8-1-2007

Differences in the College Enrollment Decision Across Race

Robert Baumann
College of the Holy Cross, rbaumann@holycross.edu

Differences in the College Enrollment Decision Across Race

By

Robert Baumann

August 2007

COLLEGE OF THE HOLY CROSS, DEPARTMENT OF ECONOMICS
FACULTY RESEARCH SERIES, PAPER NO. 07-05

Department of Economics
College of the Holy Cross
Box 45A
Worcester, Massachusetts 01610
(508) 793-3362 (phone)
(508) 793-3708 (fax)

http://www.holycross.edu/departments/economics/website

Differences in the College Enrollment Decision Across Race

By

Robert Baumann[1]
College of the Holy Cross

August 2007

Abstract

The gap in college enrollment rates between whites and blacks has remained stable since 1990, despite large increases in tuition and higher average wages for whites. We find the determinants of the enrollment decision differ greatly between whites and blacks, and within race between black males and females, but not between white males and females. These systematic differences require separate enrollment estimations for each race and for blacks each gender. Specifically, responses to changes in family income, parents' education, and school quality are vastly different across race-gender groups.

JEL Classification Codes: I21, J15

Keywords: college, enrollment, tuition, race, education

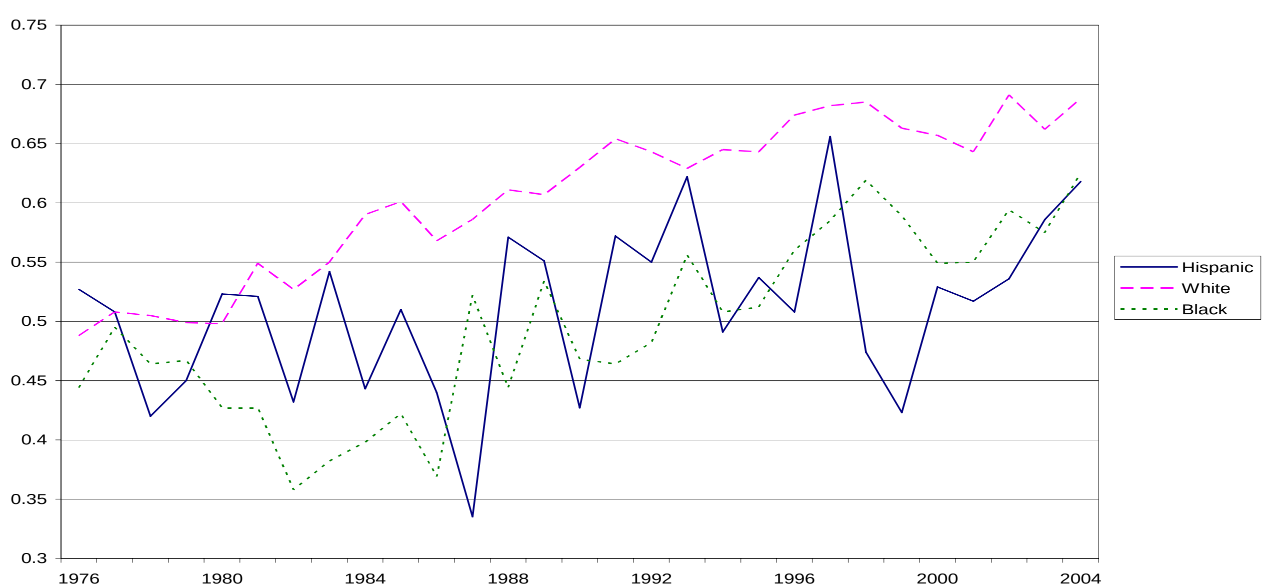
---

[1]Robert Baumann, Department of Economics, Box 192A, College of the Holy Cross, Worcester, MA 01610-2395, 508-793-3879 (phone), 508-793-3708 (fax), rbaumann@holycross.edu

INTRODUCTION

Educational attainment in the United States and several other developing countries has grown substantially in the past 50 years. Much of this growth is evident in the college enrollment data, which begins to rise in the United States around 1980. However, both the percent of those enrolled in college and the growth rates are not equal across races/ethnicities. Figure 1 graphs the percent of whites, blacks, and Hispanics between 18 and 24 who enrolled in college within 12 months of high school graduation using data from the National Center for Education Statistics (NCES). Figure 1 shows that black and Hispanic enrollment rates are consistently lower than white enrollment rates after 1980. While whites and blacks have roughly the same upward trend in education after 1990, while Hispanic enrollment rates are considerably more volatile. After hovering around 50 percent throughout the 1980s, Hispanic enrollment fluctuate around 55 percent after 1990 until the end of the sample frame.

Figure 1: Enrollment Rates



Source: National Center for Education Statistics

After the decision to attend college is made, the student must choose one institution from potentially thousands of options. One distinction among college and universities is public or private, which are imperfect substitutes. Public institutions tend to have lower costs but higher student body sizes. While larger schools tend to have more extracurricular opportunities because of scale economies (Koshal and Koshal, 2000), class sizes are also larger. Private schools usually offer smaller settings and foster student-faculty interaction, but at a higher cost.[1] Other motivations for attending a private school are religious affiliation and/or higher perceived quality; nearly all of the top 20 national universities in the *U.S. News and World Report* rankings are private.

Table 1 shows the number of white, black, and Hispanic students enrolled at a college or university, and the percent of each attending a private and public college. As expected, the level of enrollment has increased considerably for all three groups, especially blacks and Hispanics. However, the increase in enrollment at private institutions outpaced public enrollment growth for all three groups, especially blacks. The percent at private colleges exhibits the same pattern. For example, from 1976 to 2004 the number of blacks at private colleges nearly tripled (201,800 to 590,100), and the percent of black college students attending private institutions increased by almost eight percentage points (19.5% to 27.3%). In fact, the percent of black college students attending private institutions is larger than the percent of white college students attending private institutions after 2000, despite higher tuition levels and lower mean household earnings for blacks.

Many post-secondary schools have special programs to improve minority enrollment. These programs are particularly important for private institutions, which tend to have lower minority enrollment because of higher tuition. For public institutions, these programs combat criticism for under-representing minorities and those with low incomes relative to their state’s population. In

[1] A third option is an out-of-state public institution. Typically, students choosing an out-of-state public school pay tuitions on par with private schools.

addition, there are several government programs with the goal to increase educational attainment for those with below average earnings, e.g. subsidized student loans and Pell grants. Since there is a larger percentage of minorities in this population, these programs should help decrease the gap in earnings between whites and other minorities in the long run. Further, these programs are becoming increasingly more important as the wage gap between terminal high school and college graduates continues to grow.

