

©2012

La Reina J. Bates

ALL RIGHTS RESERVED

AN EVENT HISTORY ANALYSIS OF TIME TO DEGREE COMPLETION

by

LA REINA J. BATES

A Dissertation submitted to the  
Graduate School-New Brunswick  
Rutgers, The State University of New Jersey  
in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Education

written under the direction of

Dr. Douglas Penfield and Dr. Chia-Yi Chiu

and approved by

---

---

---

---

New Brunswick, New Jersey

MAY, 2012

## ABSTRACT OF THE DISSERTATION

An Event History Analysis of Time to Degree Completion

By LA REINA J. BATES

Dissertation Directors:

Dr. Douglas Penfield and Dr. Chia-Yi Chiu

In an era of increasing demand for college, declining fiscal resources, and the rising costs of undergraduate education, student retention and graduation, especially timely graduation, are important issues facing American higher education today. As state and federal lawmakers, accrediting agencies, and governing bodies demand more accountability for retention and graduation rates from college and university administrators, it is important to develop a better understanding of college student graduation behavior at the institutional level. The study of college student retention and persistence to degree completion has been plagued with methodological problems and inconsistent findings, especially when the longitudinal nature of the process is considered. Event history analysis is a regression-like technique that allows researchers to investigate the timing of graduation while addressing many of the concerns associated with the longitudinal study of college graduation behavior, such as censored cases and time-varying variables. The present study used event history analysis to understand the temporal dimensions of graduation and the factors that affect whether students succeed or fail, particularly at the study institution. Pre-enrollment, enrollment, and financial aid variables were used to model the timing of graduation for three cohorts of first-time, full-

time, degree-seeking undergraduate students for a six year period. Consistent with other studies employing event history analysis to student retention and degree completion, adding a time dimension improves our understanding of event occurrence. The present study also provides support for the strong relationship between the longitudinal effects of academic performance while in college (as measured by cumulative GPA) and graduation.

## Acknowledgements

Although the following dissertation is an individual work, many people have contributed to its production. I owe my deepest gratitude to all those people who have made this dissertation possible.

First, I would like to start off by thanking all of the members of my dissertation committee for generously giving their time and guidance to better my work. To my co-chairs, Dr. Douglas Penfield and Dr. Chia-Yi Chiu, thank you for learning about event history analysis so I could conduct my research and carefully reading and commenting on several revisions of my work. To Dr. Florence Hamrick, thank you for your willingness to serve on my committee and the insightful comments you provided. To Dr. Stephen DesJardins, thank you for lending your expertise on the subject and helping me develop a greater understanding of and appreciation for event history analysis.

Even though it has been six years since I left campus with my first postsecondary degree, thank you to all of my Wells professors that challenged me to work harder and inspired me to want more. To my undergraduate advisor, Dr. Milene Morfei, I am indebted to you for sparking my interest in quantitative research methods, which ultimately led me to Rutgers. Your continuous encouragement and guidance during my time at Wells and beyond has not been forgotten.

I am grateful to my boss, Dr. Paul Johnson, for his practical advice throughout the dissertation process. Many thanks to Ken Iuso, Phyllis Micketti, and Jean McDonald Rash for granting me access to the student data for my research. I would also like to express my sincerest gratitude to Carlos Lescano for working through several data requests to make sure my data was matched correctly. To Gary Nelson, thank you for all

the technical assistance, particularly that day when I discovered my dataset had become corrupted. To my coworkers, thank you for your continued interest in my research and celebrating with me all the little accomplishments along the way. It has made what seemed at times an impossible endeavor all the more possible.

Many friends have also helped make this dissertation reach its conclusion. Jenn, Hebbah, Karlyn, Lynne, and Nats, thank you for your support throughout my time at Rutgers. Your friendship has helped me overcome setbacks and keep moving forward. To Eric and Enid, you will forever be my New Jersey family and I have truly cherished our nightly dinners together. Your friendship has brought both stability and laughter to my life over the past year. To my Wells girls, thank you for always being there with words of encouragement and welcomed distractions all these years. Your friendship reminds me daily of what it truly means to be a Wells Woman.

Most importantly, none of this would have been possible without the love and support of my family. To my parents and siblings, thank you for encouraging me to see this process through to the end. To my nephews, Kole and Ethan, and my niece, Madelyn, thank you for your unconditional love whether I am near or far. Although you are just starting your educational careers, I expect each of you to go to college. It is here, in my greatest achievement to date, that I promise to help you in any way that I can to navigate the process of discovering who you are and what you want to be when you grow up. And finally, a very special thank you to my mom, to whom this dissertation is dedicated, for all the check-in phone calls when she hadn't heard from me in a few days.

## Dedication

For my mom.

## Table of Contents

|   |     |
|---|-----|
| ABSTRACT OF THE DISSERTATION.....   | ii  |
| Acknowledgements.....   | iv  |
| Dedication.....   | vi  |
| Table of Contents.....  | vii |
| List of tables.....   | ix  |
| List of figures.....  | x   |
| CHAPTER I. STATEMENT OF THE PROBLEM.....  | 1   |
| CHAPTER II. LITERATURE REVIEW.....  | 14  |
| College Student Retention and Theoretical Models of Student Departure.....            | 15  |
| Spady's Sociological Model of Dropout.....  | 15  |
| Tinto's Student Integration Model.....  | 18  |
| Bean's Student Attrition Model.....   | 19  |
| Converging the Student Integration Model and Student Attrition Model.....             | 22  |
| Braxton's Theory of Student Departure from Residential Colleges and Universities..... | 23  |
| The Study of College Student Departure and Persistence.....                           | 26  |
| Methodological Criticisms in the Study of College Student Retention.....              | 30  |
| Operational Definitions.....  | 30  |
| Methodological Limitations.....   | 33  |
| Using Event History Analysis to Study College Student Retention.....                  | 37  |
| Time to First Dropout.....  | 40  |
| Time to Graduation.....   | 45  |
| Summary.....  | 48  |
| CHAPTER III. STUDY DESIGN AND METHODOLOGY.....  | 49  |
| Event History Modeling.....   | 49  |
| Censoring.....  | 49  |
| Survival Function.....  | 51  |
| Hazard Function.....  | 52  |
| Time-Dependent Variables.....   | 53  |
| Structuring the Data Set.....   | 53  |
| Competing Risks.....  | 58  |
| A Statistical Model of Hazard.....  | 59  |
| Sample.....   | 62  |
| Variable Selection.....   | 65  |
| Pre-Enrollment Variables.....   | 65  |
| Enrollment Variables.....   | 67  |
| Financial Aid Variables.....  | 69  |
| Data Analysis.....  | 71  |
| CHAPTER IV. RESULTS.....  | 72  |
| Descriptive Statistics.....   | 72  |
| Survivor Function and Survivor Probabilities.....                                     | 75  |
| Hazard Function and Hazard Probabilities.....   | 77  |
| Model of Time to Degree Completion.....   | 78  |



|   |    |
|---|----|
| CHAPTER V. SUMMARY AND CONCLUSIONS.....               | 89 |
| Policy Implication.....                               | 93 |
| Limitations and Implications for Future Research..... | 95 |
| Conclusion.....                                       | 98 |
| REFERENCES.....                                       | 99 |

## Lists of tables

|   |    |
|---|----|
| Table 3.1. <i>Effective Sample Size by Cohort</i> .....                             | 65 |
| Table 3.2. <i>Time to Degree Completion Covariates</i> .....                        | 71 |
| Table 4.1. <i>Postsecondary Output of the Sample</i> .....                          | 72 |
| Table 4.2. <i>Descriptive Statistics of the Sample at Semester 1</i> .....          | 74 |
| Table 4.3. <i>Descriptive Statistics of Financial Aid Package</i> .....             | 75 |
| Table 4.4. <i>Sample Survival Probabilities for Time to Degree Completion</i> ..... | 76 |
| Table 4.5. <i>Sample Hazard Probabilities for Time to Degree Completion</i> .....   | 78 |
| Table 4.6. <i>Model Summary for Complete Time to Degree Completion Model</i> .....  | 80 |

## List of figures

|  |           |
|--|-----------|
| <i>Figure 3.1. An Example of Right Censoring.....</i>  | <i>50</i> |
| <i>Figure 3.2. Restructuring a Person-Level Data Set into a Person-Period Data Set.....</i>                        | <i>57</i> |
| <i>Figure 4.1. Sample Survivor Function for Time to Degree Completion.....</i>                                     | <i>76</i> |
| <i>Figure 4.2. Sample Hazard Function for Time to Degree Completion.....</i>                                       | <i>77</i> |
| <i>Figure 4.3. Sample Survivor Function for Time to Degree Completion by Cohort.....</i>                           | <i>82</i> |
| <i>Figure 4.4. Sample Survivor Function for Time to Degree Completion by Gender.....</i>                           | <i>83</i> |
| <i>Figure 4.5. Sample Survivor Function for Time to Degree Completion by Race/<br/>Ethnicity.....</i>              | <i>84</i> |
| <i>Figure 4.6. Sample Survivor Function for Time to Degree Completion by Residency....</i>                         | <i>85</i> |
| <i>Figure 4.7. Sample Survivor Function for Time to Degree Completion by Initial<br/>School of Enrollment.....</i> | <i>87</i> |
| <i>Figure 4.8. Sample Survivor Function for Time to Degree Completion by School-<br/>to-School Transfers.....</i>  | <i>87</i> |
| <i>Figure 5.1. Simulated Longitudinal Graduation Hazard.....</i>   | <i>95</i> |

## CHAPTER I. STATEMENT OF THE PROBLEM

In an era of increasing demand for college, declining fiscal resources, and the rising costs of undergraduate education, student retention and graduation, especially timely graduation, are important issues facing American higher education today. With politicians and higher education leaders challenging colleges and universities to do a better job of graduating their students, college graduation has also become an important part of the national agenda. During his remarks on higher education and the economy, President Barack Obama (2010) identified education as the economic issue of our time commenting that “in the coming decades, a high school diploma is not going to be enough. [Americans] need a college degree. They need workforce training. They need a higher education.” Although the United States spends more money on higher education than any other nation in the world, only about 60 percent of college students graduate from four-year institutions within six years (Schneider, 2010). Commenting on the nation’s low college graduation rate, President Obama (2010) argued that “we don’t just need to open the doors of colleges to more Americans; we need to make sure they stick with it through graduation.” Research has shown that students who leave college before completing their degree program not only personally suffer negative consequences, but they also pass those consequences on to society and the institution itself (Pascarella & Terenzini, 1991).

Higher education has traditionally been viewed as a public good in the United States, contributing to society by educating citizens, improving human capital, boosting economic development, and encouraging civic engagement (Altbach, Reisberg, & Rumbley, 2010). The attainment of a bachelor’s degree is inextricably linked to social

mobility in the United States, mediating the influence of an individual's background resources, such as family socioeconomic status, on subsequent occupation, income, and social status attainment (Pascarella & Terenzini, 1991). Today more than ever, postsecondary education is the gateway to access the status and earnings of the American middle and upper classes (Carnevale, Smith, & Strohl, 2010). In a recent analysis of workforce education requirements in the United States, Carnevale et al.'s (2010) calculations show that approximately 60 percent of jobs today require a postsecondary education and that employer demand for workers with postsecondary degrees will continue to grow over the next decade. As employers increasingly depend on advanced skills and knowledge obtained through higher education, the need to improve degree completion rates is essential to economic growth and job creation. Less likely to be unemployed and earning higher salaries, college graduates generate higher tax payments at the local, state, and federal levels over their lifetime (Baum, Ma, & Payea, 2010; Belfield & Levin, 2007; Hurley, 2009). In addition to increasing the tax base, a college-educated workforce reduces government spending on health care, criminal justice, and welfare (Belfield & Levin, 2007). Higher education is also an investment in our societal collective as college graduates are more likely than non-graduates to participate in civic responsibilities, such as community service and voting (Bureau of Labor Statistics, 2010b; U.S. Census Bureau, 2010b).

Although service to society and civic responsibility are among the most important themes of higher education, public resistance to taxes at the local, state, and federal level and higher education's inability to compete with other social priorities (e.g., health care, K-12 education) has shifted the conception of postsecondary education from a public

good to a private investment (Duderstadt & Womack, 2003). The recession is speeding up the shift in who pays for a college education, and the burden is increasingly falling on students and their families as a substantial portion of financial aid packages reflect student loans as opposed to gift aid (Baum, Payea, & Cardenas-Elliott, 2010). The average student loan debt for 2008-09 college graduates was almost \$24,000, with those attending private non-profit colleges and universities borrowing approximately \$7,000 more on average than their peers attending public institutions (Cheng & Reed, 2010). As student debt is mounting, it is imperative that college students who take on significant debt to finance their postsecondary education leave the higher education system with a college degree so they can realize the economic benefits of obtaining a degree to avoid defaulting on their student loans (Gladieux & Perna, 2005). For all demographic groups, average earnings increase measurably with higher levels of education and the gap in earnings between college graduates and those who have only completed high school is growing (Baum & Ma, 2010; Lumina Foundation for Education, 2010). In 2009, the mean salary of a full-time year-round worker 25 years old and older in the United States with a bachelor's degree was \$68,812, 72.3% more than the \$39,937 earned by the typical full-time year-round worker with only a high school diploma or GED (U.S. Census Bureau, 2010a). In addition to increasing earning potential, a college degree also mediates the likelihood of unemployment (Hurley, 2009; Lumina Foundation for Education, 2010). In 2010, high school graduates age 25 and older were twice as likely to be unemployed than college graduates with at least a bachelor's degree (10.3% vs. 4.7% respectively) (Bureau of Labor Statistics, 2010a). To the extent that degree completion provides employment opportunities and higher earnings, college graduates are more

likely to enjoy employer-provided health and pension benefits (Baum, Ma, & Payea, 2010; Belfield & Levin, 2007). College graduates also experience a better quality of life than non-graduates, and pass this onto their offspring (Belfield & Levin, 2007). For example, children of college graduates tend to do better in school and are less likely to get into trouble with the law (Belfield & Levin, 2007). As the economic and social value of a high school diploma continues to decline, the costs associated with leaving college before degree completion are significant to individuals and their families.

Graduation, especially timely graduation, is also important to institutions of higher education. Institutional costs of student departure include the loss of revenue generated by tuition and fees, decreased funding at the federal and state levels because of fluctuations in enrollment, and lowered academic profiles in a heightened accountability atmosphere. For college and university administrators, enrollment management is essential to economic stability as funding for higher education has declined dramatically over the last twenty years. Even though higher education institutions have been able to generate revenue through entrepreneurial activities, university-industry partnerships, and research activities, colleges and universities have become increasingly dependent on tuition and fees to offset diminishing state support (Altbach et al., 2010; Desrochers, Lenihan, & Wellman, 2010). This has had a significant impact on the funding structure of public colleges and universities, which are subsidized by taxpayers through state funds. Between 1980 and 2008, the share of institutional budgets financed through state appropriations decreased 7 percent (31% to 24%) while the share coming from tuition and required student fees increased 13 percent (23% to 36%) (Desrochers et al., 2010). Although the problem of shrinking appropriations and increasing reliance on tuition and

fees are a problem for public institutions across the nation, they are more acute in New Jersey. Enrollment growth at the study institution is further burdening a state plagued by structural budget deficits, high debt payments, and an underfunded pension system. In response to declining state support over the last decade, the study institution has increased tuition and fees annually such that it is now one of the most expensive public institutions in the nation. In academic year 2010-11, the average cost of total tuition, fees, and room and board for a full-time, undergraduate student living on-campus was \$23,776 for in-state residents and \$35,532 for out-of-state residents. Raising retention and graduation rates while remaining committed to access and opportunity without adequate financial means to support low-income and minority students is a growing concern among university administrators at the study institution. As cost of attendance continues to rise, the availability of financial aid becomes increasingly more important in student enrollment and persistence decisions. Without adequate financial support, students are more prone to interrupted enrollment patterns (i.e., stopouts) or leaving the institution entirely, which not only negatively impacts the institution's operating budget through the loss of tuition and fees but also lowers the institution's retention and graduation rates. Furthermore, the timely graduation of current students by keeping them enrolled continuously is essential to accommodating enrollment growth without burdening the institution's capacity to receive and serve more students with less fiscal resources.

At the same time that funding for higher education has diminished, state and federal lawmakers, accrediting agencies, and governing bodies have demanded more accountability from college and university administrators. Although retention and graduation rates have long been used as performance measures for evaluating the



institutional quality of colleges and universities in college guidebooks and national rankings such as those published by *U.S. News & World Report*, they are increasingly dominating higher education policy debates in the United States as national interest in performance-based accountability grows (Alexander, 2000; Cook & Pullaro, 2010; Burke, 2005). The public call for more accountability and better performance in higher education has increased the pressure on states to gain greater control over higher education resources, and as a result many states have introduced various types of postsecondary accountability measures based on student outcomes (Alexander, 2000; Burke, 2005). Some states have gone so far as to link state appropriations to performance on outcome measures as an incentive for their postsecondary institutions to not just enroll students, but also ensure that students complete their degree programs (Alexander, 2000; Burke, 2005). For example, policymakers in Alabama tied state funding not only to graduating more students, but graduating more students on time. For public institutions that rely heavily on state appropriations, the loss of dwindling educational resources could lead to economic instability, especially in low-performing institutions, if the basis for state funding shifts from enrollments to completions.

Reflecting the public call for more accountability in higher education, Congress passed the Student Right-to-Know and Campus Security Act in 1990 (Public Law No: 101-542) as an amendment to the Higher Education Act of 1965. As part of the Student Right-to-Know legislation, colleges and universities who receive Title IV student aid are required to disclose their institution's graduation rates for full-time, degree-seeking, undergraduate students. Initially passed to protect the educational interests of students and their families in making decisions regarding colleges and universities, educators and

policymakers have turned to these institutional graduation rates as a measure of accountability. One of the only comparable and widely recognized student outcome measures in higher education accountability, graduation rates have received a lot of scrutiny and criticism by researchers, university and higher education leaders, and state and federal policymakers (American Association of State Colleges and Universities, 2006). Reauthorizing the Higher Education Act of 1965, the Higher Education Opportunity Act of 2008 (Public Law No: 101-542) changed the disclosure and reporting requirements of graduation data under the Student Right-to-Know provision of the law. Colleges and universities receiving Title IV student aid are now required to disaggregate the graduation rate of degree-seeking, full-time, undergraduate students by gender, by racial and ethnic subgroup, by receipt of a Pell Grant, by receipt of a subsidized federal loan but not a Pell Grant, and by receipt of neither a subsidized federal loan nor a Pell Grant. By disaggregating the data, institutional graduation rates can be viewed within the context of student demographics. Although the new disclosure and reporting requirements of graduation data are still oriented toward the normal progression of traditional students from an institutional perspective, the changes are part of an attempt by federal policymakers to improve the utility and transparency of this student outcome measure for accountability purposes.

