Why Are Hispanic and African American Dropout Rates So High?

by

Magnus Lofstrom, PhD

Magnus Lofstrom, PhD

Magnus Lofstrom is Assistant Professor of Economics at the University of Texas at Dallas. His areas of expertise and research interests include education, immigration and self-employment. He received his PhD in economics from the University of California, San Diego in 1999. Prior to joining the faculty at UTD, he served as a researcher and taught at the Institute for the Study of Labor (IZA) at the University of Bonn and at the University of California, Irvine. Dr. Lofstrom also holds appointments as research fellow at IZA and research associate at the Center for Comparative Immigration Studies at the University of California, San Diego. His work has been published in numerous peer-reviewed journals including *Demography*, *Journal of Human Resources*, *Journal of Population Economics*, and *Research in Labor Economics*. He may be reached at P.O. Box 830688, GR 31, Richardson, TX 75083-0688, (972) 883-4763, e-mail: lofstrom@ utdallas.edu.

ABSTRACT

The proportion of students who do not graduate from high school is dramatically higher among the two largest minority groups, Hispanics and African Americans, compared with non-Hispanic Whites. This study utilized unique student-level data from the Texas Schools Microdata Panel (TSMP) in an attempt to determine which factors contribute to the higher minority dropout rates. The study shows that poverty is a key contributor. Lack of English proficiency among Hispanic students is linked to the higher Hispanic dropout probability. The results also suggest that neighborhood characteristics may be important in explaining the high African American dropout rates. The study also addresses the issue of surprisingly low official dropout rates reported by the Texas Education Agency (TEA) and shows that the GED program explains some of the discrepancy.

Keywords

dropout, high school, education, minority, Hispanic, African American

Acknowledgements

I would like to thank the anonymous referee for very helpful comments and suggestions.

Introduction

A substantial proportion of students do not complete high school, a problem particularly pronounced among the two largest minority groups in the United States—Hispanics and African Americans. Greene (2001) reported that only 54% of Hispanic students graduated from high school, and only a slightly higher percentage of African Americans (56%) completed high school. Among White students, the graduation rate is significantly higher (78%). Hispanic and African American students in Texas are also substantially more likely to drop out of school than White students. Using unique longitudinal student data for the class of 1999 in Texas, I have estimated that 40.3% of Hispanic, 38.5% of African American, and 26% of White public school students fail to graduate from high school by age 20.

There are a number of reasons motivating a study of dropout rates, particularly the high minority dropout rates. Chances for economic success among individuals who lack high school diplomas appear to be less likely today than at any other point in U.S. history. Evidence of this is widely reported both in the media and in academic journals. For example, Snower (1999) wrote, ". . . since the mid-1970s in the U.S., the earnings of the less educated have fallen rapidly behind those of the more educated" (p. 4). The increasing importance of skills and education is apparent for economic outcomes such as employment and earnings. Perreira, Harris, and Lee (2006) found that in 2000, high school dropouts were almost twice as likely to be unemployed as high school graduates.

Data from the 2000 U.S. Census also clearly showed that high school dropouts fared relatively poorly in the labor market (U.S. Census Bureau, n.d.). I have calculated, based on these data, that average annual earnings among male dropouts between the ages of 25 and 65 was $26,400, while male high school graduates earned, on average, close to $35,000 per year. Male college graduates in the same age group earned, on average, close to $40,000. Among women, the differences among groups with various levels of educational attainment are also notable. Moreover, the types of jobs available to high school dropouts rarely provide opportunities for significant upward mobility or benefits such as health insurance. In other words, a relatively certain road to economic long-term marginality is to not complete, at least, secondary schooling. Clearly, it is important for policymakers to know which factors and issues are related to poor educational outcomes, such as dropping out of high school.

This study had two main objectives. First and primarily, I wanted to identify factors that explain higher dropout rates among the two largest minority groups in Texas—Hispanics and African Americans. This study is based on unique longitudinal student level data, which included all Texas public school students in the cohort studied, the class of 1999. Second, I generated data about dropout rates in Texas that arguably reflect student high school outcome more accurately than official statistics and explain some reasons for the discrepancies.

Previous Research

There is a large body of research focused on identifying determinants of educational outcomes and student success. A review of this literature can be found in Haveman and Wolfe (1995). Family background, income, and parental education are factors frequently found to affect children's schooling outcomes. Other determinants are neighborhood and peer effects, as well as school characteristics. Surprisingly, research fails to find a consistent relationship between school resources and student achievement (Hanushek, 2006). Research investigating ethnic differences in school outcome, much of which has been focused on differences in test scores, has also found the above characteristics help explain minority-White differences in academic outcomes (Lavin-Loucks, 2006).

Next, I explored relevant economics of education research on ethnic differences in dropout rates. The influential recent study by Cameron and Heckman (2001) included an analysis of high dropout rates, specifically low high school graduation rates, among Hispanics and African Americans. They found that family factors (e.g., family composition, parental education, family income) explained the entire Black-White gap in high school graduation rates and most of the Hispanic-White gap. Cameron and Heckman also found that differences between White and minority scholastic ability, as measured by the Armed Forces Qualification Test (AFQT), played a prominent role in explaining the White-minority gap. Furthermore, by controlling for differences in family background and AFQT, they found that Hispanics and African Americans were more likely than Whites to graduate from high school. Cameron and Heckman concluded based on their findings that "It is early differences in resources and not later ones that matter more" (p. 492).

Given the large proportion of immigrants in Texas, particularly Mexican immigrants, research on high school completion among immigrants is also relevant. Perreira et al. (2006) found that differences in dropout rates were driven by

differences in human, cultural, school, and community capital. They also found that first-generation Hispanic immigrants were less likely to be high school dropouts than their parents, but the relative gains in schooling attainment, compared to their parents, decreased by the second and third generation. Noting that the average Hispanic immigrant had less than 9 years of schooling, and had already been out of school for at least 1 year by age 16, Betts and Lofstrom (2000) suggested low levels of educational attainment among young Hispanic immigrants, particularly those from Mexico, may be due to the possibility that they do not "drop in" to high school when they arrive in the United States.

In my analysis, I included several of the observable factors included in the Cameron and Heckman (2001) and the Perreira et al. (2006) studies, as well as school and school district characteristics. Unlike the Cameron and Heckman (2001) study, but like the Perreira et al. (2006) study, I was able to identify GED holders and have treated them as dropouts. Additionally, I was able to control for students' ability to speak English.

Data

This study utilized data from the Texas Schools Microdata Panel (TSMP), a student-level data set made available under special arrangements with the Texas Education Agency (TEA) (TEA, 2001). These unique data contain information on year, age, grade, school enrolled, gender, ethnicity, whether the student is economically disadvantaged, whether English is not the primary home language (ESL), whether the student has been identified as being Limited English Proficient (LEP), and whether the student participates in bilingual, special, or gifted education classes. For the high school graduates in the sample, I also had information on the graduation date. Furthermore, based on age, year, and grade enrolled, I calculated the number of grades a student had been held back.

