2017

Maintaining College Access in a Post Recession Era: A Multi-Level Competing Risks Model

Brendan Martin
Loyola University Chicago

LOYOLA UNIVERSITY CHICAGO

MAINTAINING COLLEGE ACCESS IN A POST RECECSSION ERA: A MULTI-LEVEL COMPETING RISKS MODEL

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

PROGRAM IN RESEARCH METHODOLOGY

BY

BRENDAN M. MARTIN

CHICAGO, IL

DECEMBER 2017

ACKNOWLEDGMENTS

To my friends and family who never questioned my unhealthy obsession with school. In particular, I’d like to thank Tim and Katie for their initial encouragement, and Bill for his good faith efforts to keep me going 8 weeks into every semester. Finally, this wouldn’t have been possible without the unwavering support, patience, and sharp insight of my wife Alli and sister Melissa.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Post-Great Recession budgets cuts and funding freezes have decreased the level of institutional resources available to recruit and retain undergraduate students. To optimize remaining expenditures in this challenging climate, new analytical approaches must be considered to evaluate and interpret pre-enrollment student data. To date, much of the higher education literature has focused on predicting enrollment using traditional fixed or mixed effects binary logistic models. While robust, these modeling approaches are constrained by standard statistical assumptions, do not account for the timing of students' enrollment decisions, and cannot efficiently incorporate censored data points or competitor information. This study applies a multi-level, competing risks model to the analysis of undergraduate application data to assess time to enrollment as a function of univariable and multivariable sociodemographic, institutional, financial, and academic factors. There are both methodological and practical strengths to the analytic approach. Conceptually, the mixed effects model applied to this sample appropriately accounts for student clustering, thereby incorporating similarities in applicants' academic preparation and backgrounds. Further, the competing risks design allows data on select competitors to enter the model, offering the opportunity to evaluate multiple institutions side-by-side.

In practice, the study uncovered differential effects across the competitive set for every sociodemographic, institutional, financial, and academic factor under review, with the exception of first choice status. The institutional and policy implications associated with these divergent

results range from a reduction in undergraduate recruitment expenditures to continued investment in student support services leading to stronger retention, higher graduation rates, and lower cohort default rates (debt delinquency). Reducing recruitment overhead will not only free up important capital to reinvest in vital student support services, including first year programming, but it will also enable administrators to maintain a focus on important post-enrollment metrics. This modeling approach provides unique insights into not only students' final decisions, but also their timelines for making those decisions. Consideration of model results within the undergraduate recruitment process will help to alleviate some of the initial budget constraints by identifying how and when certain known factors increase the probability of student enrollment, while not sacrificing on other important postsecondary measures, such as retention and graduation.

CHAPTER ONE

INTRODUCTION

The pervasive and enduring gaps in educational opportunities across traditional racial, ethnic, socioeconomic, and gender divides remain important topics of discourse (Kumar & Hurwitz, 2015). Practitioners and policymakers have offered a variety of explanations for these persistent inequities, but the fact remains that postsecondary institutions at all levels (two- and four-year public and private colleges alike) are confronting an increasingly difficult environment in their efforts to attract, retain, and graduate a diverse and qualified undergraduate student body (Harvill et al., 2012). While high school graduation rates have increased since early 2000, approximately 30% of students still do not graduate high school and only 70% of those that do enroll at a postsecondary institution (Snyder & Dillow, 2010). In addition, despite some recent incremental increases in college enrollment, minorities, lower socioeconomic status (SES) high school graduates, and first generation students (roughly 1 in 3 college-bound students) are still considerably less likely than their peers to graduate high school and pursue postsecondary education (College Board, 2010; Education Advisory Board, 2016).

These enrollment trends among traditionally under-represented groups have been further exacerbated by post-recession spending cuts and funding freezes at colleges and universities nationwide. According to the Center on Budget of Policy Priorities, a nonpartisan research and policy institute, 47 states spent less per student during the 2014-15 school year than they did at the start of the recession (Mitchell & Leachman, 2015). During that same period the cost of student recruitment increased to an all-time high. In 2015, the median cost of student recruitment at four-year private and public universities was $2,232 and $578, respectively, an increase of 15% and 45% compared to 2007 costs (Ruffalo Noel Levitz, 2016, 2009). Increasing recruitment expenditures and tuition, coupled with recent spending cuts at many public and private universities threaten to diminish student access and negatively impact a wide range of postsecondary outcomes (Fitzgerald, 2004; St. John et al., 2003). Considering these challenges, it is incumbent upon admissions staff to apportion resources to identify and recruit applicants to maximize the fit between student and institution.

Enrollment Modeling

Strategic allocation of limited recruitment budgets is, in part, informed by the collection and analysis of self-reported family and individual student data. This information is often provided throughout the recruitment, application, and financial aid processes. Predictive modeling, typically in the form of logistic regression models, is a frequently utilized technique to analyze these data to identify students with high probabilities of enrollment. Such analyses enable admissions and enrollment management staff the opportunity to target their finite resources, thereby allowing institutions to pursue multiple, sometimes competing ends (e.g. achieving baseline enrollment targets, diversification, attracting high achieving students, etc.). Often, the selection of independent variables in these predictive models depend on a combination of theoretical and practical considerations (Thomas et al., 1999). Common metrics include measures of academic achievement, financial aid, SES, first generation status, indicators of early engagement, minority status, residential status, sex, intended major, and high school context.

While these approaches are informative, additional modeling techniques are available that may provide further insight into important aspects of students' decision-making process. For example, the application of time to event models within the context of higher education offers a unique opportunity to evaluate traditional independent variables while accounting for the time dependent nature of the application cycle itself. Since the late 1990s, time to event models, or event history models as they are known in education, have been used to examine select factors that affect students' post-enrollment outcomes, such as persistence, dropout, and completion (Gross & Torres, 2010; Bahr, 2009; Scott & Kennedy, 2005; DesJardins et al., 2002, 1999, 1997; Murtaugh et al., 1999; Singer & Willett, 1991). The extension of such models to focus on undergraduate students' initial decision timelines may provide critical information to the admissions personnel tasked with recruiting them.

Although infrequently applied, the potential benefits of these techniques in the field of higher education are many and clear, especially given policymakers' renewed focus on student outcomes over the past few decades. DesJardins et al. (1999) credited such modeling approaches for helping to develop timely interventions for students at risk of dropping out, while Gross and Torres (2010) used a similar model to examine how the timing of financial aid offers affect educational attainment among minority student populations. In addition, scholars have shown that these models can be seamlessly extended to meet the demands of complex, hierarchical designs (Bahr, 2009) or even adapted to a "competing risks" framework in which the focus rests on multiple, overlapping events, such as stopouts, dropouts, and graduation (Guerin, 1997; DesJardins et al., 1999; Ronco, 1996). Despite these recent innovations, however, enrollment research remains largely limited to more widespread and traditional modeling techniques.

Proposed Analysis

The purpose of this study is to extend the current literature examining the relationship between select student- and school-level factors and undergraduate enrollment, while building on recent applications of time to event models in higher education. Specifically, the analytic approach outlined herein, a multi-level competing risks model, aims to examine what student- and high school-level factors effectively reduce time to enrollment in a competitive higher education marketplace. Since the great recession, postsecondary institutions of all types have been forced to operate in an environment of reduced or constrained budgets and increased expectations regarding student outcomes. This research seeks to determine if the important academic and sociodemographic factors that have been shown to predictor undergraduate enrollment can also effectively inform institutions' efforts to reduce recruitment expenditures by shortening students' decision timelines.

The analytic model developed in this study will build off an extensive literature as to what factors drive undergraduate enrollment, such as measures of student ability (Noel-Levitz, 2012; Avery & Hoxby, 2004; Thomas et al., 1999), financial aid (Harvill et al., 2012; Monks, 2009; Linsenmeirer et al., 2006), early outreach (Wyatt et al., 2014; Perna & Swail, 2002; Swail, 2001;), and select sociodemographic factors (Kumar & Hurwitz, 2015; Conger & Long, 2013; Hussar & Bailey, 2011). In addition, this study will augment emerging research on time to event modeling in the context of higher education, while employing a multi-level design that accounts for important differences among the high school contexts from which applicants emerge. This approach is appropriate as it incorporates variation in the outcomes driven by student clustering within secondary institutions, which play important roles in engendering the social and academic skills vital for college success.

In addition to the student-level insights provided, the model will simultaneously assess the roles of competing actors (e.g. multiple universities) in a crowded postsecondary market. This will enable institutions to directly incorporate data on their institutional peers and aspirant colleges, which will greatly inform on their enrollment management strategies. By accounting for the activities of other universities, admissions personnel can more effectively target and recruit prospective high school students, as well as accurately project freshmen enrollment. This will help to avoid unexpected budgetary shortfalls that could negatively impact future admissions and student services funding.

Study Significance

The proposed modeling techniques will also inform multiple financial and policy considerations for academic institutions. First and foremost, resource conservation across the post-recession higher education landscape will allow institutions to free up important capital by minimizing recruitment overhead. This will enable administrators to re-invest in vital student support services, first-year student programming, and other retention efforts. Second, more targeted efforts to shorten the decision timeline among a smaller pool of well-qualified and strongly matched applicants may enable institutions to cover more of the initial costs associated with the search process, such as campus visits and admitted student events. This will save admissions staff time and money in the long-run, but will also help to alleviate some of the initial cost constraints faced by otherwise qualified students and their families. Finally, early efforts to maximize student/institution match will also likely pay long-term financial dividends for students, institutions, and the Department of Education in the form of stronger retention and lower cohort default rates.

While more common in the medical literature, the proposed multi-level competing risks analysis will contribute valuable insight into the often opaque process of enrollment management. More importantly, though, it will also help identify those student and institutional factors that are key to reducing time to enrollment. The following analyses will first provide an overview of the factors that historically drive college enrollment to provide an analytic framework for the model building process. It will then outline the literature related to time to event modeling in higher education, as well as address how the data used were collected, aggregated, and de-identified. Details on the necessary methodological steps to fit an appropriate model will also be discussed. The variables employed in the univariable and multivariable models will then be defined and the characteristics of the sample described. Following the interpretation of the results, a discussion section will articulate the possible limitations and implications of these findings.

Research Questions

The research questions that will guide this analysis include:

1. Do certain sociodemographic, institutional, financial, and academic factors effectively reduce time to postsecondary enrollment?
2. Do these sociodemographic, institutional, financial, and academic factors have common effects across several similar-profile postsecondary institutions?

3. How can the magnitudes and direction of the effects select sociodemographic, institutional, financial, and academic factors have on time to enrollment at different postsecondary institutions inform an enrollment management strategy?

CHAPTER TWO

LITERATURE REVIEW

Despite more recent applications of time to event models in other educational contexts, binary logistic regression remains a standard approach to modeling enrollment data. Traditional admissions-based models often utilize student- and institution-level data from prior recruitment cycles as fixed effects to predict individual behavior (Conger & Long, 2013; Goenner & Pauls, 2006; Thomas et al., 1999; Brugglink & Gambhir, 1996). Alternatively, rational choice models focus on financial incentives, while controlling for broader macro-economic trends to identify those factors that drive enrollment (Monks, 2009; Ledesma, 2009; Linsenmeirer et al., 2006; DesJardins & Toutkoushian, 2005; Avery & Hoxby, 2004; Long, 2004; Stater, 2004; Singell, 2002; Paulsen & St. John, 2002; Paulsen, 2001 & 1990). In addition, hierarchical designs utilizing random effects to model variance at the high school level are becoming increasingly more common. (Engberg & Wolniak, 2009; Hill, 2008; Johnson, 2008; Cho, 2006; Khattab, 2005). The findings from these studies have provided profound insight into the links between the high school contexts from students emerge and their level of preparation.

Overall, binary logistic regression models provide reliable enrollment probability estimates and thereby a great deal of analytic insight into students' college choice process. Thus, these techniques have traditionally formed the basis for admissions staffs' efforts to segment their prospective student population and thereby better target their recruitment efforts (DesJardins & Lindsay, 2008; Johnson, 2008; Goenner & Pauls, 2006; DesJardins, 2002; Thomas et al., 2001). It is common for variable selection in such models to be driven by the academic and personal profile of the student sample; however, it is important that these metrics also reflect how institutional characteristics adhere to a student's preferences (Goenner & Pauls, 2006). Therefore, the proposed framework will focus on those student characteristics that are known to inform direct matriculation.

Academic Factors

Students' academic achievement, as measured by a combination of standardized test scores, advanced placement coursework, and GPA, is meaningfully associated with an array of important postsecondary measures (NCES, 2015; Ledesma, 2009; Adelman, 2006; Chang, 2006; Brugglink & Gambhir, 1996; Thomas et al., 1979). Numerous empirical studies have shown that students with a record of strong academic performance consistently outperform their lower achieving peers in terms of college enrollment rates. For example, DesJardins et al. (2002) note that students perceived academic ability typically lowers their educational costs by increasing demand from various institutional actors. Similarly, Ledesma (2009) showed that high achieving applicants tend to apply to and gain admission at multiple colleges and universities. Further, additional evidence suggests that academic achievement is not only an indicator of how well prepared students are for the rigors of postsecondary education, but also their initial college choice (NCES, 2015; Chang, 2006; Brugglink & Gambhir, 1996). Over time, a student's academic background has been found to have an even more pronounced effect on college enrollment than demographic variables such as sex, race, ethnicity, family composition, and SES (Adelman, 2006; Thomas et al., 1979).

Financial Factors

In addition to indicators of academic achievement, need-based and merit aid play significant roles in students' enrollment decisions. Offers of financial aid to admitted high school seniors often serve two purposes; to "relieve liquidity constraints" that may have undue influence on students' decision-making process and to alter students' "preference rankings" (Avery & Hoxby, 2004; DesJardins et al., 2002). Financial support typically takes many forms, including institutional scholarships; federal loans, grants, and work study; merit and need-based aid offered by external third-party lenders; and a bevy of private financing options. There are also many competing objectives postsecondary institutions consider when 'packaging' students' final financial aid offers. These can include meeting baseline enrollment goals, apportioning seats in the freshman class to address pre-specified diversity targets, or maintaining ties with particular high school networks.

Regardless of these goals, many studies have found that it's the timing, amount, and types of financial aid that ultimately affect students' college choice the most (Harvill et al., 2012; Monks, 2009; Linsenmeirer et al., 2006; DesJardins & Toutkoushian, 2005; Avery & Hoxby, 2004; Long, 2004; Stater, 2004; Singell, 2002; Thomas et al., 1999; Becker, 1993). Research has shown that students typically respond in a rational manner to financial incentives, with earlier aid offers, larger awards, and merit-based assistance tending to increase the probability of postsecondary enrollment. So strong are these causal links that evidence suggests students sometimes respond irrationally to award types. For example, Avery and Hoxby (2004) showed that students' enrollment probabilities increase at a greater rate when grants are re-positioned as 'named' scholarships, which suggests applicants can be persuaded with what are essentially "marketing gimmicks."

Another important development in the area of financial aid was the recent executive action by the Obama administration enabling students and their families to report income two years prior to their Free Application for Federal Student Aid (FAFSA) submission (Department of Education, 2015). This shift had immediate and far-reaching implications for financial aid departments and enrollment management staff. As a result of these changes, financial data are now available to university administrators earlier in the process, enabling them to estimate the potential impact of differential financial aid packaging directly in their enrollment models. In addition to easing the reporting burden on students and their families, this policy change has the potential to help postsecondary institutions provide earlier financial aid offers, adjust their communications flow, and more accurately track progress toward established enrollment goals. Nevertheless, patterns of college enrollment cannot be explained solely through a simple cost-benefit analysis.

Sociodemographic Factors

Related to financial aid, students' socioeconomic status has also been shown to have an undue influence on their college choice. Lower SES high school graduates face many impediments, or "cumulative disadvantages," to accessing higher education (Schultz & Mueller, 2006). These can include, but are often not limited to, a lack of access to information and resource networks, inequality of neighborhood resources, lack of peer/parental support for academic achievement, and ineffective high school counseling (Lin, 2011; Gándara & Bial, 2001). Consequently, these students typically record lower GPA and standardized test scores, as well as apply to relatively fewer colleges, resulting in below average postsecondary enrollment rates (Smith, 2011; Goyette, 2008). Individual students typically have varying levels of academic preparation and access to the resources necessary for success in higher education (Harvill et al., 2012). On this basis, Perna (2006) argues for a conceptual model of college enrollment that integrates aspects of students' cultural and sociological contexts.

Beyond financial resources and academic ability, the sociological factors that animate students' decision-making process often include social and cultural capital, such as parental education-level. Georgetown University's Center on Education and the Workforce (CEW) reports that approximately 1 in 3 (32%) undergraduate students in the United States is categorized as first generation (Carnevale & Strohl, 2013). First generation status is often an important indicator of postsecondary enrollment, as students whose parents have gone to college are often significantly more likely to attend college themselves (Lin, 2011; Goyette, 2008; Warburton & Nunez, 2001). An extensive literature also exists on the roles students' race and ethnicity play in their postsecondary enrollment decisions revealing, for example, a strong link between minority status and differential postsecondary enrollment patterns (Wyatt et al., 2014; Lin, 2011; Bush, 2009; Goenner & Pauls, 2006; Zarate & Gallimore, 2005; DesJardins et al., 2002). Coupled with projected demographic shifts within the broader U.S. population, early indicators of how the makeup of the higher education landscape is changing are beginning to emerge (Colby & Ortman, 2015).

Pew research shows that Hispanic and African American students have accounted for the largest gains in college enrollment over much of the past two decades (Krogstad & Fry, 2014).

For example, as of 2012, the college-going rate among 18- to 24-year old Hispanic high school graduates surpassed that of their Caucasian counterparts for the first time, by a margin of 49% to 47% (Krogstad & Fry, 2014). An enrollment gap remains, however, in part driven by lower rates of four-year college enrollment, as well as lower attendance at selective colleges. In addressing this lag, research has shown significant overlap between students' ethnicity, SES, academic achievement, language proficiency, and other important factors that often contribute to college readiness and the differential in students' postsecondary performance (Wyatt et al., 2014; Zarate & Gallimore, 2005). This complex interplay suggests a more tailored and nuanced approach to high school student recruitment may benefit institutions that currently struggle to attract and retain minority students.

Studies have also shown that there is often significant overlap among the factors that contribute to students' decisions to apply to a college and those that drive their final enrollment decision (Goenner & Pauls, 2006). For instance, students' residential status (in-state v. out-of-state) often plays an outsized role in their application and enrollment decisions. As of 2012, four in five first-time, degree-seeking undergraduate students attended a school in their state of residence (Kumar et al., 2015). This trend highlights the importance of geographical preference and, perhaps, serves as an indirect measure of the role important financial considerations play in high school students' decision-making process. Students from outside traditional recruitment areas tend to have fewer ties to an institution, may have a less clear understanding of its mission and academic reputation, and can be discouraged by the higher tuition costs and transportation expenditures associated with commuting to and from campus (Brugglink & Gambhir, 1996). This body of research suggests that for institutions that recruit heavily from particular regions, expanding their traditional footprint has both potential benefits and costs.

The economic and educational implications of the growing gender gap in college enrollment have also been well documented (Conger & Long, 2013; Cho, 2006; DesJardins et al., 2002; Card & Lemieux, 2000). In 2010, the National Center for Education Statistics (NCES) reported that only 43% of undergraduates were male (Snyder & Dillow). Further, through 2019, the NCES projected female student enrollment in colleges and universities across the country to grow by 21%, compared to just 12% for their male counterparts (Hussar & Bailey, 2011). Recent research also suggests differential performance and attendance patterns at high schools with higher college-going rates may be contributing to this existing divide (Conger & Long, 2013; Cho, 2006). For these reasons, institutions often target their resources to reduce potential imbalances in the undergraduate male to female ratio (Conger & Long, 2013; Cho, 2006; Card & Lemieux, 2000; Brugglink & Gambhir, 1996).

Institutional Factors

Early and personalized attention has also been shown to improve post-secondary outcomes. Researchers have long discussed the benefits of early outreach to college bound high school students, particularly those from impoverished backgrounds (Wyatt et al., 2014; Thomas et al., 1999). Even modest student engagement in the college preparation process has been shown to engender important postsecondary benefits often brought about by a stronger student and institution match (Thomas et al., 1999). Furthermore, scholars argue that outreach programs have evolved over time to compensate for the shortcomings of an underfunded public education system by offering a more comprehensive approach to college access (Perna & Swail, 2002; Swail, 2001). Research on the topic has also underscored the importance of early, more personalized attention as a driving factor in the college choice process. A 2011 survey of over

55,000 students from more than 100 public and private four- and two-year institutions nationwide found that early, “personalized attention prior to enrollment” was the fourth most important factor in students’ enrollment decisions, following cost, financial aid, and the academic reputation of the institution (Noel-Levitz Student Satisfaction Inventory, 2012). Over time, an understanding of the merits of early outreach and the resulting impact on direct postsecondary matriculation has emerged and become widely accepted.

Another area in which postsecondary institutions can exert more control, is their academic programming. Students’ sense of institutional fit and thus their enrollment decisions can sometimes be driven by their choice of major and the school’s perceived strength in that area (DesJardins et al., 2002). In certain instances, a college or university may even wish to attract students with interests in certain fields, as they are seen as mission-critical (Brugglink & Gambhir, 1996). A Ruffalo Noel Levitz (2016) report found that alignment with students’ intended majors was identified as one of the most effective strategies for student enrollment, retention, and completion at four-year private institutions. As a result, students’ intended major is often considered an important criterion in enrollment modeling.

In addition to these student-level factors, there is strong evidence that the high school contexts from which applicants emerge are often not only determinative of their college enrollment choice, but also closely linked to their postsecondary success (Harvill et al., 2012; Johnson, 2011; Johnson, 2008; Hill, 2008; Zarate & Gallimore, 2005). Drawing on data from the Educational Longitudinal Survey, Engberg and Wolniak (2009) argued secondary institutions play a normative role in promoting college enrollment by enabling students to acquire vital human, social, and cultural capital. Given the important function high schools play in moderating students’ enrollment decisions (Engberg & Wolniak, 2009; Johnson, 2008; Hill, 2008; Khattab, 2005), admissions models that examine the impact of student-level characteristics on college enrollment decisions must also account for different school-level variance. To this end, the mixed effects model outlined henceforth proposes a novel approach to leveraging student search services data, while still aligning closely with the rigorous methodological techniques outlined in related higher education studies (Engberg & Wolniak, 2009; Hill, 2008).

To ensure accurate projections, admissions models must account for students’ differential high school experiences and levels of preparation. Consequently, many resources exist to help enrollment managers segment their prospective student audience. One such tool is the College Board’s DescriptorPLUS service, which matches prospective students to institutions based on unique geodemographic neighborhood and high school information. High school cluster data, as it is known, provides broad descriptive characteristics upon which applicants are then grouped. These measures include, but are not limited to students’ college-going rates, advanced placement coursework, diversity, and SES. However, these important high school-level variables are often measured at a higher level of aggregation than the primary outcome of interest (e.g. student enrollment). As such, it follows that some groups of students may start from more advantageous positions and, thereby, carry higher enrollment probabilities. Oftentimes, these characteristics are modeled as a common effect across subsets of students from the same or similar types of secondary institutions (Raudenbush & Bryk, 2002). In the context of higher education, this shared effect represents a form of dependence among the enrollment probabilities of individuals from similar backgrounds (Collett, 2015; Lu & Peng, 2008; Raudenbush & Bryk, 2002).

Multi-level Analysis

Modeling institution-level effects is imperative for admissions staff when examining student-level data. For one, this can help explain situations in which a group of students who have similar values for certain explanatory variables may nonetheless be observed to have different enrollment probabilities. In a multi-level design, it is assumed that some individuals may have a greater likelihood of postsecondary enrollment than others. Student clustering is a very common phenomenon in the field of education. Typical examples include students nested within classrooms, classrooms nested within schools, and schools nested within districts. As such, individual students cannot enter the model as independent observations as their outcomes will tend to align more closely with others from similar contexts and equally differ from those in different contexts. In this analysis, a multi-level model addresses the hierarchical relationship of students nested within high school clusters (Raudenbush & Bryk, 2002).

Furthermore, in a traditional modeling approach there is typically no direct measure of an individual high school’s impact on a student’s or group of students’ college choice and postsecondary outcomes. Meaning, resource disparities, differential teacher quality, adherence to effective instructional practices, school violence, etc. are often difficult to track or quantify, and sometimes completely unavailable. Such disparate variables typically have many possible values and it would be unrealistic to build these differences into the model as fixed effects, as it would likely introduce a large number of unknown parameters. To the extent possible therefore, a multi-level design attempts to incorporate these dependent effects on students’ event times without necessitating precise measurement of each individual component. Thus, this approach incorporates the impact these common effects have on students’ observed event times through just one parameter, namely the variance of the random effects' assumed underlying distribution (Collett, 2015).

The persistent gaps in educational opportunities across traditional fault lines (e.g. race, SES, sex, etc.) have been exacerbated by recent, broader economic trends. In particular, federal and state spending cuts at many public and private universities have resulted in tuition increases and less per capita spending on student education over the past decade. These developments threaten to diminish student access and thereby negatively impact student outcomes. This proposed research seeks to inform on postsecondary institutions' efforts to reduce recruitment expenditures, while maintaining an emphasis on student-institution match and strong outcomes. As the undergraduate admissions cycle is an inherently time dependent process, this analytic approach aims to identify and quantify the impact of those student and aggregate high school factors that effectively reduce time to enrollment.

Time to Event Analysis

Time to event modeling in the context of higher education has become increasingly common, though its application remains limited (Kim, 2012, 2011; Gross and Torres, 2010; Bahr, 2009; Scott & Kennedy, 2005; DesJardins et al., 2002, 1999, 1997; Murtaugh et al., 1999; Singer & Willett, 1991). This analytic approach enables researchers to focus on the intervals between two points of interest – typically measured in semesters, quarters, or academic years (Bahr, 2009). Thus, students are said to enter the "risk set" when they enroll, for example, and are considered "at risk" until they experience a single outcome, such as graduation, or the first of several interdependent, competing events, such as graduation, transfer, or drop-out. Desjardins et al. (1999), for instance, were among the first to apply a time to event model to investigate those student and institutional factors that affect college departure. In their study, they concluded that key explanatory variables do in fact have differential effects over time. More recently, other scholars have applied similar techniques to examine undergraduate persistence and completion as well.

Time to event models are useful analytic techniques when scholars' primary interest rests not only on students' end decisions, but also on their timeline for making those decisions. For instance, prior research has shed important light on those factors that meaningfully influence students' college enrollment. However, higher education administrators often face competing goals of making and shaping their institutions' incoming freshman class. Segmenting prospective student populations enables enrollment management offices to more narrowly target messaging and recruitment activities. This can help to shorten time to enrollment for certain subsets of students, which not only reduces the burden on families, but also frees important resources for admission and student services personnel.

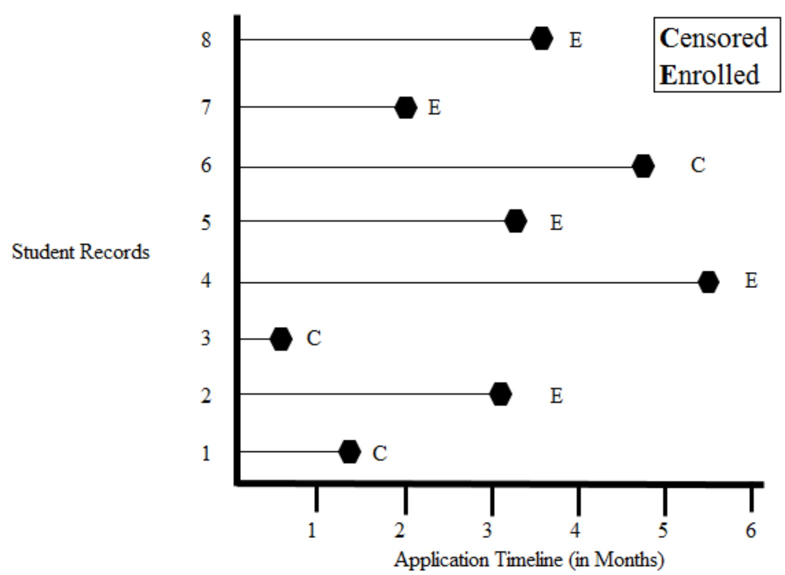
These models also provide a more nuanced picture of the admissions process and are straightforward in their application. The primary unit of measurement in time to event models is time itself – typically bracketed by a well-defined point of origin and the occurrence of a particular event or pre-specified end-point (Hosmer et al., 2008). Such models are most common in the fields of medicine, where end points may be death or cancer recurrence, as well as more applied fields, such as engineering in which stress tests typically focus on machinery failure. Since the early 2000s, time to event analyses investigating student dropout and completion have provided important evidence of the insight such approaches can lend in a higher education context. In education research, natural intervals of interest often depend on the primary outcome under investigation. For instance, if the focus is on what factors shorten time to admission, a natural starting point may be their application submission. Rather, if time to enrollment is the target, admission may be of more interest.

There are many reasons why standard statistical procedures may be inappropriate when analyzing time to event data. First, time data are typically not symmetrically distributed, but rather positively skewed. This results in certain baseline assumptions inherent to more traditional approaches being violated. Perhaps more importantly, though, is the presence of censored observations, which render standard statistical methods unsuitable (Collett, 2015; Kleinbaum & Klein, 2012; Kim, 2007). Time to event data are said to be censored when the event of interest is not observed for select individuals in the designated time frame. Three of the most common causes of censoring include: 1) a student not enrolling in any institution during the observation period; 2) the student's record being lost to follow-up, meaning no new information that could contribute to the model is known about that individual in the appropriate timeframe; or 3) there were mitigating circumstances which made immediate postsecondary enrollment impossible, such as illness, military service, or electing for a gap year.

As Scott and Kennedy (2005) noted in their investigation of competing outcomes in sub-baccalaureate enrollment among nontraditional adult students, information on some students will not always be readily available in real-life situations. In certain circumstances, data may be missing because an event, other than the one of primary interest (e.g. death or chronic illness), may occur precluding the student from finishing their studies or remaining continuously enrolled. In addition, data can also simply be lost or incomplete due to human or data entry error. Rather than excluding these elements, as is often required by standard statistical procedures, time to event models incorporate censored data efficiently, while simultaneously allowing researchers to dynamically measure the impact of specific interventions over time.

The most common form of censoring, and the method used in this analysis, is known as right censoring – when an individual enters the analysis by being admitted, but does not experience the event of interest (enrollment) by the last recorded observation. Figure 1, adapted from Collett (2015), provides a simple diagram of what right censored data might look like for a subset of eight students over the course of the application process. This example illustrates the objective of this type of analysis, namely, comparing the duration between two well-defined points in time. The start time for each student (reset to point 0 on the x-axis) corresponds to each applicants' admission date, while the end points can vary between postsecondary enrollment at one of four institutions or censoring. Thus, this analysis examines those student and institutional characteristics that inform on the primary unit of measurement, time.

Figure 1. Example Application Timeline Including Right Censored Records



In addition to their flexibility, initial evidence suggests that time to event models can also offer educational stakeholders profound insight into important student outcomes. For instance, simply adding a temporal dimension to these models can have important policy implications. DesJardins et al. (1999) argued that pinpointing the times at which students are most at risk of leaving college enables federal, state, and institutional stakeholders to target their interventions more efficiently. Similarly, Kim (2012, 2011), as well as Gross and Torres (2010) explained how the timing and type of financial aid can impact various postsecondary outcomes among minority student populations. Bahr (2009) also showed how these models could be extended to accommodate more complex designs, such as a repeated measures analysis investigating students' rate of progress through a remedial math sequence. Finally, additional analyses have also illustrated the ability of this modeling approach to assess and quantify the importance of multiple, interdependent competing events, such as graduation, transfer, or drop-out.

Competing Risks Framework

The extension of time to event modeling within the context of a "competing risks" framework is particularly important in the field of education in which overlapping and sometimes correlated events are common (Guerin, 1997; DesJardins et al., 1999, Ronco, 1996). Higher education in the United States is a complex and multilayered system where public community colleges coexist with highly selective, private four-year institutions (Kumar & Hurwitz, 2015). Each college or university plays a unique role in this setting, and students from a range of different backgrounds engage with the system at different and, often, multiple levels. For many students, the decision to enroll at a particular college is a difficult one, as they typically receive multiple acceptances requiring them to weigh the potential benefits of competing offers (Ledesma, 2009; Chang, 2006; Brugglink & Gambhir, 1996). In particular, much consideration has been given to high achieving students who, it is assumed, have many attractive admissions offers to consider and therefore start from a position of lower enrollment probabilities (Thomas et al., 1999). Further, as lower ranked institutions tend to offer more merit-based financial aid to their most desirable candidates, students must choose between going to a less prestigious institution at lower cost or attending a more selective college with a higher sticker price (Monks, 2009).

