Who Drops Out from College? A Study of Social Origin at a Midwestern Teaching University

By

Joelle Spotswood
M.A., Eastern New Mexico University, 2008
B.A., West Virginia University, 2000

Submitted to the graduate degree program in Sociology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Arts.

Chair: Dr. ChangHwan Kim

Dr. Mehrangiz Najafizadeh

Dr. Argun Saaticoglu

Date Defended: 27 March 2019

The thesis committee for Joelle Spotswood certifies that this is the approved version of the following thesis:

Who Drops Out from College? A Study of Social Origin at a Midwestern Teaching University

_____

Chair: Dr. ChangHwan Kim

Date Approved: 27 March 2019

Abstract

Previous research has examined the degree to which social origins affect college completion, but few have studied the association of family background, social class, and neighborhood contexts with regard to the rate of four-year college attrition. To fill this gap, this study utilizes rich administrative data on first-time (students who have not completed any post-secondary courses), full-time freshman cohorts (2007-2014) from a four-year Midwestern teaching university which provided information on students' demographic information, including parental education and income, academic performance, and family background via admission and Federal Application for Federal Student Aid (FASFA) applications. I supplement these with secondary data on students' county and high school socioeconomic characteristics. Linear probability and hazard models are estimated. Primary amongst the findings is that parental education is the significant predictor in dropping out of college rather than parental income, even when controlling for academic preparation and a variety of other family and neighborhood variables. Being a first-generation student, someone who does not have a parent with a college degree, significantly and substantially increases the likelihood of dropping out of college, as does being male. To a lesser though still significant extent, county unemployment also predicts retention or withdrawal, suggesting the importance of neighborhood effects. The results imply the value of cultural rather than economic capital transmission in students' college success and lend further evidence for the widening class inequality gap regarding college completion. This study is especially significant for educational sociologists and higher education retention programs, providing empirical data from which to draw to create targeted intervention for potentially at-risk freshman.

Keywords: Cultural Capital; Higher Education; College Attrition; Parental Educational Transmission; First-Generation Students; Unemployment; Academic Readiness

Table of Contents

Introduction

Little dispute remains regarding the benefits of obtaining a college degree. College graduates are employed more and earn more in the labor market (Kim & Sakamoto, 2008), are more upwardly mobile (Bailey & Dynarski, 2011), have better health outcomes (Ross & Mirowsky, 1999), and report greater degrees of happiness (Hout, 2012; Ross & Mirowsky, 1999). Graduates are more likely to be civically engaged, to vote, and to support constitutional freedoms and various non-traditional religious views (Hout, 2012; Kingston, Hubbard, Lapp, Schroeder, & Wilson, 2003). Completing college, therefore, seems to provide possible holistic improvement to the quality of graduates' lives relative to those who do not attain a college degree.

Despite all the benefits associated with attending and completing college, a large and increasing gap exists between those who go to college and those who do not. Students from less privileged economic backgrounds, who are usually, but not always, first-generation students (i.e., students whose parents did not obtain a college degree), are less likely to attend and complete college than students from families with higher incomes, who are often continuing-generations (i.e., students whose parents do hold a college degree) (Bailey & Dynarski, 2011; Zweig, 2012). Less than a third of children who are born into families in the lowest economic quintile are likely to go to college, while more than ¾ of those who are born into the top quintile are likely to attend (Goldrick-Rab, Kelchen, Harris & Benson, 2016).

Though much previous research focuses on the financial aspects of college attendance and graduation, families sometimes provide far more than simply financial resources; parents also transmit human, social, and cultural capital to their children, including motivation and expectations, comfort with authority, and the ability to seek and utilize resources for success

(Coleman, 1988; Bourdieu, 1986; Brand & Xie, 2010; Lareau, 2011; Morgan, Leenman, Todd, & Weeden, 2013; Hamilton, Roksa, & Nielsen, 2018). First-generation students often come from working-class backgrounds and may not receive the kind of parental support—specifically regarding achievement motivation—that their middle class peers receive (Kiker & Condon, 1981; Lareau, 2011). This dearth is especially significant for, as in alignment with the Wisconsin model, the influence of significant people, like parents, "is the primary mechanism which transmits the effect of family background into the college expectations that then determine college entry" (Morgan et al., 2013, p. 198). Further, first-generation, especially working-class students, are less likely to have parents who emphasize the benefits of college or who, lacking the personal experience, are not as able to assist their children in enrollment or managing the various barriers that can arise during college (Stinebrickner & Stinebrickner 2003). Thus continuing-generation students often have and are able to access support systems that first-generation students do not have and are not able to access, so first-generation students are much less likely to complete a degree (Strayhorn, 2014).

My approach employs contexts beyond immediate social class to include a robust data analysis of internal (family) and external (i.e., neighborhood socioeconomic status [SES]) characteristics to determine what conditions adversely affect completing college and, thus, further increase educational inequality between students by socioeconomic group. Further, neighborhoods, which are often segregated by income (Reardon & Buschoff, 2011), can serve to concentrate advantage or disadvantage to certain areas. One of these ways is through public schooling. Loosely, public schools receive funding based on a set of economic matrices derived from the surrounding community; thus, poor neighborhoods often have poor schools with fewer resources, narrower curriculum, and less experienced teachers while richer neighborhoods have

schools with a greater pool of economic resources, a broader curriculum, and more experienced teachers. Economic differences matter significantly because students coming from more affluent neighborhoods are more likely to graduate college (Owens, 2010), though these characteristics operate differently for African American and white students (Vartanian & Gleason, 1999). Moreover, employment rates in the county affect postsecondary pursuits: lower county unemployment levels lead to decreased college enrollment while higher unemployment rates drive people to higher education (Hillman & Orians, 2013), especially during times of recession (Card & Lemieux, 2001; Barr & Turner, 2013; Schmidt, 2018). I argue that, especially for students who are only marginally prepared academically, students who lived in neighborhoods with higher unemployment rates are more likely to persist in college, understanding the shortage of economic opportunities in their home areas. Conversely, these marginal students are more likely to drop out of college when their local unemployment rates improve.

My current research highlights how family background, neighborhood and high school characteristics, and academic preparation contribute to the contexts in which students drop out of college and the timing in which attrition occurs. While a number of studies demonstrate the importance of family background on college completion, very few examine multiple categories of variables concurrently, especially in terms of student attrition at the microlevel employed here. To help rectify this dearth, this study uses original admissions and *Free Application for Federal Student Aid (FASFA)* data from 8 first-time (students who have not completed any post-secondary courses), full-time freshman cohorts at a low-selectivity Midwestern four-year teaching university with an approximate 83% acceptance rate (data for the 2017 cohort). From the university data, I use linear probability modeling (LPM) to examine how objective family and neighborhood socioeconomic statuses, culture, and college-readiness affect a student's

college persistence and attrition. I use secondary high school and county-level government data from the U. S. Census Small Area Income and Poverty Estimates (SAIPE), the 2010 U.S. Census, the American Community Survey educational estimates, the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics (LAUS), and the National Center for Education Statistics (NCES) to supplement and analyze high school and neighborhood characteristics and evaluate these contexts for correlations to student withdrawal. Additionally, I use time-to-event and hazard modeling to better understand who drops out and under what conditions students are most likely to drop out. Given the limited availability of this type of data, this study contributes new empirical evidence to understanding the patterns, trends, and timing of college attrition.

Literature Review

Family Transmissions: Capital, Social Origins, and Postsecondary Education

Undoubtedly social class affects higher education: it predicts who is likely to attend college, where she will attend, the type of major she will select, and her chances of persisting. Questions remain, however, regarding how long and to what degree family background matters for educational attainment. Some research suggests that a student's social origin effects diminish dramatically with college attendance, and the longer a student remains in school, the less impact family background has on her continued enrollment (Torche, 2011; Mare, 1980; Hout, 1984; Hout, 1988). Conversely, other research shows that social background, especially relating to parental education, matters substantially to a student's chances of obtaining a college degree (Hansen & Mastekaasa, 2006; Coleman, 1988; Hout & Janus, 2011). The question of lasting family background effects on college attrition is central to this paper.

Family background with its differing types of capital affects the likelihood of a student attending college. Financial, human, cultural, and social capital all inform an individual's family

background (Coleman, 1988; Bourdieu, 1986), and as access to various forms of capital often reflects class, the type of cultural capital transmitted by parents is highly dependent on the family's socioeconomic standing (Bourdieu, 1986). Family social class impacts the chances a student has of going to college, where she is likely to go, what kind of selectivity the college has, the major she will select, and the chances she has of persisting to graduation. Further, as social class is highly correlated with parental education levels, continuing-generation students often receive different kinds of capital transmissions from their parents than first-generation students. Cultural capital transmission confers certain specific skills—both academic and non-academic—and benefits to continuing-generation students that first-generation students may not receive. Parents with a BA have personal, experiential narratives to share with their children—they already "know the ropes" regarding applications and deadlines, how to access help if questions or problems arise, and have knowledge negotiating the various processes associated with enrollment and selecting majors. Moreover, children of college-educated parents are more comfortable in academic settings and are more able to advocate for themselves (Lareau, 2011). When parents are able to transmit self-efficacy, making their children independent active agents, especially in educational institutions, those children are more able to operationalize the elements of their family background, like social or economic capital, to become successful in academics (Hamilton, Roksa, & Nielsen, 2018). In fact, students who have the greatest access to cultural capital perform better than their peers (Hansen & Mastekaasa, 2006). Thus, the capital conveyed to children by their parents is fundamental in understanding academic success (Brand & Xie, 2010).

Conversely, first-generation students are at a distinct disadvantage relative to their continuing-generation. They do not come to college with the same academic cultural capital and

often do not have an adequate understanding of what to expect and what they must do to be successful (Tinto, 2012). Thus, they do not perform as well academically as their continuing-generation peers (Strayhorn, 2014; Lareau, 2011). Subsequently, first-generation students are less likely to graduate from college, especially as the level of education a student's parents have affects their own educational attainment (Blau & Duncan, 1967). Students who are first-generation enroll less often in higher education and have lower persistence rates than students who have at least one parent with a four-year degree (Cataldi, Bennett, & Chen, 2018). Unfortunately, the students who are likely to benefit most from college are those students, such as first-generation, minority, and low-income students, who are least likely to attend (Brand & Xie, 2010).

While parental education as well as family income are components of family background (Coleman, 1988), these characteristics operationalize differently regarding attrition. Parental education often relates to cultural capital transmission: providing skills, for example, students need to navigate the college enrollment process. However, parental income, especially in terms of financial capital, affects the likelihood of college attendance as well as the kind of college students are likely to attend. Family income significantly affects student enrollment in higher education, especially in four-year institutions. Students from affluent families are more likely to attend college relative to students from other economic backgrounds; thus, the gap in college enrollment between high- and low- income students is widening (Kane, 2004; Alon, 2014; Zweig, 2011). Wealthier students are more likely to enroll in high-status, financially lucrative majors at more selective universities (Hout, 2012), and students who attend these universities are more likely to graduate (NCES, 2018; Hout, 2012). Interestingly, though, students who have access to the most economic capital tend to not perform as well as those students with access to

greater cultural capital, highlighting the importance of differentiating between how cultural and economic capital function in college performance. (Hansen & Mastekaasa, 2006). The combinations of privilege—financial and cultural, for example—predispose certain students who come from more advantaged backgrounds to academic and life-long financial success while increasing the inequality relative to first-generation, working-class students.

In addition, despite opportunities for financial aid packages and other monetary assistance for students from low-income families, pecuniary support may not be enticement enough to enroll in college, especially because so many students start, but do not finish, college (Bailey & Dynarski, 2011). African American students are most likely to take out loan packages but are the least likely racial group to graduate, burdening students with debt but without a degree (Merolla, 2018). Additionally, some research demonstrates a negative relationship between financial aid packages and college persistence, suggesting that more than financial support affects retention (St. John, Andrieu, Oescher, & Starkey, 1994). Students who come from less privileged socioeconomic backgrounds may not receive, or may not be able to use, the various kinds of capital transmissions they need to be successful in college, even if aid is offered them (Coark, 2013). For instance, low-income parents and parents who themselves do not have college educations may be less able to provide various navigational support for their students, including vital components like completing the Free Application for Federal Student Aid (FASFA). Social class and various capital transmissions function in multiple ways to help students get in, or works to keep them out, of college.