This paper estimates a student choice model with three options – no college, public college, and private college – to investigate the differences in the college attendance decision between whites, blacks, and Hispanics. There are at least two well-known difficulties with student choice empirical models. First, the data set needs to be sufficiently large and detailed. Omitted variables are a common problem because the college attendance decision depends on a variety of factors. For example, it is well-known that the children of wealthier parents are more likely to attend better schools. Thus, if school quality is omitted from the model, then the effects of family background will be biased upward.[2] Second, the most common estimation techniques for models with more than two outcomes are not appropriate for student choice models because of the independence of irrelevant alternatives (IIA) assumption.

I address the data concern by using the geo-coded National Longitudinal Surveys of Youth 1997, which provides a sample of high school seniors from the graduating classes of 1998 through 2003. The NLSY97 is a recent cohort of college graduates compared to the NLSY79 or the National Education Longitudinal Study of 1988, which is important given the change in both college attendance and public/private behavior displayed in Figure 1 and Table 1. More importantly, the NLSY97 provides a wealth of data on student characteristics, family background characteristics,

[2] Alternatively, Card and Krueger (1996) note that many schools devote extra resources toward remedial programs to help their weaker students. If student ability is not accounted for, it can lead to a downward bias on the effect of school resources.

and school quality. The geo-coded augmentation is a confidential file of the survey that provides detailed geographic information, including the county and state of residence, and a variety of neighborhood characteristics such as mean wages and the unemployment and homeownership rates.

I address the estimation concern by estimating a multinomial probit, which does not require the IIA assumption. In addition, the multinomial probit is also consistent with utility maximizing behavior. I find that parents' education, income, grades, and enrollment in a college prep program, impact the college attendance decision for blacks, Hispanics, and whites, but the sensitivities to these controls vary across these three groups. Any estimation that does not take this heterogeneity into account can cause misleading results for minorities since whites constitute the largest portion of the population. I also find that little is known about the decision to attend a private college for blacks, and to a lesser extent Hispanics. While preferential financial aid for minorities that is common at many colleges and universities likely impacts the individual's decision, even the rich set of observable characteristics in the NLSY97 fails to explain the large increase in private college enrollment among blacks and Hispanics.

LITERATURE REVIEW

Perhaps the most commonly cited student choice study is Fuller, Manski, and Wise (1982), which uses an expected utility setup to arrive at a multinomial logit estimation where respondents choose one of several post-secondary options. Later studies tended to focus on one aspect of the college enrollment decision: tuition, financial aid, race, school quality, and parent's education, to name a few. The effect of rising tuition on enrollment rates (i.e., the price sensitivity) receives the most attention in the literature, although the effects of financial aid, school quality, and parent's education, are also studied. Leslie and Brinkman (1987) surveys 25 of these studies and notes nearly all of them find a negative and statistically significant relationship between tuition and college enrollment. After standardizing the results of each work, Leslie and Brinkman find a mean price response of -0.7 percentage points for every $100 (in 1982-83 dollars) increase in tuition. Contrary to these findings, enrollment rates grew while tuition rose during the early 1980s. Heller (1997) updates Leslie and Brinkman, summarizing this later cohort of college enrollment studies. A large majority of papers in the Heller survey again find a negative relationship between tuition and college enrollment at roughly the same magnitude as Leslie and Brinkman.

A selection of the studies reviewed by Heller focuses on variation in tuition and financial aid sensitivities across race (Behrman et al., 1992; St. John and Noell, 1989; Jackson, 1989; Heller, 1994). From these papers, Heller concludes minorities are more sensitive to changes in tuition and price, and offers three potential explanations: 1) minority races are more price sensitive because they tend to have a lower family income;[3] 2) minorities are less willing to make financial sacrifices because they do not picture themselves as "college material"; and 3) there are different social values of attending college across races. Similar to this research, T. Kane (1994) attempts to explain the decline and subsequent rebound of black college enrollment in the early 1980s, and also notes the conflicting effects of rising tuition and increases in parental education on enrollment rates. J. Kane and Spizman (1994) find that preferential financial aid increased black college enrollment, and also point to differences in parental education. Granderton and Santos (1995) investigate Hispanic college enrollment using the High School and Beyond survey of 1980 high school graduates. They find that whites, blacks, and Hispanics have different responses to socioeconomic status.

The Supreme Court decision in *University of California Regents v. Bakke* (1978) sanctioned the use of race information in college admissions decisions, although some states have since prohibited this practice. Light and Strayer (2000) use a sequential model of college attendance and

[3] Kane (1994) and Haveman and Wolfe (1995) find that black college enrollment rates are pulled in two different directions by increases in the real cost of college and improvements in black parental education and income levels.

college graduation, and find that minorities are more likely than whites to attend and graduate college, ceteris paribus. This finding suggests that affirmative action programs increase minority enrollment, but because the obvious identification problem the authors are hesitant to state affirmative action programs increased minority enrollment. This paper and Light and Strayer (2002) use a multinomial probit, which is discussed in the next section.

Other studies focus on the differences between black and white high school students. Hurtado, et al. (1997) finds differences in college expectations, preparation, and application across races and ethnicities. Perna (2000) also finds differences in measures of social and cultural capital, such as educational expectations, parental encouragement, and parental involvement, each of which increase the probability of enrollment. These differences may indicate a college information gap for some minorities. Freeman (1997) gives the example of financing, in particular whether future earnings would offset the cost of college (tuition and opportunity). Lastly, Lucia and Baumann (2007) find differences between recent black and white high school graduates in the marginal effects of many of the explanatory variables in the college attendance decision.

Much of the literature on the impact of school quality on wages and educational attainment is also relevant here. School quality is notoriously difficult to quantify. The most common proxies include student-teacher ratios, per-pupil expenditures, length of school year, teacher salary, and teacher education. The literature is divided on the impact of school quality on educational attainment. Eric Hanushek, who has written a great deal on school quality, argues there is no systematic relationship between student performance and teacher-student ratio or teacher education (Hanushek 1986, 1996), although he does find evidence that teacher skill positively impacts student performance on standardized tests (Hanushek, 1992). David Card and Alan Krueger find school quality inputs do impact wages and educational attainment, partly because school quality impacts the marginal benefit of an additional year of schooling (Card and Krueger, 1992). Card and Krueger (1996) compares blacks and whites during and after the segregation era and finds school quality improvements for blacks after segregation ended led to increases in wages and educational attainment.