In constructing the analytic data, I began with a balanced panel of all students whom I could observe were enrolled in Texas public high schools between the ages of 15 and 20. Since the TSMP data currently include student level information from 1990 through 2001, this necessarily means the study restricted the sample to students who belonged to cohorts that would have graduated in the classes from 1993 to 1999. For example, students who would have been expected to graduate in 1993 were 15 years of age in 1990, and would not have turned 20 until 1995. Thus, students in this graduating cohort

were in my sample and represented the earliest cohort. Similarly, students expected to graduate from high school in 1999 were 20 in 2001; hence, students in this graduating cohort were in my sample and represented the latest available cohort in my data. For the sample of 15- to 20-year-old students who had expected graduation dates from 1993 through 1999, I was able to observe, for each year, whether each student was enrolled or graduated from high school. For my analysis and empirical models, I restricted the study sample to the latest cohort I could follow up until they turned 20—that is, the class of 1999. Nonetheless, I have presented dropout rates for all cohorts in order to show the trend in Texas dropout rates.

Given the structure of the data, I defined a student as a school dropout if I observed him or her as enrolled in a Texas public school at the age of 15, and in some subsequent year between the age of 15 and 20 this person was not enrolled in a Texas public school, nor had they graduated from high school. Put slightly differently, I used information on whether the student was enrolled in a Texas public school in a particular year and/or if the student was reported graduating high school in that year. If the student was not enrolled and had not graduated high school, but was observed enrolled at age 15, this study defined the student to be a school dropout.

Caveats–Definition of Dropout Rates

The focus in this paper is students' secondary educational outcome by the age of 20. I defined a person to be a dropout if he/she was observed enrolled in a Texas public school at the age of 15 and had not graduated high school by the age of 20. Students who were observed enrolled in a Texas public school at the age of 15 and who either transferred to a private school, entered home schooling, or moved out of state could not be tracked in the data. Therefore, they were incorrectly identified as dropouts in the event they actually graduated high school. Also, students who were observed enrolled at age 15 in a Texas public school and passed away before graduating high school were also incorrectly identified as dropouts in the study's data. My dropout rates may then be viewed as "upper bounds." On the other hand, my dropout rates were underestimated by the fraction of students who dropped out at age 14 or younger.

My dropout rates were higher than the dropout rates reported by the Texas Education Agency (TEA, 2001), but quite closely corresponded to the dropout

rates implied by the high school graduation rates reported by Greene (2001). Greene estimated that 67% of students in the cohort who should have graduated in 1998 in Texas actually did graduate, implying a dropout rate of 33%. My calculation yielded a dropout rate of 34.6% for the 1998 graduation cohort. I will discuss plausible factors explaining differences between TEA's reported dropout rates and ours in the next section.

Data from Other Sources–School, District, and Local Information

In my empirical models, I also incorporated information on school district characteristics, such as expenditures and revenues per pupil. These data were collected from National Center for Education Statistics' (NCES) Local Education Agency Finance Survey for the years 1996 to 2001 (NCES, n.d.).[1] I also incorporated information about individual school characteristics, including pupil-teacher ratio, enrollment, ethnic/racial student composition, and school location. The school-level data were generated from NCES' Common Core of Data and were also for the years 1996 to 2001 (NCES, n.d.). Last, I generated annual local labor market conditions such as unemployment rate, employment and earnings growth, and average weekly earnings. These data were derived from the Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW) (Bureau of Labor Statistics, n.d.). The local labor market characteristics were generated at the Metropolitan Statistical Area (MSA) in which the school district was located. If the school district was located outside an MSA, this study utilized the state average for that particular year.

Descriptive Statistics

I have presented summary statistics by race/ethnicity for the class of 1999 in Table 1. The data show large differences across groups in student characteristics. For example, while approximately 17% of White students were economically disadvantaged (defined as eligible for reduced price or free lunch), more than 70% of Hispanics and close to 56% of African Americans were economically disadvantaged. Not surprisingly, the data also showed that

[1] The period 1996 to 2001 represents the years in which the students in this study's analytical cohort, the class of 1999, were between the ages of 15 and 20.

Table 1.
Sample Means, Class of 1999, by Race/Ethnicity

	White	Hispanic	African American	Asian
Dropout Rate	0.260	0.403	0.385	0.234
Student Characteristics				
Male	0.515	0.519	0.509	0.506
Immigrant	0.009	0.116	0.013	0.161
Economically Disadvantaged	0.172	0.702	0.558	0.289
Limited English Proficiency	0.004	0.234	0.005	0.176
English as a Second Language	0.003	0.183	0.004	0.145
Gifted Program Participant	0.140	0.070	0.093	0.242
Special Education	0.132	0.124	0.165	0.030
Held Back a Grade or More	0.273	0.486	0.428	0.264
School Characteristics				
Pupil-Teacher Ratio (# students)	14.4	15.2	15.5	16.0
Total Enrollment	1,561	1,685	1,636	2,318
Enrollment Less than 700	0.241	0.149	0.144	0.047
Enrollment 700–1,400	0.198	0.195	0.264	0.086
Enrollment 1,400–1,900	0.176	0.229	0.229	0.158
Enrollment 1,900–2,500	0.200	0.283	0.228	0.330
Enrollment Greater than 2,500	0.184	0.144	0.135	0.379
Percent Free Lunch	17.8	40.8	29.4	17.7
Percentage White	66.9	26.7	34.3	50.6
Percentage Hispanic	19.7	61.8	23.2	23.3
Percentage African American	10.2	9.5	39.1	16.4
Percentage Asian	2.9	1.8	3.2	9.5
School Location				
Central City, Large City, 250K+	0.164	0.388	0.461	0.351
Central City, Midsize City, <250K	0.139	0.202	0.173	0.104
Urban fringe, Large city	0.327	0.151	0.195	0.441
Urban fringe, Midsize city	0.052	0.057	0.021	0.018
Large town	0.010	0.011	0.012	0.002
Small town	0.112	0.100	0.065	0.026
Rural	0.196	0.090	0.074	0.057

(continued, next page)

Table 1. ***continued***
Sample Means, Class of 1999, by Race/Ethnicity

	White	Hispanic	African American	Asian
School District Characteristics				
Expenditure per Pupil	6,526	6,326	6,170	6,463
Title 1 Revenue per Pupil	103	216	157	89
Local Labor Market Conditions				
Employment Growth (%)	1.8	1.9	1.8	1.9
Unemployment Rate (%)	4.4	6.2	4.4	4.3
Average Weekly Earnings ($)	639	576	655	680
Earnings Growth (%)	4.6	4.3	4.7	4.7
Sample Size	145,365	103,809	40,887	7,830

Hispanic and Asian students, as well as LEP and ESL students, were more likely to be immigrants than White or African American students.