Neighborhood Contexts and Educational Outcomes

Neighborhood contexts affect a variety of outcomes for citizens, including educational outcomes across the academic lifespan of a student. The density in a neighborhood of those who

have college experience is important to the chances that a student in that same neighborhood will attend college: living in or moving to a neighborhood in which more people have attended college increases a child's changes of attending college herself (Chetty & Hendren, 2018; Owens, 2011, Coleman, 1988). Relative economic advantage or disadvantage of a neighborhood also contributes to educational outcomes. A student living in areas of high poverty or unemployment, for example, may not receive social encouragement for school attendance and academic performance and may not have academic role models (Hicks, Handcock, Sastry, & Pebly, 2018). Conversely, students who live in economically stable neighborhoods with low levels of poverty may perceive agreement among the residents about the relative importance of college graduation (Berg, Stewart, Stewart, & Simons, 2013). Following neighborhood deprivation theories, poor neighborhoods have fewer resources from which its residents can draw, including quality schools (Hicks et al., 2018, Reardon & Bischoff, 2011). School funding is often tied to property values in a given neighborhood, so neighborhood income affects school quality, including finances for additional resources, salaries to lure better educated and more experienced teachers, and expanded curriculum. These school characteristics, combined with the greater social and familial contexts, contribute to students' educational attainment. Collectively, the educational attainment of the neighborhood residents also contribute to the individual's as well as the collective student body's academic success (Vartanian & Gleason, 1999). Instead of a single variable, multiple neighborhood contexts converge to influence the ways that students perform academically.

Specifically relating to higher education, a student's neighborhood affects the probability of aspiring to, enrolling in, and graduating from college; thus, understanding neighborhood effects is vital to deciphering patterns of college attrition. Student educational aspirations are

associated with social class, and residents who live in disadvantaged neighborhoods tend to share similar patterns and values to one another (Chetty & Hendren, 2018). As such, students who come from disadvantaged neighborhoods tend to have lower educational aspirations (Smith-Maddox, 1999; Flowers, Milner, & Moore, 2003; Stewart, Stewart, & Simmons, 2007; Wilson, 1996). Additionally, neighborhoods affect the chances a student has of completing college or the chances that she will not complete high school, and these contexts operationalize differently by race (Quillian, 2014; Vartanian & Gleason, 1999). Additionally, neighborhood racial segregation predicts a student's academic trajectory and affect African American and white students differently. Segregation by race (African American, white) in a neighborhood, for instance, decreases the chances that an African American student will graduate high school or college, but it has no predictive value for a white student's graduation (Quillian, 2014). Other neighborhood characteristics, though, function differently as predictors for high school graduation and college graduation. Notably, low socioeconomic status of residents, family construction/disruption, and the prestige of residence can influence and reduce high school dropout rates for African American students while having minimal impact on their chances of graduating from college; this is especially true for students from low-income households where the head of household did not complete high school (Vartanian & Gleason, 1999). Low-income students who attended high schools with students from mixed SES families, however, do not necessarily perform better than they would in homogenous SES schools; these students actually fare worse when in schools that are more white and more affluent, while students who themselves are from higher income families fare better academically when with white, more affluent peers (Owens, 2011). High school performance matters because high school GPA tends to be a strong predictor of college performance.

College enrollment is often counter-cyclical with the economy and labor markets. While economic downturns negatively affect the national economy, higher education often derives enrollment benefits from recessions, especially ones categorized by significant increases in unemployment or underemployment; during times in which unemployment rises, college enrollment increases (Card & Lemieux, 2001; Barr & Turner, 2013; Schmidt, 2018). People who were formerly employed may return to school to be retrained, cross-trained, or newly trained, especially during periods of recession (Hillman & Orians, 2013). Especially significant for community college enrollment, marginal students who may have been undecided regarding postsecondary education may decide the financial burden often imposed by a college education is worth the investment as the options for immediate entry to the labor force are weak (Hillman & Orians, 2013). In short, a bad economy is good for education.

Most recently, the Great Recession (2007-2009), a period marked by a volatile market and mortgage-lending collapse, led to a dramatic increase in unemployment, significant reduction in home value, and decreased spending (Hurd & Rohwedder, 2010). As consistent with prior research regarding recessions, during this period college enrollment increased, especially by minority students (Barr & Turner, 2013; Long, 2014) and by older students (Barr & Turner, 2013). However, while research shows increased enrollment across all age groups during the recession, the enrollment patterns of different age groups differed after the recession, with those under the age of 25 being increasingly likely to enroll in college (Schmidt, 2018). One could speculate that this cohort experienced more encouragement, as a cautionary tale, from their families to enroll in college to guard against another financial and housing sector and employment collapse.

Enrollment trends during periods of increased unemployment vary by type of institution, with community college and open-enrollment institutions seeing a greater increase in enrollment than other types of institutions (Hillman & Orians, 2013; Barr & Turner, 2013). When the economy stabilized after 2010, college enrollment declined (Schmidt, 2018; NCES, 2018b). However, institutions experienced different degrees of declining enrollment post-recession. After the recession, community college enrollment returned to similar prerecession numbers while 4-year college enrollments remained higher than they were prerecession, though still lower than peak recession enrollment (Schmidt, 2018). Thus it can be understood how shifts in the national and local economy affect college enrollment: as supply and demand in the labor force changes, so does supply and demand for postsecondary education.

Academic Preparation for College

Successful college students were generally successful high school students, and often college success is predicated by high school grade point average (GPA) and standardized test scores. However, current research questions whether these traditional academic predictors are accurate or broad enough to represent a student's chances of graduating college. Opponents argue that schools themselves are too varied in their expectations, rigor, and curriculum, so that high school GPA may not be the best predictor of academic success (Westrick, et al, 2015). Further, GPA and test scores fail to take into account the "non-academic" skills necessary for college success, like study habits, dispositions, and executive functioning abilities (Sommerfeld, 2011). However, traditional measures, like test scores and GPA, are more significantly predictive than other measures (Hepworth, Littlepage, & Hancock, 2018). Thus, often high school GPA is still used as a primary predictor for potential college success, as are Scholastic Aptitude Test (SAT) and ACT scores (Moore, et al, 2010; Sawyer 2013; Hepworth, Littlepage, & Hancock,

2018). Yet, GPA and ACT scores are not equally predicative; the former, GPA, rather than the latter, standardized test scores, prove better predictors for college GPA, and then by proxy, college graduation (Saunders-Scott, Braley, & Stennes-Spidahl, 2018). GPA tends to reflect the kind of non-academic skills that make students successful in college better than the ACT could, (Saunders-Scott, Braley, & Stennes-Spidahl, 2018). Even for students who must take remedial courses, high school GPA rather than the need for remediation, is a stronger predictor for college graduation (Adelman, 2004). These measures, along with including SES information about the student's high school, can also be used to predict a student's GPA in the first year of college, a prescient indicator of possible persistence or attrition (Zwick & Himelfarb, 2011).

Despite state and federal programs aimed at increasing college-readiness for high school students, many students lack preparedness for post-secondary education (Strayhorn, 2014). Unfortunately, inequality of academic preparedness often follows class and racial lines. First-generation students tend to lack the skills to be successful in college that their continuing generation peers normatively possess (Strayhorn, 2014), and children whose parents have not attended college underperform relative to children of college-educated parents (Lareau, 2011). Working-class and minority students often attend high schools with fewer resources than do white, middle-class students. Students from disadvantaged schools tend to have lower academic preparedness than students from schools with greater resources like more experienced teachers, expanded and diverse curriculums, more and better quality physical spaces, advanced technology and equipment, and more involved parents and communities. Further, minority students, like African Americans and Hispanics, are often overrepresented in economically poorer schools (Zwick & Himelfarb, 2011), and as students who come from lower SES schools tend to be less academically prepared for college, minority students are often disproportionally underprepared

for the rigors of higher education. Even more, when students *are* similarly prepared academically and grades are relatively equal, race still operationalizes as a disequalizer: having good grades benefits African American students less than the same grades would benefit whites (Davis & Otto, 2016).

Further, high schools serve as partial gatekeepers to higher education, and, whether intentionally or not, often serve to keep working-class and minority students out of college. Some students do receive a "college is for everyone" message from their schools, promoting the assumption that of course college is accessible and for "you;" however, others, often poor and minority students, receive the opposite message—college is for some, but not for "you." In these ways high schools function to both gate-keep and reproduce educational inequality (Martinez & Deil-Amen, 2015). Gatekeeping in high schools functions in a variety of ways, and while outside of the scope of this paper, it should be noted that "tracking," or placing students in certain curriculum (vo-tech versus college-bound, for example) often shapes students' educational prospects (see Kozol and Rose, for example). Students who do go to college from high school tracks outside of traditional academic pathways are often not college ready.

Academic under-preparedness adds additional temporal and financial burdens disproportionally to disadvantaged students. These students—students from low SES households, students from urban areas, and African American and Hispanic students—are more likely to enroll in remedial courses relative to their white, middle-class, suburban, and rural peers (Attewell, Levin, Domina, & Levey, 2006; Adelman, 2004). Remedial courses, which often cost the same as non-remedial courses, do not usually count towards a student's accumulation of credit hours for her degree plan, but they do count towards the amount of hours allotted for financial aid, thus adding to the financial burden of college. Further, African American students

are nearly twice as likely as white students to place into remedial courses in college (Attewell et al., 2006). This matters especially in 4-year institutions because there is a negative relationship between taking remedial/developmental courses and graduating (Attewell et al., 2006; Adelman, 2004). Additionally, taking remedial courses extends the student's time to graduation, time which further removes her from the labor force, and students who enroll in remedial/developmental courses are less likely to matriculate at all (Adelman, 2004). Thus, being academically underprepared before college increases the amount of money and time devoted to college with a decreased chance of matriculating and, thus, garnering less of a return on education.

Race and Gender: Who Stays and Who Drops Out

College enrollment has been growing rapidly since the 1950s, with significant educational expansion especially in the last 40 years (NCES, 1993; Fry, 2009), and increased enrollment reflects improved access to higher education (Bailey & Dynarski, 2011); however, access and enrollment to college do not necessarily guarantee a degree, and certain racial groups are less likely to matriculate than others. Due, at least in part, to the continual increase in economic disparities, neighborhoods are becoming more segregated and isolated along racial and socioeconomic class lines (Reardon & Bischoff, 2011), and this is especially troubling because of the disproportionate presence of African Americans in poor neighborhoods relative to whites (Johnson, 2010; Strayhorn, 2014). Educational attainment is often stratified by the socioeconomic status of high schools, which are dependent on the socioeconomic status of the neighborhoods in which they reside. Students who come from poorer neighborhoods tend to be less prepared for the rigors of academic life, and this kind of educational segregation leads to increased gaps in educational attainment by neighborhood, both by race and class (Burdick-Will,

et.al, 2011), with Latino students being especially underprepared for post-secondary education (Contreras & Contreras, 2015).

College enrollment is growing across all racial groups, but the rate of growth differs. White students enroll and graduate at the greatest rate across races. African American students enroll in college more today than 20 years ago, but they are still enrolling and graduating less often than their white peers (Haveman, Sandefur, Wolfe, & Voyer, 2004; Kane, 2004; Merolla, 2018). Hispanics fare worse than African Americans in terms of college degree completion: according to Haveman, Sandefur, Wolfe, & Voyer (2004), by 2000 35% of all whites held a Bachelor's degree, while only 20% of African Americans did, and only 18% of Hispanics did. This is not surprising: African American and Hispanic students are less prepared for college, as indicated by a variety of academic indicators, including standardized reading and math scores, self-perceived writing abilities, and senior math, relative to their white and Asian peers (Strayhorn, 2014). Even more disturbing is emerging evidence that college educated African American and Hispanic parents are not equally able to transmit the cultural benefits of their college education to their children at the same rate as their white peers, thus increasing both the minority dropout rate and intergenerational downward mobility (Alon, Domina, & Tienda, 2010).

Historically, men attended and graduated from college more than women. In the last 45 years, however, that trend has shifted. Men are now entering college less often, persisting for shorter amounts of time, and graduating at a slower rate than their female counterparts (Buchman & DiPrete, 2006; Bailey & Dynarski, 2011; Goldin, Katz, & Kuziemko, 2016). In general, men are less college-ready than their female peers, having lower grades and test scores (Goldin, Katz, Kuziemko, 2016; Strayhorn, 2014; Davis & Otto, 2016). Part of the gender shift is explained by

changing labor markets as well as how women better understand the importance of education for their future participation in and garnering benefits from the work force (Goldin, Katz, Kuziemko, 2016; Kim & Sakamoto, 2017; Buchman & DiPrete, 2006).

This move toward female college advantage is not equally applicable to all racial groups. Bailey & Dynarski (2011) find that African American men born in the last 100 years have been consistently less likely to complete college than their female counterparts, especially since the 1970s. The authors note this data reflects the incredibly low rate at which African American men complete college. In fact, the greatest college enrollment gender-gap is between African American men and women (Davis & Otto, 2016). Hispanic men are also graduating more slowly than their female counterparts, and Hispanic degree completion differs as well by Hispanic country of origin. Mexicans are significantly less likely to complete a college degree than are white students (Garcia & Bayer, 2005). Interestingly, however, Garcia & Bayer (2005) discuss the ways in which the attitude of *machismo*, or the perceived primacy of males, may culturally select men to seek more education, if the family choice is to send either a son or a daughter.

