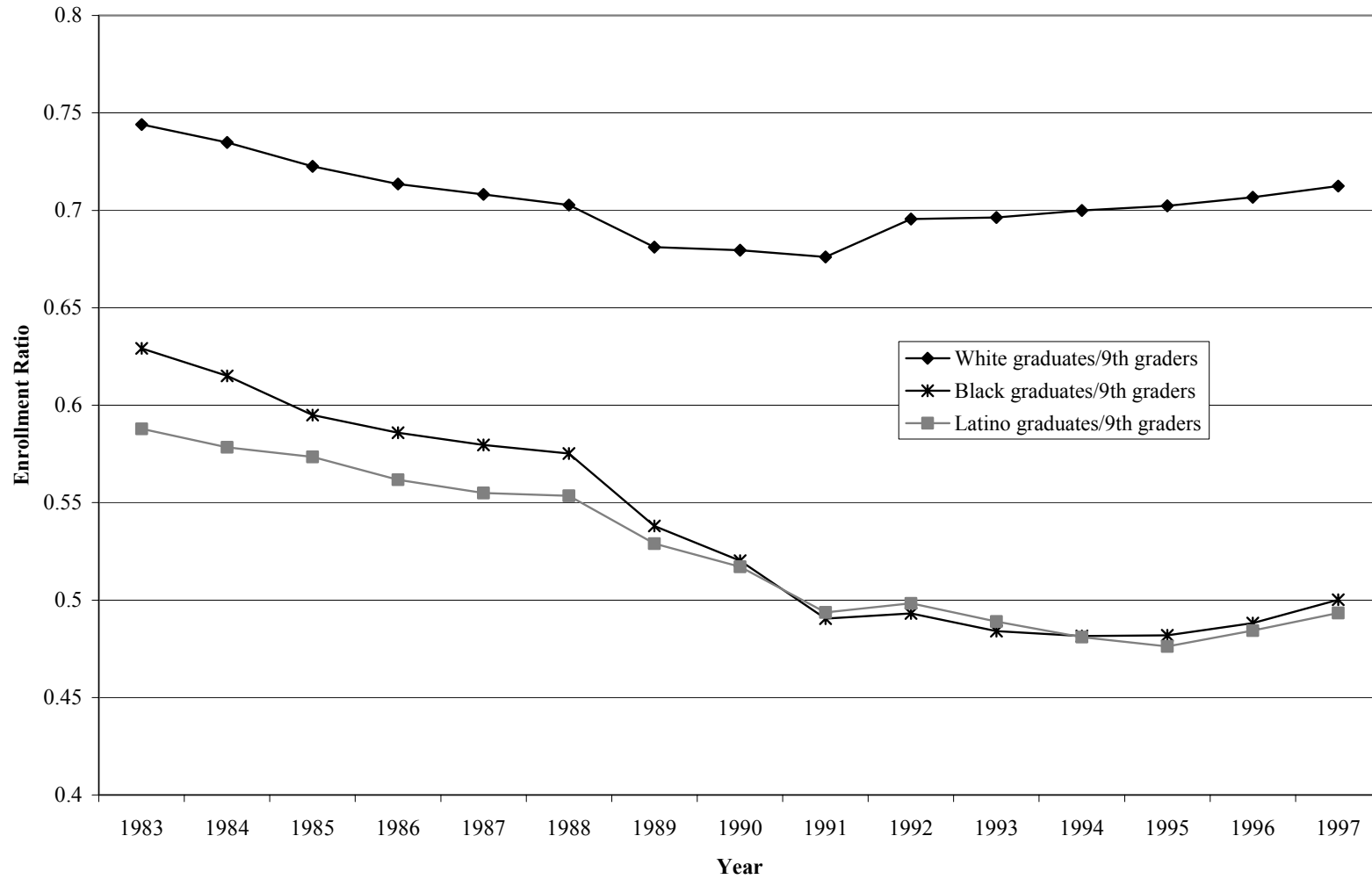


Figure 1b. Texas: 12th Grade Enrollment Compared to 9th Grade Three Years Earlier, Three-Year Rolling Averages, by Race/Ethnicity, 1983-1997