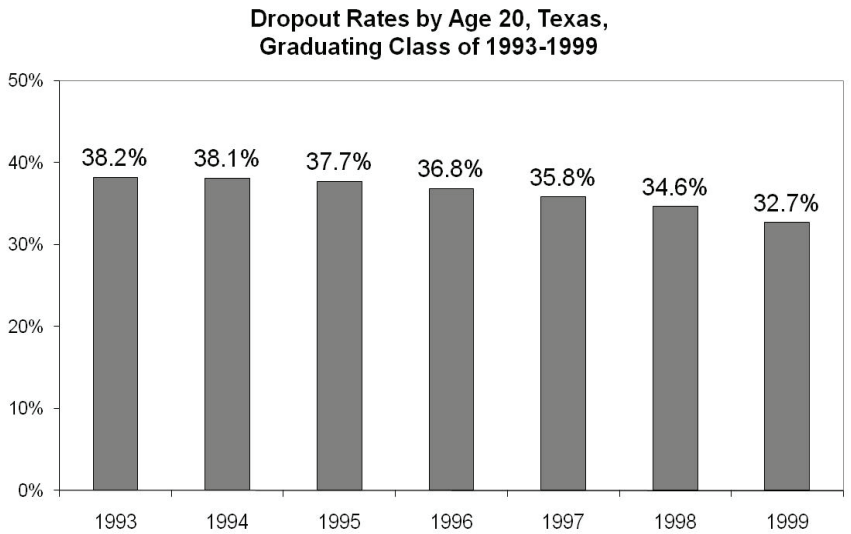
There were also differences across groups in school and school district characteristics for the schools the students attended. The data indicated that White students were more likely to attend smaller schools with lower pupil-teacher ratios. The majority of Hispanic and African American students attended schools located in central cities. Students in these two minority groups also attended schools in districts with lower expenditures per pupil. The local labor market data indicated that, among the four groups, Hispanics resided in areas with the lowest average earnings, earnings growth and employment growth, and the highest unemployment rates. Overall, the data indicated that Hispanics and African Americans attended schools and resided in areas less conducive to academic success than White and Asian students did.

Dropout Rates in Texas

Why Are TEA-Reported Dropout Rates So Low?

Dropout rates in Texas declined in the 1990s. Figure 1 shows that 38.2% of the class of 1993 did not graduate high school by age 20, while the dropout rate for students in the class of 1999 was 32.7%. TEA also reported a decline in the

Figure 1.



dropout rate during this period. However, the TEA-reported dropout rates were dramatically lower than the ones shown in Figure 1. For example, for the latest cohort available in the data and the cohort I have focused on in my analysis below—the class of 1999—TEA reported a 9% dropout rate (TEA, 2001). An obvious question is why these dropout rates differed so much. There are a number of reasons for the discrepancy. For example, TEA does not define a student who leaves school without graduating as a dropout if the student becomes incarcerated, has been expelled for criminal behavior, or has completed all course work but did not pass the Texas Assessment of Academic Skills (TAAS) exam (now the Texas Assessment of Knowledge and Skills exam, or TAKS) (TEA, n.d.).

One important reason for the difference in the dropout rate is the role of the General Educational Development (GED) program. The TEA definition of dropouts excludes all individuals who receive the GED credential instead of graduating. Furthermore, students who withdraw from school to enroll in an approved alternative program are not counted as dropouts. This includes students working toward completion of the GED certificate without necessarily passing the battery of GED exams successfully (TEA, n.d.).

The distinction between being a high school graduate and holding the GED credential (which has commonly, but incorrectly, been referred to as

the "high school equivalency" diploma) would arguably not be important if GED holders did as well in the labor market as high school graduates and pursued higher education to a similar extent. Research, however, has quite convincingly shown that neither of these holds true. Cameron and Heckman (1993) found that GED holders fared consistently worse than regular high school graduates on any number of labor market outcomes. Also, research has failed to find that the GED increases earnings among minorities relative to other minority dropouts (Tyler, Murnane, & Willett, 2000). Regarding higher education, the findings of Lofstrom and Tyler (2005) suggested that the GED was not a particularly effective route to postsecondary education relative to staying in school and obtaining a high school diploma. It appears quite clear that dropouts with the GED are not the equivalent of high school graduates, and hence should not be treated as such in official educational attainment statistics.

To illustrate the importance of the GED in explaining the discrepancy between the study's rates and the official TEA dropout rates, I generated the percentage of dropouts from the class of 1999 who attempted and passed the GED exam by the age of 20, obtained from the TSMP data. These statistics are presented in Table 2. The data show that 20% of dropouts in this cohort obtained the GED credential. If students with the GED credential are excluded, using my data for the class of 1999, the implied dropout rate was 28%, which suggests the official dropout rates were understated, at least, by 17%.[2] Table 2 also shows that slightly more than 29% of dropouts attempted the GED by the age of 20. If all these individuals are excluded from my sample, the implied dropout rates would be 25.6%, suggesting that that the official dropout rate was understated by more than 27%. It should be noted that TEA does report the number of students who graduated in each cohort. For the class of 1999 cohort, TEA reported that 78.1% graduated, implying a dropout rate of 21.9% (TEA, 2001). However, this statistic appears to receive substantially less attention than the reported so-called dropout rate.

Although these calculations make clear that the GED plays an important role in explaining the low official dropout rate, they also show

[2] The unadjusted dropout rate shown in Figure 1 was 32.7%, suggesting that the TEA dropout rate was understated by at least a factor of (32.7 - 28)/28 = 0.169.

Table 2.
Number and Percentage of Dropouts Who Passed and Attempted the GED, Class of 1999, by Race/Ethnicity

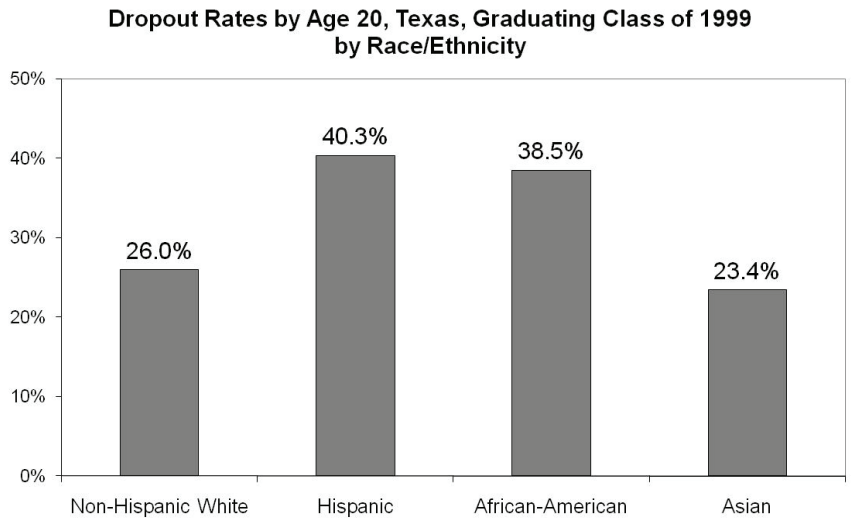
	Total # Students	Total # Dropouts	Passed GED (#)	Passed GED (%)	Attempted GED (#)	Attempted GED (%)
White	145,365	37,785	10,806	28.6	13,830	36.6
Hispanic	103,809	41,886	6,230	14.9	10,223	24.4
African American	40,887	15,751	2,196	13.9	3,882	24.6
Asian	7,830	1,834	260	14.2	348	19.0
Native American	690	293	60	20.5	82	28.0
All Students	298,581	97,549	19,552	20.0	28,365	29.1