Research Questions

Drawing on previous research regarding college attainment based on family background and neighborhood and high school socioeconomic characteristics and using data from admissions and Free Application for Federal Student Aid (FASFA) applications, I pose the following questions: Under what family and neighborhood economic conditions are first-time, full-time students most likely to drop out? Does the amount of time these students persist vary with differing economic circumstances? Does attrition vary based on gender and race? Open-enrollment and low-selectivity universities tend to have lower matriculation rates than highly-selective institutions, so understanding the conditions under which students are more likely to

drop out is important for persistence and retention initiatives at less selective universities. Knowing who is at increased risk of not completing college and what are those risks, like being first-generation, may help students in potential cost-benefit analysis, especially regarding incurring financial burden.

This original data set can provide answers to these questions through the depth and the temporal relevancy of the information gathered. As such and in response to the questions posed above, I contend the following: 1) That family socioeconomic status (SES)—that both the income of the parents and of the students separately—relate significantly to student persistence and time to dropout, namely that students who are more affluent, as determined by parental income, will persist longer in college than their working-class peers; 2) That family cultural background, analyzed using parental education, relates significantly to the time students persist in college, and students who have at least one parent with a college education operationalizes that advantage to persist longer than students who do not have at least one parent with a college degree; 3) That college persistence rates vary significantly between men and women—specifically that women persist longer than do men; and 4) That neighborhood characteristics, including high school poverty levels and county unemployment rates, predict college attrition, specifically suggesting that as poverty and unemployment rates increase, college attrition rates decrease.

Analytic Strategy

Data Set

This study utilizes a large, unique administrative application and FASFA data set for first-time, full-time freshman in the 2007-2014 cohorts at Sunflower State University (SSU), a Midwestern teaching university with a Carnegie Classification of Master's College and

Universities: Larger Programs. Founded in 1863 as a teacher's college, SSU has become an outstanding institution for a variety of disciplines, including Library and Information Management, business, and science, and has strong scholar programs, including an Honors College and a TRIO McNair Scholars program. The undergraduate student population is predominantly white (approximately 70%), in-state (86%), female (62%), and nearly 70% of the students come from rural or town areas. Approximately 4% of SSU's student body is international. SSU boasts small undergraduate class sizes, averaging 18.5 students per class as well as a low student-to-teacher ratio (18:1) and hosts programs at six off-campus locations. The university is accredited through the Higher Learning Commission and has 83% acceptance rate for approximately 3,600 undergraduates. SSU confers a number of Bachelor's and Master's degrees, one Doctoral degree, and a number of post-Baccalaureate and post-Master's certificates. More than 40% of incoming students bring transfer credits and are, therefore, not accounted as first-time, full-time freshman. SSU employs 247 full-time and 31 part-time faculty members and 119 graduate assistants on campus and enrolls more than 2,100 graduate students in approximately 38 programs.

For this study, SSU provided demographic and socioeconomic status information for students, including parental income, education, and marital status; household size and number of students in college; student income, marital, and dependency status; Pell eligibility, estimated family contributions (EFCs), and high school measures, including performance indicators. Data on courses attempted, courses passed, and college GPA for cohort members were also provided[1].

I supplement institutional information with data drawn from a variety of government sources to create neighborhood and high school variable groups. Using data gathered from the U.

[1] Not all information provided by SSU was used in this project.

S. Census Bureau Small Area Income and Poverty Estimates (SAIPE) and the Bureau of Labor Statistics Local Area Unemployment Statistics (LAUS), I construct 4-year county level averages for unemployment, median income, and overall poverty rates. Drawing on Demographic Profiles and American Community Survey estimate information from the U. S. Census Bureau, I complement the neighborhood data with county population and education information. In addition, using basic high school information from the university admissions application data then drawing data from the National Center for Education Statistics (NCES), I create a robust high school variable set, including poverty rates based on Free and Reduced Lunch Program (FRLP) eligibility, school demographics, and other institutional information. Finally, students who have withdrawn or been administratively dropped from enrollment then subsequently reenrolled are not included in this data.

Main Dependent Variables

Through access to admissions records at Sunflower State University, I obtained a rich variety of data for first-time, full-time freshman for the 2007-2014 cohorts, including demographic and family information. The main dependent variable is if a student dropped out of college, which is coded as a binary (1= dropped, 0= did not drop). This variable generates a general understanding of the circumstances under which a student is likely to fail to persist. The length of time in which a student remains enrolled is another dependent variable. Analysis with this variable suggests the timing of attrition or persistence under certain conditions. For the proportional hazard modeling, the main dependent variable again was dropping out or not, and I used a binary parental education variable (1= parent with a BA, 0= not having a parent with a BA) for the Kaplan Meier estimates.

Independent Variables

Family Background

The main independent variables are parental education (1=BA, 0= no BA) and parental income (dollars logged). Both variables represent various capital inherent in, or potentially absent from, the student's family structure. Coming from a home in which at least one parent has graduated college suggests a degree of cultural and social capital upon which a student can draw to help her complete college and, as such, is a valuable predictor in college persistence and attrition. Similarly, though functioning in a separate way, parental income serves to represent economic capital, suggesting availability of material resources from which a student can draw to complete college. Combined, parental education and parental income represent a significant portion of the capital upon which a student can rely for support during college.

Neighborhood and Community Measures

Another set of significant independent variables utilize neighborhood and high school characteristics. Using the names of the high schools the cohort members attended, I used NCES[2] to obtain high school county names for whom high school data were available. Drawing from the high school county, I gathered U. S. Census Bureau Small Area Income and Poverty Estimates (SAIPE)[3] to construct a variable for the total mean county poverty average and mean county median income averages for a 4-year period—2 years prior to the student entering the cohort, the year she entered SSU, and the year after. For estimation purposes, I divided the mean county median income by $1000. Again using the high school county name, I gathered non-seasonally adjusted county level data from the U.S. Department of Labor Bureau of Labor Statistic Local

[2] National Center for Education Statistics (NCES). Search for Public Schools. https://nces.ed.gov/ccd/schoolsearch/ Accessed 27 June 2018.
[3] U. S. Census Bureau. Small Area Income and Poverty Estimates (SAIPE). https://www.census.gov/programs-surveys/saipe/data/datasets.2007.html Accessed 27 June 2018.

Area Unemployment Statistics (LAUS)[4] for each of the high school counties and created a mean unemployment rate for each county for a 4-year period—the 2 years prior to the student entering the cohort and 2 years after. Drawing upon the U. S. Census Bureau's 2010 Demographic Profile, I added log transformed population data for each county[5]. Additionally, I included American Community Survey[6] education estimates for each cohort member (by county, by year) to estimate the percentage of people by county who are 18-24 who have a bachelor's degree (male and female) and the total number of people in the county who have at least a high school diploma (male and female). In the neighborhood and community measures and again using NCES data, I created a variable for the high school poverty rate as estimated by percentage of students eligible for the Free and Reduced Lunch (FRLP) program (drawn from 2015-2016 academic year reporting).

[Table 1]

High School Measures

For each student in the cohorts to the extent to which information was available, I used NCES to acquire information for each student's high school[7]. This information includes high school type (public or private, including charter or magnet), student/teacher ratio, total school population, ratio of white, African American, and Hispanic, and other races, and Title 1 (a federal program that provides grants to local public schools which serve a high percentage of

[4] U.S. Department of Labor. Bureau of Labor Statistics: Local Area Unemployment Statistics. https://www.bls.gov/lau/data.htm Accessed 9 July 2018.
[5] U. S. Census Bureau. American Fact Finder https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml. Information provided here derived from the Demographic Profile for each county. Accessed 12 September 2018.
[6] U. S. Census Bureau. American Fact Finder https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml. Information provided here derived from American Community Survey 5-year estimates for each county by cohort year. Accessed 12 September 2018.
[7] National Center for Education Statistics (NCES). Search for Public Schools. https://nces.ed.gov/ccd/schoolsearch/ Accessed 27 June 2018.

low-income students) eligibility. Title 1 funds can provide support either for individual schools or for schools district wide (U.S. DoE, 2015), and program eligibility is often used as a proxy for high school SES. I also used the NCES high school locale classification system to create a geographic variable with 4 basic categories (city, suburban, town, and rural) and 3 subcategories, rendering 12 separate variable designations[8]. Due to availability constraints, all high school data is drawn from the 2015-2016 academic year reporting.

Academic Information and Performance Measures

To proximally assess each student's academic readiness, I include high school grade point average (GPA) and composite ACT scores. All high school GPAs were standardized to a 4.0 scale.

Demographic/Control Variables

In all models, I control for gender (female [reference] and male), race (white [reference], African American, Hispanic, other), cohort year (2007-2014), age[9], age squared, missing parental income data, parental marital status (married, single, divorced, widowed), household size, and number of students in the household who are in college. I also include dummy variables for students who are in-state students (1= in state, 0= out-of-state) and having someone else from the same high school attend SSU in the same cohort year (1=cohort member, 0= no cohort member).

[8] https://nces.ed.gov/programs/edge/docs/LOCALE_CLASSIFICATIONS.pdf Accessed 11 August 2018

[9] Definitive categories for freshman student classification are noticeably absent from the literature. Conventionally, "traditional students" are those who enroll in college under the age of 25 (see NCES: https://nces.ed.gov/pubs/web/97578e.asp, for example); however, students who are 20-23, those with delayed enrollment, often demonstrate different educational trajectories than "traditional" students (18-19) or older, more quintessentially "non-traditional" students (24 and older), so they should be treated as a separate group (Hurtado, Kurotsuchi, & Sharp, 1996). My purpose here is not to dissect the nuances of trajectory; rather, it is to look at general patterns within the cohorts stratified by year rather than grouped by age. Thus, I censored by group to 24 and under, which is the cutoff SSU uses in some reporting.

Statistical Modeling

To analyze this unique data set, I implement linear probability modeling (LPM), which makes the assumption that the function of the independent variables is linear in relationship to the probability of the binary dependent variable. Linear models, rather than other models, like logit for example, allow easier interpretation of the coefficient and the relationship between the independent and dependent variables, even when the dependent variable is dichotomous (Hellevik, 2009). LPM can be estimated by the equation

$$\Pr_i (y=1|x) = \beta_0 + \beta_1 \boldsymbol{PE} + \beta_2 \boldsymbol{PI} + \Sigma\, \delta \boldsymbol{NB} + \Sigma\, \delta \boldsymbol{HS} + \Sigma\, \delta \boldsymbol{AP} + \Sigma\delta\, (\textbf{PE} \times \textbf{Race}) \quad (1)$$

$$+ \Sigma\delta\, (\textbf{Female} \times \textbf{Race})\, \Sigma\, \delta \boldsymbol{X} + e$$

where *i* is the probability that a student dropped out of the university, and β_0 is the constant. ***PE*** represents having at least one parent with a college degree (1= BA, 0= no BA), and ***PI*** is the log of parental income reported on the student's FASFA. ***NB*** is the vector of 9 neighborhood characteristics: county population, unemployment, poverty, and median income; high school poverty index; percentage of males and females in the county with high school diplomas and percentage of males and females in the county with college degrees. The vector of ***HS*** includes 8 high school variables: type of high school, high school Title 1 eligibility, student-teacher ratio, high school total population, percentage African American, Hispanic, and other race, and high school location (using 12 NCES locale categories). ***AP*** is the vector of academic performance of high school GPA and composite ACT scores. **PE** x **Race** is the interaction between the parental education dummy variable and race (African American, Hispanic, other), thus representing the benefits of being a continuing-generation student on the likelihood of dropping out of college

relative to a white, continuing-generation student. **Female** x **Race** is the interaction between the female dummy variable and race, representing the potential benefits for being female and African American, Hispanic, or other. ***X*** represents the vector of control variables, which includes, being female, age, age squared, race (white [reference], Hispanic, African American, or other race), cohort year, parental marital status (married, single, divorced, widowed), log student income, missing parental income, missing student income, household size, number of people in the household who are in college, living in state, and having a high school cohort member. e represents the residuals in the model.

I also execute survival analysis and Cox proportional hazard modeling, which allows a closer examination of the predictors and timing of attrition. Cox modeling focuses on individual time-to-event analysis (Cox, 1972), which in this case is dropping out of college. This kind of analysis is useful in using criteria of interest in predicting attrition probabilities (Murtaugh, Burns, & Schuster, 1999). Table 5 records the coefficients for estimated probabilities of continued enrollment at years 1, 2, 4, and 6 based on Kaplan-Meier (KM) estimations. Figures 1 and 2 show KM curve estimations for the entire cohort and by parental college degree attainment, respectively.