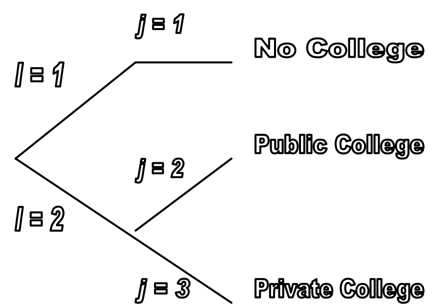
EMPIRICAL MODELS

Each observation chooses between three options: not attending college (*j* = 1), attending a public university (*j* = 2), and attending a private university (*j* = 3). Utility maximization determines the choice for each consumer. The random-utility model approach splits utility into observable and unobservable portions, and compares the differences between expected utilities (Manski, 1977). For utility $U_{i,j} = \vec{X}_i \vec{\alpha}_j + T_{i,j}\beta + \varepsilon_{i,j}$, an individual *i* chooses alternative *j* if $U_{i,j} = \max\{U_{i,j} \mid j \in J\}$. $\vec{X}_i$ is a vector of *k* variables that vary by individual such as grades, family background controls, and high school characteristics. $T_{i,j}$ represents tuition, which varies by individual and alternative. While $\vec{\alpha}_j$ is a vector of coefficients that is specific to each alternative *j*, β is not specific to each alternative because $T_{i,j}$ varies by individual and alternative.

The most common estimation techniques for discrete choice analyze the probability that one option is chosen over the other given some assumption about the error structure. One possibility is the conditional logit, which assumes the errors are have an extreme value distribution and are independently and identically distributed. Whether the conditional logit is applicable depends on the independence of irrelevant alternatives (IIA) assumption, which states the relative probability of any two outcomes is independent of the number of choices. Unfortunately, this is often an unreasonable assumption in student choice models. For example, suppose the private college option is removed from the above three-choice model. IIA requires that private college attendees will be evenly distributed among the remaining two options (no college and public college) to keep the relative probability constant, but it is more likely that the vast majority of private college attendees will choose a public college over no college. In addition, any correlation across choice-specific errors is a sufficient condition for a violation of IIA since the errors are assumed to be independent.

One alternative is the nested logit, which relaxes IIA for similar alternatives. In Figure 2, *l* represents the decision to attend college, and *j* indexes the decision to attend a public or private college. In this setup, IIA must hold across nests (in particular, between *l* = 1 and *l* = 2), but not within nests.

Figure 2: Nested Logit Decision Tree



Since no further decision is made if an observation chooses no college, *l* = 1 is called a degenerative nest.

Another alternative is the multinomial probit, which does not require the IIA assumption because it estimates the error covariance structure. Specifically, $[\vec{\varepsilon}_{j=0}, \vec{\varepsilon}_{j=1}, \vec{\varepsilon}_{j=2}] \sim N(0, \Sigma)$, where

$$\Sigma = \begin{bmatrix} \sigma_{11} & . & . \\ \sigma_{21} & \sigma_{22} & . \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$$

Although utility is a latent variable, it is possible to make inferences about utility differences. Let $v_{i,j,j'}$ be the utility difference between alternatives *j* and *j'*, or $v_{i,j,j'} = U_{i,j} - U_{i,j'}$. Substituting an expression for the utility of each alternative and simplifying yields

$$v_{i,j,j'} = \vec{X}_i(\vec{\alpha}_j - \vec{\alpha}_{j'}) + (T_{i,j} - T_{i,j'})\beta + \xi_{i,j,j'}, \quad (1)$$

where $\xi_{i,j,j'} = \varepsilon_{i,j} - \varepsilon_{i,j'}$. For example, the probability that individual *i* chooses no college (*j* = 1) is $\Pr(j=1) = \Pr(v_{i,j=2,j=1} < 0, v_{i,j=3,j=1} < 0)$, where $\xi_{i,j,j'}$ is normally distributed because it is the difference between two normally distributed variables. This probability is used in the maximum likelihood function, where each $\xi_{i,j',j=1}$ term is compared to the remaining terms in (1).

Identification depends on the normalizations imposed on Σ. I use the normalizations suggested by Train (2003), although there are other options. Suppose utility differences are always taken against the no college (*j* = 1) alternative. This is called the base, and leaves two terms for the differenced error ($\xi_{i,j=2,j=1}$ and $\xi_{i,j=3,j=1}$) with the following variance-covariance matrix:

$$\Omega = \begin{bmatrix} \omega_{22} & . \\ \omega_{32} & \omega_{33} \end{bmatrix},$$

where $\omega_{22} = \sigma_{22} + \sigma_{11} - 2\sigma_{21}$, $\omega_{22} = \sigma_{22} + \sigma_{11} - \sigma_{31} - \sigma_{21}$, and $\omega_{33} = \sigma_{33} + \sigma_{11} - 2\sigma_{31}$. After normalizing Ω by ω_{22},

$$\Omega* = \begin{bmatrix} 1 & . \\ \omega_{32}^{*} & \omega_{33}^{*} \end{bmatrix},$$

where $\omega_{32}^{*} = \dfrac{\sigma_{32} + \sigma_{11} - \sigma_{31} - \sigma_{21}}{\sigma_{22} + \sigma_{11} - 2\sigma_{21}}$ and $\omega_{33}^{*} = \dfrac{\sigma_{33} + \sigma_{11} - 2\sigma_{31}}{\sigma_{22} + \sigma_{11} - 2\sigma_{21}}$.

Because only ω_{32}^{*} and ω_{33}^{*} are estimated, not all six of the unique elements in Σ are identified. As Train (2003) points out, this is not a restriction of the model because Σ contains information about the size and scale of utility, which is irrelevant for behavior. For example, if utility for each alternative is added or multiplied by some constant, behavior would not change because the utility ranking is the same. Only differences in utility impact behavior, which is why $\Omega*$ is the relevant variance-covariance matrix and not Σ. However, one drawback is that the elements of $\Omega*$ do not have a straight-forward interpretation unlike the elements of Σ. Finally, note that not all three $\vec{\alpha}_j$ are identified because only the differences relative to the base ($\vec{\alpha}_{j=2} - \vec{\alpha}_{j=1}$ and $\vec{\alpha}_{j=3} - \vec{\alpha}_{j=1}$) are estimated. But again, this is irrelevant for estimation since only differences in utility impact behavior.