College and university administrators are increasingly more concerned about retention and graduation rates given the focus on them by regulatory agencies, their use in rankings by college guide publications, and the scrutiny given to these rates by the general public. Recently, the president of the study institution announced the university's ambition to move to the highest tier of America's public research universities. As the

university's first-year retention and graduation rates consistently lag behind aspirant peers, central administration has identified student retention and graduation as an area for improvement. The need to improve retention and graduation rates at the study institution will become increasingly important to central administration in the coming years as the university's governing board recently announced plans to start using retention and graduation rates among several other quantitative outcome measures in their annual review of the president's job performance. More structured and systematic, the revised system is viewed by the governing board as necessary to close the gap between the study institution and its aspirant peers. More vulnerable to state and federal aid cuts than its aspirant peers, and lacking the fiscal resources to invest into student retention and graduation efforts without raising tuition and fees, improving the university's retention and graduation rates to that of its aspirant peers will be a substantial challenge while maintaining diversity at the study institution without significant investment in higher education at the state and federal level. Recognizing this financial constraint, central administration at the study institution recently decided to change the All Funds Budgeting policy to increase the importance of student retention and graduation among its academic units. Starting next year, the academic units will receive graduated payments in their All Funds Budget allocations for four years instead of level payments based solely on enrollment numbers. That is, central administration will reward academic units with larger per student allocations each consecutive year a student is retained for up to four years. After four years of continuous enrollment, the payment will plateau with the assumption that students enrolling in fifth and sixth years are on course for graduating within 150 percent of the normal degree program time. Given the individual, societal, and

institutional costs associated with student departure before degree completion, it is important to develop a better understanding of the longitudinal process of student retention and degree completion, especially at the institutional or campus level.

With a body of literature spanning over 75 years, student retention has been extensively studied in higher education (Braxton, 2000). Empirical research on student retention and persistence has primarily focused on three categories of variables: demographic variables related to student and their families, precollege academic preparation variables, and college variables. Demographic variables related to students and their families have included gender, ethnicity/race, age, residency, distance from home, parents' level of education, and family income/socioeconomic status. Demographic variables generally account for only a small percentage of explained variance associated with student departure before degree completion (Kennedy & Sheckley, 1999), and their impact on student retention and persistence has not been consistent. Precollege academic preparation variables have included high school grade point average, class rank, standardized test scores, and intensity of the high school academic curriculum. According to the body of research, precollege academic preparation variables can be useful predictors of student retention and persistence, but do not explain all of the variation in student departure (Murtaugh, Burns, & Schuster, 1999). College variables have included semester and cumulative grade point averages, academic major, enrollment patterns, course load, the accumulation of credits, and financial aid. Although useful predictors of student retention, the use of cross-sectional data and traditional static methodologies to study a longitudinal process has been particularly

problematic for ascertaining the effect of these variables on student retention and persistence to degree completion (DesJardins, 2003; Willett & Singer, 1991).

Some researchers have also attempted to incorporate measures of academic and social integration, or institutional fit, as defined in Tinto's (1975) and Bean's (1980, 1982, 1983) models of student departure. Although general support has been found for the relationship between the student and the institution, particularly in the testing of these student departure models, the extent to which the practical application of that research is possible for college and university administrators is extremely limited (Kennedy & Sheckley, 1999). Extensive surveys or data collection techniques are required to collect information regarding the psychological constructs espoused in these models. Although the research to date provides insight into the student retention issue in higher education, its practical application at the institutional level is nearly impossible as these variables refer to realities that lie beyond the control of those who can best steer students toward degree completion (Adelman, 1999). With increasingly limited resources, and to avoid methodological limitations associated with low survey response rates, most institutional researchers conduct single-institution studies using information that is collected for the majority, if not all, of their students.

Some of the discrepancies in the literature can be attributed to the various methodological approaches employed by researchers studying student retention. Although researchers have been successful in identifying the many factors associated with student departure before graduation, most of these studies have ignored the temporal nature of student retention and persistence to degree completion (DesJardins, Ahlburg, & McCall, 1999; Willett & Singer, 1991). Studying a temporal process like college student

retention and graduation with cross-sectional data and traditional statistical designs is very problematic (DesJardins, 2003; Willett & Singer, 1993). Willett and Singer (1993) identified five problems with the traditional approaches to studying event occurrence in longitudinal studies. First, the outcome is inextricably linked to the particular time frame chosen for data collection and analysis, which is rarely substantively motivated. Second, contradictory conclusions can result from variation in the particular time frames studied. Third, traditional analytic methods offer no systematic mechanism for dealing with censored observations, or those individuals who do not experience the event of interest during the time frame of the study. Fourth, observed differences in rates of event occurrence may be attributable to nothing more than research design. As risk periods vary across people and time, individuals followed for longer periods of time have a greater likelihood of experiencing the target event than those followed for shorter periods of time. Finally, few mechanisms exist for the inclusion of time-varying predictors, or variables whose values vary from one time period to another. DesJardins (2003) also noted several problems with using cross-sectional designs to study longitudinal processes. First, the assumption of statistical equilibrium is violated. Student retention to degree completion is not a time-invariant process, and neither are the variables that explain the process time-invariant. Second, establishing the direction of causality using a “snapshot” of time is almost impossible, and doing so with traditional statistical methods can result in a misleading picture of the longitudinal process being studied. Third, cross-sectional designs make it difficult to control non-random processes, such as self-selection and time-related selectivity issues, because information is available for only one point in time. Fourth, cross-sectional designs cannot distinguish age and cohort effects. Fifth,

cross-sectional data and traditional static analytic methods often make it difficult to untangle reciprocal effects that take place over time. For these reasons, Willett and Singer (1993) and DesJardins (2003) argue event history analysis (also known as survival analysis) is a more appropriate statistical technique for studying the longitudinal processes that take place within higher education than traditional static methods.

Event history analysis is a regression-like technique originating out of biostatistics that allows researchers to answer research questions about the occurrence and timing of events (Allison, 1982; DesJardins, 2003; Singer & Willett, 1991). For example, event history analysis has long been used to study survival and relapse rates for diseases such as cancer. The extension of event history analysis to educational research accompanied by new developments in statistical computing has allowed researchers studying student retention to reframe the question from whether students leave a particular institution to when are students most at risk of leaving the institution or higher education in general (Willett & Singer, 1991). There is a small but growing body of published literature using event history analysis to study college student retention (Chen & DesJardins, 2008; DesJardins et al., 1999; DesJardins, Ahlburg, & McCall, 2002a; DesJardins, Ahlburg, & McCall, 2002b; DesJardins, Ahlburg, & McCall, 2006; DesJardins, McCall, Ahlburg, & Moye, 2002; Ishitani, 2003; Ishitani, 2006; Ishitani & DesJardins, 2002; Murtaugh et al., 1999). Although the time to first departure was the main criterion of interest in earlier studies using event history analysis, researchers were still plagued with the difficulty of operationalizing a consistent definition of attrition while the number of college students with non-traditional college trajectories continued to increase (e.g., multiple stopouts, enrolling in two or more postsecondary institutions). According to Summerskill (1962),

college student attrition has been defined from the higher education system, institutional, and departmental perspectives, and what constitutes student departure varies among each of these constituencies. Without a consistent definition for the criterion variable, no solid body of evidence could be established. Calling for a whole new way of thinking about student retention, Adelman (1999) argued that degree completion should be the criterion of interest because:

Degree completion is the true bottom line for college administrators, state legislators, parents, and most importantly students – not retention to the second year, not persistence without a degree, but completion. (p. v)

Following Adelman's (1999) suggestion, more recent studies have focused on graduation as the outcome of interest and have increased the time frame of data collection and analysis to account for increasing time-to-degree rates (e.g., Chen & DesJardins, 2008; DesJardins et al., 2002a; DesJardins et al., 2006; Ishitani, 2003; Ishitani, 2006).

Although the body of literature using event history analysis is limited, it is clear that understanding the temporal dimensions of graduation and the factors that affect whether students succeed or fail is becoming increasingly important, particularly at the study institution.

The purpose of the present study is to use event history analysis to understand the temporal dimensions of graduation and the factors that affect whether students succeed or fail, particularly at the study institution. As previously mentioned, central administration at the study institution is increasingly concerned about the university's student retention and graduation rates as these measures consistently lag behind aspirant peers. Furthermore, the university's governing boards will now use these measures during the president's annual performance evaluation.



## CHAPTER II. LITERATURE REVIEW

An important issue facing American higher education today, the study of college student retention has been part of the higher education literature for over 75 years (Braxton, 2000). In one of the earliest reviews of the literature, Summerskill (1962) pointed out that 40 years of research had failed to adequately explain why approximately half the students attending American colleges and universities left before obtaining an undergraduate degree. Pantages and Creedon (1978) reached a similar conclusion after reviewing 25 years of research conducted between 1950 and 1975. Although much of the early research on student departure was limited to monitoring enrollments on individual college campuses to ensure institutional survival, rapid enrollment growths during the 1950s and 1960s increased the importance of student retention as college and university administrators looked to improve the economic efficiency of their institutions (Berger & Lyon, 2005; Summerskill, 1962). As institutions grew in size and complexity, the need for more systematic research on student retention became imperative as college and university administrators needed to manage their enrollment numbers to stabilize operational budgets driven primarily by student tuition and fees (or state appropriations on a per student basis). Increasing pressure on colleges and universities to retain and graduate their students by federal and state policymakers has increased the need for empirically based research on student retention. As funding for higher education diminishes and the call for more accountability increases, these administrative and economic concerns continue to be driving factors for higher education stakeholders to commit a great deal of time, energy, and resources to improve institutional retention and graduation rates.

### **College Student Retention and Theoretical Models of Student Departure**

Over the years, a large body of research on college student retention and persistence has emerged, and much of this research has focused on the development and testing of two theoretical models of student departure: Tinto's Student Integration Model (1975, 1982, 1987, 1988, 1993) and Bean's Student Attrition Model (1980, 1982, 1983, 1985). Applying sociological and organizational theories to the study of student departure, these models attempt to clarify the processes linking student-related factors with institutional ones. While Tinto and Bean remain the early pioneers in student departure models, Braxton, Hirschy, and McClendon (2004) advance a revised theory of student departure that places greater emphasis on the influence of social integration, especially as it pertains to student departure from residential colleges and universities.

#### **Spady's Sociological Model of Dropout**

Recognizing the need to move the field beyond periodic reviews of the empirical research, Spady (1970) proposed the first conceptual model of the college dropout process. In his model, Spady (1970) postulated that the process by which a student leaves a particular college or university, or the higher education system entirely, parallels the social process of Durkheim's (1951) theory of suicide. According to Durkheim (1951), an individual's desire to break ties with a social system stems from a lack of social integration between the individual and the larger society. While acknowledging that dropping out of college is less drastic than committing suicide, Spady (1970) suggested that social integration, or the lack thereof, could be a useful concept for explaining a student's decision to leave college. Spady (1970) expands Durkheim's social integration process to include both the academic and social systems of the college or university.

In Spady's (1970) model, students enter college with goals, attitudes, and personality dispositions shaped by their family background and high school experiences. It is the interaction between the student and the institution that impacts the student's ability to assimilate successfully into the academic and social systems of the college environment. Successful integration is influenced by at least two factors in each system. Success within the academic system is achieved through grades and intellectual development. Grades are the extrinsic reward of the academic system, whereas intellectual development represents an intrinsic reward. Within the social system, the student must experience normative congruence and friendship support for successful social integration. Normative congruence is the compatibility between the student's attitudes, interests, and personality attributes and the norms of the college environment (Spady, 1970). Friendship support is the establishment of close relationships with others in the system (Spady, 1970). These four independent variables contribute directly to social integration, which indirectly influences the dropout decision through two intervening variables (Spady, 1970). Satisfaction is the first intervening variable, and represents the student's satisfaction with his or her college experience. The second intervening variable is institutional commitment, which represents the student's commitment to the social system (i.e., the college or university). Spady (1970) argues that a student's level of satisfaction and degree of institutional commitment emerge from the social integration process. A student's grades are also assumed to have a direct effect on the dropout decision, as poor academic performance could lead to an academic dismissal from the institution.

Putting his model to the test, Spady (1971) used a sample of 683 freshmen who enrolled at the University of Chicago in fall 1965 to examine the effect of social integration and related sociological influences on college attrition. Spady (1971) combined data on the students' family background and high school experiences, their perceptions of environmental and social influences, and institutional GPA and retention data. Based on the empirical results, Spady (1971) revised his model by adding a new variable called "structural relations." Friendship support became a subset of structural relations because it was found to be "directly dependent on elements in both the family background and normative congruence clusters" (Spady, 1971, p. 58). Finding several gender differences, Spady (1971) also modified some of the directional arrows and structural paths connecting variables in his model, particularly in relation to females. For example, males based their dropout decision primarily on their grade performance. Focused on meeting the formal performance standards set by the faculty, males were more willing to tolerate the environmental conditions imposed on them. In contrast, females initially based their dropout decision primarily on institutional commitment and social integration than academic performance. That is, interpersonal needs appeared to dominate the decision-making process for females initially. However, as achievement and persistence became more synonymous over time, formal academic performance became the dominating factor accounting for student attrition among both sexes.

Although criticized for its descriptive nature, Spady's (1970, 1971) sociological model is the first known model of student departure. His attempt to explain the process of student departure as an interaction between the student and the college environment provided the basis for more advanced theoretical models.

### **Tinto's Student Integration Model**

Building on Spady's (1970, 1971) work and Durkheim's (1951) Theory of Suicide, Tinto (1975, 1987, 1993) formulated and advanced a theory of student departure explaining the process that motivates individuals to leave college before graduating. Tinto's (1975) theory asserts that student departure from a particular higher education institution, or in some cases from higher education completely, is a longitudinal process that results from insufficient integration into the academic and social systems of the college or university.

According to Tinto (1975, 1987, 1993), students enter a college or university with a variety of personal and pre-enrollment characteristics (e.g., individual attributes, precollege experiences, and family background), each of which have direct and indirect effects on goal commitment (i.e., degree completion) and institutional commitment. Once enrolled, these commitments are continuously modified and reformulated through a longitudinal series of interactions between the individual and the structures and members of the academic and social systems of the institution. The greater the institution's ability to integrate the student into the formal and informal academic and social systems of the college or university, the stronger the student's commitment to graduating, and more specifically graduating from that particular institution, will be. Viewing these commitments as important predictors of and reflections of the student's experiences in the college environment, Tinto (1975) asserts that it is the interaction between the student's commitment to the goal of completing college and his or her commitment to the specific educational institutional that ultimately determines whether or not the student drops out.

In proposing an institutional model of dropout behavior, Tinto (1975) reconceptualized college student retention as a “longitudinal process of interactions that lead differing persons to varying forms of persistence and/or dropout behavior” (p. 93). As a result of this reconceptualization, and at Tinto’s recommendation, researchers abandoned cross-sectional methods in favor of longitudinal studies when studying college student behavior. Although Tinto’s Student Integration Model has served as a conceptual framework for numerous studies, contradictory findings on the impact of precollege, goal and institutional commitments, and integration factors on college persistence have been attributed to the type of institution (e.g., two-year vs. four-year institutions, public vs. private colleges), to inconsistencies in the measurement of the constructs across studies, particularly academic integration, and the lack of control for variables external to the institution (Braxton & Lee, 2005; Braxton & Lien, 2000; Braxton, Sullivan, & Johnson, 1997; Cabrera, Castaneda, Nora, & Hengstler, 1992; Cabrera, Nora, & Castaneda, 1993).

### **Bean’s Student Attrition Model**

Based on Price’s (1977) process model of organizational turnover and models of attitude-behavior interactions, Bean (1980) proposed the Student Attrition Model as an alternative model of student departure to explain the college persistence process. Bean (1980) argues that students leave institutions of higher education for reasons similar to those that cause employees to leave work organizations. Viewing student departure as analogous to employee turnover in work organizations, Bean’s (1980) theoretical framework stresses the importance of a student’s behavioral intentions as important predictors of his or her persistence behavior. According to Bean (1980), a student’s behavioral intentions are shaped by a process in which beliefs shape attitudes and

attitudes, in turn, shape behavioral intentions. A student's experiences with the various components of a particular institution are presumed to influence his or her beliefs. Bean's (1980) model also recognizes the role factors external to the institution can have on a student's attitudes and persistence decisions.

Seeking to model student attrition, defined as the cessation of individual student membership in an institution of higher education, Bean's (1980) original model included three sets of independent variables: background variables, organizational determinants, and intervening variables. Background variables (e.g., past academic achievement, socioeconomic status, state residence, distance from home, and hometown size) contribute to organizational determinants. Organizational determinants (e.g., routinization, student development, practical value, institutional quality, student integration, university GPA, goal commitment, communication, distributive justice, centralization, student's advisor, staff/faculty relationship, campus job, area of major, certainty of major, housing, campus organizations, and opportunities for alternative roles) influence the level of student satisfaction with the institution, which in turn influences the level of institutional commitment. Institutional commitment has a direct impact on the student's decision to drop out, such that the higher the student's level of institutional commitment, the less likely he or she will drop out of college. Following this theoretical framework, Bean's (1980) empirical study produced two final models of student attrition, one for female students and one for male students. In both models, institutional commitment was the most important indicator of dropping out. Other important variables included academic performance, routinization, student development, and college GPA. There were two major differences between the models: (1) the female model had 13

variables that significantly influenced dropout, whereas the male dropout model had only 7; and (2) student satisfaction was a significant intervening variable for female students, but not for male students.

Eliminating the background variables from the analysis, Bean (1982) reduced his causal model of student attrition to 10 independent variables: intent to leave, practical value, certainty of choice, loyalty, grades, courses, educational goals, major and job certainty, opportunity, and family approval. In this model, Bean (1982) found that the intent to leave accounted for the largest proportion of explained variance in dropout. Adjusting his original model of student attrition to reflect this new finding and the revised model of worker turnover developed by Price and Mueller (1981), Bean (1983) proposed the industrial model of student attrition. Bean's (1983) industrial model of student attrition differed from past models in four ways: (1) background variables were excluded because they did not appear in the Price and Mueller model; (2) the specification of intent (not institutional commitment) as the immediate precursor of attrition; (3) a clearly specified one-way causal ordering of the variables; and (4) the identification of specific student organizational interactions as determinants of satisfaction.