In the creation of Table 6, I utilized the following model for a Cox proportional hazard regression:

$$h(t)=h_0(t)\ \exp[\beta_1 X_1+\dots\ \beta_p X_p] \qquad (2)$$

The Cox proportional hazard model (1972) is used to predict the hazard that an event will occur based on a set of predictors, with *t* representing the amount of time students persist and *h(t)* representing the hazard function that someone will experience attrition. In Equation 2, $\beta_{1\dots}\ \beta_p$

represent the predictor coefficients of $X_{1...}X_p$, which are similar to the ones used in Equation 1. Because my primary question involves how family and neighborhood socioeconomic characteristics correlate to college attrition, I only considered parental education and parental income for family background and the full vector of neighborhood variables used in Equation 1, omitting the high school vector present in my previous analyses. To test the proportionality assumption, I used the Schoenfeld and scaled Schoenfeld residual and the log-rank test of equality to help determine the most appropriate variables in the final model.

Results

Descriptive Statistics

This study examines family and neighborhood socioeconomic and high school characteristics of the 2007-2014 first-time, full-time freshman cohorts at SSU, a Midwestern teaching university. Table 2 provides an overview of some of the censored cohort characteristics. Sociological demographics were consistent with current trends: women significantly out-enrolled men, 62.9% to 37.1%, and men dropped out at a greater rate (12.4%) than women.

The vast majority of SSU students in this cohort are white (78.6%), which is about 8% above the university's average, and graduated at the highest rate (42.2%) among the races included. African Americans enrolled at the lowest rate (5.9%) and were least likely to matriculate (25.9%). These statistics are consistent with previous work on college enrollment and graduation. In addition, the graduation and attrition data on family income is also similar to previous research, with student attrition operating on a negative gradient. Students from families below 1 standard deviation of the mean dropped out more than 30% more often than students from above 1 standard deviation of the mean. These students fared best statistically; odds were even that affluent students would graduate or dropout.

[Table 2]

Inequality is present, though perhaps least significant, in students' academic preparation. First-generation students were slightly less prepared academically than their continuing-generation peers, having lower ACT and lower GPA averages, though the differences were small, especially between high school GPA for first-generation and continuing-generation students who graduated. Taken together, policy makers should consider program requirements that help support low-income students who are coming to college less prepared, and are, unsurprisingly, dropping out at a higher rate than their more affluent peers; I take this up later in the discussion section.

Who is Dropping Out?

To understand who is dropping out of SSU and under what conditions, I utilize linear probability modeling (Table 3; Appendix Table 2). All models are restricted to cohort members 16-24 years and for whom there was available information on parental education (N=4,433). Model 1 includes my main variables of interest—having a parent with a BA and parental income—as well as all significant family background and demographic variables. Being a first-generation student, being Hispanic, and being male are all highly and significantly associated (≤.001) with the likelihood of dropping out. Being a race other than white, African American, or Hispanic is moderately associated with dropping out (.089 at ≤.01), and being African American is only modestly related to a student's chances of attrition. In this model, log parental income is also meaningfully related to attrition, though the coefficient is slightly less than that for parental education (-.048 and -.057 respectively).

Model 2 adds several neighborhood characteristic variables to those in Model 1. In this model a number of variables remain highly significant, though the addition of neighborhood

variables nominally reduces the strength of most of the coefficients. Parental education and log parental income lose .001 and .005, respectively, while being female loses .003, and being in the racial group of other modestly loses coefficient strength (.005). The strength of the coefficient for the cohort year diminishes by .002 though the variable's statistical power remains unchanged at ≤.001 across all models. The addition of neighborhood variables, however, does nominally increase the strength of the coefficient, though not the significance, of being African American; in Model 2, African Americans are .013 more likely to drop out when controlling for county and high school characteristics. Being Hispanic in these models also increases a student's chances of dropping out, though only by .002 from Model 1. No additional variables, however, are themselves significant in Model 2.

I add academic preparation—high school GPA and composite ACT scores—in Model 3. As when I added neighborhood controls in Model 2, a number of coefficients lose strength, though unlike in Model 2, some of these changes are slightly more intense and more variables lose statistical power. The coefficient for having a parent with a BA loses .013, reducing its coefficient to -.043, and its statistical significance is reduced slightly as well (≤.01). Parental income loses both strength and significance: the coefficient decreases by .014 (-.029) and becomes only liminally significant. Though retaining its statistical power, the strength of being female weakens to -.085, a .021 reduction from Model 2. Being African American loses all statistical power, and being Hispanic is reduced in both strength (down .030 to .068) and significance (≤.05). The coefficient for being a race other than white, African American, or Hispanic drops by .023 to .051 and is reduced to minimal significance. Further, being female remains highly significant, with women being more than 8% less likely to drop out than men.

[Table 3]

The addition of academic performance alters a few of the neighborhood variables added in Model 2. County unemployment becomes significant at ≤.01, suggesting that as local unemployment increases by 1%, the chances a student from that county will drop out decreases by nearly 2 percentage points. In the county, the number of both men and women with high school diplomas become marginally significant but in very different ways. As the number of men in the student's home county with high school diplomas increases, the likelihood that student will drop out decreases by 49%, net of other factors. Conversely, as the number of women with high school diplomas in the county increases, the likelihood of the student from that county dropping out increases to 75%, again when controlling for all other variables. Both elements of academic performance—high school GPA and composite ACT scores—are highly significant, though they vary greatly in strength. For each 1 point increase in high school GPA, students are 9.6% less likely to drop out. Contrariwise, ACT scores count for less than 1% change in likelihood in retention. These data remain nearly identical across Models 3, 4, and 5. Finally, in this model, the R^2 value jumps from .071 in Model 2 to .115 here, suggesting a better fit in the analysis of the relationship of academic preparation to college persistence and retention.

Models 4 and 5 (see Appendix: Table 2) both include the addition of interaction variables. Model 4 introduces an interaction between parental education and race. Adding this variable returns both strength and magnitude to the main dependent variable, parental education; students who have at least one parent with a college degree are 5.8% less likely to drop out than their first-generation peers (≤.001). The effect of parental income remains identical from Model 2, county unemployment remains nearly unchanged, and the coefficients associated with the percentage of people in the county with high school diplomas reduce only slightly and remain marginally significant. Academic performance is nearly the same as in Model 3; high school

GPA loses only .001 in the coefficient, and the ACT composite score remains unchanged. The significance of race *singularly* drops off for both being Hispanic and other; however, the interaction between being Hispanic and having one parent have a college degree is both strong and significant at .195 (≤.001). While the main effect alone of being Hispanic is not significant, the net effect of being Hispanic and having one parent with a BA is .206. In Model 4, the R^2 very slightly increases by .002 to .117.

Model 5 adds an interaction variable between race and being female. Outside the main female variable, very little changes. Parental education remains nearly the same, and log parent income is identical across models. The coefficient for the number of people in the county who have high school diplomas changes slightly for both men and women (from -.484 to -.478 for men and .735 to .726 for women), and the significance remains liminal. The net effect of being Hispanic changes a bit, including the race and female interaction increases the total to .317, up .111 from the previous model. Adding this interaction variable also alters the coefficient for the main effects of being female to -.104 (a gain of .019), though no interaction terms themselves are significant. The R^2 increases only incrementally to .118.

Between the Years: Comparing Dropout Characteristics by Attrition Timing

In the previous section, I explored the broad characteristics of who is dropping out from college. In this section, I use the same linear probability models as described previously to compare dropout characteristics, but this time I do so estimating for drop out timing characteristics. Table 4 and Appendix: Table 3 show the comparative estimations for students who dropped out after the 1st or 2nd year, after the 3rd or 4th year, or after year 5. Model 1 suggests two notable factors in predicting attrition. First, the number of people in students' counties who have completed high school figures significantly in the likelihood of both dropping

out and persisting, though it functions differently between men and women. As the percentage of men in the county with high school diplomas increases, the likelihood of dropping out decreases by almost 6 percent. Conversely, as the number of women in the county who have high school diplomas increases, the likelihood of dropping out after year 1 or year 2 also increases, this time by nearly 9 percent. Additionally, academic preparation is also significant, though weaker. As student GPA increases by 1 point, the likelihood that a student will dropout decreases by 6.9%. While statistically strong, the effects of ACT score are nominal. ACT effects continue to diminish in strength and significance across the models.

[Table 4]

Model 2 examines students who dropped out after years 3 or 4. In this model three new factors becoming significant to a student's chances of attrition, with county unemployment average becomes the strongest predictor. As the average unemployment rate in a student's home county increases by 1%, the chances of that student dropping out strongly reduces by 5%. Compared to Model 1, in which being female was only marginally significant as a predictor, in Model 2 being female is strongly and significantly associated with persisting, with women having a nearly 6% chance of returning after years 3 or 4 than men. In Model 2 the cohort year become highly significant, suggesting larger, external forces become determiners of persistence and retention. As in Model 1, high school GPA is significant though weaker, the coefficient reducing by .029, suggesting the diminishing effects of academic preparedness and the increasing effects of other forces influencing a student's chances of attrition. Similarly, composite ACT scores nominally lose strength and significance.

Finally, Model 3 shows the estimation of effects if a student persists to year 5, where being Hispanic is most strongly predictive of persisting through year 5 then dropping out without

matriculating (attrition is 16.6 % more likely relative to a white student). Another change from Model 2 to 3 is the loss of unemployment as a significant predictor; it loses all statistical power and strength. The same occurs for cohort year. From Model 2 to Model 3, the chances of a woman persisting, relative to men, increases to -.071, a gain of .012, and the effect of high school GPA increases by .031 to -.071. As from Model 1 to Model 2, composite ACT scores lose strength and power in Model 3, becoming nearly insignificant. The R^2 jumps from .067 to .129 from Model 1 to Model 3. Appendix: Table 4 shows complete estimation for this model, including the shifting importance of the cohort year. In Model 1, the cohort year is barely significant, and in Model 3, cohort year is not significant at all. However, when estimating the likelihood of attrition after years 3 or 4, the cohort year is highly significant, affecting the chances of dropping out by 4.8%.

Who is Surviving? How Long?

Kaplan-Meier survivor function estimates in Table 5 show the estimates of persistence at years 1, 2, 4, and 6. Consistent with LPM, parental education is strongly related to persisting. The probability gap between students who have two parents with education and those whose parents have no college degree is fairly constant, averaging about 10 percentage points between year 1 and year 4. While the magnitude is smaller, the probabilities gap between students with one parent who has graduated college and those who do not have a parent who has graduated is similarly consistent across enrollment years at an average of 6%.

[Table 5]

Also consistent with other findings, being female decreases a student's probability of dropping out. The probability of female retention at the end of year one is 3.8% greater than that

of men, and that difference in probability of continued enrollment continues to grow. At the end of year 2, the difference increases to 6.3% and is 8.4% at the end of year 4.

[Figure 1]

Figure 1 is the Kaplan-Meier curve estimating the survival estimates for all members of the SSU 2007-2014 cohort by parental education status and academic year, as indicated by intervals of 3 semesters. Consistent with previous research as well as my other findings, students who are continuing generation, represented by the red line, persist longer and drop out at lower rates than first-generation students.

Figure 2 is the assessment of proportional hazards using one parent with a college degree as the predictor. After a relatively tight initial drop, the gap between dropout rates between first-generation and continuing generation students is fairly consistent until year 5.

[Figure 2]

Table 6 Model 1 shows the proportional hazards of college attrition. Results in univariate analysis are quite similar to those found in LPM. Being first-generation strongly increases a student's chances of attrition, especially relative to having two parents who have college degrees. As in my other analyses, Parental income, neighborhood effects, race and gender also contribute to the risk of dropping out. Different from other findings, however, parental marital status contributes to attrition hazards.

[Table 6]

I use the same set of variables in multivariate proportional hazard analysis (Table 6 Model 2) as in linear probability modeling. Consistent with LPM results, parental education is more significant to a student's risk of attrition than is parental income. Similar to Table 4 Model 1, the number of males and females in the county who have at least a high school diploma is

related to the chances of dropping out—again, the chances are greater when more women have high school diplomas than men, and in this model the strength is remarkable. As demonstrated in earlier analyses, being female also greatly reduces the risk of dropping out of college relative to men. Additionally, Appendix: Table 5 shows the results for the complete univariate and multivariate proportional hazard estimates. Consistent with other findings, cohort year is strongly significant related to relative dropout risk. However, unlike in other models, the percent of non-white, non-African American, non-Hispanic students at a student's high school is strong and moderately significant. This finding warrants further investigation.

Robustness Check

As I created the linear probability models, I tried fitting a number of interaction variables with parental education, parental income, academic preparation, gender, and race. Only the results that were significant in and of themselves or changed the coefficients and/or statistical significance of the other variables in a meaningful way were included. Additionally, I attempted a number of combinations for the timing of dropping out in Table 4, including running the years individually and in a variety of timing groupings. The results were similar across all models, and I chose the current modeling for conciseness. See Appendix for additional modeling.