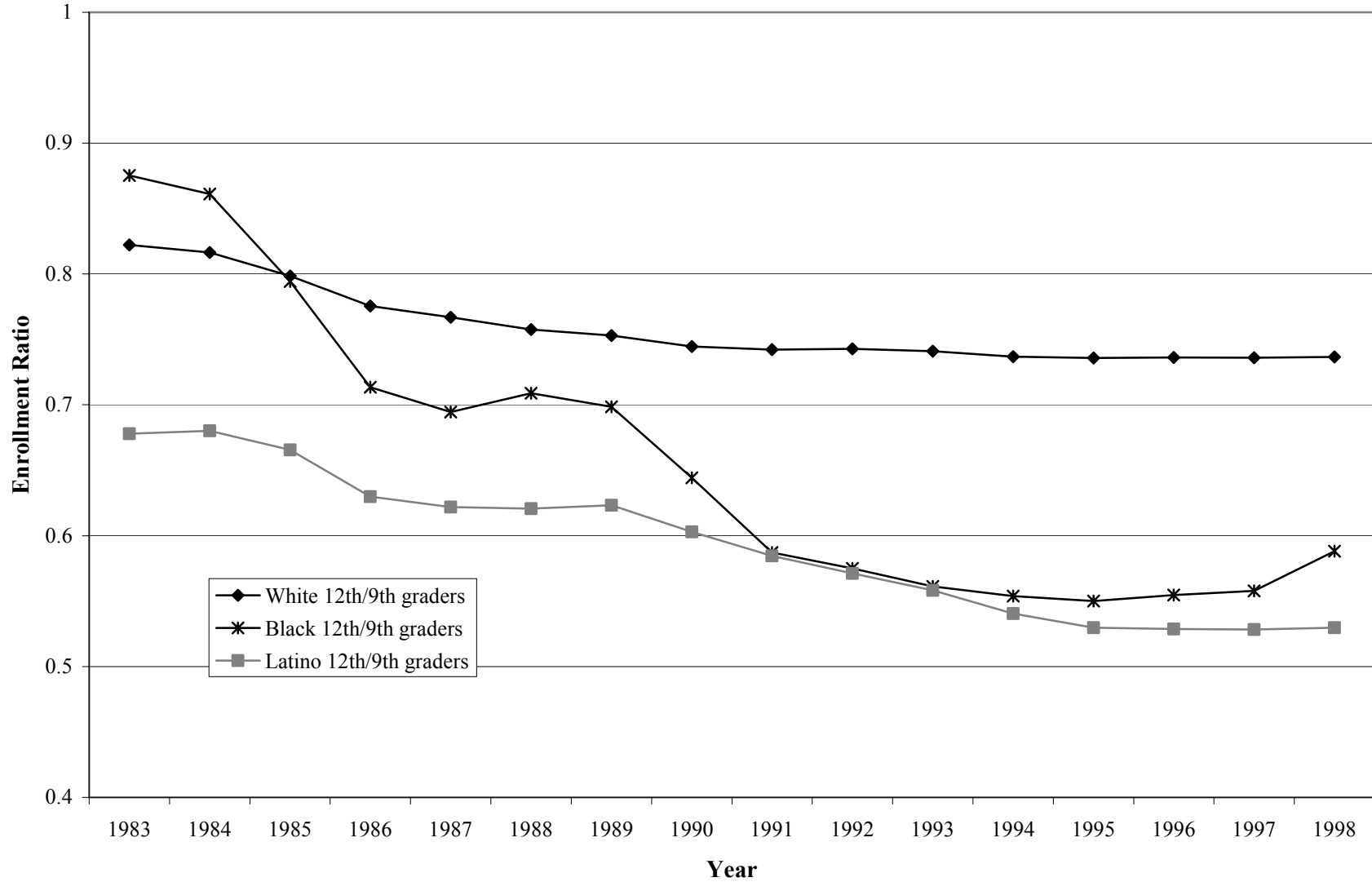


Figure 2. Texas: 9th Grade Enrollment Compared to 8th Grade Enrollment, Three-Year Rolling Averages, by Ethnicity, 1981-1998