One difficulty is the lack of a closed-form solution. The GHK algorithm (Geweke, 1989; Hajivassiliou and McFadden, 1998; and Keane, 1994) is used to approximate the multivariate normal distribution. The GHK algorithm uses the Cholesky factorization of Ω^* to rewrite $\xi_{i,j=1,j=0}$ and $\xi_{i,j=2,j=0}$ in terms of uncorrelated standard normal variables which are simulated. This is advantageous because the simulated draws do not require an a priori assumption of the error correlations, and also because the Cholesky decomposition ensures that Ω^* will be positive definite.

NLSY97 DATA

The data set is the geo-coded National Longitudinal Surveys of Youth 1997 (NLSY97). The NLSY97 provides a large and recent sample, and the geo-coded augmentation adds to an already data rich set. The NLSY97 is a sample of 8,984 people born between 1980 and 1984. Each sample member is interviewed annually beginning in 1997, which produces a sample of high school seniors from the graduating classes of 1998 through 2003.

To create a sample of high school graduates, I omit those who did not complete a high school diploma because the enrollment rate for this population is zero. I also drop those who receive a general equivalency diploma because of the evidence that this population differs substantially from high school graduates (Cameron and Heckman, 1993). Although the NLSY97 is a panel data set, I only use data from the student’s final year of high school. This captures the most recent set of information, and also includes time invariant factors that impact the college decision such as parent's education level. Identifying the year involves comparing the high school graduation date to the interview dates.[4]

I define enrollment as attending either a two- or four-year postsecondary institution within four months of high school graduation.[5] Although there are respondents who will enroll in college years after completing high school, their decision to attend college after, say, a spell in the labor force, is not the same as the enrollment decision soon after completing high school. For example, it is likely that the influence parents' income, grades, standardized test scores, etc. diminishes as the respondent waits to attend college. Light (1995) details this 'interrupted schooling' decision and finds substantial differences between those who return to school after a spell in the labor force and those who do not. In addition, there appears to be little sensitivity in the results to changes in the enrollment cutoff to 6, 12, or 18 months. Finally, college attendees are further separated into those that attended a public college and those that attended a private college.

Table 2 presents brief descriptions and sample means of each variable. After removing All variables are from the NLSY97 except for tuition. Tuition data are from the Digest of Education Statistics, which provides annual data on the state average of tuition, room, and board for full time students. These tuition figures are available at the state-level and separately for public and private colleges, and are matched to NLSY97 respondents using their state of residence. After making the sample deletions listed above and removing those with missing data, there are 3,079 respondents in the sample which includes 683 blacks, 501 Hispanics, and 1895 whites.[6]

RESULTS

[4] In the NLSY97, interviews typically happen between November and January. In my data set, the data are typically from this period during senior year for an observation that graduates in May or June.

[5] I pool two-year and four-year college attendees for two reasons. First, I am not modeling college completion (e.g., Light and Strayer, 2000), and often attending a two-year institution is a stepping stone to a four-year institution. Second, I attempted to estimate a separate "attend a two-year college" decision, but these results did not provide any additional insight.

[6] The NLSY97 classifies all respondents into non-overlapping race/ethnicity categories. One of these categories is mixed race/non-Hispanic. These respondents are omitted.

Table 3 presents the estimates from a binary probit where the alternatives are no college attendance and college attendance. These results are presented for comparison with the three alternative model estimates forthcoming in Table 4. In general, the estimates from the binary probit have the expected sign. High school grades, household income, enrollment in a college prep program, ASVAB score, and whether each parent attended college all have positive impacts on college attendance. In addition, blacks and females are more likely to attend college than whites and males, respectively, holding the other variables constant. However, the number of AP courses, student-teacher ratio, and public tuition do not have statistically significant effects.

Not surprisingly, a Hausman (1978) specification test suggests the conditional logit estimation of the three alternative model does not satisfy the IIA assumption. The Hausman test compares the estimates of the full model to a second set of estimates where one of the alternatives is removed. The nested logit, which groups the public and private college options as in Figure 2, allows correlation between these two error terms. However, IIA must hold across nests, or between the no college ($l = 1$) and college ($l = 2$). Nevertheless, a Hausman test that removes the no college option suggests the nested logit does not satisfy IIA. Because the Hausman tests suggest the conditional and nest logits are not appropriate for the model, I omit their results.

Table 4 presents the estimates for from the multinomial probit estimation of the three-alternative model, where the base is the no college ($j = 1$) alternative. Many of the statistically significant variables from the binary probit remain important in the multinomial probit estimation. While the private college estimates are generally larger in absolute value, the confidence intervals between the public and private college estimates overlap. Therefore, this model separates public attendees from those that don’t attend college and also private attendees from those that don’t attend college, but there is little that can be said about the public/private decision.

Table 5 presents results from a binary probit estimation on three sub-samples of the data: blacks, Hispanics, and white. The impetus for this cut is twofold. First, the binary probit using the pooled sample suggests blacks are more likely to attend college, ceteris paribus, while Hispanics have a weakly negative effect. More importantly, these minority groups are targeted by colleges and governments, who use financial aid packages to entice enrollment. There are likely several other interesting cuts of these data, e.g. male and female, but I limit the focus to these three racial/ethnic groups because of existing college and government policy.

These results exhibit the problems of interpreting estimates from data that pools blacks, Hispanics, and whites together. For example, black college enrollment is influenced by the educational attainment of the mother but not the father, while white college enrollment is affected by the educational attainment of the father but not the mother. Meanwhile, parents' educational attainment has only a weakly positive effect on Hispanic college attendance. Not all of this difference is a result of fewer black fathers in the household. In fact, including this control does not substantially change the marginal effects of parents' education for any of the three samples. A second difference across these groups is that whites have larger responses to increases in grades and enrollment in a college prep program. However, the effects of household income and tuition do not have large differences, suggesting that there are not different responses to these proxies of affordability.