Discussion

Through analysis of original administrative data for the 2007-2014 first-time, full-time freshman cohorts at a Midwestern teaching university, I have examined the continuing importance of social class, family background, neighborhood contexts, and academic readiness on the chances of dropping out from college before completing a degree. While many of the results are consistent with other higher education and sociological research, including social class and female advantages, the results of this analysis provide some new insight to the study of

college attrition. Despite other findings in studying the sociology of higher education, in this study social class matters, especially as suggested by parental education, in predicating the relatively likelihood of a student dropping out of college. A student's social origin continues to exert influence on her, even in college, and inequalities between advantaged and disadvantaged groups, i.e., working-class students and students of color, are evident in the disparity of dropping out of college.

In addition to social origin, academic readiness is a strong predictor of college attrition: the worse a student's high school GPA, the more likely she is to drop out. This is especially significant for open-enrollment and low-selectivity colleges and universities as these institutions are likely to enroll students with low academic preparation, thus admitting students who are at high risk for dropping out. While race alone was not significant in predicting a student's chances for not completing a degree, minority students are most likely to come to college ill-prepared for the rigors of academic life, demonstrated namely by having low high school GPAs. Students who come in with low GPAs often need remedial courses and additional forms of academic support to keep them enrolled, persist, and graduate. Without further support, these students enroll and dropout, thus decreasing the overall graduation rates for open-enrollment and low-selectivity universities and incur additional financial burdens without the benefit of a degree.

Consistent with other literature, being male strongly predicts college attrition, and this finding is consistent across all models. Overall, men are approximately 9% more likely to drop out of college than their female peers. Since men are enrolling in institutions of higher education less often than are women, this increased chance of dropping out further exacerbates the widening gender gap in college attainment.

Parental education is the single most consistent family background measure on college retention or likelihood of withdrawal. First-generation students enrolled less frequently (33%) than their peers who had at least one parent complete college (47%) and dropped out faster and at a higher rate (68%) than their peers who had at least one parent with a BA (57%); thus, continuing-generation students are less likely to withdraw and are more likely to persist longer in college than their first-generation peers. Similarly, when controlling for a student's high school characteristics, being a first-generation student increases the likelihood of dropping out by 7.8%.

In educational literature, parental income is often the lynchpin for educational achievement. The preceding results do not support that assertion. Because parental income was not statistically significant, potential accessibility to financial and economic resources do not seem to relate to college persistence or attrition, which is contrary to other findings (Tinto, 1993). While students who have parents with college degrees live in counties with higher median incomes, neighborhood characteristics did not demonstrate statistical significance in full linear probability models. Other studies have demonstrated the importance of neighborhood characteristics on college outcomes (Owens, 2010; Vartanian & Gleason 1999) though analysis of this data does not support those findings. High school measures were not significant relating to college graduation either, neither singularly nor when controlling for other variables.

My results contribute to sociology of higher education research in a number of meaningful ways. First, even after controlling for a rich variety of other variables, parental education is a stronger predictor than parental income for risk of attrition. More students who have at least one parent with a BA have higher GPAs than their first-generation peers. If we accept parental education as a critical element of social class, working-class students are coming to college already disadvantaged and underprepared; they lack the capital transmissions that their

more advantaged peers have received, and the results are evident in the rate of attrition for these students.

This finding is especially significant for open-enrollment and low-selectivity colleges and universities as these institutions are more likely to enroll working-class, first-generation students, who have low academic preparation and are at high risk for dropping out, characteristics which have significant financial implications for both students and institutions. Students admitted with low GPAs often need remedial courses and additional forms of academic support for continued enrollment. Remedial courses mean incurring a greater financial burden for more credit hours and, presumably, additional time outside the labor force. While student income was never a strong predictor for attrition, more than 41% of the students in the cohort held a paying job in the year in which they enrolled, and other research suggests that students who work more academically achieve less (Stinebrickner & Stinebrickner, 2003*b*). Institutions in which these students enroll, namely open-enrollment and low-selectivity schools, have lower retention and graduation rates than other types of colleges and universities, and the findings provided by this study offer some suggestions as to why.

Second, in addition to its relationship to academic preparedness, parental education in its own right is a strong, social class predictor of a student enrolling then graduating or dropping out of college. First-generation students enrolled less frequently than their peers who had at least one parent complete college and dropped out faster and at a higher rate than their peers who had at least one parent with a BA. These persistence gaps remained even after controlling for a number of other effects, including interaction effects and academic performance.

Finally, my results contribute significantly to understanding some of the contexts under which students drop out, especially after reaching year 4. Students who complete 4 years of

college and then dropout appear to be differentially affected by SES neighborhood contexts. Students who attended a Title 1 eligible high school were more likely to persist to year 4 then fail to graduate than those from a higher SES school. Unemployment rates also seem to be significant predicators for attrition and retention: students who came from counties with higher instances of unemployment were less likely to persist to year 4 then drop out, underscoring the potentially lasting effects of neighborhood SES on college completion.

Throughout this study, I have demonstrated the continuing influence of social class and neighborhood SES on students who enroll but fail to complete college. The unequal preparedness with which students enroll and persist in college is inherently related to parental education and various capital transmission. The inequalities with which a student enrolls increases the inequality in likelihood of college completion. Working-class students and students of color, who are already disadvantaged socially and academically, experience further barriers which prevent them, more than their continuing-generation and white peers, from graduating college.

To help alleviate some of the stratification in college graduation rates among various groups, higher education administrators must consider what policy measures could be adopted to support at-risk students (i.e., working-class students and students of color) academically. Increased academic support centers can provide tutoring in a variety of disciplines and coaching in other areas in which students may struggle. SSU implemented such a center with a soft launch in Fall 2015, and a more comprehensive rollout in January, 2016. The center offers academic one-on-one and group peer tutoring in high DFW courses as well as workshops to bolster what Sommerfeld (2011) calls "non-academic" factors, those skills which students need to be successful, like study skills and time management. The SSU Writing Center and English Language Learner Lab operates within the academic support center as well, providing

undergraduate and graduate writing process support for domestic and international students. Data are not yet available to fully determine the effectiveness of the center on student retention, but preliminary information on retention rates show increased retention for the 2015 and 2016 cohorts compared to the previous 5 year averages. 1st term retention has increased by 1.44%, 3rd term by 2.7%, and 2nd year by 3.85%.

Limitations

While contributing to the research on attrition in higher education, this study has a few limitations. First, I was not able to distinguish between students who dropped out and who transferred to another university. Additionally, I used the most current NCES high school data (2015-2016) which did not exactly reflect the characteristics of the schools at the time of the cohort member's attendance. However, it seems unlikely that major shifts in demographics occurred between the earliest cohort's attendance and when the data was compiled.

Implications

The data show that as local unemployment rates increase, students are less likely to drop out of college. However, the local unemployment rates matter for more than just college enrollment. Workers do not flock to, nor return to, areas in which prospects in the labor force are low. For instance, the state of Indiana trains more than sufficient workers, but those workers are not returning to their home communities when better market opportunities lie elsewhere (Hicks, 2013). County level policy makers must consider more broadly ways to entice workers to an area which has unstable employment. Further, considering the manifold benefits of a college degree to an individual, her employer, and those around her, business owners need to examine policies, like subsidies or other enticements, to keep students in school when employment rates do go up.

In addition to the academic support, many universities, SSU included, utilize other programing to support first-year students, including freshman seminar courses and majors-based learning communities, both of which are high impact learning components[10] as well as social programs designed to increase the students' sense of community and belonging. Having a sense of belonging is important in student retention, and those who do not believe they belong in college are more likely to drop out (Tinto, 2017; Tinto, 1993). This is exceptionally important to consider as men and women access support networks differently. Men and women communicate differently (see Tannen, for example), and the ways in which men participate in and experience college campus life are different as well.

Men might fail to persist as long as women in college for a number of reasons, most notably, however, might be the disparity of opportunities in the labor market. Men have more opportunities to participate in paid labor without a college degree than do women; thus, women have a greater return to college than do men. If we want to truly create an egalitarian environment on open-enrollment and low-selectivity campuses, policies and programs need to address ways to keep men in school. Much research, notably Tinto, demonstrates the importance of community and belonging on student retention. If women are the ones being most retained with these initiatives, what might work to keep men enrolled?

This study demonstrates the importance and significance of academic preparation in college retention. For open-enrollment and low-selectivity universities, this finding highlights the importance of the availability and usage of academic support units, especially to increase male persistence. Further, increased availability of developmental courses in math and reading can

[10] See Kuh G. D. (2008). High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter, for example.

offer a transition between high school and college level rigor for students who may be weak in these academic areas. Mandatory participation in freshman seminar courses have also been shown to increase student retention, as have the creation of smaller major communities. These latter two operate under the principle that when students are engaged in a community, they report a greater sense of belonging and are, therefore, likely to remain in school (Tinto, 2017; Tinto, 1993).

While this study has contributed to the literature on who is dropping out of college, it leaves some questions for future research as well. First, if, as this study has demonstrated, parental college graduation is significant in predicting college retention and attrition, yet parental income is not, what specifically do college educated parents transmit to their children that is not economically related but makes them academically successful? Is that unidentified transmission teachable? Could a program secondarily supplement what continuing-generation students receive from their families but first-generation students do not? Alternatively, is the something transmitted so implicitly cultural that external support may be insufficient? The second set of questions requires us to step back from college and look at improving academic performance at the high school level, which demands examining contexts both inside and outside of school. While the quality of education a student receives at secondary school is certainly important, situations external to school often affect learning more than what goes on inside the classroom. Family poverty, inadequate, inconsistent, or no housing, neighborhood violence, and lack of access to adequate healthcare are all social barriers to academic success. To prepare students for long-term academic success, namely graduate from college, policy makers need to find ways to support students and their families long before they enter a college campus.

REFERENCES

Adelman, C. (2004). *Principal indicators of student academic histories in post-secondary education, 1972–2000.* Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Alon, S. (2014). Continuing to build a theory of inequality in higher education: Claims, evidence, and future directions. *American Sociological Review*, *79*(4), 817-824.

Alon, S., Domina, T., & Tienda, M. (2010). Stymied mobility or temporary lull? The puzzle of lagging Hispanic college degree attainment. *Social Forces*, *88*(4), 1807–1832.

Attewell, P., Lavin, D., Domina, T., & Levey, T. (2006). New evidence on college remediation. *Journal of Higher Education*, *77*(5), 886–924.

Bailey, M.J. & Dynarski, S.M. (2011). Inequality in postsecondary education. In G.J. Duncan & R. J. Murnane (eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (117-131). New York, NY: Russell Sage Foundation.

Barr, A. & Turner, S. (2013). Expanding enrollments and contracting state budgets: The effect of the Great Recession on higher education. *The Annals of the American Academy of Political and Social Science, 650*, 168-193.

Berg, M. T., Stewart, E. A., Stewart, E., & Simons, R. L. (2013). A multilevel examination of neighborhood social processes and college enrollment. *Social Problems*, *60*(4), 513–534.

Blau P. M. & Duncan, O. D. (1967). *The American occupational structure*. New York, NY: Wiley.

Bourdieu, P. (1986). The forms of capital. In J.G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (241-58). New York, NY: Greenwood.

Brand, J., & Xie, Y. (2010). Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education. *American Sociological Review, 75*(2), 273-302.

Buchman, C., & DiPrete, T. A. (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review*, *71*(4), 515–541.

Burdick-Will, J., Ludwig, J., Raudenbush, S. W., Sampson, R. J., Sanbonmatsu, L. & Sharkey, P. (2011). Converging evidence for neighborhood effects on children's test scores: An experimental quasi-experimental, and observational comparison. In G.J. Duncan & R. J. Murnane (eds.), *Whither opportunity? Rising inequality, schools, and children's life chances (*255-276). New York, NY: Russell Sage Foundation.

Card, D. & Lemieux, T. (2001). Dropout and enrollment trends in the postwar period: What went wrong in the 1970s? In J. Gruber (ed), *Risky behavior among youths: An economic analysis* (439-82). Chicago, IL: University of Chicago Press.

Cataldi, E. F., Bennett, C. T., & Chen, X. (2018). First-generation students: College access, persistence, and postbachelor's outcomes. U.S. Department of Education: Statistics in brief. NCES 2018-421.

Chetty, R., & Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *Quarterly Journal of Economics*, *133*(3), 1107–1162.

Coleman, J. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*, S95-S120.

Contreras, F., & Contreras, G. J. (2015). Raising the bar for Hispanic serving institutions: An analysis of college completion and success rates. *Journal of Hispanic Higher Education*, *14*(2), 151–170.

Cox, D. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B, 34*(2), 187-220.

Davis, T., & Otto, B. (2016). Juxtaposing the black and white gender gap: Race and gender differentiation in college enrollment predictors. *Social Science*, *97*(5), 1245–1266.

Flowers, L. A., Milner, H. R., & Moore, J. L., III. (2003). Effects of locus of control on African American high school seniors' educational aspirations: Implications for preservice and inservice high school teachers and counselors. *High School Journal*, *87*(1), 39–50.

