Are Students Dropping Out or Are They Simply Dragging Out the College Experience?

Persistence at the Six-Year Mark

By

Leslie S. Stratton

James N. Wetzel*

February 2012

RH: Dropping Out or Dragging Out College?

Leslie S. Stratton, Professor, Department of Economics, Virginia Commonwealth University, 301 W. Main Street., P.O. Box 844000, Richmond, VA 23284-4000. Phone: 804-828-7141. FAX: 804-828-9103. E-Mail: lsstratt@vcu.edu. Research Fellow at IZA, Bonn, Germany. Corresponding author.

James N. Wetzel, Professor, Department of Economics, Virginia Commonwealth University, 301 W. Main Street., P.O. Box 844000, Richmond, VA 23284-4000. E-Mail: jnwetzel@vcu.edu.

* This material is based upon work supported by the Association for Institutional Research, the National Center for Education Statistics, the National Science Foundation, and the National Postsecondary Education Cooperative, under Association for Institutional Research Grant Number RG10-128. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Association for Institutional Research, the National Center for Education Statistics, the National Science Foundation, or the National Postsecondary Education Cooperative. We thank participants at the 2011 AIR Forum and the University of Cergy Pontoise for helpful suggestions.

Are Students Dropping Out or Simply Dragging Out the College Experience?

Persistence at the Six-Year Mark

ABSTRACT

Standard analyses of college outcomes look at six-year graduation rates, treating all non-graduates alike as 'failures.' However, 36% of these non-graduates are still enrolled. Using a micro-level data set that includes rich information on demographic and academic background, we employ a multinomial logit model to distinguish between graduates, persisters, and dropouts. We find there are significant differences between persisters and dropouts. Not surprisingly, poor academic ability/background, particularly high school GPA, increases the probability of dropout substantially, but poorly prepared students are also more likely to drag out the college experience. Furthermore, there are still substantial differences by socioeconomic status, as well. Hispanics are significantly more likely to persist, while first generation college students are at greater risk of dropping out. Evidence suggests that as many as half of those persisting at the six-year mark will eventually graduate. Differences between persisting and dropping out should not to be ignored.

JEL: I24, I2

Are Students Dropping Out or Simply Dragging Out the College Experience?

Persistence at the Six-Year Mark

1. INTRODUCTION

The traditional approach to measuring 'success' in college by focusing purely on graduation rates may be misleading. While four-year graduation rates were once the metric of choice, six-year rates are now the norm. However, even looking at six-year graduation rates fails to recognize the substantial heterogeneity amongst those who have not graduated. Some of these individuals are best classified as dropouts. They gave college life a try but left higher education, frequently relatively quickly, and do not seem to be looking back. Others have persisted and may simply be taking longer to graduate. It is important to know who graduates in six years; it is also important to know what happens to the non-graduates. We extend the standard literature on college graduation to distinguish between non-graduates who are and are not still enrolled.

Of those who have not graduated at the end of six years, we find that 36% are actually still enrolled. This suggests that treating all non-graduates alike may be statistically misleading and miss important information. Identifiable differences between graduates and non-graduates have been highlighted by policy makers interested in boosting graduation rates and ensuring equal opportunity to all. Further distinguishing between persisters and dropouts may be just as valuable both for future economic growth and for equity considerations, especially given policy interest in enrolling and graduating those subsets of the population historically underrepresented in the pool of college graduates. Unfortunately most research utilizes a standard logit specification that by definition treats all those who have not graduated as 'failures,' as one

homogeneous population. We present alternative estimates that employ a multinomial logit specification to distinguish between persisters and dropouts.

The paper follows with a brief review of prior research and a description of the nationally representative cohort of first-time undergraduates used here. Raw graduation rates are shown, with a particular focus on those from less advantaged socioeconomic classes who have been historically underrepresented in college. We then utilize a standard logit to model six-year graduation rates with and without controls for academic background and ability so as to establish a base case and to compare our data/results with the extant literature. Next, we move into the heart of the paper, estimating a multinomial logit model that distinguishes among graduation, persistence, and drop out. This second approach enables us to determine whether persistence and dropout are distinct states and what characteristics are associated with each outcome. Of particular interest are the results for those of low socioeconomic status. Finally, we present some evidence regarding the likelihood of these persons eventually graduating and discuss policy implications.

Our standard logit results match those reported elsewhere in the literature. We find that controlling for academic preparation/ability substantially reduces the gap in graduation rates between more and less advantaged populations. There remains, however, a significant 6-11 percentage point differential in graduation rates for students from lower income families as well as for students whose parents have less formal education. More importantly, our multinomial logit results indicate that those who are still enrolled six years following matriculation are statistically different from those who are no longer enrolled as well as from those who have graduated. Even after controlling for academic ability/background, the marginal impact of socioeconomic status on persistence differs across the population. For example, Hispanics who

have not graduated are significantly more likely to still be enrolled than non-Hispanics. On the other hand, first generation college students have a higher probability of non-enrollment at the six year mark as compared to students with more highly educated parents. A substantial fraction of persisters are enrolled part-time and supplementary information on time-to-degree suggests that a substantial fraction may eventually graduate. Hence, further knowledge concerning the distinction between persistence and dropout may have useful policy implications.