Fortunately, the allure of these competing options can be directly modeled. Subjects in these studies are typically followed until the occurrence of one of several pre-specified events or a predetermined end to the observation period. In such instances, the occurrence of the first event is said to preclude the occurrence of other events of interest. For example, Scott and Kennedy (2005) conducted an event history analysis in a discrete-time setting by modeling the *odds* (known as the *hazard* or risk in such models) of graduation, in the context of two competing risks, transfer and dropout. If admissions staff fail to account for the activities of other postsecondary institutions in their own yield models, they risk making decisions within an unrealistic vacuum devoid of competing options. Further, in these instances, the standard product limit, or Kaplan Meier (KM) method of estimating the distribution of the underlying time to event probabilities, by ignoring events of all other types, may yield biased or misleading results (Austin et al., 2016; So et al., 2014; Kim, 2007; Harrell, 2001).

In a competing risks analysis, the influence of covariates can be evaluated in relation to the cause-specific hazard or cumulative incidence of students' different enrollment decisions (Austin et al., 2016; Dignam et al., 2012). The choice of model has implications for how the results of the analysis can and should be interpreted. Competing risks analysis is becoming increasingly common in biomedical research, a field in which multiple, potentially overlapping outcomes is fairly common (Rodriguez et al., 2015; Coleman, 2014; Haller et al., 2013; Noordzji et al., 2013; Abdollah et al., 2011; Gillam et al., 2010; Glynn & Rosner; 2005). For instance, there are many studies that assess competing risks in the field of clinical cancer research where local/distant cancer recurrences, new cancer diagnoses, and death are important and competing events of interest (Dasgupta et al., 2016; de Glas et al., 2016; Bianchi et al., 2014; Lughezzani et al., 2011; Kim, 2007). Despite widespread application in the medical field, this modeling approach has gained less traction in the field of higher education (Scott and Kennedy, 2005; Guerin, 1997; DesJardins et al., 1999; Ronco, 1996).

While there are many enrollment options from which applicants can choose, it is not always possible to discern to what extent dependence between event times may exist in such models. Through many simulations, Dignam et al. (2012) showed that when covariate effects are ‘shared’ among competing events, it may be the case that none achieves statistical significance when modeled on the cumulative incidence scale. In addition, scholars argue that cause specific hazard ratio (CHR) estimates are often far better suited for addressing etiologic questions when these covariates, or common effects, are available for modeling (Austin et al., 2016; Allison, 2010). Given these statistical and practical considerations, the current analysis will initially investigate modeling CHR estimates for the four institutions under review (Kleinbaum & Klein, 2012).

Study Significance

Since the early 2000s, time to event modeling has been used to examine critically important issues, such as student completion and graduation. Nonetheless, the bulk of enrollment modeling remains limited to more traditional modeling techniques, such as binary logistic regression. The proposed model will build on the extensive undergraduate enrollment literature, while simultaneously augmenting and extending the field's emergent interest in time to event models. The multi-level design will also appropriately account for variation driven by aggregate high school-level characteristics. Finally, this approach will simultaneously assess the roles competitors play in a crowded higher education market, thereby enabling institutions to incorporate important information on the appeal of similar profile colleges into their own yield models.

The main objective of this sort of model is to identify those covariates that are related to and drive students' enrollment decisions. By delineating between the effects these factors have on students' enrollment times, admissions professionals can gain crucial insight into students' enrollment probabilities over time (Hosmer et al., 2008). For instance, certain student- and institution-level variables may shorten or lengthen students' enrollment timelines, and the effects of these factors may differ across institutions and over time. Given the complex interplay between such variables, a primary goal of the proposed research will be to leverage as much student data as possible.

To the author's knowledge, the statistical approach outlined henceforth has not been formally applied to the analysis of undergraduate enrollment preferences in a multi-level, competing risks framework. The proposed analysis will provide an empirical measure of the determinants of undergraduate enrollment in the context of a large and competitive postsecondary marketplace. By doing so, it accounts for the ways in which students must engage with the complexity of the sprawling and ever-changing U.S. higher education system, as well as the roles played by often very similar institutional actors.

CHAPTER THREE

METHODOLOGY

The research questions that will guide this analysis include:

1. Do certain sociodemographic, institutional, financial, and academic factors effectively reduce time to postsecondary enrollment?
2. Do these sociodemographic, institutional, financial, and academic factors have common effects across several similar-profile postsecondary institutions?
3. How can the magnitudes and direction of the effects select sociodemographic, institutional, financial, and academic factors have on time to enrollment at different postsecondary institutions inform an enrollment management strategy?

Data

The sample for this analysis consisted of over 69,960 de-identified undergraduate application records drawn from a single mid-sized, private not-for-profit institution located in the Midwest between 2013 and 2015. Institutional data from this single site served as a baseline reference for all summary and statistical estimates presented throughout this analysis. Application data included measures of high school seniors' academic ability and major preferences, as well as select geodemographic and sociodemographic factors. Using student data from a single source ensured consistency in how important metrics, such as high school GPA, were recorded, thus providing a common reference point for interpretation.

These application elements were then compared across three similar-profile peer competitors using enrollment information appended to the original dataset. As institution-specific metrics were not available for these three institutions, only parameters that were not subject to change from one institution to another were included in this analysis. Specifically, standardized test scores, sex, race, student's geographic location, etc. (Table 1). These three institutions were selected due to the similar academic profiles, geographic proximity, and overlapping recruiting footprints. All four institutions are Doctoral/Research universities, according to their Carnegie classification, with average annual enrollments around 10,000 students. The institutions represented range from selective to highly selective private-not-for-profit universities. Each institution included in this study also draws a plurality of its undergraduate enrollment from in-state applicants, but maintains national recruiting profiles.

The use of archived undergraduate application records for this project was initially sanctioned by the university's Enrollment Management division. The project was also submitted to the college's Institutional Review Board (IRB) and was found to be exempt in July 2017. Student criteria selected for inclusion in the analysis were aggregated from multiple internal data sources using MySQL. Interim checks to ensure accurate and reproducible results were implemented at multiple steps throughout the process for quality assurance purposes. Initial coding decisions and syntax were vetted by appropriate database administrators and university personnel. Finally, a sample of individual student records in the final dataset were then examined manually to confirm consistent reporting across each of the internal systems.

In accordance with Family Educational Rights and Privacy Act (FERPA) guidelines, only archived data were referenced for this analysis and all personally identifiable student information was removed. At a minimum, these include student names or identification numbers, as well as date of birth and detailed geodemographic records. As additional safeguards, further steps to remove all extraneous variables from the sample were taken to ensure only those aggregate criteria necessary for modeling were retained. As such, only fifteen student-level variables were included, each of which was tracked in the most discrete manner possible to still provide analytic insight (see Variable Selection section below). Finally, the results of this analysis are only reported in summary or statistical format.

The final dataset is a combination of both internal and external sources. Data on aggregate high school characteristics were drawn from the College Board's DescriptorPLUS services and merged with institutional data. Further, federal financial aid eligibility, which was determined from information students provided on the Free Application for Federal Student Aid (FAFSA), also supplemented this analysis. Finally, students' college choice was confirmed and appended to the aggregate dataset using information from the National Student Clearinghouse (NSC). Each of these databases are described in further detail below. As the focus of this project is enrollment yield, this analysis will examine the direct matriculation patterns of admitted students. More information is generally available on this subset of students, thereby increasing the likelihood of accurate predictions (Thomas et al., 1999).

Free Application for Federal Student Aid (FAFSA)

Among the FAFSA data reviewed for this analysis were students' Estimated Family Contribution (EFC), which served as an adjusted proxy for their socioeconomic status. Other important variables included the number of other institutions to which applicants submitted their FAFSA information and parent education level. This information was supplemented with

College Board's DescriptorPLUS service, which segments prospective students into high school clusters based on various academic and geodemographic factors.

College Board

The College Board's DescriptorPLUS services utilize High School Cluster tagging to segment soon-to-be high school graduates according to academic, financial, and geographical measures (2006). By leveraging the geodemographic data of over four million students from more than 28,000 high schools, the College Board has generated 28 descriptive high school clusters (see Appendix B). These clusters group secondary institutions based on students' prior academic achievement, rates of extracurricular activity, college enrollment preferences, diversity, SES, and so on. Aggregate high school characteristics included in the analysis reflect the academic quality, poverty levels, and racial/ethnic composition of the student populations.

National Student Clearinghouse (NSC)

Further, this sample was augmented by data drawn from the NSC. The NSC is a nonprofit and nongovernmental organization that provides educational reporting, data exchange, verification, and research services to participating postsecondary member institutions. Since its inception, over 3,600 colleges and universities have participated in the Clearinghouse to report enrollment and degree information – accounting for 98% of all students enrolled in public and private U.S. institutions. Working with partner institutions, the NSC is designed to facilitate compliance with FERPA and The Higher Education Act, among other applicable laws. The NSC was the primary source of information on the destinations for admitted students who enrolled at any institution, including two- and four-year public/private institutions.

Variable Selection

The selection of student-level characteristics was informed by prior empirical studies on the topic of college choice (see Literature Review section). The following individual- and institution-level pre-collegiate characteristics are included in the analysis (Table 1).

Table 1. Student & Institutional Variables

Sociodemographic Factors
Sex
Race
Ethnicity
Residential Status
U.S. Region
First Generation Status
Institutional Factors
Intended Major
First Choice College
Number of College Applications
Target of Early Outreach
Financial Factors
Number of Kids in College
Pell Grant Eligibility
Merit Aid
Academic Factors
Cumulative High School GPA
ACT Test Scores

Variable Definitions

Sociodemographic Factors

Student sex is a binary indicator variable based on students' application responses. In the model, the value "Female" serves as the referent. Students' race is a derived multinomial variable with five distinct levels: Asian, Black or African-American, Multi-Racial, White, and

Other (Not Specified). Students were assigned into racial categories that aligned with the information provided on their college application. Students who self-identified as descending from more than one racial background were reassigned into a 'multi-racial' domain. Student ethnicity was tracked separately as a binary indicator variable. Students' with Hispanic heritage were recorded as "Hispanic," and the value "Non-Hispanic" serves as the referent.

Similar to sex and ethnicity, residency status is a binary indicator variable based on students' application responses. Residency was a derived variable based on students' entry for their state of origin. The value "Out-of-State" served as the referent, so the impact of being "In-State" could be modeled directly. Similarly, state information was further categorized based on U.S. Census Bureau regions. The four regions included in the univariable analysis were Midwest, Northeast, South, and West (including Pacific).

First generation status is a derived binary indicator variable based on students' responses across multiple parental education fields on the FAFSA. Specifically, to qualify as a first generation student, an applicant had to indicate that both their mother and father did not complete any 'college or beyond' (level 3 on the FAFSA form). As a result, only students who reported that neither of their parents completed grades beyond high school were tracked as first generation. The value "Not First Generation" served as the referent, so the impact of being "First Generation" could be modeled directly.

Institutional Factors

Students' intended major is also a derived multinomial variable taking six distinct levels: Business, Communication, Education, Liberal Arts, STEM fields, and Undecided. These categories were generated based on students' responses to two questions on their undergraduate applications. The first was the school into which the students planned to matriculate, which was an aggregated field one level higher than the major category itself. These responses served as broad categories, which when appropriate, were retained for the analysis.

To provide additional insight into students' major preferences, a more nuanced approach was taken for their responses to the major question itself. Across all four institutions included in this analysis, broader categorizations, such as Arts & Sciences, were utilized. As a result, students' responses to the major question on their applications were used to delineate between the Liberal Arts and STEM fields. Coding decisions were cross validated with the Bureau of Labor Statistics' STEM designation index (STEM Index, 2016). Those students who indicated they were undecided about their intended major or college on both questions were grouped together as "Undecided."

An a priori decision was also made to evaluate the impact of being admitted into a first choice school. First choice school designation is a derived variable using information drawn from both the FAFSA and NSC database. On the FAFSA, students are asked to designate up to 10 schools to which they want their financial information disclosed. Prior research suggests that most students list the schools in order of preference, and nearly two-thirds of applicants enroll in their first choice school if admitted (CNN Money, 11/24/2015).

To quantify the impact of this first choice preference, students' FAFSA data were supplemented by information downloaded from the NSC. Specifically, the name of the institution into which admitted students enrolled, as well as the matriculation date were cataloged and merged. Aligning this information, a first choice indicator variable was created. Therefore, First Choice is a binary variable based on an amalgamation of student information from the

FAFSA and NSC. The value “Not First Choice” served as the referent, so the impact of being admitted into a “First Choice” college or university could be modeled directly.

Similarly, recent research has shown that the number of colleges to which students apply often affects their college enrollment decisions (Smith, 2011). As one in four high school graduates who apply to four-year colleges still do not enroll in one, the number of college applications was identified as a potentially important predictor of time to enrollment (Avery & Kane, 2004). The number of schools to which students submitted their FAFSA information was tracked as an ordinal count, ranging from 0 to 10. This served as a proxy for the number of schools to which students applied. Evidence shows that increasing the number of college applications can increase a student’s probability of enrolling at a four-year college by as much as 40 to 50 percent (Smith, 2011).

Early outreach and contact are also strong indicators of student engagement. To measure the effect of early outreach efforts, the date of students’ first contact of record was identified and coded as a binary variable. Those students with whom schools had contact prior to the fall semester of their senior year of high school were designated as early outreach targets. Those students whose first contact was after the start of their fall semester of their senior year were grouped as part of the regular communications flow. This was a derived indicator variable based on student recruitment and marketing logs. The value “Normal Communication Flow” served as the referent, so the impact of “Early Outreach” could be modeled directly.

Financial Factors

Several additional variables from students’ FAFSA submission were also used to create indicators of financial need or burden. The first of these was a variable designating if a student’s family had other children in college at the same time. An a priori hypothesis was that students whose family had multiple children in college may delay their enrollment decision to maximize the amount of financial aid they were offered. To assess the impact of this factor, the number of kids of in college was tracked as a binary variable with zero as the referent, so the impact of having any additional children (>= 1) in college could be modeled directly.

Using information students and their families reported on the FAFSA, the Department of Education also derives what's known as an Estimated Family Contribution (EFC). This serves as index number that college financial aid staff use to determine how much financial aid a student would be eligible for if they were to attend their school (Department of Education). This variable is a continuous measure that serves as an adjusted proxy for students' socioeconomic status. Each year, the government establishes a threshold below which students are Pell Grant eligible. This figure has typically ranged from $5,000 to $6,000 over the past few years. The primary purpose of the Federal Pell Grant Program is to provide need-based grants to low-income undergraduate students (Department of Education). Using EFC estimates for each of the three years, as well information collected from the Department of Education, Pell Grant eligibility status was calculated for each of the application cycles included in this study. Pell Grant Eligibility is a binary variable based on EFC data drawn from students' FAFSA submission. The value "Not Pell Grant Eligible" served as the referent, so the impact of being "Pell Grant Eligible" could be modeled directly.

Another important measure of students' financial status is whether they were eligible for and received an offer of merit aid. Consideration for merit aid is based on a review of various student reported criteria including, but not limited to their academic ability and standardized test scores. Merit Aid is a binary variable based on institutional data drawn from students' application submission. The value "No Merit Aid" served as the referent, so the impact of receiving "Merit Aid" could be modeled directly.

Academic Factors

Two measures of students' academic ability were also evaluated in this analysis. The first was students' cumulative high school grade point average (GPA). This variable was tracked as a continuous measure on students' application based on a review of their official high school transcripts. On univariable analysis, the effect of high school GPA was measured in two ways. First, unit increases of 0.50 (equivalent to a one standard deviation increase) on the variable's continuous scale were reviewed. Second, students' high school GPA was also binned into quartiles: low, low middle, high middle, and high. The impact of a unit increase on this more discrete ordinal scale was then also modelled directly.

The second measure was students' standardized ACT test scores. This variable was also tracked as a continuous measure on students' application based on official score reports. Valid scores ranged from 0 to 36. On univariable analysis, the effect of ACT scores was measured in two ways. First, unit increases of 4 (equivalent to a one standard deviation increase) on the variable's continuous scale were reviewed. Second, similar to GPA, students' ACT scores were also binned into quartiles: low, low middle, high middle, and high. The impact of a one unit increase on this more discrete ordinal scale was then also modelled directly. Ultimately, continuous measurements of GPA and ACT, using predetermined unit increases, provided a stronger model fit. Therefore, the quartile approach for both factors were not reported.

In line with prior research (Johnson, 2008), special consideration was given to information derived from the FAFSA in the modeling process. Information on First Generation status, number of kids in college, number of college applications, First Choice, and Pell Grant eligibility is not available for students who did not file a FAFSA. Consideration was then given to multiple imputation as a method for addressing missing data elements. However, while simulation studies have shown that multiple imputation can perform well, under certain circumstances, for variables with up to 50% missing observations, larger amounts of missing information can lead to estimation problems and are generally not recommended (Allison, 2002; Johnson and Young, 2011). Thus, models including these variables were limited to student records for which complete FAFSA information was available.

Additional Model Parameters

As time to enrollment is the primary unit of measurement in this analysis, particular attention was paid to how the time variable was calculated. Accurate and detailed time logs were available for all major stages of the application cycle. Using customer relationship management (CRM) software (Technolutions Slate), specific dates and times for the point of application, admission, and enrollment were reviewed. In addition, records of students being denied admission and the timing of their first point of contact were also tracked.

The primary outcome of interest was enrollment in one of four similar profile institutions, thus an a priori decision was made to limit this analysis to only admitted students. This both adhered to the tenets of a traditional enrollment funnel (e.g. only admitted students can enroll), but also meant that complete information on important model parameters was also often available for most retained student records. As a result, students denied entry, as well as though who simply did not progress past the point of application were omitted, and the date of admission was used as time zero for the analysis. Students' date of enrollment or last known follow-up bookended the observation period. All students who withdrew from the application process or had no additional contact after their admission date were therefore treated as censored variables.

Due to the presence of competing risks, NSC data were also referenced to derive a multinomial outcome variable. In addition to censored records, there were four universities tracked in this study. As a result, the primary dependent variable had five levels. A value of zero indicates the admitted student did not enroll at any of the four institutions, while each school is assigned a value ranging from 1 to 4 to indicate enrollment at one of the institutions during the observation period.

Time to Event Models

There are two main approaches to conducting a time to event analysis, standard parametric and non-parametric procedures (Austin et al., 2016; Collett, 2015). Models in which a pre-specified probability distribution is assumed for the underlying time to event estimates are known as parametric models. Such techniques often require a thorough review of the modified Cox-Snell residuals to ascertain which of many possible probability distributions (Weibull, Exponential, etc.) presents the best fit. Non-parametric or "distribution-free" methods are far more common and do not require specific a priori assumptions to be made about the underlying distribution of students' enrollment times (Collett, 2015). The Cox regression model is perhaps the best-known extension of traditional non-parametric procedures.

In standard Cox regression, the primary objective is often to explore the relationship between a set of explanatory variables and the time to a single event of interest, enrollment in this instance (Collett, 2015). Allison (2010) provides the basic model form below:

$$h_i(t) = \exp(\beta' x_i) h_0(t)$$

Here $h_0(t)$ represents the baseline hazard, x_i is the vector of values for the independent variables for the *i*th individual, and β is the vector of their coefficients (Collett, 2015). The goal of the modelling process is to determine which combination of potential explanatory variables affect the form of the underlying probability distribution (Collett, 2015; Kleinbaum & Klein, 2012; Allison, 2010). However, in many instances, a student's enrollment decision is not solely driven by a set of clearly defined explanatory variables, but rather also influenced by the activities and outreach of other universities. In the presence of these "competing risks," enrollment at a competitor institution is said to preclude direct matriculation at the institution of primary interest, and these activities have implications for the data analysis (Austin et al., 2016).

Competing Risks Modeling

Competing risks observations provide important context when evaluating time dependent processes. In a competing risks model the cause specific hazard (CHR) heuristically represents the probability of enrollment at one institution at a moment in time (*t*), given that the student has not already enrolled at another institution (Austin et al., 2016; Dignam et al., 2012; Belot et al., 2010; Dignam & Kocherginsky, 2008). Collett (2015) specifies the following model as an extension of the standard Cox regression:

$$h_{ij}(t) = \exp({\beta'}_j x_i) h_{0j}(t)$$

Here $h_{0j}(t)$ represents the baseline hazard for the *j*th cause, x_i is the vector of values for the independent variables for the *i*th individual, and β_j is the vector of their coefficients for the *j*th cause (Collett, 2015). In a competing risks setting, an individual can potentially enroll at any of several institutions, but only the time to event for the first of these is observed (Dignam & Kocherginsky, 2008). Importantly, though, even when only one event is observed per student, partial information on enrollment at other colleges is available due to censoring.

There are many technical and practical advantages to this modeling approach. In real-life situations, modeling CHR estimates provides important predictive value as only the earliest enrollment time and at most one enrollment type is observed (Austin et al., 2016; Belot et al., 2010; Peterson, 1976; Gail, 1975; Tsiatis, 1975). Further, CHR estimates from these models are largely interpreted in the same way as the hazard ratio derived from a tradition Cox regression in the absence of competing risks. When interest lies in identifying those variables that inform directly on the event of interest, CHRs indicate the odds (known as hazards or risks in such models) of a student enrolling at any given time at one of several institutions as a function of univariable and multivariable individual and institutional factors. By accounting for the common effects identified throughout the literature, one can confidently rely on functions of the observed CHRs for interference when analyzing data with respect to the first enrollment event (Dignam et al., 2012).

Multi-level Modeling

When data are nested, as in this case, the assumption that observations contribute to the model independently is violated. Thus, the standard errors produced by such models are often too small, which can lead to a higher probability of Type II error (incorrectly rejecting the null hypothesis, or a false positive) than if the observations were truly independent (Raudenbush & Bryk, 2002). To account for this clustering effect in the proposed time to event model, a mixed effects model of the following form is proposed:

$$h_{ij}(t) = z_i \exp(\beta'_j x_i) h_{0j}(t)$$

Substituting $z_i = \exp(u_i)$ provides an alternative representation of the clustering effect, which is generally considered more convenient for modeling purposes (Collett, 2015):

$$h_{ij}(t) = \exp\left(\beta'_j x_i + u_i\right) h_{0j}(t), \qquad u_i = 0 \; corresponds\ to\ no\ clustering\ effect$$

In this model, u_i is a random effect in the linear component of the proportional hazards model. This model includes student-level information as fixed effects, but also allows for random intercepts at the level of high school cluster. In total, there were 27 unique clusters representing nearly 4,500 individual high schools included in the analysis. This list of secondary institutions included a range schools, such as select enrollment magnet and smaller rural schools, as well as general public and private entities. Incorporating high school cluster information in the model as a random effect based on externally validated and aggregated criteria efficiently and appropriately accounts for the correlation among students who come from secondary institutions assumed to share certain characteristics (socioeconomic, academic ability, college preparation, etc.).

Further, the random effects in this model introduce a degree of dependence across students' time to enrollment estimates, thus anticipating and accounting for important variation at the high school level (Collett, 2015; Kleinbaum & Klein, 2012; Allison, 2010). These random effects are assumed to have levels drawn from a "population of possible values, where the actual levels are representative of that population" (Collett, 2015). While the effects corresponding to student level factors, the fixed effects in such models, may remain largely unchanged, these multi-level models incorporate important institution-level variation that might otherwise be difficult or unwieldy to incorporate in a single model.

Model Building Process

Based on an extensive review of the existing literature, various measures of students' academic ability and SES, as well as sociodemographic factors, intended major, and school choice were selected for inclusion in the exploratory univariable models. As this analysis involves only one response variable per observation (e.g. enrollment), univariable is thus defined as a model that employs a single explanatory variable (Hidalgo et al., 2013; Tsai, 2013; Peters, 2008). These analyses were conducted to individually assess the relationship between each explanatory variable and students' enrollment patterns. The objective of this preliminary stage is to determine which variables independently affect students' likelihood of enrollment at any given time. Any explanatory variables that are found to be marginally or meaningfully associated with enrollment outcomes (p <.10) in the univariable analyses will be considered for inclusion in the final multivariable model. As before, this analysis involves only one response variable per observation, so multivariable is defined as a model that employs multiple explanatory variables simultaneously (Hidalgo et al., 2013; Tsai, 2013; Peters, 2008).

Goodness of Fit Diagnostics

On multivariable analysis, assessment of model fit will be conducted using Akaike's Information Criterion (AIC). Unlike $-2\log\hat{L}$, the value of the AIC is penalized and will tend to increase when additional, unnecessary terms are added to the model (Collett, 2015; Allison, 2010). The formula for AIC is as follows, where q represents the number of unknown β-parameters:

$$AIC = -2\log\hat{L} + 2q$$

Utilizing AIC avoids overfitting models, which results in the estimated values of some of the beta coefficients being highly dependent on the actual data, thus limiting the generalizability of the results. An additional benefit of comparisons on the basis of AIC is that sequential models need not be nested. For interpretation purposes, smaller AIC values indicate a better fitting model. Specifically, when AIC decreases by more than two points upon removing an independent variable, the results indicate the more parsimonious model provide better estimates of the true expected values. In the event the AIC value remains unchanged or increases, the omitted variable should be retained in the final analysis as the more complex model provides a better approximation of the true relationship between the parameters (Agresti, 2007).

Statistical Assumptions

In any multivariable model, an issue of multicollinearity could arise. Multicollinearity in regression exists when two or more explanatory variables are highly correlated with each other, resulting in unstable regression coefficients (Weisberg, 2005). One approach to diagnosing collinear variables is to review variance inflation factor (VIF) estimates, with any VIF value > 10 indicating a potential problem. When issues of multicollinearity are detected, the highly correlated explanatory variables will be removed from the model and re-entered one at a time. In the final multivariable model, only those explanatory variables that best minimize AIC will be retained.

Further, in standard non-parametric Cox regression models, there is the assumption that the hazard of each event type is proportional over time. For this analysis, the proportional hazards assumption will be assessed as described by Cox and Oakes (1984). Cox and Oakes (1984) proposed a parametric test of the proportional hazards hypothesis using the following model:

$$\log h_j(t) = \alpha_0(t) + \alpha_j + \beta_j t, \qquad j = 1, 2, 3 \ldots$$

They showed that if all $\beta_j = \beta$ for all j, then the proportional hazards hypothesis is satisfied. Graphical evaluation of the Martingale residuals for each predictor will also be examined as described by Lin, Wei, and Ying (1993).

If the proportional hazards assumption is retained, it means the log-hazards for any two events (e.g. enrollment at institutions 1 and 2) are parallel, or proportional, at any given time *t*. In this analysis, with four competing events, the equation proposed by Cox and Oakes (1984) implies a multinomial logistic regression model with a generalized logit link.

```
*Parametric Proportional Hazards Test;
PROC LOGISTIC DATA=CompetingRisks;
CLASS Enrollment_Outcome (REF="1");
WHERE Enrollment_Outcome NE 0;
MODEL Enrollment_Outcome = DaysFromAdmitToFinalDecision / LINK=GLOGIT;
RUN;
```

In the above syntax, the four event types serve as levels of the dependent variable, while the measurement of time serves as the independent variable. Censored variables are excluded for this test using the "Where" statement. Under the proportional hazards hypothesis, the coefficient for time will be 0; therefore, a non-significant Type 3 effect for this test of proportionality indicates the null hypothesis should be retained.

Upon review, the proportional hazards assumption was retained, with a non-significant Type 3 effect (p = .39). These results confirm a multinomial Competing Risks model, an extension of the non-parametric Cox regression model, is an appropriate choice for the analysis. Further, the small chi-square statistics and beta estimates near 0 suggest that the hazard functions for all four event types were nearly identical (Tables 2 & 3).

Table 2. Output for Parametric Proportional Hazards Test

Type 3 Analysis of Effects			
Effect	**Degrees of Freedeom**	**Wald Chi-Square**	***P***
DaysFromAdmitToFinalDecision	3	3.0335	0.3865

Table 3. Output for Parametric Proportional Hazards Test

Type 3 Analysis of Effects						
Parameter	**Outcome**	**Degrees of Freedeom**	**Estimate**	**Standard Error**	**Wald Chi-Square**	***P***
Intercept	2	1	-1.4461	0.0455	1054.372	<.0001
Intercept	3	1	-3.2319	0.098	1086.6077	<.0001
Intercept	4	1	-3.9164	0.1403	778.7428	<.0001
DaysFromAdmitToFinalDecision	2	1	-0.00027	0.000501	0.2929	0.5884
DaysFromAdmitToFinalDecision	3	1	0.00152	0.00105	2.0759	0.1496
DaysFromAdmitToFinalDecision	4	1	-0.00115	0.00162	0.5008	0.4791

CHAPTER FOUR

RESULTS

A total of 35,434 students who were admitted into a single postsecondary institution between 2013 and 2015 were included in the final analysis. This total accounted for 50.6% of all applicant records (*N* = 69,962), with the remaining records being lost to follow-up, withdrawn, or denied. Students' academic profile, as well as breakdowns of their demographic, socioeconomic, and geodemographic characteristics are listed in Table 4. Overall, two-thirds of all admitted applicants were female (69%). The mean age among admitted students was 17, with a standard deviation (*SD*) of 0.53. The sample was predominately Caucasian (64%). While nearly four in five admitted students reported living in the Midwest (83%), only half of the students reported being in-state (50%).

According to the College Board's DescriptorPLUS ratings, nine in ten students originated from high school clusters that aligned with high college-going rates, students with strong academic ability, and high levels of extracurricular participation. Most students came from three of the 27 unique clusters: 79 (30%), 68 (22%), and 70 (21%). In line with College Board data, the mean high school GPA for admitted students was 3.80 (*SD* = 0.48) and the average ACT score was 27 (*SD* = 3.68). Most students indicated they intended to major in a STEM field (38%), followed by Undecided (19%), Business (16%), and the Liberal Arts (10%). On average, admitted students took just under two months to make their enrollment decisions, with a narrow range of one and a half to three months across the competitive set.

Two in three admitted students (67%) submitted FAFSA information. Of those students, approximately one in five (18%) were categorized as first generation status. A third of the sample (35%) reported that their family had one or more additional children in college and a nearly equal percentage (32%) were Pell Grant eligible. The median count of colleges to with admitted students applied was six, with an interquartile range (*IQR*) of four to nine. Only 18% of admitted students reported enrolling in their first choice institution.

Table 4. Descriptive Statistics

	Institution 1 (*N = 6,898*)	Institution 2 (*N = 2,372*)	Institution 3 (*N = 361*)	Institution 4 (*N = 159*)	Total (*N = 35,434*)
Race					
Asian	17%	11%	22%	11%	14%
Black or African American	4%	7%	10%	10%	6%
Caucasian	72%	75%	59%	69%	73%
Multiracial	6%	6%	8%	9%	6%
Other	1%	2%	2%	1%	1%
Ethnicity					
Hispanic	16%	23%	30%	28%	19%
Non-Hispanic	84%	77%	70%	72%	81%
# of Applications (*Mdn* , *IQR*)	3 (1 - 6)	1 (0 - 5)	3 (0 - 7)	2 (0 - 9)	4 (0 - 7)
First Generation	18%	27%	19%	17%	18%
# Kids in College	64%	67%	66%	71%	65%
Early Outreach	69%	67%	61%	60%	56%
Pell Grant Eligible	26%	29%	28%	18%	23%
Merit Aid	84%	55%	85%	79%	83%
Sex					
Male	33%	33%	32%	39%	31%
Female	67%	67%	68%	61%	69%
Residency (*% In-State*)	58%	70%	77%	76%	50%
ACT (*Mean, SD*)	26.4 (3.4)	25.1 (3.7)	31.5 (2.6)	31.8 (2.7)	26.6 (3.7)
GPA (*Mean* , *SD*)	3.76 (0.49)	3.65 (0.55)	4.40 (0.40)	4.36 (0.39)	3.80 (0.50)
Major					
Business	15%	24%	7%	4%	16%
Communication	6%	10%	7%	1%	6%
Education	3%	4%	2%	1%	3%
Liberal Arts	9%	11%	7%	13%	10%
STEM	51%	31%	60%	63%	45%
Undecided	17%	20%	17%	18%	20%
First Choice	49%	75%	83%	79%	88%
Region					
Midwest	86%	90%	92%	86%	83%
Northeast	3%	2%	2%	2%	4%
South	4%	4%	2%	6%	5%
West	7%	4%	4%	6%	8%

Note: IQR = Interquartile range.