Fry, R. (2009). College enrollment his all-time high, fueled by community college surge. Pew Research Center: A social & demographic trends report.

Garcia, L. M., & Bayer, A. E. (2005). Variations between Latino groups in US post-secondary educational attainment. *Research in Higher Education*, *46*(5), 511–533.

Goldin, C., Katz, L. F., & Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, *20*(4), 133-A4.

Goldrick-Rab, S., Kelchen, R., Harris, D. N. ., & Benson, J. (2016). Reducing income inequality in educational attainment: Experimental evidence on the impact of financial aid on college completion. *American Journal of Sociology*, *121*(6), 1762–1817.

Hansen, M., & Mastekaasa, A. (2006). Social origins and academic performance at university. *European Sociological Review, 22*(3), 277-291.

Haveman, R., Sandefur, G., Wolfe, B., & Voyer, A. (2004) Trends in children's attainments

and their determinants as family income inequality has increased. In K. M. Neckerman (ed), *Social inequality* (149-188). New York, NY: Russel Sage Foundation.

Hellevik, O. (2009). Linear versus logistic regression with the dependent variable is dichotomy. *Quality & Quantity 43*(1), 59–74.

Hepworth, D., Littlepage, B., & Hancock, K. (2018). Factors influencing university student academic success. *Educational Research Quarterly*, *42*(1), 45–61.

Hicks, A. L., Handcock, M. S., Sastry, N., & Pebley, A. R. (2018). Sequential neighborhood effects: The effect of long-term exposure to concentrated disadvantage on children's reading and math test scores. *Demography*, *55*(1), 1–31.

Hicks, M. J. (2013). Notes on: Labor markets after the Great Recession: Unemployment and policy for Indiana. *Labor Law Journal*, *64*(2), 103–113.

Hillman, N., & Orians, E. (2013). Community colleges and labor market conditions: How does enrollment demand change relative to local unemployment rates? *Research in Higher Education*, *54*(7), 765–780.

Hout, M. (1984). Status, autonomy, and training in occupational mobility. *American Journal of Sociology, 89*(6), 1379-1409.

Hout, M. (1988). More universalism, less structural mobility: The American occupational structure in the 1980s. *American Journal of Sociology, 93*(6), 1358-1400.

Hout, M. (2012). Social and economic returns to college education in the United States. *Annual Review of Sociology, 38*, 379-400.

Hout, M. & Janus, A. (2011). Educational mobility in the United States since the 1930s. In G.J. Duncan & R. J. Murnane (eds.), *Whither opportunity? Rising inequality, schools, and children's life chances (*91-115). New York, NY: Russell Sage Foundation.

Hurd, M. D., & Rohwedder, S. (2010). Effects of the financial crisis and Great Recession on American households. *National Bureau of Economic Research, Inc*, NBER Working Papers: 16407.

Hurtado, Kurotsuchi, K. & Sharp, S. (1996). Traditional, delayed entry, and nontraditional students. Presented at the annual meeting of the American Educational Research Association, 1996.

Johnson Jr., O. (2010). Assessing neighborhood racial segregation and macroeconomic effects in the education of African Americans. *Review of Educational Research*, *80*(4), 527–575.

Kane, T. J. (2004). College-going and inequality. In K. M. Neckerman (ed), *Social inequality* (310-353). New York, NY: Russel Sage Foundation.

Kiker, B., & Condon, C. (1981). The influence of socioeconomic background on the earnings of young men. *The Journal of Human Resources, 16*(1), 94-105.

Kim, C., & Sakamoto, A. (2008). The Rise of intra-occupational wage inequality in the United States, 1983 to 2002. *American Sociological Review, 73*(1), 129-157.

Kim, C., & Sakamoto, A. (2017). Women's progress for men's gain? Gender-specific changes in the return to education as measured by family standard of living, 1990 to 2009-2011. *Demography*, *54*(5), 1743–1772.

Kingston, P., Hubbard, R., Lapp, B., Schroeder, P., & Wilson, J. (2003). Why education matters. *Sociology of Education, 76*(1), 53-70.

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life* (2nd ed). Berkeley, CA: University of California Press.

Long, B. T. (2014). The financial crisis and college enrollment: How have students and their families responded? In J. R. Brown & C. M. Hoxby (Eds.), *How the financial crisis and Great Recession affected higher education (*209-233). Chicago, IL: University of Chicago Press.

Mare, R. (1980). Social background and school continuation decisions. *Journal of the American Statistical Association, 75*(370), 295-305.

Martinez, G. F., & Deil-Amen, R. (2015). College for all Latinos? The role of high school messages in facing college challenges. *Teachers College Record*, *117*(3), 1–50.

Merolla, D. M. (2018). Completing the educational career: High school graduation, four-year college enrollment, and bachelor's degree completion among Black, Hispanic, and white students. *Sociology of Race and Ethnicity*, *4*(2), 281–297.

Moore, G. W., Slate, J. R., Edmonson, S. L., Combs, J. P., Bustamante, R., & Onwuegbuzie, A. J. (2010). High school students and their lack of preparedness for college: A statewide study. *Education & Urban Society*, *42*(7), 817–838.

Morgan, S., Leenman, T., Todd, J., & Weeden, K. (2013). Occupational plans, beliefs about educational requirements, and patterns of college entry. *Sociology of Education, 86*(3), 197-217.

Murtaugh, P.A, Burns, L. D., & Schuster, J. (1999). Predicting the retention of university students. *Research in Higher Education, 40*(3), 355-371.

National Center for Education Statistics (1993). 120 years of American education: A statistical portrait. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=93442

National Center for Education Statistics (2018). Undergraduate enrollment. https://nces.ed.gov/programs/coe/indicator_cha.asp

Owens, A. (2010). Neighborhoods and schools as competing and reinforcing contexts for educational attainment. *Sociology of Education, 83*(4), 287-311.

Quillian, L. (2014). Does segregation create winners and losers? Residential segregation and inequality in educational attainment. *Social Problems*, *61*(3), 402-426.

Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, *116*(4), 1092–1153.

Ross, C., & Mirowsky, J. (1999). Refining the association between education and health: The effects of quantity, credential, and selectivity. *Demography, 36*(4), 445-460.

Saunders-Scott, D., Braley, M. B., & Stennes-Spidahl, N. (2018). Traditional and psychological factors associated with academic success: Investigating best predictors of college retention. *Motivation & Emotion*, *42*(4), 459–465.

Sawyer, R. (2013). Beyond correlations: Usefulness of high school GPA and test scores in making college admissions decisions. *Applied Measurement in Education*, *26*(2), 89–112.

Schmidt, E. P. (2018). Postsecondary enrollment before, during, and since the Great Recession. *Current population reports*. U. S. Census: P20-580.

Smith-Maddox, R. (1999). The social networks and resources of African American eighth graders: Evidence from the National Education Longitudinal Study of 1988. *Adolescence*, *34*(133), 169–183.

Sommerfeld, A. (2011). Recasting non-cognitive factors in college readiness as what they truly are: Non-academic factors. *Journal of College Admission*, (213), 18–22.

St. John, E. P., Andrieu, S., Oescher, J., & Starkey, J.B. (1994). The influence of prices on within-year persistence by traditional college-age students in four-year colleges. *Research in Higher Education, 35*(4), 455-480.

Stewart, E. B., Stewart, E. A., & Simons, R. L. (2007). The effect of neighborhood context on the college aspirations of African American adolescents. *American Educational Research Journal*, *44*(4), 896–919.

Stinebrickner, R., & Stinebrickner, T. R. (2003). Understanding educational outcomes of students from low-income families: Evidence from a liberal arts college with a full tuition subsidy program. *Journal of Human Resources*, *38*(3), 591-617.

Stinebrickner, R., & Stinebrickner, T. R (2003b). Working during school and academic performance. *Journal of Labor Economics, 21*(2): 473–491.

Strayhorn, T. L. (2014). Modeling the determinants of college readiness for historically underrepresented students at 4-year colleges and universities: A national investigation. *American Behavioral Scientist*, *58*(8), 972–993.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed). Chicago, IL: University of Chicago Press.

Tinto, V. (2012). *Completing college: Rethinking institutional action*. Chicago, IL: University of Chicago Press.

Tinto, V. (2017). Through the eyes of students. *Journal of College Student Retention: Research, Theory & Practice*, *19*(3), 254–269.

Torche, F. (2011). Is a college degree still the great equalizer? Intergenerational mobility across levels of schooling in the United States. *American Journal of Sociology, 117*(3), 763-807.

U. S. Department of Education (2015). Programs: Improving basic programs operated by local education agencies (Title 1, Part A). https://www2.ed.gov/programs/titleiparta/index.html

Vartanian, T. P. & Gleason, P. M. (1999). Do neighborhood conditions affect high school dropout and college graduation rates? *The Journal of Socio-Economics* 28(*1*), 21-41.

Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M. R., & Schmidt, F. L. (2015). College performance and retention: A meta-analysis of the predictive validities of ACT scores, high school grades, and SES. *Educational Assessment*, *20*(1), 23–45.

Zweig, M. (2012). *The working class majority: America's best kept secret.* 2nd ed. Ithica, NY Cornell UP.

Zwick, R., & Himelfarb, I. (2011). The Effect of high school socioeconomic status on the predictive validity of SAT scores and high school grade-point average. *Journal of Educational Measurement*, *48*(2), 101–121.

Tables and Figures

Table 1. Description of variables.

Variable	Source	Measurement
Parental education	SSU Application /FASFA	Binary. 1= BA, 0= no BA
Parental income	SSU Application /FASFA	Log dollars
Gender	SSU Application /FASFA	Binary. 1= female, 0= male
Race	SSU Application /FASFA	White (reference), African American, Hispanic, other
Age	SSU Application /FASFA	Continuous
Student income	SSU Application /FASFA	Log dollars
Parental marital status	SSU Application /FASFA	Single, married, divorced, widowed
Household size	SSU Application /FASFA	Discrete
# in college	SSU Application /FASFA	Discrete
Cohort year	SSU Application /FASFA	Discrete
High school GPA	SSU Application /FASFA	Continuous 4.0 scale
ACT score (composite)	SSU Application /FASFA	Raw ACT score
In-state	SSU Application /FASFA	Binary. 1= in-state, 0= out-of-state
High school cohort member	SSU data (author compiled)	Binary. 1= cohort member, 0= no cohort member
Unemployment	U.S. Department of Labor. Bureau of Labor Statistics: Local Area Unemployment Statistics (author compiled)	4-year average (percent)
% males in county with BA	U. S. Census Bureau: American Community Survey	Percentage by cohort year
% females in county with BA	U. S. Census Bureau: American Community Survey	Percentage by cohort year
% males in county with high school diploma	U. S. Census Bureau: American Community Survey	Percentage by cohort year
% females in county with high school diplomas	U. S. Census Bureau: American Community Survey	Percentage by cohort year
County population	U. S. Census Bureau: American Community Survey	Discrete (cohort year)
County poverty	U.S. Census Bureau: Small Area Income and Poverty Estimates (author compiled)	4-year average (percent)
County median income	U.S. Census Bureau: Small Area Income and Poverty Estimates (author compiled)	Dollars/1000 (cohort year)
Missing parental income	Imputed	Dollars (log)
Missing student income	Imputed	Dollars (log)
Type of high school	National Center for Education Statistics	Public (reference), private, home school, GED
High school poverty	National Center for Education Statistics	Percentage of students eligible for the Free and Reduced Lunch Program
High school Title 1 eligible	National Center for Education Statistics	Binary. 1= Title 1 eligible, 0 = not Title 1 eligible
Student-teacher ratio	National Center for Education Statistics	Ratio
High school total population	National Center for Education Statistics	Discrete
High school location	National Center for Education Statistics	1= City, Large; 2= City, Midsize; 3= City, Small; 4= Rural, Distant; 5= Rural, Fringe; 6= Rural, Remote; 7=Suburb, Large; 8= Suburb, Midsize; 9= Suburb, Small; 10= Town, Distant; 11= Town, Fringe; 12=Town, Remote
HS % African American	National Center for Education Statistics	Percentage of students who are not white, Hispanic, or other
HS % Hispanic	National Center for Education Statistics	Percentage of students who are not white, African American, or other
HS % other	National Center for Education Statistics	Percentage of students who are not white, African American, or Hispanic

Note: **SSU**- Sunflower State University

Table 2. Select descriptive characteristics of the 2007-2014 first-time, full-time freshman cohorts.

