What Matters Most?

Analyzing the Gap in College Dropout Rates Between Black and White Students

Christina Ciocca
Columbia University

Thomas A. DiPrete
Columbia University

April 6, 2015

Abstract

Black students have lower rates of degree completion than do white students. They also have lower rates of persistence to degree than do white students, conditional on beginning a postsecondary degree program. This paper analyzes three factors that contribute to these existing disparities: individual background and academic performance, which produce racial differences in the risk of dropout; group distinctions in the probability of high school graduates entering a degree program, given any particular *a priori* risk of dropout; and specific sorting by race into different levels of postsecondary institutions, which impacts students' probability of persistence to degree. We find that black students with a higher predicted probability of dropout are more likely to enter four-year degree programs than are white students with similarly high dropout likelihood. This "academically risky" behavior produces a higher rate of dropout for black students than for white students. But it also increases the rate of degree completion for black students by nearly 5 percent, on average, relative to what it would be if black students entered four-year colleges in the same manner as white students.

1 Introduction

It is well known that degree completion rates for black students are lower than those for white students in postsecondary education at both the two- and four-year level. The National Center for Education Statistics (Snyder and Dillow 2012: Table 379) reports that while 63 percent of white first-time beginners in four-year institutions receive bachelor's degrees six years after college entry, only 41 percent of Black students achieve the same. Gaps also exist at the two-year level, with 17 percent of white students and 11 percent of Black students receiving associate's degrees, and 13 and 5 percent of each group transferring to four-year schools and achieving a bachelor's degree, respectively. These gaps reflect many decades of unequal educational opportunities for white and black students, provoking numerous scholarly inquiries (see Reardon et al. 2015 for a recent overview). However, it remains unclear exactly what combination of factors creates these discrepancies, or how to isolate the relative impact of drivers such as demographic and compositional factors, student background and experiences, prior academic achievement, college enrollment decisions, and college experiences.

We provide insight into these questions by focusing on one of the most important behaviors underlying differential degree completion by race and ethnicity, college dropout. While dropout and degree non-completion at first may seem synonymous, they in fact are not. Postsecondary non-completion rates are largely a function of who chooses to apply and then enroll in higher education, whereas dropout can only occur once students have been accepted and have made the initial decision to enroll. It is clear from national statistics that dropout is more pervasive among black students than white students, with only 21 percent of white students failing to achieve a bachelor's degree six years after four-year college entry, versus 34 percent of black students. In two-year schools, 45 percent of white students and 50 percent for black students drop out in the same time period.[1]

In examining these discrepancies in college dropout, our work builds on a body of literature that has examined various stages of the college-going process to identify individual factors that best explain the higher dropout likelihood of traditionally underrepresented racial and eth-

[1]These figures do not take into account those students who remain in school after six years in both four- and two-year schools, creating the possibility for these gaps to become even greater over time. In BA programs, 10 percent of white students and 20 percent of Black students are still in school, while in AA programs, 16 percent of white students and 22 percent of Black students remain.

nic minorities (Light and Strayer 2002; Swail 2003; Carter 2006; Bowen et al. 2009). We know from these studies that socioeconomic background and academic preparation matter tremendously for both student selection into college and persistence; it also has been found that these factors differ significantly by race and ethnicity (Adelman 2006; Kuh et al. 2006). Enrollment patterns by race and ethnicity also have changed over time, with an increasing number of black, Hispanic, and other minority students enrolling as the college degree premium has increased, impacting the overall distribution of the college-going population (Black and Sufi 2002; Bound et al. 2010). Some institutional factors at the college level additionally seem to make a difference as to whether students persist, such as selectivity in four-year schools (Light and Strayer, 2000; Alon and Tienda, 2005; Bowen et al., 2009) or the clarity of academic and career-oriented pathways in two-year schools (Rosenbaum et al. 2007; Karp and Hughes 2008; Scott-Clayton 2011).

However, few studies have contextualized these individual determinants within the broader process of deciding to enter a postsecondary institution – which creates the opportunity to graduate and the risk to dropout – and the sorting of students into colleges with differing efficacy when it comes to degree completion. While some notable exceptions in the literature have explored these relationships (Venti and Wise 1983; Light and Strayer 2002; Bound et al. 2010), important gaps remain. First, past analyses have tended to focus on four-year entrants, which is a notable omission considering that 45 percent of all undergraduates today begin at two-year colleges (Knapp et al.2012; AACC 2014), and that these institutions tend to enroll a disproportionate number of non-white and high-risk students (Goldrick-Rab 2010). Second, past studies of the selection process have relied on a basic set of measures of academic preparation, family background, and high school experiences, which provide an inadequate understanding of the sources of racial difference in enrollment and dropout. Third, past work has not provided a granular analysis of the ways in which gender interacts with race to shape differential dropout rates between black and white students, limiting our understanding of the primary demographic forces driving the dropout gap. Finally, the studies that do the best job of simultaneously considering entry and dropout behavior fail to distinguish among types of minority students.

In response, our study seeks to address these omissions by making five distinct contributions. First, we estimate enrollment and dropout behavior for students who enter both two- and four-year colleges. Second, we use a comprehensive set of academic and attitudinal high school

measures from the Educational Longitudinal Study (ELS 2002), including intensity of students' high school curriculum, parental engagement in college-going decisions, college decisions of peers, and student self-expectations and priorities. Third, we estimate the proportion of postsecondary degrees that result from the race-specific entry and dropout patterns of black secondary school graduates, in particular, enabling us to assess the implications of racial differences in college entry and departure. Fourth, we analyze gender-specific impacts on the dropout gap for both four-year and two-year postsecondary institutions, allowing us to assess the primary demographic drivers underlying the higher dropout rate of black students. Fifth, we provide a direct comparison between white and black students, rather than aggregating students of color into a comprehensive "underrepresented minority" category – an important distinction considering that recent studies have found that black and Hispanic students experience college quite differently (Lee 2002; Greene et al. 2008).

Our findings indicate that the most important factor contributing to differential college dropout between black and white students is the distribution of initial dropout risk, which accounts for nearly 80 percent of the gap in dropout rates between white and black students who begin college in four-year colleges and universities and over 80 percent of the gap for students who start in two-year colleges. This finding suggests that the college dropout gap between white and black students is primarily a byproduct of decisive educational, social, and economic patterns that occur prior to enrollment in higher education. However, we also find that race-specific patterns of college entry and dropout, resulting from a combination of academic and socioeconomic forces, have a clear impact on persistence to degree, accounting for the remaining portion of the persistence gap. Specifically, if black students were to enter and depart four-year colleges using the same decision models as white students, their overall dropout rate would fall by 10 percent. Yet the proportion of black students earning a bachelor's degree also would decrease by nearly 5 percent (.290 versus .276). The two-year case is more complex, as the dropout rate for black students in fact would rise by 2 percent if these students were to follow the white entry and departure models. This finding supports our conviction that black students' tendency to select into four-year colleges at higher risk levels may increase their overall rate of degree completion. The combination of these results reveals a clear "tradeoff" dynamic for black students in the interconnected process of college entry and departure, with greater risk-taking on entry opening the possi-

bility for higher degree completion rates. Accordingly, we also find that postsecondary experiences are a crucial mediator of the black-white dropout gap, independently and apart from pre-college factors. It is the combination of the pre-college distribution of risk and race-specific decisions concerning college entry and departure that generate the observable gap in the dropout rate between black and white students.

We proceed with a review of relevant literature, followed by a description of our data and methods, and a discussion of our findings. We finish with our conclusions and a brief outline of areas of further investigation we intend to pursue.

2 Review of Relevant Literature

Racial and ethnic gaps in educational attainment are well-documented at all levels of the educational system and the life course (Lee 2002; Kao and Thompson 2003; Maralani 2013). In the context of higher education, most studies tend to focus exclusively on racial and ethnic discrepancies in college enrollment or dropout, working to identify the various mechanisms contributing to gaps in each phase of the college-going process individually. The enrollment literature suggests that discrepancies can be explained by a combination of demographic and economic shifts over time (Black and Sufi 2002; Kane 1994), financial and social resource gaps (Beattie 2002; Mumper 2003; Bastedo and Jaquette 2011), school effects at the high school level (Hill 2008; Roderick et al. 2011; Belasco 2013; Jennings et al. 2015), and college admissions processes (Bowen and Bok 2000; Alon and Tienda 2007; Alon 2009; Grodsky et al. 2008). The dropout literature, by contrast, focuses primarily on college-level factors such as institutional selectivity (Small and Winship 2007; Bowen et al. 2009), financial aid packages (Kim 2007; Chen and DesJardins 2010; Alon 2011) or integration and support within the collegiate environment (Deil-Amen and Rosenbaum 2003; Swail 2003; Hausmann et al. 2007). While most studies of college dropout introduce statistical controls to account for pre-college experiences, few explicitly model the impact of selection patterns by race on subsequent gaps in college persistence or degree completion rates. The result of this omission is a less comprehensive and precise understanding of differential dropout by race than can be provided by accounts that take enrollment patterns as a critical contributor to subsequent dropout decisions and degree attainment rates.

Light and Strayer (2002)provide an important exception, using the 1979 National Longitudinal Study of Youth (NLSY) data to study the differences between white and minority students in college enrollment and completion. Building on the foundation laid by Venti and Wise (1983), arguably the first to use a joint discrete-continuous utility maximization model to evaluate the impact of selection on college dropout, Light and Strayer employ a similar analytical strategy to make race the focal point of their study. The authors argue that without jointly estimating college enrollment and dropout, it would be relatively easy to mistake various unobserved factors in the enrollment stage for racial impacts on dropout. They also incorporate selectivity categories into their initial enrollment choice framework, helping to account for the impacts of increasing horizontal stratification within the higher education sector (Gerber and Cheung 2008). However, as mentioned earlier, their study is constrained in some important ways, including its restriction to four-year college students, its limited pre-college controls, its lack of granularity in exploring the mechanisms underlying differences in degree completion between racial and ethnic groups, and its decision to group all minority students together. Perhaps as a result of these limitations, the authors ultimately attribute the higher rate of degree completion among white students to unobservable factors.