Underscoring the near parity of the competitors included in the analysis, these breakdowns remained consistent across each of the four institutions. Between 61% and 69% of admitted students were female, with a mean age of 17, across all four institutions. A majority of students admitted to each institution were Caucasian and reported being in-state applicants (58% - 77%) originally from the Midwest (81% - 92%). Nearly all emerged from high school clusters (79, 68, and 70) with high college-going rates, reflected by their mean high school GPA (3.80 - 4.40) and ACT scores (26 - 32). A plurality of students indicated they intended to major in a STEM field in college.

As noted early, a majority of students admitted to each institution submitted FAFSA information. Of those students who submitted a FAFSA, approximately one in five (17% - 27%) were categorized as first generation status. A third of the sample (29% - 35%) reported their family had one or more additional children in college and a similar percentage (27% - 40%) were Pell Grant eligible. The median count of colleges to with admitted students applied ranged from four to nine (*IQR*: 2 – 10). In line with the overall sample, approximately one in five admitted students reported enrolling in their first choice institution.

Univariable Analysis

Measures of students' academic ability and SES, as well as sociodemographic factors, intended major, and school choice were selected for inclusion in the exploratory univariable models. These analyses were conducted to individually assess the relationship between each explanatory variable and students' enrollment patterns. The objective of this preliminary stage is to determine which variables independently affect students' likelihood of enrollment at any given time.

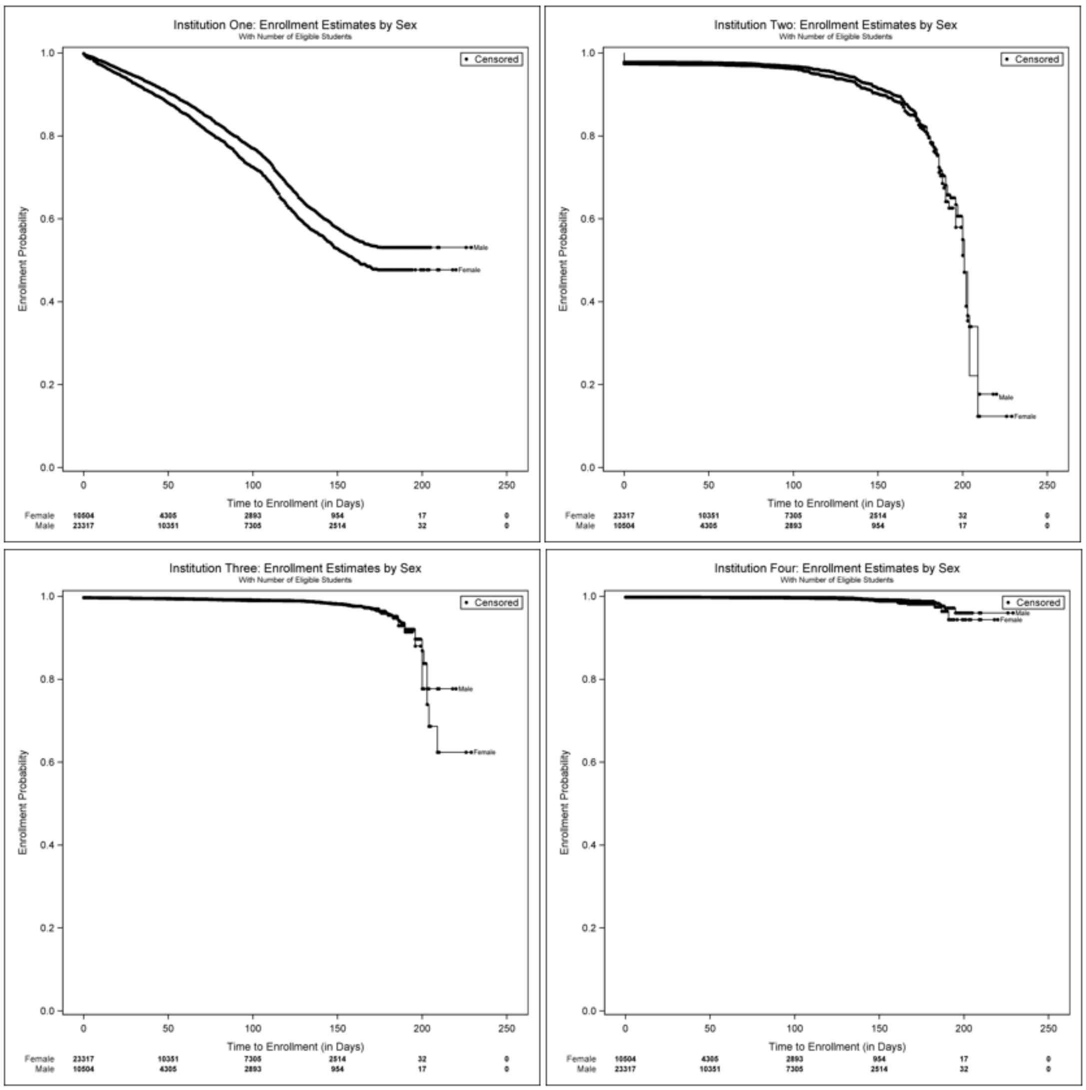
Sociodemographic Factors

Table 5. Univariable Models Assessing Sociodemographic Factors

		Institution 1		Institution 2		Institution 3		Institution 4	
	Valid N	**HR (95% CI)**	***p***	**HR (95% CI)**	***p***	**HR (95% CI)**	***p***	**HR (95% CI)**	***p***
Sex (*Ref*= 'Female')	31,821	1.21 (1.14 - 1.28)	<.0001	1.23 (1.10 - 1.37)	<.001	1.10 (0.85 - 1.41)	.47	1.49 (1.03 - 2.14)	.03
Race	29,169		<.0001		<.0001		<.0001		<.0001
Asian		1.29 (1.20 - 1.39)	<.0001	0.69 (0.58 - 0.82)	<.0001	1.98 (1.45 - 2.70)	<.0001	0.82 (0.43 - 1.54)	.53
Black or African American		1.05 (0.92 - 1.21)	.46	1.04 (0.81 - 1.34)	.74	3.23 (2.14 - 4.89)	<.0001	2.87 (1.59 - 5.18)	<.0001
Caucasian (*Ref*)		-		-		-		-	
Multiracial		1.24 (1.11 - 1.38)	<.001	0.94 (0.74 - 1.19)	.61	1.78 (1.11 - 2.88)	.02	1.92 (0.99 - 3.71)	.05
Other		1.04 (0.79 - 1.36)	.80	1.19 (0.75 - 1.88)	.45	3.04 (1.34 - 6.90)	.01	0.92 (0.13 - 6.68)	.94
Ethnicity	31,015	1.06 (0.98 - 1.14)	.13	1.36 (1.19 - 1.56)	<.0001	3.09 (2.39 - 4.00)	<.0001	2.28 (1.53 - 3.38)	<.0001
Residency	31,821	1.47 (1.39 - 1.56)	<.0001	1.94 (1.73 - 2.18)	<.0001	3.52 (2.64 - 4.71)	<.0001	1.71 (1.17 - 2.49)	.01
Region	31,821		.004		<.0001		<.001		.80
Midwest (*Ref*)		-		-		-		-	
Northeast		0.95 (0.82 - 1.11)	.53	0.58 (0.41 - 0.84)	.004	0.28 (0.09 - 0.88)	.03	0.71 (0.22 - 2.23)	.56
South		0.78 (0.68 - 0.90)	<.001	0.79 (0.60 - 1.03)	.08	0.26 (0.10 - 0.69)	.01	1.33 (0.65 - 2.73)	.44
West		0.94 (0.84 - 1.04)	.24	0.61 (0.47 - 0.78)	<.001	0.38 (0.19 - 0.77)	.01	1.07 (0.54 - 2.11)	.85
First Generation	22,222	1.21 (1.13 - 1.31)	<.0001	1.80 (1.57 - 2.07)	<.0001	1.28 (0.87 - 1.89)	.21	1.10 (0.62 - 1.97)	.74

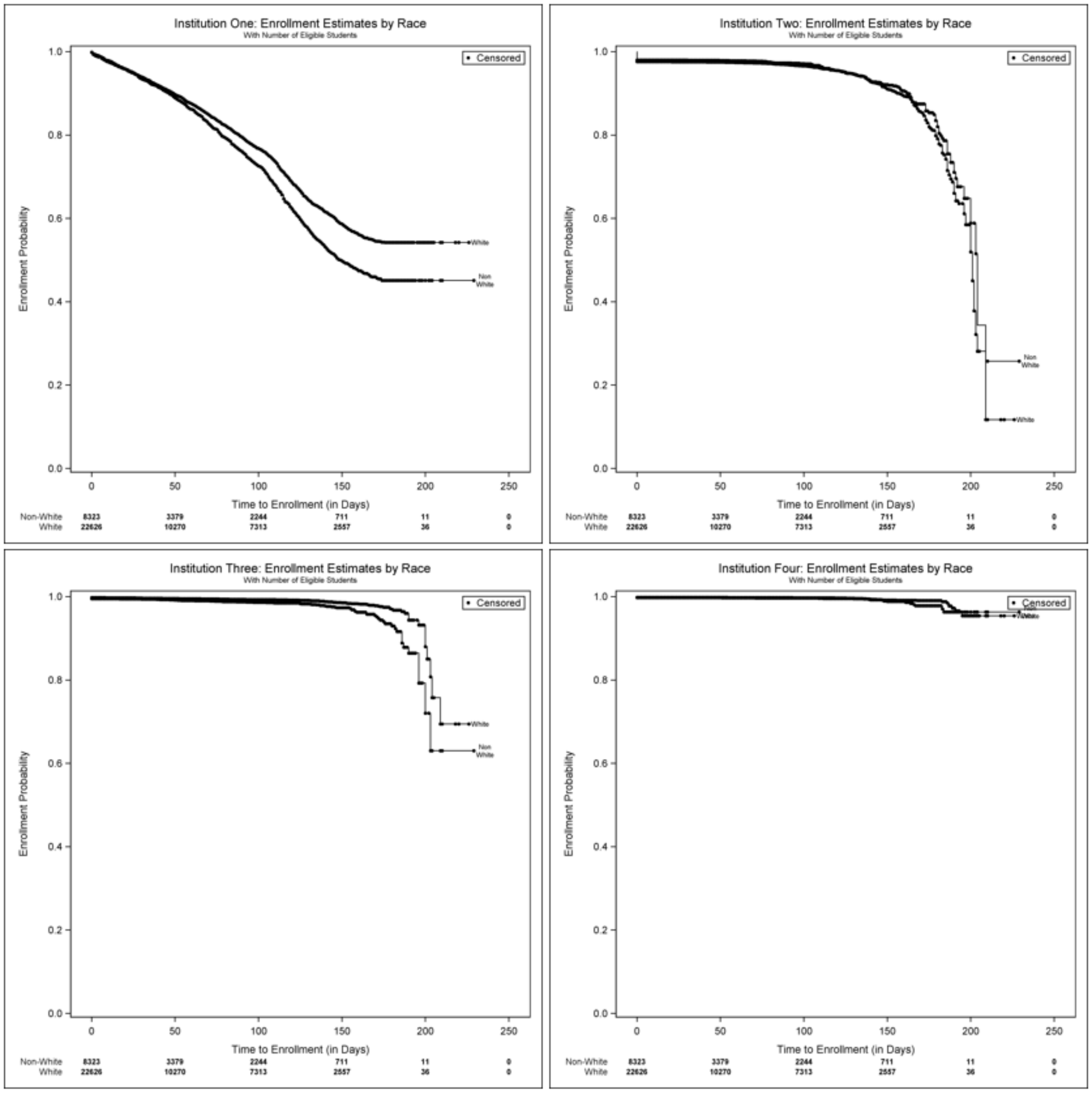
A student's sex was independently associated with the instantaneous odds of enrollment at three of the four institutions under review. Specifically, male students were 1.21 (95% CI: 1.14-1.28, p < .0001) and 1.23 (95% CI: 1.10-1.37, p < .001) times more likely to enroll at any given time at institutions one and two, respectively, compared to female applicants (Table 5). Male students were also 49% (Hazard Ratio [*HR*] = 1.49, 95% CI: 1.03-2.14, p = .03) more likely to enroll at any given time at institution four. The Kaplan-Meier results presented in Figure 2 indicate how students' enrollment probabilities changed over time based on their sex. These paneled findings align closely with the model output included in Table 5. For each institution, male students recorded higher enrollment probabilities throughout the application timeline. Nonetheless, for institution one there was clear separation between male and female students, while the differences between institutions two and four were less pronounced, but still evident. By contrast, the enrollment patterns by student sex for institution three were nearly identical.

Figure 2. Kaplan Meier Curves for Student Sex



In addition, students' race was also meaningfully associated with the odds of enrollment at each of the four institutions, but the magnitude and direction of the relationship varied across the four universities. Asian students were 1.31 (95% CI: 1.21-1.41, $p < .0001$) and 2.58 (95% CI: 1.85-3.61, $p < .0001$) times more likely to enroll at any given time at institutions one and three, but were 32% less likely to enroll at institution two ($HR = 0.68$, 95% CI: 0.57-0.81, $p < .0001$) compared to Caucasian students (Table 5). Black or African American students were 4.46 (95% CI: 2.85-6.97, $p < .0001$) and 3.68 (95% CI: 1.96-6.90, $p < .0001$) times more likely than their Caucasian counterparts to enroll at any given time at institutions three and four, respectively. Multi-racial students were also 24% ($HR = 1.24$ 95% CI: 1.11-1.38, $p < .001$) and 78% ($HR = 1.78$, 95% CI: 1.11-2.88, $p = .02$) more likely to enroll at any given time at institutions one and three compared to Caucasian admitted students. For the Kaplan-Meier results presented in Figure 3, the findings related to student race were simplified by dichotomizing student characteristics to White v. Non-White. This allowed for a clearer presentation of how students' enrollment probabilities changed over time based on the general racial breakdown of the student sample. When combined in this manner, broader trends emerged, but the results were equally divergent across all four institutions as observed in Table 5. Overall, white students were more likely to enroll at institutions one and three, but less likely compared to their minority counterparts at institutions two and four.

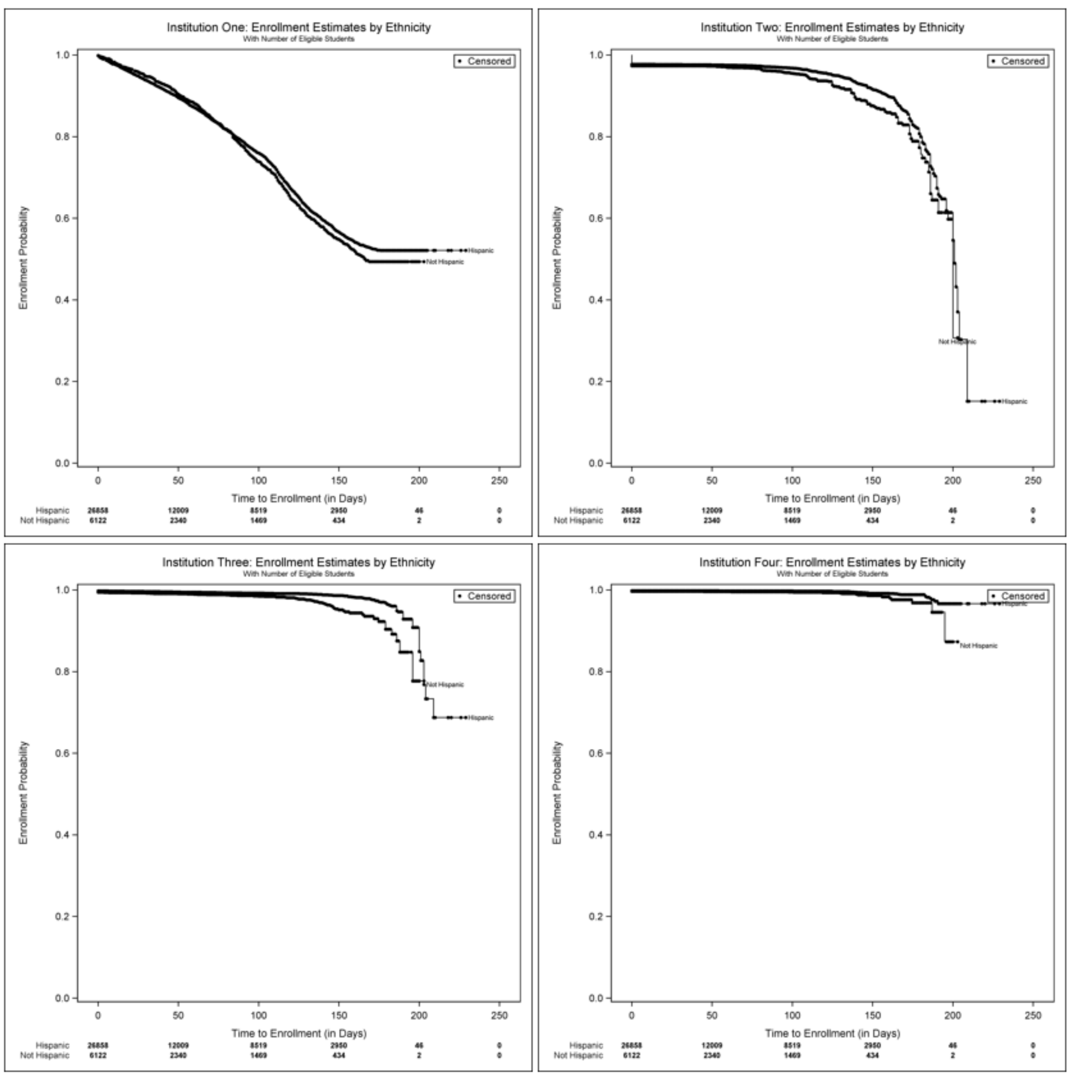
Figure 3. Kaplan Meier Curves for Student Race



Student ethnicity followed a similar pattern, as students who identified as Hispanic exhibited increased odds of instantaneous enrollment at three of the four institutions. This trend was most pronounced at institutions three and four. Hispanic students were 4.31 (95% CI: 3.06-6.05, $p < .0001$) and 3.25 (95% CI: 2.00-5.27, $p < .0001$) times more likely to enroll at any given time at institutions three and four (Table 5). Hispanic students were also 36% (*HR* = 1.36, 95%

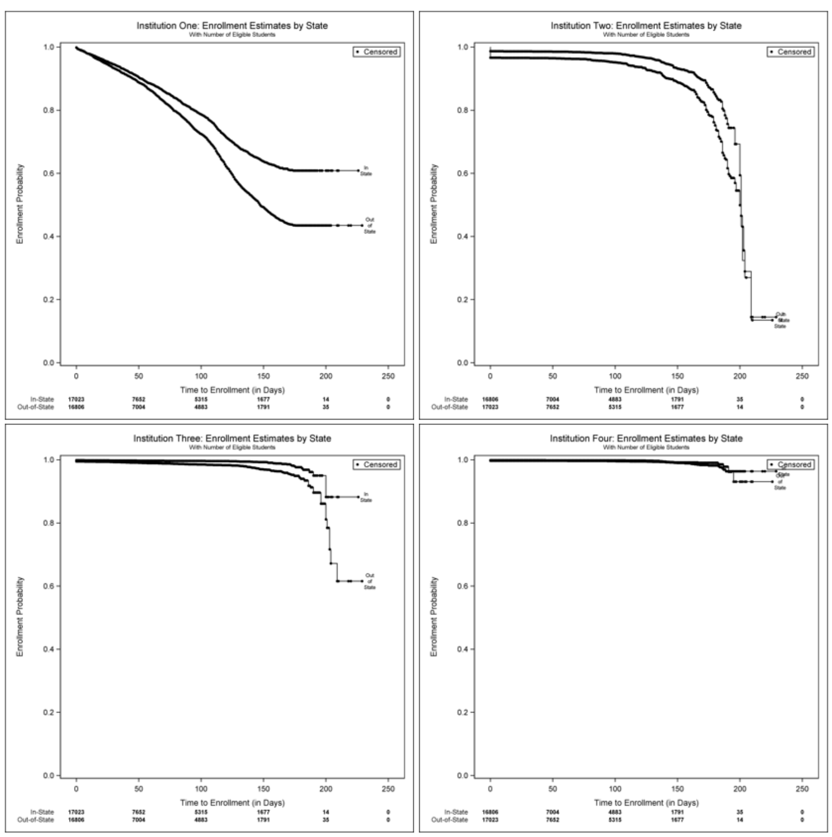
CI: 1.19-1.56, $p < .0001$) more likely to enroll at any given time at institution two. By comparison, students' ethnic identity was not meaningfully associated the odds of enrollment at institution one ($HR = 1.06$, 95% CI: 0.98-1.14, $p = .13$). The Kaplan-Meier results presented in Figure 4 closely align with the model output included in Table 5. For each institution, there was a trend toward increased enrollment probabilities among Hispanic students throughout the application cycle. However, these trends were most pronounced for institutions three and four.

Figure 4. Kaplan Meier Curves for Student Ethnicity



Residency status was also independently associated with an increase in the odds of enrollment at any given time at all four institutions. The most pronounced association was for institution three, followed by institutions two, four, and one, respectively. In-state applicants were most likely to enroll at any given time (*HR* = 3.52, 95% CI: 2.64-4.71, $p < .0001$) at institution three (Table 5). However, in-state applicants were also nearly two times (*HR* = 1.94, 95% CI: 1.73-2.18, $p < .0001$) more likely to enroll at institution two compared to out-of-state applicants. In-state applicants were also 71% (*HR* = 1.71, 95% CI: 1.17-2.49, $p = .01$) and 47% (*HR* = 1.47, 95% CI: 1.39-1.56, $p < .0001$) more likely to enroll at any given time at institutions four and one, respectively. The Kaplan-Meier results provided in Figure 5 reflect the pronounced effect sizes detailed in Table 5. For each institution, there was a clear trend toward increased enrollment probabilities among in-state residents throughout the application cycle.

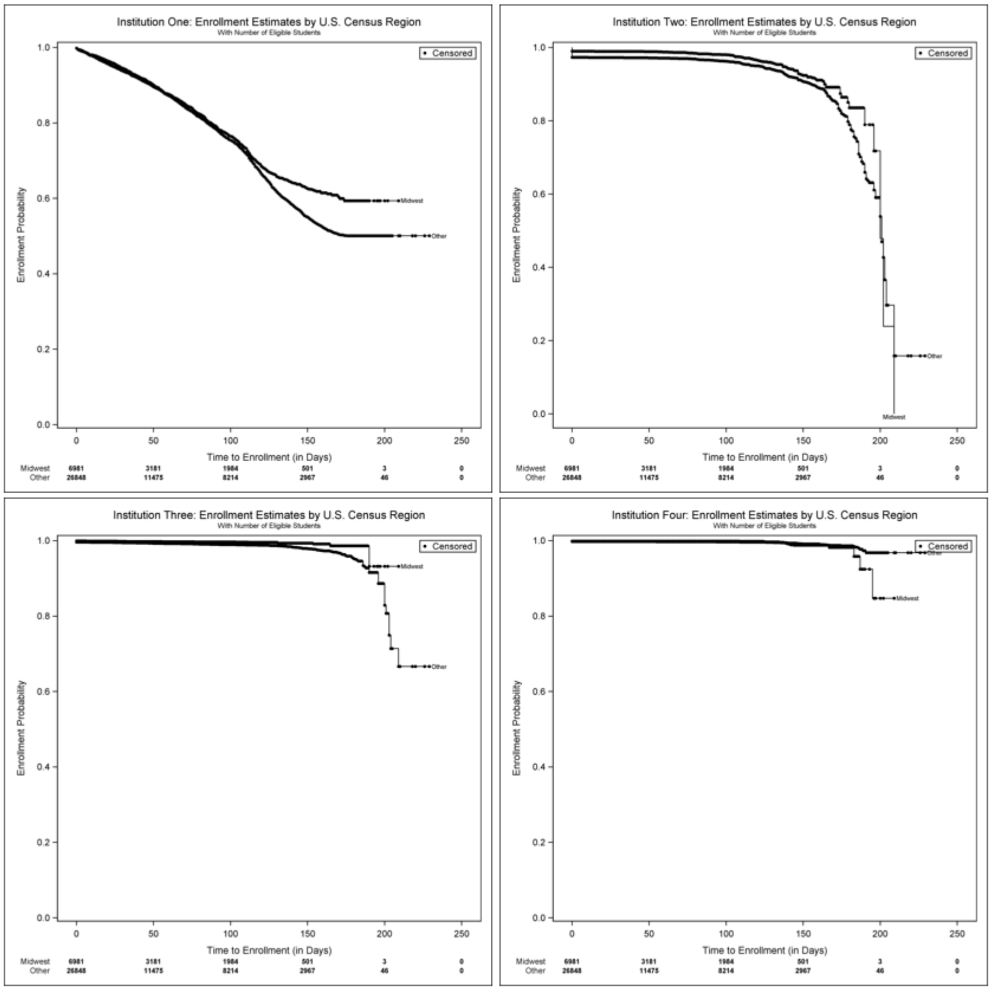
Figure 5. Kaplan Meier Curves for Residency Status



Similar to residency status, univariable results underscored students' strong regional preferences for institutions located in the Midwest. Overall, the region of the country in which students resided was significantly associated with the odds of enrollment at institutions one (p = .004), two (p < .0001), and three (p < .0001), but not four (p = .80). Compared to students from the Midwest, students from the Northeast were 42% (HR = 0.58, 95% CI: 0.41-0.84, p = .004) and 72% (HR = 0.28, 95% CI: 0.09-0.88, p = .03) less likely to enroll at any given time at institutions two and three (Table 5). Similarly, students from the South were 22% (HR = 0.78, 95% CI: 0.68-0.90, p < .001) and 74% (HR = 0.26, 95% CI: 0.0.10-0.69, p = .01) less likely to enroll at any given time at institutions two and three, respectively. Further, students from the West region were 39% (HR = 0.61, 95% CI: 0.47-0.78, p < .001) and 62% (HR = 0.38, 95% CI: 0.19-0.77, p = .01) less likely to enroll at any given time at institutions two and three.

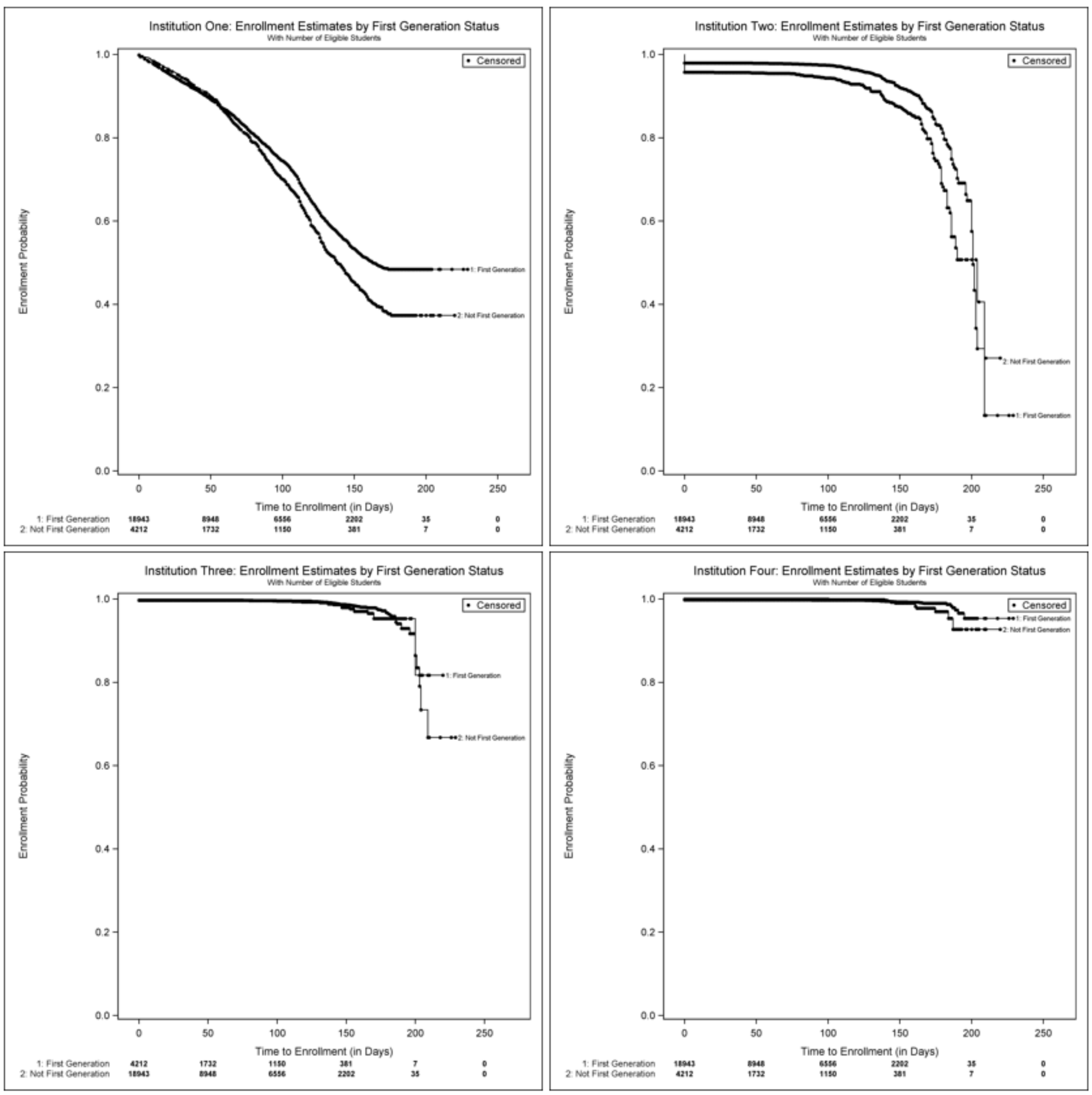
For the Kaplan-Meier results presented in Figure 6, the findings related to U.S. Census Region were simplified by dichotomizing student residency to Midwest v. Other. This allowed for a clearer presentation of how students' enrollment probabilities changed over time based on the region of the country from which they originated. It also aligned with results presented in Table 4, which indicated a clear majority of applicants came from the Midwest. When combined in this manner, broader trends emerged, but the patterns largely mirrored the more detailed findings presented in Table 5. Overall, students from the Midwest were more likely to enroll at institutions one, two, and three, but somewhat less likely to enroll compared to students from other regions at institution four.

Figure 6. Kaplan Meier Curves for U.S. Census Region



In line with other sociodemographic factors, first generation status was also associated with increased odds of enrollment at select institutions. Specifically, first generation status was predictive of the instantaneous odds of enrollment at institutions one ($p < .0001$) and two ($p < .0001$), but not meaningfully associated with outcomes at three ($p = .21$) or four ($p = .74$). Compared to applicants whose parents attended college, first generation students were 21% (*HR* = 1.21, 95% CI: 1.13-1.31, $p < .0001$) more likely to enroll at any given time at institution one (Table 5). By comparison, first generation students were 80% (*HR* = 1.80, 95% CI: 1.57-2.07) more likely to enroll at any given time at institution two ($p < .0001$). The Kaplan-Meier results presented in Figure 7 closely align with the model output included in Table 5. For each institution, there was a trend toward increased enrollment probabilities among first generation students throughout the application cycle.