that other factors are important. For example, as long as a student withdraws from school with the intent to enroll elsewhere, the student is an "official other leaver" and will not be included in the statistics used to calculate the official dropout rate. Put differently, these students are not part of the denominator used to derive the official dropout rate. An indication that this issue is important in explaining the differences between my dropout rates and the official TEA dropout rates is the discrepancy in the reported size of the class of 1999. TEA reported that the class of 1999 grade 7 cohort included 240,865 students, while the class of 1999 used in this paper contained 298,581 students (TEA, 2001). Again, the GED and the number of students attempting the GED was a factor explaining the discrepancy. The number of GED test takers in the class of 1999, according to my data, was 28,365, which accounts for a difference of more than 29,000 students between TEA's reported cohort size and ours. It appears unlikely that 29,000 students from the class of 1999 (or about 10% of the cohort) who were enrolled in Texas public schools at the age of 15 enrolled in private schools, out-of-state schools, out-of-country schools, or became home schooled (and eventually graduated). Hence, the official dropout rates arguably do not accurately reflect secondary school attainment among students in Texas.

Differences in Dropout Rates Across Groups

The main objective of this paper is to analyze differences in dropout rates and particularly to identify factors explaining the high dropout rates among Hispanic and African American students. My calculated dropout rates varied substantially across ethnic groups, as shown in Figure 2, and were especially high among Hispanics and African Americans. According to the data, slightly more than 40% of Hispanic students in the class of 1999 who were enrolled in Texas public schools at age 15 did not graduate from high school by age 20. The dropout rate was almost as high for African Americans, 38.5%. Both White and Asian students dropped out to a significantly lower extent, 26% and 23.4% respectively. In my empirical analysis below, I have investigated the causes for these large differences.

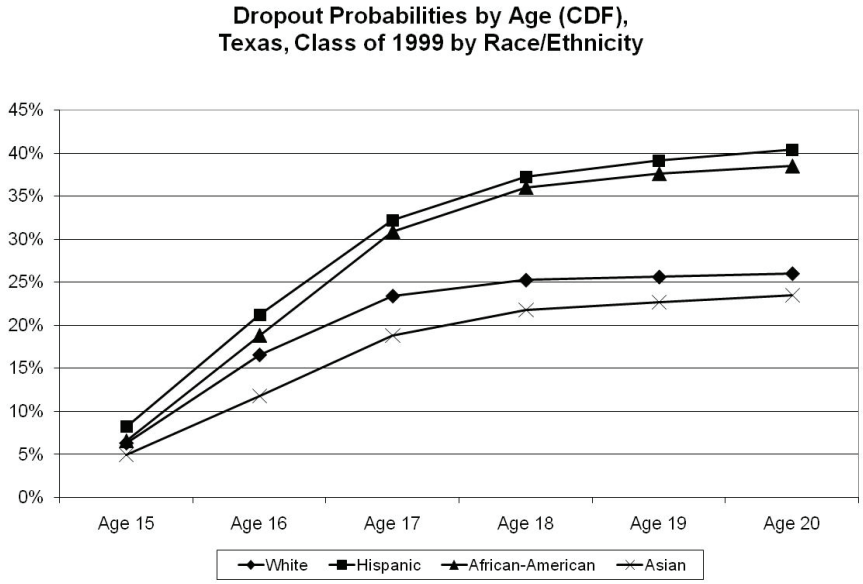
Figure 2.



Of obvious interest is also the timing of dropping out of school. That is, at what age and/or grade are students most likely to drop out? Figure 3 shows unconditional dropout probabilities by age for each ethnic/racial group—in other words, group-specific cumulative density functions (cdf). The figure indicates relatively few students were dropouts by age 15, and that there were only small differences across groups at this young age. This is consistent with the findings of Cameron and Heckman (2001), who noted that "few males quit

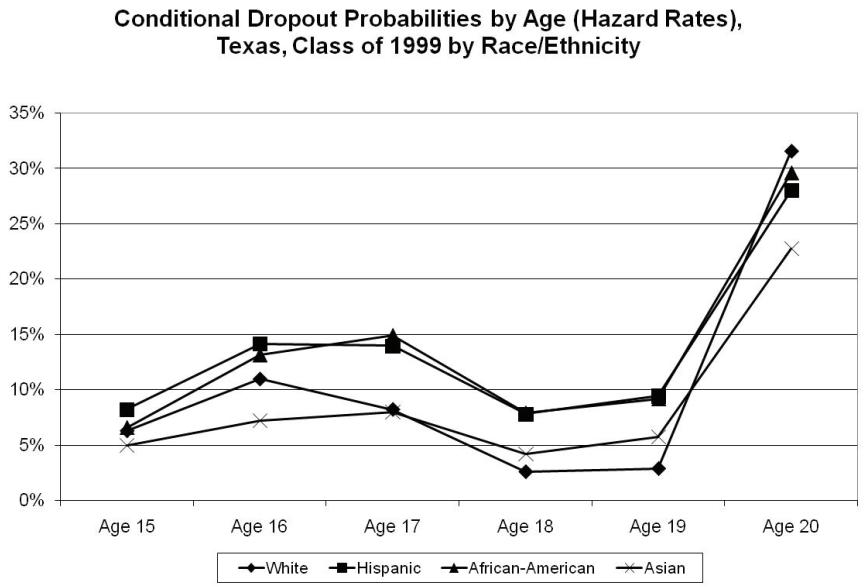
school before age 16" (p. 459). Furthermore, they noted that "this finding is due in part to laws about compulsory schooling attendance and to the lack of labor market opportunities for people younger than 16" (p .461). The largest difference in dropout rates in Texas at age 15 was between Hispanic and Asian students (3%). However, the dropout gap between these two groups increased with age, with a considerable increase during the year students turned 17. Overall, Figure 3 suggests that at one end of the spectrum, the changes in cumulative dropout rates over age were quite similar and relatively low for Whites and Asians, while on the other end, dropout rates for Hispanics and African Americans were alarmingly high.

Figure 3.



The longitudinal data utilized here can also be used to generate conditional dropout probabilities. The conditional dropout probability—the hazard rate—is the probability that a student will drop out by the end of the year, conditional upon being enrolled during the year. Unlike the unconditional dropout probability, the hazard rate does not need to be a monotonically increasing function with age. It also has the advantage of more clearly illustrating the timing of differences in dropout probabilities across groups. I present the group-specific hazard rates in Figure 4.