Bound et al. (2010) help to address some of the remaining questions left by Light and Strayer (2002) in their study of the overall decline in college persistence rates over time. In particular, they choose to include two-year students in their analysis, demonstrating the ways in which expansion of higher education, and especially the two-year college sector, over time has impacted college completion rates. Using a counterfactual framework that enables a comparison between students with comparable academic preparation and socioeconomic background in 1972 and 1997, they find that the increased willingness among 1997 students – and especially men – to attend two-year and less selective four-year colleges is a key factor in explaining the overall decline in postsecondary degree receipt. By contrast, students with similar backgrounds in 1972 were less likely to enroll in college at all. This finding once again focuses attention on the importance of evaluating the selection process when studying college dropout rates, while also demonstrating that two-year colleges must be included in such studies to provide a comprehensive understanding of the postsecondary landscape. Despite these key contributions, Bound et al. are not particularly focused on racial differences in college completion, instead drawing on academic preparation

in high school as their primary strategy for dividing the student population. More generally, the authors are most interested in determining whether the demand-side (students) or the supply-side (institutions and institutional funders, like states) is more important in explaining persistence trends rather than the ways in which these factors impact racial and ethnic differences in persistence.

Since Bound et al.'s analysis, there has been no study to our knowledge that has focused explicitly on connecting the college selection process to persistence decisions. In addition, we have not found an example beyond Light and Strayer's (2003) work that uses such an analytical strategy to shed light on the question of differential dropout by race. We also note that both of these prior analyses have been driven by the disciplinary perspective of economics, whereas sociologists have tended to focus more intensively on either enrollment or dropout, exclusively. It is our goal to build on these foundations, presenting a counterfactual analysis focused on connecting the relative independent impacts of students' family backgrounds, academic achievement, college enrollment decisions, and postsecondary experiences on the black-white gap in college persistence.

3 Data

We draw our data from the Educational Longitudinal Survey (ELS), collected by the National Center for Educational Statistics (NCES). ELS includes a nationally representative sample of nearly 15,000 students through four waves of data collection: the 2002 base round, at which point all students were in tenth grade; a first follow-up in 2004, when most students were in twelfth grade; and a second follow-up in 2006, which traces students into either their first two years of postsecondary education or the workforce, and a third follow-up in 2012, which captures information about students' collegiate experiences and outcomes, as well as workforce experiences. The initial ELS sample was drawn from a representative distribution of 750 high schools, from which students were selected at random to participate in the survey. The ELS data, being a cluster sample of high school students, does not contain clustered data at the post-secondary level. However, because data collection began when the students were in high school, we are able to gather quite detailed information on pre-college academics in addition to postsecondary enrollment decisions, experiences, and persistence outcomes over time.

As a result of the rich information we gain from the ELS about students prior to college enrollment, we are able to include information on the intensity of high school curriculum, performance in high school, communication between students and parents regarding academics, high school dropout behavior of friends, and students' own expectations and priorities regarding higher education and other life outcomes in addition to basic demographic information. We also include key pieces of information concerning students' postsecondary experiences in our following analyses of college impacts on dropout, such as college selectivity, the availability and use of financial aid, remedial course-taking, major field of study, contact with faculty and advisors, and social engagement in the school community. We employ an appropriate panel weight, constructed by the NCES, to account for the longitudinal design of the study and to make population inferences for our focal sample. While most ELS variables that comprise our initial regressions are coded in a straightforward ordinal manner, the curriculum intensity variable has been constructed using raw data on students' course-taking patterns. It follows the scheme developed in a prior education dataset, the National Education Longitudinal Survey (NELS), and subsequently employed by Adelman (2006), to generate 30 levels of curricular intensity. These levels have been condensed to five in our model to improve statistical power. The list of all pre-college and post-college-entry variables can be found in Appendix 1.

For the purposes of this draft, we limit our sample to cases in which all necessary variables for our analyses are present, using list-wise deletion. We will pursue a multiple imputation strategy to restore missing cases in a future draft of the paper. As a point of comparison, we demonstrate in Table 1 that the ELS data prior to list-wise deletion closely parallels the high school / GED completion rates and college enrollment rates reported by the NCES, providing confidence in the representativeness of our findings once we have undertaken multiple imputation. Our focal population is all black and white students who have completed a high school diploma or GED; we exclude students who have dropped out of high school, since these students would be largely ineligible for college entry. Our sample therefore consists of 6,209 students, 54 percent of whom begin college in a four-year institution, 31 percent in a two-year school, and 15 percent in neither (see Table 2). Of those students who begin in a four-year college, 23 percent do not complete a college degree; for those who start in two-year schools, 47 do not complete a college degree. We define "dropout" as occurring when students have not completed a bachelor's or associate's degree

by 2012, the censoring point in the ELS data.

4 Methods

We employ counterfactual analysis as our primary empirical strategy. To do so, we first fit two different logistic regressions, one for college entry and the second for college dropout. For college entry, we define three mutually exclusive binary entry states: 1. Four-year college or university entry; 2. Two-year college entry, given that a students initially enrolls in either a two-year or four-year college; and 3. Two-year college entry, given that a student has not enrolled in a four-year college or university. The same independent variables are used in each of these analyses and include the pre-college socioeconomic, academic, and expectation factors described in Appendix 1. To define these states, we do not draw on any variables that reflect students' experiences following college entry, as we wish to understand entry as a function of pre-college experiences, alone. We run separate models for black and white students in order to isolate race-specific coefficient values on college entry probabilities. We therefore exclude race as an independent variable from the initial estimation models.

Next, using the same independent variables as utilized in the entry models, we run three separate logistic regressions to asses students' probability of college dropout based on pre-college factors. Nearly 85 percent of students within the ELS sample receive their high school diploma or GED in 2004, and since over 95 percent of four-year beginners and 70 percent of two-year beginners in this sample start college within one year of receiving a high school diploma or GED, we generally gain an eight-year window to observe potential dropout behavior. This window equates to 400 percent of the typical program time for two-year college entrants and 200 percent for four-year college entrants, providing a strong gauge of actual dropout behavior over time. As with our entry models, we estimate students' likelihood of dropout in three separate ways: 1. For all students who have started college in four-year schools; 2. For all students who have started in two-year schools, given college entry; and 3. For all students who have started in two-year schools, given that they have not initially enrolled in four-year schools. Once again, we run separate models for black and white students, using these initial models of dropout to generate three separate "dropout risk indices" for all students in our sample, regardless of whether or not they actually

have attended college. The first index represents four-year dropout risk ("Risk Index 1"); the second, two-year dropout risk given college entry ("Risk Index 2"); and the third, two-year dropout risk given no four-year college entry ("Risk Index 3"). All three indices are generated as a linear combination of coefficient values on each control variable in our model. We are able to extend these indices to students who do not enter the focal type of college via out-of-sample prediction.

These indices subsequently enable us to calculate enrollment and dropout rates as a function of race and dropout risk, alone, rather than as a function of individual covariate values, as we did in the first case. To accomplish this task, we substitute each dropout risk index as the sole independent variables in our initial logistic regression models for college entry. So, we first calculate four-year college entry as a function of four-year dropout risk (in original, squared, cubed, and quadrupled forms), then two-year entry as a function of four-year dropout risk, then two-year entry as a function of two-year dropout risk. We run these equations separately for black and white students in order to assign new entry probability scores based on dropout risk values, alone. We then can use these new entry probabilities to calculate dropout as a function of initial risk. In doing so, we produce mean entry and dropout estimates that reflect students' behavior in light of comprehensive risk rather than individual covariate values.

We next perform a series of counterfactual analyses to tease apart the impact of each race's initial distribution of pre-college risk factors, decision-making model for entry, and decision-making model for dropout on the gap in the dropout rate between black and white students. Because we can define entry and dropout models by using covariates, alone, or by using the comprehensive risk index, we also can construct counterfactual simulations in light of each type of model. We therefore pursue six different counterfactual simulations:

1. Black students' entry rate based on black students' variables and white students' entry decision model as a function of covariates;

2. Black students' entry rate based on black students' variables and white students' entry decision model as a function of comprehensive risk;

3. Black students' dropout rate based on black students' variables, black students' dropout decision model, and white students' entry decision model as a function of covariates;

4. Black students' dropout rate based on black students' variables, black students' dropout decision model, and white students' entry decision model as a function of comprehensive risk;

5. Black students' dropout rate based on black students' variables, white students' entry decision model, and white students' dropout decision model as a function of covariates;

6. Black students' dropout rate based on black students' variables, white students' entry decision model, and white students' dropout decision model as a function of comprehensive risk.

We generate these counterfactual simulations for all three of our initial entry states, and also calculate the reverse counterfactual – e.g. if white students were to use black students' entry model based on covariates, etc.

As a final step, we develop an index of "pure" postsecondary dropout risk, in that the index has been purged of the influence of all pre-college variables. We do this by regressing the postsecondary variables on the pre-college variables, residualizing, and then using all of the residuals as the independent variables in a logistic regression with "dropout" as the dependent variable. From this regression, we generate a predicted postsecondary dropout value for each student who has started in the focal college level. Because of the high dimensionality of many of the postsecondary variables, we use an initial factor analysis to provide values for the postsecondary variables that can be regressed against the pre-college variables. We are confident in this choice because the correlation between postsecondary dropout indices created using individual postsecondary covariates versus those generated using postsecondary factors is greater than 0.7, on average.