	N=4,433	Graduated	Dropped	Total
Gender				
Female	2,789	43.7%	56.3%	100.0%
Male	1,644	31.3%	68.7%	100.0%
Race				
White	3484	42.2%	57.8%	100.0%
African American	263	25.9%	74.1%	100.0%
Hispanic	307	26.4%	73.6%	100.0%
Other	379	30.1%	69.9%	100.0%
Family measures				
Parental education				
BA+	2,695	43.0%	57.0%	100.0%
No BA	1,738	33.1%	66.9%	100.0%
Parental income				
Below 1 s.d. of average	328	18.6%	81.4%	100.0%
-1 s.d. to mean	2327	36.2%	63.8%	100.0%
Mean to +1 s.d.	1746	46.7%	53.3%	100.0%
Above 1 s.d.	32	50.0%	50.0%	100.0%
Academic performance				
Composite ACT				
Continuing-generation		23.25 (3.84)	21.89 (3.89)	22.48 (3.92)
First-generation		22.35 (3.46)	20.39 (3.71)	21.07 (3.74)
High school GPA				
Continuing-generation		3.49 (.64)	3.11 (.74)	3.27 (.72)
First-generation		3.42 (.62)	2.96 (.70)	3.12 (.71)

Table 3. Linear probability models for dropping out.

	Model 1		Model 2		Model 3	
	Coeff	**(se) sig**	**Coeff**	**(se) sig**	**Coeff**	**(se) sig**
Parent with a BA	-.057	(.015)***	-.056	(.016)***	-.043	(.015)**
Parental income (log)	-.048	(.012)***	-.043	(.012)***	-.029	(.012)*
Neighborhood characteristics						
County population			.003	(.009)	.002	(.009)
Unemployment			-.019	(.007)	-.019	(.007)**
County poverty			.002	(.004)	.002	(.004)
County median income			-.001	(.002)	-.001	(.002)
High school poverty			.066	(.067)	.068	(.067)
% males in county with BA			-.183	(.207)	-.060	(.204)
% females in county with BA			.071	(.214)	-.002	(.210)
% males in county with hs diploma			-.031	(.200)	-.494	(.205)*
% females in county with hs diploma			.049	(.302)	.750	(.311)*
High school characteristics						
High school Title 1 eligible			.016	(.022)	.010	(.022)
HS % African American			-.142	(.106)	-.170	(.103)
HS % Hispanic			-.028	(.076)	-.021	(.074)
HS % other			.373	(.175)	.339	(.166)*
Academic performance						
High school GPA					-.096	(.009)***
ACT score (composite)					-.009	(.001)***
Demographics						
Female	-.109	(.015)***	-.106	(.015)***	-.085	(.015)***
Race						
African American	.073	(.031)*	.086	(.036)*	.020	(.036)
Hispanic	.096	(.027)***	.098	(.029)***	.068	(.029)*
Other	.079	(.025)**	.074	(.026)**	.051	(.025)*
Constant	.393	(.234)	.518	(.253)	1.151	(.260)***
R^2	.067		.071		.115	

Notes: Attending college in state, having someone from the same high school graduating class attend in the same college cohort, cohort year, age, age^2, parental marital status, household size, number of students in the household in college, student income, variables controlling for missing parental and student income, high school type, high school total population, high school location, high school student-teacher ratio are also included in the models but are not shown here.

*≤.05, **≤.01, ***≤.001

Table 4. Linear probability models by drop out timing.

	Model 1 Dropout after years 1 or 2 **N=4,433**		Model 2 Dropout after years 3 or 4 **N=2,657**		Model 3 Dropout after year 5 **N=2,041**	
	Coeff	**(se) sig**	**Coeff**	**(se) sig**	**Coeff**	**(se) sig**
Parent with a BA	-.020	(.016)	-.041	(.018)*	-.029	(.017)
Parental income (log)	-.024	(.012)	-.008	(.014)	-.016	(.015)
Neighborhood characteristics						
County population	.005	(.009)	.012	(.009)	-.010	(.008)
Unemployment	.007	(.007)	-.050	(.008)***	.002	(.007)
County poverty	-.007	(.004)	.006	(.004)	.002	(.004)
County median income	-.004	(.002)	.002	(.002)	.000	(.002)
High school poverty	.082	(.071)	.003	(.082)	.052	(.082)
% males in county with BA	.137	(.211)	-.093	(.222)	-.160	(.196)
% females in county with BA	-.043	(.217)	.041	(.231)	.084	(.201)
% males in county with hs diploma	-.588	(.207)**	-.116	(.199)	-.288	(.173)
% females in county with hs diploma	.890	(.313)**	.178	(.302)	.440	(.262)
High school characteristics						
High school Title 1 eligible	-.015	(.023)	.044	(.026)	-.026	(.022)
HS % African American	-.135	(.108)	.044	(.122)	-.245	(.111)*
HS % Hispanic	-.016	(.077)	.005	(.088)	.021	(.085)
HS % other	.402	(.170)*	.293	(.221)	-.017	(.197)
Academic performance						
High school GPA	-.069	(.009)***	-.040	(.009)***	-.071	(.010)***
ACT score (composite)	-.008	(.001)***	-.005	(.002)**	-.004	(.002)*
Demographics						
Female	-.031	.(015)*	-.059	(.017)***	-.071	(.018)***
Race						
African American	-.052	(.040)	.039	(.047)	.061	(.054)
Hispanic	-.045	(.032)	.020	(.037)	.166	(.045)***
Other	-.009	.(027)	.036	(.032)	.084	(.035)*
Constant	.853	(.302)	.062	(.388)	.319	(.440)
R^2	.067		.107		.128	

Notes: Attending college in state, having someone from the same high school graduating class attend in the same college cohort, cohort year, age, age^2, parental marital status, household size, number of students in the household in college, student income, variables controlling for missing parental and student income, high school type, high school total population, high school location, high school student-teacher ratio are also included in the models but are not shown here.

*≤.05, **≤.01, ***≤.001

Table 5. Kaplan-Meier coefficients of the estimated survivor functions at select years.

	1 Year	2 Years	4 Years	6 Years
Parental education				
No college	.698	.560	.406	.232
One college	.747	.622	.475	.288
Both college	.790	.667	.509	.295
Parental income by s.d.				
Below 1 s.d.	.667	.667	.167	----
- 1s.d. to mean	.710	.581	.427	.249
To 1 s.d. above	.785	.657	.511	.327
> + 1s.d.	.656	.625	.515	----
Gender				
Female	.748	.626	.487	.298
Male	.710	.563	.403	.232
Race				
White	.738	.609	.471	.302
African American	.697	.531	.337	.160
Hispanic	.719	.602	.422	.234
Other	.723	.585	.415	.196

Figure 1. Kaplan-Meier curve for all sample members by parental education status.

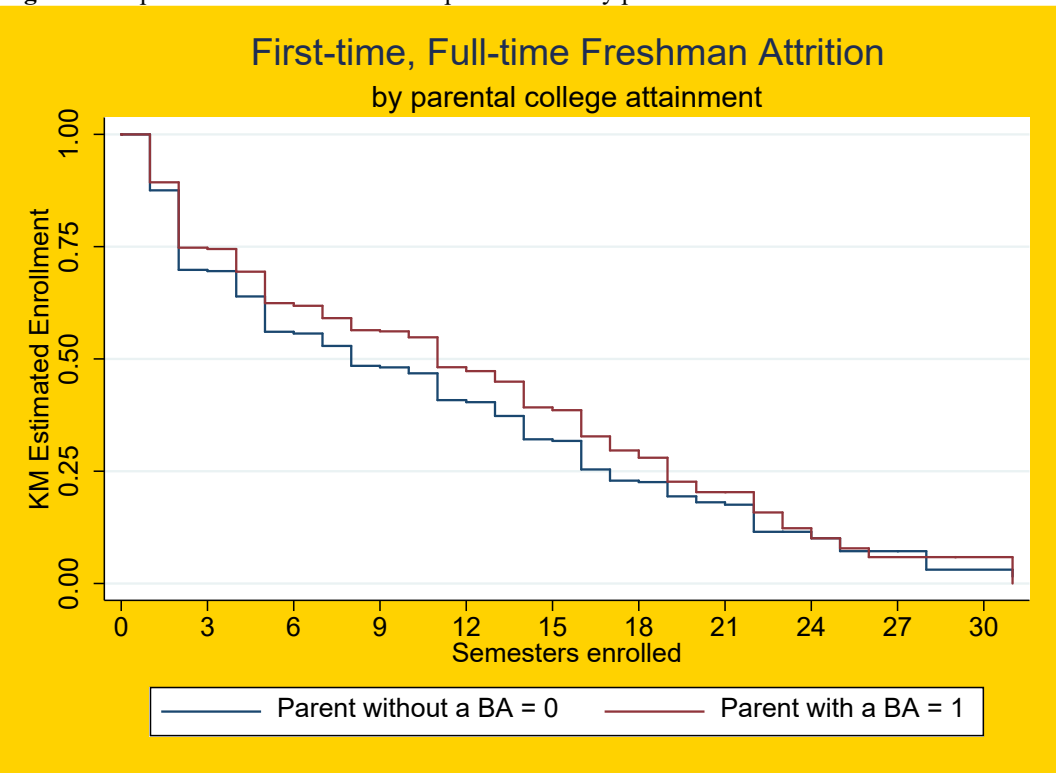
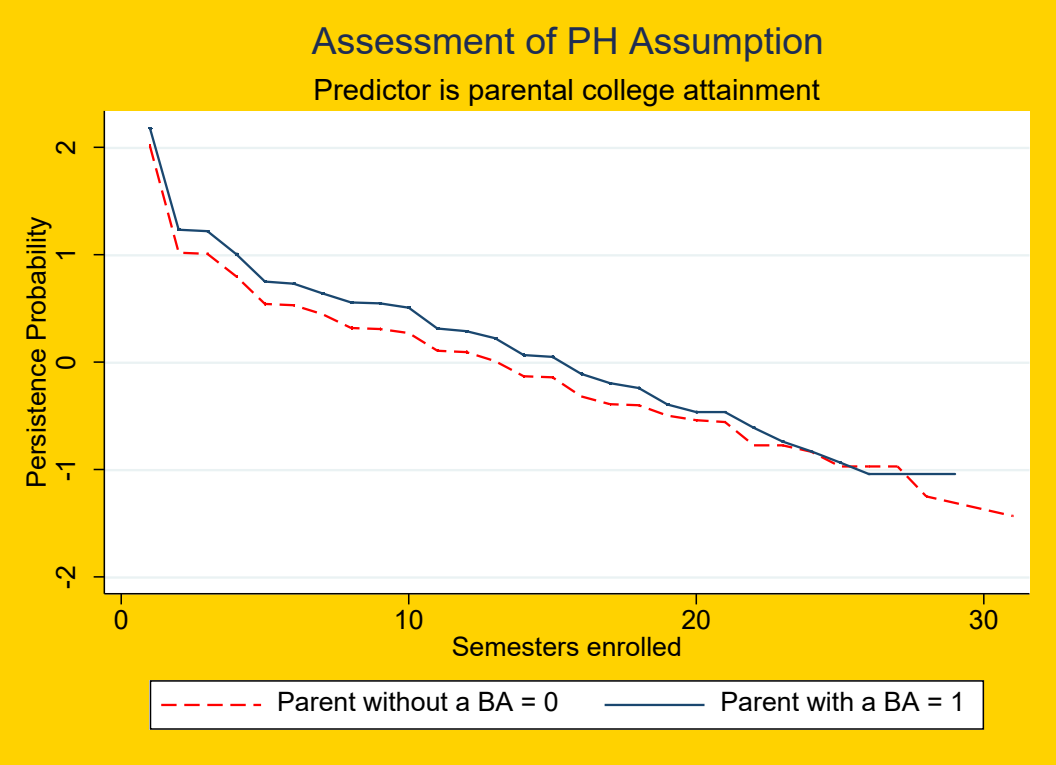


Figure 2. Proportional Hazard assumption by parental education status.



Note: Adjusted for female and race

Table 6. Univariate and multivariate proportional hazard estimates.

	Model 1 Univariate Coeff (se) sig		Model 2 Multivariate Stratified by cohort year Coeff (se) sig	
Parental education				
No college	1.195	(.047)***		
One college	.836	(.033)***	.899	(.036)**
Both college	.781	(.035)***		
Parental income	.985	(.004)***	.899	(.036)
Neighborhood effects				
Unemployment	1.016	(.015)	.899	(.036)
Poverty average	1.020	(.004) ***	.899	(.036)
County median income	.996	(.002)**	.899	(.036)
High school poverty	1.639	(.179)***	1.302	(.232)
% males in county with hs diploma	1.002	(.002)	.194	(.113)**
% females in county with hs diploma	1.002	(.003)	11.986	(10.551)**
Demographics				
Female	.811	(.030)***	.871	(.033)***
Male	1.233	(.045)***	--	--
Race				
White	.834	(.034)***	--	--
African American	1.301	(.095)***	.958	(.083)
Hispanic	1.088	(.074)	.961	(.072)
Other	1.144	(.063)*	1.066	(.065)
Parental marital status				
Married	.829	(.036)***	.905	(.116)
Divorced	1.146	(.057)**	.923	(.118)
Single	1.348	(.099)***	.965	(.129)
Widowed	.826	(.139)	.699	(.129)

Notes: Attending college in state, having someone from the same high school graduating class attend in the same college cohort, cohort year, age, age^2, parental marital status, household size, number of students in the household in college, student income, variables controlling for missing parental and student income, county population, living in-state, high school type, high school Title 1 eligibility, high school total population, high school location, high school student-teacher ratio, high school racial percentages (African American, Hispanic, other), high school GPA, and ACT scores (composite) are also included in the models but are not shown here. Estimates in parenthesis indicate standard deviation.