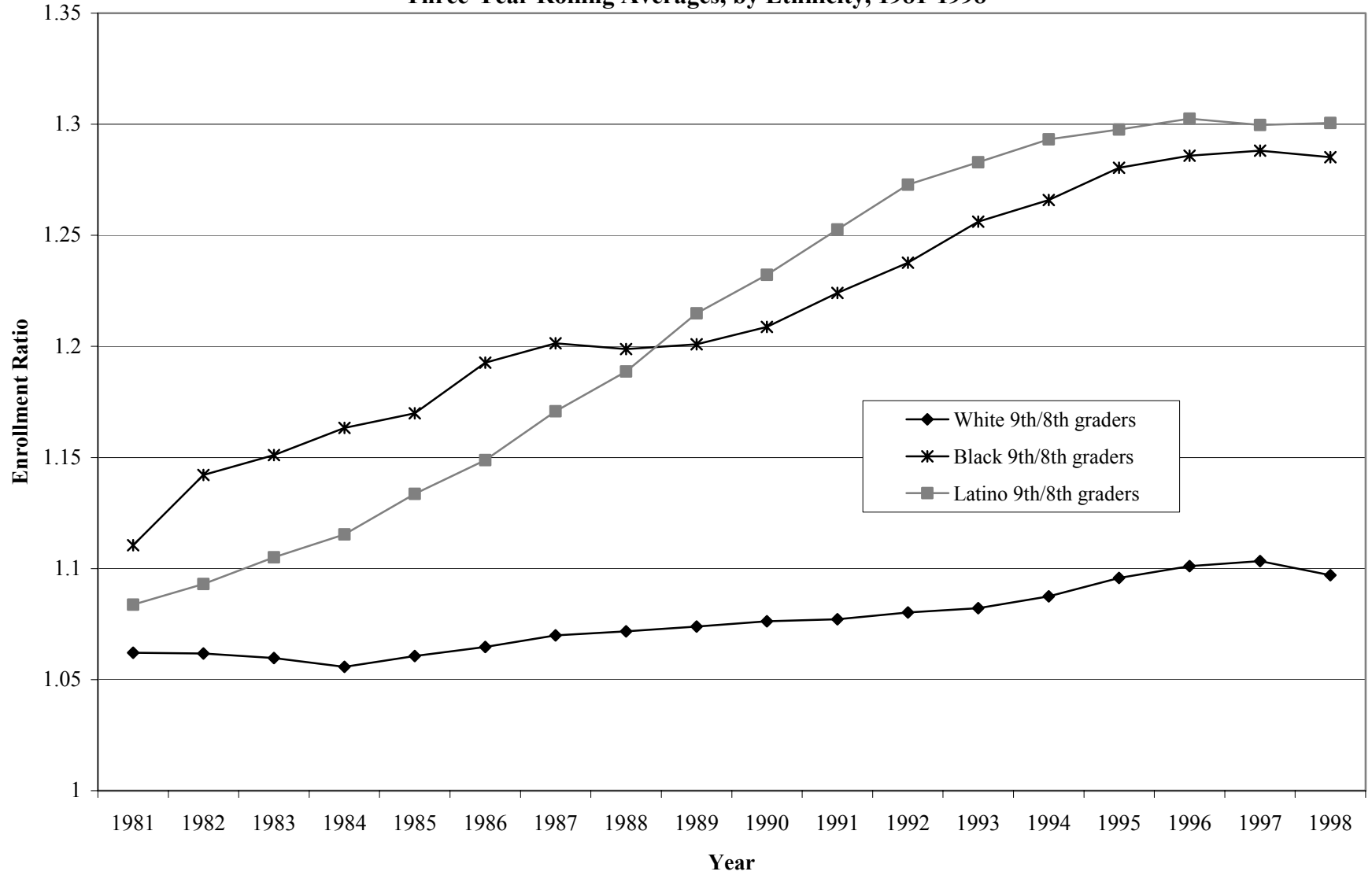


Figure 3. Texas: 12th Grade Enrollment Compared to 8th Grade Four Years Earlier, Three-Year Rolling Averages, by Race/Ethnicity, 1983-1997