Bean (1985) also developed a conceptual model of student attrition that combined components of the student integration and student attrition models. In this model of dropout, student background variables and environmental variables are emphasized as having a direct influence on academic and social integration. Bean (1985) proposed that students' intentions were shaped by their belief and attitudes related to the institution, faculty, and friends. Similar to Tinto's student integration model, the conceptual model argues that positive academic and social experiences result in positive beliefs and



attitudes, and these positive beliefs and attitudes result in the intention to persist in college. Negative academic and social experiences result in negative beliefs and attitudes, and these negative beliefs and attitudes result in the intention to leave college which causes the act of dropping out. By focusing less on social integration and more on environmental variables, Bean's (1985) conceptual model offers a different perspective on college persistence than proposed in Tinto's (1975) Student Integration Model.

### **Converging the Student Integration Model and Student Attrition Model**

Although the two theoretical frameworks offer different perspectives on what variables have the strongest effects on college persistence to degree completion, Hossler (1984) pointed out that the Tinto's Student Integration Model and Bean's Student Attrition Model share several commonalities. Comparing Tinto's (1975) and Bean's (1980) models, Cabrera et al. (1992) noted that the two models: (1) regard persistence as a complex set of interactions that occur over time within a given institution, (2) acknowledge the importance of pre-college characteristics, and (3) argue that persistence is influenced by the level of fit between the student and the institution. Exploring the possibility of merging the two theoretical models, Cabrera et al. (1992) simultaneously tested the predictive validity of the two models and preliminary results suggested that a more comprehensive understanding of student departure could be achieved if the two models were combined. This finding was later confirmed in a follow-up study by Cabrera, Nora, and Castaneda (1993), in which the researchers tested an integrated model of student attrition.

Nora and Cabrera (1996) developed a Student Adjustment Model that proposes the experiences of college students are represented in two domains: a social domain,

involving experiences with fellow students, and an academic domain, involving experiences with faculty and academic staff at the institution. According to Nora and Cabrera (1996), these collective experiences enhance the affective and cognitive development of the student, which leads to academic and intellectual development and greater commitment to both the institution and attaining a college degree. Based on the theoretical frameworks of Tinto (1987) and Bean and Metzner (1985), the Student Adjustment Model presupposes that academic experiences and social integration are not independent, as “positive experiences in one domain are seen as conducive of positive experiences in the other domain” (p. 123).

In addition to effectively combining the Student Integration Model and the Student Attrition Model, Nora and Cabrera’s (1996) Student Adjustment Model addresses four assertions made regarding the nature of factors involved in the persistence process of both minority and nontraditional students. First, academic preparedness does not exert a stronger effect among minority students than among White students. Second, attachments with family, friends, and past communities play a key role to the successful transition of students to college. Third, the perceptions of prejudice and discrimination do not have the overwhelming effect on the college persistence process among minority students as presumed in the literature. Finally, the study found that existing conceptual models of college persistence are useful in explaining the persistence of minority students.

### **Braxton’s Theory of Student Departure from Residential Colleges and Universities**

Recognizing the near “paradigmatic status” of Tinto’s theory among scholars studying college student departure, Braxton et al. (1997) reviewed the empirical research

and found strong support for only five of the thirteen propositions ( 1, 9, 10, 11, and 13) of Tinto's theoretical framework in residential colleges and universities. The thirteen propositions of Tinto's theory are as follows:

1. Student entry characteristics affect the level of initial commitment to the institution.
2. Student entry characteristics affect the level of initial commitment to the goal of graduation from college.
3. Student entry characteristics directly affect the student's likelihood of persistence in college.
4. Initial commitment to the goal of graduation from college affects the level of academic integration.
5. Initial commitment to the goal of graduation from college affects the level of social integration.
6. Initial commitment to the institution affects the level of social integration.
7. Initial commitment to the institution affects the level of academic integration.
8. The greater the level of academic integration, the greater the level of subsequent commitment to the goal of graduation from college.
9. The greater the level of social integration, the greater the level of subsequent commitment to the institution.
10. The initial level of institutional commitment affects the subsequent level of institutional commitment.
11. The initial level of commitment to the goal of graduation from college affects the subsequent level of commitment to the goal of college graduation.
12. The greater the level of subsequent commitment to the goal of college graduation, the greater the likelihood of student persistence in college.
13. The greater the level of subsequent commitment to the institution, the greater the likelihood of student persistence in college.

(Braxton et al., 1997, p. 112). Based on the lack of empirical support for many of the theoretical propositions, especially those assertions pertaining to the influence of academic integration on subsequent institutional commitment and persistence, Braxton and Lien (2000) called for a substantial revision of Tinto's theory. Braxton et al. (2004) developed revised models of student departure from residential and commuter institutions. As the study institution for this dissertation is primarily residential, this discussion will focus on the revised theory for student departure from residential colleges

and universities (see Braxton et al., 2004; Braxton & Hirschy, 2005 for their discussion of the theory of student departure from commuter colleges and universities).

Using four of the five strongly supported propositions for Tinto's theory and six factors that have a statistically significant influence on social integration, the revised model for student departure advanced by Braxton et al. (2004) argue that student entry characteristics have a direct impact on a student's decision to leave a postsecondary institution before degree completion (or the higher education system entirely) and an indirect impact on persistence through the initial goal and institutional commitments. Entry characteristics include the student's gender, racial or ethnic background, socioeconomic status, academic ability, high school academic preparation, parental education, and ability to pay for college (Braxton & Hirschy, 2005). The greater a student's initial goal and institutional commitments, the greater a student's social integration and subsequent institutional commitment and persistence (Braxton & Hirschy, 2005; Braxton et al., 2004).

The student's initial institutional commitment influences his or her perception of institutional commitment to the welfare of students, or the enduring concern for student growth and development (Braxton et al., 2004). The student's initial institutional commitment also influences the student's perception of institutional integrity, or the degree of congruency between the actions of faculty, administrators, and staff members of a college or university community and the institution's stated mission (Braxton et al., 2004). It also influences his or her perception of communal potential, or the extent to which a student perceives that a subgroup of students exists within the college community which share similar values, beliefs, and goals. Braxton et al. (2004) argue that

greater initial commitment to the institution leads to more favorable perceptions of these three institutional characteristics, which leads to greater levels of social integration and persistence. In addition to influencing perceptions of institutional commitment to the welfare of students, institutional integrity, and communal potential, the student's initial commitment to the institution also affects a student's proactive social adjustment and psychosocial engagement. Proactive social adjustment refers to a student's propensity to positively adjust to the demands and pressures of social interaction in a college community (Braxton et al., 2004). Psychosocial engagement represents the psychological energy a student invests in interactions with college peers and involvement in extracurricular college activities. Greater initial commitment to the institution leads to greater use of proactive social adjustment and psychosocial engagement, which leads to greater levels of social integration (Braxton & Hirschy, 2005; Braxton et al., 2005).

According to Braxton and colleagues (2004, 2005), social integration has a direct effect on student persistence in college, such that the greater the social integration of a student the more likely he or she will persist. Student integration is also hypothesized to have an indirect effect on student persistence by affecting the level of subsequent institutional commitment. Braxton and Hirschy (2005) argue that higher levels of social integration lead to greater subsequent institutional commitment, and the greater the level of subsequent institutional commitment the more likely the student will persist in college.

### **The Study of College Student Departure and Persistence**

As college and university administrators seek to increase student retention and persistence to degree completion and scholars seek explanations to the student departure puzzle, the need to understand this phenomenon has generated a plethora of research in

the higher education literature (Bean, 1980; Bean, 1982; Bean, 1983; Bean, 1985; Braxton et al., 1997; Braxton et al., 2004; Cabrera et al., 1992; Cabrera et al., 1993; Nora & Cabrera, 1996; Pascarella & Terenzini, 1991; Pascarella & Terenzini, 2005; Tinto, 1975; Tinto, 1987; Tinto, 1993). Empirical research on student departure and persistence in higher education has primarily focused on three categories of variables: (1) demographic variables related to students and their families; (2) precollege academic preparation variables; and (3) college variables. Demographic variables related to students and their families have included gender, ethnicity, age, residency, distance from home, parents' level of education, and family income/socioeconomic status. Demographic variables generally account for only a small percentage of explained variance associated with student attrition (Kennedy & Sheckley, 1999), and their impact on student retention and persistence has not been consistent. Precollege academic preparation variables have included high school GPA, class rank, standardized test scores, and intensity of the high school academic curriculum. According to the body of research, precollege academic preparation variables can be useful predictors of student retention but do not explain all of the variation in college student attrition rates (Murtaugh et al., 1999; Pantages & Creedon, 1978). College variables have included semester and cumulative GPA, academic major, enrollment patterns, course load, the accumulation of credits, and financial aid. Although useful predictors of student retention, the use of cross-sectional methodologies to study a longitudinal process has been particularly problematic for ascertaining the effect of these variables on student retention (Willett & Singer, 1991).

Some researchers have also attempted to incorporate measures of academic and social integration, or institutional fit, as defined in Tinto's (1975) and Bean's (1980, 1982, 1983) models of student departure. Although general support has been found for the relationship between the student and the institution, particularly in the testing of these student departure models, the extent to which the practical application of that research is possible for college and university administrators is extremely limited (Kennedy & Sheckley, 1999). Extensive surveys or data collection techniques are required to collect information regarding the psychological constructs espoused in these models. Although the research to date provides insight into the retention issue in higher education, its practical application at the institutional level is nearly impossible (Adelman, 1999).

Commenting on this issue, Adelman (1999) wrote:

Both research traditions place an extraordinary emphasis on psychological variables: intentions, attitudes, influences, commitment, perceptions ... These variables unfortunately refer to realities that lie beyond the control of those who can best steer students toward degree completion. (p. 27)

With increasingly limited resources, and to avoid methodological limitations associated with low survey response rates, most institutional researchers conduct single-institution studies using information that is collected for the majority, if not all, of their students.

Numerous studies that have focused on testing the academic and social integration, or institutional fit, components of Tinto's (1975) and Bean's (1980, 1982, 1983) models of student departure have reported moderate to significant effects of various forms and measures of academic and social integration on year-to-year persistence (Pascarella & Terenzini, 2005). Research suggests that some degree of integration in the collegiate setting is necessary for persistence, as students with lower levels of academic and social integration tend to drop out. However, it is not clear

whether the effects on persistence are direct or indirect and what the exact relationship between academic and social integration is (Pascarella & Terenzini, 2005). Also, the effects of pre-college characteristics on persistence have been inconsistent across empirical studies (Pascarella & Terenzini, 2005).

Some of the discrepancies in the literature can be attributed to the various methodological approaches employed by researchers studying student attrition. Gekowski and Schwartz (1961) pointed out that many of the early studies on student attrition utilized cross-sectional methods that focused on the characteristics of either persisting students or those that dropped out, and drew conclusions without the proper use of a comparison group from the other category. Gekowski and Schwartz (1961) also criticized the univariate or bivariate nature of early studies of student attrition as limiting. Assuming that multiple factors operate concurrently to produce student attrition, Gekowski and Schwartz (1961) argued for a more multivariate approach to studying student attrition. Jex and Merrill (1962), Eckland (1964) and Marks (1967) criticized the heavy reliance on ex-post facto methodology. Advocating the use of longitudinal studies over cross-sectional studies, Jex and Merrill (1962) argued that the longitudinal approach provides a clearer view on the complex interaction of factors on student attrition as they occur over time. Furthermore, researchers studying student attrition at a particular institution with longitudinal methods would be able to distinguish between stopouts and dropouts; a major limitation of cross-sectional methods (Jex & Merrill, 1962). Pantages and Creedon (1978) argued that the traditional two-way analysis distinguishing between dropouts and persisters, combined with ex-post facto methodology in attrition research, obscured many important details of student withdrawal and inflated estimates of attrition



rates. Pantages and Creedon (1978) also warn researchers against combining attrition data derived from different institutions given the enormous variation in attrition rates of individual institutions.

### **Methodological Criticisms in the Study of College Student Retention**

#### **Operational Definitions**

Difficulties in operationalizing a consistent definition of attrition and dropout has resulted in contradictory findings on the factors associated with college student departure. Without a consistent definition for the criterion variable, no solid body of evidence could be established. In one of the earliest reviews of the literature, Summerskill (1962) pointed out that 40 years of research on student attrition had seen considerable variability in the definition of “attrition.” Attrition had been defined as students lost to a particular academic unit within a college, lost to the college or university as a whole, or lost to the higher education system altogether. Summerskill (1962) also noted that the criterion by which a student was classified as a dropout or nondropout had also been particularly problematic for student retention researchers. Nondropouts had been defined as those students who graduated in four years, those students who graduated or are still enrolled after four years, or those students who eventually graduated from college. Furthermore, distinctions such as “involuntary vs. voluntary withdrawal” have also added to the complexity of operationally defining the criterion (Pantages & Creedon, 1978; Summerskill, 1962). Inconsistencies in defining these terms and differentiating these groups, both theoretically and operationally, not only affect the research findings but also affect the usefulness of the study to other researchers and educators (Pantages & Creedon, 1978).

Reviewing student attrition research between 1950 and 1975, Pantages and Creedon (1978) noted that very different definitions of the criterion variable were employed despite the increase in longitudinal studies of student attrition. Researchers studying student attrition have defined dropout using a higher education systems approach, an institutional perspective, or a departmental perspective (Summerskill, 1962). According to Tinto (1982), it is important to consider how the definition of dropout may vary among different constituencies concerned with the character of dropout from higher education. Academic and institutional dropout have typically been of interest to institutional researchers and stakeholders concerned with local attrition rates, whereas dropout from the higher education system as a whole has been the focus of researchers interested in policies at the state and national levels (Tinto, 1982). Although most institutional studies defined “dropout” as the loss of a student from a particular college or university, the definition of dropout remains specific to the perspective of the researcher (Pantages and Creedon, 1978). The validity of summarizing and combining the findings across studies is questionable if inconsistent definitions have been used because the studies are measuring different phenomena (Panos & Astin, 1968; Pantages & Creedon, 1978). Questionable findings also resulted from the cross-sectional nature of many of the studies, which failed to distinguish between stopouts and dropouts. These studies typically defined “dropout” as any student who was previously enrolled at the institution but not enrolled at the time of the study (Pantages & Creedon, 1978). Thus in a cross-sectional design, a student exhibiting an interrupted enrollment behavior (i.e., stopout) could be classified as a dropout or a persister depending on the chosen time frame of the study. The inability to differentiate between permanent and temporary dropouts is a

major limitation of using cross-sectional designs to study a longitudinal process (Pantages & Creedon, 1978).

In addition to distinguishing between dropouts and nondropouts, some researchers advocated for the distinction between the varying forms of dropout, particularly voluntary and involuntary withdrawal (Johansson & Rossmann, 1973; Rossmann & Kirk, 1970; Starr, Betz, & Menne, 1972; Tinto, 1975). Although these researchers advocated for separate methodological approaches for studying voluntary and involuntary departure, Pantages and Creedon (1978) argued that such efforts be abandoned because: (1) it is less confusing in the long run to regard academic success or failure as well as academic dismissal as intervening variables that lead to withdrawal rather than as part of the dependent variable; (2) the distinction implies that a “voluntary” dropout, such as withdrawal with good grades, is any less determined by social forces than is a “nonvoluntary” dropout, such as withdrawal which follows academic dismissal; and (3) labeling withdrawals following dismissal as “nonvoluntary” distracts researchers and educators from a crucial issue: why students who qualified for admissions to college get poor grades in the first place (p. 92).

In studying the persistence and bachelor’s degree completion of students enrolled in four-year institutions, Adelman (1999) argued that graduation, not persistence, should be the criterion of interest to researchers studying student attrition from college.

According to Adelman (1999),

Degree completion is the true bottom line for college administrators, state legislators, parents, and most importantly students – not retention to the second year, not persistence without a degree, but completion. (p. v)

The shift in the conceptualization of the criterion variable to degree completion eliminates much of the theoretical and operational issues in defining dropout.

### **Methodological Limitations**

Researchers studying college student retention and graduation have traditionally used logistic regression or structural equation modeling to assess the associations various factors have on college enrollment behavior. Cross-sectional techniques such as these involve defining a particular cohort of interest, choosing a time period of study, and then comparing the enrollment status of those who have persisted or graduated with those who have not (Gekowski & Schwartz, 1961). Prospective two-wave studies that compare the enrollment status of a cohort of students at time 1 and time 2 have generally employed arbitrary points of time such as a single semester or academic year (Ishitani & DesJardins, 2002; Willett & Singer, 1991). One of the major limitations of employing cross-sectional designs is that the statistical results are inextricably linked to the particular time frame chosen for the study (Willett & Singer, 1993). As previously mentioned, the inability of cross-sectional methods to differentiate between stopouts and dropouts can result in students who experience interrupted enrollment as being identified as a dropout in one semester and a persister in another. In their review of 25 years of college student attrition research, Pantages and Creedon (1978) indicated that the most meaningful research on student departure was provided by longitudinal studies covering a period of more than four years. Although researchers have been successful in identifying the many factors associated with student departure before graduation, most of these studies have ignored the temporal dimension of student departure, particularly the timing of dropout (DesJardins, 2003; Ishitani & DesJardins, 2002; Willett & Singer, 1991).

According to Tinto (1988), college student retention is not time invariant, especially across different groups. Empirical specifications based on a “snapshot” of time can result in a misleading picture of the longitudinal process being studied when the assumption of statistical equilibrium is violated (DesJardins, 2003). In fact, any inferences made using cross-sectional data and designs will likely be ambiguous because static statistical techniques such as regression and structural equation modeling only explain the net differences in the effects of explanatory variables, not how relationships change over time (DesJardins, 2003, p. 425).