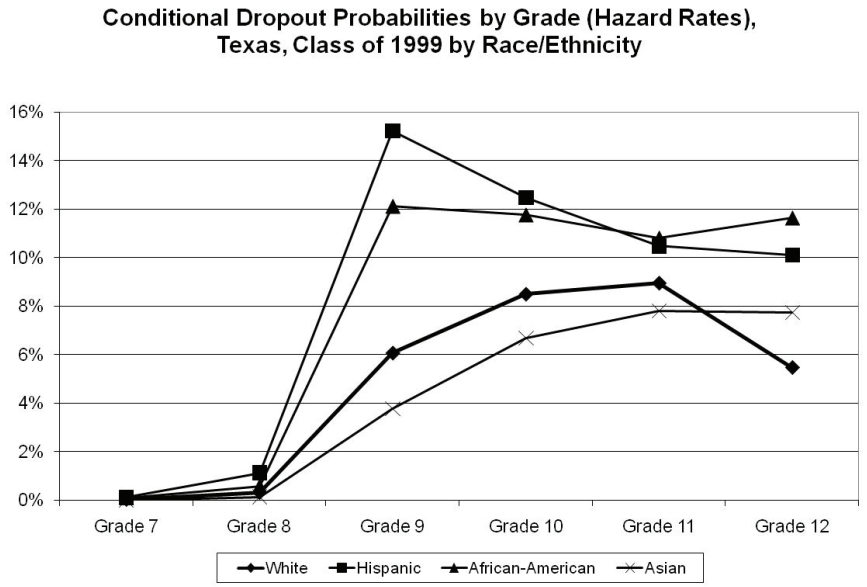
Figure 4.



It appears that the large differences in dropout rates across groups start at age 16. Figure 4 shows that, conditional on having stayed in school until age 16, Hispanics and African Americans were approximately twice as likely as Asian students to drop out by the end of the year. This roughly held true for students who continued to be enrolled the following year, when they turned 17. Approximately 8% of Asian and White students dropped out during this year, while 14% of Hispanic students dropped out. Among African American students, 15% dropped out. It is not surprising that conditional dropout rates declined for all groups when students stayed enrolled until age 18, the age at which most students graduate if they do graduate. Although the gaps in the conditional dropout probability between both Hispanics and Whites and African Americans and Whites decreased to 5%, these minority groups were still more than twice as likely to leave school without graduating at this age than White students. Overall, Figure 4 suggests that the higher dropout rates among Hispanics and African Americans, compared to Whites and Asians, develop at ages 16 and 17.

Arguably, it is of greater relevance to determine at what grade, as opposed to age, the dropout differences arise. To shed light on this, I generated group-specific conditional dropout probabilities, or hazard rates, by grade. These are presented in Figure 5. The figure quite clearly shows that these differences are

Figure 5.



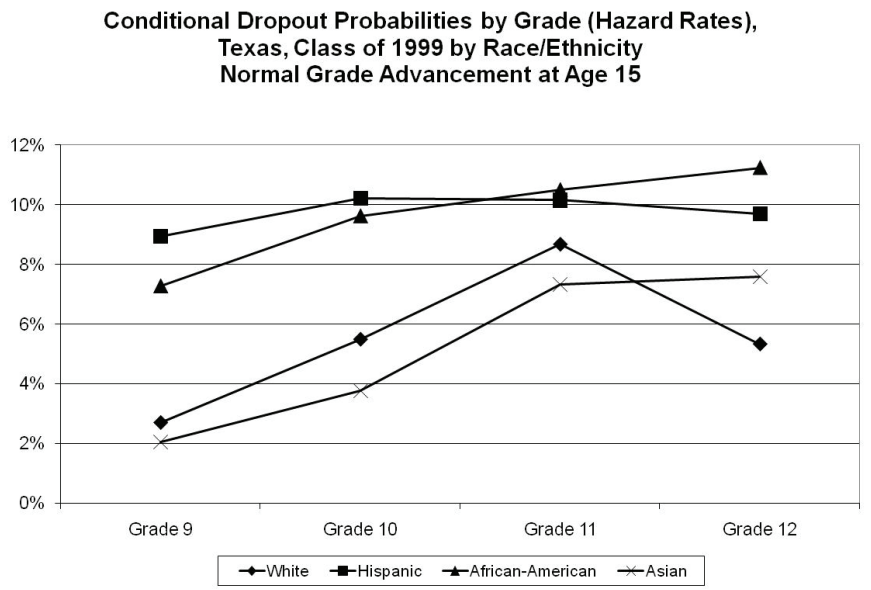
particularly generated in the 9th grade, and that relatively few students drop out in earlier grades. Approximately 12% of African American students who enrolled in 9th grade did not complete this grade. Among Hispanic students, an even larger proportion of students did not finish 9th grade—slightly more than 15%. This is in stark contrast to White and Asian students. For these two groups, 6% and 4% respectively dropped out in 9th grade. Figure 5 also shows quite clearly that these differences declined in subsequent grades, with the exception of the difference between Whites and African Americans in 12th grade. Surprisingly, the conditional dropout rates among African American students enrolled in 12th grade were higher than the rates among African American students enrolled in the 11th grade. In fact, the conditional dropout probability rates remained relatively constant among African American students from grades 9 through 12, while they declined among Hispanic students in these grades.

The importance of the 9th grade as an explanation for the differences in dropout rates between the groups can also be illustrated by showing the group-specific proportion of all dropouts who dropped out in 9th grade. About 16% of Asian dropouts left school in the 9th grade, while about 23% of White dropouts enrolled in 9th grade without completing it. Among Hispanic and African American dropouts, 37% and 31% respectively dropped out in 9th grade.

One plausible explanation for the substantially higher Hispanic and African American 9th grade conditional dropout probabilities, compared to Whites, is differential grade retention across ethnic groups. If a student is progressing through school at a normal rate—that is, has not been held back a grade—I would expect him or her to be enrolled in 9th grade by age 15. If, however, a student has been held back one grade, I would expect the student to be in 9th grade the year he or she turns 16. An important observation, as noted above, is that this is also the year in which students, according to compulsory attendance rules, may leave school. Furthermore, Cameron and Heckman (2001) reported that minority groups were significantly more likely to have fallen behind in school by age 15 than Whites.

To investigate whether normal grade progression, or lack thereof, is a key reason for the large minority-White difference in the 9th grade conditional dropout probabilities, I generated the grade-specific conditional dropout probabilities by ethnic groups for the subsample of students who were on track at age 15, and were hence enrolled in 9th or higher grades by this age. The hazard rates for this restricted sample are shown in Figure 6. The figure reveals that although the differences were

Figure 6.



smaller, large differences remained, and the largest minority-White gaps, specifically Hispanic-White and African American-White, were still observed for the 9th grade. It appears that other factors besides differences in grade retention across ethnic groups are the main determinants. Although beyond the scope of this paper, further research is needed to determine specific causes for the increase in the dropout probability at this point, and also reasons why minorities have a harder time transferring to high school than Whites.

Empirical Dropout Probability Model

Table 1 suggests there are large differences in student, school, and district characteristics across the four ethnic/racial groups. A key objective of this paper is to determine how these factors relate to the differences in dropout rates across the groups.