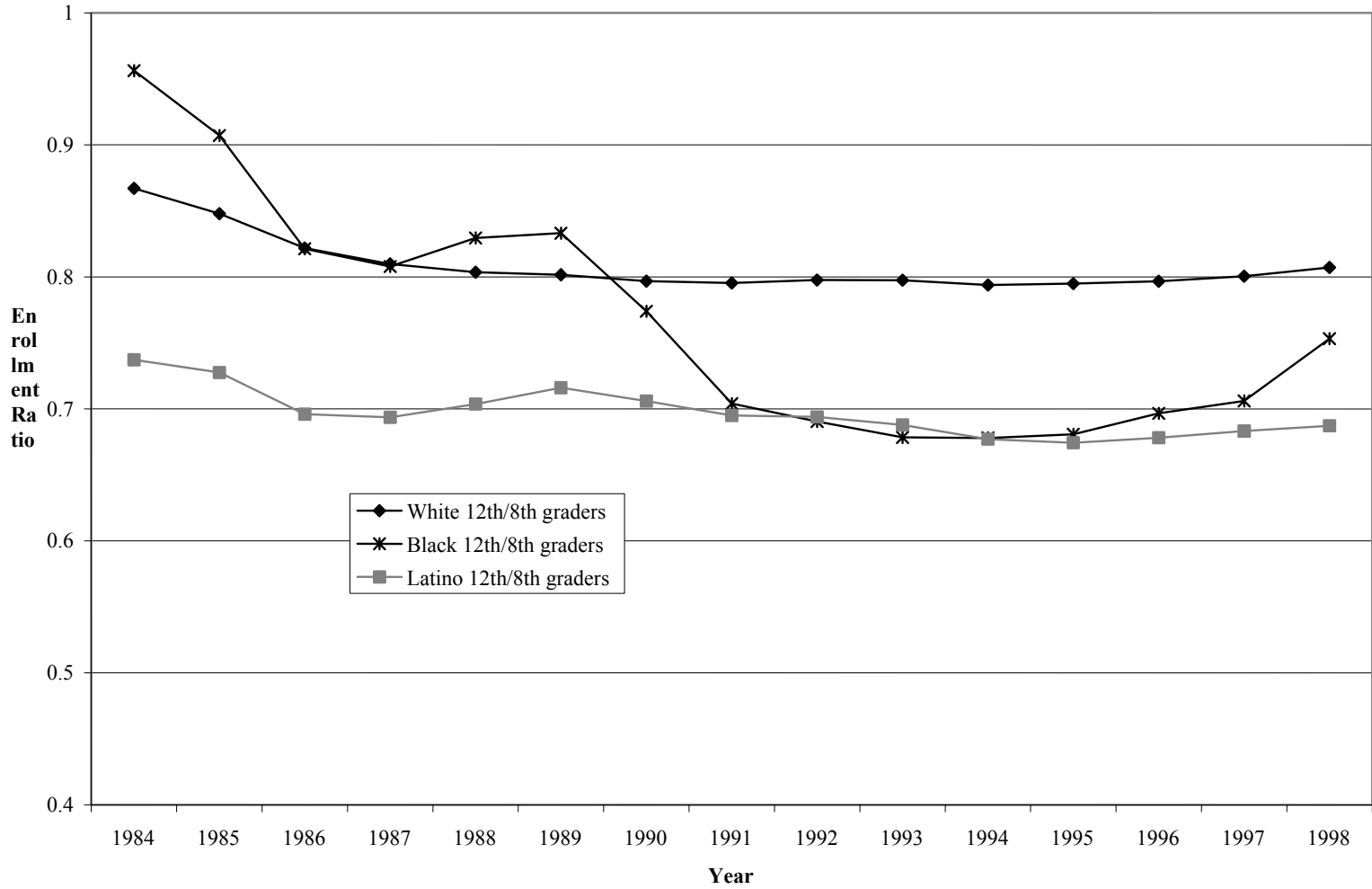


Figure 4. Texas: 12th Grade Enrollment Compared to 10th Grade Two Years Earlier, Three-Year Rolling Averages, by Race/Ethnicity, 1983-1997