Tables 6, 7, and 8 present the estimates from separate multinomial probit estimations for blacks, Hispanics, and whites, respectively. These estimates show that little is known about the private college decision for blacks and Hispanics. None of the estimates for private college choice are statistically significant for the black sample. For Hispanics, an increase in grades or the ASVAB score has a larger marginal effect on private college than public college choice. However, several of the coefficients for public and private choice are statistically significant in the white sample. Similar to the pooled estimation, the private college estimates for whites are generally larger in absolute value, but there is overlap in the confidence intervals for the public and private college estimates. Finally, the effect of tuition is not statistically significant for all three groups, and is even positive for the black and Hispanic sample, suggesting the difference between public and private tuition does not impact the attendance decision for any of the three groups.

These estimates illustrate a subtle difference between the three samples. For controls that only vary individual (all expect tuition), the multinomial probit estimates the difference between the coefficients of each option, or $\vec{\alpha}_j - \vec{\alpha}_{j'}$. Given that none of the private college estimates are statistically significant, this means none of the controls differentiate between blacks that do not attend college and blacks that attend private colleges in the presence of a public college option. However, several of the controls differentiate between blacks that do not attend college and blacks that attend public colleges, namely female, ASVAB, grades, income, and whether the student was in a college prep program in high school. A similar statement can be made for Hispanics, although to a lesser extent because two of the controls are significant for the private college option. For whites, the problem is differentiating between public and private college attendees. But of all three alternatives, a priori this pair is the most likely to be positively correlated.

Table 9 presents selected marginal effects from each of the three multinomial probit models. The marginal effects are calculated using the means of each sample. In the case of binary independent variables, the marginal effects are the difference in the predicted probabilities caused by changing the binary variable from zero to one. Standard errors of the marginal effects are calculated using the delta method. Nearly all of the private college marginal effects are smaller in absolute value than the public college marginal effects, but this is partly because the predicted probability of private college is smaller than the predicted probability of public college.

For the black sample, the marginal effects reinforce several of the points above. Several of the controls have statistically significant marginal effects on public college attendance, but the controls are not good predictors for black private college choice. For Hispanics, none of the controls have statistically significant marginal effects. In other words, the multinomial probit can identify statistically significant effects relative to the base (as in Table 7), but not in the marginal effects of an individual choice.

The marginal effects for the white sample reveal some interesting aspects of college choice. Both public and private tuition have statistically significant effects, but have the opposite of the expected sign. For example, an increase in public tuition has a positive effect on choosing a public college and a negative effect on choosing a private college. The same is true for private tuition; higher private tuition increases the probability of choosing a private college. It is possible that students use quality to determine the public/private decision. For example, students in states with high quality public institutions (which also are likely to have above-average public tuition levels) are more likely to choose the public option because the quality-adjusted tuition is lower than the private options. Unfortunately, attempts to include proxies for public college quality do not improve the fit of the model.[7] This is likely a byproduct of imperfectly measuring school quality; a problem that has generated its own literature and debates.

CONCLUSIONS

[7] I use data from the *U.S. News & World Report* College Rankings to construct the average public college ranking and minimum ranking (i.e., the ranking of the 'best' public school) for each state. Neither of these controls, whether included separately of jointly, are statistically significant and do not substantially impact the log likelihood at the estimated coefficients.

The results show that black, Hispanic, and white high school graduates differ greatly in the college enrollment decision. Some of these differences are attributable to observable factors, but there is much that is unknown about the college attendance and public/private decisions for blacks and Hispanics. Because blacks and Hispanics are minorities, failure to account for this heterogeneity produces misleading results for these groups. Dichotomous controls for race or ethnicity are not sufficient because minorities have different responses to observable factors. In addition, it seems likely that other ethnicities, such as Asian-Americans, also have unique college decision processes. While these differences complicate policy, they are likely to improve minority attendance rates.

More importantly, these findings suggest the decision to attend college is made long before the senior year of high school, regardless of race or ethnicity. The most influential factors, such as parents' education, income, grades, and enrollment in a college prep program, are reflections of household preferences for education rather than responses to changes in exogenous factors. For white high school students, these factors also influence the public/private decision. However, these factors do not influence the public/private decision for minority students, despite increasing private enrollment for blacks and Hispanics in both absolute and percentage terms.

One missing piece is the influence of preferential financial aid for minorities, which are common at many colleges and universities. These are especially prevalent at private institutions, which have a harder time attracting minorities in part because of higher tuition levels. This research does not suggest these programs are ineffective for individual institutions. Rather, these findings suggest that average tuition of public and private institutions does not impact the amount of college-bound high school seniors. In other words, preferential financial aid at individual institutions affects how minorities sort into colleges, but does not increase the amount of college-bound students.

REFERENCES

Behrman, J.R., Kletzer, L.G., McPherson, M.S., and M.O. Schapiro. 1992. "The College Investment Decision: Direct and Indirect Effects of Family Background on Choice of Postsecondary Enrollment and Quality." Williamstown, MA: Williams Project on the Economics of Higher Education Discussion Paper No. 18.

Cameron, S. and J. Heckman. 1993. "The Nonequivalence of High School Equivalence." *Journal of Labor Economics* 11 (1): 1-47.

Card, D. and A. B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools." *The Journal of Political Economy* 100 (1): 1-40.

Card, D. and A. B. Krueger. 1996. "Labor Market Effects of School Quality." In Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success edited by Gary Burtless. p. 97-140. Brookings Institution Press, Washington, D.C.

Freeman, K. 1997. "Increasing African Americans' Participation in Higher Education: African American High-School Students' Perspectives. *Journal of Higher Education* 68: 523-550.

Fuller, W., Manski, C., and D. Wise. 1982. "New Evidence on the Economic Determinants of Post-Secondary Schooling Choices." *Journal of Human Resources* 17 (4): 477-498.

Geweke, J. 1989. "Bayesian Inference in Econometric Models Using Monte Carlo Integration." *Econometrica* 57: 1317-1339.

Granderton, P. T. and R. Santos. 1995. "Hispanic College Attendance and Completion." *Economics of Education Review*, 14(1): 35-46.

Hajivassiliou V. and D. McFadden. 1998. "The Method of Simulated Scores for the Estimation of LDV Models." *Econometrica* 66: 863-896.

Hanushek, E. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24 (3): 1141-1177.

Hanushek, E. 1992. "The Trade-Off Between Child Quantity and Quality." *Journal of Political Economy* 100 (1): 84-117.