5 Discussion of Findings

Note: For the purposes of this draft, we will focus exclusively on our results for the four-year entry state. Future drafts will detail our two other focal states, probability of two-year attendance given college entry and probability of two-year attendance given no four-year attendance.

5.1 Distribution of Pre-College Dropout Risk

To begin understanding the impact of pre-college experiences on college entry and departure behavior for black and white students, we first examine the distribution of initial dropout risk (Figure 1a). We find that black dropout risk is more heavily concentrated in the "high-risk" portion of the distribution while white dropout risk peaks at a much lower risk level, concretizing the distributional inequity between the two groups of students given pre-college experiences. The low-risk concentration of white students as compared to black students is even more notable when we focus on four-year beginners: while nearly the entire distribution of white students in this group is condensed between a predicted dropout probability of 0 and 0.4, that of black students stretches the length of the risk spectrum and only begins to decline at the 0.6 level. Though black students who begin college in four-year schools are also more focused in the low-risk portion of the distribution than the entire population of black students, this difference only closes the gap slightly between the pre-college dropout risk levels of black and white students. In examining the separate graphs of pre-college dropout risk for men (Figure 1b) and women (Figure 1c), it becomes clear that the gap between black and white women drives that which we observe in the overall distribution, whereas the differences between black and white men in their distribution of pre-college risk is less severe. As our later results will confirm, this discrepancy matters significantly for overall racial differences in entry and departure patterns.

5.2 College Entry

Turning to college entry, we initially examine entry as a function of pre-college covariates (Figure 2). Figure 2a demonstrates that the main difference between white and black students is that, for all values larger than about 0.3 on the x-axis (predicted probability of dropout), black students are more likely to enter college than are white students with identical risk. We observe this same pattern for black men (Figure 2b) and black women (Figure 2c), though the "crossover" point is higher for men at nearly 0.4 and lower for women at approximately 0.2. We also learn from the separate graphs by gender that the lowest-risk women from each race respond to pre-college dropout risk in a more comparable way than the lowest-risk men when it comes to four-year college entry. This pattern resonates with prior findings regarding the heightened disadvantage of

black men in the college-going process (*DiPrete and Buchmann 2013*) and suggests that low-risk black women may not experience as substantial a disadvantage in relation to their white peers with similar dropout risk levels.

Our counterfactual analysis of entry trends helps to clarify these patterns. The results for entry likelihood in response to dropout risk defined by covariates demonstrate that if black students were to enter four-year colleges using the same entry model as white students, their average probability of entry would fall by 10 percent, from 0.452 to 0.407 (see Table 3, Column 2) . Meanwhile, the entry rate for white students would increase by 10 percent, from 0.553 to 0.616. Similar trends emerge for women and men, though it is worthwhile to note that black men enter four-year college at a higher rate than do black women in this ELS sample. Since we know that black women complete nearly two-thirds of all bachelor's degrees among black students *(DiPrete and Buchmann 2013*), it is likely that the ELS statistic reflects the fact that our sample is limited to students who have earned high school diplomas and GEDs, rather than all students who begin high school, in which a higher proportion of black men drop out prior to diploma receipt than do women.

More generally, it is worthwhile to conceptualize what these counterfactual values mean as explicitly as possible. Using the counterfactual value for all black students as an example, a useful way to think about the number is: "How would a white student think about college entry if she observed the pre-colleges experiences of a similar black student as her own?" Importantly, the unchanging factor in all black counterfactual numbers is the value of variables: the variables always reflect the family background, high school achievement, and college expectations of black students, even as we shift to change the lens of the viewer interpreting these experiences from black to white in the counterfactual analysis. We therefore can translate the black counterfactual value to mean that an average white student would be less likely to make the decision to enter a four-year college if she had the same pre-college resources as her black peers. In this sense, the response of white students to black students' pre-college experiences represents a "resource devaluation" – depicted numerically in Table 4. Table 4 demonstrates that, when we divide the distribution of white students into deciles, white students are more likely to enter college within every decile if responding to their own resources versus the resources of their black peers. This pattern holds for women and men, alike.

At first glance, it seems possible that this finding simply could reflect the fact that on average, black students tend to have fewer resources than white students – a trend immediately apparent in the distribution of pre-college dropout risk (Figure 1). However, the results of the counterfactual analysis of four-year entry probability as a function of comprehensive pre-college dropout risk suggests that the dynamic is far more nuanced (Table 3, Column 3). To be clear, predicted entry based on "comprehensive pre-college dropout risk" simply means that our logistic regressions are run to make college entry a function of the dropout risk score, alone, rather than all covariates. We discover from this analysis that, not only do the counterfactual entry values differ markedly from those produced by entry response to individual variables, but that within-gender differences serve as the primary driver of these discrepancies. In the case of all students, for example, the counterfactual result for black students only drops by 1 additional percentage point to 0.397, while that for white students dips by over 5 added percentage points to 0.562. Using the analytical framework detailed above, this result indicates that when white students assess entry based on the combination of multiple risk factors within the black resource pool, they are slightly more cautious about entering four-year colleges or, looking at it a different way, they devalue black resources slightly more. Yet when black students assess white resources in the comprehensive risk framework, they become far more cautious about college entry, suggesting that a similar, though more significant, devaluation process has taken place in the conversion of white students' singular experiences to an amalgamated indicator of risk.

However, considering the foundational point that white students tend to have greater resources than black students, this devaluation makes little intuitive sense. So we turn to the entry probabilities by gender to try to disentangle this finding. Looking separately at women, we find that the counterfactual value for black women drops by over 7 percentage points to 0.321, while that for white women stays relatively stable, only increasing by 1 percentage point to 0.617. For men, the counterfactual entry value for black students rises by 15 percentage points to 0.529, while that for white men diminishes to 0.512 – also a 15 percentage point drop. These gender-specific results indicate that women and men of both races interpret and respond to comprehensive dropout risk quite differently from one another. More specifically, while the lower counterfactual entry rate for black women suggests that white women perceive fewer resources when evaluating black women's comprehensive dropout risk in comparison to variable-based risk, the

higher counterfactual entry rate for black men indicates that white men in fact perceive black men's resources based on comprehensive dropout risk as far greater than those present in the covariate scenario. Shifting to consider the opposite perspective, the relative stability of the white women's counterfactual entry value suggests that black women perceive of white women's resources as a function of dropout risk as about equally valuable to those derived from individual variables, whereas the lower counterfactual value for white men indicates that black men view the risk-based resources of white men as far less valuable.

The language of currency conversion, in combination with Figure 3, can help us to interpret these findings. Figure 3 depicts the relationship between students' predicted probability of dropout and their percentile rank on these same dropout indices. The dashed line reflects the black counterfactual dropout rate that would result as a direct consequence of the black counterfactual entry rate (so, the black dropout rate if white students were to assess black resources in the entry decision-making process). As such, these dropout curves can be viewed as the conversion of pre-college entry behavior into the metric of dropout risk. A numeric representation of these patterns can be found in Table 5. Bearing this in mind, we observe in Figure 3a that at the lowest portion of the dropout distribution, black women's race-specific "dropout risk currency" is equally – if not slightly more – valuable than white women's. However, beginning at about the 18th percentile of the dropout risk distribution, black women's dropout risk currency diminishes in value as compared to white women's, thereby increasing black women's predicted probability of dropout substantially, especially at the very highest levels of risk. The story for men (Figure 3b) is somewhat different, in that black men's dropout risk currency is more valuable than white men's up until the 35th percentile of the distribution, at which point white men's dropout risk becomes more valuable, but substantially less so than white women's over black women's. Extending the currency analogy, it is as though black students on average hold U.S. dollars; white women, British pounds,' and white men, only euros.

The counterfactual curves help to make sense of what this valuation means for the interpretation of comprehensive dropout risk by gender. The counterfactual for women indicates that white women perceive their pounds to be much more valuable than dollars. In the case that white women all of a sudden realized that they in fact only held dollars rather than pounds– as the counterfactual scenario asks of us – they would behave more cautiously in the face of "currency

devaluation," choosing to enter college less often and in turn, lowering the counterfactual dropout rate of black women. As such, although black women's dropout rate would lessen consequently, the devaluation process also would serve to widen the entry gap between white and black women, in a way over-exaggerating the actual dropout risk of black women prior to college entry. Decomposing this devaluation process down into socioeconomic and academic components (Figures 4a and 5a), it is clear that socioeconomic differences between black and white women are the more important force in driving the overall devaluation.

For men, on the other hand, the conversion process looks very different, with white men gaining awareness of the comparative strength of black men's currency and black men coming to realize that white men's currency in fact is not much stronger than their own. These processes help to explain the increase in the counterfactual entry rate for black men, translating directly into the higher dropout rate for a large portion of the black distribution depicted by the counterfactual curve in Figure 3b, versus the lower counterfactual entry rate for white men. Comparing the counterfactual for white men based on covariates versus comprehensive dropout risk, it appears as though the general power of white men's resources lies in their amount rather than their value. When black men assess the amount of white men's resources, captured in the covariate counterfactual, they are much more willing to enter college than they are when assessing the value, as depicted by the risk counterfactual. In this way, the currency conversion process for men seems to represent the exact opposite from that for women, in that white men's currency in fact is devalued in conversion while black men's increases in value. Looking at Figures 4b and 5b, it seems the devaluation of white men's resources is rooted in academics: white men do not possess academic resources that are as valuable in comparison to black men. Together, the results for men and women convey the importance of socioeconomic resources in widening the advantage of white students in the college entry process over their black peers, seeming to override the value of academic experiences.