Figure 7. Kaplan Meier Curves for First Generation Status



Institutional Factors

Table 6. Univariable Models Assessing Institutional Factors

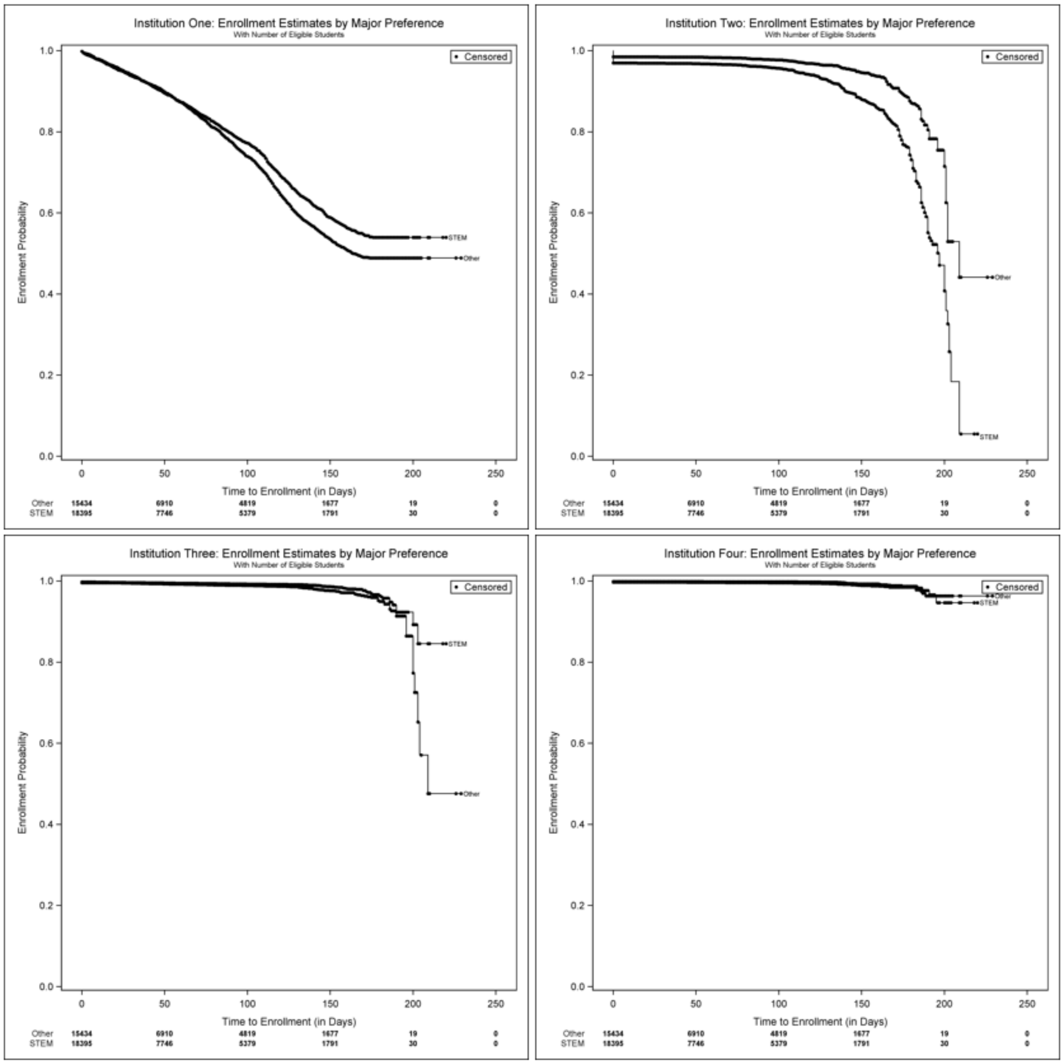
		Institution 1		Institution 2		Institution 3		Institution 4	
	Valid N	**HR (95% CI)**	***p***	**HR (95% CI)**	***p***	**HR (95% CI)**	***p***	**HR (95% CI)**	***p***
Major	31,821		<.0001		<.0001		<.001		.001
Business		1.01 (0.92 - 1.11)	.86	1.62 (1.40 - 1.89)	<.0001	0.53 (0.32 - 0.88)	.01	0.34 (0.14 - 0.85)	.02
Communication		1.05 (0.93 - 1.19)	.42	1.88 (1.56 - 2.27)	<.0001	1.47 (0.89 - 2.44)	.13	0.15 (0.02 - 1.15)	.07
Education		0.95 (0.80 - 1.12)	.51	1.15 (0.86 - 1.53)	.35	0.77 (0.33 - 1.80)	.54	0.30 (0.04 - 2.26)	.24
Liberal Arts		0.93 (0.83 - 1.03)	.15	1.00 (0.82 - 1.21)	.98	0.79 (0.47 - 1.32)	.36	1.86 (1.01 - 3.44)	.05
STEM		1.12 (1.04 - 1.21)	.002	0.56 (0.49 - 0.65)	<.0001	1.38 (1.01 - 1.90)	.04	1.40 (0.86 - 2.28)	.17
Undecided (*Ref*)		-		-		-		-	
First Choice	31,821	6.05 (5.74 - 6.37)	<.0001	4.65 (4.18 - 5.17)	<.0001	1.65 (1.24 - 2.21)	<.0001	2.84 (1.90 - 4.25)	<.0001
# of Applications	31,821	0.95 (0.94 - 0.96)	<.0001	0.97 (0.96 - 0.99)	<.0001	0.96 (0.93 - 0.99)	.02	1.06 (1.01 - 1.11)	.02
Early Outreach	31,821	1.84 (1.74 - 1.94)	<.0001	1.56 (1.41 - 1.74)	<.0001	1.29 (1.02 - 1.64)	.04	1.19 (0.83 - 1.70)	.34

Students' intended major was meaningfully associated with the instantaneous odds of enrollment at each of the four institutions, but the magnitude and direction of the relationship varied across the four universities. Compared to undecided applicants, students interested in business were 62% (*HR* = 1.62, 95% CI: 1.40-1.89, $p < .0001$) more likely to enroll at any given time at institution two (Table 6). By contrast, business students were 47% (*HR* = 0.53, 95% CI: 0.32-0.88, $p = .01$) and 66% (*HR* = 0.34, 95% CI: 0.14-0.85, $p = .02$) less likely to enroll at institutions three and four, respectively. However, a preference for business studies was not meaningfully associated with enrollment at institution one (*HR* = 1.01, 95% CI: 0.92-1.11, $p = .86$). Similarly, students interested in the field of communication were 88% (HR = 1.88, 95% CI: 1.56-2.27) more likely to enroll at any given time at institution two ($p < .0001$). Communication, as a major preference, did not affect students' decision timelines as they relate to institutions one ($p = .42$), three ($p = .13$), or four ($p = .07$).

Students intending to major in education were no more likely to enroll at any of the four institutions: one ($p = .51$), two ($p = .35$), three ($p = .54$), and four ($p = .24$). Nonetheless, those admitted students interested in the liberal arts were 86% (*HR* = 1.86, 95% CI: 1.01-3.44, p = .05) more likely to enroll at any given time at institution four. Similar to communication studies, a liberal arts major did not lead to an increase in the instantaneous odds of enrollment at any of the remaining institutions. Compared to undecided applicants, students who indicated they intended to major in a STEM field were more likely to enroll at institutions one (*HR* = 1.12, 95% CI: 1.04-1.21, $p = .002$) and three (*HR* = 1.38, 95% CI: 1.01-1.90, $p = .04$). By comparison, STEM students were 44% less likely to enroll at institution two (*HR* = 0.56, 95% CI: 0.49-0.65, $p < .0001$). A preference for a STEM major was not significantly associated with time to enrollment at institution four ($p = .17$).

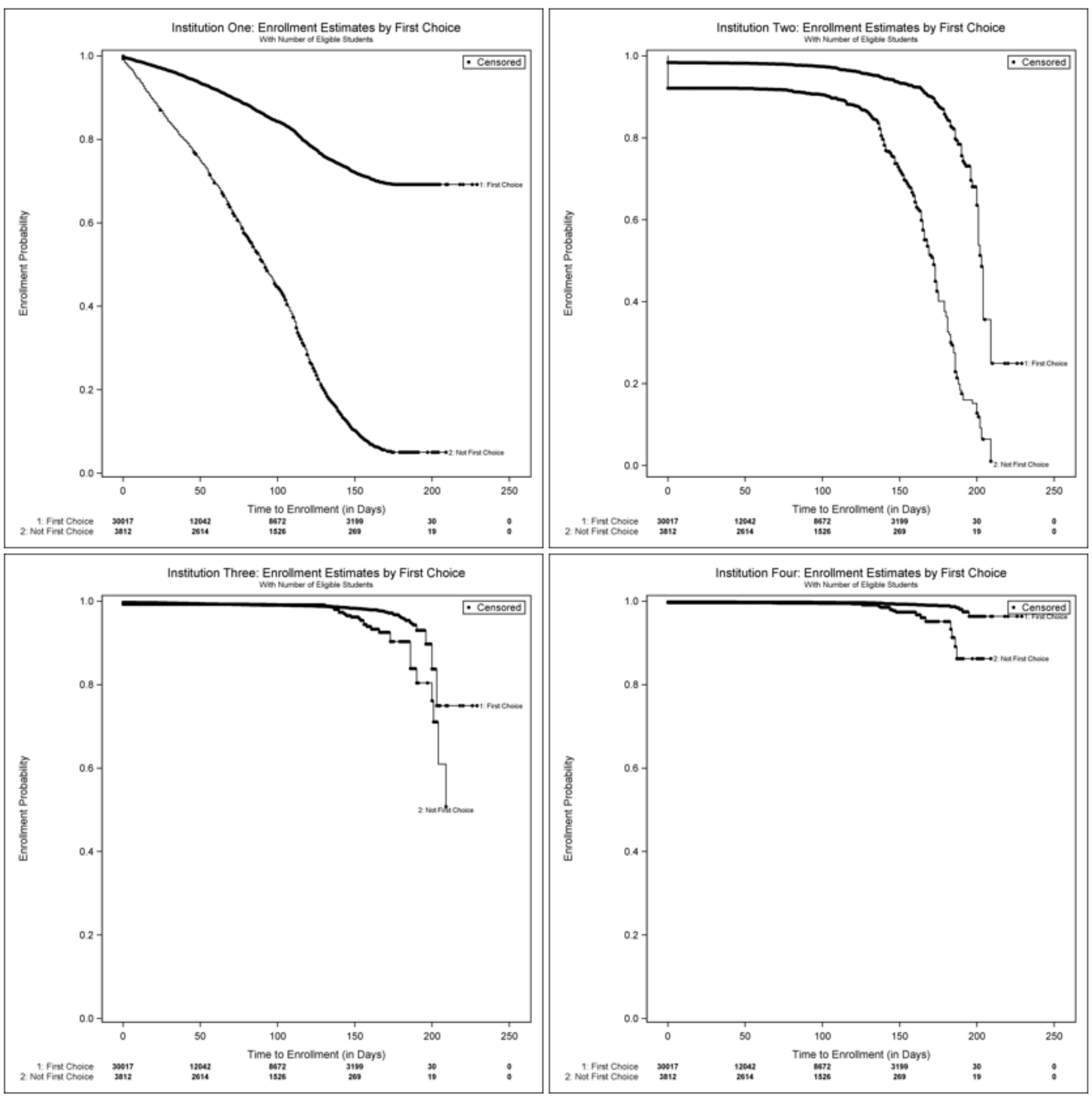
For the Kaplan-Meier results presented in Figure 8, the results related to major preference were simplified by dichotomizing the results to STEM v. non-STEM. This allowed for a clearer presentation of how students' enrollment probabilities changed over time based on major preference. It also aligned with results presented in Table 4, which indicated a clear plurality of applicants indicated a preference for STEM majors. When combined in this manner, broader trends emerged, but the patterns largely mirrored the more detailed findings presented in Table 6. Overall, STEM students were more likely to enroll at institutions one, three, and four, but somewhat less likely to enroll at institution two compared to the students with other major preferences.

Figure 8. Kaplan Meier Curves for Intended Major



In every instance, first choice status was independently associated with an increase in the instantaneous odds of enrollment at all four institutions (Figure 9). For example, students who regarded university one as their "First Choice" school were significantly more likely to enroll if they were admitted. This pattern held for each of the four competitors, but the most pronounced associations were for institutions one and two. First choice applicants were six times (*HR* = 6.05, 95% CI: 5.74-6.37, $p < .0001$) more likely to enroll at institution one, while first choice applicants were nearly five times (*HR* = 4.65, 95% CI: 4.18-5.17, $p < .0001$) more likely to enroll at institution two (Table 6). By comparison, first choice applicants were 2.84 (95% CI: 1.90-4.25, $p < .0001$) times more likely to enroll at institution four and 65% (HR = 1.65, 95% CI: 1.24-2.21, $p < .0001$) more likely to enroll at institution three. The Kaplan-Meier results presented in Figure 9 closely align with the model output included in Table 6. For each institution, there was a pronounced trend toward increased enrollment probabilities among first choice students throughout the application cycle.

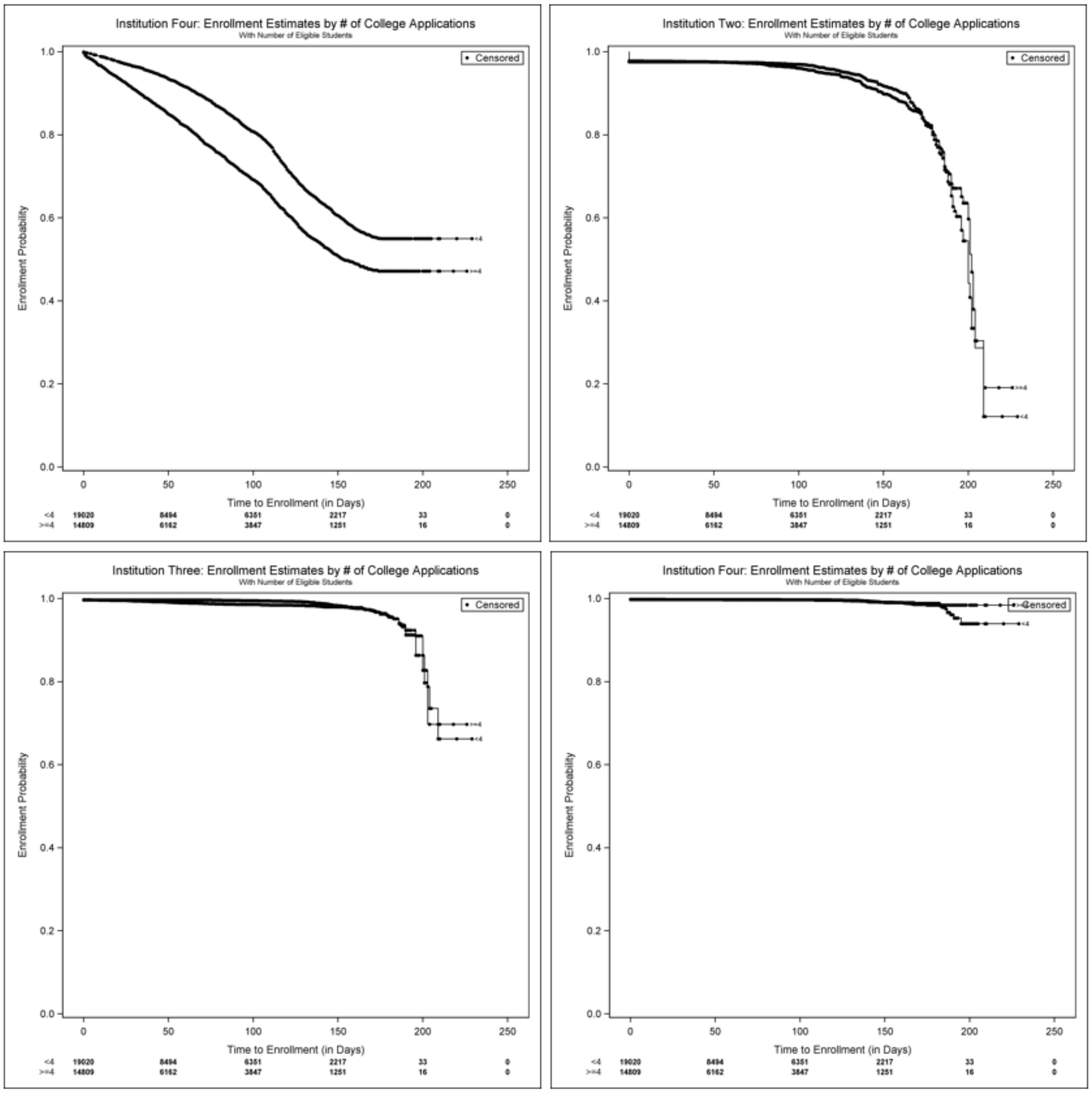
Figure 9. Kaplan Meier Curves for First Choice



The number of applications students submitted was also significantly associated with time to enrollment at each of the four institutions (Figure 10). However, the odds of enrollment at three of the four institutions decreased for every additional application an admitted student submitted. For every one additional application, admitted students were 5% (*HR* = 0.95, 95% CI: 0.94-0.96, $p < .0001$) less likely to enroll at any given time at institution one (Table 6). Similarly, for every one additional application, admitted students were 3% (*HR* = 0.97, 95% CI: 0.96-0.99, $p < .0001$) and 4% (*HR* = 0.96, 95% CI: 0.93-0.99, $p = .02$) less likely to enroll at any given time at institutions two and three, respectively. Only at institution four did the odds of enrollment at any given time increase as the number of applications a student submitted increased. For every one additional application, admitted students were 6% (*HR* = 1.06, 95% CI: 1.01-1.11, $p = .02$) more likely to enroll at any given time at institution four.

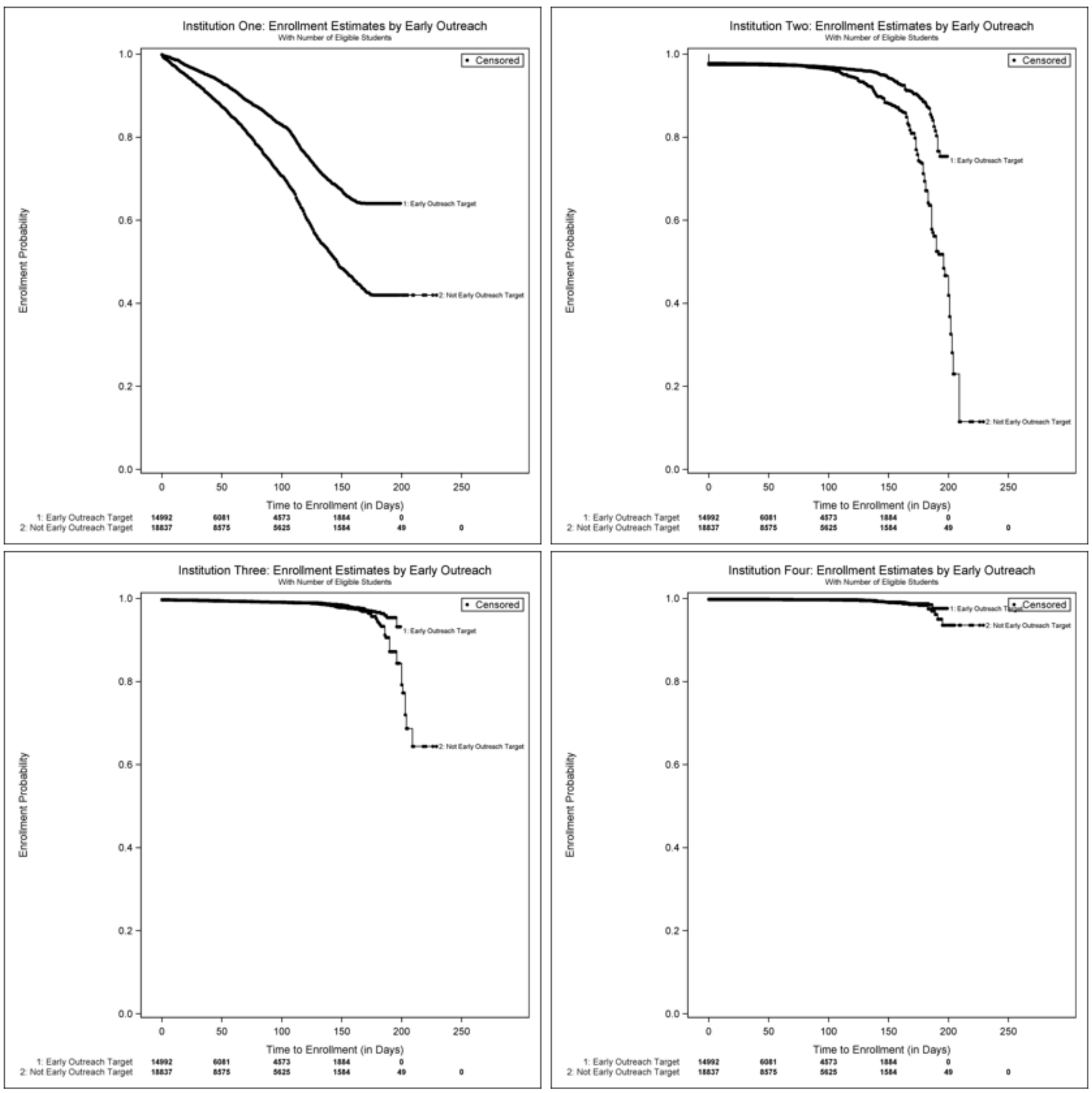
For the Kaplan-Meier results presented in Figure 10, the results related to number of applications students submitted were simplified by dichotomizing the results using the sample median as the cut point. This allowed for a clearer presentation of how students' enrollment probabilities changed over time based on number of applications they submitted. It also aligned with results presented in Table 4. Overall, students who submitted more applications were less likely to enroll at institutions one, two, and three, but somewhat more likely to enroll at institution four compared to those students who submitted fewer applications.

Figure 10. Kaplan Meier Curves for # of Applications (Cut at Median)



In line with other institutional factors, early outreach was also an important indicator of the odds of enrollment at any given time at three of the four institutions (Figure 11). Early outreach targets were 84% (HR = 1.84, 95% CI: 1.74-1.94, $p < .0001$) more likely to enroll at institution one (Table 6). Comparatively, early outreach targets were 56% (HR = 1.56, 95% CI: 1.41-1.74, $p < .0001$) and 29% (HR = 1.29, 95% CI: 1.02-1.64, $p < .0001$) more likely to enroll at institutions two and three, respectively. However, early outreach efforts were not independently associated with the instantaneous odds of enrollment at institution four despite a positive trend (HR = 1.19, 95% CI: 0.83-1.70, $p = .34$). The Kaplan-Meier results presented in Figure 11 closely align with the model output included in Table 6. For three of the four institutions, there was a pronounced trend toward increased enrollment probabilities among students who were the targets of early outreach throughout the application cycle.

Figure 11. Kaplan Meier Curves for Early Outreach



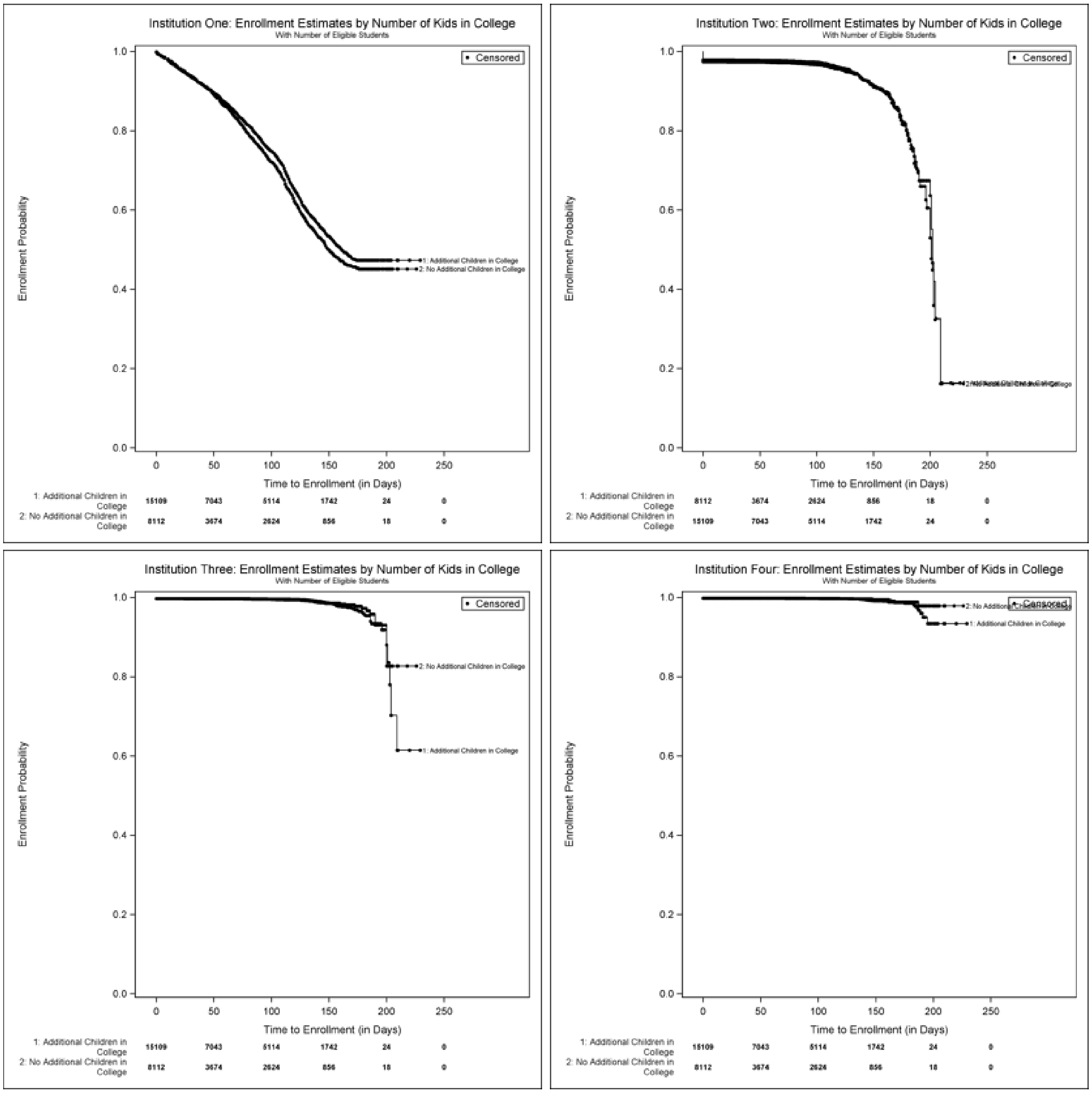
Financial Factors

Table 7. Univariable Models Assessing Financial Factors

	Valid N	Institution 1 HR (95% CI)	p	Institution 2 HR (95% CI)	p	Institution 3 HR (95% CI)	p	Institution 4 HR (95% CI)	p
# Kids in College	22,287	1.09 (1.03 - 1.16)	.003	0.92 (0.81 - 1.04)	.19	0.93 (0.67 - 1.29)	.68	0.80 (0.50 - 1.29)	.36
Pell Grant Eligible	30,599	1.34 (1.26 - 1.42)	<.0001	1.43 (1.27 - 1.61)	<.0001	1.48 (1.11 - 1.98)	.01	0.87 (0.55 - 1.39)	.57
Merit Aid	31,821	1.44 (1.34 - 1.55)	<.0001	2.34 (2.07 - 2.64)	<.0001	3.62 (1.59 - 25.83)	<.001	16.17 (2.26 - 115.79)	.01

Univariable model results confirmed the importance of select financial factors in students' final enrollment decisions as well. Specifically, admitted students whose families reported additional children in college was an important indicator of enrollment at institution one (Figure 12). Admitted students who came from families with more than one child in college were 9% (*HR* = 1.09, 95% CI: 1.03-1.16, *p* = .003) more likely to enroll at any given time at institution one (Table 7). By comparison, the number of children a family had in college was not meaningfully associated with the odds of enrollment at institutions two (*p* = .19), three (*p* = .68), or four (*p* = .36). In each instance, however, there was a trend toward reduced odds of enrollment: two (*HR* = 0.92, 95% CI: 0.81-1.04), three (*HR* = 0.93, 95% CI: 0.67-1.29), or four (*HR* = 0.80, 95% CI: 0.50-1.29). The Kaplan-Meier results presented in Figure 12 indicate how students' enrollment probabilities changed over time based on the number of additional children the applicant's family has in college. These paneled findings align closely with the model output included in Table 7. For institution one, applicants from families with additional children in college recorded higher enrollment probabilities throughout the application timeline. By contrast, this trend flipped for institutions two, three, and four for which there was less separation between the two applicant subgroups.

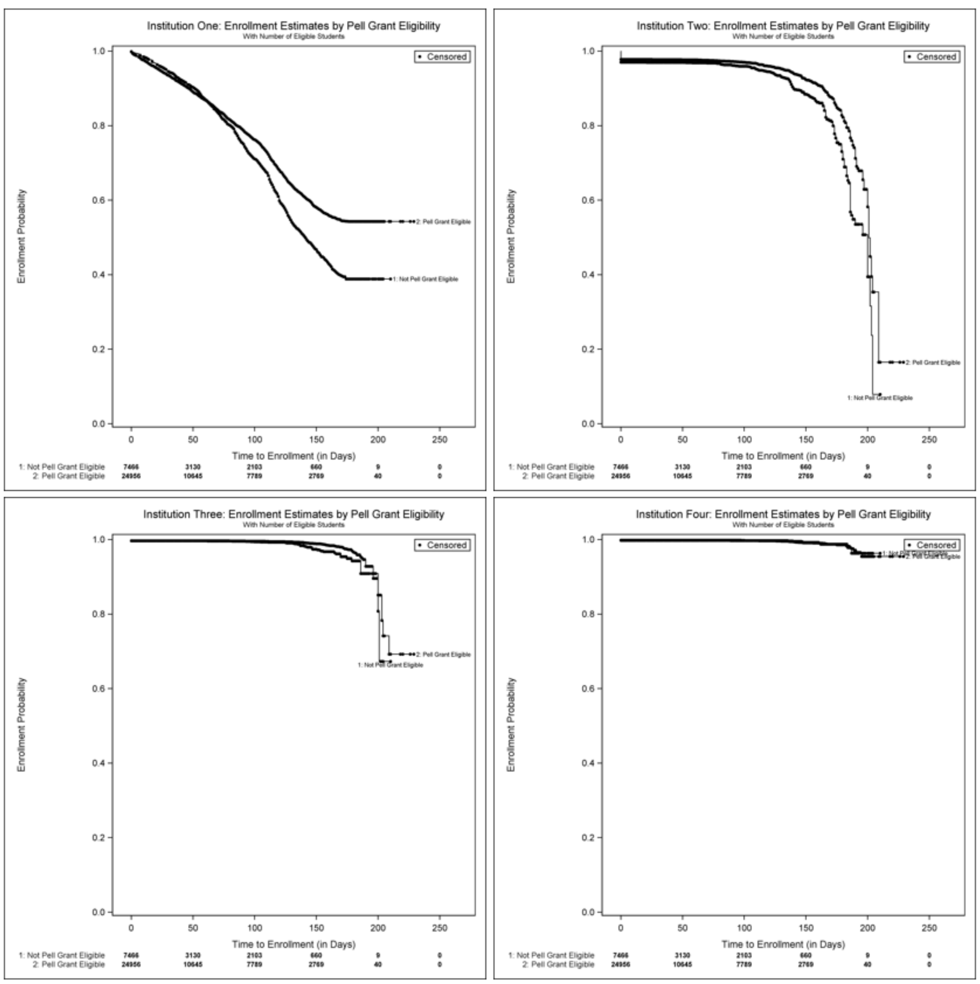
Figure 12. Kaplan Meier Curves for Additional Children in College



Pell Grant eligibility, as determined by students' derived estimated family contribution on their FAFSA submission, was also an independent predictor of the instantaneous odds of enrollment at three of the four institutions under review (Figure 13). Pell Grant eligible students were 34% (*HR* = 1.34, 95% CI: 1.26-1.42, $p < .0001$) more likely to enroll at institution one

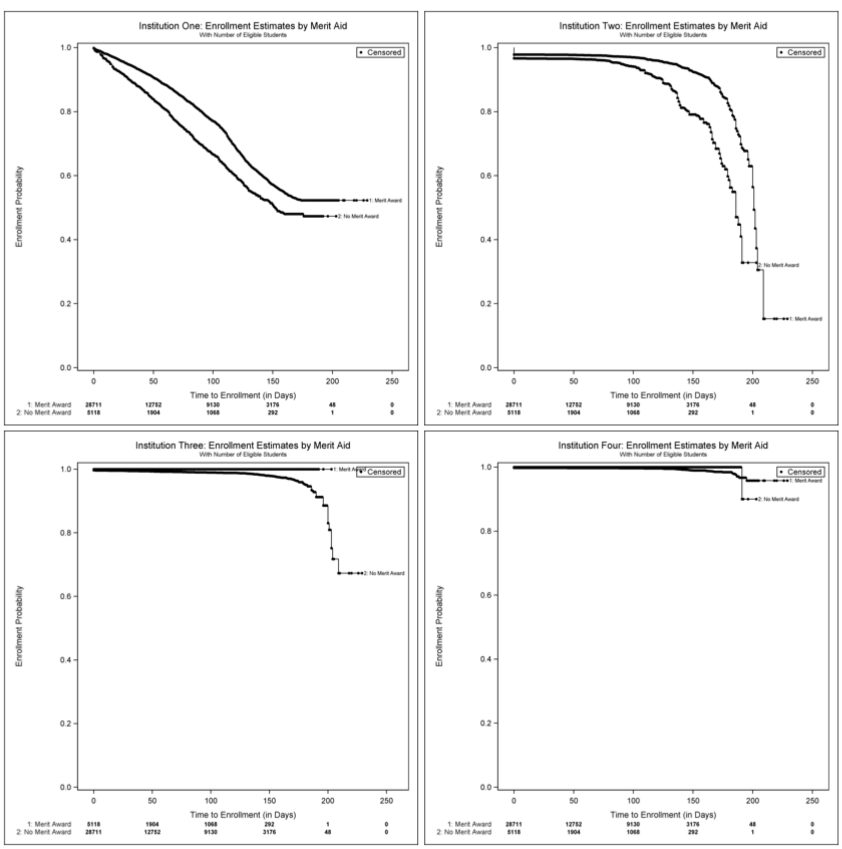
(Table 7). Similarly, Pell Grant eligible students had nearly equal odds of enrollment at institutions two (*HR* = 1.43, 95% CI: 1.27-1.61, $p < .001$) and three (*HR* = 1.48, 95% CI: 1.11-1.98, $p = .01$) over time. By contrast, Pell Grant eligibility was not meaningfully associated with time to enrollment at institution four (*HR* = 0.87, 95% CI: 0.55-1.39, $p = .57$). The Kaplan-Meier results are presented in Figure 13. For institutions one, two, and three there was a trend toward increased enrollment probabilities among Pell Grant eligible students throughout the application cycle. By contrast, this relationship was flipped, but also less pronounced for institution four.