Hanushek, E. 1996. "School Resources and Student Performance." In Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success edited by Gary Burtless. p. 43-73. Brookings Institution Press, Washington, D.C.

Hausman, J. 1978. "Specification Tests in Econometrics." *Econometrica* 49: 1377-1398.

Haveman, R. and B. Wolfe. 1995. "The Determinants of Children's Attainments: A

Review of Methods and Findings." *Journal of Economic Literature,* 33(4) (December): 1829-1878.

Heller, D. E. 1994. "Tuition Prices, State Grants, and Public College Enrollments: Are Minority Youth Losing Ground?" Cambridge, MA: Monograph, Harvard Graduate School of Education.

Heller, D. E. 1997. "Student Price Response in Higher Education: An Update to Leslie and Brinkman." *Journal of Higher Education,* 68(6) (November-December): 624-659.

Hurtado, S., Inkelas, K. K., Briggs, C., & Rhee, B. S. 1997. Differences in College Access and Choice Among Racia/Ethnic Groups: Identifying Continuing Barriers. *Research in Higher Education* 38 (1): 43-75.

Keane, M. 2003. "A Computationally Practical Simulation Estimator for Panel Data." *Econometrica* 62: 95-116.

Koshal, R.K. and M. Koshal. 2000. "Do Liberal Arts Colleges Exhibit Economies of Scale and Scope?" *Education Economics*, 8(3): 209-220.

Jackson, G.A. 1989. "Responses of Black, Hispanic, and White Students to Financial Aid: College Entry Among Recent High School Graduates." College Park, MD: National Center for Postsecondary Governance and Finance, University of Maryland.

Kane, T. J. 1994. "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education." *Journal of Political Economy*, 102(5) (October): 878-911.

Kane, J. and L.M Spizman. 1994. "Race, Financial Aid Awards, College Attendance." *American Journal of Economics and Sociology*, 53(1) (January): 85-97.

Leslie, L. L. and P. T. Brinkman. 1987. "Student Price Response in Higher Education: The Student Demand Studies." *Journal of Higher Education*, 58(2) (March-April): 181-204.

Light, A. 1995. "The Effects of Interrupted Schooling on Wages." *Journal of Human Resources* 30 (3): 472-502.

Light, A. and W. Strayer. 2000. "Determinants of College Completion: School Quality or Student Ability?." *Journal of Human Resources* 35 (2): 299-332.

Light, A. and W. Strayer. 2002. "From Bakke to Hopwood: Does Race Affect College Attendance Student Ability?." *The Review of Economics and Statistics* 84 (February): 34-44.

Lucia, K. and R. Baumann. 2007. "Differences in the College Enrollment Decision Across Race." Manuscript, College of the Holy Cross.

Manski, C. 1977. “The Structure of Random Utility Models,” Theory and Decision, Vol. 8, 229-254.

Perna, L.W. 2000. “Differences in the Decision to Attend College among African Americans, Hispanics, and Whites.” *The Journal of Higher Education* 71 (2): 117-141.

St. John, E.P, and J. Noell. 1989. “The Effects of Student Financial Aid on Access to Higher Education: An Analysis of Progress with Special Consideration of Minority Enrollment.” *Research in Higher Education*, 30(6): 563-581.

Train, K. 2003. “Discrete Choice Methods With Simulation” Cambridge, U.K: Cambridge University Press.

TABLE 1: PUBLIC AND PRIVATE COLLEGE ENROLLMENT

	White, non-Hispanic	*Black, non-Hispanic*	*Hispanic*
1976			
public enrollment	7,094,500 (78.2%)	831,200 (80.5%)	336,800 (87.8%)
private enrollment	1,981,600 (21.8%)	201,800 (19.5%)	47,000 (12.2%)
1980			
public enrollment	7,656,100 (77.9%)	876,100 (79.2%)	406,200 (86.1%)
Private enrollment	2,176,900 (22.1%)	230,700 (20.8%)	65,600 (13.9%)
1990			
Public enrollment	8,385,400 (78.2%)	976,400 (78.3%)	671,400 (85.8%)
Private enrollment	2,337,000 (21.8%)	270,600 (21.7%)	111,000 (14.2%)
2000			
public enrollment	7,963,400 (76.1%)	1,319,200 (76.2%)	1,229,300 (84.1%)
private enrollment	2,498,700 (23.9%)	411,100 (23.8%)	232,500 (15.9%)
2001			
public enrollment	8,214,000 (76.2%)	1,397,100 (75.5%)	1,308,800 (83.9%)
private enrollment	2,560,500 (23.8%)	453,300 (24.5%)	251,800 (16.1%)
2002			
public enrollment	8,490,500 (76.2%)	1,487,200 (75.2%)	1,388,700 (83.6%)
private enrollment	2,649,800 (23.8%)	491,600 (24.8%)	273,100 (16.4%)
2003			
public enrollment	8,531,400 (75.7%)	1,533,500 (74.1%)	1,414,600 (82.4%)
private enrollment	2,744,100 (24.3%)	535,300 (25.9%)	301,400 (17.6%)
2004			
public enrollment	8,546,300 (74.8%)	1,574,600 (72.7%)	1,477,400 (81.6%)
private enrollment	2,876,500 (25.2%)	590,100 (27.3%)	332,200 (18.1%)

Source: National Center for Education Statistics

TABLE 2: SAMPLE MEANS

Variable	Description	sample mean (sample s.d.)
No College	did not enroll within four months of college graduation	0.380
Attends Public College	enrolled in public college within four months of college graduation	0.494
Attends Private College	enrolled in private college within four months of college graduation	0.126
Black	=1 if black, =0 otherwise	0.222
Hispanic	=1 if Hispanic, =0 otherwise	0.163
female	=1 if female, =0 otherwise	0.518
ASVAB	percentile score on math and verbal sections of Armed Services Vocational Aptitude Battery, adjusted for age	54.19 (28.0)
Mother attended college	=1 if mother attended college, =0 otherwise	0.508
Mother education data missing	=1 if education data is missing for the mother, =0 otherwise	0.023
Father attended college	=1 if father attended college, =0 otherwise	0.505
Father education data missing	=1 if education data is missing for the father, =0 otherwise	0.089
grades	=1 if "mostly below D's" =2 if "mostly D's" =3 if "about half C's and half D's" … and so on until… =9 if "mostly A's"	6.08 (1.38)
household income, first quartile	household income less than $25,000 in 1997	0.235
household income, second quartile	household income between $25,000 and $50,000 in 1997	0.316
household income, third quartile	household income between $50,000 and $100,000 in 1997	0.353
household income, fourth quartile	household income greater than $100,000 in 1997	0.096
college prep program	=1 if enrolled in college prep, =0 otherwise	0.444
completed AP course	=1 if completed at least one, =0 otherwise	0.117
student-teacher ratio	=1 if student-teacher ratio less than 14 =2 if student-teacher ratio between 14 and 18 =3 if student-teacher ratio between 18 and 22 =4 if student-teacher ratio greater than 22	2.12 (1.02)
public tuition	state average of full-time public tuition, includes room and board	8621.80 (1590.71)
private tuition	state average of full-time private tuition, includes room and board	19463.08 (4026.25)