2. LITERATURE REVIEW

While theory (Becker 1964) suggests that only students who expect the benefits of pursuing a college degree to exceed the costs will enroll in college, statistics clearly indicate that initial enrollment does not guarantee graduation. Lee and Rawls (2010) report that on average only 56.1% of those entering college with the intent of earning a bachelor's degree graduate within six years. Graduation rates are even lower for socioeconomic groups that have been historically underrepresented at college. Socioeconomic status is captured here by race/ethnicity, parental education, and family income. The six year graduation rate for African Americans is 40.5%, for Hispanics 46.8%, for those whose parents have no more than a high school degree 43.4%, and for those with low family income 50.3%.[1]

Researchers have typically cited changing perceptions about the benefits and costs associated with higher education to help explain these low graduation rates. Expectations may change as students obtain new information about their expected returns (see Altonji 1993 for a model of such decision making under uncertainty and Manski 1989). New information could arrive regarding their relative academic ability, college costs, or future returns. Students coming from less advantaged households may have had less accurate information when making the

initial decision to enroll and thus may be more likely to update their expected net benefits as they obtain more information. As parental education is in general lower for these populations, their knowledge of the process and its net benefits is probably less accurate as compared to those from more advantaged backgrounds.

There is a substantial literature pertaining to graduation in the education field (see Kuh et al. 2006 for a review). These studies often employ data on students from only a single institution (see for example, DesJardin, Ahlburg, and McCall 1999), even though research suggests that as many as 60% of all undergraduates attend multiple institutions (Adelman 2006). Some notable exceptions include Adelman (2006) who uses NELS data and Cragg (2009) who uses the same data employed here. Both Adelman and Cragg demonstrate the importance of controlling for academic background. Adelman in particular argues that controls for test scores, high school grades, and high school curriculum are all important and that the impact of academic background dominates the impact of socioeconomic status, particularly race and ethnicity.[2]

Economics articles addressing college completion are less common and often focus on aggregate data looking for time trends. Turner (2004) documents that graduation rates have declined in recent years and reports some circumstantial evidence that student achievement, as measured by math scores and GED completion rates, may explain some of the differential. Later work by Bound, Lovenheim, and Turner (2010) suggests that differential preparedness explains much of the declining graduation rate for those entering two-year institutions, but less so for those entering four year institutions, which is not to say that academic preparation is not important. Bailey and Dynarski (2011) demonstrate a substantial and increasing association between family income and college completion, but similarly argue that while cognitive ability is

an important determinant of success, changes in reported cognitive ability do not explain the changing relation between family income and graduation.[3]

Graduation, in much of this literature, is modeled as a binary outcome occurring within a fixed time frame or by a certain age. Turner (2004) looks at twenty-three year olds. Bound, Lovenheim, and Turner (2010) focus on college completion rates eight years following high school graduation. Bailey and Dynarski (2011) look at college completion by age twenty-five. Those who have not graduated within that time frame are treated as a homogenous population.

Work in the related persistence literature suggests this assumption may be invalid. The persistence literature examines dropout within the pivotal first year following matriculation. Researchers typically use a logit model to distinguish between those enrolled and those not enrolled one year or one term following matriculation. Stratton, O'Toole, and Wetzel (2008), however, use a multinomial logit specification to distinguish among those enrolled one year following matriculation, those not enrolled at the one year mark who reenroll within one year, and those not enrolled at the one year mark who remain unenrolled for over a year. They find that a substantial fraction of those not enrolled at the one year mark do reenroll quickly and that there are significant differences between those who reenroll within the subsequent year and those who do not. Thus, there is evidence from the persistence literature of heterogeneity among the population of non-persisters. Furthermore, both Turner (2004) and Bound, Lovenheim, and Turner (2010) report evidence that time-to-degree has increased, suggesting that relying on fixed time horizons may paint an incomplete picture of college degree receipt.

Our primary contribution to the literature is the introduction of a multinomial logit specification (MNL) to model college success. This specification enables us to distinguish between non-graduates who are still enrolled in college and those who are not. Furthermore, the

representative national data sample of younger college students we use follows students as they move between institutions, giving us a more accurate outcome measure as compared to the common institution-specific analyses conducted in the education field.