The educational outcome analyzed in this paper was whether or not a student of the class of 1999 who was observed enrolled in a Texas public school at age 15 was observed not completing high school by the age of 20, and hence defined as a dropout. Let the outcome variable y_{ijk} equal zero if student i in school j in school district k was observed graduating from high school by age 20, and let y_{ijk} equal one if the student was observed dropping out between the ages of 15 and 20.[3]

We may think of the observed outcome y_{ijk} as being the result of a latent process in which a student compares the marginal benefits to the marginal costs of continuing schooling. In this case, let the continuous latent variable y_{ijk}, which represents the value an individual i in school j in district k receives from his or her particular decision for each school year in which the student is still enrolled and between the ages of 15 and 20, be specified as:

$$y^*_{ijk} = \alpha_0 + \alpha_1 Hisp_{ijk} + \alpha_2 Black_{ijk} + \alpha_3 Asian_{ijk} + \mathbf{X}_{ijk}\boldsymbol{\beta} + \mathbf{W}_{jk}\boldsymbol{\delta} + \mathbf{Z}_k\boldsymbol{\gamma} + \varepsilon_{ijk}$$

$$y_{ijk} = 1(y^*_{ijk} > 0)$$

[3] Given the approach and data, the outcome analyzed is identical to not graduating high school by age 20. One implication, and advantage, of this is that students who leave school but return at a later point are only defined to be dropouts if they did not graduate by age 20.

Where 1(•) is an indicator function equal to one if the enclosed statement is true—that is, the student opted to drop out—and zero otherwise. The corresponding, linear dropout probability model is given by equation 1:

$$\mathbf{P}\left[y_{ijk} = 1 \mid \mathbf{X}, \mathbf{W}, \mathbf{Z}, \text{Race/Ethnicity} \right] = \alpha_0 + \alpha_1 Hisp_{ijk} + \alpha_2 Black_{ijk} + \alpha_3 Asian_{ijk} + \mathbf{X}_{ijk}\beta + \mathbf{W}_{jk}\boldsymbol{\delta} + \mathbf{Z}_k\boldsymbol{\gamma} + \varepsilon_{ijk}$$

Where *Hisp*, *Black*, and *Asian* are indicator variables for Hispanics, African American, and Asian students respectively, and matrices **X**, **W**, and **Z** are defined as:

$\mathbf{X}_{ijk}$ = Matrix containing student characteristics, such as gender, whether the student was designated as economically disadvantaged, reported English as a Second Language (ESL), was designated as Limited English Proficient (LEP), indicator variables for whether the student had been held back a grade or more by age 15, was identified as an immigrant, and whether he or she participated in special or gifted education.

$\mathbf{Z}_{jk}$ = Matrix containing controls for school characteristics, such as pupil-teacher ratio, enrollment, school location, percentage of students who received free lunch, and percentage of students who were White, Hispanic, African American, and Asian.

$\mathbf{W}_k$ = Matrix containing controls for school district characteristics and local labor market conditions by the MSA in which the district is located. The matrix includes variables for expenditure per pupil and per student federal Title 1 revenues. The local labor market characteristics included are employment growth, the unemployment rate, average weekly earnings, and earnings growth.

The coefficients of particular interest in the above specification are α_1, α_2, and α_3, which represent the differences in the dropout probability between White and minority students. An alternative, and far more complex, approach would be to model the dropout probability as a dynamic process, such as Cameron and Heckman (1998) and Colding (2006) did. Given that my empirical approach does not account for possible dynamic selection bias, the results should be interpreted with some caution.

Empirical Results

The model specifications defined above are estimated using a sample of students who belonged to the class of 1999 cohort. My approach was to start with a parsimonious model specification and subsequently add controls for factors expected to affect the dropout decision and for group differences in the dropout probability. This will shed light on what specific factors contributed to the White-minority gaps and to what extent. The estimated coefficients and *t*-statistics are shown in Table 3.[4]

I begin by presenting the Linear Probability Model (LPM) results from my first specification, Model 1, which only controlled for gender and nativity. The objective here was to investigate whether the high dropout rates for Hispanic students were partially driven by the higher proportion of immigrants, who may face difficulties adjusting to U.S. schools. The results show that, although immigrant students were significantly more likely to drop out of school, the higher proportion of foreign-born students among Hispanics only explains a small proportion of this group's overall higher dropout rate. Native-born Hispanic students were still about 13 percentage points more likely to drop out than White students, compared with the unadjusted difference of 14.4 percentage points. Given the low proportion of immigrants among African Americans, it is no surprise that the controls for gender and nativity did not affect the estimated Black-White dropout gap, which remained at 12.5 percentage points.

Minority students studied were more likely to be economically disadvantaged, as can be seen in Table 1, a factor likely to impact educational attainment. To assess the role of poverty on group differences in the dropout probability, I added the measure of students who came from economically disadvantaged families, shown as Model 2. The estimates clearly illustrate the influence of poverty on both the dropout probability and on differences between minority and White students. A student who was economically disadvantaged was approximately 12 percentage points more likely to drop out of school than other students. My estimates also show that,

[4] Given the approach and data, the outcome analyzed is identical to not graduating high school by age 20. One implication, and advantage, of this is that students who leave school but return at a later point are only defined to be dropouts if they did not graduate by age 20.

Table 3.

Linear Probability Models of the Probability of Dropping Out of School, Class of 1999

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Hispanic	**0.131**	**0.069**	**0.044**	**0.043**	**0.039**	**0.039**	**0.036**	**0.031**	**0.028**
	(16.78)	**(9.03)**	**(5.99)**	**(6.20)**	**(8.05)**	**(8.25)**	**(7.62)**	**(6.62)**	**(11.14)**
African American	**0.125**	**0.079**	**0.083**	**0.070**	**-0.015**	**-0.014**	**-0.009**	**-0.009**	**-0.005**
	(12.99)	**(8.29)**	**(8.73)**	**(7.05)**	**(2.82)**	**(2.52)**	**(1.92)**	**(1.79)**	**(1.55)**
Asian	-0.042	-0.054	-0.071	-0.021	-0.053	-0.052	-0.053	**-0.051**	**-0.054**
	(4.53)	(6.39)	(8.47)	(2.27)	(7.08)	(6.92)	(7.80)	**(7.48)**	**(11.14)**
Native American	0.162	0.142	0.142	0.126	0.111	0.105	0.087	**0.086**	**0.104**
	(7.75)	(7.02)	(7.00)	(6.41)	(5.75)	(5.45)	(4.64)	**(4.77)**	**(5.98)**
Student Characteristics									
Immigrant	0.111	0.095	0.011	0.026	0.013	0.017	0.041	0.037	-0.006
	(6.38)	(5.77)	(0.83)	(2.31)	(1.08)	(1.69)	(2.58)	(2.35)	(1.33)
Male	0.071	0.070	0.068	0.056	0.055	0.056	0.052	0.048	0.046
	(29.40)	(28.75)	(28.66)	(26.77)	(26.57)	(26.66)	(23.53)	(21.91)	(28.76)
Economically Disadvantaged		0.121	0.109	0.082	0.076	0.076	0.071	0.066	0.077
		(20.22)	(17.82)	(14.24)	(15.04)	(15.15)	(13.32)	(12.46)	(37.08)
Limited English Proficiency			0.073	0.056	0.053	0.046	0.048	0.042	0.046
			(3.83)	(3.62)	(3.70)	(2.70)	(1.95)	(1.79)	(7.04)
English as a Second Language			0.130	0.137	0.136	0.142	0.128	0.120	0.130
			(5.90)	(7.44)	(8.19)	(7.55)	(4.78)	(4.68)	(17.90)

Note. The data rows with bold print capture the estimated difference in the dropout probability between White students and Hispanic/African American students, holding the factors included in the specification constant.