*≤.05, **≤.01, ***≤.001

Appendix

Appendix: Table 1. Additional select descriptive characteristics of the 2007-2014 first-time, full-time freshman cohorts.

	N=4,433	Graduated	Dropped
Gender			
Female	2,789	70.3%	58.2%
Male	1,644	29.7%	41.8%
	Total	*100.0%*	*100.0%*
Race			
White	3484	84.8%	74.6%
African American	263	3.9%	7.2%
Hispanic	307	4.7%	8.4%
Other	379	6.6%	9.8%
	Total	*100.0%*	*100.0%*
Family measures			
Parental education			
BA+	2,695	66.8%	57%
No BA	1,738	33.2%	43%
	Total	*100.0%*	*100.0%*
Parental income			
Below 1 s.d. of average	328	3.5%	9.9%
-1 s.d. to mean	2327	48.6%	55.0%
Mean to +1 s.d.	1746	47.0%	34.5%
Above 1 s.d.	32	0.9%	0.6%
	Total	*100.0%*	*100.0%*

Appendix: Table 2. Linear probability models for dropping out.

	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coeff	(se) sig	Coeff	(se) sig	Coeff	(se) sig	Coeff	(se) sig	Coeff	(se) sig
Parent with a BA	-.057	(.015)***	-.056	(.016)***	-.043	(.015)**	-.058	(.017)***	-.057	(.017)***
Parental income (log)	-.048	(.012)***	-.043	(.012)***	-.029	(.012)*	-.029	(.012)*	-.029	(.012)*
Neighborhood characteristics										
County population			.003	(.009)	.002	(.009)	.001	(.009)	.002	(.009)
Unemployment			-.019	(.007)	-.019	(.007)**	-.018	(.007)**	-.019	(.007)**
County poverty			.002	(.004)	.002	(.004)	.002	(.004)	.002	(.004)
County median income			-.001	(.002)	-.001	(.002)	-.001	(.002)	-.001	(.002)
High school poverty			.066	(.067)	.068	(.067)	.061	(.067)	.063	(.067)
% males in county BA			-.183	(.207)	-.060	(.204)	-.076	(.204)	-.066	(.204)
% females in county with BA			.071	(.214)	-.002	(.210)	.020	(.210)	.005	(.210)
% males in county with hs diploma			-.031	(.200)	-.494	(.205)*	-.484	(.205)*	-.478	(.205)*
% females in county with hs diploma			.049	(.302)	.750	(.311)*	.735	(.311)*	.726	(.310)*
High school characteristics										
Type of high school			.006	(.020)	.001	(.020)	.000	(.020)	-.001	(.020)
High school Title 1 eligible			.016	(.022)	.010	(.022)	.011	(.022)	.010	(.022)
Student-teacher ratio			-.001	(.002)	-.001	(.001)	-.001	(.001)	-.001	(.001)
High school total population			.000	(.000)	.000	(.000)	.000	(.000)	.000	(.000)
High school location			-.003	(.002)	-.001	(.002)	-.001	(.002)	-.001	(.002)
HS % African American			-.142	(.106)	-.170	(.103)	-.166	(.103)	-.170	(.102)
HS % Hispanic			-.028	(.076)	-.021	(.074)	-.006	(.075)	-.005	(.075)
HS % other			.373	(.175)	.339	(.166)*	.333	(.166)*	.319	(.168)
College variables										
Living in state			-.062	(.030)	-.014	(.030)	-.015	(.030)	-.022	(.030)
High school cohort member			.036	(.022)	.017	(.022)	.018	(.022)	.017	(.022)
Academic performance										
High school GPA					-.096	(.009)***	-.095	(.009)***	-.095	(.009)***
ACT score (composite)					-.009	(.001)***	-.009	(.001)***	-.009	(.001)***
Demographics										
Female	-.109	(.015)***	-.106	(.015)***	-.085	(.015)***	-.085	(.015)***	-.104	(.017)***
Cohort year	.025	(.004)***	.023	(.004)***	.030	(.004)***	.030	(.004)***	.030	(.004)***
Age	.058	(.009)***	.050	(.010)***	.016	(.010)	.016	(.010)	.015	(.010)
Age2	.000	(.000)***	.000	(.000)**	.000	(.000)**	.000	(.000)**	.000	(.000)*
Student income (log)	-.018	(.013)	-.017	(.013)	-.015	(.012)	-.015	(.012)	-.015	(.012)
Missing parental income	-.459	(.128)***	-.404	(.0129)**	-.259	(.127)*	-.258	(.128)*	-.262	(.128)*
Missing student income	-.118	(.104)	-.114	(.104)	-.102	(.102)	-.101	(.102)	-.099	(.102)
Household size	-.040	(.072)	-.021	(.073)	-.034	(.074)	-.035	(.074)	-.040	(.074)
# in college	.011	(.016)	.008	(.016)	.013	(.016)	.013	(.016)	.014	(.016)
Race										
African American	.073	(.031)*	.086	(.036)*	.020	(.036)	-.011	(.048)	-.059	(.054)
Hispanic	.096	(.027)***	.098	(.029)***	.068	(.029)*	.011	(.034)	-.040	(.046)
Other	.079	(.025)**	.074	(.026)**	.051	(.025)*	.041	(.034)	-.016	(.045)

Parental marital status										
Married	-.090	(.051)	-.091	(.051)	-.030	(.052)	-.029	(.052)	-.028	(.052)
Divorced	-.079	(.051)	-.075	(.051)	-.029	(.051)	-.030	(.051)	-.030	(.051)
Single	-.046	(.053)	-.042	(.053)	-.009	(.053)	-.008	(.053)	-.010	(.053)
Widowed	-.171	(.076)*	-.164	(.076)	-.119	(.071)	-.120	(.071)	-.120	(.071)
Interaction terms										
Parental education interaction with race										
BA* African American							.048	(.056)	.051	(.056)
BA* Hispanic							.195	(.055)***	.191	(.055)***
BA * Other							.014	(.048)	.016	(.048)
Female interaction with race										
Female* African American									.093	(.058)
Female * Hispanic									.086	(.052)
Female * Other									.091	(.049)
Constant	.393	(.234)	.518	(.253)	1.151	(.260)***	1.162	(.260)	1.198	(.260)***
R^2	.067		.071		.115		.117		.118	

Note: *≤.05, **≤.01, ***≤.001

Appendix: Table 3 Linear probability models by drop out timing.

	Model 1 Dropout after years 1 or 2 **N=4,433**		Model 2 Dropout after years 3 or 4 **N=2,657**		Model 3 Dropout after year 5 **N=2,041**	
	Coeff	**(se) sig**	**Coeff**	**(se) sig**	**Coeff**	**(se) sig**
Parent with a BA	-.020	(.016)	-.041	(.018)*	-.029	(.017)
Parental income (log)	-.024	(.012)	-.008	(.014)	-.016	(.015)
Neighborhood characteristics						
County population	.005	(.009)	.012	(.009)	-.010	(.008)
Unemployment	.007	(.007)	-.050	(.008)***	.002	(.007)
County poverty	-.007	(.004)	.006	(.004)	.002	(.004)
County median income	-.004	(.002)	.002	(.002)	.000	(.002)
High school poverty	.082	(.071)	.003	(.082)	.052	(.082)
% males in county with BA	.137	(.211)	-.093	(.222)	-.160	(.196)
% females in county with BA	-.043	(.217)	.041	(.231)	.084	(.201)
% males in county with hs diploma	-.588	(.207)**	-.116	(.199)	-.288	(.173)
% females in county with hs diploma	.890	(.313)**	.178	(.302)	.440	(.262)
High school characteristics						
Type of high school	.015	(.024)	-.035	(.034)	-.002	(.037)
High school Title 1 eligible	-.015	(.023)	.044	(.026)	-.026	(.022)
Student-teacher ratio	-.001	(.002)	.000	(.002)	-.002	(.002)
High school total population	.000	(.000)	.000	(.000)	.000	(.000)
High school location	-.003	(.003)	.002	(.003)	-.001	(.003)
HS % African American	-.135	(.108)	.044	(.122)	-.245	(.111)*
HS % Hispanic	-.016	(.077)	.005	(.088)	.021	(.085)
HS % other	.402	(.170)*	.293	(.221)	-.017	(.197)
College variables						
Living in state	-.021	(.032)	-.035	(.034)	.017	(.035)
High school cohort member	.010	(.023)	.043	(.024)	-.030	(.022)
Academic performance						
High school GPA	-.069	(.009)***	-.040	(.009)***	-.071	(.010)***
ACT score (composite)	-.008	(.001)***	-.005	(.002)**	-.004	(.002)*
Demographics						
Female	-.031	(.015)*	-.059	(.017)***	-.071	(.018)***
Cohort year	.010	(.004)*	.048	(.005)***	.003	(.005)
Age	.012	(.012)	.028	(.017)	.027	(.019)
Age2	.000	(.000)	.000	(.000)	.000	(.000)
Student income (log)	-.009	(.013)	-.020	(.013)	-.003	(.013)
Missing parental income	-.200	(.138)	-.020	(.165)	-.168	(.170)
Missing student income	-.085	(.104)	-.141	(.106)	.012	(.106)
Household size	-.009	(.082)	-.042	(.091)	-.012	(.107)
# in college	.029	(.016)	-.008	(.017)	.004	(.016)
Race						
African American	-.052	(.040)	.039	(.047)	.061	(.054)
Hispanic	-.045	(.032)	.020	(.037)	.166	(.045)***
Other	-.009	(.027)	.036	(.032)	.084	(.035)*
Parental marital status						
Married	-.063	(.062)	.057	(.084)	-.085	(.100)
Divorced	-.041	(.062)	.062	(.085)	-.130	(.098)
Single	-.056	(.063)	.081	(.088)	-.043	(.104)
Widowed	-.179	(.078)*	-.038	(.092)	-.033	(.106)
Constant	.853	(.302)	.062	(.388)	.319	(.440)
R^2	.067		.107		.128	

Note: *≤.05, **≤.01, ***≤.001

Appendix: **Table 4**. Univariate and multivariate proportional hazard estimates.

	Model 1 Univariate Coeff (se) sig		**Model 2 Multivariate Stratified by cohort year Coeff (se) sig**	
Parental education				
No college	1.195	(.047)***		
One college	.836	(.033)***	.899	(.036)**
Both college	.781	(.035)***		
Parental income	.985	(.004)***	.899	(.036)
Neighborhood effects				
County population			1.010	(.022)
Unemployment	1.016	(.015)	.899	(.036)
Poverty average	1.020	(.004) ***	.899	(.036)
County median income	.996	(.002)**	.899	(.036)
High school poverty	1.639	(.179)***	1.302	(.232)
% males in county with BA			1.356	(.755)
% females in county with BA			.908	(.512)
% males in county with hs diploma	1.002	(.002)	.194	(.113)**
% females in county with hs diploma	1.002	(.003)	11.986	(10.551)**
High school characteristics				
Type of high school			1.054	(.060)
High school Title 1 eligible			.986	(.057)
Student-teacher ratio			.999	(.005)
High school total population			1.000	(.000)
High school location			.997	(.006)
HS % African American			.692	(.183)
HS % Hispanic			.895	(.174)
HS % other			2.568	(.911)**
College variables				
Living in state			.920	(.071)
High school cohort member			1.054	(.060)
Academic performance				
High school GPA			.827	(.017)
ACT score (composite)			.979	(.004)
Demographics				
Female	.811	(.030)***	.871	(.033)***
Male	1.233	(.045)***	--	--
Cohort year			1.110	(.012)***
Age			1.081	(.030)**
Age2			.999	(.000)**
Student income (log)			.958	(.032)
Missing parental income			.670	(.220)
Missing student income			.726	(.201)
Household size			.961	(.192)
# in college			1.059	(.043)
Race				
White	.834	(.034)***	--	--
African American	1.301	(.095)***	.958	(.083)
Hispanic	1.088	(.074)	.961	(.072)
Other	1.144	(.063)*	1.066	(.065)
Parental marital status				
Married	.829	(.036)***	.905	(.116)
Divorced	1.146	(.057)**	.923	(.118)
Single	1.348	(.099)***	.965	(.129)
Widowed	.826	(.139)	.699	(.129)