3. DATA

Specifically, the data employed in this analysis come from the restricted access 1996-2001 Beginning Postsecondary Survey (BPS) collected by the National Center for Educational Statistics (NCES) of the Department of Education. These data constitute a nationally representative sample of students who first matriculated to a postsecondary institution in the 1995-1996 academic year. We restrict our analysis to those individuals interviewed through Spring 2001 so that we have adequate time to track progress and focus on academic programs culminating in a Baccalaureate degree. Per common practice, initial enrollment is limited to those attending four-year institutions, so as to avoid problems introduced by the heterogeneity of intent amongst community college students. These restrictions yield a sample of about 6190 individuals.[4]

Information on academic preparation and student ability is critical for this analysis. These data are missing for a substantial fraction of older students and those not from the United States. As a result, students from abroad and students age 23 and older are excluded from the analysis.[5] These restrictions leave a final estimation sample of about 5820 individuals. Sample statistics for this population are reported in Table 1. Given the stratified nature of the sample, all the results reported here utilize the BPS longitudinal weights and all statistical estimates are corrected for the BPS's complex survey design.

The outcome measures for our analysis are derived using information on Baccalaureate degree receipt and college enrollment at the conclusion of Spring 2001. The six-year graduation rate as calculated from this sample is 63.2%. By comparison, the NCES reports that the six-year graduation rate for this cohort was 55.4%.[6] The difference between these rates is primarily driven by NCES counting as graduates only those who earn their degrees from the post-secondary institution first attended. The fraction graduating from the institution first attended in the sample employed here is a very similar 54.9%. From a national policy perspective, it is important to recognize the substantial numbers graduating from other institutions.

A key concern here is the enrollment status of those who have not graduated. If few of these students are still pursuing a degree, further analysis is unnecessary as a standard logit would suffice. In fact, however, we find that 36% of those who have not graduated, or 13.3% of the entire sample, are still enrolled in Spring 2001. On average the enrollment patterns of those who are and are not still enrolled are quite different. About 50% of those not enrolled in Spring 2001 have enrolled for no more than two of the six years following matriculation. Conversely, this is true for only 3% of those who are still enrolled in Spring 2001. Eighty-three percent of those who are enrolled in Spring 2001 have been enrolled for four or more academic years; less than 25% of those who are no longer enrolled were enrolled that long. Twenty percent (almost 40%) of those not enrolled in Spring 2001 have not been enrolled since the first (the second) calendar year after they matriculated. Further analysis of those still enrolled, especially relative to those who have left the higher education system, requires use of a multinomial rather than a standard logit.[7]

In addition to enrollment data, detailed personal information is available for every respondent. This includes information on gender, race, ethnicity, and age; state of residence;

marital and parental status; and parental education and income. State of residence is used to match the state's unemployment rate in the year 1999 to the sample. Higher unemployment rates imply a lower opportunity cost associated with college enrollment and may attract a different population of students. Regional dummies are also incorporated. Parental education is identified based on the reported education of the most educated parent. College degree receipt is the modal response. Very few parents were high school dropouts. We distinguish between those parents with no more than a high school degree, those with some college, and those with a post-graduate degree using dummy variables. First generation college students are variously defined in the literature as either those whose most educated parent has no more than a high school degree or those whose most educated parent has less than a college degree: our specification accommodates both definitions. A dummy variable is used to identify respondents who declare they are independent of their parents, and income dummies that approximately split the population into quartiles are employed to allow a non-linear income effect. The highest income quartile is treated as the base case.

Academic preparation/ability is captured using a number of different variables, as suggested by Adelman (2006). A dummy variable to indicate high school degree receipt is incorporated to identify graduation and perhaps the character trait 'persistence'. Less than 2% of our sample has no high school degree. A measure of the most advanced math course the student plans to take in high school is included to capture the rigor of the student's high school curriculum. Trigonometry is the base case. Alternative specifications using the NCES coding of the quality of the student's high school curriculum yield substantially the same results. Standardized combined SAT test scores and self-reported high school GPA are used to assess individual ability. Students taking the ACT are identified with a dummy variable and their ACT

scores converted to SAT scores using a concordance table published by the College Board (1999). Grades are self-reported, since high school transcripts were not available, and such reports are likely biased upward (more students report an A average than any other outcome). Some students fail to report these academic background/ability measures and dummy variables are used to identify these cases. However, missing data are far less of a problem with the BPS data than with the NELS or NLSY72 data used by Bound, Lovenheim, and Turner (2010). Where those data sets were missing some measure of academic background for 40% of their samples, we are missing some measure for only 12% of our sample.

4. ANALYSIS FRAMEWORK

As these restricted access data are not generally used in economics, we begin by demonstrating that these data yield results comparable to those observed in the prior literature. First, we present raw statistics regarding the impact of socioeconomic status on graduation rates. As described above, we measure socioeconomic status using race, ethnicity, parental education, and household income. Second, we run a standard logit model that controls simultaneously for gender, race, ethnicity, parental education, household income, age, unemployment rate, region of residence, and marital and parental status. Using these results we recalculate the marginal impact of socioeconomic status on the graduation rate. Finally, we add controls for academic background/ability and recalculate the marginal impact of socioeconomic status to determine the degree to which academic preparedness rather than socioeconomic status per se influences graduation rates in the standard logit.