Figure 13. Kaplan Meier Curves for Pell Grant Eligibility



In addition, the offer of merit aid was significantly associated with increased odds of enrollment at any given time at all four institutions (Figure 14). This effect was most pronounced at institution four (*HR* = 16.17, 95% CI: 2.26-115.79, *p* = .01) where the offer of merit aid resulted in a sixteen-fold increase in the instantaneous odds of enrollment (Table 7). Students offered merit aid were also 44% (*HR* = 1.44, 95% CI: 1.34-1.55) more likely to enroll at institution one ($p < .0001$). By comparison, offers of merit aid also doubled and nearly quadrupled the instantaneous odds of enrollment at institutions two (*HR* = 2.34, 95% CI: 2.07-2.64) and three (*HR* = 3.62, 95% CI: 1.59-25.83), respectively. The Kaplan-Meier results presented in Figure 14 closely align with the model output included in Table 7. For each institution, there was a pronounced trend toward increased enrollment probabilities among merit aid recipients throughout the application cycle.

Figure 14. Kaplan Meier Curves for Merit Aid



Academic Factors

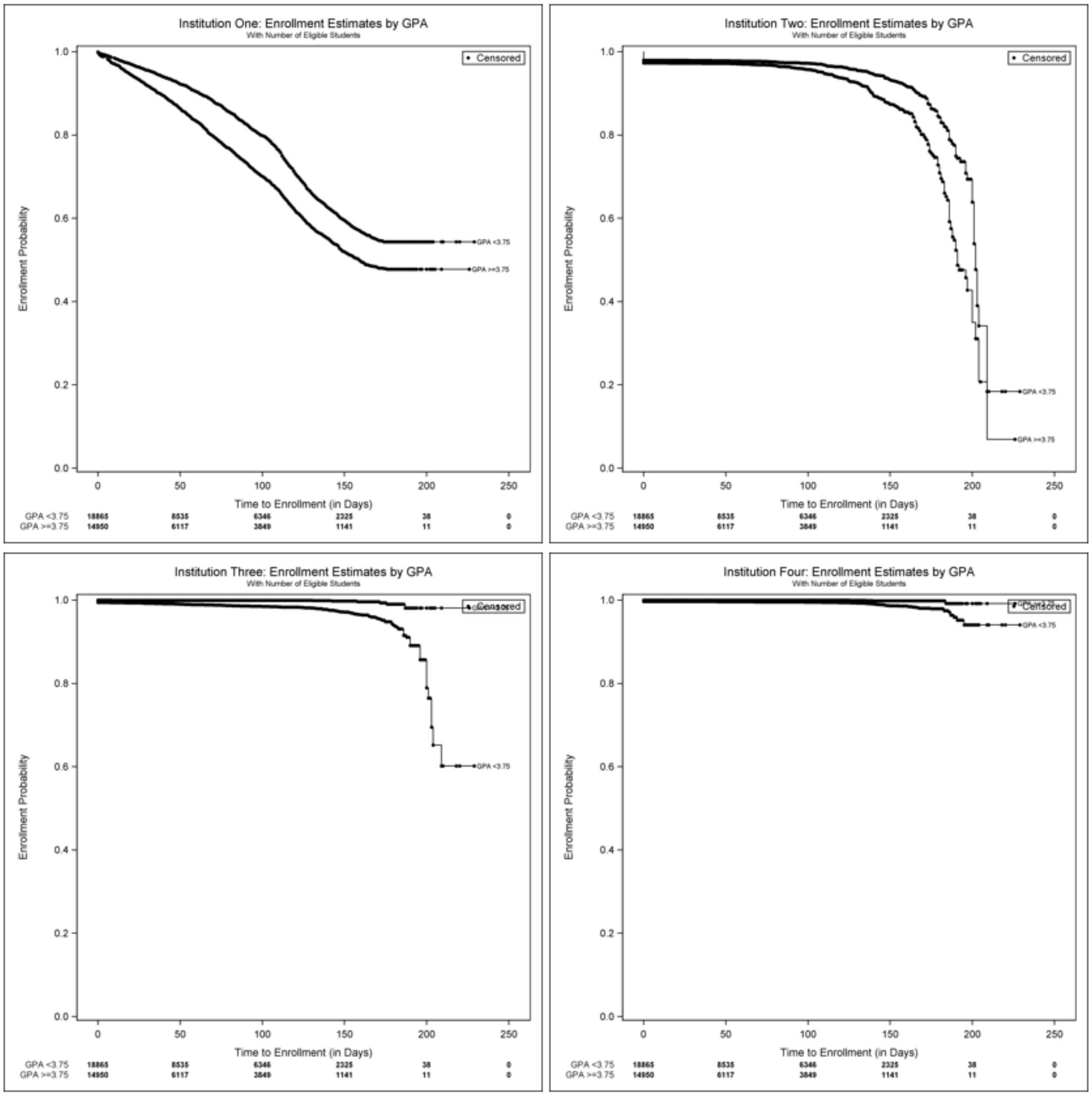
Table 8. Univariable Models Assessing Academic Factors

	Valid N	Institution 1 HR (95% CI)	p	Institution 2 HR (95% CI)	p	Institution 3 HR (95% CI)	p	Institution 4 HR (95% CI)	p
GPA (*Unit* =0.5)	31,810	0.79 (0.76 - 0.81)	<.0001	0.71 (0.67 - 0.75)	<.0001	2.42 (2.22 - 2.62)	<.0001	2.26 (2.01 - 2.55)	<.0001
ACT (*Unit* =4)	31,702	0.74 (0.72 - 0.77)	<.0001	0.60 (0.56 - 0.63)	<.0001	4.94 (4.18 - 5.84)	<.0001	6.02 (4.58 - 7.92)	<.0001

An applicant's cumulative grade point average (GPA) was also significantly associated with the instantaneous odds of enrollment at each of the four institutions (Figure 15). Nonetheless, the direction of this relationship was different for institutions one and two compared to institutions three and four. For every one standard deviation (units=0.5) increase in students' cumulative GPA, admitted students were 21% (HR = 0.79, 95% CI: 0.76-0.81, p < .0001) and 29% (HR = 0.71, 95% CI: 0.67-75, p < .0001) less likely to enroll at any given time at institutions one and two, respectively (Table 8). By contrast, for every half unit increase in students' cumulative GPA, admitted students were over two times more likely to enroll at any given time at institutions three (*HR* =2.42, 95% CI: 2.22-2.62, p < .0001) and four (*HR* =2.26, 95% CI: 2.01-2.55, p < .0001).

For the Kaplan-Meier results presented in Figure 15, the findings related to student GPA were simplified by dichotomizing the results using the sample median as the cut point. This allowed for a clearer presentation of how students' enrollment probabilities changed over time based on their GPA. It also aligned with results presented in Table 4. When combined in this manner, broader trends emerged, but the patterns largely mirrored the more detailed findings presented in Table 8. Overall, students with higher GPAs were less likely to enroll at institutions one and two, but more likely to enroll at institutions three and four compared to those students with lower GPAs.

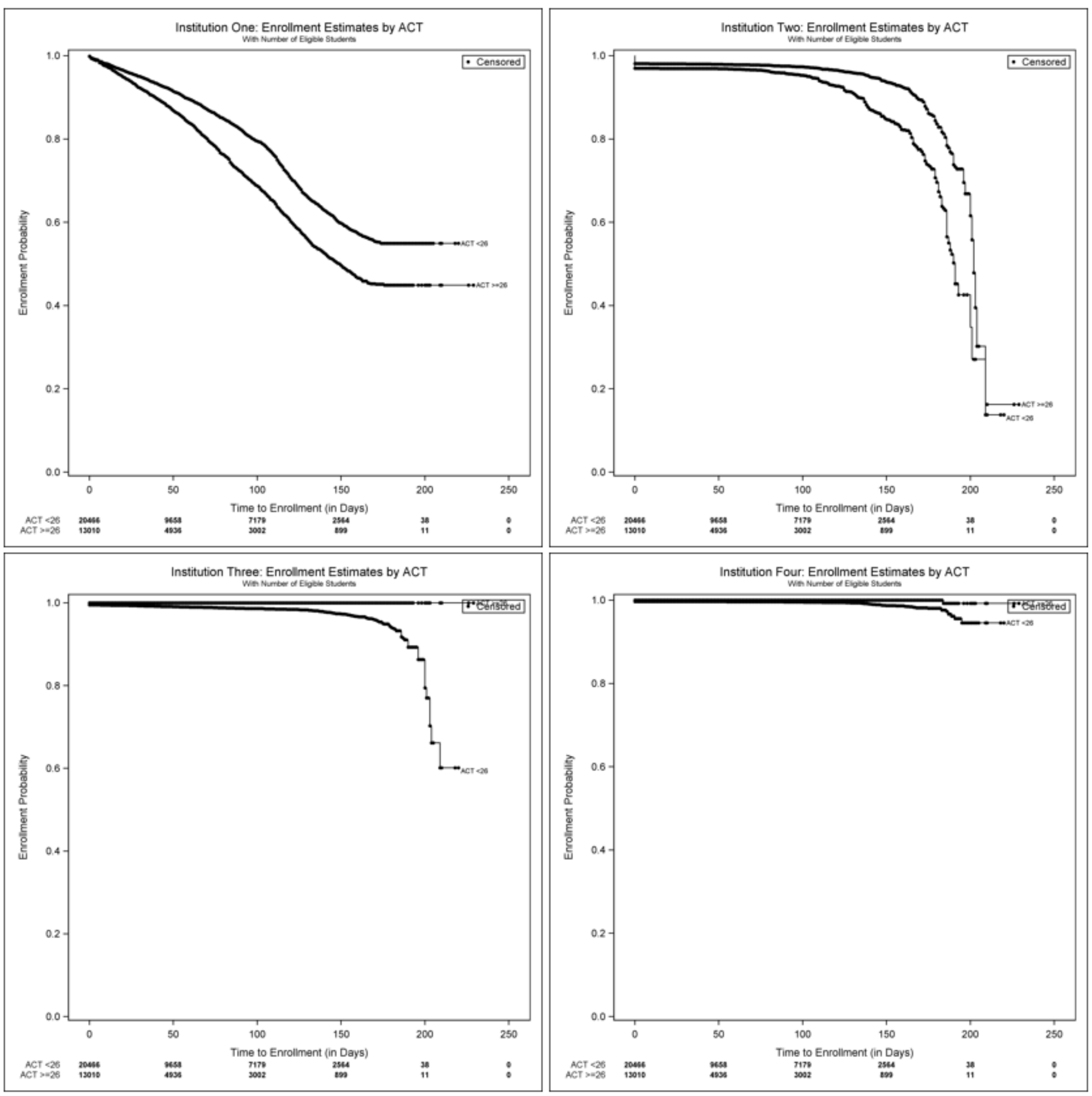
Figure 15. Kaplan Meier Curves for Cumulative GPA (Cut at Median)



A similar pattern to students' GPA results emerged for their standardized test scores as reported by the ACT (Figure 16). For every one standard deviation (units=4) increase in students' ACT scores, admitted students were 26% (HR = 0.74, 95% CI: 0.72-0.77, p < .0001) and 40% (HR = 0.60, 95% CI: 0.56-0.63, p < .0001) less likely to enroll at any given time at institutions one and two, respectively (Table 8). By contrast, for every four-unit increase in students' ACT scores, admitted students were five to six times more likely to enroll at any given time at institutions three (HR = 4.94, 95% CI: 4.18-5.84, p < .0001) and four (HR = 6.02, 95%

CI: 4.58-7.92, $p < .0001$). For the Kaplan-Meier results presented in Figure 16, the results related to student ACT scores were simplified by dichotomizing the results using the sample median as the cut point. This allowed for a clearer presentation of how students' enrollment probabilities changed over time based on their ACT scores. It also aligned with results presented in Table 4. Overall, students with higher ACT scores were less likely to enroll at institutions one and two, but more likely to enroll at institutions three and four compared to those students with lower ACT scores.

Figure 16. Kaplan Meier Curves for ACT (Cut at Median)



Multivariable Analysis

Any sociodemographic, institutional, financial, or academic factors found to be significant on univariable analysis were considered for inclusion in the final multivariable model. Multicollinearity diagnostics, including Variance Inflation Factor (VIF) and correlation estimates (see Appendix B), were assessed to identify and remove correlated independent variables in the final analysis (Weisberg, 2005). In the final multivariable model, only the "Region" variable was excluded due to its overlap with residency status. Based on the univariable results and theoretical considerations, all remaining explanatory variables were initially evaluated. For interpretation purposes, the effect of each individual variable included in the final analysis should be interpreted as though one is holding all other variables in the model line constant (Table 9).

For the formal model building process, AIC values for sequential models were assessed to ensure the results reflected the most parsimonious predictive estimates. Utilizing AIC avoids overfitting models, which results in the estimated values of some of the beta coefficients being highly dependent on the actual data, thus limiting the generalizability of the results. An additional benefit of comparisons on the basis of AIC is that sequential models need not be nested. For interpretation purposes, smaller AIC values indicate a better fitting model. Specifically, when AIC decreases by more than two points upon removing an independent variable, the results indicate the more parsimonious model may provide better estimates of the true expected values. In the event the AIC value remains unchanged or increases, the omitted variable should be retained in the final analysis as the more complex model provides a better approximation of the true relationship between the parameters.

Table 9. Multivariable Analysis

	Institution 1		Institution 2		Institution 3		Institution 4	
	Adjusted HR (95% CI)	***p***	**Adjusted HR (95% CI)**	***p***	**Adjusted HR (95% CI)**	***p***	**Adjusted HR (95% CI)**	***p***
Sex (*Ref*= 'Female')	1.14 (1.06 - 1.21)	<.001						
Race		.002		.02		<.0001		<.0001
Asian	1.18 (1.09 - 1.28)	<.0001	0.73 (0.59 - 0.89)	.003	1.89 (1.15 - 3.08)	.01	0.82 (0.43 - 1.57)	.55
Black or African American	1.01 (0.86 - 1.17)	.93	0.75 (0.54 - 1.03)	.07	13.99 (8.05 - 24.32)	<.0001	12.10 (6.53 - 22.43)	<.0001
Caucasian (*Ref*)	-		-		-		-	
Multiracial	1.06 (0.94 - 1.20)	.36	0.91 (0.68 - 1.22)	.53	1.47 (0.67 - 3.22)	.33	2.21 (1.14 - 4.28)	.02
Other	0.91 (0.67 - 1.24)	.55	1.18 (0.71 - 1.95)	.52	3.30 (1.27 - 8.59)	.01	0.98 (0.13 - 7.17)	.98
Ethnicity					7.68 (5.07 - 11.64)	<.0001	5.63 (3.54 - 8.96)	<.0001
Residency	0.84 (0.80 - 0.88)	<.001	1.58 (1.35 - 1.86)	<.0001	1.60 (1.06 - 2.40)	.03		
First Generation			1.41 (1.20 - 1.66)	<.0001				
Major		<.0001		<.0001		.12		
Business	0.96 (0.86 - 1.07)	.47	1.65 (1.36 - 1.99)	<.0001	0.58 (0.26 - 1.29)			
Communication	1.03 (0.89 - 1.18)	.76	1.84 (1.45 - 2.33)	<.0001	1.93 (0.91 - 4.10)			
Education	1.05 (0.87 - 1.27)	.59	1.05 (0.73 - 1.52)	.79	0.19 (0.01 - 3.34)			
Liberal Arts	1.02 (0.90 - 1.15)	.75	1.14 (0.89 - 1.45)	.31	1.00 (0.50 - 1.98)			
STEM	1.22 (1.12 - 1.32)	<.0001	0.62 (0.52 - 0.75)	<.0001	0.79 (0.49 - 1.25)			
Undecided (*Ref*)	-		-		-		-	
First Choice	5.32 (4.98 - 5.69)	<.0001	5.47 (4.75 - 6.29)	<.0001	4.50 (3.07 - 6.59)	<.0001	4.36 (2.86 - 6.66)	<.0001
# of Applications	0.92 (0.91 - 0.93)	<.0001	1.05 (1.02 - 1.08)	<.001	1.09 (1.02 - 1.17)	.02	1.00 (0.95 - 1.05)	.91
Early Outreach	1.42 (1.33 - 1.51)	<.0001	1.25 (1.09 - 1.43)	.002	1.61 (1.11 - 2.33)	.01		
# Kids in College	1.06 (0.99 - 1.13)	.06	1.04 (0.91 - 1.19)	.55	1.35 (0.94 - 1.96)	.10		
Pell Grant Eligible	0.93 (0.87 - 0.99)	.04	1.14 (0.98 - 1.32)	.09	2.17 (1.47 - 3.13)	<.0001		
Merit Aid	1.05 (0.95 - 1.17)	.41	0.74 (0.61 - 0.91)	.01	7.92 (0.48 - 129.52)	.95	3.19 (0.44 - 22.99)	.25
GPA (*Unit*=0.5)	0.90 (0.87 - 0.94)	<.0001	0.95 (0.87 - 1.03)	.22	1.81 (1.54 - 2.12)	<.0001	1.58 (1.31 - 1.90)	<.0001
ACT (*Unit*=4)	0.84 (0.80 - 0.88)	<.0001	0.76 (0.69 - 0.83)	<.0001	6.55 (4.87 - 8.81)	<.0001	7.12 (5.11 - 9.92)	<.0001

Note: Valid N = 19,649.

Sociodemographic Factors

After controlling for students' race, ethnicity, residency, first generation status, major preference, admission into a first choice school, number of applications, early outreach, number of siblings in college, Pell Grant eligibility, merit aid, GPA, and ACT, sex was still a significant determinant in the odds of enrollment at any given time at institution one. Male students were 14% (HR = 1.14, 95% CI: 1.06-1.21) more likely to enroll at any given time at institution one compared to female applicants ($p < .001$). By contrast after controlling for all additional covariates, a student's sex was no longer meaningfully associated with the instantaneous odds of enrollment at institution four. In line with univariable findings, male students' odds of enrollment did not significantly increase for either institutions two or three. Further, for the models assessing institutions two, three, and four, the variable sex was removed from the final analyses based on an evaluation of fit statistics (AIC).

On multivariable analysis, students' race remained significantly associated with the odds of enrollment at all four institutions over time. As before, though, the magnitude and direction of the relationship varied widely across each institutions. After controlling for students' sex, ethnicity, residency, first generation status, major preference, admission into a first choice school, number of applications, early outreach, number of siblings in college, Pell Grant eligibility, merit aid, GPA, and ACT, Asian students were 18% (*HR* = 1.18, 95% CI: 1.09-1.28, $p < .0001$) and 89% (*HR* = 1.89, 95% CI: 1.15-3.08, $p = .01$) more likely to enroll at any given time at institutions one and three compared to Caucasian students. By contrast, after controlling for other important factors, Asian students were 27% less likely to enroll at institution two (*HR* = 0.73, 95% CI: 0.59-0.89, $p = .003$).

Adjusting for other covariates, Black or African American students were 13.99 (95% CI: 8.05-24.32, $p < .0001$) and 12.10 (95% CI: 6.53-22.43, $p < .0001$) times more likely than their Caucasian counterparts to enroll at any given time at institutions three and four, respectively. Multi-racial students' instantaneous odds of enrollment also increased 2.21 (95% CI: 1.14-4.28) times at institution four after controlling for select sociodemographic, institutional, financial, and academic factors ($p = .02$). By contrast, Black or African American students were marginally ($HR = 0.75$, 95% CI: 0.54-1.03, $p = .07$) less likely to enroll at any given time at institution two compared to Caucasian students. On multivariable analysis, students who identified with 'Other' racial groups were 3.30 (95% CI: 1.27-8.59) times more likely to enroll at any given time at institution three compared to Caucasian students ($p = .01$).

Controlling for select sociodemographic, institutional, financial, and academic factors, students' ethnicity emerged as a meaningful indicator of time to enrollment at two of the four institutions. Hispanic students were over seven times ($HR = 7.68$, 95% CI: 5.07-11.64) more likely to enroll at any given time at institution three ($p < .0001$). Hispanic students were also over five times ($HR = 5.63$, 95% CI: 3.54-8.96) more likely to enroll at any given time at institution four ($p < .0001$). However, on multivariable analysis, ethnicity was no longer meaningfully associated with time to enrollment at institution two. In line with univariable results, students' ethnic identity was also not meaningfully associated the time to enrollment at institution one. For the models assessing institutions one and two, the variable ethnicity was removed from the final analyses based evaluation of the fit statistics.

After controlling for students' sex, race, ethnicity, first generation status, major preference, admission into a first choice school, number of applications, early outreach, number of siblings in college, Pell Grant eligibility, merit aid, GPA, and ACT, residency status was still significantly associated with an increase in the odds of enrollment at any given time at three of the four institutions. In-state applicants were 58% (*HR* = 1.58, 95% CI: 1.35-1.86, $p < .0001$) more likely to enroll at institution two at any given time compared to out of state students. Similarly, in-state applicants were also 60% (*HR* = 1.60%, 95% CI: 1.06-2.40, $p = .03$) more likely to enroll at institution three. However, after adjusting for other covariates, in-state applicants were 16% (*HR* = 0.84, 95% CI: 0.80-0.88, $p < .001$) less likely to enroll at any given time at institution one. For the model assessing institution four, the variable residency was removed from the final analysis based on an evaluation of fit statistics.

Controlling for the aforementioned sociodemographic, institutional, financial, and academic factors, first generation status only remained a significant predictor of time to enrollment at institution two. On multivariable analysis, first generation status was not meaningfully associated with enrollment outcomes over time at intuitions one, three, or four. In each instance, fit statistics were evaluated and the variable first generation was removed from the final analyses. Compared to applicants whose parents attended college, first generation students were still 41% (*HR* = 1.41, 95% CI: 1.20-1.66, $p < .0001$) more likely to enroll at any given time at institution two, even after adjusting for additional covariates.

Institutional Factors

After controlling for students' sex, race, ethnicity, residency, first generation status, admission into a first choice school, number of applications, early outreach, number of siblings in college, Pell Grant eligibility, merit aid, GPA, and ACT, students' major preference remained meaningfully associated with the instantaneous odds of enrollment at two of the four institutions.

Compared to undecided applicants, students interested in business were 65% (HR = 1.65, 95% CI: 1.36-1.99, $p < .0001$) more likely to enroll at any given time at institution two. After adjusting for important covariates, students interested in the field of communications were also 84% (HR = 1.84, 95% CI: 1.45-2.33) more likely to enroll at any given time at institution two ($p < .0001$). By contrast, a preference for business or communication studies was not meaningfully associated with enrollment at any of the other institutions.

Controlling for select factors, education and liberal arts majors were no longer predictive of time to enrollment at any of the four institutions, compared to undecided students. However, an important difference between outcomes at two of the institutions was the magnitude and direction of the association between STEM designated fields. On multivariable analysis, students who indicated they intended to major in a STEM field were 22% more likely to enroll at institution one ($HR = 1.22$, 95% CI: 1.12-1.32, $p < .0001$). By contrast, STEM students were 38% less likely to enroll at any given time at institution two ($HR = 0.62$, 95% CI: 0.52-0.75, $p < .0001$). A preference for a STEM major was not significantly associated with time to enrollment at institutions three or four. The variable major preference was removed from the final model assessing institution four based on an evaluation of the fit statistics.

Even after adjusting for all other covariates, first choice status remained significantly associated with an increase in the instantaneous odds of enrollment at each of the four institutions. In every instance, first choice designation was associated with a four to five-fold increase in the instantaneous odds of enrollment among admitted students across the four universities. Unlike univariable results in which institutions one and two clearly separated from three and four, each of the competitors exhibited a pronounced association with first choice preference in the multivariable model. First choice applicants were over five times more likely to enroll at institutions one (*HR* = 5.32, 95% CI: 4.98-5.69, $p < .0001$) and two (*HR* = 5.47, 95% CI: 4.75-6.29, $p < .0001$). Similarly, first choice applicants were 4.50 (95% CI: 3.07-6.59, $p < .0001$) times more likely to enroll at institution three and 4.36 (95% CI: 2.86-6.66, $p < .0001$) times more likely to enroll at institution four.

Similarly, the number of applications students submitted remained significantly associated with the odds of enrollment at three of the four institutions after controlling for important covariates. While the odds of enrollment decreased at institution one for every additional application an admitted student submitted, it increased at institutions two and three. For every additional application, admitted students were 5% (*HR* = 1.05, 95% CI: 1.02-1.08, $p < .001$) more likely to enroll at any given time at institution two. Similarly, on multivariable analysis, for every additional submitted application, admitted students were 9% (*HR* = 1.09, 95% CI: 1.02-1.17) more likely to enroll at any given time at institutions three ($p = .02$). By contrast, for every additional application, admitted students were 8% (*HR* = 0.92, 95% CI: 0.91-0.93) less likely to enroll at any given time at institution one ($p < .0001$). Only at institution four did the odds of enrollment at any given time remain unchanged as the number of applications a student submitted increased after adjusting for important covariates.

In line with other institutional factors, early outreach emerged as an important indicator of the odds of enrollment at any given time at three of the four institutions on univariable analysis. This association remained unchanged even after controlling for students’ sex, race, ethnicity, residency, first generation status, major preference, admission into a first choice school, number of applications, number of siblings in college, Pell Grant eligibility, merit aid, GPA, and ACT. On multivariable analysis, early outreach targets were 42% (*HR* = 1.42, 95% CI: 1.33-1.51) more likely to enroll at institution one ($p < .0001$). Comparatively, early outreach targets were 25% (*HR* = 1.25, 95% CI: 1.09-1.43, $p < .002$) and 61% (*HR* = 1.61, 95% CI: 1.11-2.33, $p = .01$) more likely to enroll at institutions two and three, respectively. As in univariable analysis, early outreach efforts were not meaningfully associated with the instantaneous odds of enrollment at institution four after adjusting for select covariates. In fact, for the model assessing institution four, the variable early outreach was removed from the final analysis based on an evaluation of the fit statistics.

Financial Factors

Unlike univariable findings that confirmed the importance of select financial factors in students' final enrollment decisions, multivariable model results found their effects substantially moderated. After controlling for select sociodemographic, institutional, financial, and academic factors, admitted students whose families reported additional children in college was no longer an important indicator of time to enrollment at any of the institutions. Despite this finding, there remained a marginal association for students who enrolled at institutions one and three. Admitted students who came from families with more than one child in college were 6% (*HR* = 1.06, 95% CI: 0.99-1.13, $p = .06$) more likely to enroll at any given time at institution one. By comparison, similar admitted students were 35% (*HR* = 1.35, 95% CI: 0.94-1.96, $p = .10$) more likely to enroll at any given time at institution three. For the model assessing institution four, the indicator variable for the number of children a family had in college was removed from the final analysis based on an evaluation of fit statistics.

After controlling for students' sex, race, ethnicity, residency, first generation status, major preference, admission into a first choice school, number of applications, early outreach number of siblings in college, merit aid, GPA, and ACT, Pell Grant eligibility remained an important predictor of the instantaneous odds of enrollment at two of the four institutions under review. Pell Grant eligible students were 7% (*HR* = 0.93, 95% CI: 0.87-0.99, $p = .04$) less likely to enroll at institution one. Conversely, Pell Grant eligible students were 2.17 (95% CI: 1.47-3.13, $p < .001$) times more likely to enroll at any given time at institution three after adjusting for select factors. Similarly, Pell Grant eligibility maintained a marginal positive association with increased enrollment at institution two (*HR* = 1.14, 95% CI: 0.98-1.32, $p = .09$). On multivariable analysis, Pell Grant eligibility was not meaningfully associated with increased enrollment at institution four and was removed from the final model after evaluation of AIC.

By contrast to univariable findings, the offer of merit aid remained significantly associated with the odds of enrollment at only one of the four institutions. Students offered merit aid were 26% (*HR* = 0.74, 95% CI: 0.61-0.91) less likely to enroll at any given time at institution two ($p = .01$). This result may be indicative of the competitive market for higher achieving admitted students, as evidenced by this subset of four like-profile institutions. Regardless, offers of merit aid were no longer significantly associated with time to enrollment at institutions one (*HR* = 1.05, 95% CI: 0.95-1.17, $p = .41$), three (*HR* = 7.92, 95% CI: 0.48-129.52, $p = .95$), or four (*HR* = 3.19, 95% CI: 0.44-22.99, $p = .25$) on multivariable analysis.

Academic Factors

After adjusting for select covariates, students' cumulative GPAs remained significantly associated with the odds of enrollment at three of the four institutions. In line with univariable findings, however, the direction of this relationship was different for institution one compared to institutions three and four. For every one standard deviation (units=0.5) increase in students' cumulative GPA, admitted students were 10% (HR = 0.90, 95% CI: 0.87-0.94, $p < .0001$) less likely to enroll at any given time at institution one on multivariable analysis. By contrast, for every half unit increase in students' cumulative GPA, admitted students were 81% (*HR* =1.81, 95% CI: 1.54-2.12, $p < .0001$) and 58% (*HR* =1.58, 95% CI: 1.31-1.90, $p < .0001$) more likely to enroll at any given time at institutions three and four.

Controlling for select factors, students' standardized test scores emerged as significant predictors of time to enrollment for each of the four universities. As with other important factors, though, the magnitude and direction of these associations varied across institutions. For every one standard deviation (units=4) increase in students' ACT scores, admitted students were 16% (*HR* = 0.84, 95% CI: 0.80-0.88, $p < .0001$) and 24% (*HR* = 0.76, 95% CI: 0.69-0.83, $p < .0001$) less likely to enroll at any given time at institutions one and two, respectively. Conversely, for every four unit increase in students' ACT scores, admitted students were six to seven times more likely to enroll at any given time at institutions three (*HR* = 6.65, 95% CI: 4.87-8.881, $p < .0001$) and four (*HR* = 7.12, 95% CI: 5.11-9.92, $p < .0001$).

CHAPTER FIVE

DISCUSSION

The primary objective of this study was to ascertain if select sociodemographic, institutional, financial, and academic factors effectively reduce time to postsecondary enrollment. Further, this analysis also looked to extend prior research by identifying any common effects that may exist across several similar-profile postsecondary institutions. To address these aims, this study applied a multi-level, competing risks model to the analysis of undergraduate application data. The results revealed differential effects across the competitive set for every parameter under review, except first choice status. These findings can not only be used to inform a single institution's enrollment management strategy, but there are also numerous policy implications associated with these divergent results.

As the primary data source for this analysis was a single institution, the results and their implications should be interpreted in that context. This analysis provides important insight on not only the profile of student that an institution typically attracts, but also the segments with which it may struggle compared to similar profile peers. This information has the potential to inform on a multitude of institutional aims. First, it can help to tailor and appropriate align institutional services to meet the needs of the incoming undergraduate class. Second, it identifies the student subgroups among whom additional or different outreach may be beneficial. Finally, building on this, it can also help to pinpoint the specific competitors to which these students are drawn, providing a roadmap for what additional services, programs, and messages may be effective. While institution-specific data from all included competitors would provide a more nuanced understanding of select metrics, such as the impact of competing financial aid offers, the results of this analysis remain informative and actionable on an institution by institution basis.