5.3 Impact of College Entry Patterns on College Dropout Rates

Yet how do these entry patterns impact college dropout decisions? Table 6 summarizes the first phase this analysis, in which we estimate black and white students' dropout rates as a function of entry response to covariates and comprehensive dropout risk. The counterfactuals in this ta-

ble reflect the combination of counterfactual figures for entry and actual, race-specific figures for dropout. Looking first at the counterfactual dropout given entry based on covariates (Column 2), it is somewhat surprising that the average effect on dropout rates for all students are modest when entry rates for both groups of students have changed notably in the counterfactual entry state. If white students were to respond to black resources based on covariates, the dropout rate for black students would fall only by 1 percent, from 0.358 to 0.355. Meanwhile, the white dropout rate would only rise by about 5 percent, from 0.217 to 0.228, in the equal but opposite counterfactual scenario. Though there is slightly more movement in the counterfactual number for white men – likely because the entry counterfactual inflates the male entry rate to such a great extent, thereby enlarging the pool of potential non-completers – the patterns observed within each gender generally support the overall trend. We can attribute these patterns largely to the distribution of dropout risk by race together with race-specific models of dropout. For black students, following the white entry curve would cause students with low dropout risk to enter four-year college at a higher rate, whereas students with high dropout risk would enter at a lower rate (see Figure 2). However, we also know that a far smaller proportion of black students can be found in the low-risk portion of the distribution than the high-risk portion. In addition, we learn from plotting actual and counterfactual dropout risk for black students that those with low dropout risk tend to drop out at higher rates than would be expected, while those with higher risk actually drop out at lower rates than anticipated (see Figure 6). These two trends therefore appear to offset one another in the when considering the counterfactual dropout rate based on variables, leaving the black dropout rate rather level. The white scenario reflects a similar offsetting effect, though in the opposite direction.

We again gain further insight into these patterns by observing the counterfactual dropout rates that would result from counterfactual entry based on comprehensive dropout risk, which include the race- and gender-specific conversion processes discussed above. Table 6, Column 3 depicts the consequences of the conversion process on dropout rates, which differ substantially from those resulting from the covariate case and once again are driven by gender differences. For women, the dropout rates represent logical extensions of the counterfactual entry rates. The smaller proportion of black entrants emerging from the black counterfactual entry model translates directly into a far lower dropout rate at 0.176, especially because those women who would no longer

enter would be the highest-risk women based on the devaluation process in combination with the white entry curve. On the whole, this lower counterfactual figure demonstrates that the black women who would enter college according to the counterfactual represent high performers not prone to dropout. Meanwhile, the white counterfactual dropout rate climbs to 0.244 as a result of the higher proportion of women that would start college according to the white entry counterfactual, especially higher-risk women. Because these higher-risk women tend to perform quite poorly in the white case, it follows that the counterfactual dropout rate increases.

Though it would seem intuitive for this same logic to apply to men, the fact that the entry models for men and women differ so significantly, combined with the divergent conceptualization of devaluation, appears to change the nature of expected dropout rates. Here, the higher counterfactual entry rate of black men in fact translates to an extremely low dropout rate of 0.191, whereas the lower counterfactual entry rate of white men produces a far higher dropout rate of 0.358. We can attribute this pattern to the fact that a higher entry rate for black students would mean that a much higher proportion of the lowest-risk black men would enter college according to the white entry model, and that these low-risk men in fact perform far better in the "comprehensive dropout risk" universe than the covariate universe. Conversely, a higher proportion of the highest-risk white men, together with a lower proportion of low-risk white men, would enter college according to the black entry model. In addition, all of these white men in fact perform worse than expected in space of comprehensive risk. As a result, the dropout rates based on comprehensive dropout risk look very different from those produced based on covariates.

As one additional step, we can learn more about the nature of the dropout gap between white and black students when we examine the counterfactuals based on both counterfactual entry and departure models (Table 7), which represent the full expression of race- and gender-specific differences in four-year college entry and departure. The covariate scenario (Column 2) reflects the tendency of white women to persist at higher rates than black women, as the counterfactual dropout rate for black women falls to 0.264 while than for white women rises to 0.262. Yet we also observe the stronger performance of black men as compared to white men, in that the black counterfactual rate rises to 0.389 while that for white men falls to 0.287. These results suggest that once in college, black men in fact persist at higher rates than white men, especially white men deemed "high risk" based on their pre-college characteristics. Interestingly, while the

counterfactual results based on risk have differed substantial from those emerging from covariates in both of our initial analyses, they reflect very similar patterns here. Black women persist at lower rates than do white women, whereas black men persist at much higher rates than do white men. We can attribute the uniformity of the results in this analysis to the fact that race-specific departure patterns are more closely aligned with race-specific entry patterns, no matter whether we use covariates or comprehensive dropout risk as the baseline. It is also worthwhile to note that the counterfactual results based on comprehensive risk for black students are a direct reflection of the counterfactual curves depicted in Figure 3.

5.4 Postsecondary Experiences

As a final step, we examine the purged postsecondary component of dropout risk to gain insight into the impact of postsecondary experiences on dropout, net of pre-college factors. Confirming our analysis of dropout rates as a function of counterfactual entry and departure behavior, Figures 7 and 8 demonstrate that postsecondary experiences in fact matter substantially above and beyond pre-college experiences. If they did not matter, there would be little or no overlapping in the distributions of Figure 7 and Figure 8 would simply contain parallel lines at the mean dropout rates for all black and white four-year beginners. The fact that there are notable slopes supports our finding, represented by Figure 6, that the entry model is not destiny for students when it comes to college completion and departure. Instead, students' experiences within postsecondary education independently shape their likelihood of completion. As we see from Figures 7 and 8, a large proportion of black students are performing better than would be anticipated by their initial dropout risk, especially among male students. Yet male students also account for the most extreme dropout behavior in the high-risk portion of the postsecondary risk distribution. Conversely, black women with low postsecondary risk drop out at a might higher rate than would be anticipated. These two surprising trends that take shape after collegiate enrollment merit substantial future investigation.

6 Conclusions

In this paper, we make a contribution to the literature on the gap between white and black students in college degree completion by focusing on the critical issue of dropout. Rather than limiting our study to either the issue of enrollment or dropout, we analyze these decision-making moments in tandem, seeking to understand more comprehensively the impacts of the pre-college distribution of student dropout risk and students' selection preferences on observed differences by race in dropout. We also include an analysis of postsecondary experiences, net of pre-college factors, to capture the ways in which these experiences interact with students' initial dropout risk to either divert or confirm their decision to leave. We therefore are able to present an account of the entire dropout process, from pre-college experiences through to the dropout decision.

Our study builds on a small body of past literature that attempts to implement a similarly thorough analytical approach (Light and Strayer 2002; Bowen et al. 2009; Bound et al. 2010). However, because we define our students not only by race but also by their placement on three, separate but interconnected risk scales, we gain a more granular understanding of which experiences during the dropout process matter most in explaining differential dropout. We find that the initial distribution of dropout risk, in combination with students' race-specific entry decisions in the high-risk portion of the dropout risk distribution, in particular, are most central to the dropout gap. Specifically, higher-risk black students are more likely to enter four-year colleges than are higher-risk white students. The college entry process therefore represents a tradeoff for black students: while entering college at higher risk levels appears to drive up the dropout rate, it also provides more students with the opportunity to achieve a degree. Quantifying this tradeoff, we learn from our counterfactual analysis that if black students were to use the white model of college entry, the dropout rate would fall by 1 percent according to the covariate model or 30 percent based on the comprehensive risk model. However, utilizing the white entry model also would decrease the number of four-year degrees completed by 10 percent in the case of the covariate model and, while the degree completion rate would increase by about 3 percent drawing on the risk model, far fewer black students would have the opportunity to gain from the experience of accruing some college. If we use the white counterfactual model for entry and dropout, black students' overall dropout rate would fall by 10 percent drawing on the covariate model, but the

number of bachelor's degrees earned would decrease by nearly 5 percent. Finally, using the comprehensive risk model, the dropout rate in fact would rise by 9 percent. Taking all of these scenarios into account, it becomes clear that black students' rationally "risky" entry behavior in fact yields worthwhile rewards on the whole.

More generally, our comparative analyses of counterfactual entry and departure based on covariates versus comprehensive risk provide insights into black and white students' perceptions of dropout risk in relation to the amount and nature of pre-college resources. Here, we discover that, as compared to white men's, black men's resources are in fact much more valuable than represented by individual variables such as family socioeconomic background or high school grades. As a result, black men seem to persist to degree in four-year colleges at a much higher rate than would be expected by their pre-college dropout risk. Meanwhile, white men's pre-college resources appear less valuable than their individual variable values would suggest, leading to a higher-than-anticipated dropout rate in college. White men's advantage in the college-going process therefore instead derives from their access to a relatively large amount of pre-college resources, rather than the value of their resources. By contrast, white women's comparative advantage over black women results more directly from the quality of their socioeconomic resources rather than the amount, alone. Accordingly, white women persist at a much higher rate than black women, an advantage that also exists in relation to black men and white men. Finally, it is worth noting that for both white women and men, socioeconomic factors appear more important in generating an advantageous college-going posture over their black peers than do academic factors.

In subsequent versions of this paper, we will more intensively study the sorting process into college levels, analyzing the impact of this process on overall degree completion rates by race. We also recognize that our analytical approach could be extended usefully to examine changing enrollment trends between racial and ethnic groups over time or differences between the dropout behavior of U.S. students versus those from other countries. We hope to pursue both of these avenues in the future.

References

AACC. 2014. "Community College Fact Sheet." Technical report, American Association of Community Colleges.