Our primary concern, however, remains the enrollment status of those who have not graduated by Spring 2001. To take into consideration the substantial fraction that are still

enrolled six years following matriculation, we utilize a multinomial logit model (MNL) that differentiates between the non-graduates who are and are not still enrolled. The application is much like that in Stratton, O'Toole, and Wetzel (2008) who use a MNL specification to distinguish between continued enrollment, stopout, and dropout in the first year of college. This analysis will allow us to determine whether some less advantaged populations might have lower graduation rates, not because they are no longer enrolled, but rather because they are simply taking longer to graduate. Finally we conduct some sensitivity analysis and use information from the 2001 Baccalaureate and Beyond Longitudinal Study (B&B:2000/01) to assess the probability with which those who have not graduated may do so.

5. CONVENTIONAL ANALYSIS: SIMPLE LOGIT

Figure 1 summarizes the results of a conventional analysis of graduation rates, showing the marginal impacts on six-year graduation rates of three common socioeconomic classifications: race/ethnicity, parental education, and household income. African Americans are compared to whites and Hispanics to non-Hispanics. The base case for parental education constitutes individuals whose most educated parent has a post-graduate degree. The base case for household income constitutes those with a family income over $75,000.[8] The darkest bars show raw results when one factor, such as income, is considered in isolation. These raw statistics are often discussed in the press, but may be subject to misinterpretation as they fail to take into account the fact that household income and parental education are not uniformly distributed by race or ethnicity. The second bar shows the results of the standard logit that controls for all these various socioeconomic factors jointly, as well as gender, age, marital and parental status, region of residence, and the unemployment rate. The third bar illustrates the marginal effect of these

socioeconomic indicators after additionally controlling for academic ability/background. Detailed statistical results are available upon request.

The raw statistics indicate substantial differences. African Americans are over 20% less likely to graduate than whites, while those from the lowest income quartile are more than 25% less likely to graduate than those from the highest income quartile. These raw differentials, particularly that for income, decline substantially once one controls for these demographic characteristics simultaneously, but remain sizeable (at around 10%) and statistically significant. Adding controls for academic ability/background, however, actually causes the marginal effects of race and ethnicity to approach zero. The marginal effects of parental education and household income also decline but not as much. These two historically underrepresented groups remain significantly less likely to graduate and, in the case of parental education, the differential is still on the order of 15%. These results are comparable to those reported by Adelman (2006) and Cameron and Heckman (2001).

6. ADVANCED ANALYSIS: MNL

The use of a MNL specification enables us to distinguish between those who did not graduate but are still enrolled in Spring 2001 (also called ‘persisters’) and those who did not graduate and are no longer enrolled in Spring 2001 (the ‘not enrolled’ or ‘drop outs’). The fraction of the entire sample that is persisting in the final term is 13%, but this fraction is usually higher for those from less advantaged socioeconomic backgrounds. For example, 19% of African Americans and 17% of those with the lowest household income are still enrolled. The lower graduation rate observed for these populations six years following matriculation may be offset in the long run, if these students do subsequently graduate.

Results from a MNL including not just demographic, but also academic ability/background controls are reported in Table 2. The first row indicates the predicted probability, given the base case characteristics, of graduating, persisting, and not enrolling respectively. Thus, an individual with base case characteristics has on average a 73% probability of having graduated, an 11% probability of still being enrolled, and a 16% probability of no longer being enrolled six years following matriculation. The rows that follow report the numerical marginal effects associated with each characteristic on each of these outcomes. As expected, the predicted marginal impact of each characteristic on the probability of graduating is almost exactly that generated by the standard logit specification underlying Figure 1. Thus, we focus the discussion on the two outcomes not addressed by the standard logit model, namely persistence and dropout. The discussion centers on whether a characteristic is associated with a differential persistence rate or a differential dropout rate or both. Which behavior is impacted the most is a useful finding, especially if students with different characteristics follow different paths. The marginal differences are all relative to the base case probability.

We begin by noting the results clearly indicate that the factors included in our model are jointly statistically significant in distinguishing between graduation and both persistence and non-enrollment (p-values both 0.00). More importantly, non-enrollment and persistence are also found to be statistically distinct outcomes (p-value 0.00). These findings suggest that the MNL specification is to be preferred over the standard logit.