Arguing that the timing of withdrawal was critical for understanding the student attrition phenomenon, Barger and Hall’s (1965) study was one of the first attempts to circumvent the inevitable analytic problems associated with censored events in studying the timing of event occurrence with traditional research methodologies. Without access to a coherent methodology to model the relationship between when students dropped out and known explanatory variables, Barger and Hall (1965) were forced to focus exclusively on the noncensored individuals in their sample. Exploring the timing of student withdrawal, Barger and Hall (1965) dichotomized dropouts as early or late based on the week during which a student withdrew from college. Students who quit before the tenth week of the first semester were identified as early dropouts and those who subsequently quit by the end of that semester were identified as late dropouts. By dichotomizing dropouts this way, Barger and Hall (1965) were able to incorporate the timing of dropout into their study. However, such a dichotomization eliminated those students who were still enrolled at the institution from the sample. By focusing exclusively on students who had dropped out, Barger and Hall’s (1965) study illustrates

the difficulty early researchers had addressing censored events in studying the timing of student attrition. Those students who were still enrolled at the end of the data collection period had not yet experienced the event of interest, dropping out. The inability to categorize enrolled students using the dichotomization of early and late dropouts ultimately led to the exclusion of this population from the study. Recognizing that such a dichotomization sacrifices considerable variability in the outcome, other researchers have attempted to correct the methodological issue of censored event times by imputing time-to-event as the total time of schooling completed by the end of the data collection period (Willett & Singer, 1991). Collecting student data only for the first trimester illustrates a second methodological limitation of early researchers: the inadequate attention paid to the longitudinal nature of student attrition when determining the length of the data collection period (Willett & Singer, 1993). It is expected that the time-to-dropout would be greater than one trimester for those students who enrolled in subsequent semesters but dropped out before completing their degree program.

For studying college student departure, Willett and Singer (1991) recommend following one or more cohorts of new students over multiple years, noting each student's enrollment status along the way. Longitudinal data gathered in multiwave studies permit a more refined and realistic view of student attrition, the ability to track factors associated with a student's decision to drop out or persist, and increase statistical power (Willett & Singer, 1991). Longitudinal studies that follow students over time present several methodological issues, especially when the temporal nature of college enrollment behaviors, such as dropping out or graduating, is of interest. The most notable problem has been defining time for censored observations when the timing of the event occurrence

is the criterion (Willett & Singer, 1991). For example, a researcher interested in the timing of first dropout follows a cohort of students semester by semester for several years. The researcher is confronted with the analytic dilemma of what measure of time to assign to those students who have not yet experienced their first dropout at the close of data collection. Censored cases such as these need to be treated in a manner that allows for the inclusion of their information even though the knowledge about event occurrence is imprecise (DesJardins, 2003). Event history analysis is a statistical technique that allows researchers studying educational transitions such as student departure and graduation to appropriately frame questions about the timing of an event.

Willett and Singer (1991) propose event history analysis as a new method for studying college student dropout because it is specifically designed to study longitudinal processes. A major advantage of event history analysis is the ability to include the entire cohort of students under study, both those who experience the event of interest and those who do not. In addition to addressing the censoring issue that has long plagued researchers studying student attrition and retention, this methodological approach allows researchers to reframe the retention question from whether students leave a particular institution to when are students most at risk of leaving the institution (Willett & Singer, 1991). Compared to the cross-sectional designs of earlier studies, event history analysis is a statistical technique well-suited for studying student retention but has been used very infrequently in educational research (Ishitani & DesJardins, 2002; DesJardins, 2003). Willett and Singer's (1991, 1993) efforts to extend event history analysis to educational research accompanied by new developments in statistical computing has increased its application to the study of student retention. There is now a small but growing body of

published literature using event history analysis to study the timing of student departure (Chen & DesJardins, 2008; DesJardins et al., 1999; DesJardins et al., 2002a; DesJardins et al., 2002b; DesJardins et al., 2006; DesJardins et al., 2002; Ishitani & DesJardins, 2002; Ishitani, 2003; Ishitani, 2006; Murtaugh et al., 1999). Although the time to first dropout was the main criterion of interest in earlier studies of student retention using event history analysis, more recent studies have taken Adelman's (1999) suggestion to focus on graduation as the outcome of interest and have increased the time frame of data collection and analysis to account for increasing time-to-degree rates (e.g., Chen & DesJardins, 2008; DesJardins et al., 2006; Ishitani, 2003; Ishitani, 2006).

### **Using Event History Analysis to Study College Student Retention**

Although researchers have been successful in using longitudinal studies to identify factors associated with student departure before graduation, most of these studies have ignored the temporal nature of student retention and graduation because of methodological limitations (DesJardins, 2003; Willett & Singer, 1991). Arguing that much more could be learned about educational transitions by answering research questions about whether events occur by modeling when events occur, Willett and Singer (1991) proposed event history analysis as a new method for studying student dropout behavior. Willett and Singer (1993) identified five problems with the traditional approaches to studying event occurrence in longitudinal studies.

First, traditional statistical summaries are inextricably linked to the particular time frame chosen for data collection and analysis, which is rarely substantively motivated. For example, a researcher studying the graduation rate of a cohort of students after six years of enrollment is simply describing the cumulative difference in graduation until that



time. Without accounting for the timing of graduation during that six year period, all other variation in graduation over the six year time period is lost. The failure to document variation in event occurrence over time limits the ability of traditional methodologies to discover what predicts event occurrence.

Second, contradictory conclusions can result from variation in the particular time frames studied. For example, a researcher studying graduation behavior on a cohort of students may come to different conclusions when the time frame of the study is four years versus six years, especially at large public universities where students are less likely to graduate in four years. In the study of college student retention and graduation, it is particularly problematic to understanding the problem if conclusions change for no other reason than differences in the time frames studied.

Third, traditional analytic methods offer no systematic mechanisms for dealing with censored observations, or those individuals who do not experience the event of interest during the time frame of the study. If all censoring occurs at the same point in time, then traditional methods are adequate for collapsing the sampled individuals into two groups: those who experienced the event before the censoring point and those who did not experience the event at the censoring point. If the timing of the event is believed to result in two groups that are systematically different, then the dichotomization approach employed by traditional methods conceals such differences.

Fourth, if censoring times vary across individuals, the risk periods vary as well. For example, in studies involving multiple cohorts and a single fixed time period, some individuals are followed for longer periods of time and others for shorter periods of time. Individuals followed for longer periods of time have a greater likelihood of experiencing

the event of interest than those individuals who are followed for shorter periods of time. When the censoring of observations does not occur at the same time for everyone under study, then observed differences in rates of event occurrence may be attributable to nothing more than research design.

Fifth, few mechanisms exist for the inclusion of time-varying predictors, or those variables whose values vary from one time period to another. Researchers studying the influence of time-varying predictors with traditional methods may ignore information either by using the value of the predictor at a single time point or by pooling predictor values over time (e.g., average). In studying student retention, such approaches result in the loss of much information about the impact of enrollment and financial aid variables which typically vary over time.

DesJardins (2003) also noted several problems with using cross-sectional data and traditional methodological approaches to study event occurrence for longitudinal processes. First, assessing a substantive process at a single point in time assumes statistical equilibrium exists. This assumption is violated when studying complex longitudinal processes such as college enrollment behavior. Student retention to degree completion is not a time-invariant process, and neither are the variables that explain the process time-invariant. Inferences made about a complex longitudinal process based on traditional methodological approaches do not assess how relationships change over time which is an important part of explaining a temporal process. Second, establishing the direction of causality using a “snapshot” of time approach is almost impossible, and doing so with traditional statistical methods can result in a misleading picture of the temporal process being studied. Third, non-random processes, such as self-selection and

time-related selectivity issues, are difficult to control because information is available for only one point in time. Fourth, they cannot distinguish age and cohort effects. Fifth, it is difficult to untangle reciprocal effects that take place over time.

For these reasons, Willett and Singer (1993) and DesJardins (2003) argue event history analysis is a more appropriate statistical technique for studying college student enrollment behavior. Although limited, event history analysis has been applied to the study of college student retention and graduation (Chen & DesJardins, 2008; DesJardins et al., 1999; DesJardins, Ahlburg, & McCall, 2002a; DesJardins, Ahlburg, & McCall, 2006; DesJardins, McCall, Ahlburg, & Moye, 2002; Ishitani, 2003; Ishitani, 2006; Murtaugh, Burns, & Shuster, 1999). Earlier studies on student retention using event history analysis have primarily focused on the time to first dropout as the main criterion of interest. Following Adelman's (1999) suggestion that degree completion should be the criterion of interest when studying student attrition, more recent studies have focused on graduation as the outcome of interest and have increased the time frame of data collection and analysis to account for increasing time-to-degree rates (e.g., Chen & DesJardins, 2008; DesJardins et al., 2006; Ishitani, 2006).

### **Time to First Dropout**

Applying event history analysis to the study of retention of 8,867 undergraduate students at Oregon State University between 1991 and 1996, Murtaugh et al. (1999) examined the following variables on student retention: age of first enrollment, sex, ethnicity/race, residency, college at first enrollment, high school GPA, SAT score, first quarter GPA, participation in Educational Opportunities Program, and Enrollment in Freshman Orientation Program. Murtaugh et al. (1999) reported the following findings:

student attrition increased with age of enrollment, student attrition decreased with increasing academic performance as measured by high school GPA and first-quarter GPA; out-of-state students had higher attrition rates than their in-state and international peers; African American students were less likely to withdraw than White students; and graduation rates decreased with age. Murtaugh et al. (1999) also reported significant effects for college of first enrollment. Combining data from all five cohorts and using a single fixed time period, fall 1996, to end data collection and analysis is extremely problematic to their study of number of years enrolled. Without an adequate time period for the later cohorts, the risk set for each time period is inflated by the subsequent cohort, which can seriously bias the results.

DesJardins et al. (1999) used event history analysis to model the time to first dropout of 3,975 students at the University of Minnesota who enrolled for the first time in fall 1986 at the Minneapolis campus. The researchers collected twenty-two terms of data on these students from various institutional sources. Commenting on the inclusion of only institutional data in their model, DesJardins et al. (1999) argue that the use of data readily available to institutional researchers may be advantageous compared to the attitudinal data espoused in the models of student departure put forth by Tinto (1975), Bean (1980), and Cabrera et al. (1993). As event history analysis requires the researcher to delete all records with missing information, non-response to surveys can result in a biased sample and loss of statistical precision (DesJardins et al., 1999; DesJardins, 2003). The inclusion of variables identified in the existing literature as well as disaggregated financial aid variables makes this study one of the most comprehensive studies of college student departure. In terms of dropout behavior, the researchers reported the following findings

for their model of time to first departure: African American students were more likely to dropout than White students in year three, but no significant relationship was found between Hispanics and White students; no gender differences in dropout; disabled students were less likely than the general population to leave in year two, but were more likely to exit in year four; high school rank percentile had no significant effect on dropout, but students who scored high on the ACT test were less likely to depart in year two; students from the local area were less likely to drop out in year one; age at enrollment was positively associated with dropout in year one, suggesting older students have a much more difficult time adjusting to their academic careers; semester GPA had a negative effect on dropout, but this effect was also found to diminish over time; student loans were more likely to reduce the risk of dropping out in year three, year five, and beyond; the effect of student employment was quite constant and reduced dropout risk at year two and years four through seven; in contrast, work study reduced the risk of dropping out in year one only; scholarships helped reduce the risk of dropping out in year three; and no significant effect of grants on dropping out.

Using national survey data and event history analysis, Ishitani and DesJardins (2002) investigated the departure behavior of 3,450 first-time college freshman over a 5-year period. To examine whether including time-varying effects would improve model fit, the researchers built two event history models: (1) an exponential model, and (2) a time-varying model with time-dependent variables. Although adding a time-dimension to the study of student departure, the exponential model was similar to early student attrition and retentions studies in that it assumed the effects of the explanatory variables were constant over time (i.e., their parameters do not vary over time). Relaxing the exponential

assumption, the time-varying model with time-dependent variables allowed the parameters of the explanatory variables to vary over time. Comparing the log-likelihood statistics for the two models, the researchers found the latter model significantly improved model fit. In addition to providing empirical evidence that explanatory factors affecting student departure have effects that change over time, the researchers found that the amount and timing of student financial aid had varying effects on student attrition rates. For example, not only were low-income students more likely to depart from college but coming from a low-income family had a more adverse effect in the second and third years of college than in the first year.

In a follow-up study using the same sample of students from the University of Minnesota as in their 1999 study, DesJardins et al. (2002b) took the empirical results of their hazard model to simulate how financial aid affects students' departure decisions. Modeling the time to first stopout, the results showed that the effects of financial aid vary temporally and by type of aid. The findings indicated that all forms of aid except grants were associated with decreased stopout. Although the second largest source of federal financial aid for college, it appears that at the study institution grants allow students to attend college but do not have a statistically significant effect on student retention. Scholarships were found to have the largest impact on student retention every year. Work study had the next largest impact in the first two years of college but its effect appeared to wane in later years when earnings from campus employment had the larger impact on student retention. Student loans were found to have less impact on student retention than either forms of employment. Using these empirical results, DesJardins et al. (2002b) conducted policy simulations to quantify the effects of changes in financial aid packaging

on stopout behavior using three simulations. The first simulation “original financial aid package versus no financial aid package” was designed to test the effect that the provision of these different forms of aid had on stopout over time. Not surprising, the researchers found that the provision of financial aid reduces stopout over time relative to providing no financial aid at all. The second simulation, “reallocating loans and scholarships” was designed to examine how a policy like that implemented at Princeton University whereby loans are replaced by institutional grants or scholarships would affect stopout at the study institution. The researchers found evidence that the Princeton approach increased retention by reducing stopout, particularly in years three and four, and had effects that are time-varying. The third simulation, “frontloading scholarship aid” was designed to evaluate the policy of providing gift aid to students to entice them to enroll in a particular institution, then, once they are enrolled, require the students to finance their education through work or loans. Both models of frontloading aid were found to increase persistence, although the observed effect was larger in years two and three. The research conducted by DesJardins et al. (1999, 2002b) contributed to the literature using event history analysis to study student departure in two substantial ways. First, their work highlighted the finding that not all forms of financial aid have the same impact and that the impact of a particular form of aid can vary over time. Second, their work on examining the temporal dimensions of financial aid policies on student departure decision-making calls for future researchers to model the time dimensions of aid changes rather than relying on the time-invariant approaches typically used to study student departure.

Ishitani (2003) used event history analysis to study the student attrition behavior among first-generation college students. Ishitani (2003) used a sample of 1,747 undergraduate students who enrolled at a 4-year comprehensive public university in the Midwest in fall 1995 to study the time to first spell of departure. Examining the time varying effects of precollege characteristics, variables in this study included gender, race, parent's education, annual family income, size of hometown, and high school GPA. The primary outcome of this study was the confirmatory finding that first-generation students were more likely to depart the institution than their peers with one college-educated parent or two college-educated parents.

### **Time to Graduation**

Early research using event history analysis to study student departure had primarily focused on dropout and stopout behaviors. In one of the most comprehensive studies of graduation, Adelman (1999) advocated the use of degree completion as the criterion of interest. While discussing Tinto's (1975) Student Integration Model and Bean's (1980) Student Attrition Model, Adelman (1999) also emphasizes the use of variables readily available to institutional researchers and administrators. Although Adelman's (1999) study employed logistic regression modeling in an attempt to identify what contributes most to bachelor's degree completion for students who attend four-year institutions at any point in their college career, the fact that he followed the 1980 High School and Beyond/Sophomore cohort for 13 years reiterated the need for researchers to conduct longitudinal studies that follow one or more cohorts for an extended period of time.



Extending his original focus on first-generation college students to include degree completion, Ishitani (2006) used data from two national data sets sponsored by the National Center for Education Statistics to study 4,427 students who initially enrolled in public and private four-year institutions between 1991 and 1994. Ishitani (2006) again reported that first-generation college students exhibited higher risks of departure than did students of college-educated parents. Moreover, first-generation college students were less likely to complete their degree programs in a timely manner. Compared to first-generations students whose parents never attended college, first-generations students with parents who had some college education were slightly more likely to graduate in a timely manner. As precollege characteristics were central in projecting the odds of timely college graduation among first-generation students in his sample, Ishitani (2006) noted that diverse precollege characteristics exist within the group of first-generation college students and that these precollege characteristics can have prolonging effects on students' time to degree completion.

DesJardins et al. (2002a) applied an event history model to a sample of 3,070 students who matriculated to the University of Minnesota-Twin Cities campus as freshman in fall 1991. Using institutional data, the researchers jointly estimated stopout and graduation behavior as competing events for 6 years. By using a competing risk model, DesJardins et al. (2002a) found that the effect of some variables may indirectly influence a particular outcome in the college enrollment process. In particular, the researchers found that financial aid did not appear to increase graduation in their model, but that these variables facilitated postsecondary degree attainment by reducing stopout behavior. Using a multiple-spells competing risk model of stopout, dropout, reenrollment,

and graduation behavior, DesJardins et al. (2006) examined the effects of interrupted enrollment on graduation from college. Advocating for the use of competing risks models, DesJardins et al. (2006) argue that this approach allows the researcher to more completely capture the complexity of student behavior. One of the primary results of their research was the finding that “students who experience a stopout are more likely to experience subsequent stopouts, and that such a pattern of enrollment behavior is detrimental to the student’s chances of graduation, especially in a timely fashion” (p. 388). Again, high school rank percentile was found to have very little impact on student outcomes when other factors were controlled for, whereas ACT test score had an independent effect. Another consistent finding with previous research was the influence of college performance, as measured by grades, on timely progression toward a degree and eventual degree attainment.

Using national survey data and event history analysis, Chen and DesJardins (2008) explored the effects of financial aid on the relationship between parental income and student dropout behavior. Their findings confirm a gap in dropout rates for low-income students compared to their upper income peers. Also, the researchers found that some types of aid were associated with lower risks of dropout behavior. To avoid main-effect bias, the researchers incorporated interaction effects between financial aid type and parental income into their model. Receipt of a Pell grant appeared to narrow the dropout gap between low- and middle-income students, although the interaction between receiving a Pell grant and income was not significant. Similar effects on student dropout across all groups were reported for student loans and work study.

## Summary

Although a relatively new methodology to educational researchers studying student retention and graduation, the temporal nature of event history analysis has already proven beneficial to researchers interested in the time-varying effects of known predictor variables. By reframing the student retention question from whether students leave a particular institution to when are students most at risk of leaving the institution, event history analysis is argued to be more informative than past studies utilizing cross-sectional methods. Such a method also allows researchers to address a major methodological issue – the censoring of individuals who have not yet experience the event of interest during the time frame of the study.

Based on previous literature, advancements in research methodology, and assumptions about using student record data readily available to institutional researchers, this study proposes the study of a single institution using event history analysis to model the time to degree completion for multiple cohorts. Variable selection will be driven largely by the existing literature and its subsequent availability in the University's data bases. Chapter III provides a more complete description of the study design and research methodology.