Adelman, Clifford. 2006. "The Toolbox Revisited: Paths to Degree Completion from High School Through College." Technical report, U.S. Department of Education.

Alon, Sigal. 2009. "The Evolution of Class Inequality in Higher Education: Competition, Exclusion, and Adaptation." *American Sociological Review* 74:731–755.

Alon, Sigal. 2011. "Who Benefits Most from Financial Aid? The Heterogeneous Effect of Need-Based Grants on Students' College Persistence*." *Social Science Quarterly* 92:807–829.

Alon, Sigal and Marta Tienda. 2005. "Assessing the "Mismatch" Hypothesis: Differences in College Graduation Rates by Institutional Selectivity." *Sociology of Education* 78:294–315.

Alon, Sigal and Marta Tienda. 2007. "Diversity, Opportunity, and the Shifting Meritocracy in Higher Education." *American Sociological Review* 72:487–511.

Bastedo, Michael N. and Ozan Jaquette. 2011. "Running in Place: Low-Income Students and the Dynamics of Higher Education Stratification." *Educational Evaluation and Policy Analysis* 33:318–339.

Beattie, Irenee R. 2002. "Are All "Adolescent Econometricians" Created Equal? Racial, Class, and Gender Differences in College Enrollment." *Sociology of Education* 75:19–43.

Belasco, Andrew S. 2013. "Creating College Opportunity: School Counselors and Their Influence on Postsecondary Enrollment." *Research in Higher Education* 54:781–804.

Black, Sandra E. and Amir Sufi. 2002. "Who Goes to College? Differential Enrollment by Race and Family Background." Working Paper 9310, National Bureau of Economic Research.

Bound, John, Michael F. Lovenheim, and Sarah Turner. 2010. "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources." *American economic journal. Applied economics* 2:129–157.

Bowen, William G and Derek Bok. 2000. *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions*. Princeton University Press.

Bowen, William G, Matthew M Chingos, and Michael S McPherson. 2009. *Crossing the finish line : completing college at America's public universities*. Princeton, N.J.: Princeton University Press.

Carter, Deborah Faye. 2006. "Key issues in the persistence of underrepresented minority students." *New Directions for Institutional Research* 2006:33–46.

Chen, Rong and Stephen L. DesJardins. 2010. "Investigating the Impact of Financial Aid on Student Dropout Risks: Racial and Ethnic Differences." *The Journal of Higher Education* 81:179–208.

Deil-Amen, Regina and James E. Rosenbaum. 2003. "The Social Prerequisites of Success: Can College Structure Reduce the Need for Social Know-How?" *The ANNALS of the American Academy of Political and Social Science* 586:120–143.

DiPrete, Thomas A. and Claudia Buchmann. 2013. *The Rise of Women: The Growing Gender Gap in Education and What It Means for American Schools*. Russell Sage Foundation.

Gerber, Theodore P. and Sin Yi Cheung. 2008. "Horizontal Stratification in Postsecondary Education: Forms, Explanations, and Implications." *Annual Review of Sociology* 34:299–318.

Goldrick-Rab, Sara. 2010. "Challenges and Opportunities for Improving Community College Student Success." *Review of Educational Research* 80:437–469.

Greene, Thomas G., C. Nathan Marti, and Kay McClenney. 2008. "The Effort-Outcome Gap: Differences for African American and Hispanic Community College Students in Student Engagement and Academic Achievement." *The Journal of Higher Education* 79:513–539.

Grodsky, Eric, John Robert Warren, and Erika Felts. 2008. "Testing and Social Stratification in American Education." *Annual Review of Sociology* 34:385–404.

Hausmann, Leslie R. M., Janet Ward Schofield, and Rochelle L. Woods. 2007. "Sense of Belonging

as a Predictor of Intentions to Persist Among African American and White First-Year College Students." *Research in Higher Education* 48:803–839.

Hill, Diane. 2008. "School Strategies and the "College-Linking" Process: Reconsidering the Effects of High Schools on College Enrollment." *Sociology of Education* 81:53–76.

Jennings, Jennifer L., David Deming, Christopher Jencks, Maya Lopuch, and Beth E. Schueler. 2015. "Do Differences in School Quality Matter More Than We Thought? New Evidence on Educational Opportunity in the Twenty-first Century." *Sociology of Education* 88:56–82.

Kane, Thomas J. 1994. "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education." *Journal of Political Economy* 102:878–911.

Kao, Grace and Jennifer S. Thompson. 2003. "Racial and Ethnic Stratification in Educational Achievement and Attainment." *Annual Review of Sociology* 29:417–442.

Karp, Melinda Mechur and Katherine L. Hughes. 2008. "Information networks and integration: Institutional influences on experiences and persistence of beginning students." *New Directions for Community Colleges* 2008:73–82.

Kim, Dongbin. 2007. "The Effect of Loans on Students' Degree Attainment: Differences by Student and Institutional Characteristics." *Harvard Educational Review* 77:64–100,127.

Knapp, Laura G., Kelly-Reid Janice E. and Scott A. Ginder. 2012. "Enrollment in Postsecondary Institutions, Fall 2011; Financial Statistics, Fiscal Year 2011; and Graduation Rates Selected Cohorts, 2003-2008." Technical report, National Center for Education Statistics.

Kuh, George D.; Kinzie, Jillian; Buckley Jennifer A.; Bridges Brian K.; Hayek John C. 2006. "What Matters to Student Success: A Review of the Literature." Technical report, National Postsecondary Education Cooperative.

Lee, Jaekyung. 2002. "Racial and Ethnic Achievement Gap Trends: Reversing the Progress Toward Equity?" *Educational Researcher* 31:3–12.

Light, Audrey and Wayne Strayer. 2000. "Determinants of College Completion: School Quality or Student Ability?" *Journal of Human Resources* 35:299–332.

Light, Audrey and Wayne Strayer. 2002. "From Bakke to Hopwood: Does Race Affect College Attendance and Completion?" *Review of Economics and Statistics* 84:34–44.

Maralani, Vida. 2013. "The Demography of Social Mobility: Black-White Differences in the Process of Educational Reproduction." *American Journal of Sociology* 118:1509–1558.

Mumper, Michael. 2003. "The Future of College Access: The Declining Role of Public Higher Education in Promoting Equal Opportunity." *Annals of the American Academy of Political and Social Science* 585:97–117. ArticleType: research-article / Issue Title: Higher Education in the Twenty-First Century / Full publication date: Jan., 2003 / Copyright © 2003 American Academy of Political and Social Science.

Reardon, Sean F., Joseph P. Robinson-Cimpian, and Ericka S. Weathers. 2015. "Patterns and Trends in Racial/Ethnic and Socioeconomic Academic Achievement Gaps." In *Handbook of Research in Education Finance and Policy*, edited by Helen F. Ladd and Margaret E. Goertz. Routledge.

Roderick, Melissa, Vanessa Coca, and Jenny Nagaoka. 2011. "Potholes on the Road to College High School Effects in Shaping Urban Students' Participation in College Application, Four-year College Enrollment, and College Match." *Sociology of Education* 84:178–211.

Rosenbaum, James E., Regina Deil-Amen, and Ann E. Person. 2007. *After admission: From college access to college success*. Russell Sage Foundation.

Scott-Clayton, Judith E. 2011. "The Shapeless River: Does a Lack of Structure Inhibit Students' Progress at Community Colleges?" .

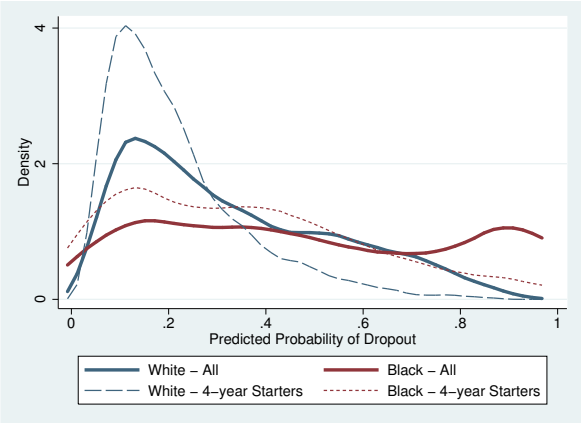
Small, Mario L. and Christopher Winship. 2007. "Black students' graduation from elite colleges: Institutional characteristics and between-institution differences." *Social Science Research* 36:1257–1275.

Snyder, Thomas D. and Sally A. Dillow. 2012. "Digest of Education Statistics 2012." Technical report, U.S. Department of Education.

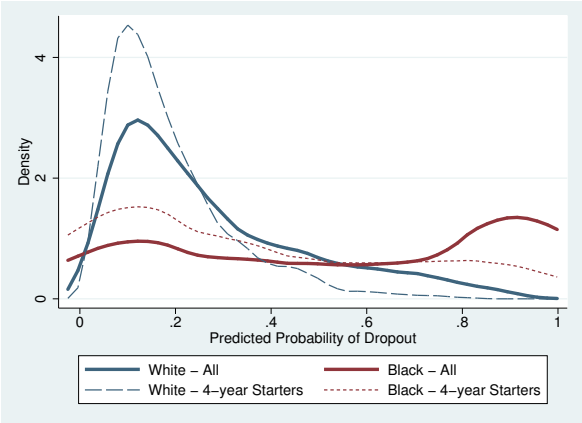
Swail, Watson Scott. 2003. *Retaining Minority Students in Higher Education: A Framework for Success. ASHE-ERIC Higher Education Report. Jossey-Bass Higher and Adult Education Series.* Jossey-Bass.