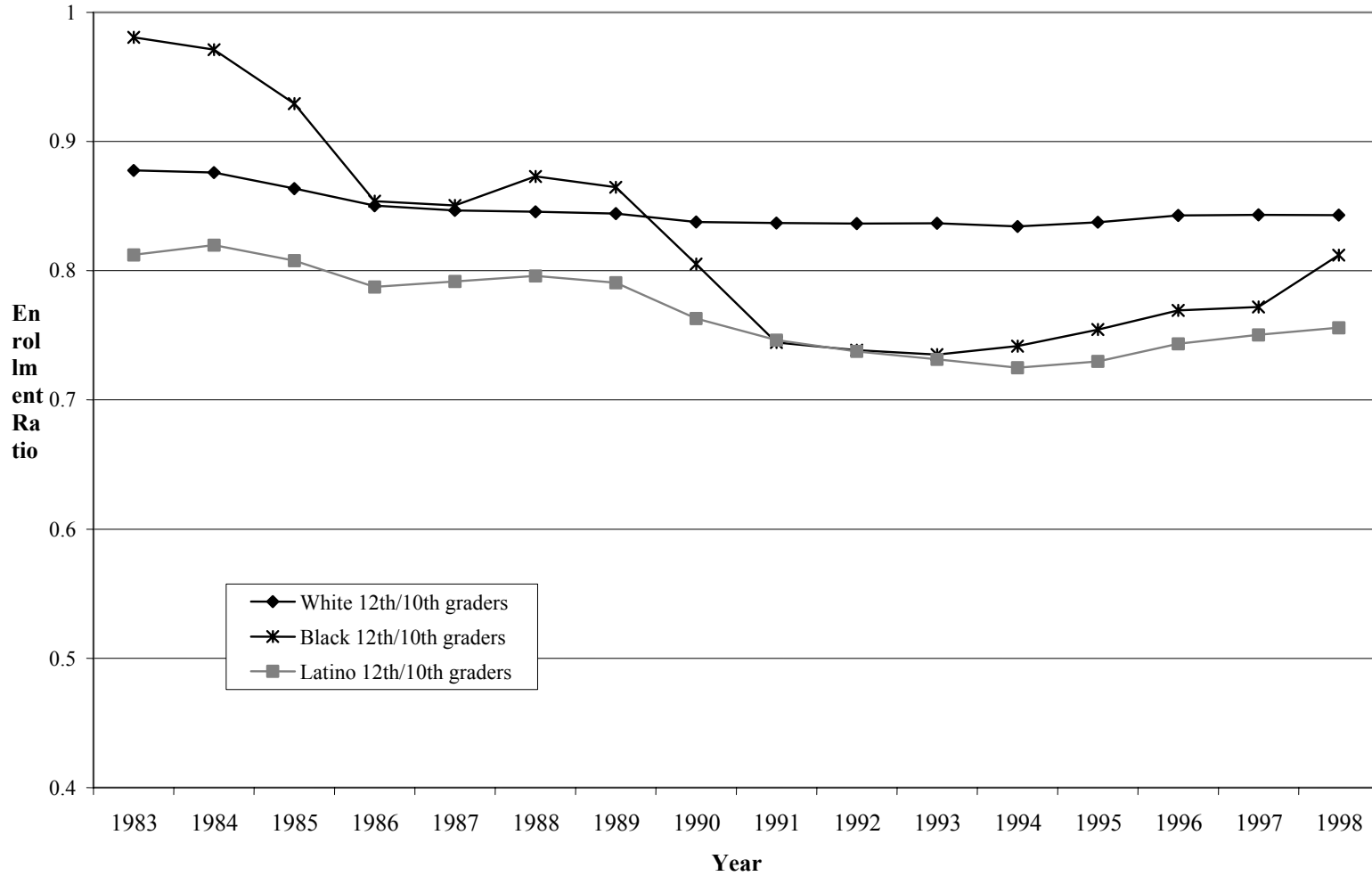


Figure 5. Texas: Proportion of High School Graduates with College Plans, by Race/Ethnicity, 1979-1998