While this modeling approach provides unique insights into students' timelines for making their final enrollment decisions, identifying individual and institutional characteristics that drive students' decision-making is only the first step. What is just as, if not more important is how an institution puts these findings into action. To that end, institutions that employ such analytic techniques must be prepared to use the evidence they uncover to inform strategic decision-making (Alkin, 2013; Patton, 2008). This will likely necessitate, among other steps, an internal evaluation of sorts, through which future research endeavors are tailored to address the information needs of key organization members (Russ-Eft & Preskill, 2009). In this manner, the findings of this analysis are but the first step in a larger process of strategically clarifying institutional priorities, identifying when and where new or updated services are necessary, and aligning corresponding recruitment activities to achieve well-defined and broadly accepted institutional goals. Consideration of model results within the undergraduate recruitment process will help to alleviate some of the initial budget constraints by identifying how and when certain known factors increase the probability of student enrollment, while not sacrificing on other important postsecondary measures, such as retention and graduation.

Sociodemographic Factors

Select sociodemographic factors have significant influence on students' college choice. In particular, an extensive literature exists on the roles students' race and ethnicity play in their postsecondary enrollment decisions (Wyatt et al., 2014; Lin, 2011; Bush, 2009; Goenner & Pauls, 2006; Zarate & Gallimore, 2005; DesJardins et al., 2002). For example, minority high school graduates, as well as those from more impoverished backgrounds often face many impediments, or "cumulative disadvantages," to accessing higher education (Schultz & Mueller, 2006). These can include, but are often not limited to, a lack of access to information and resource networks, inequality of neighborhood resources, and lack of peer/parental support for academic achievement (Lin, 2011; Gándara & Bial, 2001). Consequently, these students typically record lower GPA and standardized test scores, as well as apply to relatively fewer colleges, resulting in below average postsecondary enrollment rates (Smith, 2011; Goyette, 2008). While Pew research shows that Hispanic and African American students have accounted for the largest gains in college enrollment since 2000, enrollment gaps remain due to lower rates of four-year college enrollment, as well as lower attendance at selective colleges (Krogstad & Fry, 2014).

In this analysis, students' race remained significantly associated with the instantaneous odds of enrollment at all four institutions on multivariable analysis, whereas their ethnicity emerged as a meaningful indicator of enrollment at only two of the four institutions. Despite these overall trends, the magnitude and direction of these relationships often varied widely across each of the universities, underscoring the importance of incorporating competitor data into an institution's enrollment model. These concurrent evaluations provide institution-specific insight that could be used to support positive enrollment trends or provide further insight into potential strategies to bolster outreach efforts among students that would otherwise choose a competing institution.

Several institutions included in the analysis appear well positioned to attract minority students vis-à-vis their peers. For example, the odds of instantaneous enrollment increased significantly among Asian students at institution three ($p = .01$), while Black or African American students were also over ten times more likely to enroll at any given time at institutions three and four ($p < .0001$). Similarly, these same institutions have clearly made significant inroads among Hispanic applicants, who were five to seven times more likely to enroll at any given time ($p < .0001$). For these institutions, an internal evaluation of current outreach efforts and student services could prove beneficial to help identify areas of strength. In doing so, they could ensure these current trends are not only sustained, but perhaps replicated among other applicant subgroups that may warrant additional consideration.

By contrast, evidence indicates institution two may face significant challenges in attracting Asian and Black or African American students admitted to other universities this subset. For institutions that struggle to attract and retain minority students, it is important to identify and accentuate those institutional characteristics that increase the likelihood of enrollment very early in the process. To this end, key university stakeholders must investigate strategies to enhance the coordination of current student services (admissions, first year programming, advising, etc.), as well as look for opportunities to bolster targeted outreach efforts. Therefore, one goal of an internal evaluation would be to assess students' awareness of and satisfaction with current and future services aimed toward improving minority student recruitment and retention at the undergraduate level. Although results varied by institution, the findings of this analysis underscore the crucial role a student's racial and ethnic identity play in the timing of their final enrollment decisions. This sort of information coupled with projected demographic shifts within the broader U.S. population, suggest a more tailored and nuanced approach to high school student recruitment may benefit institutions looking to diversify their undergraduate student populations (Colby & Ortman, 2015).

Whether or not an institution has an established track record of success recruiting minority students, ongoing efforts to enhance current strategies are key to sustained success. As administrators look to diversity for their undergraduate student population, it is imperative that they consider the viewpoints of front-line staff, student workers (e.g. resident assistants), and other undergraduate students who may serve as informal brand ambassadors. Such targeted outreach could include surveys to establish an empirical measure of student sentiment regarding current or proposed services, social spaces, and academic support. These quantitative findings could then be supplemented through focus groups and/or one-on-one interviews to elicit feedback regarding more detailed strategies to address issues related to the on-campus social and cultural climate. An inclusive strategy to address issues related to minority student recruitment will likely not only pay dividends in terms of direct matriculation, but also increased persistence and graduation rates.

In addition to race and ethnicity, there are a wide range of economic and educational implications resulting from the growing gender gap in college enrollment (Conger & Long, 2013; Cho, 2006; DesJardins et al., 2002; Card & Lemieux, 2000). Through 2019, the NCES projects female student enrollment in colleges and universities across the country will grow by 21%, compared to just 12% for their male counterparts (Hussar & Bailey, 2011). While sex remains an important determinant in postsecondary enrollment, the results of this analysis found that it was often not a significant driver of their time to enrollment. After adjusting for select sociodemographic, institutional, financial, and academic factors, male students were 14% (HR = 1.14, 95% CI: 1.06-1.22) more likely to enroll at any given time at institution one compared to female applicants ($p < .001$). By contrast, a student's sex was no longer meaningfully associated with the odds of enrollment at institutions two ($p = .14$), three ($p = .67$), or four ($p = .56$) on multivariable analysis. While it is important institutions continue to target their resources to reduce gender imbalances, the findings of this study indicate this criterion does not inform on the timing of students' enrollment decisions (Conger & Long, 2013; Cho, 2006; Card & Lemieux, 2000; Brugglink & Gambhir, 1996). As such, this factor is less subject to the timing of administrators' outreach, as well as the potential influence of competitor activities to address any previously identified deficits. Overall, these findings reveal that sex is not a time-sensitive factor in students' postsecondary enrollment after accounting for other important criteria for a majority of the institutions currently under review.

By contrast, the results of this study confirmed residency status was meaningfully associated with an increase in the odds of enrollment at any given time at three of the four institutions on multivariable analysis. This is in line with Kumar et al. (2015), who found a majority of undergraduates attend a school in their state of residence. In this analysis, in-state applicants were 54-56% more likely to enroll at institutions two and three at any given time compared to out of state students ($p < .05$). As a result of their success among in-state applicants, these institutions are particularly well positioned to expand their traditional recruiting footprint compared to their peers. For administrators at institutions two and three, a review of primary and secondary research, including institutional records and third-party search service white papers, could inform participation at select college fairs and outreach to high school counselor groups in new cities and states. Brugglink and Gambhir (1996) found that students from outside traditional recruitment areas tend to have a less clear understanding of a school's mission or academic reputation. As a result, identifying amenable student audiences and recognizing/testing messages that may resonate with these new groups would be of the utmost importance before a full-scale investment recruitment resources would be warranted.

However, after accounting for the draw of select similar-profile institutions, the results suggest that in-state recruitment is more of a challenge for institution one than others ($p < .001$). Recognizing these geographical patterns and adjusting recruitment efforts accordingly, early in the process, may help to avoid an overreliance on out-of-state/region applicants, who are both more time and resource intensive targets. For universities similar to institution one, additional research among prospective in-state students could prove vital. One strategy would be a mixed methods approach that would incorporate survey administrations among students, high school personnel, and university administrators alike, as well as follow-up interviews and document analyses. These steps could also be supplemented by secondary research on in-state student retention efforts at similar profile institutions across the country, such as policies for living at home and commuter student services. A complete review and, if necessary, evidence-based reshaping of in-state recruitment efforts would likely pay long-term dividends.

Finally, parental education-level is among the most important sociodemographic factors that typically animate students' decision-making process. In fact, numerous studies have found that first generation status is a crucial indicator of postsecondary enrollment and performance (Lin, 2011; Goyette, 2008; Warburton & Nunez, 2001). Nonetheless, after controlling for other important sociodemographic, institutional, financial, and academic factors, first generation status only remained a significant predictor of time to enrollment at a single institution in this study. Specifically, first generation students were 41% more likely to enroll at any given time at institution two ($p < .0001$). Given the potential cultural and academic deficits with which these students may enter, it is imperative that this institution, and others like it, maintain existing and possibly fund new support services to address any and all shortfalls (Carnevale & Strohl, 2013). Such steps could include the establishment or possible expansion of federally funded TRIO programs. What is critical, is that institutions that attract a disproportionate percentage of first generation students coordinate all necessary levels of support before and after enrollment to ensure these undergraduate students are positioned for postsecondary success.

Institutional Factors

Early, personalized attention has also been shown to improve post-secondary enrollment rates. Even modest levels of engagement early in the application cycle have been shown to engender important postsecondary benefits, particularly among those students from impoverished backgrounds (Wyatt et al., 2014; Thomas et al., 1999). In this study, early outreach also emerged as an important indicator of the odds of enrollment at any given time at three of the four institutions on multivariable analysis. Specifically, early outreach targets were 25-64% more likely to enroll at institutions one ($p < .0001$), two ($p < .002$), and three ($p = .01$) after adjusting for select covariates. While the merits of early outreach and the resulting impact on direct postsecondary matriculation are generally well accepted, these findings further suggest that such efforts can significantly shorten students' decision timelines, saving both families and institutions money. Engaging prospective students before their senior year of high school produces sustained benefits throughout the application cycle. By targeting qualified candidates earlier in their academic career, administrators can simultaneously increase the likelihood of enrollment, while also reducing future recruitment overhead.

To enhance early outreach efforts, however, administrators should implement strategies to critically examine what types of communications impact students' decision timelines the most. For instance, prospective randomized studies could be employed to investigate if broader topics, such as reputation, campus location, sports, etc. resonate better with prospective students earlier in their high school careers. Consideration of the timing of such communications and messaging content could also provide additional avenues for future research. Specifically, do messages with particular subject lines lead to more email opens, are there particular days that generate broader readership, and are there topics that are more effective among student or parent audiences. Each of these options provides examples of how these and other institutions could capitalize on areas of perceived strength. While these findings confirm the importance of establishing relationships with prospective students and their families early in the process, further research could help to identify the mechanisms by which administrators could capitalize on and magnify these advantages.

Another area in which postsecondary institutions can exert a modicum of control, is their academic programming and how they market such offerings to prospective students. Students' sense of institutional fit and thus their enrollment decisions can sometimes be driven by their choice of major and the school's perceived strength in that area (DesJardins et al., 2002). In this analysis, students' major preferences were found to be meaningfully associated with the instantaneous odds of enrollment at two of the four institutions after controlling for select student and institutional factors. Targeting marketing efforts to strategically align messaging with students' academic preferences is an important tactic in shortening students' decision timelines. Further, this type of outreach could also be used to strengthen ties with select high schools or high school networks that have particular affinities, such as magnet STEM schools or secondary institutions that incorporate significant Advanced Placement (AP) or International Baccalaureate (IB) coursework.

Furthermore, major preference has also been identified as one of the most effective strategies for promoting student retention and completion at four-year private institutions (Ruffalo Noel Levitz, 2016). Major preference is one of the most accessible pieces of information available on prospective students. Related information is often repeatedly reported via multiple channels, including the application itself, requests for information (RFI), open houses, campus visits, and college fairs. Leveraging this information to micro-target marketing efforts and personalize student and parent outreach has the potential to reduce the time spent recruiting applicants that present specific academic profiles, which align with institutional strengths.

For institution two, at which business and communication majors are more likely to enroll at any given time ($p < .0001$), this information could be used in support of expanded programming in target fields. It could also be paired with other information, such as outcomes data, internship placement rates, and networking events to build upon a track record of established success. Similarly, institution one could parlay its success among prospective STEM majors ($p < .0001$) into new private and public investment opportunities with various industry, state, and federal actors. This would have the potential to further cement the institution's standing in the STEM community, but also attract the grants necessary to pursue the critical research that could further position it as a leader in the field. By contrast, institution two's observed disadvantage among prospective STEM majors ($p < .0001$) could spur further funding in areas in which the university wishes to expand or could be used as confirmation of its orientation and commitment to other academic areas. Importantly, in each instance, major preference information is readily available and important to students' decision-making throughout the application process.

Another critical factor in students' decision timelines is their admission into a first choice school. Prior research suggests that most students list schools in order of preference on their FAFSA submission, and nearly two-thirds of applicants enroll in their first choice school if admitted (CNN Money, 2015). In line with these findings, first choice designation was associated with a four- to eight-fold increase in the instantaneous odds of enrollment at each of the four universities, after adjusting for other important covariates ($p < .0001$). First choice applicants are clearly among the most amenable to an institution's recruitment efforts. The sustained magnitude of these observed effects, however, effectively provides administrators with a level of flexibility, as these prospects present options for both immediate enrollment or as targets for later efforts to make class (e.g. achieve predetermined enrollment goals) should other, potentially more difficult to attract, enrollment targets fall through.

Leveraging financial aid data throughout the application process is a critical component to the success of any recruitment strategy. By doing so, institutions have access to a wider variety of student characteristics than would otherwise be available through the application process alone. While such information provides vital data for modeling students' merit and financial aid, it also includes valuable criteria, such as indicators of institutional preference. First choice and related variables are likely most useful in concert with other important data points. For instance, cross tabulation of first choice applicants against other desirable criteria, such as race, measures of students' academic ability, etc., could be the difference between achieving or falling short of enrollment targets among certain subgroups of applicants. What is clear from the evidence presented throughout this analysis, is that first choice preference is among the most predictive factors associated with students' time to enrollment. Collecting this information early in the process will provide significant flexibility in any institutions' broader enrollment management strategy.

Researchers have also documented associations between other broad application factors and undergraduate enrollment trends. For instance, Smith (2011) showed that the more college applications a student submits leads to a corresponding increase in their probability of enrolling at a four-year college by as much as 40-50% (Smith, 2011). As one in four high school graduates who apply to four-year colleges still do not enroll in one, this criterion can play a potentially vital role in predicting time to enrollment (Avery & Kane, 2004). The findings of this analysis confirm that for each additional application submitted, there is a meaningful and corresponding shift in students' time to enrollment at three of the four institutions. For every additional application submitted, admitted students were 5-31% more likely to enroll at any given time at institutions two ($p < .001$), three ($p = .02$), and four ($p < .0001$). These three institutions were particularly successful at attracting students in an otherwise competitive undergraduate recruitment market. By contrast, the odds of instantaneous enrollment declined at institution one after adjusting for other important covariates ($p < .0001$).

The institutions to which students release their FAFA data can be instructive as to what the market considers an institution's peer and, potentially, aspirant set. Institutions which struggle to recruit applicants with more college options may need to evaluate their position vis-à-vis this core group of similar profile competitors, known as FAFSA overlap schools. A full landscape analysis including an inventory of major offerings, location, cost of attendance, outcomes, and other institutional factors should be considered as part of this review. Conversely, for institutions which appear to thrive in a crowded market, a similar review has the potential to uncover new strategies and services that may further the universities' perceived advantages. Informed by this secondary research, targeted primary research can then be employed to test new messaging, evaluate demand for new programming, and generally solicit feedback about current and proposed university services. At its core, the application cycle is a highly competitive process, which is subject to change from a variety of inputs. Through a better understanding of who students consider to be an institution's competitors, a university can begin the process of truly assessing its strengths and weaknesses against a well-defined set of peers.

Financial Factors

Offers of financial aid to admitted high school seniors often serve two purposes; to "relieve liquidity constraints" that may have undue influence on students' decision-making process and to alter students' "preference rankings" (Avery & Hoxby, 2004; DesJardins et al., 2002). Research has shown that students typically respond in a rational manner to financial incentives, with earlier aid offers, larger awards, and merit-based assistance tending to increase the probability of postsecondary enrollment. Regardless, the multivariable model results from this study found the effects of select financial factors were substantially moderated. For instance, offers of merit aid were no longer significantly associated with the instantaneous odds of enrollment at institutions one ($p = .28$), three ($p = .96$), or four ($p = .41$) on multivariable analysis. Similarly, Pell Grant eligibility only remained an important predictor of the instantaneous odds of enrollment at two of the four institutions.

As this analysis focused on those students admitted to selective, private four-year colleges, it is possible that competing offers of merit aid were widely available and largely cancelled each other out as meaningful factors within the decision making process. At a minimum, the high academic ability of the applicants under review could suggest offers of merit aid were likely more available and thus not a distinguishing factor among any of the institutions included in this analysis. Similarly, three quarters of admitted students across the four institutions did not qualify for Pell Grant funding. In such instances, universities must 1) accurately identify the parameters that define its prospective student base, and 2), decide on alternative strategies that may set it apart for high achieving college applicants. One possible approach is to message on honors programming or other tailored academic options, such as learning communities, which serve not only to acknowledge students' past achievements and abilities, but also to differentiate an institution from its peers. Another possible route might be to offer selective benefits in the form of early registration or move-in times, coupled with opportunities for undergraduate research. Further evaluation of current and prospective student services and secondary analyses of competitor offerings would serve to both inform and guide these activities and related marketing.

With the recent executive action by the Obama administration, students' financial data are now available early enough in the application cycle to be meaningfully incorporated into yield models (Department of Education, 2015). As a result of this shift, university administrators are now able to estimate the potential impact of differential financial aid packaging directly in their enrollment projections. In addition to easing the reporting burden on students and their families, this policy change has the potential to help postsecondary institutions provide earlier financial aid offers, adjust their communications flow, and more accurately track progress toward established enrollment goals. While financial factors did not emerge as significant distinguishing factors between institutions in this analysis, it is important to recognize that highly desirable candidates will still expect and will likely receive competing offers of financial aid. In such instances, accurately calibrating the thresholds at which the probability of timely enrollment increase or decrease is of paramount importance.

Academic Factors

Students' academic achievement has consistently been found to be meaningfully associated with an array of important postsecondary measures (Ledesma, 2009; Chang, 2006; Brugglink & Gambhir, 1996; Thomas et al., 1979). Numerous empirical studies have shown that students with a record of strong academic performance consistently outperform their lower achieving peers in terms of college enrollment rates (NCES, 2015; Adelman, 2006). In this analysis, a student's cumulative GPA was also found to inform on their time to enrollment at three of the four institutions, even after adjusting for important covariates. Specifically, students with higher cumulative GPAs were 62-80% more likely to enroll at any given time at institutions three and four ($p < .0001$). Meaning, at any point throughout the recruitment cycle, the likelihood a student with a higher GPA would enroll at institutions three and four increased substantially, even after controlling for an array of important sociodemographic, institutional, financial, and academic factors.

By contrast, admitted students with higher GPAs were also found to be 10% ($p < .0001$) less likely to enroll at any given time at institution one on multivariable analysis. This trend is likely indicative of the competitive nature of the higher education marketplace. Ledesma (2009) showed that high achieving applicants tend to apply to and gain admission at multiple colleges and universities. Thus, qualified students are typically confronted with a wider range of enrollment options from which they must delineate between often subtle and subjective measures of institutional quality. The competing risks framework applied in this study appropriately captures this conflict, highlighting the importance of not only measuring the association between student-level factors and institutional enrollment, but also emphasizing the interplay between such measures and competitor activities. Thus, it provides important insight into how postsecondary institutions can position themselves vis-à-vis their closest peers to appeal to as qualified and broad a prospective student audience as possible.

This research also confirmed that after controlling for select sociodemographic, institutional, financial, and academic factors, students' standardized test scores were significant predictors of time to enrollment. Increasing ACT scores aligned with a 16-25% drop in the likelihood of enrollment across the application timeline at institutions one and two ($p < .0001$). Conversely, the analysis found that admitted students with higher ACT scores were between five and six times more likely to enroll at any given time at institutions three and four ($p < .0001$). These results support evidence that suggests that academic achievement is not only an indicator of how well prepared students are for the rigors of postsecondary education, but also their initial college choice (NCES, 2015; Chang, 2006; Brugglink & Gambhir, 1996). In this large and diverse student sample, several consistent patterns emerged suggesting that higher achieving students would prove to be significantly more difficult for institutions similar to one and two to recruit within the context of a competitive educational market with several like-profile peers.

These findings support evidence that students' academic ability not only increases demand from various institutional actors (DesJardins et al., 2002), but also, perhaps as importantly, that such measures can also inform on the timing of their decisions. In this study, higher achieving admitted students' enrollment patterns exhibited a wide range of often divergent outcomes. The evidence presented throughout this analysis confirmed that a student's academic background has a profound effect on the timing of their college choice, even after controlling for other important factors (Adelman, 2006; Thomas et al., 1979). In addition, it also underscores the importance of strategic allocation of institutions' recruitment budgets. As the enrollment outcomes associated with students' academic ability will be sustained throughout the application cycle, it is incumbent upon enrollment management personnel to balance competing goals of making and shaping each freshman class. Identifying which other student and institutional characteristics align with more desired institutional outcomes (e.g. enrollment) and then micro-targeting when further investment of additional capital is appropriate could lead to a higher percentage of stronger academic candidates enrolling over time.

Data Consideration for Future Analyses

Strategic allocation of limited recruitment budgets is, in part, informed by the collection and analysis of self-reported family and individual student data. This information is often provided throughout the recruitment, application, and financial aid processes. However, as with all analyses, predictive modeling, in any form, is limited by the data elements available for the analysis. Important metrics available for this analysis were intentionally limited to select sociodemographic, institutional, financial, and academic measures. Nevertheless, other possible covariates, such legacy status, were not readily available, limiting the scope of the current study. Further, as this initial analysis served as a proof of concept of sorts, interactions between the parameters employed in these models were not investigated. However, this may provide an interesting and useful line of inquiry for future analyses. Moving forward, there are also several noteworthy practical limitations to the methods employed in the analysis.

First, the models outlined throughout this study require application of multiple, often sophisticated statistical techniques for which institutions may not have adequate personnel. This particular approach combines three advanced methodologies, each of which, in isolation, requires the extension of more standard statistical models. One, the mixed effects model accounts for high school-level variation, or the shared effects representing a form of dependence among the enrollment probabilities of individuals from similar backgrounds (Collett, 2015; Lu & Peng, 2008; Raudenbush & Bryk, 2002). Two, a time to event analysis incorporates important aspects of the time dependent nature of the application cycle. For such analyses, the unit of measurement is time itself, as interest lies on the odds of an event occurring over time. Finally, a competing risks framework simultaneously assesses enrollment at multiple, similar-profile universities. This enables institutions to directly incorporate data on their institutional peers and aspirant colleges, which will greatly inform on their enrollment management strategies. While the extension and combination of such models provides critical information to the admissions personnel tasked with recruiting them, it also may make replication of these methods increasingly more difficult for some institutions.

In addition, an a priori decision was also made in this analysis to evaluate the impact of being admitted into a first choice school. To date, students are asked to designate up to 10 schools to which they want their financial information disclosed on the FAFSA. As this and other studies have shown, these “overlap schools,” as they are commonly referred to, can provide important analytic and practical insight into students’ decision timelines. Specifically, prior research has shown that nearly two-thirds of applicants enroll in their first choice school and the findings of this study indicated a four to five-fold increase in the instantaneous odds of enrollment at each of the four universities. Despite these results, regular access to this information is not always readily available and will likely require personnel that can combine data from disparate sources.

Finally, timely access to important sociodemographic, institutional, and financial factors was only guaranteed by the recent executive action by the Obama administration to enable students to report income two years’ prior to their FAFSA submission (Department of Education, 2015). As a result of this shift, university administrators are now able to estimate the potential impact of differential financial aid packaging directly in their enrollment models with enough time to adjust their communications flow and more accurately track progress toward established enrollment goals. If future government actions reverse or limit this access, several important variables will no longer be available to estimate important parameters early enough to proactively and appropriately adjust yield models.

Implications

Since the early 2000s, time to event modeling has been used to examine critically important issues, such as student completion and graduation. However, the bulk of enrollment modeling remains limited to more traditional modeling techniques, such as fixed and mixed effects binary logistic regression. The model outlined herein builds on the extensive undergraduate enrollment literature, while simultaneously augmenting and extending the field's emergent interest in time to event models. The multi-level design also appropriately accounts for variation driven by aggregate high school-level characteristics. Further, the competing risks framework assesses the roles competitors play in a crowded higher education market, thereby enabling institutions to incorporate important information on the appeal of similar profile colleges into their own yield models.

The potential benefits of these techniques in the field of higher education are many and clear, especially given policymakers' renewed focus on student outcomes over the past few decades. DesJardins et al. (1999) credited related modeling approaches for helping to develop timely interventions for students at risk of dropping out, while Gross and Torres (2010) used a similar model design to examine how the timing of financial aid affects educational attainment among minority student populations. In addition, these findings support prior work that seamlessly extend such techniques to meet the demands of complex, hierarchical designs (Bahr, 2009) or even those adapted to a "competing risks" framework (Guerin, 1997; DesJardins et al., 1999; Ronco, 1996).

Another important and potentially overlooked strength of this approach is that it shares a common objective with most prior models and utilizes readily available student data provided throughout the recruitment, application, and financial aid processes. Specifically, the main objective of this approach is to identify common covariates that are related to and drive students' enrollment decisions. Importantly, though, by delineating between the effects these factors have on students' enrollment timelines, admissions professionals can gain crucial insight into students' enrollment probabilities over time (Hosmer et al., 2008). For instance, the results of this analysis confirmed that sex, race, ethnicity, residency, first generation status, early outreach, major preference, first choice, number of applications, Pell Grant eligibility, cumulative GPA, and ACT scores are not only significantly associated with students' time to enrollment estimates, but also that these effects often differ across institutions. Given the complex interplay between such variables, the primary goal of this research was to present preliminary evidence on how such data could be leveraged to provide further insight into those factors that impact students' decision-making throughout the application cycle.

The statistical approach outlined herein provides evidence as to how a multi-level, competing risks framework can be formally applied to the analysis of undergraduate enrollment preferences. Practically, the analysis presents an empirical measure of the determinants of undergraduate enrollment in the context of a large and competitive postsecondary marketplace. By doing so, it appropriately accounts for the ways in which students from a range of academic and socioeconomic backgrounds must engage with a complex and ever-changing U.S. higher education system.

Further, the results of this model also inform on multiple financial and policy considerations across the postsecondary education system. First, institutions that place a premium on resource conservation after the recent financial crisis, have a roadmap to minimizing recruitment overhead and, thereby, freeing up important capital for other initiatives. For example, administrators can re-invest in vital student support services and first-year student programming or pay down outstanding debt and, as necessary, address pension shortfalls.

Second, more targeted efforts to shorten the decision timeline among a smaller pool of well-qualified and strongly matched applicants enables institutions to cover more of the initial costs, such as campus visits and admitted student events, associated with the search process. This will save admissions staff time and money in the long-run, but will also help to alleviate some of the initial cost constraints faced by otherwise qualified students and their families. Finally, early efforts to maximize student/institution match will also pay long-term financial dividends for students, institutions, and the Department of Education in the form of stronger retention, higher graduation rates, and lower cohort default rates.

In 2015, the Bureau of Labor Statistics reported that most of the 22 major occupational groups projected through 2024 will require significant levels of education, including postsecondary training and beyond (Hogan & Roberts). Furthermore, multiple studies have shown that college graduates earn twice as much and accumulate nearly two and half times the wealth of their less educated peers (DesJardins et al., 2002; Diaz-Jiminez et al., 1997; Murphy & Welch, 1993). These trends notwithstanding, the U.S. Census Bureau recently found that just one in three adults (33%) reported having a bachelor’s degree or more education (Ryan & Bauman, 2016). In addition, evidence suggests its economically disadvantaged high school graduates that tend to disproportionately pursue non-traditional enrollment options, delaying the social and economic benefits of postsecondary education (Goldrick-Rab, 2006; Reynolds et al., 2006; Hearn, 1992). Given the inextricable links between college readiness, retention, degree completion, and career preparation, such shortfalls have significant implications for higher education administrators, in particular those involved in the enrollment management process.

Recent trends toward increasing recruitment expenditures and tuition threaten to diminish student access and negatively impact a wide range of postsecondary outcomes (Fitzgerald, 2004; St. John et al., 2003). The Great Recession has only exacerbated the pervasive gaps in educational opportunities across traditional racial, ethnic, socioeconomic, and gender divides. According to the Center on Budget of Policy Priorities, 47 states spent less per student during the 2014-15 school year than they did at the start of the recession (Mitchell & Leachman, 2015). Coupled with these funding shortfalls, postsecondary institutions across the spectrum are facing increasing pressure to attract, retain, and graduate an ever more diverse and qualified undergraduate student body (Harvill et al., 2012). Despite recent growth in minority, lower SES, and first generation student enrollment, these subgroups are still considerably less likely than their peers to graduate high school and pursue postsecondary education (College Board, 2010; Education Advisory Board, 2016). Considering these challenges, admissions staff must apportion resources to identify and recruit applicants to maximize the fit between student and institution.

Despite possible limitations, this study succeeded in extending the current literature examining the relationship between select student- and school-level factors and undergraduate enrollment. It builds on recent applications of time to event models in higher education, while extending these approaches to a multi-level competing risks framework. The results confirm that such an approach will facilitate higher education administrators' efforts to identify those factors that effectively reduce time to enrollment in a competitive higher education market. Since the Great Recession, postsecondary institutions of all types have been forced to operate in an environment of reduced or constrained budgets and increased expectations regarding student outcomes. This research shows that important student and institutional factors that have been shown to predict undergraduate enrollment can also inform institutions' efforts to reduce recruitment expenditures by shortening students' decision timelines. This will enable universities sufficient flexibility to re-invest in vital student services and, by doing so, begin to regain some of the security lost during the recent financial collapse.