TABLE 3: BINARY PROBIT ESTIMATION

variable	**Estimate (standard error)**
black	0.1364* (0.0703)
Hispanic	-0.0568 (0.0742)
female	0.1798*** (0.053)
ASVAB	0.0104*** (0.0012)
Mother attended college	0.2095*** (0.0573)
Mother education data missing	-0.2940* (0.1682)
Father attended college	0.1991*** (0.0615)
Father education data missing	-0.2932*** (0.0965)
grades	0.1902*** (0.0217)
household income, second quartile	0.1257* (0.0670)
household income, third quartile	0.3397*** (0.0739)
household income, fourth quartile	0.8486*** (0.1281)
college prep program	0.4747*** (0.0557)
completed AP course	0.0911 (0.0868)
student-teacher ratio	-0.0374 (0.0253)
public tuition	0.000018 (0.000016)
constant	-2.1345*** (0.2122)

* indicates a *p*-value between 0.05 and 0.1.
** indicates a *p*-value between 0.01 and 0.05.
*** indicates a *p*-value below 0.01.

TABLE 4: MULTINOMIAL PROBIT ESTIMATION

variable	public college $(\alpha_{j=2} - \alpha_{j=1})$	private college $(\alpha_{j=3} - \alpha_{j=1})$
black	0.1832* (0.0994)	0.1834* (0.1047)
Hispanic	-0.0611 (0.1048)	-0.0559 (0.1113)
female	0.2560*** (0.0746)	0.2761*** (0.1048)
ASVAB	0.0149*** (0.0015)	0.0163*** (0.0049)
Mother attended college	0.2974*** (0.0808)	0.3183*** (0.1108)
Mother education data missing	-0.4027* (0.2367)	-0.3876 (0.2566)
Father attended college	0.2951*** (0.0867)	0.3427* (0.1843)
Father education data missing	-0.4355*** (0.1360)	-0.4784** (0.2059)
grades	0.2683*** (0.0308)	0.2948*** (0.0985)
household income, second quartile	0.1662* (0.0976)	0.1577 (0.1063)
household income, third quartile	0.4750*** (0.1043)	0.4653** (0.1159)
household income, fourth quartile	1.1344*** (0.1804)	1.2392*** (0.2019)
college prep program	0.6810*** (0.0788)	0.7104*** (0.1233)
completed AP course	0.1597 (0.1224)	0.1726 (0.1358)
student-teacher ratio	-0.0656* (0.0359)	-0.0901 (0.0897)
constant	-2.8545*** (0.3149)	-3.4900 (2.1325)
tuition (β)	0.0000072 (0.0000246)	
ω_{32}^{*}	-2.0410 (3.4900)	
ω_{33}^{*}	1.5564*** (0.4761)	

* indicates a *p*-value between 0.05 and 0.1, ** indicates a *p*-value between 0.01 and 0.05, and *** indicates a *p*-value below 0.01.

TABLE 5: BINARY PROBIT ESTIMATION, BY RACE/ETHNICITY

variable	black	Hispanic	white
female	0.3013*** (0.1108)	0.1330 (0.1242)	0.1511** (0.0699)
ASVAB	0.0121*** (0.0026)	0.0069*** (0.0026)	0.0106*** (0.0015)
Mother attended college	0.4818*** (0.1186)	0.2005 (0.1396)	0.1063 (0.0754)
Mother education data missing	-0.3196 (0.2693)	0.0961 (0.5560)	-0.3851 (0.2391)
Father attended college	0.0412 (0.1451)	-0.0239 (0.1463)	0.3069*** (0.0789)
Father education data missing	-0.0132 (0.1899)	-0.3800 (0.2430)	-0.3753*** (0.1332)
grades	0.1214*** (0.0457)	0.1866*** (0.0528)	0.2114*** (0.0283)
household income, second quartile	0.0893 (0.1215)	0.1426 (0.1441)	0.1213 (0.1082)
household income, third quartile	0.4142*** (0.1561)	0.4387*** (0.1603)	0.3052*** (0.1079)
household income, fourth quartile†	0.6912** (0.1804)		0.8028** (0.1552)
college prep program	0.4121*** (0.1158)	0.3368*** (0.1313)	0.5518*** (0.0740)
completed AP course	0.2534 (0.1765)	0.0662 (0.1935)	0.0283 (0.1118)
student-teacher ratio	-0.0747 (0.0553)	-0.0474 (0.0556)	-0.0237 (0.0338)
public tuition	-0.000013 (0.000034)	0.0000005 (0.00004)	0.000032 (0.000022)
constant	-1.4972*** (0.4097)	-1.7014*** (0.5037)	-2.4146*** (0.2777)
log L	-387.5367	-303.8934	-884.1443
N	683	501	1895

† All Hispanics in the fourth quartile of household income (36 out of 489) attend college. This variable is dropped from the estimation, and Hispanics in the fourth quartile of household income are grouped with those in the third quartile.