Looking at the results in Table 2, we note striking differences in the predicted distribution of behaviors by those who have not graduated by socioeconomic status. First, holding all else constant, Hispanics are significantly more likely to still be persisting than non-Hispanics – on average 4 percentage points or almost forty percent more likely. Overall, it appears that

Hispanics who have not graduated in six years may not have dropped out, but are rather dragging out the time-to-degree. Second, African Americans have a somewhat higher probability of persisting as compared to whites, but also a larger relative chance of not enrolling. Neither impact, however, is statistically significant at conventional levels once academic ability is taken into account. Third, first generation college students are on average significantly less likely to have graduated than non-first generation college students because they are so much more likely to not be enrolled (9 to 10 percentage points or about 60% more likely). First-generation college students are not significantly more likely to persist than those with more educated parents. Fourth, individuals from households with lower income are less likely to graduate than those with higher income. The strongest effect is for those from the second quartile who are 5% less likely to graduate and 4.6% more likely to have dropped out than those from the third quartile. While those from the lowest quartile are significantly less likely to graduate than those from the highest quartile, they are not significantly less likely to graduate than those from the third quartile and are neither significantly more likely to have persisted or to have dropped out. Analysis incorporating controls for first year financial aid packages suggests these results may be related to grant aid receipt. Controlling for grant aid receipt, which has a significant positive association with graduation and a significant negative association with drop out , those from the lowest income quartile are significantly less likely to graduate and significantly more likely not to be enrolled as compared to those from the third quartile.

Looking at other demographic variables, the results in Table 2 indicate that on average women are significantly more likely to graduate than men (5 percentage points more likely). Since women are more likely to graduate, the question then becomes are women are less likely to drop out and/or are women are less likely to drag out the college experience relative to men. In

fact, women are less likely both to drop out and to persist relative to men, but the marginal effect on dropout behavior is over 50% larger than the marginal effect on persistence. Finally, age is significantly and substantially correlated with six year outcomes. Although we capture individuals no older than age 22 when they enter college, the results clearly indicate that older individuals are less likely to graduate because they are more likely to drop out.

As was the case with the standard logit analysis of graduation, educational ability/background plays a critical role. The probability of graduating drops 9 percentage points for individuals moving from the highest to the lowest expected high school math curriculum, 31 percentage points for individuals moving from the highest to the lowest high school GPA, and 10 percentage points for individuals moving from the highest to the lowest SAT score. Only parental education has as large an impact (15 percentage points). The effect of high school curriculum on graduation is solely offset by changes in non-enrollment. High school curriculum has no significant impact on the probability of persisting (p-value 0.22). While high school GPA has a larger impact on non-enrollment (23 percentage points) than on persistence (8 percentage points), its association with each is highly statistically significant (both p-values are 0.00). Finally, while lower than average test scores are primarily associated with higher rates of non-enrollment, higher than average test scores are primarily associated with lower persistence rates. These differences further emphasize the importance of utilizing a multinominal logit versus the traditional binomial logit to evaluate college outcomes.

To test the robustness of our results and to see if any patterns arise using different observation windows, we reran the analysis using (1) sixth year outcomes allowing enrollment at any point during the sixth year to identify persistence and (2) fifth year outcomes (results available upon request). Reestimating the MNL model with these alternative definitions of the

dependent variable does not substantially change our results. If anything they show that academic background/ability has a larger effect on graduation and persistence at the five than at the six year cutoff. This result may be due to the fact that less able students who persevere eventually graduate: persistence does pay off. We also tested for interaction effects between race/ethnicity and income/parental background. No significant effect was identified. From this we infer that the effects of low income or first generation status do not differ by race or ethnicity.

7. WILL PERSISTERS GRADUATE?

We have demonstrated thus far that persistence at the six-year mark is neither unusual nor randomly distributed. Yet to be addressed is whether these persisters are ultimately successful. We approach this question from two angles. First, we exploit the information available in the BPS regarding fifth and sixth year outcomes as well as some more detailed information regarding sixth year enrollment status. Second, we use data from the 2001 Baccalaureate and Beyond Longitudinal Study (B&B 2001) to examine time to a degree conditional upon completion to identify the fraction of non-graduates who may eventually graduate.

While the BPS does not follow students beyond their sixth year, we can look at those who were persisting at the end of their fifth year and see how they progressed in the sixth year. Of those who had not yet graduated but were still enrolled in the final term of their fifth year, 26% had graduated and 52% were still enrolled at the end of year six. If the progression from year five to year six is any indication of future trends, many of those classified as persisting in year six may well complete their baccalaureate degree within an additional year or two. Further analysis of enrollment status indicates that many of those still enrolled in Spring 2001 either are or have been enrolled part-time. In general, part-time enrollment is not unusual. Almost 22% of

the sample enrolls part-time for at least one term. However, one-third of those still enrolled in Spring 2001 are enrolled part-time in that term and over half have at some point been enrolled part-time. Clearly part-time enrollment drags out the college experience and could help explain a long time-to-degree.

To further address this concern, we use B&B 2001 to assess the probability with which those who have not graduated in six years will do so soon. These data report time-to-degree for those who graduated with a Bachelor's degree in the 1999-2000 academic year. To match our sample as closely as possible, we restrict the analysis to those who began college before the age of 23, who resided in the United States, and who first attended a four-year institution. We further restrict the analysis to those who complete within twelve years. (Note: Fully 5.2% of the B&B sample reports taking over twelve years to complete their degree!) Eighty-nine percent of this similar sample report taking six or fewer years to graduate. If our 63.2% graduation rate based on initial enrollment corresponds to this 89% based on who ultimately graduates, then 71% of our sample should graduate within twelve years. This calculation suggests that another 7.8% of our sample (or over half of those still enrolled in Spring 2001) will graduate within the next six years. The highest concentration (40%) of these delayed graduates graduate in year seven. It is highly likely that these individuals are actively enrolled during year six.