Figure 1: Initial Distribution of Four-year Dropout Risk by Race and Gender

(a) All Students



(b) Women



(c) Men

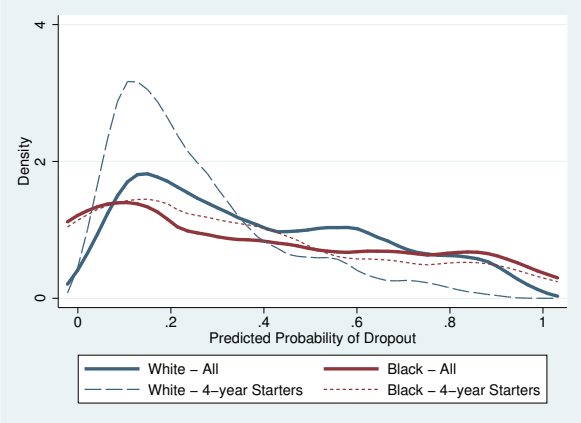
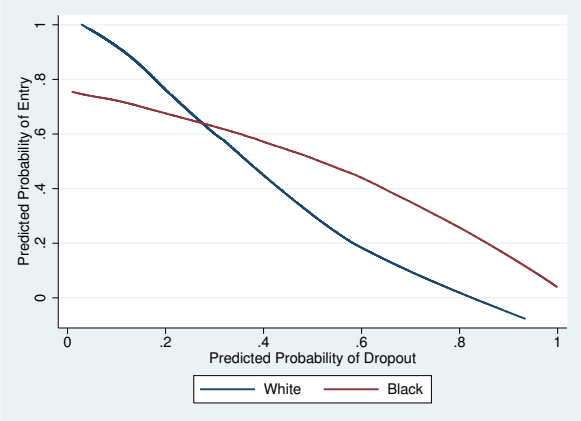
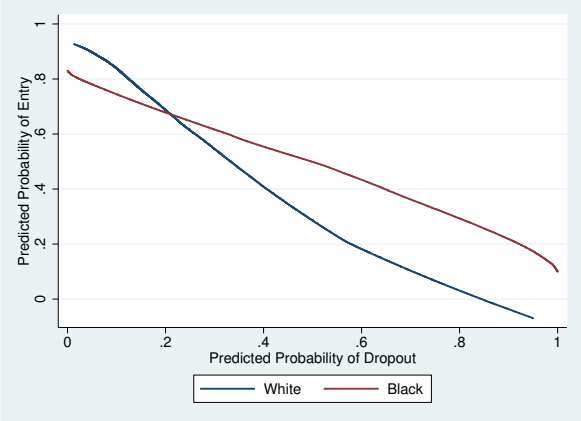


Figure 2: Four-year Entry Probability as a Function of Covariates, Given Pre-College Dropout Risk By Race and Gender

(a) All Students



(b) Women



(c) Men

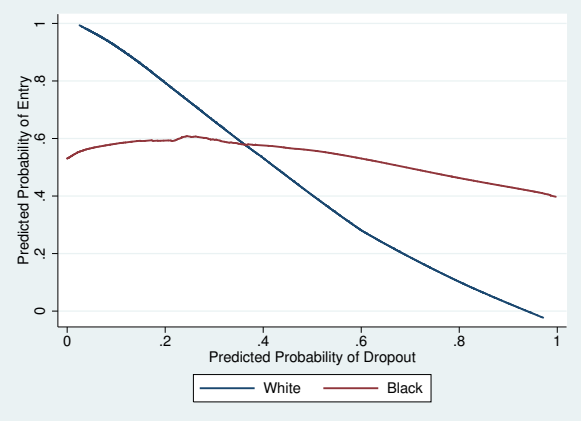
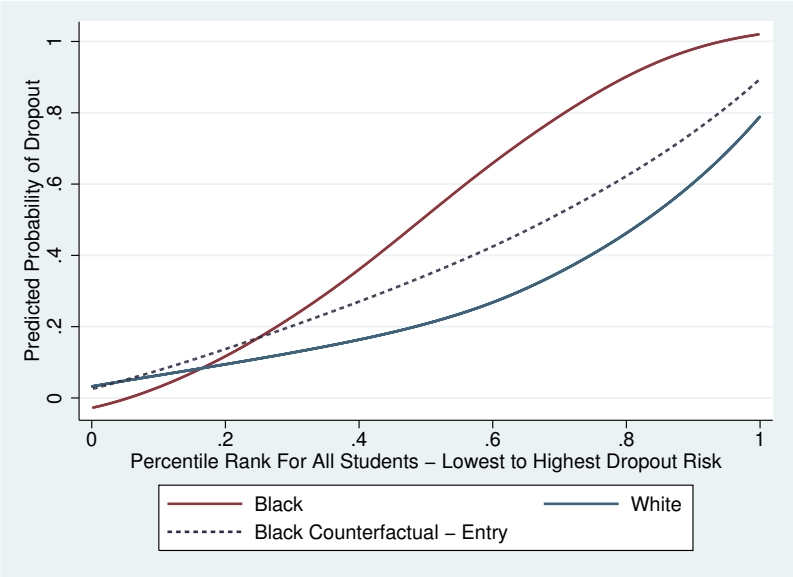


Figure 3: Percentile Ranking of Black and White Students by Actual Dropout Risk Values and Counterfactual Dropout Risk Values for Black Students

(a) Women



(b) Men

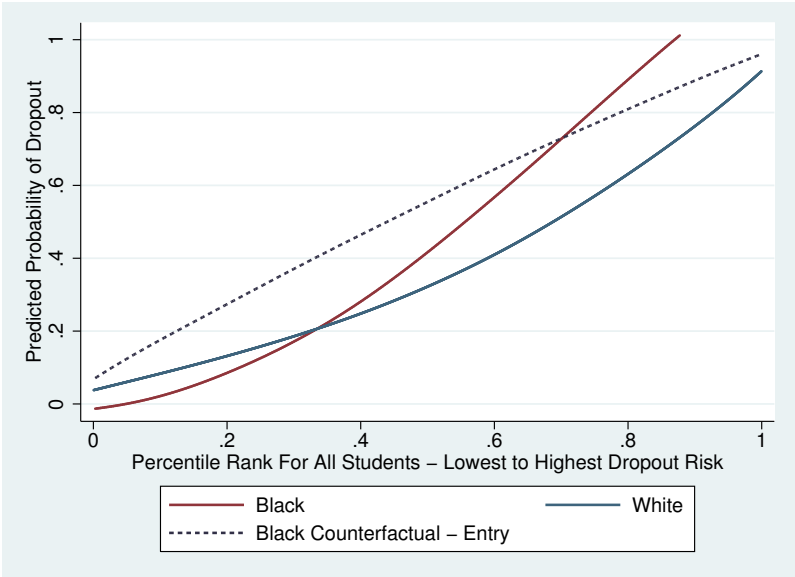
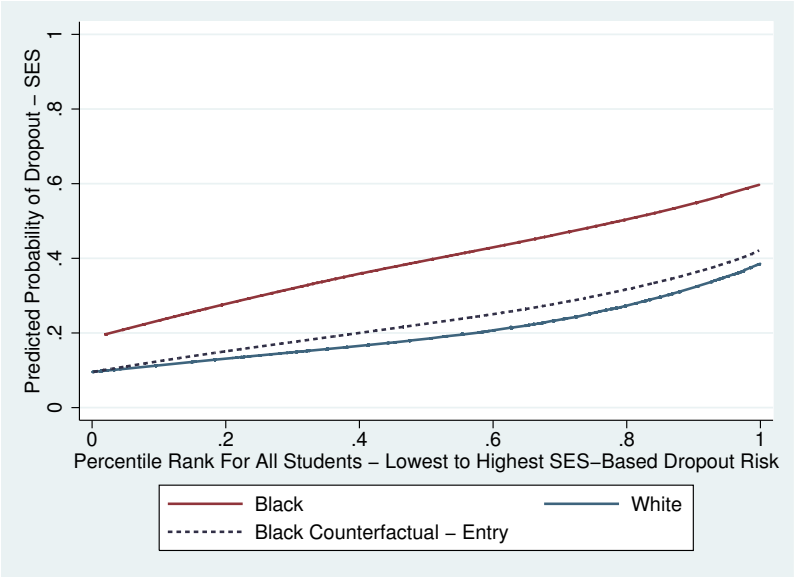


Figure 4: Percentile Ranking of Black and White Students by Actual SES-Based Dropout Risk Values and Counterfactual SES-Based Dropout Risk Values for Black Students