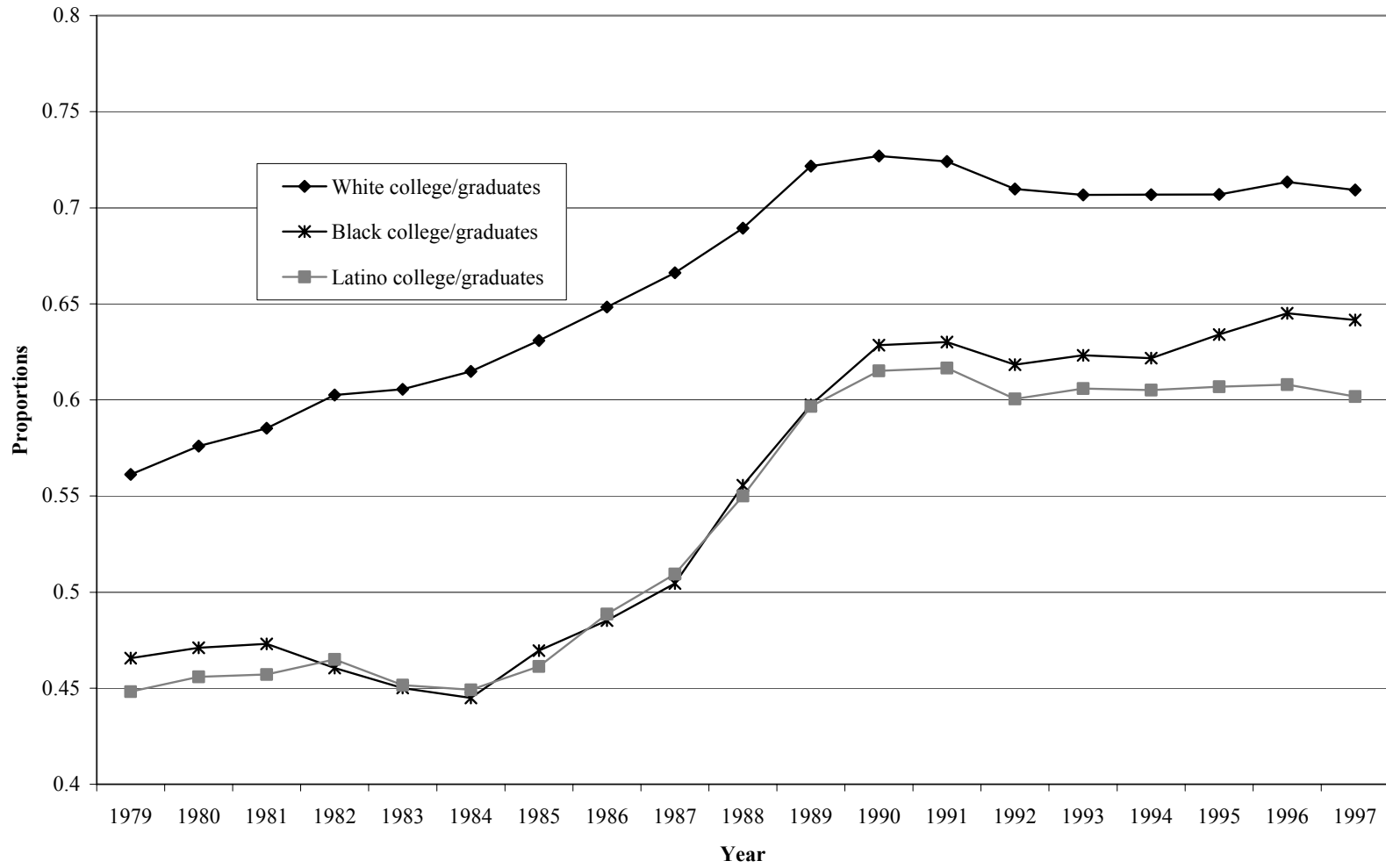


Figure 6a: Texas White Enrollment Trends By Grade