Analysis by socioeconomic status using the B&B further supports our findings. While 21% of Hispanics who graduate do so in more than six years, only 15% of African Americans do so. Seventeen percent of first-generation college students and 17% from the lowest income quartile take between six and twelve years to graduate, while only 5 to 7% of those with the most educated parents or from the highest income quartile do so. Dragging out the college experience is far more common for individuals who are from historically underrepresented populations.

Unfortunately it is not possible to control for academic background/ability in making these calculations using the B&B.

8. CONCLUSION & DISCUSSION

In this study we make two primary contributions to the literature on college outcomes. First, we examine the relation between commonly used socioeconomic factors and graduation using micro-level data. Previous economic analysis has focused on college attendance not outcomes or has used aggregate data. The rich data we use allow us to include a broad array of controls for individual academic preparation/ability. We replicate earlier results in the field showing that higher education differentials by race and ethnicity are primarily driven by differences in academic preparation/ability. However, differentials by parental education and household income remain significant and substantial. Second, and substantially different from prior studies, we distinguish between those non-graduates who are still enrolled six years following matriculation and those who are not still enrolled by using a multinomial logit specification. Standard logit analysis treats all non-graduates the same, and hence in some sense as 'failures.' Results from the more detailed and hence more useful multinomial logit specification, indicate that persistence and non-enrollment are statistically distinct outcomes.

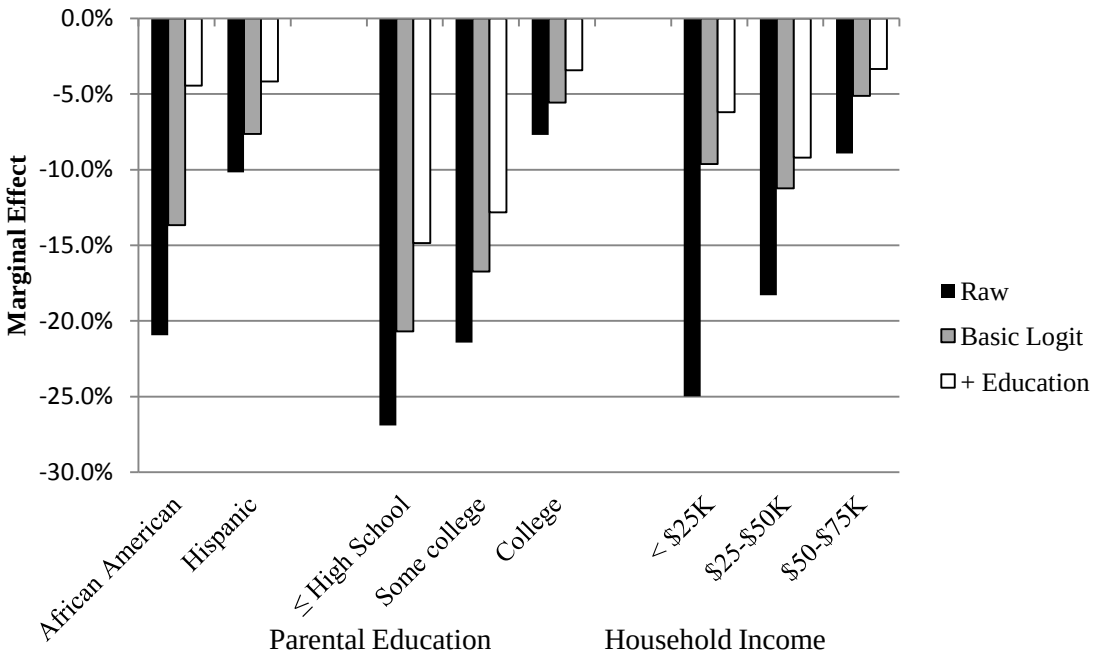
Overall 36% of those who had not graduated in six years were still enrolled when last observed: persistence at the six year point is substantial. Evidence from those persisting at the five year mark as well as from data on time-to-degree from a survey of graduates suggests a good fraction of those still enrolled after six years may in fact go on to graduate – perhaps half of them. Such persisters may just be taking longer to graduate. If students from more

disadvantaged backgrounds are disproportionately likely to persist, then significant differences in raw graduation rates by socioeconomic status may diminish and possibly disappear over time.

The more complex MNL analysis does in fact reveal significant differences in the marginal impact of socioeconomic status on the probability of persisting. Those of Hispanic descent are significantly more likely to persist than non-Hispanics, but are not significantly more likely to drop out. Conversely, first generation college students are significantly more likely to not be enrolled, but not significantly more likely to persist than non-first generation college students. African American students/ those from lower income households have higher probabilities of both persisting and not enrolling, as compared to their white/higher income counterparts. Studying the distinctions between persisters and dropouts may suggest alternative policies that can speed time-to-degree. In general, persistence is substantial, significant, and oftentimes ultimately rewarding and should not to be ignored in future analyses of college outcomes.