(a) Women



(b) Men

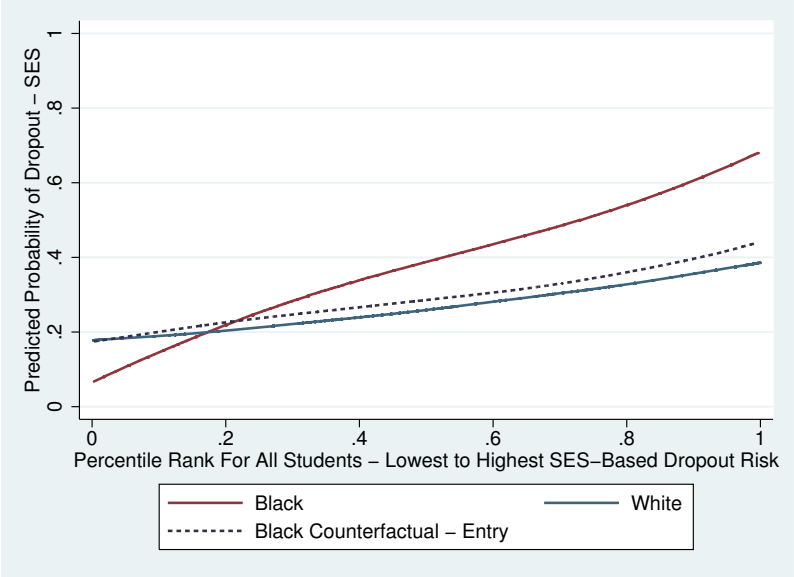
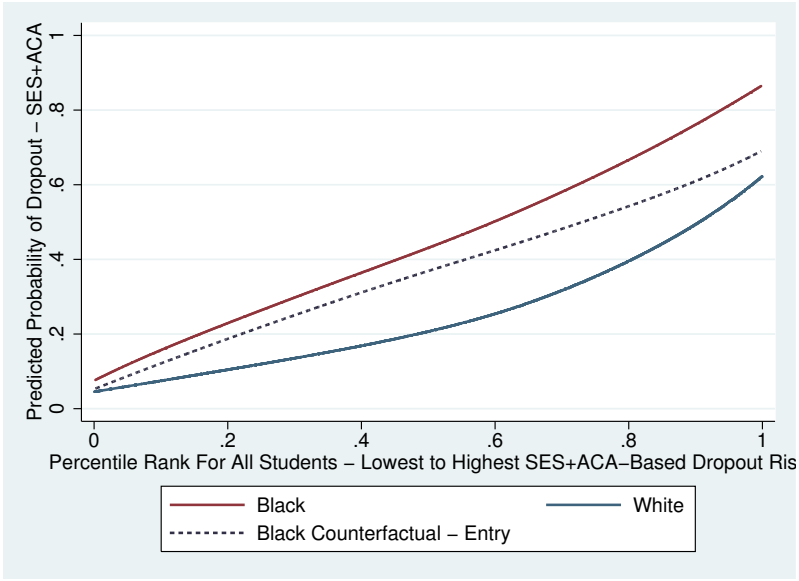


Figure 5: Percentile Ranking of Black and White Students by Actual SES- and Academic-Based Dropout Risk Values and Counterfactual SES- and Academic-Based Dropout Risk Values for Black Students

(a) Women



(b) Men

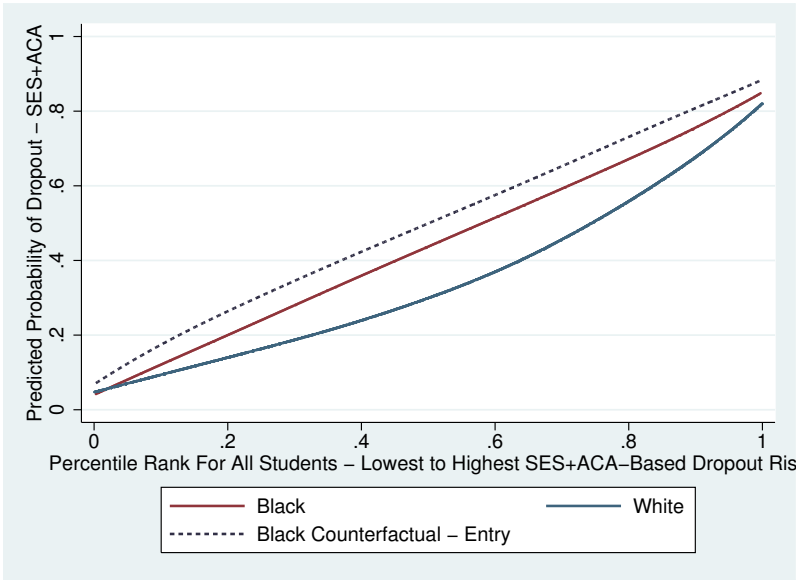
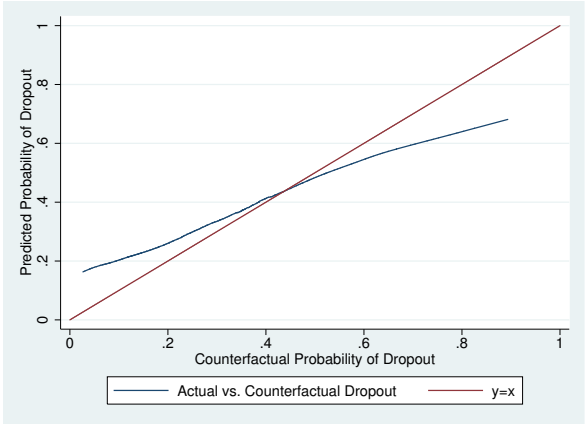
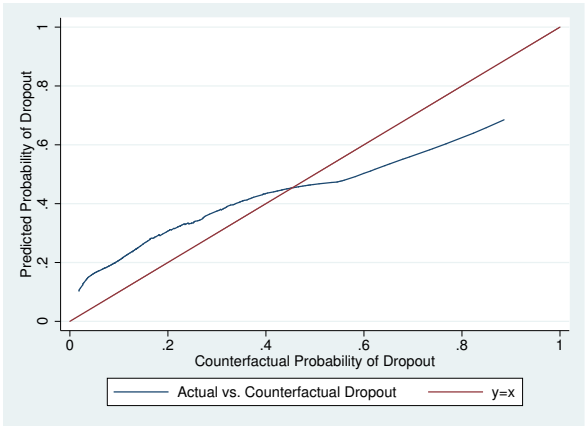


Figure 6: Actual and Counterfactual Dropout Risk for Black Students

(a) All Students



(b) Women



(c) Men

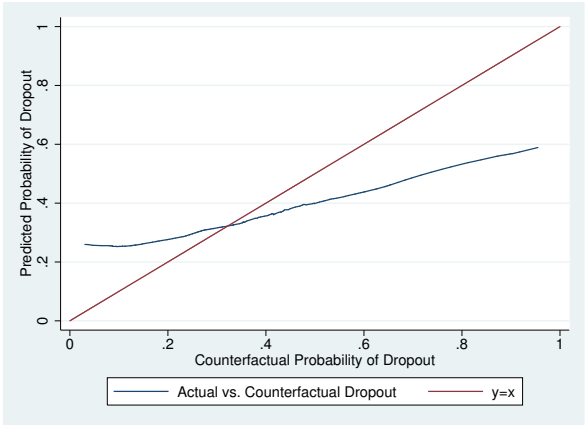
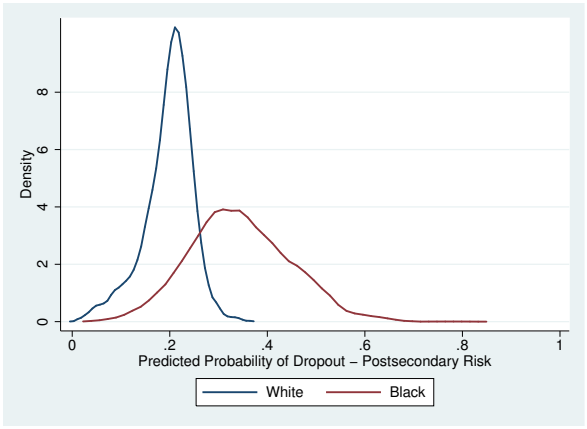
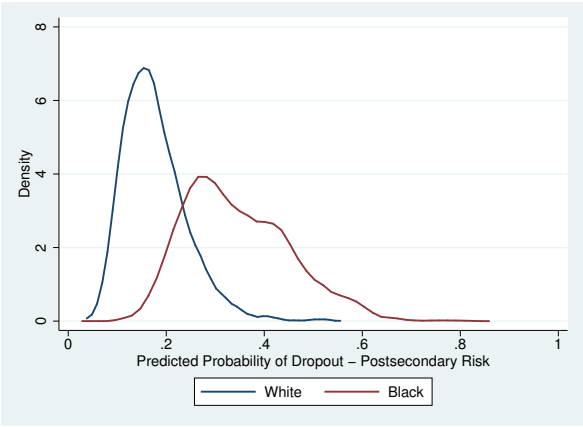


Figure 7: Distribution of Purged Postsecondary Dropout Risk by Race and Gender

(a) All Students



(b) Women



(c) Men

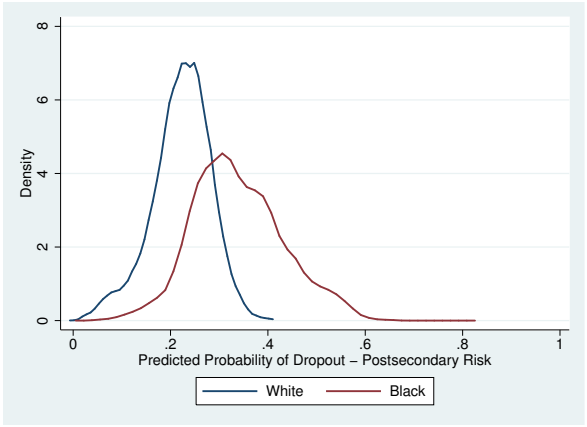
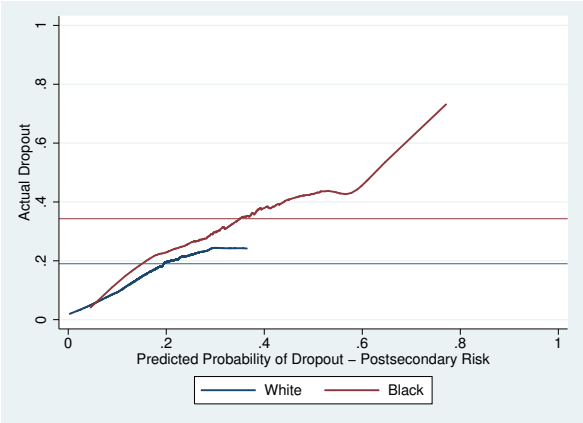
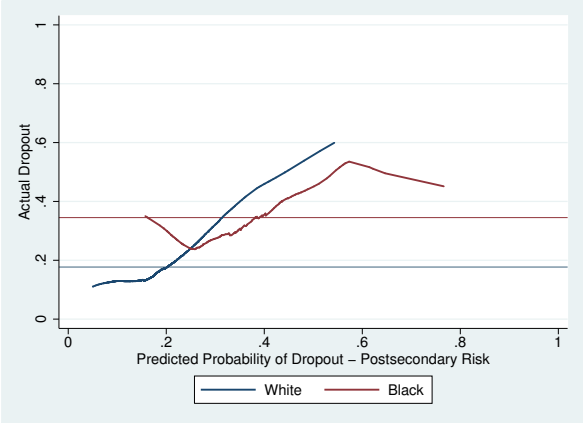