Table 3 continued, next page.

Table 3. ***continued***

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Student Characteristics, ***continued***									
Gifted Program Participant				-0.227 (30.99)	-0.240 (25.77)	-0.232 (24.61)	-0.181 (19.51)	-0.175 (19.60)	-0.234 (117.60)
Special Education				0.074 (14.54)	0.076 (15.76)	0.078 (16.20)	0.074 (11.54)	0.062 (9.76)	0.060 (22.59)
School Characteristics									
Pupil-Teacher Ratio				0.008 (3.12)	0.006 (2.85)	0.006 (2.77)	0.005 (2.43)	0.005 (2.46)	
Total Enrollment				-0.005 (0.21)	-0.033 (1.52)	-0.044 (2.05)	-0.052 (2.53)	-0.047 (2.30)	
Enrollment 700-1,400				0.024 (1.23)	-0.007 (0.41)	-0.023 (1.21)	-0.001 (0.07)	-0.003 (0.18)	
Enrollment 1,400-1,900				-0.013 (0.40)	-0.059 (1.87)	-0.060 (1.92)	-0.022 (0.65)	-0.015 (0.45)	
Enrollment 1,900-2,500				-0.073 (1.77)	-0.103 (2.62)	-0.100 (2.56)	-0.073 (1.85)	-0.062 (1.56)	
Enrollment Greater than 2,500				-0.070 (1.15)	-0.084 (1.46)	-0.081 (1.42)	-0.025 (0.41)	-0.016 (0.27)	
Percentage Free Lunch					0.002 (3.40)	0.002 (4.31)	0.001 (1.58)	0.001 (0.83)	

Table 3 continued, next page.

Table 3. ***continued***

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
School Characteristics, ***continued***									
Percentage Hispanic					-0.0012 (3.09)	-0.001 (1.45)	-0.002 (3.04)	-0.002 (2.74)	
Percentage African American					0.001 (2.74)	0.001 (2.95)	0.003 (5.55)	0.003 (5.81)	
Percentage Asian					0.003 (2.00)	0.003 (2.03)	0.008 (4.12)	0.007 (3.90)	
School Location									
Central City, Large City (250K+)					0.191 (10.93)	0.162 (8.94)	0.150 (5.18)	0.136 (4.76)	
Central City, Mid-size City(<250K)					0.155 (10.34)	0.129 (7.65)	0.015 (0.38)	0.003 (0.07)	
Urban fringe, Large city					0.140 (10.20)	0.117 (8.33)	0.140 (6.61)	0.130 (6.28)	
Urban fringe, Mid-size city					0.032 (1.78)	0.021 (1.05)	-0.173 (4.25)	-0.173 (4.36)	
Large town					0.080 (2.01)	0.044 (1.01)	0.045 (0.64)	0.040 (0.56)	
Small town					0.033 (3.05)	0.025 (1.95)	0.033 (1.94)	0.032 (1.94)	

Table 3 continued, next page.

Table 3. ***continued***

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
School District Characteristics									
Expenditure Per Pupil						-0.037	-0.033	-0.032	
						(2.96)	(2.90)	(2.89)	
Title 1 Revenue Per Pupil						-0.342	-0.485	-0.447	
						(4.39)	(4.24)	(3.96)	
Local Labor Market Conditions									
Unemployment Rate							0.024	0.023	
							(5.27)	(5.32)	
Weekly Earnings							-0.002	-0.002	
							(13.55)	(13.88)	
Employment Growth							0.064	0.064	
							(8.48)	(8.63)	
Weekly Earnings Growth							0.064	0.064	
							(8.48)	(8.63)	
Grade Retention									
Held Back a Grade at Age 15								0.349	0.302
								(37.66)	(61.04)
Number of Students					**298,581**				
R-Squared	0.031	0.043	0.052	0.091	0.113	0.128	0.259	0.275	0.279

other factors constant, Hispanic students were approximately 7 percentage points more likely than White students to drop out of school, while African American students were roughly 8 percentage points less likely to complete high school. These results imply that my simple poverty measure explains almost half of the Hispanic-White dropout probability difference and more than a third of the Black-White disparity.

Another potentially important factor contributing to high minority dropout rates, mainly among Hispanic students, is English proficiency (Valenzuela, 2006). I added controls for English as a Second Language (ESL) and Limited English Proficiency (LEP) to my models in Model 3. I found that both of these variables affected the dropout probability. The results also show that the higher dropout rates among immigrant students were due to relatively low English proficiency—the estimated immigrant coefficient became statistically insignificant. Limited English proficiency status was also a contributor to the Hispanic-White difference in the dropout probability, which decreased to about 4.4 percentage points with the added ESL and LEP variables. The two language factors explained slightly more than a third of the estimated dropout probability difference between Hispanic and White students, compared with my Model 2 specification estimate.

Next, I turned to the role of school characteristics, and particularly pupil-teacher ratio and school size. The results are presented as Model 4. Although I found that a school's average pupil-teacher ratio affected the dropout probability positively and significantly, I did not find that school size significantly altered students' decision to leave school before graduation in this model specification. Furthermore, these factors did not appear to explain much of the higher dropout probabilities among Hispanic and African American students, relative to White students.

The next model specification, Model 5, was intended to investigate the role of school location and the ethnic/racial composition of its students, the latter capturing, to some extent, peer effects. Table 1 shows that Hispanic and African American students were much more likely to attend schools located in the center of large cities than White students. The data also show that Texas schools are segregated in the sense that White, Hispanic, and African American students were most likely to attend schools where their ethnic/racial group is the largest ethnic/racial group. Clearly, school composition is a factor of area—or neighborhood—characteristics, and the effects of composition and location on educational outcome are difficult to disentangle.

The results for Model 5 indicated that, holding all other factors constant, students who attended schools in central large- and mid-sized cities were more likely to drop out of school than students who attended schools in any other location. The estimates also indicate that students who attended schools with a higher proportion of economically disadvantaged students, all else being equal, were somewhat less likely to graduate from high school. The results also point to a positive association between the dropout probability and the proportion of African American and Asian students. Interestingly, the estimates also imply a negative correlation between the dropout probability and the proportion of Hispanic students. An important finding is that my estimates showed location-related factors strongly affected the difference in the dropout probability between African American and White students. Once I added these controls to the specification, holding all factors constant, African American students were no more likely to drop out of school than White students. However, Hispanic students were still more likely to drop out than White students.

Other potentially influential dropout factors were the amount spent per student—that is, per-pupil expenditure—and revenues received from Title I. Title I funds are specifically intended to improve academic achievement of the disadvantaged. It is possible that differences across school districts in the amounts received from Title I, as well as per-student expenditure, may affect ethnic group differences in the dropout probability. The results, when I added these two factors to the specification in Model 6, suggest that higher spending and Title I revenues per student are associated with lower student dropout probability. However, there was no indication that differences in spending or revenues explain high minority dropout rates. Similarly, when I added the controls for local labor market conditions, as shown in Model 7, I found that while these conditions appear to be linked to the probability of completing high school, they do not seem to be strong contributors to group differences in the dropout probability.