Furthermore a common thread throughout this discussion is the overwhelming importance of academic preparation from the pre-college years on college success. Colleges work with the raw material they receive. As policy makers evaluate actions to improve performance outcomes for K-12 education, they should also take into account the long term impact such improvements will have on college outcomes. Finally to the extent that there still remains a significant difference in college outcomes by socioeconomic class, even after controlling for academic background, policy changes at the college level to help those specific historically underrepresented populations may also be in order. If one believes in equal opportunity for all, then differences due to the luck of the draw at birth need to be eliminated.

Figure 1
Marginal Effect of Socioeconomic Status on 6-Year Graduation Rate



Note: African Americans are compared to whites; Hispanics are compared to non-Hispanics; the various levels of parental education are as compared to those having a parent with a post-graduate degree; the various income groups are as compared to those having a family income over $75,000.

Note: The marginal effect is calculated as that compared to a white, non-Hispanic, never married and childless, 18 year old male, with a high school degree, who lives in New England, in an area with a sample average unemployment rate, has a parent with a college degree and family income in the third quartile, expects to complete trigonometry, has a high B average in high school, and an SAT score of between 800 and 1100.

Table 1
Sample Means
(% except where noted)

Variables	Mean	Std. Dev.
Basic Specification		
Female	0.550	0.498
White	0.776	0.417
African American	0.109	0.311
Other race	0.115	0.320
Hispanic	0.083	0.276
Parental Education		
High school	0.305	0.012
Some college	0.124	0.329
College	0.251	0.434
Post-graduate	0.264	0.441
Missing	0.055	0.229
Family Income		
Independent	0.028	0.166
Income ($000s)	60.648	54.651
< $25,000	0.224	0.417
$25-$50,000	0.262	0.440
$50-$75,000	0.245	0.430
>= $75,000	0.269	0.443
Age - 17	1.412	0.756
Ever married male	0.004	0.063
Ever married female	0.007	0.083
Father	0.004	0.061
Mother	0.010	0.101
Unemployment rate in state of residence	5.494	1.194
Measures of Academic Preparation/Ability		
No high school diploma	0.011	0.103
Highest level of math:		
Algebra II or less	0.229	0.420
Trigonometry	0.163	0.370
Pre-calculus	0.230	0.421
Calculus	0.259	0.438
Missing	0.119	0.324
Standardized Test Information		
SAT score of 800-	0.186	0.389

SAT score of 800-1000	0.468	0.499
SAT score of 1100+	0.317	0.465
Took ACT test	0.306	0.461
Missing test score	0.029	0.169
High school GPA		
B- or lower	0.088	0.283
B- to B	0.142	0.349
B to A-	0.270	0.444
A- or higher	0.384	0.486
Missing	0.117	0.322
Number of Observations	~5820	

Eight regional dummies are also incorporated in each specification.

Table 2
Multinomial Logit Results
Marginal Effects

	Graduated	Still Enrolled	Not Enrolled
Base Probability	73.41%	10.76%	15.83%
African American	-4.44%	1.68%	2.76%
	(0.131)	(0.394)	(0.193)
Hispanic	-4.12%	3.91%	0.21%
	(0.189)	(0.040)	(0.927)
Parental Education (vs. College)			
<= High School	-11.56%	2.14%	9.42%
	(0.000)	(0.119)	(0.000)
Some College	-9.66%	-0.16%	9.82%
	(0.003)	(0.917)	(0.001)
Post Graduate	3.52%	-0.68%	-2.83%
	(0.135)	(0.586)	(0.109)
Household Income (vs. 3rd Quartile)			
< $25,000	-2.41%	1.57%	0.84%
	(0.275)	(0.275)	(0.633)
$25-50,000	-5.18%	0.56%	4.62%
	(0.015)	(0.684)	(0.006)
$75,000+	3.39%	-1.46%	-1.94%
	(0.173)	(0.220)	(0.365)
Female	5.09%	-1.94%	-3.15%
	(0.001)	(0.020)	(0.007)
Age -17	-2.84%	0.69%	2.14%
	(0.013)	(0.255)	(0.020)
Highest Math Expected			

Algebra II or less	-5.54%	2.33%	3.22%
	(0.012)	(0.110)	(0.088)
Pre-Calculus	3.01%	0.02%	-3.03%
	(0.198)	(0.991)	(0.086)
Calculus	3.82%	-0.23%	-3.59%
	(0.152)	(0.896)	(0.036)
High School GPA (vs high B)			
B- or lower	-20.91%	5.68%	15.22%
	(0.000)	(0.057)	(0.000)
B- to B	-11.97%	3.86%	8.11%
	(0.000)	(0.052)	(0.001)
A	10.42%	-2.52%	-7.90%
	(0.000)	(0.033)	(0.000)
SAT Equivalent Test Score			
≥ 1100	2.62%	-2.98%	0.35%
	(0.255)	(0.016)	(0.860)
≤ 800	-7.22%	2.48%	4.73%
	(0.006)	(0.167)	(0.019)

P-values in parentheses.