## CHAPTER III. STUDY DESIGN AND METHODOLOGY

### **Event History Modeling**

The present study used event history analysis to examine the temporal nature of college graduation behavior. Event history analysis is a regression-like technique originating out of biostatistics that allows researchers to answer research questions about the occurrence and timing of events (Allison, 1982; DesJardins, 2003; Singer & Willett, 1991). Although used infrequently in educational research, event history modeling is a method specifically designed to study complex longitudinal process such as graduation. The extension of event history modeling to educational research accompanied by new developments in statistical computing has allowed researchers studying student retention to examine the relationship between the timing of events and the factors that affect those events. By including a time dimension to the study of college student retention, researchers can focus on the time periods when college students are most “at-risk” for a particular event of interest (e.g., dropout, graduation). The basic concepts underlying event history analysis are presented to orientate the reader to the methodology.

**Censoring.** In studying the time to event occurrence for longitudinal processes, it is likely that the target event is not determinable for some individuals within the observation period. This phenomenon is known as censoring, and can cause estimation problems such as severe bias or loss of information when analyzing longitudinal events with traditional statistical techniques (DesJardins, 2003). There are two types of censoring: right and left. Right censoring occurs when an individual does not experience the target event, or has experienced some other terminating event, by the end of the data collection period (Allison, 1982; DesJardins, 2003; Singer & Willett, 2003; Willett &

Singer, 1991; Willett & Singer, 1993). The event histories of such individuals are described as right censored because the researcher has incomplete information about event occurrence and knows only that, if the person ever experiences the event, it is after data collection ends (DesJardins, 2003; Singer & Willett, 1991; Singer & Willett, 2003; Willett & Singer, 1993). That is, the event of interest is to the right of the last data collection point,  $t_n$  (Singer & Willett, 1991). Less common, left censoring occurs when the fundamental outcome, time to the target event, is indeterminable because the origin of time is unknown for an individual. As the start time is known for every college student in the study, left censoring is not a concern and therefore the remainder of the discussion will focus on right censoring. Figure 3.1 below illustrates the concept of right censoring.

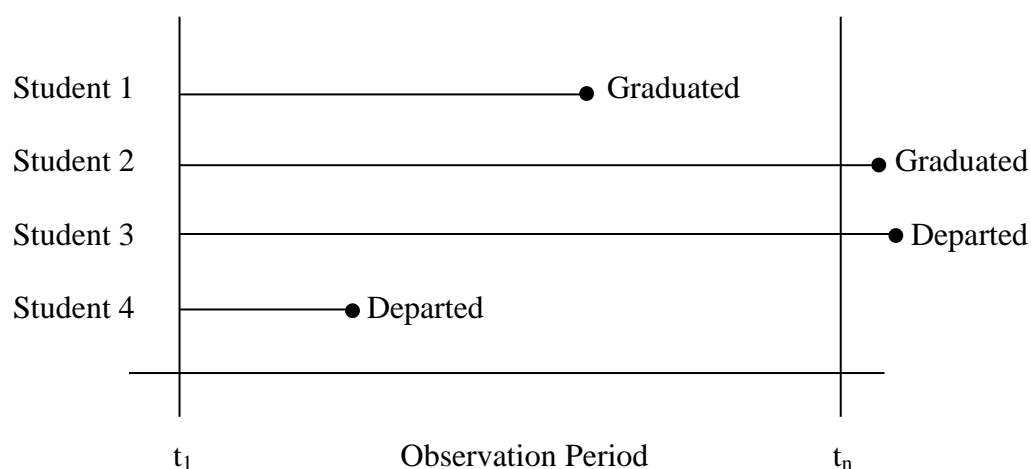


Figure 3.1. An example of right censoring.

If graduation is defined as the outcome of interest, Student 1 is the only unit with a known event time for time to degree completion. Although Students 2 and 3 experience different events (i.e., one graduated and one departed the institution), both are considered right censored because the target (or terminating) event occurred after the data collection period had ended. That is, at the end of the observation period both individuals were still at-risk for graduating. Student 4 is also considered right censored because the individual

experienced a terminating event (i.e., departing the institution) before the target event occurred (i.e., graduating). By departing the institution before the end of the observation period, Student 4 is no longer at risk for graduating. Unlike traditional analytic techniques such as regression and structural equation modeling, event history analysis allows researchers to easily incorporate information about right censored cases (DesJardins, 2003). Event history modeling addresses right censoring by the way in which time to event occurrence is developed when structuring the data set, and will be discussed in a subsequent section.

**Survivor function.** Event history analysis begins with the survivor function. The survival function is a chronologically ordered plot of survival probabilities over time that illustrates a cumulative longitudinal summary of the proportion of participants who have not experienced the event of interest (Singer & Willett, 1991; Willett & Singer, 1993). If  $T$  is the time interval of the event where  $T = \{1, 2, \dots, J\}$ , then the survival probability,  $S_j$ , is the probability of “surviving” beyond the time interval  $j$ , i.e., the probability that the event occurs after interval  $j$ :

$$S_j = P(T > j) \quad (3.1)$$

Until the first censored event time, survival probabilities can be computed directly.

As the data for the present study are collected institutionally on a semester-by-semester basis, the survival probabilities are computed by summing the number of students who had not experienced the event of graduating at the end of a particular semester and dividing it by the total number of students in the study. Therefore, the survivor function in the present study is a monotonically decreasing plot of the survival probabilities for each semester over a six-year period.

**Hazard function.** Although the survivor function is important for summarizing the occurrence of events over time by providing a cumulative longitudinal summary of the proportion of students who have not graduated, it cannot effectively capture the distribution of risk across time because it confounds information about graduation for each semester with cumulative information from the prior semesters (Singer & Willett, 1991). According to Singer and Willett (1991), the fundamental quantity representing the risk of event occurrence in each time period is the hazard. The hazard,  $h_j$ , is a conditional probability of the event occurring in the time interval  $j$ , provided the event has not occurred prior to  $j$ :

$$h_j = P(T = j | T \geq j)$$
(3.2)

In the present study, hazard probabilities are computed by summarizing the number of students who graduated at the end of a particular semester and dividing it by the number of students enrolled at the beginning of that semester. Since the hazard represents the conditional probability that a student will graduate during the current semester given they have not graduated in any prior semester, the hazard function is able to capture the “risk” of graduating over time. The hazard function is a chronologically ordered plot of probabilities over time that illustrates a risk profile for the outcome of interest (Singer & Willett, 1991). According to Willett and Singer (1993), the hazard function is the cornerstone of event history analysis for several reasons: (1) the sample hazard function tells us exactly what we want to know – whether and, if so, when events occur; the sample hazard function includes both noncensored and censored cases; (2) the sample hazard probabilities are computed in every time period that an event occurs – no

information is ignored or pooled; and (3) the sample hazard function can be used to estimate the sample survivor function indirectly in time periods that censoring precludes its direct computation.

**Time-dependent variables.** Researchers studying student retention and graduation with traditional methods have had difficulty accommodating time-dependent variables such as semester grades, financial aid, and enrollment histories (DesJardins et al., 1999). Taking the average across all semesters in the time frame of the study and including that average as a time-independent variable masks the temporal effects of predictor variables that are actually time-dependent (Willett & Singer, 1993). By manipulating the structure of the longitudinal data set (to be discussed in the next section), event history analysis allows researchers to include time-dependent variables into the analysis when data are collected at discrete time periods. In the present study, all of the time-dependent variables have semester-by-semester values.

### **Structuring the Data Set**

For the present study, student data was extracted from institutional databases for six years for the fall 2001, fall 2002, and fall 2003 entering cohorts, excluding winter session, for a total of 18 semesters for each cohort. A person-period data set was created by merging pre-enrollment, enrollment, financial aid, and payroll data. In the person-period data set, each semester a student enrolled was represented by a separate row in the data set, such that right censored observations and time-dependent variables are addressed in the structure of the data set (Singer & Willett, 1991). The records in the person-period data set note what happened to each student during each discrete-time



period when the event of interest could have occurred, until it did occur, or until data collection ended (whichever comes first) (Singer & Willett, 1991).

To address the issue of event occurrence, three additional variables are added to the person-period data set: PERIOD, STATUS, and EVENT. The PERIOD variable specifies the time period  $j$  that the record describes, the STATUS variable indicates whether the student graduated or not (did not graduate = 0, graduated = 1) during the 6-year time period, and the EVENT variable indicates whether the event occurred in that time period (event did not occur = 0, event occurred = 1). For students who graduated, the EVENT variable is used to identify the semester in which they graduated. In the person-period data set, the value of EVENT is 0 for all semesters except the semester during which the student graduated, in which the value for that semester becomes 1. For those students who did not graduate, the value of EVENT remains zero for all semesters enrolled. In the present study, the EVENT variable is the outcome variable.

By restructuring the data set from person-level to person-period, the unit of analysis changes from the individual to the individual's semesters of enrollment. Expanding the number of records per person artificially reduces the variability from the person (0 or 1) to the time period in which the event occurred for EVENT (a string of 0's followed by a value of 1 if the event occurred). This artificial reduction of variability across time periods poses a significant problem when fitting a logistic model on the person-period data set. In order to fit a model for EVENT, the researcher must specify a baseline function by using a set of "time indicators" to index the particular time period a given record represents (Singer & Willett, 2003). As an event can occur in one of  $J$  time periods, the standard dummy variable representation of PERIOD is adequate and yields  $J$

time indicators,  $D_{1ij}$ ,  $D_{2ij}$ , ...,  $D_{Jij}$  (Singer & Willett, 2003). According to Singer and Willett (2003), it is when we refer to the collective set of time indicators using the conceptual label “time” that an important paradox of the event history model is highlighted: although time is a conceptual outcome, it is actually the fundamental predictor of EVENT. Singer and Willett (2003) argue that:

This seeming anomaly reflects our reformulation of the research questions from “What is the relationship between event times and predictors?” to “What is the relationship between the risk of event occurrence in each time period and predictors?” This reformulation is vital, for it is by answering the second question that we answer the first. (p. 371)

Therefore, the constructed person-period data set must also include time indicators in order to fit the logistic regression model.

Figure 3.2 illustrates how the individual records in a person-level data set are restructured into multiple records in the person-period data set for event history analysis. As can be seen in Figure 3.2, the two cases in the person-level data set have been expanded into 21 records in the person-period data set. The 21 records represent three semesters for Student 1 and eighteen semesters for Student 2. For all individuals,  $D_1$  is 1 in the record for the first period,  $D_2$  is 1 in the record for the second period,  $D_3$  is 1 in the record for the third period, and so forth, with all other values set to 0. Student 1 departed the institution after his third semester and therefore has no enrollment history for the remaining 15 semesters. As Student 1 did not experience the event of interest, both STATUS and EVENT have zero values for those three time periods. Student 2 was enrolled for 18 semesters before graduating, and the STATUS and EVENT variables are coded to reflect a graduated enrollment history. The 18 time indicators identify the time period being referenced in the record. Also shown in the restructuring process is the

differential treatment of time-independent and time-dependent variables when constructing the person-period data set. The values for the time-independent variable GENDER remain the same for each student by semester. For example, Student 1 has the value M recorded in all three of his person-period records and Student 2 has the value F recorded in all 18 of her person-period records. The time-dependent predictors SEMESTER 1 GPA, SEMESTER 2 GPA, ... , SEMESTER 3 GPA become a single column of values called SEMESTER GPA, with values appropriate to each time period. To assess the effect of SEMESTER GPA, researchers working in a person-level data set would have to take the average of the 18 semester GPAs or include the entire vector of semester GPAs in their model – neither of which is methodologically optimal. In contrast, the structure of the person-period data set permits the value of SEMESTER GPA to change from semester to semester for each student, and thus allows the researcher to study its temporal effect.

Person-Level Data Set

| ID | TIME | CENSOR | GENDER | SEMESTER 1<br>GPA | SEMESTER 2<br>GPA | SEMESTER 3<br>GPA | SEMESTER 4<br>GPA | ... | SEMESTER 18<br>GPA |
|----|------|--------|--------|-------------------|-------------------|-------------------|-------------------|-----|--------------------|
| 1  | 3    | 0      | Male   | 2.50              | 2.75              | 2.80              | .                 | ... | .                  |
| 2  | 18   | 1      | Female | 3.80              | 4.00              | 3.75              | 3.85              | ... | 3.75               |

Person-Period Data Set

| ID | PERIOD | D <sub>1</sub> | D <sub>2</sub> | D <sub>3</sub> | D <sub>4</sub> | D <sub>5</sub> | D <sub>6</sub> | D <sub>7</sub> | D <sub>8</sub> | D <sub>9</sub> | D <sub>10</sub> | D <sub>11</sub> | D <sub>12</sub> | D <sub>13</sub> | D <sub>14</sub> | D <sub>15</sub> | D <sub>16</sub> | D <sub>17</sub> | D <sub>18</sub> | STATUS | EVENT | GENDER | SEMESTER<br>GPA |
|----|--------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|--------|-------|--------|-----------------|
| 1  | 1      | 1              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0      | 0     | M      | 2.50            |
| 1  | 2      | 0              | 1              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0      | 0     | M      | 2.75            |
| 1  | 3      | 0              | 0              | 1              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0      | 0     | M      | 2.80            |
| 2  | 1      | 1              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 3.80            |
| 2  | 2      | 0              | 1              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 4.00            |
| 2  | 3      | 0              | 0              | 1              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 3.75            |
| 2  | 4      | 0              | 0              | 0              | 1              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 3.80            |
| 2  | 5      | 0              | 0              | 0              | 0              | 1              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 3.95            |
| 2  | 6      | 0              | 0              | 0              | 0              | 0              | 1              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 4.00            |
| 2  | 7      | 0              | 0              | 0              | 0              | 0              | 0              | 1              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 3.50            |
| 2  | 8      | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 1              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 3.75            |
| 2  | 9      | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 1              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 3.00            |
| 2  | 10     | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 1               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 2.80            |
| 2  | 11     | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 1               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 3.00            |
| 2  | 12     | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 1               | 0               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 3.50            |
| 2  | 13     | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 1               | 0               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 3.50            |
| 2  | 14     | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 1               | 0               | 0               | 0               | 0               | 1      | 0     | F      | 4.00            |
| 2  | 15     | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 1               | 0               | 0               | 0               | 1      | 0     | F      | 3.95            |
| 2  | 16     | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 1               | 0               | 0               | 1      | 0     | F      | 3.80            |
| 2  | 17     | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1               | 0               | 1      | 0     | F      | 3.00            |
| 2  | 18     | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0              | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 0               | 1               | 1      | 1     | F      | 3.75            |

Figure 3.2. Restructuring a Person-Level Data Set into a Person-Period Data Set

## Competing Risks

In a standard event history analysis, the researcher investigates event occurrence as an individual's transition from one "state" to another "state," where each individual under study can occupy one, and only one, of two possible states. For example, Willett and Singer (1991) studied teacher attrition as the transition from working as a teacher (state 1) to leaving the teaching profession permanently (state 2). However, there are naturally occurring situations in which an individual can occupy three or more possible states and these two states are referred to as competing risks (DesJardins, 2003; Singer & Willett, 2003). For example, when studying college student careers, researchers track first-year students enrolled at a particular institution as they transition from being enrolled in school (state 1) to one of two competing alternatives: institutional departure (state 2) or graduation (state 3). "Conceptually, the occurrence of a competing event acts like a form of censoring – it removes an individual from the risk set for all other events" (Singer & Willett, 2003, p. 588). In the present study, institutional departure is a competing risk with graduation because premature departure, whether voluntary or involuntary, removes the student from being at-risk of graduating.

Although event history analysis allows for simultaneous analysis of competing risks, the estimation and interpretation of the survival and hazard functions is much more complex. For example, the event-specific survivor function describes the probability that individuals survive given that they have not previously experienced this, or any other, competing event (Singer & Willett, 2003). Because of this complexity, Allison (1995) and Singer and Willett (2003) suggest running separate analyses for each competing event to estimate the event-specific survivor and hazard rates. According to Allison

(1995), the benefits of conducting separate analyses lie in the ability to estimate models only for those events of interest with no loss of statistical precision. In order to run separate event-specific analyses, the person-period data set requires slight modification of the STATUS and EVENT variables before running a logistic regression on the criterion of interest (i.e., graduation). Both of these variables will be coded with a value of 1 to indicate graduation and a 2 to indicate institutional departure.

In this study, the appropriate analysis of graduation will be accomplished by deleting the single record where EVENT has a value of 2. This will result in the elimination of the last semester for any student who permanently left the institution prematurely. In Figure 3.2, the event history model for graduation would eliminate the third semester for Student 1 because the student did not return to the institution. Students who stopped out for a least one semester during the fall or spring semesters and were granted readmission to the institution will remain in the sample as long as they have an enrollment history or experienced the event of interest at the end of the six year observation period.

### **A Statistical Model of Hazard**

In event history analysis, researchers use the person-period data set to model the relationship between the occurrence of the event of interest and one or more predictor variables (Singer & Willett, 1991). Unlike linear regression, the outcome of the event history model is an entire hazard profile (Willett & Singer, 1991). Because the hazard profile is a set of conditional probabilities, each bounded by 0 and 1, a transformation is necessary to reparameterize the hazard probabilities so they are logistically dependent on the predictors and time periods (Singer & Willett, 1993; Willett & Singer, 1991). When

the outcome is a probability, the logit (or log odds) transformation is mathematically and conceptually appealing as the appropriate link function because it improves the distributional behavior of variables, prevents specification of inadmissible values due to the bounded nature of probabilities, and renders the distance between hazard functions more comparable over time (Allison, 1984; Singer & Willett, 2003). To develop the logit-hazard function, the probabilities of the hazard function are first transformed to odds through the formula  $odds = hazard / (1 - hazard)$  and then the natural log of the odds are computed (Willett & Singer, 1993). This transformation is shown in the following equation:

$$\begin{aligned} \text{logit } h_{ij} &= \log_e \left( \frac{h_{ij}}{1 - h_{ij}} \right) \\ &= (\alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \dots + \alpha_J D_{Jij}) + (\beta_1 X_{1ij} + \beta_2 X_{2ij} + \dots + \beta_P X_{Pij}) \end{aligned} \quad (3.3)$$

where for each individual  $i$  and time period  $j$ ,  $h_{ij}$  represents the entire hazard function,  $D_{Jij}$  is a set of  $J$  dummy variables referred to as time indicators,  $\alpha_j$  are the intercept parameters representing the baseline hazard at each time period  $j$ ,  $X_{Pij}$  is a set of predictor variables, and  $\beta_P$  are the slope parameters that describe the effects of the predictors on the baseline hazard function. As a set, the  $\alpha$ 's represent the baseline hazard function (Singer & Willett, 2003). This transformed function is the one modeled in event history analysis and refers to the log odds of event occurrence in any time period, given that the event did not occur in an earlier time period (Willett & Singer, 1993). Hazard models such as Equation 3.3 closely resemble what is traditionally described through logistic regression (Singer & Willett, 1991).