APPENDIX A

STUDENT VARIABLES

Ascribed Characteristics		
Variable	***Measure Type***	***Definition***
Sex	Binary	Dummy code, 1 = Male, 0 = Female
Race	Multinomial	Multilevel categorical variable with seven distinct levels: (1) Native American or Alaska Native; (2) Asian; (3) Black or African American; (4) Multi-Racial; (5) Native Hawaiian or other Pacific Islander; (6) Caucasian; or (7) Other
Ethnicity	Binary	Dummy code, 1 = Hispanic 0 = Not Hispanic
Residential Status	Binary	Dummy code, 1 = In-State 0 = Out-of-State
U.S. Region	Multinomial	Multilevel categorical variable with four distinct levels: (1) Midwest; (2) Northeast; (3) South; (4) West (*including Pacific*)
First Generation Status	Binary	Dummy code, 1 = Yes, 0 = No
Intended Major	Multinomial	Multilevel categorical variable with seven distinct levels: (1) Business; (2) Communications; (3) Education; (4) Liberal Arts; (5) Social Work; (6) STEM or (7) Undecided
First Choice College	Binary	Dummy code, 1 = Yes, 0 = No
Number of College Applications	Ordinal	Ordinal count, ranging from 0 to 10
Target of Early Outreach	Binary	Dummy code, 1 = Yes, 0 = No
Number of Kids in College	Binary	Dummy code, 1 = Greater than or equal to one, 0 = Family has no additional kids in college
Pell Grant Eligible	Binary	Dummy code, 1 = Yes, 0 = No
Merit Aid	Binary	Dummy code, 1 = Yes, 0 = No
Cumulative High School Grade Point Average	Continuous	Raw high score GPA scores
ACT Test Scores	Ordinal	Standardized test score ranging from 0 to 36
High School Cluster	Multinomial	College Board utilizes 28 unique high school clusters to group college applicants by various attributes

APPENDIX B

VARIANCE INFLACTION FACTOR AND CORRELATION ESTIMATES

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	B	0.08	0.02	3.36	0.00	.	0.00
ACT	1	0.00	0.00	2.26	0.02	0.65	1.54
HS GPA	1	0.00	0.01	-0.73	0.47	0.63	1.59
Asian	1	0.03	0.01	2.27	0.02	0.30	3.33
Black	1	0.00	0.01	-0.03	0.98	0.50	2.02
Multi	1	0.02	0.01	1.36	0.17	0.52	1.93
Other	1	-0.02	0.02	-1.12	0.26	0.87	1.14
White	1	0.01	0.01	0.85	0.39	0.21	4.69
First Choice	1	0.83	0.01	138.23	<.0001	0.84	1.19
Sex	1	0.00	0.00	-0.06	0.95	0.93	1.08
Early Outreach	1	0.06	0.00	14.88	<.0001	0.98	1.02
Pell Grant	1	0.00	0.01	-0.62	0.54	0.81	1.23
Merit Aid	1	-0.01	0.01	-1.84	0.07	0.68	1.48
College Kids	1	0.00	0.00	-0.30	0.76	0.98	1.02
First Generation	1	0.01	0.01	2.12	0.03	0.83	1.20
Hispanic	1	0.00	0.01	0.63	0.53	0.59	1.69
Resident	1	0.03	0.00	5.62	<.0001	0.81	1.23
Biz	B	0.02	0.01	2.64	0.01	0.66	1.53
Comm	B	0.05	0.01	4.48	<.0001	0.81	1.24
Edu	B	-0.01	0.01	-0.70	0.48	0.88	1.13
LA	B	0.02	0.01	2.43	0.02	0.72	1.38
STEM	B	0.02	0.01	3.08	0.00	0.54	1.85
Undecided	0	0.00	.	.	.	.	.
App Count	1	0.00	0.00	-5.25	<.0001	0.82	1.22

	ACT	HS GPA	Asian	Black	Multi	Other	White	First Choice	Sex	Early Outreach	Pell Grant
ACT	1.00	0.40	0.03	-0.15	0.01	-0.05	0.17	-0.02	0.08	0.07	-0.21
		<.0001	<.0001	<.0001	.14	<.0001	<.0001	.001	<.0001	<.0001	<.0001
HS GPA	0.40	1.00	0.04	-0.07	0.00	0.02	0.00	0.01	-0.14	0.02	0.04
	<.0001		<.0001	<.0001	.54	.002	.45	.23	<.0001	<.0001	<.0001
Asian	0.03	0.04	1.00	-0.09	-0.09	-0.04	-0.55	0.01	0.03	-0.02	0.05
	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	.07	<.0001	<.001	<.0001
Black	-0.15	-0.07	-0.09	1.00	-0.05	-0.03	-0.33	-0.02	-0.02	-0.03	0.11
	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	.001	<.0001	<.0001
Multi	0.01	0.00	-0.09	-0.05	1.00	-0.03	-0.33	0.00	-0.01	0.03	0.01
	.14	.54	<.0001	<.0001		<.0001	<.0001	.44	.27	<.0001	.36
Other	-0.05	0.02	-0.04	-0.03	-0.03	1.00	-0.16	0.00	0.00	-0.02	0.05
	<.0001	.002	<.0001	<.0001	<.0001		<.0001	.81	.47	0.0001	<.0001
White	0.17	0.00	-0.55	-0.33	-0.33	-0.16	1.00	0.02	-0.01	0.04	-0.20
	<.0001	.45	<.0001	<.0001	<.0001	<.0001		.004	.14	<.0001	<.0001
First Choice	-0.02	0.01	0.01	-0.02	0.00	0.00	0.02	1.00	0.00	0.09	0.09
	.001	.23	.07	<.0001	.44	.81	.004		.83	<.0001	<.0001
Sex	0.08	-0.14	0.03	-0.02	-0.01	0.00	-0.01	0.00	1.00	0.01	-0.02
	<.0001	<.0001	<.0001	.001	.27	.47	.14	.83		.05	<.0001
Early Outreach	0.07	0.02	-0.02	-0.03	0.03	-0.02	0.04	0.09	0.01	1.00	-0.01
	<.0001	<.0001	<.001	<.0001	<.0001	0.0001	<.0001	<.0001	.05		.19
Pell Grant	-0.21	0.04	0.05	0.11	0.01	0.05	-0.20	0.09	-0.02	-0.01	1.00
	<.0001	<.0001	<.0001	<.0001	.36	<.0001	<.0001	<.0001	<.0001	.19	
Merit Aid	0.41	0.53	0.00	-0.08	0.00	0.00	0.06	0.03	-0.06	-0.03	-0.03
	<.0001	<.0001	.38	<.0001	.98	.83	<.0001	<.0001	<.0001	<.0001	<.0001
College Kids	0.01	0.00	0.01	0.02	0.02	0.01	-0.05	0.00	0.00	0.01	-0.01
	.40	.74	.03	.02	.01	.28	<.0001	.62	.45	.06	.40
First Generation	-0.20	0.03	-0.02	-0.01	-0.03	0.06	-0.11	0.02	-0.02	-0.03	0.31
	<.0001	<.0001	.001	.22	<.0001	<.0001	<.0001	.01	.0001	<.0001	<.0001
Hispanic	-0.27	-0.02	-0.17	-0.07	-0.04	0.13	-0.18	-0.02	-0.01	-0.04	0.20
	<.0001	.0002	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	.17	<.0001	<.0001
Resident	-0.08	0.16	0.15	0.02	-0.02	0.05	-0.20	0.13	0.02	0.00	0.22
	<.0001	<.0001	<.0001	.001	.003	<.0001	<.0001	<.0001	<.0001	.65	<.0001
Biz	-0.05	-0.11	-0.02	-0.01	-0.01	-0.01	0.04	-0.01	0.17	-0.01	-0.06
	<.0001	<.0001	.001	.01	.01	.01	<.0001	.08	<.0001	.03	<.0001
Comm	-0.04	-0.07	-0.05	0.01	0.00	-0.01	0.05	0.01	-0.06	0.01	-0.03
	<.0001	<.0001	<.0001	.17	.66	.19	<.0001	.34	<.0001	.10	<.0001
Edu	-0.05	-0.02	-0.03	-0.02	-0.02	0.00	0.03	-0.01	-0.07	0.01	0.01
	<.0001	<.001	<.0001	<.0001	.004	.39	<.0001	.30	<.0001	.25	.23
LA	0.02	-0.03	-0.07	0.01	0.00	0.00	0.04	-0.01	-0.04	0.01	-0.01
	.001	<.0001	<.0001	.06	.50	.38	<.0001	.28	<.0001	.01	.29
STEM	0.13	0.16	0.11	0.02	0.02	0.00	-0.08	0.03	-0.01	0.01	0.03
	<.0001	<.0001	<.0001	<.0001	<.001	.91	<.0001	<.0001	.25	.14	<.0001
Undecided	-0.08	-0.03	-0.02	-0.02	-0.01	0.01	0.00	-0.02	-0.05	-0.02	0.03
	<.0001	<.0001	<.001	<.001	.33	.03	.78	<.0001	<.0001	<.001	<.0001
App Count	0.08	0.13	-0.01	0.07	0.01	0.02	-0.05	-0.02	-0.01	0.00	0.26
	<.0001	<.0001	.01	<.0001	.03	<.001	<.0001	<.0001	.11	.50	<.0001

	Merit Aid	College Kids	First Generation	Hispanic	Resident	Biz	Comm	Edu	LA	STEM	Undecided	App Count
ACT	0.41	0.01	-0.20	-0.27	-0.08	-0.05	-0.04	-0.05	0.02	0.13	-0.08	0.08
	<.0001	.40	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	.001	<.0001	<.0001	<.0001
HS GPA	0.53	0.00	0.03	-0.02	0.16	-0.11	-0.07	-0.02	-0.03	0.16	-0.03	0.13
	<.0001	.74	<.0001	<.001	<.0001	<.0001	<.0001	<.001	<.0001	<.0001	<.0001	<.0001
Asian	0.00	0.01	-0.02	-0.17	0.15	-0.02	-0.05	-0.03	-0.07	0.11	-0.02	-0.01
	.38	.03	.001	<.0001	<.0001	<.001	<.0001	<.0001	<.0001	<.0001	<.001	.01
Black	-0.08	0.02	-0.01	-0.07	0.02	-0.01	0.01	-0.02	0.01	0.02	-0.02	0.07
	<.0001	.02	.22	<.0001	.001	.01	.17	<.0001	.06	<.0001	<.001	<.0001
Multi	0.00	0.02	-0.03	-0.04	-0.02	-0.01	0.00	-0.02	0.00	0.02	-0.01	0.01
	.98	.01	<.0001	<.0001	.003	.01	.66	.004	.50	.0003	.33	.03
Other	0.00	0.01	0.06	0.13	0.05	-0.01	-0.01	0.00	0.00	0.00	0.01	0.02
	.83	.28	<.0001	<.0001	<.0001	.01	.19	.39	.38	.91	.03	<.001
White	0.06	-0.05	-0.11	-0.18	-0.20	0.04	0.05	0.03	0.04	-0.08	0.00	-0.05
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	.78	<.0001
First Choice	0.03	0.00	0.02	-0.02	0.13	-0.01	0.01	-0.01	-0.01	0.03	-0.02	-0.02
	<.0001	.62	.01	<.0001	<.0001	.08	.34	.30	.28	<.0001	<.0001	<.0001
Sex	-0.06	0.00	-0.02	-0.01	0.02	0.17	-0.06	-0.07	-0.04	-0.01	-0.05	-0.01
	<.0001	.45	<.001	.17	<.0001	<.0001	<.0001	<.0001	<.0001	.25	<.0001	.11
Early Outreach	-0.03	0.01	-0.03	-0.04	0.00	-0.01	0.01	0.01	0.01	0.01	-0.02	0.00
	<.0001	.06	<.0001	<.0001	.65	.03	.10	.25	.01	.14	<.001	.50
Pell Grant	-0.03	-0.01	0.31	0.20	0.22	-0.06	-0.03	0.01	-0.01	0.03	0.03	0.26
	<.0001	.40	<.0001	<.0001	<.0001	<.0001	<.0001	.23	.29	<.0001	<.0001	<.0001
Merit Aid	1.00	0.00	-0.02	-0.07	0.01	-0.07	-0.04	-0.03	-0.01	0.12	-0.03	0.09
		.92	<.001	<.0001	.07	<.0001	<.0001	<.0001	.19	<.0001	<.0001	<.0001
College Kids	0.00	1.00	0.06	0.06	0.04	-0.02	-0.01	-0.01	0.02	0.02	-0.01	0.05
	.92		<.0001	<.0001	<.0001	.01	.26	.15	.003	.004	.04	<.0001
First Generation	-0.02	0.06	1.00	0.25	0.21	-0.03	-0.05	0.01	-0.02	0.01	0.05	-0.01
	<.001	<.0001		<.0001	<.0001	<.0001	<.0001	.13	<.001	.17	<.0001	.29
Hispanic	-0.07	0.06	0.25	1.00	0.12	0.00	-0.02	0.02	0.02	-0.03	0.03	0.06
	<.0001	<.0001	<.0001		<.0001	.37	0.0001	<.0001	.0004	<.0001	<.0001	<.0001
Resident	0.01	0.04	0.21	0.12	1.00	-0.10	-0.10	0.03	-0.07	0.13	0.03	0.03
	.07	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Biz	-0.07	-0.02	-0.03	0.00	-0.10	1.00	-0.11	-0.08	-0.14	-0.40	-0.22	-0.05
	<.0001	.01	<.0001	.37	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Comm	-0.04	-0.01	-0.05	-0.02	-0.10	-0.11	1.00	-0.04	-0.08	-0.23	-0.12	0.00
	<.0001	.26	<.0001	<.001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	.50
Edu	-0.03	-0.01	0.01	0.02	0.03	-0.08	-0.04	1.00	-0.06	-0.16	-0.09	-0.01
	<.0001	.15	.13	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	.21
LA	-0.01	0.02	-0.02	0.02	-0.07	-0.14	-0.08	-0.06	1.00	-0.30	-0.16	0.02
	.19	.003	.001	<.001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001
STEM	0.12	0.02	0.01	-0.03	0.13	-0.40	-0.23	-0.16	-0.30	1.00	-0.46	0.02
	<.0001	.004	.17	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001
Undecided	-0.03	-0.01	0.05	0.03	0.03	-0.22	-0.12	-0.09	-0.16	-0.46	1.00	0.01
	<.0001	.04	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		.21
App Count	0.09	0.05	-0.01	0.06	0.03	-0.05	0.00	-0.01	0.02	0.02	0.01	1.00
	<.0001	<.0001	.29	<.0001	<.0001	<.0001	.50	.21	<.0001	<.0001	.21	

APPENDIX C

DESCRIPTORPLUS HIGH SCHOOL CLUSTER DESCRIPTIONS

(Source: College Board Educational Cluster Key)

High School Cluster 51

These high schools are predominantly public and serve traditional, blue-collar communities with very low home values. Families are mature and own their homes but have relatively low incomes. Students often will be the first in their family to graduate from college and have modest curricular preparation, below average test scores, and low degree aspirations. They submit relatively few applications and set their sights on low cost, less selective institutions and local community colleges within their home state. Many will be applying for financial aid, particularly if they are going away to school.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	457	22	% of Students 1st generation	71%	7	Ave cost targeted colleges (x $1000)	$8.3	29
Mean SAT Math Score	462	23	Ave Admit Rate at Targeted Colleges	61%	3	% of students non-White	33%	22
Mean SAT Writing Score	445	22	Ave Number of AP Exams per Student	0.91	22	% of families below poverty	14%	8
Ave Number of Advanced Courses	0.55	18	% likely to apply out of state	14%	29	% interested in Financial Aid	68%	12

Dominant Cluster Factors

Focused/Early Decision | Few AP/Honors | College Interest: Local Technical | Cost Not an Object

High School Cluster 52

The high schools in this cluster are primarily religious or private and serve populations which are well-educated with a significant Hispanic influence. Although incomes are only slightly above average, families tend to own their own homes. Frequently dealing with English as a second language, students have access to good academic curricula and take advantage of AP/honors coursework but have slightly below average test scores. They are highly mobile and aspire to high levels of educational attainment generally at selective private or flagship public institutions with relatively high costs…financial aid is seen as a must.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	496	16	% of Students 1st generation	35%	23	Ave cost targeted colleges (x $1000)	$22.4	3
Mean SAT Math Score	489	17	Ave Admit Rate at Targeted Colleges	48%	24	% of students non-White	99%	1
Mean SAT Writing Score	487	16	Ave Number of AP Exams per Student	1.15	4	% of families below poverty	9%	16
Ave Number of Advanced Courses	1.15	13	% likely to apply out of state	81%	1	% interested in Financial Aid	77%	3

Dominant Cluster Factors

Puerto Rican/Caribbean ESL | Strong Academic Curriculum | College Interest: National Selective | Weak Standardized Testers

High School Cluster 53

These high schools are often religiously affiliated and serve middle class communities with a mix of professional, managerial and blue-collar households. Most of the families have some acquaintance with college although only a modest proportion includes a college graduate. Although students tend to get good grades, their test scores are below average and their involvement in AP and honors courses is minimal. Their degree aspirations are quite low and their college choices tend to less selective and lower cost church-related institutions close to home. Many will be applying for financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	484	18	% of Students 1st generation	51%	13	Ave cost targeted colleges (x $1000)	$11.4	26
Mean SAT Math Score	471	22	Ave Admit Rate at Targeted Colleges	61%	4	% of students non-White	38%	19
Mean SAT Writing Score	470	18	Ave Number of AP Exams per Student	0.54	23	% of families below poverty	8%	17
Ave Number of Advanced Courses	0.78	20	% likely to apply out of state	28%	25	% interested in Financial Aid	62%	16

Dominant Cluster Factors

Religious Curriculum | Few AP/Honors | College Interest: Local Technical | Lower Ability

High School Cluster 54

These high schools serve predominantly rural, working-class African-American and Hispanic families at the lowest end of the economic scale. Few parents have any experience with college. Students have access to a general curriculum which has few AP or honors opportunities; their test scores are at or near the bottom. Although they are willing to look out of state and to apply to moderately selective institutions, as well as local two-year and technical colleges, students from these schools seem to have low aspirations and little guidance or information regarding financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	371	28	% of Students 1st generation	91%	1	Ave cost targeted colleges (x $1000)	$13.5	21
Mean SAT Math Score	376	29	Ave Admit Rate at Targeted Colleges	55%	15	% of students non-White	96%	3
Mean SAT Writing Score	366	29	Ave Number of AP Exams per Student	0.32	27	% of families below poverty	22%	1
Ave Number of Advanced Courses	0.28	28	% likely to apply out of state	33%	20	% interested in Financial Aid	38%	24

Dominant Cluster Factors

Primarily African-American | Black Inner City | College Interest: Small Private | Not Athletic Participant

High School Cluster 55

The high schools in this cluster are primarily private or religiously affiliated and serve predominantly male, racially mixed populations from homes with modestly above average incomes. Most parents have attended college and hold predominantly professional or managerial positions. Although education is a community value, student participation in AP and honors courses, standardized test scores, and aspirations beyond high school are all below average. Willing to consider going out-of-state, students tend toward moderately priced and relatively selective institutions.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	481	19	% of Students 1st generation	46%	18	Ave cost targeted colleges (x $1000)	$16.1	16
Mean SAT Math Score	489	18	Ave Admit Rate at Targeted Colleges	56%	11	% of students non-White	66%	14
Mean SAT Writing Score	469	19	Ave Number of AP Exams per Student	0.47	24	% of families below poverty	8%	18
Ave Number of Advanced Courses	0.65	21	% likely to apply out of state	39%	13	% interested in Financial Aid	44%	22

Dominant Cluster Factors

College Prep School | Coed | College Interest: Lower Cost Satellite public | Relatively Low Grades

High School Cluster 56

These high schools, sometimes religious, serve solidly middle class, racially mixed, and slightly older communities with a mix of professional, managerial and blue-collar households and may have a strong athletic traditions. Most families have a parent with at least some college experience. Although not involved in many AP or honors courses, students have access to a math science curriculum and perform at an above average level on standardized tests. While not applying to many institutions they tend towards selective privates with higher costs, quite often outside their home state. Interest in financial aid is moderate.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	508	12	% of Students 1st generation	40%	19	Ave cost targeted colleges (x $1000)	$20.5	7
Mean SAT Math Score	536	9	Ave Admit Rate at Targeted Colleges	51%	20	% of students non-White	43%	17
Mean SAT Writing Score	505	12	Ave Number of AP Exams per Student	0.44	25	% of families below poverty	9%	13
Ave Number of Advanced Courses	0.62	22	% likely to apply out of state	73%	3	% interested in Financial Aid	49%	19

Dominant Cluster Factors

Athletic Achievements | Relatively Low Grades | Not Community Oriented | Few AP/Honors

High School Cluster 57

The high schools in this cluster are overwhelmingly public and serve predominantly low income, urban, African-American communities. Although there are some professionals, these are blue-collar families with very only a few college graduates among them. Students tend to be active in their schools, and avail themselves of AP and honors opportunities, although their standardized test performance is below average. Highly dependent on financial aid, they are likely to stay in state and apply to less selective publics.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	435	24	% of Students 1st generation	69%	8	Ave cost targeted colleges (x $1000)	$11.1	27
Mean SAT Math Score	434	26	Ave Admit Rate at Targeted Colleges	57%	9	% of students non-White	82%	10
Mean SAT Writing Score	425	24	Ave Number of AP Exams per Student	0.80	15	% of families below poverty	14%	7
Ave Number of Advanced Courses	1.34	11	% likely to apply out of state	29%	22	% interested in Financial Aid	79%	2

Dominant Cluster Factors

Primarily African-American | Ethnic Activities | Black Inner City | Weak Standardized Testers

High School Cluster 58

These high schools often serve very wealthy non-Christian religious communities which place a high value on education. Parents are most often professionals and have at least a baccalaureate degree. Students have high aspirations and take advantage of the AP and honors coursework offered. Their standardized test scores are well above average. They apply to a number of institutions, mostly highly selective privates pretty evenly divided between in-state and out-of-state. There is only a moderate interest in financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	577	5	% of Students 1st generation	40%	20	Ave cost targeted colleges (x $1000)	$20.5	6
Mean SAT Math Score	592	7	Ave Admit Rate at Targeted Colleges	45%	26	% of students non-White	27%	24
Mean SAT Writing Score	580	3	Ave Number of AP Exams per Student	0.82	12	% of families below poverty	4%	28
Ave Number of Advanced Courses	1.44	8	% likely to apply out of state	51%	9	% interested in Financial Aid	32%	26

Dominant Cluster Factors

Jewish Culture | Professional and Affluent | College Interest: Private Selective | Coed

High School Cluster 59

These are public high schools serving older, economically depressed, white, blue-collar, suburban communities. While a majority of parents have some college, of the small proportion who have earned degrees most have also earned graduate degrees and are professionals or managers. Students tend to perform well in the classroom, take modest advantage of the advanced courses offered, and have very modest aspirations and standardized test scores. They don't apply to many institutions, but tend to favor less selective publics and community colleges in their home state. Financial aid will be a large factor.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	489	17	% of Students 1st generation	62%	10	Ave cost targeted colleges (x $1000)	$11.5	25
Mean SAT Math Score	499	15	Ave Admit Rate at Targeted Colleges	65%	1	% of students non-White	18%	29
Mean SAT Writing Score	473	15	Ave Number of AP Exams per Student	0.79	16	% of families below poverty	9%	12
Ave Number of Advanced Courses	1.35	10	% likely to apply out of state	19%	28	% interested in Financial Aid	74%	6

Dominant Cluster Factors

Relatively High Grades | Working Class | College Interest: Lower Cost Satellite public | Non-Sectarian Curriculum

High School Cluster 60

The high schools in this cluster are primarily private or sectarian; serving mostly women with professional, college-educated parents who are often from non-Christian communities. Household incomes and home values are above average. Students are academically oriented and perform well in class and on standardized tests, although they are generally uninvolved in AP and honors coursework. They tend to make a few focused applications, usually to moderately priced, relatively selective privates. They tend to stay close to home and have a below average interest in financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	549	6	% of Students 1st generation	61%	11	Ave cost targeted colleges (x $1000)	$17.2	14
Mean SAT Math Score	523	12	Ave Admit Rate at Targeted Colleges	55%	14	% of students non-White	30%	23
Mean SAT Writing Score	544	8	Ave Number of AP Exams per Student	0.42	26	% of families below poverty	9%	14
Ave Number of Advanced Courses	0.59	24	% likely to apply out of state	23%	27	% interested in Financial Aid	33%	25

Dominant Cluster Factors

Jewish Culture | Single Gender | Cost Not an Object | Not Athletic Participant

High School Cluster 61

These are predominantly private high schools serving older, racially-mixed, inner-city communities where some of the population deals with English as a second language. There is an almost equal split between professional, managerial, and blue-collar occupations. Students are exposed to college prep curricula but not AP and honors courses. Standardized test scores are below average and lowest on language-related sections. They aspire beyond the baccalaureate and apply to a small number of moderately selective, private schools. They seem disinterested in financial aid despite very average family incomes, which may suggest that only the most affluent go on to college.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	370	29	% of Students 1st generation	85%	2	Ave cost targeted colleges (x $1000)	$19.2	10
Mean SAT Math Score	485	19	Ave Admit Rate at Targeted Colleges	51%	21	% of students non-White	89%	6
Mean SAT Writing Score	408	26	Ave Number of AP Exams per Student	0.14	29	% of families below poverty	12%	9
Ave Number of Advanced Courses	0.08	29	% likely to apply out of state	33%	19	% interested in Financial Aid	9%	29

Dominant Cluster Factors

Strong Academic Curriculum	Small Private	Other Than Christian Culture	Not Athletic Participant

High School Cluster 62

The high schools in this cluster serve predominantly lower middle class, bilingual Hispanic families with strong traditional values. Many parents have had some experience in higher education which is reflected in a mix of professional, managerial, and blue-collar occupations. Students take a range of college prep offerings and frequently have access to AP and honors level courses, but their standardized test results are below average. Moderately mobile, they tend towards lower cost, relatively selective privates where financial aid will be important.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	473	20	% of Students 1st generation	59%	12	Ave cost targeted colleges (x $1000)	$14.8	19
Mean SAT Math Score	474	20	Ave Admit Rate at Targeted Colleges	47%	25	% of students non-White	92%	5
Mean SAT Writing Score	463	20	Ave Number of AP Exams per Student	0.93	8	% of families below poverty	11%	10
Ave Number of Advanced Courses	1.71	4	% likely to apply out of state	34%	18	% interested in Financial Aid	67%	13

Dominant Cluster Factors

Hispanic	Diverse Low Income	Other Than Mexican	Non-Religious Activities

High School Cluster 63

These public high schools serve an inner-city mix of non-white populations about half of whom speak English as a second language. Often with younger children, the parents have below average incomes, generally do not own their homes, have completed high school or some college, and are in blue-collar or lower level professional jobs. Students have moderate educational goals and are involved in some AP and honors coursework, but score consistently below average on admission tests. They tend to look at in-state publics or reasonably priced and moderately selective privates, from which they will expect financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	427	25	% of Students 1st generation	77%	6	Ave cost targeted colleges (x $1000)	$15.0	18
Mean SAT Math Score	440	25	Ave Admit Rate at Targeted Colleges	52%	18	% of students non-White	85%	9
Mean SAT Writing Score	422	25	Ave Number of AP Exams per Student	0.73	17	% of families below poverty	16%	6
Ave Number of Advanced Courses	0.94	17	% likely to apply out of state	28%	24	% interested in Financial Aid	72%	8

Dominant Cluster Factors

Hispanic	African-American	Less Educated	Relatively Low Grades

High School Cluster 64

The high schools in this cluster are mostly public and serve predominantly younger, Asian families, many of whom are bilingual. The parents have broad experience with higher education, well above average incomes, and hold professional or managerial positions. Students pursue both math/science and liberal arts curricula, take full advantage of AP and honors courses, and score well on standardized tests. Although not overly mobile and with only an average interest in financial aid, they are cost conscious in their consideration and will likely apply at many different colleges across a range of selectivity.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	542	9	% of Students 1st generation	50%	15	Ave cost targeted colleges (x $1000)	$16.8	15
Mean SAT Math Score	606	2	Ave Admit Rate at Targeted Colleges	49%	23	% of students non-White. of students non-White	89%	7
Mean SAT Writing Score	547	7	Ave Number of AP Exams per Student	1.24	2	% of families below poverty families below poverty	7%	21
Ave Number of Advanced Courses	1.45	7	% likely to apply out of state	37%	14	% interested in Financial Aid:rested in Financial Aid	57%	18

Dominant Cluster Factors

Large Asian ESL population	College Prep School	College Interest: Private Selective	Non-Sectarian

High School Cluster 65

These public schools serve relatively diverse, close-in suburbs where affluent younger families with above average incomes have recently moved from the city. Parents are primarily professionals and managers, although there also is a sizeable blue-collar population; most have at least some college experience. Students have modest aspirations and standardized test score but pursue solid academic curricula including a good number of AP and honors courses. They tend towards selective public institutions including in-state flagships and have an average interest in financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	503	14	% of Students 1st generation	47%	17	Ave cost targeted colleges (x $1000)	$11.6	24
Mean SAT Math Score	515	13	Ave Admit Rate at Targeted Colleges	59%	7	% of students non-White	43%	16
Mean SAT Writing Score	491	15	Ave Number of AP Exams per Student	1.01	7	% of families below poverty	7%	22
Ave Number of Advanced Courses	1.65	5	% likely to apply out of state	28%	23	% interested in Financial Aid	65%	14

Dominant Cluster Factors

College Prep Culture | Large Families | Non-Sectarian | New/ Highly Mobile

High School Cluster 66

The high schools in this cluster serve racially-mixed middle class communities with younger children. Most parents have some acquaintance with, if not a degree from, high education and hold jobs from professional to blue-collar. Students are disproportionately women and are involved in a number of extra-curricular activities. They have an academic orientation but do not evidence strong disciplinary interests or educational aspirations; their standardized test performances are not much beyond average. Their college choices are generally less selective, modestly priced privates where they will most likely be seeking financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	515	11	% of Students 1st generation	50%	16	Ave cost targeted colleges (x $1000)	$13.5	22
Mean SAT Math Score	498	16	Ave Admit Rate at Targeted Colleges	61%	5	% of students non-White	37%	20
Mean SAT Writing Score	494	13	Ave Number of AP Exams per Student	0.72	18	% of families below poverty	9%	15
Ave Number of Advanced Courses	1.15	14	% likely to apply out of state	37%	15	% interested in Financial Aid	73%	7

Dominant Cluster Factors

Activist/Community Achievements | Few AP/Honors | Not Athletic Participant | Not Leadership Oriented

High School Cluster 67

The schools in this cluster are most often religiously affiliated and predominantly serve women from older upper middle class communities. Most parents have at least some college and are either professionals or managers. Students are academically oriented and involved in a number of activities, their curricula are solid in both math/science and AP/honors, and they score above average on standardized tests. They have fairly high educational aspirations, are relatively mobile, and apply to a good number of selective privates where financial aid will be sought by many.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	546	7	% of Students 1st generation	37%	21	Ave cost targeted colleges (x $1000)	$20.1	8
Mean SAT Math Score	526	11	Ave Admit Rate at Targeted Colleges	54%	16	% of students non-White	41%	18
Mean SAT Writing Score	547	6	Ave Number of AP Exams per Student	1.08	5	% of families below poverty	7%	20
Ave Number of Advanced Courses	1.84	3	% likely to apply out of state	48%	10	% interested in Financial Aid	69%	11

Dominant Cluster Factors

Single Gender | College Interest: National Selective | Leadership/Organizational Achievements | Art Achievements

High School Cluster 68

Almost exclusively religious, and predominantly Catholic, these high schools serve communities with extensive home ownership and household incomes well above average. Almost all parents have some college and most are either professionals or managers. Students are active in their communities and athletics, and tend to have moderate educational aspirations, solid involvement in AP and honors coursework, and good above average test scores. They apply to a fair number of schools, more in-state than out, and mostly selective, moderately-priced privates and sectarian colleges. Financial aid is on the minds of a majority.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	540	10	% of Students 1st generation	31%	25	Ave cost targeted colleges (x $1000)	$19.0	11
Mean SAT Math Score	541	8	Ave Admit Rate at Targeted Colleges	56%	12	% of students non-White	26%	25
Mean SAT Writing Score	537	9	Ave Number of AP Exams per Student	0.91	10	% of families below poverty	5%	25
Ave Number of Advanced Courses	1.60	6	% likely to apply out of state	41%	12	% interested in Financial Aid	62%	17

Dominant Cluster Factors

Catholic Culture | Financially Constrained | Highly Educated | Coed

High School Cluster 69

These high schools serve very low income, predominantly African-American communities. Although the largest proportion of parents hold blue collar job and have only a high school education, there is also a noticeable professional and managerial presence. Students tend to be active in school and have an academic orientation, although participation in advanced course work is quite low and test scores are near the bottom. Some students will look out of state at somewhat selective, moderately priced privates, but many will choose a public two or four college close to home. Financial aid will be essential for most.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	395	27	% of Students 1st generation	79%	4	% of students non-White	92%	4
Mean SAT Math Score	390	28	Ave Admit Rate at Targeted Colleges	56%	13	% of families below poverty	19%	3
Mean SAT Writing Score	390	28	Ave Number of AP Exams per Student	0.60	21	% interested in Financial Aid	74%	5
Ave Number of Advanced Courses	0.49	26	% likely to apply out of state	36%	16	Ave cost targeted colleges (x $1000)	$14.2	20

Dominant Cluster Factors

Primarily African-American | Single Gender | Black Inner City | Less Educated

High School Cluster 70

These primarily public schools serve established, very affluent suburban communities. Parents overwhelmingly are in professional or managerial positions, with over half having a degree beyond the baccalaureate. Students have access to strong curricula, take advantage of AP and honors coursework, are active and involved, and perform very well on standardized tests. Overwhelmingly committed to earning a degree, they send applications to many highly selective, public and private colleges both in and out of state. Despite the costs associated with their college choices, slightly less than one-half will seek financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	581	3	% of Students 1st generation	22%	29	Ave cost targeted colleges (x $1000)	$21.0	5
Mean SAT Math Score	595	5	Ave Admit Rate at Targeted Colleges	50%	22	% of students non-White	33%	21
Mean SAT Writing Score	580	2	Ave Number of AP Exams per Student	1.23	3	% of families below poverty	4%	29
Ave Number of Advanced Courses	1.91	2	% likely to apply out of state	56%	8	% interested in Financial Aid	48%	20

Dominant Cluster Factors

Professional and Affluent | Good Standardized Testers | Activist/Community Achievements | National Selective

High School Cluster 71

The high schools in this cluster, about one-third of which are private or religiously affiliated, serve very low income Hispanic communities with lots of children. Most parents have had at least some college but largest proportion is in blue-collar occupations. Students tend towards softer coursework but perform well in them; a few get involved with AP and honors courses. Their standardized test scores are near the bottom. They tend to be rather focused in their college choices often looking at either public flagships or somewhat selective, moderately-priced privates where financial aid would be a must.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	412	26	% of Students 1st generation	68%	9	Ave cost targeted colleges (x $1000)	$13.2	23
Mean SAT Math Score	400	27	Ave Admit Rate at Targeted Colleges	52%	19	% of students non-White	98%	2
Mean SAT Writing Score	392	27	Ave Number of AP Exams per Student	0.72	19	% of families below poverty	18%	4
Ave Number of Advanced Courses	0.49	27	% likely to apply out of state	57%	7	% interested in Financial Aid	80%	1