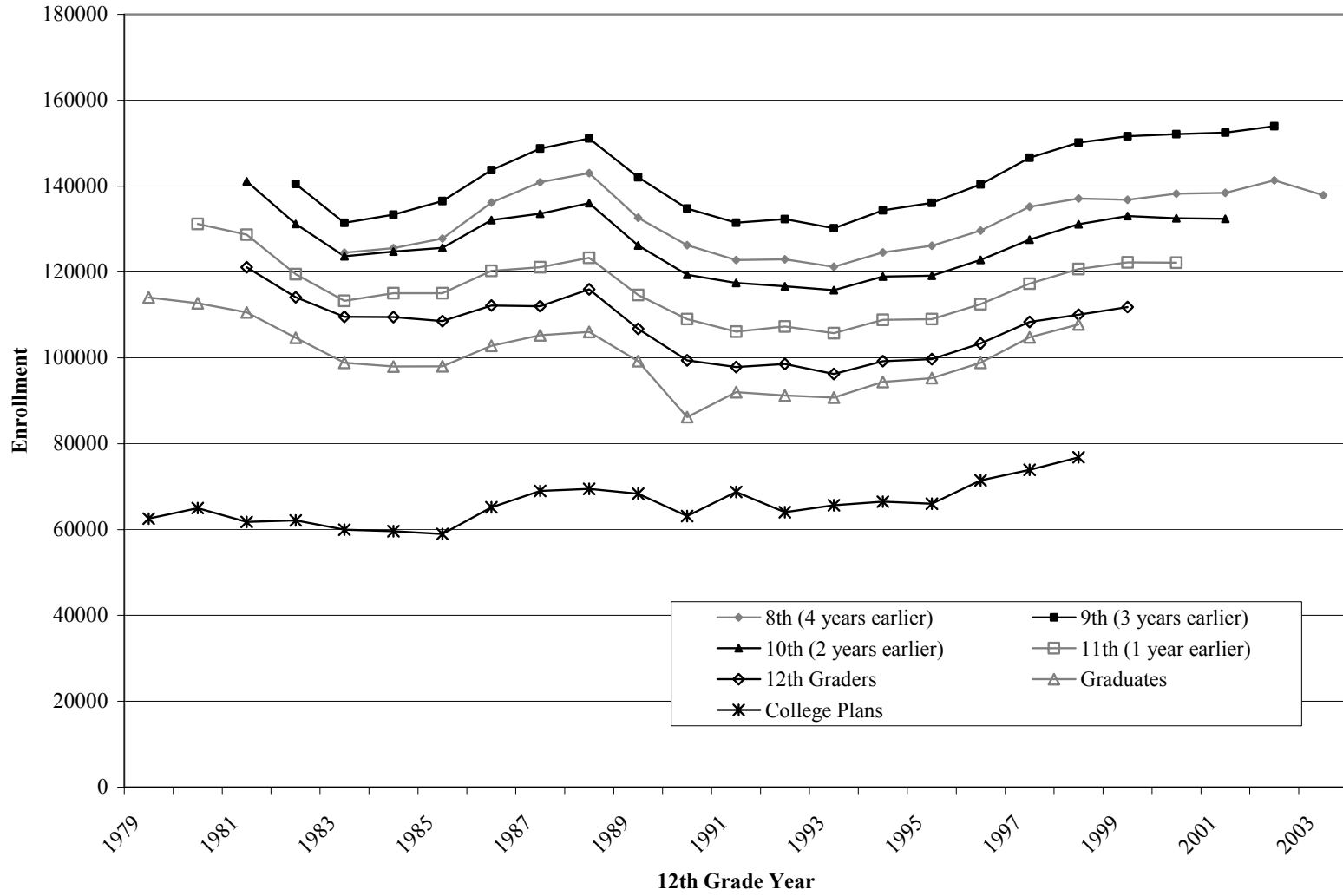


Figure 6b: Texas Black Enrollment Trends By Grade