To illustrate, Equation 3.4 below represents the entire hazard function being modeled for the time-independent variable GENDER and time-dependent variable SEMESTER GPA:

$$\text{logit } h_{ij} = (\alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \dots + \alpha_j D_{jij}) + (\beta_1 \text{GENDER} + \beta_2 \text{SEM GPA}_{2ij}) \quad (3.4)$$

It is the addition of predictor variables to the set of  $\alpha$ 's that change the identity of the hazard function (Singer & Willett, 2003). When a predictor variable is time-independent, it will shift the logit hazard function the same amount for each time period. In this example, when the value of GENDER is equal to 1, the logit hazard function shifts by  $\beta_1$  regardless of the time period because the value of gender does not vary from semester to semester. The inclusion of the subscript  $j$ , indexing time periods, on the time-independent variable GENDER is dropped because it is redundant when the value of the predictor does not change from one time period to the next (Singer & Willet, 2003). When a predictor variable is time-dependent, the shift in the logit hazard function is a function of the value of the variable multiplied by its beta coefficient. In this example, SEMESTER GPA would be multiplied by  $\beta_2$  and the resulting shift in the hazard function would vary from one semester to another if there was an observed change in GPA between semesters.

When fitting an event history model, parameters are estimated using maximum likelihood methods. The likelihood function for the event history model in Equation 3.3 is:

$$L = \prod_{i=1}^n \prod_{j=1}^{J_i} h_{ij}^{\text{EVENT}_{ij}} (1 - h_{ij})^{1 - \text{EVENT}_{ij}} \quad (3.5)$$

where EVENT is defined as 0 if individual  $i$  does not experience the event in time period  $j$  and 1 if individual  $i$  does experience the event in time period  $j$ . The likelihood function



in Equation 5 expresses the probability of observing the sample data on event occurrence that we actually observed as a function of unknown population parameters (Singer & Willett, 2003). Taking the logarithm of Equation 3.5 yields the log-likelihood function:

$$LL = \sum_{i=1}^n \sum_{j=1}^{J^i} EVENT_{ij} \log(h_{ij}) + (1 - EVENT_{ij}) \log(1 - h_{ij}) \quad (3.6)$$

When the logistic regression procedure is applied to the person-period data set, we get parameter estimates of the  $\alpha$ 's and the  $\beta$ 's (and hence  $h_{ij}$ ) of the event history model that maximizes the log-likelihood in Equation 6 (Singer & Willett, 1993). As each discrete time period for every individual is treated as the unit of analysis in the person-period data set, the logistic regression procedure essentially pools these semester observations and computes the maximum likelihood estimates (Allison, 1982; Allison, 1984).

### **Sample**

The present study used event history analysis to model time to degree completion at Rutgers University for three cohorts of first-time, full-time, degree-seeking undergraduates enrolled at the New Brunswick campus. Today, Rutgers University is a large, doctoral-granting public AAU in the Northeast enrolling over 58,000 students across three regional campuses: New Brunswick, Newark, and Camden. Only students attending the flagship New Brunswick campus were used in the analysis to eliminate potential effects attributable to individual campus locations (e.g., SAT scores).

Originating as one of the first colonial colleges and later established as New Jersey's land-grant institution, Rutgers University has a history both past and present of reorganization while maintaining a highly decentralized structure among the academic

units. For example, the New Brunswick campus was reorganized from a federation of colleges, each with their own faculty, to a comprehensive research university in 1980 and operated for over 25 years with four distinct undergraduate liberal arts colleges – Douglass College, Livingston College, Rutgers College, and University College – each of which had different academic expectations and graduation requirements (Transforming Undergraduate Education [TUE], 2005). The perceived superiority of one of the liberal arts colleges not only made it difficult for the other colleges on the New Brunswick campus to recruit high achieving students, the structure of having four distinct liberal arts colleges impeded the university's efforts to distribute resources and opportunities equitably, prevented student movement between colleges and between departments, and discouraged faculty and student interactions beyond those located in the major (TUE, 2005). Such findings led to the creation of the School of Arts and Sciences in 2007, a new school that combined the four liberal arts colleges under a coherent set of standards and policies. Another outcome of the TUE initiative was the reorganization of Cook College, a professional school prominent in agricultural and environmental sciences, into the School of Environmental and Biological Sciences. Having such a recent dramatic reorganization not only makes Rutgers University an unusual case compared to its peer institutions but also impacts the university's historical data. For example, when reporting at the academic unit level Rutgers combines the historical data for the four liberal arts colleges prior to the reorganization in 2007 to produce retroactive enrollment numbers and retention and graduation rates for the School of Arts and Sciences. This approach was used in the current study when constructing the data set.

Student data was extracted from the institution's student information system (SIS) for six years starting with the student's entering fall semester for a total of 18 semesters (fall, spring, and summer) for the fall 2001, fall 2002, and fall 2003 entering cohorts. Although students can also enroll in courses at Rutgers during winter break, these semesters will be excluded from the analysis because enrollment in winter session is extremely atypical, accounting for approximately one percent of the semesters of enrollment for a cohort of students over six years. For students who graduated during the winter term, the event time for graduation was recorded on the previous semester of enrollment. Of the 15,047 students in the 2001, 2002, and 2003 entering cohorts who were assigned a Rutgers University Identification Number (RUID), 50 students were deleted from the sample because they took classes at Rutgers prior to matriculating as a first-time, full-time, degree-seeking undergraduate on the New Brunswick campus. Another 149 students were deleted from the sample because they did not complete the enrollment process for the fall semester of their entering cohort, 19 students were deleted for dropping below full-time status during their first semester, and 115 students were deleted because of missing information. An additional 380 students were excluded from the time to degree completion model through the competing risks framework (see Competing Risks section on p. 57), resulting in an effective sample size of 14,336, or 95.3% of the original sample. The number of students from each cohort in the sample are displayed in Table 3.1. As previously discussed, each semester a student enrolled is represented as a separate record in the person-period data set. There were 124,397 person-period records created for the 14,336 students in the sample.

Table 3.1  
*Effective Sample Size by Cohort*

| Cohort | Number of<br>Students<br>Extracted | Number of<br>Students Excluded | Number of<br>Students Included | Number of<br>Person-Period<br>Records |
|--------|------------------------------------|--------------------------------|--------------------------------|---------------------------------------|
| 2001   | 5,389                              | 268                            | 5,121                          | 44,238                                |
| 2002   | 5,024                              | 242                            | 4,782                          | 41,541                                |
| 2003   | 4,634                              | 201                            | 4,433                          | 38,618                                |
| Total  | 15,047                             | 711                            | 14,336                         | 124,397                               |

### **Variable Selection**

Variables for the present study were selected based upon a review of the literature and their availability at the institution under study. An underlying assumption of the present study is that student retention and graduation is a campus-based phenomenon, and for that reason research must be conducted at the institutional level using variables that are readily available to institutional researchers for the vast majority of their students. Three categories of variables are included in the present study: pre-enrollment, enrollment, and financial aid.

**Pre-enrollment variables.** Incorporating both student demographic characteristics as well as common measures of student academic ability collected during the admissions process, the pre-enrollment variables in the present study include gender, enrollment age, entering cohort, first generation college student, race/ethnicity, residency, and SAT scores (math and verbal), and transfer credits. Gender is included in the model to examine whether there are longitudinal differences in graduation by gender at the study institution. Gender is specified by inclusion of a dummy variable with males as the reference group. Age at the time of enrollment is also included in the model to examine whether there are longitudinal differences in graduation for non-traditionally aged students at the study institution. Although age increases with each passing year, only the

initial age of the student is included in the model to distinguish between traditional and non-traditional aged students. Entering cohort is included in the model to test for longitudinal differences in graduation for students who enrolled in college during fall 2001, fall 2002, and fall 2003. The cohort variable will be dummy coded with the 2001 cohort as the reference group. First generation college student status is included in the model to test for longitudinal differences in graduation for students who come from a family in which neither parent has obtained a 4-year college degree. The first generation variable is comprised of a set of two dummy variables (first generation and unknown). The reference group is students who do not classify as first generation college students because at least one of their parents has a 4-year college degree. Residency is included in the model to examine whether there are longitudinal differences in graduation by residency at the study institution. As a public university, Rutgers charges New Jersey residents a discounted tuition price relative to the non-resident tuition it charges out-of-state and international students. Although previous studies generally compare in-state and out-of-state students, the present study also includes international students as they are a population of interest at the study institution as the university aims to increase diversity and tuition revenue while capping the size of the entering first-year class. Residency is entered into the model by inclusion of two dummy variables (out-of-state and international). The reference group is in-state students. Race/ethnicity is included in the empirical specification of the model because research shows that minority students tend to have lower probabilities of graduation than non-minority students. Race/ethnicity is specified in the model by a series of dummy variables, which include Asian/Pacific Islander, Black/African American, Hispanic, and Other. The reference group is White.

SAT math and verbal scores are included in the model to provide a standardized measure of a student's pre-enrollment math and verbal abilities. The highest SAT math and verbal scores reported to the university will be used. The number of transfer credits at time of matriculation is included in the model to test for longitudinal differences in graduation for students who enrolled at the university with some prior college experience. Although the fall 2001, fall 2002, and fall 2003 entering cohorts entered the university as first-time, full-time, degree-seeking undergraduates, a portion of these students were awarded university credit toward their degree program for having some prior college experience. For example, many high school students will earn college credit by taking Advanced Placement courses while in high school. Although many studies also include the student's high school percentile rank as an additional measure of academic ability, there has been a growing trend among high schools over the years to stop providing class ranking information to colleges and universities. As a result, approximately 30 percent of the students in the entering fall 2001, fall 2002, and fall 2002 cohorts at the study institution do not have a high school percentile rank. To avoid excluding one-third of the sample for missing information, high school percentile rank is not included in the model specification as this information is no longer readily available for all of the students. All pre-enrollment variables are treated as time-independent as their values do not change once a student enrolls in college.

**Enrollment variables.** Enrollment variables represent information available after students have enrolled at the university. Most of these variables are time-dependent as their values vary from semester to semester. Enrollment variables include initial school of enrollment, student housing, cumulative grade point average, credits attempted,

percentage of credits completed, and school-to-school transfers. The study institution considers an undergraduate student full-time if he or she carries a credit load of 12 or more credits each fall or spring term. Although taking courses during the summer is optional, a student is considered studying full-time during summer session at 6 or more credits. Enrollment status is specified by inclusion of a dummy variable indicating whether the student was studying full-time or part-time during that semester of enrollment. Full-time status will be used as the reference group. The initial school of enrollment of a student is included in the model as time-independent to examine whether there are college specific environmental factors that help to explain graduation. Rutgers College, Douglass College, University College, and Livingston College will be recoded as the School of Arts and Sciences to reflect the 2007 merger of Rutgers' undergraduate liberal arts colleges. Cook College will also be recoded to the School of Environmental and Biological Sciences to reflect the 2007 name change. Initial school of enrollment will be specified by inclusion of a series of dummy variables (School of Environmental and Biological Sciences, School of Engineering, Ernest Mario School of Pharmacy, and Mason Gross School of the Arts). Enrolling the largest percentage of undergraduate students, the School of Arts and Sciences will be the reference group. Due to the shortage of on-campus housing after the first year and the high cost associated with living in on-campus residences, student housing will be included in the model and allowed to vary by term to examine whether there are longitudinal differences in graduation by changes in a student's housing situation while in college. Student housing will be specified by including a dummy variable indicating whether the student is living on-campus or commuting. As Rutgers is a primarily residential institution, students living in on-campus

housing will be coded as the reference group. A student's cumulative grade point average (GPA) for each term of enrollment is included to control for variations in academic performance. The number of credits attempted and the percentage of credits completed each term are included in the model to control for variations in credit loads and the subsequent completion of those credit hours by the student. As the federal government tightens the satisfactory academic progress regulations regarding Title IV aid eligibility, it is important to develop a better understanding of the longitudinal effect of credit completion on graduation. School-to-school transfers is included in the model as a time-independent variable to determine the effect of changing academic schools of study within the institution. Consistent with federal reporting, no differentiation is made between students transferring between schools on the New Brunswick campus and those transferring to a school on one of the regional campuses. The number of school-to-school transfers was categorically coded as no transfers and at least one transfer, with no school-to-school transfers coded as the reference group.

**Financial aid variables.** Financial aid variables are included in the present study to examine the effects of paying for college on graduation. All of the financial aid variables are treated as time-dependent. Although financial aid is packaged for an academic year at the study institution, the values are included for each term as a change in a student's enrollment status (e.g., full-time to part-time status) changes their eligibility for award amounts and borrowing limits. Although full-time status at the time of enrollment was a prerequisite for inclusion in the sample, students remained in the sample if they dropped to part-time status during subsequent semesters. Disaggregating the total financial aid offered to a student by aid amounts and types is useful for



examining the differential effects that various types of aid have on graduation behavior. In the present study, financial aid is disaggregated into four variables: grants, scholarships, student loans, and federal work study earnings. Although many event history models at the institutional level also include on-campus (non-federal work-study) earnings as a financial aid variable, this data is currently not stored in a readily available format in the old payroll system at Rutgers and the Payroll department was too overwhelmed with implementing a new payroll administrative information system to fulfill the data request within the study timeline. Table 3.2 lists all of the predictor variables included in the present study, identifies whether it the variable was treated as time-independent or time-varying in the model, and specifies the reference group if the variable is categorical.

Table 3.2  
*Time to Degree Completion Covariates*

| Variables                        | Variable Classification | Reference Group         |
|----------------------------------|-------------------------|-------------------------|
| Pre-Enrollment                   |                         |                         |
| Cohort                           | Time-independent        | 2001                    |
| Gender                           | Time-independent        | Male                    |
| Enrollment Age                   | Time-independent        | -                       |
| Race/Ethnicity                   | Time-independent        | White                   |
| First Generation College Student | Time-independent        | Not first generation    |
| Residency                        | Time-independent        | In-state                |
| SAT Math                         | Time-independent        | -                       |
| SAT Verbal                       | Time-independent        | -                       |
| Transfer Credits                 | Time-independent        | -                       |
| Enrollment                       |                         |                         |
| Initial School of Enrollment     | Time-independent        | Sch. of Arts & Sciences |
| Student Housing                  | Time-varying            | On-Campus               |
| Credits Attempted                | Time-varying            | -                       |
| Percentage of Credits Earned     | Time-varying            | -                       |
| Cumulative GPA                   | Time-varying            | -                       |
| School-to-School Transfer        | Time-independent        | None                    |
| Financial Aid                    |                         |                         |
| Grants                           | Time-varying            | -                       |
| Scholarships                     | Time-varying            | -                       |
| Student Loans                    | Time-varying            | -                       |
| Work Study Earnings              | Time-varying            | -                       |

### Data Analysis

Descriptive statistics are provided to characterize the relationship between the pre-enrollment, enrollment, and financial aid factors and graduation behavior. The survivor and hazard probabilities for the 18 time periods were computed and then plotted to obtain the survivor and hazard functions for the sample. To model the time to degree completion, binary logistic regression was used on the person-period data set to estimate associations between the pre-enrollment, enrollment, and financial aid variables and *J-1* time indicators in SPSS. Model fit was evaluated using the likelihood ratio test.

## CHAPTER IV. RESULTS

In this chapter, a detailed discussion is given of the results obtained following the application of the methodology outlined in Chapter III. Specifically, descriptive statistics are provided for the criterion and predictor variables, followed by the survival and hazard probabilities for the 18 time periods and the plotted survivor and hazard functions for the sample. Results from the binary logistic regression on the person-period data are then presented, as well as interpretation of the individual parameter estimates and model fit.

### Descriptive Statistics

A summary of the postsecondary educational output of the sample at the end of the 6-year observation period is provided in Table 4.1. As can be seen from Table 4.1, 20.7 percent of the sample departed the institution without graduating within the six-year time period. Of the 10,948 students that graduated within 6 years, the majority of the students graduated within four years.

Table 4.1  
*Postsecondary Educational Output of the Sample*

| Enrollment Status After 18 Semesters | Count  | Percent |
|--------------------------------------|--------|---------|
| Graduated                            | 10,948 | 76.4    |
| Departed Institution                 | 2,963  | 20.7    |
| Still Enrolled                       | 425    | 3.0     |
| Degree Completion Behavior           |        |         |
| Graduated Within 4 Years             | 7,297  | 66.6    |
| Graduated in 5th Year                | 2,809  | 25.7    |
| Graduated in 6th Year                | 843    | 7.7     |

Descriptive statistics of the sample of first-time, full-time, degree-seeking undergraduates are displayed in Table 4.2. Students in the sample were largely in-state residents, with 89.6 percent coming from the institution's home state. The sample is predominantly traditionally aged, with a mean age of 18.22 years. The majority of the

students lived residentially on campus their first semester, with approximately 15 percent commuting to campus for classes. About 54 percent of the sample was female, and 44.3 percent identified as non-White. Approximately 14 percent of the students were identified as first generation college students whose parents did not graduate from a four-year college. The largest initial school of enrollment was the School of Arts and Sciences, with 69.4 percent of the entering first-year students in the sample. Over 90 percent of the students remained in their initial school of enrollment during their time at the study institution. The majority of the students in the sample enrolled at Rutgers with no college credit, and those students who did transferred an average of 9.25 credits. The sample enrolled with an average SAT Math score of 606.82 and an average SAT Verbal score of 578.41. Students enrolled in an average of 15 credit hours, and earned approximately 83 percent of the credits in which they attempted. The sample had an average cumulative grade point average of 2.83 for the first semester.

Descriptive statistics of the financial aid package awarded are displayed in Table 4.3. Approximately 75 percent of the sample received some form of financial aid (grants, scholarships, student loans, federal work study) entering Rutgers University and the average aid awarded during their first year was \$9,331. Of the 10,723 students who received financial aid, 42.9 percent received one or more grants and the average grant awarded the first semester was \$3,316. Approximately 15 percent received one or more scholarships and the average scholarship amount applied to their financial aid package their first semester was \$1,293. About 70 percent of the sample received a student loan their first semester and the average student loan disbursed was \$1,460. Approximately 17

percent were employed part-time on-campus through the Federal Work Study Program and the average amount earned during their initial semester was \$569.