The base probability is for a single, childless, 18 year old white, non-Hispanic male with a college educated parent and a household income of $50-$75,000, who lives in New England, in a state with a sample average unemployment rate, has a high school diploma, expects to complete trigonometry, has a high B average in high school, and has an SAT score of between 800 and 1100.

Also included in the model is the unemployment rate in the state of residence and dummy variables for 8 regions, those of 'other race', those missing parental education or academic ability/background information, those who are no longer dependents, those without a high school degree, those who took the ACT, and a constant term.

References Cited

Adelman, Clifford. 2006. *The toolbox revisited: Paths to degree completion from high school through college.* Washington, D.C.: U.S. Department of Education.

Altonji, Joseph G. 1993. "The demand for and return to education when education outcomes are uncertain." *Journal of Labor Economics*, 11 (1): 48-83.

Bailey, Martha J. and Susan M. Dynarski. 2011. "Gains and Gaps: Changing Inequality in U.S. College Entry and Completion." NBER Working Paper No. 17633.

Becker, Gary. 1964. *Human capital*. New York: National Bureau of Economic Research.

Belley, Phillipe and Lance Lochner. 2007. "The Changing Role of Family Income and Ability in Determining Educational Achievement." NBER Working Paper No. 13527.

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(3): 129-57.

Cameron, Stephen V. and James J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy*, 109(3): 455-99.

The College Board. 1999. "Concordance Between SAT I and ACT Scores for Individual Students." Research Notes. Office of Research and Development. RN-07.

Lee, Michael Jr. and Anita Rawls. 2010. "The College Completion Agenda: 2010 Progress Report." The College Board Advocacy and Policy Center. Accessed at http://completionagenda.collegeboard.org/sites/default/files/reports_pdf/Progress_Report_2010.pdf on August 10, 2011.

Cragg, Kristina M. 2009. "Influencing the probability for graduation at four-year institutions: A multi-model analysis." *Research in Higher Education*, 50(4): 394-413.

DesJardin, Stephen L., Dennis A. Ahlburg, and Brian P. McCall. 1999. "An Event History Model of Student Departure." *Economics of Education Review*, 18(3): 375-90.

Digest of Education Statistics, 2010. 2010. Accessed at http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2011015 on January 2, 2012.

Kuh, George D., Jillian Kinzie, Jennifer A. Buckley, Brian K. Bridges, and John C. Hayek. 2006. *What matters to student success: A review of the literature*. Final report for the National Postsecondary Education Cooperative and National Center for Education Statistics. Bloomington, IN: Indiana University Center for Postsecondary Research. http://nces.ed.gov/npec/papers.asp.

Manski, Charles F. 1989. "Schooling as Experimentation: A Reappraisal of the Postsecondary Dropout Phenomenon." *Economics of Education Review*, 8(4): 305-12.

Stratton, Leslie S., Dennis M. O'Toole, and James N. Wetzel. 2008. "A multinomial logit model of college attrition that distinguishes between stopout and dropout behavior." *Economics of Education Review*, 27(3): 319-31.

Turner, Sarah. 2004. "Going to College and Finishing College: Explaining Different Educational Outcomes." In C. Hoxby, ed., *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It.* Chicago, University of Chicago Press.

ENDNOTES

[1] Initial enrollment probabilities are also lower for these population groups, though they have improved considerably.

[2] Cameron and Heckman (2001) find similar results looking at college enrollment. Controlling for long-term family income, academic ability, and family background (including parental education) explains the substantial raw differences by race and ethnicity.

[3] Belley and Lochner (2007) find similar results looking at college enrollment.

[4] The NCES requires sample size be rounded to the nearest ten.

[5] A handful of individuals are excluded due to missing age or other characteristics of interest.

[6] This figure is reported in the Digest of Education Statistics (2010, Table 341) and reflects the six-year graduation rate for all degree seeking students beginning as full-time fall semester students at a four-year institution in 1996.

[7] Measuring persistence using enrolment status in the Spring 2001 term alone may appear restrictive, but the fraction who have not graduated who are persisting rises only to 40% when allowing enrollment at any time during the 2000-2001 academic year.

[8] In the multivariate analysis, the base case is that of a white, non-Hispanic, single, childless, 17 year old male with a college educated parent and household income of $50-$75,000, who is from New England and a residence with a sample average unemployment rate. Academic preparation and ability are assumed to be approximately modal with the highest expected level of math being trigonometry, high school GPA being between a B and an A-, and SAT test scores falling between 800 and 1100, all for respondents with a high school degree.