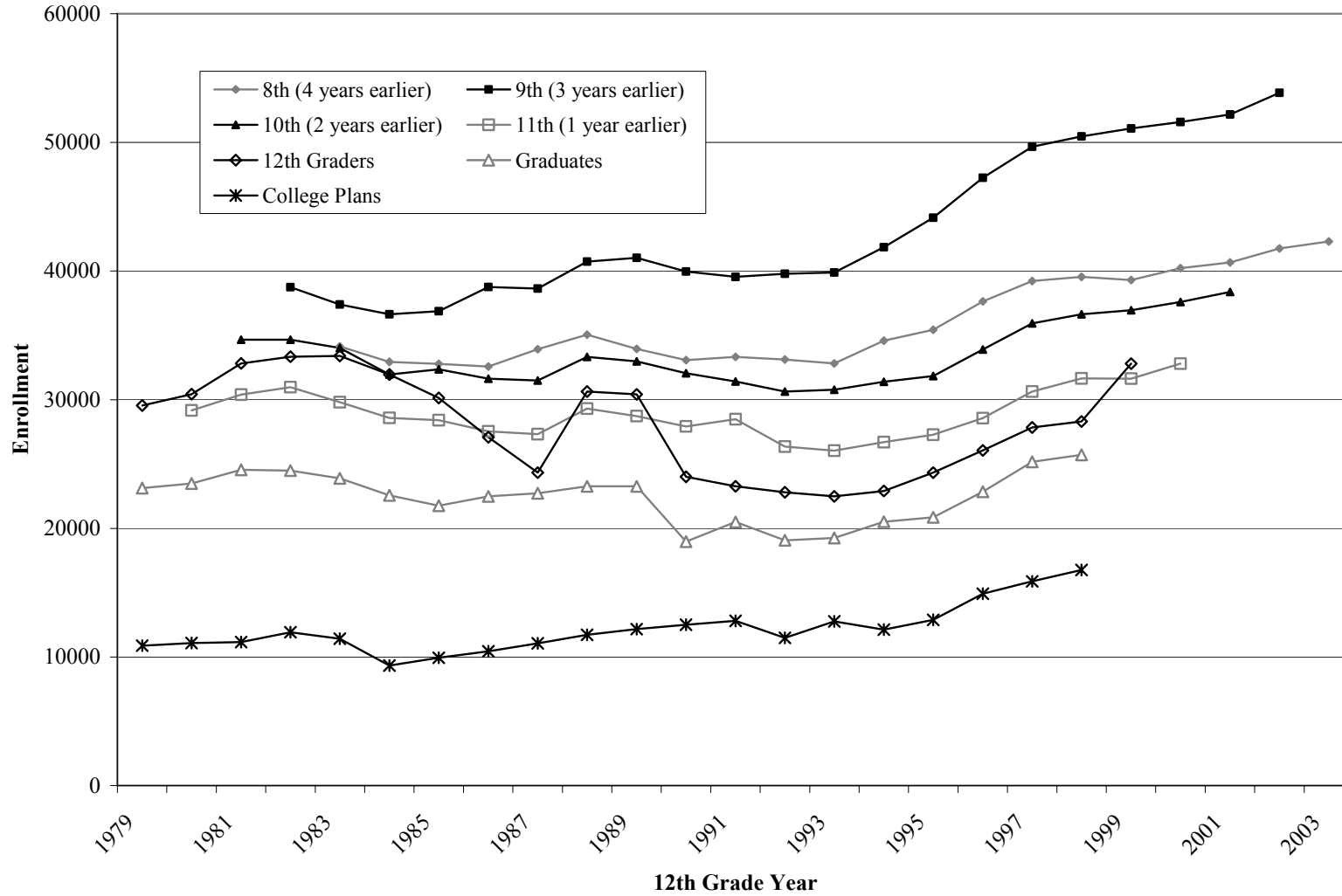


Figure 6c: Texas Latino Enrollment Trends By Grade