Table 4.2  
*Descriptive Statistics of the Sample at Semester 1*

| Variable                         | Label               | <i>n</i> | %    | <i>M (SD)</i> |
|----------------------------------|---------------------|----------|------|---------------|
| Cohort                           | 2001                | 5,121    | 35.7 | -             |
|                                  | 2002                | 4,782    | 33.4 | -             |
|                                  | 2003                | 4,433    | 30.9 | -             |
| Gender                           | Male                | 6,671    | 46.5 | -             |
|                                  | Female              | 7,665    | 53.5 | -             |
| Race/Ethnicity                   | White               | 7,979    | 55.7 | -             |
|                                  | Asian/Pac. Islander | 3,395    | 23.7 | -             |
|                                  | Black               | 1,223    | 8.5  | -             |
|                                  | Hispanic            | 1,294    | 9.0  | -             |
| First Generation College Student | Other               | 445      | 3.1  | -             |
|                                  | No                  | 11,646   | 81.2 | -             |
|                                  | Yes                 | 1,957    | 13.7 | -             |
|                                  | Unknown             | 733      | 5.1  | -             |
| Residency                        | In-State            | 12,843   | 89.6 | -             |
|                                  | Out-of-State        | 1,317    | 9.2  | -             |
|                                  | International       | 176      | 1.2  | -             |
| Initial School of Enrollment     | SAS                 | 9,951    | 69.4 | -             |
|                                  | SEBS                | 1,795    | 12.5 | -             |
|                                  | Engineering         | 1,604    | 11.2 | -             |
|                                  | Pharmacy            | 610      | 4.3  | -             |
|                                  | Mason Gross         | 376      | 2.6  | -             |
| Student Housing                  | On-Campus           | 12,197   | 85.1 | -             |
|                                  | Commuter            | 2,139    | 14.9 | -             |
| School-to-School Transfers       | None                | 13,046   | 91.0 | -             |
|                                  | At Least One        | 1,290    | 9.0  | -             |
| Transfer Credits                 | No college credit   | 13,310   | 92.8 | -             |
|                                  | Some college credit | 1,026    | 7.2  | 9.25 (7.461)  |

| Variable                     | Min  | Max   | <i>M (SD)</i>   |
|------------------------------|------|-------|-----------------|
| Enrollment Age               | 12   | 26    | 18.22 (.516)    |
| SAT Math                     | 290  | 800   | 606.82 (82.964) |
| SAT Verbal                   | 210  | 800   | 578.41 (78.824) |
| Cumulative GPA               | 0.00 | 4.00  | 2.83 (.754)     |
| Credits Attempted            | 12.0 | 27.0  | 15.2 (1.729)    |
| Percentage of Credits Earned | 6.06 | 100.0 | 83.19 (17.598)  |

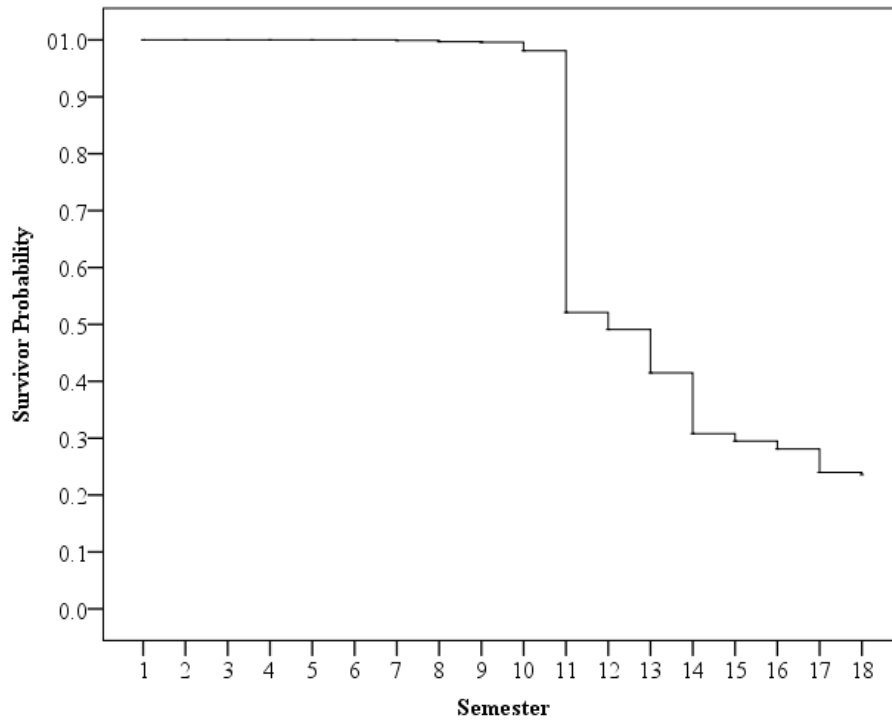
Table 4.3  
*Descriptive Statistics of Financial Aid Package*

|                          |             | <i>n</i> | %    | Average Total Aid<br>Year 1 |
|--------------------------|-------------|----------|------|-----------------------------|
| Financial Aid            | Awarded     | 10,723   | 74.8 | \$9,330.98                  |
|                          | Not Awarded | 3,613    | 25.2 |                             |
| Financial Aid Recipients |             |          |      | Average Award<br>Semester 1 |
| Grants                   | Yes         | 4,596    | 42.9 | \$3,315.96                  |
|                          | No          | 6,127    | 57.1 |                             |
| Scholarships             | Yes         | 1,611    | 15.0 | \$1,293.40                  |
|                          | No          | 9,112    | 85.0 |                             |
| Loans                    | Yes         | 7,536    | 70.3 | \$1,460.00                  |
|                          | No          | 3,187    | 29.7 |                             |
| Work Study Earnings      | Yes         | 1,860    | 17.3 | \$569.12                    |
|                          | No          | 8,863    | 82.7 |                             |

*Note.* Average award shown for grants, scholarships, student loans, and work study earnings are calculated for only those students who received that specific type of aid their first semester.

### **Survivor Function and Survivor Probabilities**

As previously discussed, the survival function is a chronologically ordered plot of survival probabilities over time that illustrates a cumulative summary of the proportion of students who have not graduated. The survivor function for the present study is displayed in Figure 4.1, and the corresponding survival probabilities of this plot are shown in Table 4.4. As shown in Table 4.4, Semester 7 was the first semester in which students graduated and the largest number of students graduated in Semester 11. Approximately 50 percent of the sample had graduated by the end of four years (Semester 12). By the end of the study period, 23.6 percent of the sample had not graduated.



*Figure 4.1.* Sample Survivor Function for Time to Degree Completion. Survivor function shown is unconditional (i.e., not based on regressors).

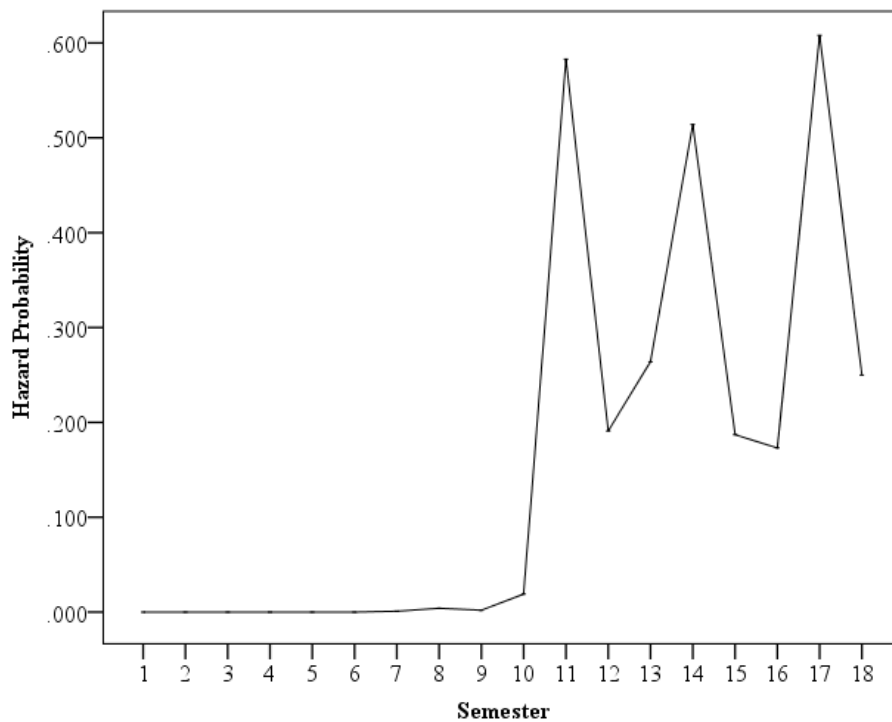
Table 4.4

*Sample Survival Probabilities for Time to Degree Completion*

| Semester | Term   | Survivors<br>(Not Graduated) | Total<br>Graduated | Survival<br>Probability |
|----------|--------|------------------------------|--------------------|-------------------------|
| 1        | Fall   | 14,336                       | 0                  | 1.000                   |
| 2        | Spring | 14,336                       | 0                  | 1.000                   |
| 3        | Summer | 14,336                       | 0                  | 1.000                   |
| 4        | Fall   | 14,336                       | 0                  | 1.000                   |
| 5        | Spring | 14,336                       | 0                  | 1.000                   |
| 6        | Summer | 14,336                       | 0                  | 1.000                   |
| 7        | Fall   | 14,331                       | 5                  | .999                    |
| 8        | Spring | 14,289                       | 47                 | .997                    |
| 9        | Summer | 14,281                       | 55                 | .996                    |
| 10       | Fall   | 14,062                       | 274                | .981                    |
| 11       | Spring | 7,472                        | 6,864              | .521                    |
| 12       | Summer | 7,039                        | 7,297              | .491                    |
| 13       | Fall   | 5,950                        | 8,386              | .415                    |
| 14       | Spring | 4,422                        | 9,914              | .308                    |
| 15       | Summer | 4,230                        | 10,106             | .295                    |
| 16       | Fall   | 4,032                        | 10,304             | .281                    |
| 17       | Spring | 3,438                        | 10,898             | .240                    |
| 18       | Summer | 3,388                        | 10,948             | .236                    |

### Hazard Function and Hazard Probabilities

The hazard function is a chronologically ordered plot of hazard over time that illustrates a risk profile for graduation. The hazard function for the present study is displayed in Figure 4.2, and the corresponding hazard probabilities of this plot are shown in Table 4.5. As shown in Table 4.5, the substantial risk periods for graduating occurred in the last 3 spring semesters (11, 14, and 17), with the greatest risk occurring in Semester 17. As hazard is dependent on the risk set, the spike in hazard during Semester 17 is attributed to the School of Pharmacy, which is a 6-year professional program. The majority of the students in the sample enrolled during the sixth year are Pharmacy students, and they make up 65 percent of the semester 17 graduates.



*Figure 4.2.* Sample Hazard Function for Time to Degree Completion. Hazard function shown is unconditional (i.e., not based on regressors).



Table 4.5  
*Sample Hazard Probabilities for Time to Degree Completion*

| Semester | Term   | Graduated | Risk Set<br>(Students<br>Enrolled) | Hazard<br>Probability |
|----------|--------|-----------|------------------------------------|-----------------------|
| 1        | Fall   | 0         | 14,336                             | .000                  |
| 2        | Spring | 0         | 14,140                             | .000                  |
| 3        | Summer | 0         | 3,893                              | .000                  |
| 4        | Fall   | 0         | 13,060                             | .000                  |
| 5        | Spring | 0         | 12,531                             | .000                  |
| 6        | Summer | 0         | 4,975                              | .000                  |
| 7        | Fall   | 5         | 12,073                             | .001                  |
| 8        | Spring | 42        | 11,780                             | .004                  |
| 9        | Summer | 8         | 4,924                              | .002                  |
| 10       | Fall   | 219       | 11,638                             | .019                  |
| 11       | Spring | 6,590     | 11,302                             | .583                  |
| 12       | Summer | 433       | 2,263                              | .191                  |
| 13       | Fall   | 1,089     | 4,124                              | .264                  |
| 14       | Spring | 1,528     | 2,973                              | .514                  |
| 15       | Summer | 192       | 1,027                              | .187                  |
| 16       | Fall   | 198       | 1,144                              | .173                  |
| 17       | Spring | 594       | 977                                | .608                  |
| 18       | Summer | 50        | 200                                | .250                  |

### **Model of Time to Degree Completion**

The model summary and individual parameter estimates of the logistic regression analysis on the person-period data are displayed in Table 4.6. To determine how well the model fit the sample, a likelihood ratio test was conducted between a model with no predictors other than the time effect (individual parameter estimates not shown) and the model with all the predictors. For the model of time to degree completion, the likelihood ratio test is used to test the significance of the coefficients in the model. The -2 Log Likelihood statistic (-2LL) has a chi-square distribution under the null hypothesis that all coefficients in the model are zero. The difference in fit between two nested models is assessed by looking at the change in -2LL, with degrees of freedom equal to the difference between the number of parameters in the two models. In the present study, the

likelihood ratio test is significant,  $\chi^2 (28, N=124,397) = 6,188.96, p < .001$ . The omnibus tests of model coefficients provide a test for the null hypothesis that all beta coefficients are equal to 0. As can be seen from the omnibus tests for the complete model in Table 4.6, the null hypothesis is rejected at the  $p < .05$  level with  $\chi^2 (45, N = 124) = 48,018.89$ . This means that at least one of the covariates in the complete model is significant. A review of the individual parameter estimates for variables in the equation reveals that several variables in the model are significant and allows us to examine the direction and magnitude of the effects of these covariates on the time to degree completion at the study institution.

In Table 4.6, the Wald statistic tests the statistical significance of each covariate's coefficient ( $\beta$ ) in the model. If the Wald statistic is significant, then we conclude that the coefficient differs from zero. The odds ratio, or  $\exp(\beta)$ , predicts the odds of graduating for each unit increase in the covariate. When  $\exp(\beta)$  is close to 1.0, a discrete variable has only a very small effect on graduation. If the value is less than 1.0, the direction of the effect is toward reducing the hazard rate. For continuous variables, the interpretation of the odds ratio is augmented by expressing the odds ratio as a percentage change in the risk of graduation (Allison, 1995). That is, the conversion of the odds ratio for continuous variables is  $100 (\exp(\beta) - 1)$ .

Table 4.6  
*Model Summary for Complete Time to Degree Completion Model*

| Omnibus Tests of Model Coefficients |            |    |      |  |  |  |
|-------------------------------------|------------|----|------|--|--|--|
|                                     | $\chi^2$   | df | Sig. |  |  |  |
| Block 1 – Time only                 | 41,829.923 | 17 | .000 |  |  |  |
| Block 2 – Predictors added          | 6,188.963  | 28 | .000 |  |  |  |
| Model – Time with predictors        | 48,018.886 | 45 | .000 |  |  |  |

| Variables in the Equation         |         |         |           |    |      |        |
|-----------------------------------|---------|---------|-----------|----|------|--------|
| Variable                          | B       | SE      | Wald      | df | Sig. | Exp(B) |
| Time                              |         |         | 7,898.287 | 17 | .000 |        |
| Semester 1                        | -20.766 | 306.451 | .005      | 1  | .946 | .000   |
| Semester 2                        | -20.809 | 319.191 | .004      | 1  | .948 | .000   |
| Semester 3                        | -21.519 | 603.190 | .001      | 1  | .972 | .000   |
| Semester 4                        | -20.927 | 331.468 | .004      | 1  | .950 | .000   |
| Semester 5                        | -21.020 | 340.413 | .004      | 1  | .951 | .000   |
| Semester 6                        | -21.684 | 539.261 | .002      | 1  | .968 | .000   |
| Semester 7                        | -7.732  | .487    | 251.648   | 1  | .000 | .000   |
| Semester 8                        | -5.621  | .247    | 516.708   | 1  | .000 | .004   |
| Semester 9                        | -7.039  | .401    | 308.913   | 1  | .000 | .001   |
| Semester 10                       | -3.946  | .205    | 371.796   | 1  | .000 | .019   |
| Semester 11                       | 1.047   | .192    | 29.903    | 1  | .000 | 2.850  |
| Semester 12                       | -1.230  | .193    | 40.521    | 1  | .000 | .292   |
| Semester 13                       | -.012   | .192    | .004      | 1  | .950 | .988   |
| Semester 14                       | 1.704   | .193    | 77.980    | 1  | .000 | 5.496  |
| Semester 15                       | -.249   | .205    | 1.475     | 1  | .225 | .780   |
| Semester 16                       | .066    | .206    | .104      | 1  | .748 | 1.069  |
| Semester 17                       | 2.624   | .202    | 168.140   | 1  | .000 | 13.792 |
| Cohort                            |         |         | 16.648    | 2  | .000 |        |
| 2002                              | .106    | .038    | 7.825     | 1  | .005 | 1.111  |
| 2003                              | .157    | .040    | 15.294    | 1  | .000 | 1.170  |
| Female                            | .248    | .034    | 51.935    | 1  | .000 | 1.281  |
| Race/Ethnicity                    |         |         | 17.382    | 4  | .002 |        |
| Asian Pacific Islander            | .059    | .041    | 2.121     | 1  | .145 | 1.061  |
| Black/African American            | -.129   | .060    | 4.551     | 1  | .033 | .879   |
| Hispanic                          | -.157   | .058    | 7.260     | 1  | .007 | .854   |
| Other                             | -.152   | .089    | 2.918     | 1  | .088 | .859   |
| Enrollment Age                    | -.055   | .031    | 3.218     | 1  | .073 | .946   |
| First Generation College Student  |         |         | .406      | 2  | .816 |        |
| Yes                               | -.030   | .047    | .401      | 1  | .526 | .971   |
| Unknown                           | -.002   | .072    | .000      | 1  | .982 | .998   |
| Residency                         |         |         | 21.194    | 2  | .000 |        |
| Out-of-State                      | .262    | .060    | 19.311    | 1  | .000 | 1.299  |
| International                     | -.191   | .152    | 3.064     | 1  | .080 | .827   |
| Transfer Credits at Matriculation | .011    | .006    | 1.580     | 1  | .209 | 1.011  |

|                                |        |      |           |   |      |       |
|--------------------------------|--------|------|-----------|---|------|-------|
| SAT Math                       | .000   | .000 | .644      | 1 | .422 | 1.000 |
| SAT Verbal                     | -.002  | .000 | 74.244    | 1 | .000 | .998  |
| Initial School of Enrollment   |        |      | 1,637.578 | 4 | .000 |       |
| Environ. & Biological Sciences | -.401  | .047 | 71.400    | 1 | .000 | .670  |
| Engineering                    | -.108  | .053 | 4.123     | 1 | .042 | .898  |
| Pharmacy                       | -3.086 | .077 | 1,603.397 | 1 | .000 | .046  |
| Mason Gross School of the Arts | -.128  | .112 | 1.301     | 1 | .254 | .880  |
| Housing Status – Off-Campus    | -.006  | .040 | .025      | 1 | .874 | .994  |
| At Least One Transfer          | -.637  | .051 | 157.317   | 1 | .000 | .529  |
| Credits Attempted              | -.056  | .005 | 116.754   | 1 | .000 | .946  |
| Percentage of Credits Earned   | .017   | .001 | 162.341   | 1 | .000 | 1.017 |
| Cumulative GPA                 | 1.922  | .040 | 2,273.566 | 1 | .000 | 6.837 |
| Grants                         | .000   | .000 | 144.614   | 1 | .000 | 1.000 |
| Scholarships                   | .000   | .000 | 5.157     | 1 | .023 | 1.000 |
| Student Loans                  | .000   | .000 | 168.270   | 1 | .000 | 1.000 |
| Work Study Earnings            | .000   | .000 | .009      | 1 | .924 | 1.000 |
| Constant                       | -4.880 | .639 | 58.422    | 1 | .000 | .008  |