TABLE 6: MULTINOMIAL PROBIT ESTIMATION, BLACK SAMPLE

variable	public college $(\alpha_{j=2} - \alpha_{j=1})$	private college $(\alpha_{j=3} - \alpha_{j=1})$
female	0.4541*** (0.1605)	0.0690 (0.6429)
ASVAB	0.0175*** (0.0036)	0.0290* (0.0171)
Mother attended college	0.7107*** (0.1693)	1.2709 (0.8083)
Mother education data missing	-0.4611 (0.3915)	0.3355 (1.3629)
Father attended college	0.0542 (0.2062)	0.7701 (0.9100)
Father education data missing	-0.0359 (0.2694)	-0.4334 (0.9610)
grades	0.1606** (0.0651)	0.5737 (0.4473)
household income, second quartile	0.1245 (0.1725)	0.1767 (0.6280)
household income, third quartile	0.5884*** (0.2223)	0.9180 (0.7798)
household income, fourth quartile	1.1018** (0.4959)	1.7372 (1.4451)
college prep program	0.5683*** (0.1679)	0.5299 (0.5514)
completed AP course	0.3993 (0.2486)	0.1133 (0.8811)
student-teacher ratio	-0.1272 (0.0788)	0.1237 (0.3380)
constant	-1.8846*** (0.5662)	-9.6087 (7.5321)
tuition (β)	-0.000040 (0.000045)	
ω_{32}^{*}	0.7171 (1.1169)	
ω_{33}^{*}	3.6528 (2.7642)	

* indicates a *p*-value between 0.05 and 0.1, ** indicates a *p*-value between 0.01 and 0.05, and *** indicates a *p*-value below 0.01.

TABLE 7: MULTINOMIAL PROBIT ESTIMATION, HISPANIC SAMPLE

variable	public college $(\alpha_{j=2} - \alpha_{j=1})$	private college $(\alpha_{j=3} - \alpha_{j=1})$
female	0.1458 (0.1846)	0.3789 (0.3282)
ASVAB	0.0087** (0.0041)	0.0166* (0.0089)
Mother attended college	0.2909 (0.2007)	0.2246 (0.3089)
Mother education data missing	0.1863 (0.8007)	-4.7451 (5.4485)
Father attended college	-0.0560 (0.2258)	0.2872 (0.3741)
Father education data missing	-0.5088 (0.3647)	-0.9761 (0.7482)
grades	0.2491*** (0.0828)	0.4325** (0.2093)
household income, second quartile	0.2198 (0.2076)	0.1784 (0.3148)
household income, third quartile	0.6597*** (0.2313)	0.4596 (0.3610)
college prep program	0.4796*** (0.1874)	0.3520 (0.2997)
completed AP course	0.1092 (0.2759)	-0.0331 (0.4012)
student-teacher ratio	-0.0500 (0.0828)	-0.1579 (0.1365)
constant	-2.7215*** (0.5671)	-5.5618** (2.7595)
tuition (β)	0.000037 (0.000031)	
ω_{32}^{*}	-0.0820 (0.7068)	
ω_{33}^{*}	1.2564 (0.9103)	

* indicates a *p*-value between 0.05 and 0.1, ** indicates a *p*-value between 0.01 and 0.05, and *** indicates a *p*-value below 0.01.

TABLE 8: MULTINOMIAL PROBIT ESTIMATION, WHITE SAMPLE

variable	**public college** $(\alpha_{j=2} - \alpha_{j=1})$	**private college** $(\alpha_{j=3} - \alpha_{j=1})$
female	0.2089*** (0.0984)	0.2999* (0.1564)
ASVAB	0.0157*** (0.0022)	0.0204*** (0.0060)
Mother attended college	0.1561 (0.1064)	0.2160 (0.1466)
Mother education data missing	-0.4947 (0.3380)	-0.5470 (0.4247)
Father attended college	0.4398*** (0.1110)	0.6197** (0.2405)
Father education data missing	-0.5523*** (0.1879)	-0.6808** (0.2764)
grades	0.3035*** (0.0403)	0.3779*** (0.0995)
household income, second quartile	0.1674 (0.1528)	0.1133 (0.2018)
household income, third quartile	0.4275*** (0.1522)	0.3839* (0.1983)
household income, fourth quartile	1.1390*** (0.2184)	1.1646*** (0.2540)
college prep program	0.7849*** (0.1046)	0.9393*** (0.2048)
completed AP course	0.0312 (0.1664)	0.1219 (0.2203)
student-teacher ratio	-0.0375 (0.0480)	-0.1517 (0.1377)
constant	-3.3145*** (0.3984)	-5.0976** (2.2885)
tuition (β)	0.000027 (0.000030)	
ω_{32}^{*}	-0.9326 (1.3554)	
ω_{33}^{*}	1.9064*** (0.5654)	

* indicates a *p*-value between 0.05 and 0.1, ** indicates a *p*-value between 0.01 and 0.05, and *** indicates a *p*-value below 0.01.

TABLE 9: SELECTED MARGINAL EFFECTS FROM MULTINOMIAL PROBITS

Marginal effect on public college choice	black	Hispanic	white
public tuition	-0.000013 (0.000014)	0.000012 (0.000013)	0.000015* (0.000007)
private tuition	0.000001 (0.000002)	-0.000003 (0.000021)	-0.000009*** (0.000002)
Mother attended college	0.1722*** (0.0475)	0.0772 (0.0798)	0.0159 (0.0273)
Father attended college	-0.0153 (0.0571)	-0.0402 (0.1868)	0.0430 (0.0331)
grades	0.0280 (0.0181)	0.0481 (0.1894)	0.0439*** (0.0103)
household income, third quartile	0.1467** (0.0605)	0.1788 (0.1112)	0.1091*** (0.0378)
household income, fourth quartile	0.2473** (0.1093)		0.1937*** (0.0404)
college prep program	0.1585*** (0.0457)	0.1284 (0.0980)	0.1239*** (0.0328)
mean predicted probability of private college choice	0.4540	0.4175	0.6203
Marginal effect on private college choice	**black**	**Hispanic**	**white**
public tuition	0.000013 (0.000014)	-0.000003 (0.00002)	-0.0000089*** (0.0000023)
private tuition	-0.000002 (0.000002)	0.000004 (0.00003)	0.0000090*** (0.0000023)
Mother attended college	0.0248 (0.0244)	0.0025 (0.0417)	0.0196 (0.0184)
Father attended college	0.0314 (0.0278)	0.0368 (0.2965)	0.0579** (0.0245)
grades	0.0175 (0.0082)	0.0283 (0.2472)	0.0247*** (0.0072)
household income, third quartile	0.0151 (0.0306)	-0.0008 (0.0349)	-0.0140 (0.0269)
household income, fourth quartile	0.0316 (0.0617)		0.0089 (0.0335)
college prep program	-0.0010 (0.0457)	0.0020 (0.0439)	0.0515** (0.0238)
mean predicted probability of private college choice	0.0661	0.0644	0.1267