Dominant Cluster Factors

Puerto Rican/Caribbean ESL | Focused/Early Decision | College Interest: Flagship Public | Few AP/Honors

High School Cluster 72

These schools, which are predominantly Christian affiliated and may include homeschoolers, serve upper middle class communities where most families own their homes. Parents work in a variety of vocations across the spectrum and almost all have at least some experience with higher education. Students generally have are exposed to good to above average curricula, are involved in AP and honors coursework, and attain above average standardized test scores. Their educational aspirations are very modest; they apply to fewer schools than most and generally consider less selective, private, church-related institutions. Their interest in financial aid is about average.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	544	8	% of Students 1st generation	32%	24	Ave cost targeted colleges (x $1000)	$15.6	17
Mean SAT Math Score	528	10	Ave Admit Rate at Targeted Colleges	63%	2	% of students non-White	25%	27
Mean SAT Writing Score	527	10	Ave Number of AP Exams per Student	0.82	11	% of families below poverty	6%	24
Ave Number of Advanced Courses	1.10	15	% likely to apply out of state	35%	17	% interested in Financial Aid	64%	15

Dominant Cluster Factors

Religious Activities | Christian Culture | College Interest: Sectarian | Relatively High Grades

High School Cluster 73

The schools in this cluster are generally public and serve urban families with modest incomes and lots of children. Although there is some diversity, families are largely blue collar, with large Mexican and other Hispanic populations, speak English as a second language, and have little or no experience with college. Although they test below average, students avail themselves of academic opportunities and frequently seek out AP and honors coursework. They apply to a reasonable number of public two and four year colleges, mostly within their home state, along with some less selective and relatively low cost privates. Financial aid is seen as being a key to attendance.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	438	23	% of Students 1st generation	80%	3	Ave cost targeted colleges (x $1000)	$10.6	28
Mean SAT Math Score	451	24	Ave Admit Rate at Targeted Colleges	58%	8	% of students non-White	89%	8
Mean SAT Writing Score	435	23	Ave Number of AP Exams per Student	1.03	6	% of families below poverty	17%	5
Ave Number of Advanced Courses	1.37	9	% likely to apply out of state	24%	26	% interested in Financial Aid	76%	4

Dominant Cluster Factors

Mexican | Large Families | Less Educated | Diverse Low Income

High School Cluster 74

These schools are most often private and serve highly educated, relatively small, middle class families. They are more likely to be professional than blue collar, and the largest ethnic group is Asian. Students seek out strong curricula, although their involvement in AP and honors courses is modest. They have extremely high aspirations and score at or near the top on standardized tests. They are highly mobile and apply to a number of institutions, generally to some of the most selective and expensive private colleges. Despite only modest income levels, their interest in financial aid is slightly below average.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	579	4	% of Students 1st generation	30%	26	Ave cost targeted colleges (x $1000)	$28.1	1
Mean SAT Math Score	654	1	Ave Admit Rate at Targeted Colleges	34%	29	% of students non-White	80%	11
Mean SAT Writing Score	579	4	Ave Number of AP Exams per Student	0.80	14	% of families below poverty	10%	11
Ave Number of Advanced Courses	0.61	23	% likely to apply out of state	76%	2	% interested in Financial Aid	46%	21

Dominant Cluster Factors

Large Asian ESL population | College Interest: Private Selective | Higher Ability | Leadership/Organizational Achievements

High School Cluster 75

The schools in this cluster are overwhelmingly public and represent well established small town and rural communities where almost everyone owns a home and households have comfortable incomes. Most parents have traditional values, some experience with college and represent the breadth of the vocational spectrum. Students tend towards basic college prep curricula and only modestly get involved in AP and honors level coursework. Their educational aspirations are low and they are very average testers. They tend to seek colleges close to home that are somewhat selective and moderately priced, where financial aid will be available.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	502	15	% of Students 1st generation	51%	14	Ave cost targeted colleges (x $1000)	$17.4	13
Mean SAT Math Score	514	14	Ave Admit Rate at Targeted Colleges	57%	10	% of students non-White	20%	28
Mean SAT Writing Score	494	14	Ave Number of AP Exams per Student	0.71	20	% of families below poverty	5%	27
Ave Number of Advanced Courses	1.32	12	% likely to apply out of state	31%	21	% interested in Financial Aid	71%	10

Dominant Cluster Factors

Multiple Apps | Athletic Achievements | College Interest: Sectarian | Lower Ability

High School Cluster 76

Overwhelmingly private, the schools in this cluster serve, somewhat racially/ethnically mixed, upper income families with few children. Parents are almost all professionals or managers and highly educated. Students have good curricula which include solid math and science, and some AP and honors level courses. They have test scores at or near the top and generally aspire to education beyond the baccalaureate. They are willing to travel and consider a large number of colleges, generally highly selective and expensive privates where only some will apply for financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	584	2	% of Students 1st generation	32%	23	Ave cost targeted colleges (x $1000)	$27.6	2
Mean SAT Math Score	600	3	Ave Admit Rate at Targeted Colleges	41%	28	% of students non-White	50%	13
Mean SAT Writing Score	587	1	Ave Number of AP Exams per Student	0.81	13	% of families below poverty	5%	26
Ave Number of Advanced Courses	1.06	16	% likely to apply out of state	72%	4	% interested in Financial Aid	28%	27

Dominant Cluster Factors

College Prep School | Affluent | College Interest: Small Private | Older Retired

High School Cluster 77

The schools in this cluster are mostly private and serve highly mobile, mixed inner-city and urban immigrant neighborhoods with significant diversity. Although a significant proportion of the community has very low incomes, parents are most frequently professionals and have college degrees. Students, often dealing with English as a second language, tend towards humanities and social studies courses with some AP and honors work. Their language-based test scores are average, while their math scores are near the top. They have high educational aspirations, are willing to go away to college, and apply to mainly highly selective, private institutions where many will apply for financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	508	13	% of Students 1st generation	30%	27	Ave cost targeted colleges (x $1000)	$22.1	4
Mean SAT Math Score	595	4	Ave Admit Rate at Targeted Colleges	45%	27	% of students non-White	75%	12
Mean SAT Writing Score	518	11	Ave Number of AP Exams per Student	0.92	9	% of families below poverty	19%	2
Ave Number of Advanced Courses	0.80	19	% likely to apply out of state	64%	5	% interested in Financial Aid	39%	23

Dominant Cluster Factors

New/ Highly Mobile | Large Asian ESL population | Black Inner City | Higher Ability

High School Cluster 78

These schools serve small towns and outlying middle class suburbs with little diversity where many people own homes of moderate value. Parents have had some exposure to higher education and, although a significant proportion is blue collar, they pursue a variety of vocations. Students tend towards very traditional curricula without much AP/honors content, have relatively low educational aspirations, and score below average on standardized tests. They will look beyond their home state and do apply to a number of less selective privates. They express relatively little interest in financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	468	21	% of Students 1st generation	79%	5	Ave cost targeted colleges (x $1000)	$18.1	12
Mean SAT Math Score	473	21	Ave Admit Rate at Targeted Colleges	60%	6	% of students non-White	43%	15
Mean SAT Writing Score	458	21	Ave Number of AP Exams per Student	0.24	28	% of families below poverty	8%	19
Ave Number of Advanced Courses	0.53	25	% likely to apply out of state	48%	11	% interested in Financial Aid	22%	28

Dominant Cluster Factors

College Prep School | Other Than Christian Culture | Small Private | Not Athletic Participant

High School Cluster 79

10.1%

The schools in this cluster are almost all public and serve solid, highly educated, middle class communities, perhaps with a local high education presence. Parents almost all have some college experience with many holding post-baccalaureate degrees and pursuing professional or managerial vocations. Students tend towards leadership opportunities while pursuing good grades in curricula with lots of AP and honors coursework. They have high aspirations that often include post-graduate work, and score at or near the top on standardized tests. They are focused in their college choices, willing to venture out of state, and concentrating on selective privates and flagship publics with good financial aid.

Values & Rankings of Key Attributes

	value	rank		value	rank		value	rank
Mean SAT Critical Reading Score	585	1	% of Students 1st generation	27%	28	Ave cost targeted colleges (x $1000)	$19.3	9
Mean SAT Math Score	594	6	Ave Admit Rate at Targeted Colleges	53%	17	% of students non-White	26%	26
Mean SAT Writing Score	570	5	Ave Number of AP Exams per Student	1.94	1	% of families below poverty	6%	23
Ave Number of Advanced Courses	2.55	1	% likely to apply out of state	61%	6	% interested in Financial Aid	71%	9

Dominant Cluster Factors

Higher Ability | College Interest: National Selective | College Interest: Flagship Public | Leadership/Organizational Achievements

REFERENCES

Abdollah, F., Sun, M., Thuret, R., Jeldres, C., Tian, Z., Briganti, A., ... & Karakiewicz, P. I. (2011). A competing-risks analysis of survival after alternative treatment modalities for prostate cancer patients: 1988–2006. *European urology*, *59*(1), 88-95.

Adelman, C. (1999). Answers in the Tool Box. Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Adelman, C. (2006). The Toolbox Revisited: Paths to Degree Completion From High School Through College. *US Department of Education*.

Agresti, A. (2007). *An introduction to categorical data analysis*. John Wiley.

Ahlburg, D., McPherson, M., & Schapiro, M. O. (1994). Predicting higher education enrollment in the United States: an evaluation of different modeling approaches. *DP 26. Williams Project on the economics of higher education*.

Ahlburg, D. A., McCall, B. P., & Na, I. G. (2002). Time to dropout from college: A hazard model with endogenous waiting. *Unpublished paper. Industrial Relations Center, University of Minnesota*.

Akin, B. A. (2011). Predictors of foster care exits to permanency: A competing risks analysis of reunification, guardianship, and adoption. *Children and Youth Services Review*, *33*(6), 999-1011.

Allison (2002). Missing Data. Sage Publications.

Allison (2012). Handling Missing Data by Maximum Likelihood. SAS Global Forum: Statistics and Data Analysis.

Allison, P. D. (2010). *Survival analysis using SAS: a practical guide*. SAS Institute.

Alkin, M. (2013). *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks, CA: Sage.

Austin, P. C., Lee, D. S., & Fine, J. P. (2016). Introduction to the analysis of survival data in the presence of competing risks. *Circulation*, *133*(6), 601-609.

Avery, C., & Hoxby, C. M. (2004). Do and should financial aid packages affect students' college choices?. In *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 239-302). University of Chicago Press.

Avery, C. and Thomas, K. (2004). “Student Perceptions of College Opportunities: The Boston COACH Program.” *In College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*, ed. Caroline M. Hoxby, 355–393. Chicago: University of Chicago Press.

Baer, H. J., Glynn, R. J., Hu, F. B., Hankinson, S. E., Willett, W. C., Colditz, G. A., Sampfer, M., & Rosner, B. (2011). Risk factors for mortality in the nurses’ health study: a competing risks analysis. *American journal of epidemiology, 173*(3), 319-329.

Bahr, P. R. (2009). Educational attainment as process: Using hierarchical discrete-time event history analysis to model rate of progress. *Research in Higher Education*, *50*(7), 691-714.

Bain, R. K., & Anderson, J. G. (1974). School context and peer influences on educational plans of adolescents. *Review of Educational Research*, *44*(4), 429-445.

Becker, G. S. (1993). *Human capital: A theoretical and empirical analysis with special reference to education* (3rd ed.). Chicago, IL: University of Chicago Press.

Belot, A., Abrahamowicz, M., Remontet, L., & Giorgi, R. (2010). Flexible modeling of competing risks in survival analysis. *Statistics in medicine*, *29*(23), 2453-2468.

Bianchi, M., Gandaglia, G., Trinh, Q. D., Hansen, J., Becker, A., Abdollah, F., Tian, Z., Lughezzani, G., Roghmann, F., Briganti, A., Montorsi, F., Karakiewicz, P., & Montorsi, F. (2014, January). A population-based competing-risks analysis of survival after nephrectomy for renal cell carcinoma. In *Urologic Oncology: Seminars and Original Investigations* (Vol. 32, No. 1, pp. 46-e1). Elsevier.

Bruggink, T. H., & Gambhir, V. (1996). Statistical models for college admission and enrollment: A case study for a selective liberal arts college. *Research in Higher Education*, *37*(2), 221-240.

Bush, V. B. (2009). *A predictive model of Hispanic participation in Texas higher education: Inferences drawn from institutional data in prevalent Hispanic states* (Doctoral dissertation, University of North Texas).

Calcagno, J. C., Crosta, P., Bailey, T., & Jenkins, D. (2007). Stepping stones to a degree: The impact of enrollment pathways and milestones on community college student outcomes. *Research in Higher Education*, *48*(7), 775-801.

Carnevale, A. P., & Strohl, J. (2013). Separate and unequal: How higher education reinforces the intereneratіonal reproduction of white racial privilege. *Washington DC: Georgetown University Center on Education and the Workforce*.

Card, D., & Lemieux, T. (2001). Dropout and enrollment trends in the postwar period: What went wrong in the 1970s?. In *Risky behavior among youths: An economic analysis* (pp. 439-482). University of Chicago Press.

Chang, L. (2006). Applying data mining to predict college admissions yield: A case study. *New Directions for Institutional Research*, *2006*(131), 53-68.

Chen, R. (2012). Institutional characteristics and college student dropout risks: A multilevel event history analysis. *Research in Higher Education*, *53*(5), 487-505.

Cho, D. (2007). The role of high school performance in explaining women's rising college enrollment. *Economics of Education Review*, *26*(4), 450-462.

Colby, S. L., & Ortman, J. M. (2015). Projections of the Size and Composition of the US Population: 2014 to 2060. *US Census Bureau*, *Ed*, 25-1143.

Collett, D. (2015). *Modelling survival data in medical research*. CRC press.

Colman, M., Choi, L., Chen, A., Crossett, L., Tarkin, I., & McGough, R. (2014). Proximal femoral replacement in the management of acute periprosthetic fractures of the hip: a competing risks survival analysis. *The Journal of arthroplasty*, *29*(2), 422-427.

Conger, D., & Long, M. C. (2013). Gender Gaps in College Enrollment The Role of Gender Sorting Across Public High Schools. *Educational Researcher*, 42(7), 371-380.

Dasgupta, P., Turrell, G., Aitken, J. F., & Baade, P. D. (2016). Partner status and survival after cancer: A competing risks analysis. *Cancer epidemiology*, *41*, 16-23.

de Glas, N. A., Kiderlen, M., Vandenbroucke, J. P., de Craen, A. J., Portielje, J. E., van de Velde, C. J., Liefers, G.J., Bastiaannet, E., & Le Cessie, S. (2016). Performing survival analyses in the presence of competing risks: A clinical example in older breast cancer patients. *Journal of the National Cancer Institute*, *108*(5), djv366.

DesJardins, S. L. (2002). An analytic strategy to assist institutional recruitment and marketing efforts. Research in Higher Education 43(5): 531–553.

DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (1999). An event history model of student departure. *Economics of education review*, *18*(3), 375-390.

DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (2002). A temporal investigation of factors related to timely degree completion. *Journal of Higher Education*, 555-581.

DesJardins, S. L., & Lindsay, N. K. (2008). Adding a statistical wrench to the "toolbox". *Research in Higher Education*, *49*(2), 172-179.

DesJardins, S. L., & Toutkoushian, R. K. (2005). Are students really rational? The development of rational thought and its application to student choice. In J. C. Smart (Ed.), *Higher* *education: Handbook of theory and research* (Vol. 20, pp. 191-240). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Diaz-Jiminez, J., Quadrini, V., & Rios-Rull, J. V. (1997, Spring). Dimensions of Facts on the U.S. distribution of earnings, income, and wealth. Federal Bank of Minneapolis Quarterly Review, 3-21.

Dignam, J. J., & Kocherginsky, M. N. (2008). Choice and interpretation of statistical tests used when competing risks are present. *Journal of Clinical Oncology*, *26*(24), 4027-4034.

Dignam, J. J., Zhang, Q., & Kocherginsky, M. (2012). The use and interpretation of competing risks regression models. *Clinical Cancer Research*, *18*(8), 2301-2308.

Doyle, W. R. (2006). Adoption of merit-based student grant programs: An event history analysis. *Educational evaluation and policy analysis*, *28*(3), 259-285.

Edwards, K., Sawtell, E., Perfetto, G., Shapiro, D. T., & Kelly, P. J. (2013). Landscape of Higher Education: Human Capital. *College Board*.

Engberg, M. E., & Wolniak, G. C. (2010). Examining the effects of high school contexts on postsecondary enrollment. *Research in Higher Education*, 51(2), 132-153.

ESTELA ZARATE, M. A. R. I. A., & Gallimore, R. (2005). Gender differences in factors leading to college enrollment: A longitudinal analysis of Latina and Latino students. *Harvard Educational Review*, 75(4), 383-408.

Finance, N. E., Schultz, J. L., & Mueller, D. (2006). Effectiveness of programs to improve postsecondary education enrollment and success of underrepresented youth.

Fitzgerald, B. (2004). Federal financial aid and college access. In E. P. St. John (Ed.), *Public policy and college access: Investigating the federal and state roles in equalizing postsecondary opportunity* (Vol. 19, pp. 1-28). New York, NY: AMS Press, Inc.

Gail, M. (1975). A review and critique of some models used in competing risk analysis. *Biometrics*, 209-222.

Gándara, P., & Bial, D. (2001). *Paving the way to postsecondary education: K-12 interventions for underrepresented youth*. Washington, D.C.: National Center for Education Statistics.

Guerin, K. J. (1997, May 18-21). Tracking student stop-out, transfer, and graduation: An event history analysis of competing risks. Paper presented at the 37th annual of the Association for Institutional Research, Lake Buena

Gillam, M. H., Ryan, P., Graves, S. E., Miller, L. N., de Steiger, R. N., & Salter, A. (2010). Competing risks survival analysis applied to data from the Australian Orthopaedic Association National Joint Replacement Registry. *Acta orthopaedica*, *81*(5), 548-555.

Glynn, R. J., & Rosner, B. (2005). Comparison of risk factors for the competing risks of coronary heart disease, stroke, and venous thromboembolism. *American journal of epidemiology*, *162*(10), 975-982.

Godfrey, K. E., & Matos-Elefonte, H. (2010). Key Indicators of College Success: Predicting College Enrollment, Persistence, and Graduation. *College Board*.

Goenner, C. F., & Pauls, K. (2006). A predictive model of inquiry to enrollment. *Research in Higher education*, *47*(8), 935-956.

Goldrick-Rab, Sara, 2006. Following their every move: an investigation of social-class differences in college pathways. Sociology of Education 79 (1), 61–79.

Goyette, K. A. (2008). College for some to college for all: Social background, occupational expectations, and educational expectations over time. *Social Science Research, 37*(2), 461-484.

Gross, J. P., & Torres, V. (2010). Competing Risks or Different Pathways? An Event History Analysis of the Relationship between Financial Aid and Educational Outcomes for Latinos. *Center for Enrollment Research, Policy, and Practice*.

Haller, B., Schmidt, G., & Ulm, K. (2013). Applying competing risks regression models: an overview. *Lifetime data analysis*, *19*(1), 33-58.

Harvill, J. L. (2012). An Introduction to Mathematical Statistics and Its Applications: Richard J. Larsen and Morris L. Marx. Boston: Prentice Hall, 2011.

Harvill, E. L., Maynard, R. A., Nguyen, H. T., Robertson-Kraft, C., & Tognatta, N. (2012). Effects of College Access Programs on College Readiness and Enrollment: A Meta-Analysis. *Society for Research on Educational Effectiveness*.

Hearn, J. C. (1992). Emerging variations in postsecondary attendance patterns: An investigation of part-time, delayed, and nondegree enrollment. *Research in Higher Education*, 33(6), 657-687.

Hidalgo, B., & Goodman, M. (2013). Multivariate or multivariable regression?. *American journal of public health*, *103*(1), 39-40.

Hill, D. H. (2008). School strategies and the "college-linking" process: Reconsidering the effects of high schools on college enrollment. *Sociology of education*, 81(1), 53-76.

Hogan, A., & Roberts, B. (2015). Occupational employment projections to 2024. *Monthly Lab. Rev.*, *138*, 1. United States Department of Labor: Bureau of Labor Statistics.

Hurwitz, M., & Kumar, A. (2015). Supply and Demand in the Higher Education Market: College Admission and College Choice. Research Brief. *College Board*.

Hussar, W. J., & Bailey, T. M. (2011). Projections of education statistics to 2019. US Department of Education. *National Center for Education Statistics*.

Jaffe, B. (2005). Older and wiser: An event history analysis of women's adult college enrollment behavior.

Jia, P., & Maloney, T. (2015). Using predictive modelling to identify students at risk of poor university outcomes. *Higher Education*, *70*(1), 127-149.

Johnson, I. (2008). Enrollment, Persistence and Graduation of In-State Students at a Public Research University: Does High School Matter?. *Research in Higher Education*, 49(8), 776-793.

Johnson, L. K. (2011). High School Influences on College Enrollment and College Graduation of Dallas Independent School District 2002 Graduates.

Johnson and Young (2011). Towards Best Practices in analyzing Datasets with Missing Data: Comparisons and Recommendations. Journal of Marriage and Family, 73(5): 926-45.

Khattab, N. (2005). The effects of high school context and interpersonal factors on students' educational expectations: A multi-level model. *Social Psychology of Education*, *8*(1), 19-40.

Kim, H. T. (2007). Cumulative incidence in competing risks data and competing risks regression analysis. *Clinical Cancer Research*, *13*(2), 559-565.

Kim, J. Y. (2011). *An Analysis of the Effects of State Financial Aid Policy on the Timing of Postsecondary Enrollment: A Focus on Income and Race Differences* (Doctoral dissertation, University of Michigan).

Kim, J. (2012). Exploring the relationship between state financial aid policy and postsecondary enrollment choices: A focus on income and race differences. *Research in Higher Education*, *53*(2), 123-151.

Kleinbaum, D. G., & Klein, M. (2006). Survival analysis: a self-learning text. Springer Science & Business Media.

Krogstad, J. M., & Fry, R. (2014). More Hispanics, Blacks enrolling in college, but lag in bachelor's degrees. *Washington, DC: Pew Research Center. Retrieved from http:// www. pewresearch. org/ fact-tank*, *4*, 24.

Ledesma, R. (2009). Predictive modeling of enrollment yield for a small private college. *Atlantic Economic Journal*, *37*(3), 323-325.

Lin, M. M. (2011). *Intersections of Race, SES, and First-Generation College Student Status in* Understanding the Factors Affecting Undergraduate Academic Persistence: A

Psychosociocultural Approach. ProQuest LLC. 789 East Eisenhower Parkway, PO Box 1346, Ann Arbor, MI 48106.

Linsenmeier, D., Rosen, H., & Rouse, C. (2006). Financial aid packages and college enrollment decisions: An econometric case study. *Review of Economics and Statistics*, *88*(1), 126–145.

Long, B. (2004). Does the format of a financial aid program matter? The effect of state-in-kind tuition subsidies. *Review of Economics and Statistics*, *86*(3), 767–782.

Lu, W., & Peng, L. (2008). Semiparametric analysis of mixture regression models with competing risks data. *Lifetime data analysis*, *14*(3), 231-252.

Lughezzani, G., Sun, M., Shariat, S. F., Budäus, L., Thuret, R., Jeldres, C., Liberman, D., Montorsi, F., Perrrotte, P., & Karakiewicz, P. I. (2011). A population- based competing-risks analysis of the survival of patients treated with radical cystectomy for bladder cancer. *Cancer*, *117*(1), 103-109.

Mitchell, M., & Leachman, M. (2015). Years of cuts threaten to put college out of reach for more students. *Center on Budget and Policy Priorities*, 1-26.

Monks, J. (2009). The impact of merit-based financial aid on college enrollment: A field experiment. *Economics of Education Review*, *28*(1), 99-106.

Murphy, K., & Welch, F. (1993). Inequality and relative wages. American Economic Re-view, 83, 104-109.

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

Nandeshwar, A. R. (2010). *Longitudinal Study of First-Time Freshmen Using Data Mining*. ProQuest LLC. 789 East Eisenhower Parkway, PO Box 1346, Ann Arbor, MI 48106.

Noordzij, M., Leffondré, K., van Stralen, K. J., Zoccali, C., Dekker, F. W., & Jager, K. J. (2013). When do we need competing risks methods for survival analysis in nephrology?. *Nephrology Dialysis Transplantation*, *28*(11), 2670-2677.

Patton, M.Q. *Utilization-Focus Evaluation*. 4th Edition, Thousand Oaks, CA: Sage 2008.

Paulsen, M. B. (1990). *College Choice: Understanding Student Enrollment Behavior. ASHE-ERIC Higher Education Report No. 6*. ASHE-ERIC Higher Education Reports, The George Washington University, One Dupont Circle, Suite 630, Dept. RC, Washington, DC 20036-1183.

Paulsen, M. B. (2001). The economics of human capital and investment in higher education. *The finance of higher education: Theory, research, policy, and practice*, 55-94.

Paulsen, M. B., & St John, E. P. (2002). Social class and college costs: Examining the financial nexus between college choice and persistence. *The Journal of Higher Education*, *73*(2), 189-236.

Perna, L. W. (2006). Studying college access and choice: A proposed conceptual model. In *HIGHER EDUCATION:* (pp. 99-157). Springer Netherlands.

Perna, L.W., & Swail, W.S. (2002). Pre-college outreach and early intervention programs. In D.E. Heller (Ed.), *Condition of access: Higher education for lower income students* (pp. 97-112). Westport, CT: American Council on Education and Praeger Publishers.

Peters, T. J. (2008). Multifarious terminology: multivariable or multivariate? univariable or univariate?. *Paediatric and perinatal epidemiology*, *22*(6), 506-506.

Peterson, A. V. (1976). Bounds for a joint distribution function with fixed sub-distribution functions: Application to competing risks. *Proceedings of the National Academy of Sciences*, *73*(1), 11-13.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods (Vol. 1). Sage.

Reynolds, John R., Stewart, Michael, MacDonald, Ryan, Sischo, Lacey, 2006. Have adolescents become too ambitious? high school seniors educational and occupational plans, 1976 to 2000. Social Problems 53 (2), 186–206.

Rodríguez-Camacho, E., Pita-Fernández, S., Pértega-Díaz, S., López-Calviño, B., & Seoane-Pillado, T. (2015). Clinical-Pathological Characteristics and Prognosis of a Cohort of Oesophageal Cancer Patients: a Competing Risks Survival Analysis. *Journal of Epidemiology*, *25*(3), 231.

Ronco, S. L. (1996). How enrollment ends: Analyzing the correlates of student graduation, transfer, and dropout with a competing risks model. AIR Professional File, Number 61.

Ruffalo Noel Levitz (2009). 2009 Cost of Recruiting Report: Comparative Benchmarks for Two-Year and Four-Year Institutions. Cedar Rapids, Iowa: Ruffalo Noel Levitz. Retrieved from www.RuffaloNL.com/BenchmarkReports.

Ruffalo Noel Levitz (2009). *Retooling the Enrollment Funnel: Strategies and Metrics for a New Era*. Cedar Rapids, Iowa: Ruffalo Noel Levitz. Retrieved from www.RuffaloNL.com/BenchmarkReports.

Ruffalo Noel-Levitz (2012). Why did they enroll? The factors influencing college choice. Coralville, Iowa: Author. Retrieved from www.noellevitz.com/Factors2012.

Ruffalo Noel Levitz (2015). 2015 student retention and college completion practices report for four-year and two-year institutions. Coralville, Iowa: Ruffalo Noel Levitz. Retrieved from www.noellevitz.com/BenchmarkReports.

Ruffalo Noel Levitz (2016). *2016 marketing and student recruitment practices benchmark report* for *four-year colleges and universities*. Cedar Rapids, Iowa: Ruffalo Noel Levitz. Retrieved from www.RuffaloNL.com/BenchmarkReports.

Ruffalo Noel Levitz (2016). 2016 report: cost of recruiting an undergraduate student for four-year and two-year institutions. Cedar Rapids, Iowa: Ruffalo Noel Levitz. Retrieved from www.RuffaloNL.com/BenchmarkReports.

Ruffalo Noel Levitz (2016). *How to respond to new FAFSA rules for prior-prior year (PPY) income tax information taking effect October 1, 2016*. Coralville, Iowa: Ruffalo Noel Levitz. Retrieved from www.RuffaloNL.com/PriorPriorYear

Russ-Eft D. & Preskill, H. (2009). *Evaluation in Organizations: A Systematic Approach to Enhancing Learning, Performance, and Change*. 2nd Edition. New York, NY: Basic Books.

Ryan, C. L., & Bauman, K. (2016). Educational Attainment in the United States: 2015. United States Department of Commerce: U.S. Census Bureau.

Schuh, R., Kaider, A., Windhager, R., & Funovics, P. T. (2015). Does competing risk analysis give useful information about endoprosthetic survival in extremity osteosarcoma?. *Clinical Orthopaedics and Related Research®, 473*(3), 900-906.

Scott, M. A., & Kennedy, B. B. (2005). Pitfalls in pathways: Some perspectives on competing risks event history analysis in education research. *Journal of Educational and Behavioral Statistics*, *30*(4), 413-442.

Singell, L. (2002). Merit, need, and student self-selection: Is there discretion in the packaging of aid at a large public university? *Economics of Education Review*, *21*(5), 445–454.

Smith, J. (2011). Can Applying to More Colleges Increase Enrollment Rates? Research Brief. *College Board*.

Snyder, T. D., & Dillow, S. A. (2010). *Digest of Education Statistics, 2009* (NCES 2010-013). Washington, DC: National Center for Education Statistics.

Stater, M. (2004). *Strategic tuition and financial aid policies: Implications for enrollment and graduation*. Ph.D. Dissertation, University of Oregon.

St. John, E. P., Musoba, G. D., & Simmons, A. B. (2003). Keeping the promise: The impact of Indiana's Twenty-first Century Scholars Program. *The Review of Higher Education, 27*(1), 103-123.

Swail, W.S. (2001). Educational opportunity and the role of pre-college outreach programs. *College Board Outreach Program Handbook.* Washington, D.C.: Educational Policy Institute. Retrieved July 25, 2006, from http://www.educationalpolicy.org/pdf/OutreachHandbookEssays.pdf

Thomas, E., Alexander, K. L., and Eckland, B.K. 1979. "Access to Higher Education: The Importance of Race, Sex, Social Class, and Academic Credentials." School Review, vol. 87, pp. 133–156.

Thomas, E., Dawes, W., and Reznik, G. (2001). Using predictive modeling to target student recruitment: Theory and practice. AIR Professional File 78 (Winter): 1–8.

Thomas, E., Reznik, G., & Dawes, W. (1999). Using Predictive Modeling To Target Student Recruitment: Theory and Practice. AIR 1999 Annual Forum Paper.

Tsai, A. C. (2013). Achieving consensus on terminology describing multivariable analyses. *American journal of public health*, *103*(6), e1.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, *72*(1), 20-22.

Warburton, E. C., Bugarin, R., & Nunez, A. M. (2001). Bridging the Gap: Academic Preparation and Postsecondary Success of First-Generation Students. Statistical Analysis Report. Postsecondary Education Descriptive Analysis Reports.

Weisberg, S. (2005). *Applied linear regression* (Vol. 528). John Wiley & Sons.

Wyatt, J. N., Smith, K., & Proestler, N. (2014). *The benefits of early engagement in the college preparation process: Implications for practitioners* (pp. 367-381). College Board Research Report 2014-1). New York: The College Board. Morgan, DL (2012). College placement testing of entering students. In C. Secolsky & DB Denison (Eds.), Handbook on measurement, assessment, and evaluation in higher education.

Yang, F. (2000). Using Survival Analysis To Analyze and Predict Students' Achievement from Their Status of Developmental Study.

VITA

Brendan Martin graduated from the University of Notre Dame in South Bend, Indiana in 2007 with a Bachelor of Arts degree in Political Science. In 2009 he earned a Master's degree from the University of Chicago. Brendan began his doctoral studies in the Research Methodology program at Loyola University Chicago in the fall of 2013. Brendan is an analyst with more than nine years' experience in quantitative modeling and consulting. His professional background includes three years in communications and market research in the private sector, followed by nearly seven years in higher education and clinical research. Currently he is a Senior Biostatistician in Loyola University Chicago's Clinical Research Office in Maywood, IL.