Figure 8: Purged Postsecondary Dropout Risk and Actual Dropout by Race and Gender

(a) All Students



(b) Women



(c) Men

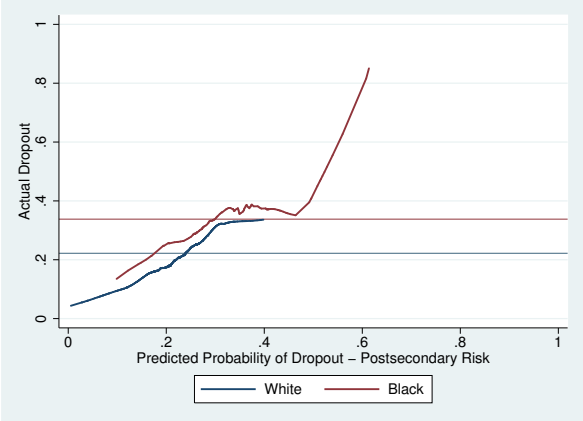


Table 1: Comparison of High School Completion and Recent College Entry Rates Reported By ELS and NCES

	ELS	NCES
	High School or GED Completion	
Black	.933	.925
White	.976	.957
	Enrollment in Two- or Four-year College	
Black	.563	.625
White	.710	.688

Note: NCES figures for high school or GED completion taken from Table 219.70; figures for enrollment in two- or four-year colleges taken from Table 302.20. "Recent" is defined as college entry within one year of high school diploma or GED receipt and is measured as of October 2004.

Table 2: Trajectories of High School Diploma and GED Recipients in the ELS (*calculations are conditional on high school diploma or GED and include entry at any age observable in the ELS*)

	Two-year College Beginner	Four-year College Beginner	Transfer from Two- to Four-year College
Black	.349	.452	.126
White	.305	.553	.113

Note: Numbers reflect list-wise deletion.

Table 3: Probability of Entry into Four-year College and Counterfactual Probability Using Other Race's Response to Own Pre-College Situation (*calculations are conditional on high school diploma or GED and include entry at any age observable in the ELS*)

	Actual	Other Race Entry Response to Own Covariates	Other Race Entry Response to Own Pre-College Risk
		All Students	
Black Students, White Entry Decision Model	.452	.407	.362
White Students, Black Entry Decision Model	.553	.616	.486
		Women	
Black Students, White Entry Decision Model	.433	.396	.419
White Students, Black Entry Decision Model	.572	.604	.512
		Men	
Black Students, White Entry Decision Model	.474	.383	.337
White Students, Black Entry Decision Model	.533	.661	.478

Notes: Own covariates are pre-college sociodemographic, academic and behavioral measures from the ELS. Pre-entry dropout risk is computed via race-specific logistic regressions. Counterfactual estimates are the mean of individual-level predictions.

Table 6: Actual Probability of Degree Non-Completion Given Four-year College Entry and Counterfactual Probability Based on Own Race Characteristics and Other Race Entry Behavior (*calculations are conditional on high school diploma or GED and include entry at any age observable in the ELS*)

	Actual	Other Race Entry Response to Own Covariates	Other Race Entry Response to Own Pre-College Risk
		All Students	
White Entry Decision Model, Black Students' Variables	.358	.355	.377
Black Entry Decision Model, White Students' Variables	.217	.228	.279
		Women	
White Entry Decision Model, Black Students' Variables	.361	.364	.454
Black Entry Decision Model, White Students' Variables	.186	.188	.231
		Men	
White Entry Decision Model, Black Students' Variables	.355	.345	.343
Black Entry Decision Model, White Students' Variables	.252	.289	.380

Notes: Own covariates are pre-college sociodemographic, academic and behavioral measures from the ELS. Pre-entry dropout risk is computed via race-specific logistic regressions. Counterfactual estimates are the mean of individual-level predictions.

Table 4: Average Probability of Entry into Four-year Colleges By Decile: White Entry Decision Model

	1	2	3	4	5	6	7	8	9	10
	All Students									
White Students' Decision Model, White Students' Variables	0.018	0.064	0.191	0.489	0.706	0.815	0.883	0.924	0.950	0.972
White Students' Decision Model, Black Students' Variables	0.009	0.028	0.081	0.193	0.366	0.529	0.669	0.790	0.884	0.948
	Women									
White Students' Decision Model, White Students' Variables	0.014	0.061	0.214	0.511	0.718	0.836	0.899	0.936	0.959	0.978
White Students' Decision Model, Black Students' Variables	0.007	0.023	0.063	0.174	0.350	0.515	0.681	0.805	0.897	0.955
	Men									
White Students' Decision Model, White Students' Variables	0.018	0.091	0.251	0.457	0.636	0.761	0.847	0.903	0.941	0.972
White Students' Decision Model, Black Students' Variables	0.009	0.036	0.086	0.189	0.301	0.433	0.571	0.720	0.851	0.941

Table 5: Actual Dropout Rates of Black and White Students and Counterfactual Dropout Risk for Black Students by Decile

	1	2	3	4	5	6	7	8	9	10
	All Students									
Black Students' Entry Decision Model, Black Students' Variables	0.677	0.768	0.648	0.523	0.508	0.422	0.443	0.320	0.220	0.181
White Students' Entry Decision Model, Black Students' Variables	0.750	0.642	0.578	0.590	0.520	0.419	0.362	0.290	0.185	0.110
White Students' Entry Decision Model, White Students' Variables	0.677	0.494	0.422	0.413	0.303	0.245	0.180	0.152	0.113	0.075
	Women									
Black Students' Entry Decision Model, Black Students' Variables	0.845	0.756	0.736	0.670	0.581	0.492	0.469	0.341	0.169	0.129
White Students' Entry Decision Model, Black Students' Variables	0.686	0.558	0.402	0.525	0.411	0.349	0.307	0.248	0.153	0.094
White Students' Entry Decision Model, White Students' Variables	0.583	0.381	0.368	0.323	0.246	0.206	0.171	0.139	0.112	0.075
	Men									
Black Students' Entry Decision Model, Black Students' Variables	0.444	0.242	0.518	0.508	0.486	0.372	0.308	0.297	0.314	0.262
White Students' Entry Decision Model, Black Students' Variables	0.871	0.787	0.788	0.697	0.594	0.490	0.462	0.345	0.290	0.132
White Students' Entry Decision Model, White Students' Variables	0.750	0.641	0.546	0.457	0.344	0.286	0.210	0.160	0.121	0.072

Table 7: Actual Probability of Degree Non-Completion Given Four-year College Entry and Counterfactual Probability Based on Own Race Characteristics, Other Race Entry Behavior, and Other Race Dropout Behavior (*calculations are conditional on high school diploma or GED and include entry at any age observable in the ELS*)

	Actual	Other Race Entry and Dropout Response to Own Covariates	Other Race Entry and Dropout Response to Own Pre-College Risk
		All Students	
White Entry and Dropout Decision Models, Black Students' Variables	.358	.321	.304
Black Entry and Dropout Decision Models, White Students' Variables	.217	.289	.280
		Women	
White Entry and Dropout Decision Models, Black Students' Variables	.361	.264	.248
Black Entry and Dropout Decision Models, White Students' Variables	.186	.262	.240
		Men	
White Entry and Dropout Decision Models, Black Students' Variables	.355	.389	.369
Black Entry and Dropout Decision Models, White Students' Variables	.252	.287	.296

Notes: Own covariates are pre-college sociodemographic, academic and behavioral measures from the ELS. Pre-entry dropout risk is computed via race-specific logistic regressions. Counterfactual estimates are the mean of individual-level predictions.

Appendix 1: ELS Variables

Pre-College	*Post-College Entry*
Gender	Applied for Aid
Age	Received Aid
Race	Total Student Loans Accrued
Family/Individual Income	Most Selective College Where Admitted
Father's Highest Education	Selectivity of Current College
Mother's Highest Education	First College Out of State
Parents Married	Enrollment Mostly Part-Time
Public High School	Took Remedial Math
High School GPA	Took Remedial English
Level of High School Math	Major Field of Study
Level of High School English	Held Job at Some Point During College
Level of High School Science	Participated in "High Impact" Activities
Level of High School Soc. Stud.	Number of Negative Life Events
Academic Intensity Index	Participated in Extracurriculars
Talk About High School Courses with Parents	Participated in Varsity Sports
Highest Education Expected by Student	Participated in Intramural Sports
Level of College Entry Expected by Student	Used School Library for Studying
No Plans for College	Use School Website and Library for Materials
Number of Friends Dropped Out of High School	Talked with Advisor about Academic Plans
Importance of College Financial Aid	Talked with Faculty Outside of Class
Importance of College Course Offerings	
Importance of Living at Home During College	
Importance of College Job Placement Record	
Importance of Academic Reputation at College	
Importance of Easy Admission into College	
Importance of Racial Diversity in College	
Importance of Earning Lots of Money	
Importance of Getting Steady Work	
Importance of Getting Away from Home	
Importance of a Good Education	