The most general specification presented so far, Model 7, suggests that the higher dropout probability among African American students, relative to White students, can be entirely explained by my set of observable factors. Observationally similar Hispanic students are, however, still predicted to be more likely to drop out of school than White students by approximately 3.6 percentage points.

In the last model specification, Model 8, I added an indicator variable for whether a student had been held back one grade or more by age 15. I based

this variable on whether the student was observed enrolled in 8th grade or lower during the year in which he or she turned 15. Grade retention is arguably simply a symptom, or predictor, of poor educational outcomes resulting from academic and personal difficulties experienced by the student. Viewed from this perspective, the finding that students who were held back one grade or more were more likely to drop out of school is neither surprising nor extremely useful, because it is a manifestation of these negative experiences, but does not indicate which specific ones caused the student to drop out. Nonetheless, in the specification with the grade retention variable (Model 8), my estimated minority-White differences in the dropout probability tells us what these differences are, conditional on grade retention—that is, whether a Hispanic student who was held back a grade was more likely to drop out than a White student who was held back a grade, all other factors held constant.

My estimates show that differences in grade retention between Hispanic and White students were a relevant factor. Students held back a grade or more were substantially more likely to drop out, by close to 35 percentage points, compared with students who stayed on track. In other words, grade retention is a relatively good predictor of dropping out of school. The Model 8 estimates suggest that, compared with Whites, lack of successful grade progression among Hispanic students contributed to the higher Hispanic dropout rates. The estimates also indicate that Hispanic students who made normal grade progression and were enrolled in at least 9th grade by age 15 were still more likely to drop out (by about 3 percentage points) than White students who also had not been held back a grade by age 15.

Finally, to determine whether other school characteristics, beyond the ones specified above and for which the study does not have information, help explain the difference in dropout probability between Hispanic and White students, I also estimated a school fixed-effects specification, shown as Model 9. Note that in this specification I cannot include the observable school and school district variables, as they are perfectly multicolinear with the school fixed effects. The estimated Hispanic-White difference decreased only marginally to 2.8 percentage points. These results suggest that the school characteristics included in Models 7 and 8 are most relevant for explaining the Hispanic-White dropout probability difference. Last, the fixed effects results also serve as a robustness check of my estimated individual student effects. The generally small differences in the estimates between Model 8 and 9

indicate that student-level estimates are not driven by unobserved school heterogeneity, which may be correlated with the student characteristics.

Summary and Conclusions

This paper analyzes the high dropout rates among Hispanic and African American students in Texas. Utilizing unique longitudinal student data, this study shows that Hispanic and African American students are about 14 and 12 percentage points, respectively, more likely to drop out of school than White students. This study also shows that differences in the dropout rates appear to develop in 9th grade.

I have estimated multivariate dropout probability models to assess what factors affect the student dropout probability and differences across groups. Factors associated with higher minority dropout rates differ somewhat between Hispanics and African Americans. However, one determinant in common for these minority groups is poverty. Almost half of the difference in dropout probability between Hispanic and White students stems from the higher likelihood of Hispanic students being economically disadvantaged. More than one third of the African American and White student difference in the dropout probability was linked to my simple measure of student poverty. This study also finds that lack of English proficiency is a key factor in explaining the high Hispanic dropout rate. The results suggest that neighborhood characteristics, here simply measured by the location of the school attended and the school's student race/ethnicity composition, contribute more to the high African American dropout probability than school characteristics, such as pupil-teacher ratio and expenditure per pupil. This suggests the need for future research that not only focuses on school characteristics, but also attempts to determine the role of neighborhood characteristics on student outcomes.

A secondary objective of this paper was to attempt to explain the discrepancy between official dropout rates reported by the TEA (TEA, 2001) and rates derived using appropriate and unique student data, such as the ones utilized here. I have shown that the GED program is an important variable in the differences, because dropouts with GEDs or students who enrolled in GED preparation programs are not included in the TEA dropout statistics. Previous research findings have strongly indicated that individuals with the GED credential are less successful in the labor market and have lower post-

secondary educational attainment than high school graduates; therefore, this study suggests that these dropouts should not be treated and viewed as the equivalent of high school graduates.

REFERENCES

Betts, J. R., & Lofstrom, M. (2000). The Educational attainment of immigrants: Trends and implications. In G. J. Borjas (Ed.), *Issues in the economics of immigration* (pp. 51-115). Chicago: University of Chicago Press, for National Bureau of Economic Research.

Bureau of Labor Statistics. (n.d.). *Quarterly census of employment and wages (QCEW)*. Retrieved October 13, 2007, from http://www.bls.gov/cew/

Cameron, S. V., & Heckman, J. J. (1993). The nonequivalence of high school equivalents. *Journal of Labor Economics*, *11*, 1–47.

Cameron, S. V., & Heckman, J. J. (1998). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy*, *106*, 262–333.

Cameron, S. V., & Heckman, J. J. (2001). The dynamics of educational attainment for Black, Hispanic, and White males. *Journal of Political Economy*, *109*, 455–499.

Colding, B. (2006). A dynamic analysis of educational progression of children of immigrants. *Labour Economics*, *13*, 479-492.

Greene, J. P. (2001, November). *High school graduation rates in the United States*. Retrieved May 1, 2007, from http://www.manhattan-institute.org/html/cr_baeo.htm

Hanushek, E. A. (2006). School resources. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (pp. 867–908). Amsterdam: Elsevier.

Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature*, *33*, 1829–1878.

Lavin-Loucks, D. (2006, July). *The academic achievement gap*. Dallas: J. McDonald Williams Institute.

Lofstrom, M., & Tyler, J. H. (2005). *Is the GED an effective route to postsecondary education?* Paper presented at the Association for Public Policy Analysis and Management Fall Research Conference 2005, Washington, DC.

National Center for Education Statistics (NCES). (n.d.) *Common core of data* (1996–2001). Retrieved October 13, 2007, from http://nces.ed.gov/ccd/

National Center for Education Statistics (NCES). (n.d.) *Local education agency finance survey (1996–2001)*. Retrieved October 13, 2007, from http://nces.ed.gov/ccd/f33agency.asp

Perreira, K., Harris, K. M., & Lee, D. (2006). Making it in America: High school completion by immigrant and native youth. *Demography*, *43*, 511–536.

Snower, D. J. (1999, January). Causes of changing earnings inequality. *IZA Discussion paper No. 29*. Retrieved May 1, 2007, from http://ssrn.com/abstract=166562

Texas Education Agency. (2001, August). *Secondary school completion and dropouts in Texas public schools 1999-00*. (Doc. No. GE01 601 09). Austin, TX: Author

Texas Education Agency. (n.d.). Website. http://www.tea.state.tx.us/

Tyler, J. H., Murnane, R. J., & Willett, J. B. (2000). Estimating the labor market signaling value of the GED. *Quarterly Journal of Economics*, *115*, 431–468.

U.S. Census Bureau. (n.d.) *United States Census 2000*. Retrieved October 13, 2007, from http://www.census.gov/main/www/cen2000.html

Valenzuela, A. (2006). The disappearance of high school English language learners from Texas high schools. *Williams Review*, *1*, 170–200.