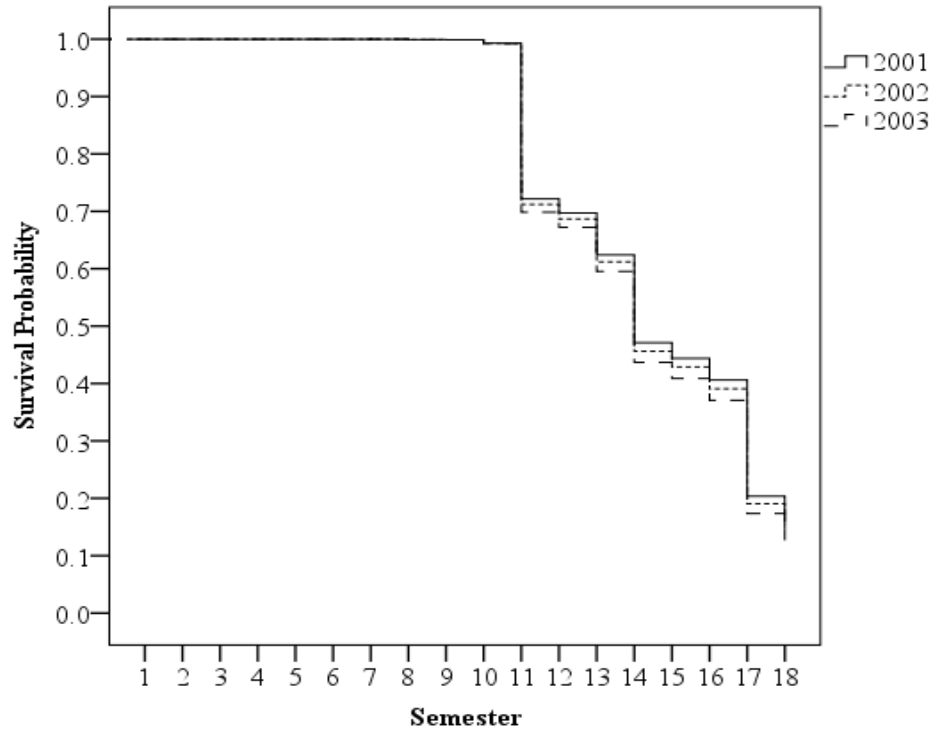
-2 Log Likelihood (time only model): 32,287.295

-2 Log Likelihood (time with predictors model): 26,098.331

Time is not only significant but the fundamental predictor of time to degree completion at the study institution,  $\chi^2 (17, N = 124,397) = 7,898.29$   $p < .001$ . As can be seen in Table 4.6, all time periods before the first event occurrence (Semester 7) are not significant. Although Semester 7 is significant, the magnitude of the effect is very small,  $\exp(\beta) = .000$ . Semesters 11, 14, and 17 are substantial periods for graduation,  $\exp(\beta) = 2.850, 5.496$ , and  $13.792$  respectively.

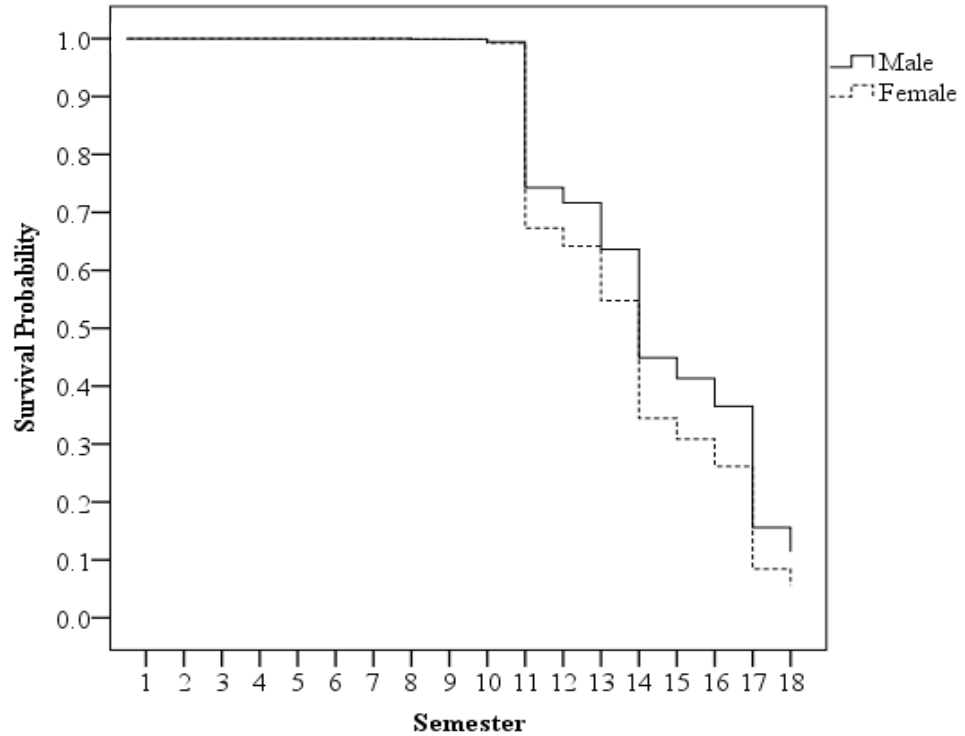
Cohort is significant,  $\chi^2 (2, N = 124,397) = 16.65$ ,  $p < .001$ . Based upon the odds ratio, the 2002 and 2003 cohorts are 1.110 and 1.170 times more likely to graduate than the 2001 cohort within the 6-year observation period. Although the magnitude of the effect is small, .11 and .17 respectively, the cohort effect suggests gradual but positive movement toward increasing graduation rates at the study institution. Figure 4.3 provides the sample survivor function for cohort. As can be seen in Figure 4.3, the 2002 and 2003 cohorts look very similar to the 2001 cohort with regards to time to graduation. However,

the slightly larger percentage of students in the 2002 and 2003 cohorts graduating each time period is enough to cause a small but significant cohort effect.



*Figure 4.3.* Sample Survivor Function for Time to Degree Completion by Cohort. Survivor function shown is unconditional (i.e., not based on regressors).

Gender, specifically being female, is significant,  $\chi^2(1, N = 124,397) = 51.94$ ,  $p < .001$ . Based upon the odds ratio, females are 1.281 times more likely than males to graduate. Figure 4.4 provides the sample survivor function for gender. As can be seen in Figure 4.4, females are more likely than males to graduate, and graduate sooner.



*Figure 4.4.* Sample Survivor Function for Time to Degree Completion by Gender. Survivor function shown is unconditional (i.e., not based on regressors).

Race/ethnicity is significant,  $\chi^2(4, 124,397) = 17.382, p = .002$ . Since race/ethnicity is a categorical variable with more than two groups, the significance of the overall Wald statistic suggests that at least one of the effect coefficients differs from zero. According to Table 4.6, Black/African American and Hispanic are significant racial/ethnic groups,  $\chi^2(1, 124,397) = 4.551, p = .033$  and  $\chi^2(1, 124,397) = 7.260, p = .007$  respectively. Based upon the odds ratio, Black/African Americans .879 and Hispanics .854 times less likely to graduate than their White peers. Figure 4.5 provides the sample survivor function for race/ethnicity. For visual purposes, the Asian/Pacific Islander and Other racial/ethnic groups are not shown in the sample survivor function for race/ethnicity. As can be seen in Figure 4.5, White students are more likely to graduate than their Black/African American and Hispanic peers.

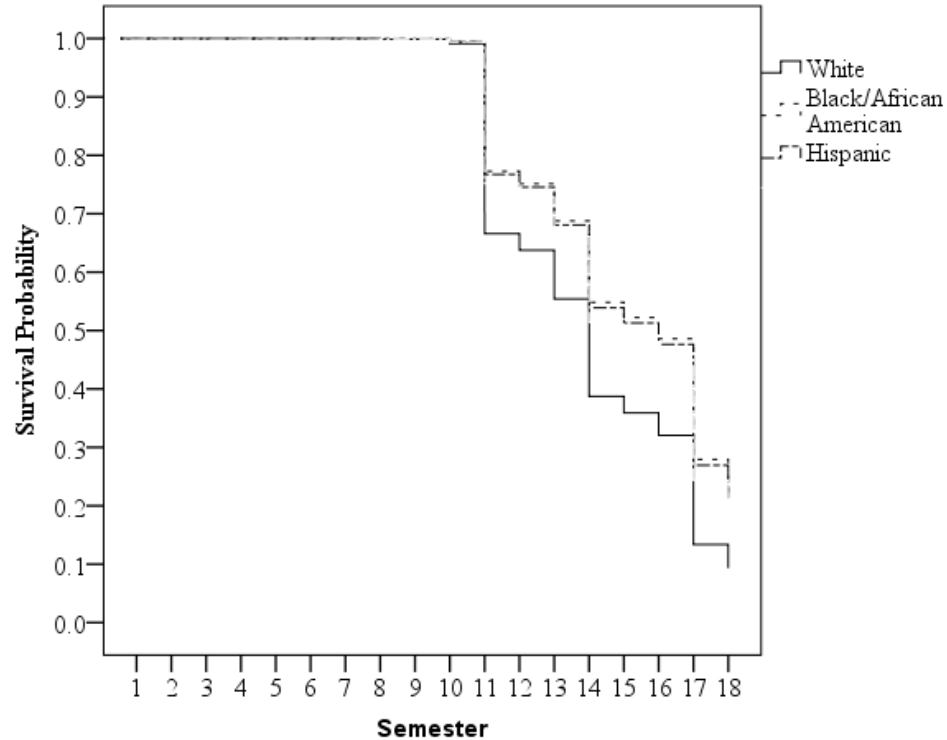


Figure 4.5. Sample Survivor Function for Time to Degree Completion by Race/Ethnicity. Survivor function shown is unconditional (i.e., not based on regressors).

The overall residency variable is significant,  $\chi^2(2, 124,397) = 21.194, p < .001$ .

The effect coefficient for out-of-state students is significant,  $\chi^2(1, 124,397) = 19.311, p < .001$ . Figure 4.6 provides the sample survivor function for residency. As can be seen in Figure 4.6, out-of-state students graduate earlier than their New Jersey resident and international peers. Charged almost double the tuition than in-state residents, there is a strong financial incentive for out-of-state students to finish their undergraduate degree within 4 years. Although international students are also charged the same tuition differential as out-of-state students, there is no financial incentive for international students to graduate sooner as these students generally come from very wealthy families or are sponsored by their government.

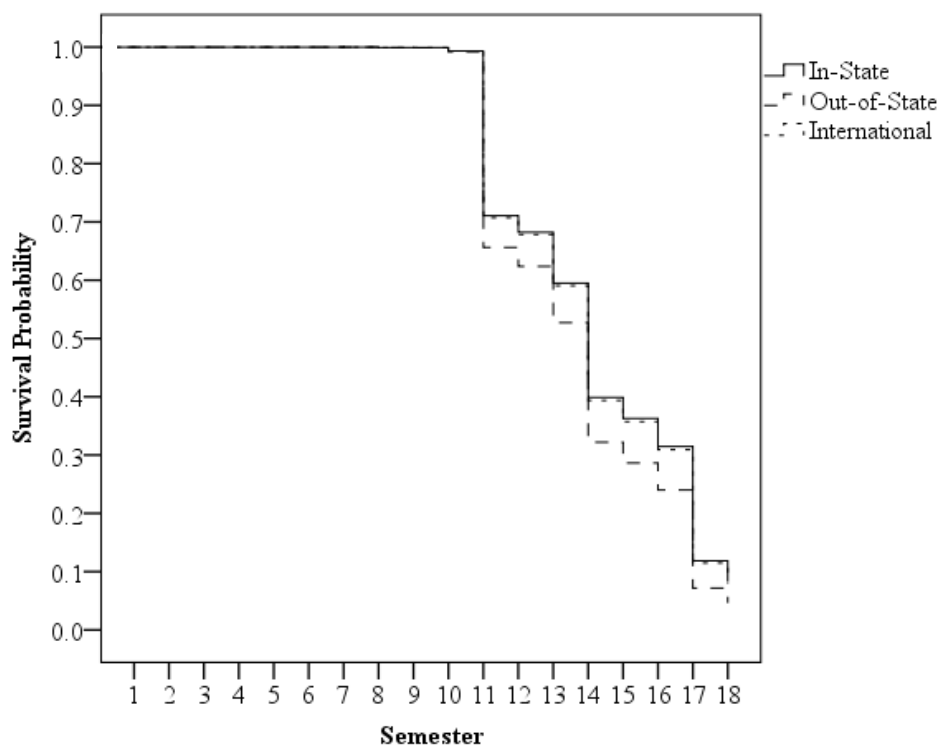


Figure 4.6. Sample Survivor Function for Time to Degree Completion by Residency. Survivor function shown is unconditional (i.e., not based on regressors).

Although SAT Math is not significant, SAT Verbal is significant,  $\chi^2(1, 124,397) = 74.244, p < .001$ . Since SAT scores are continuous, the expression of the odds ratio as a percentage in the risk of graduation is -.2 percent for SAT Verbal. For every unit increase in the SAT Verbal score there is a .2 percent decrease in the hazard for graduation.

Although significant, the effect of SAT Verbal is very small. For example, a 100 point increase in a student's SAT Verbal score would only result in a 2 percent decrease in the hazard of graduating.

The remaining pre-enrollment variables are not significant, which includes enrollment age, transfer credits at time of matriculation, and first generation college student status. Although several studies have found a significant effect for first generation college students, this is not substantiated at the study institution. Comprising almost 14 percent of the sample, first generation college students seem to fair well at the study



institution compared to their peers who have at least one parent with a 4-year college degree. As Rutgers has a deep commitment to access and opportunity, the data suggests the university is serving its first generation college student population quite well in regards to removing the educational and financial barriers to a postsecondary education.

According to Table 4.6, all of the enrollment variables except housing status are significant in the model. The overall initial school of enrollment variable is significant,  $\chi^2(4, 124,397) = 1,637.578, p < .001$ . The effect coefficients for the School of Environmental and Biological Sciences (SEBS), School of Engineering, and Ernest Mario School of Pharmacy are significant. Figure 4.7 provides the sample survivor function for initial school of enrollment. As can be seen in Figure 4.7 students enrolled in the SEBS, Engineering, and Pharmacy take longer to graduate than students enrolled in the School of Arts and Sciences (SAS) and Mason Gross School of the Arts (MGSA). One noticeable feature in Figure 4.7 is the dramatic drop in the survivor function for Pharmacy during the sixth year. As Pharmacy is a 6-year professional program, these results are expected.

School-to-school transfers, specifically having at least one, is significant,  $\chi^2(1, 124,397) = 157.317, p < .001$ . According to the odds ratio, students who transferred at least once between the academic schools were .529 times less likely to graduate than their peers who remained in their initial school of enrollment. Figure 4.8 provides the sample survivor function for school-to-school transfers. As can be seen in Figure 4.8, students who remained in their initial school of enrollment were more likely to graduate and to graduate sooner than their peers who transferred at least once to another school within the institution.

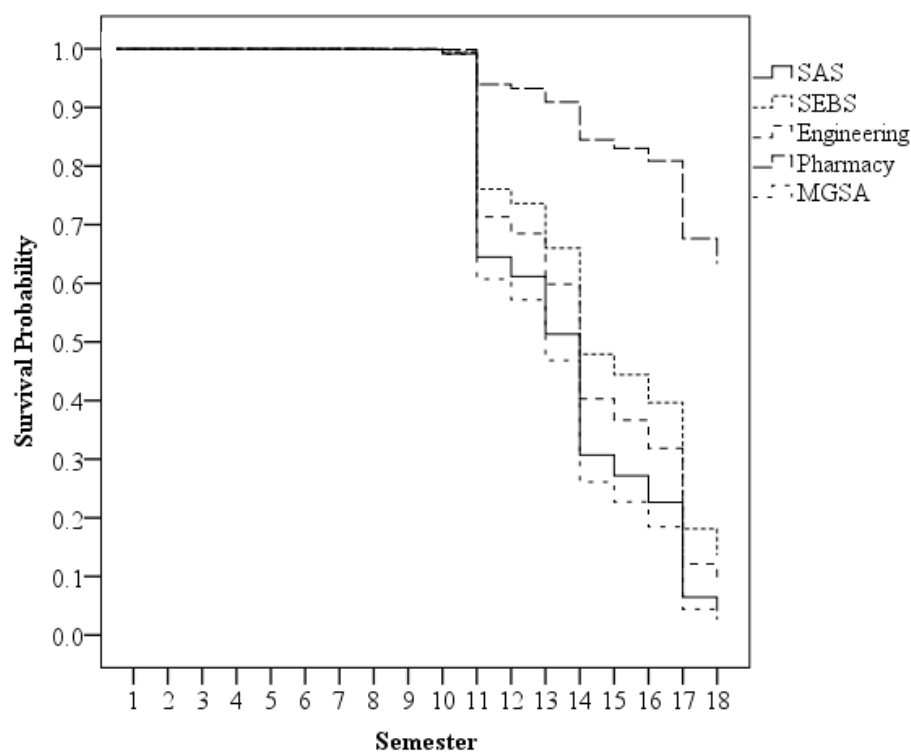


Figure 4.7. Sample Survivor Function for Time to Degree Completion by Initial School of Enrollment. Survivor function shown is unconditional (i.e., not based on regressors).

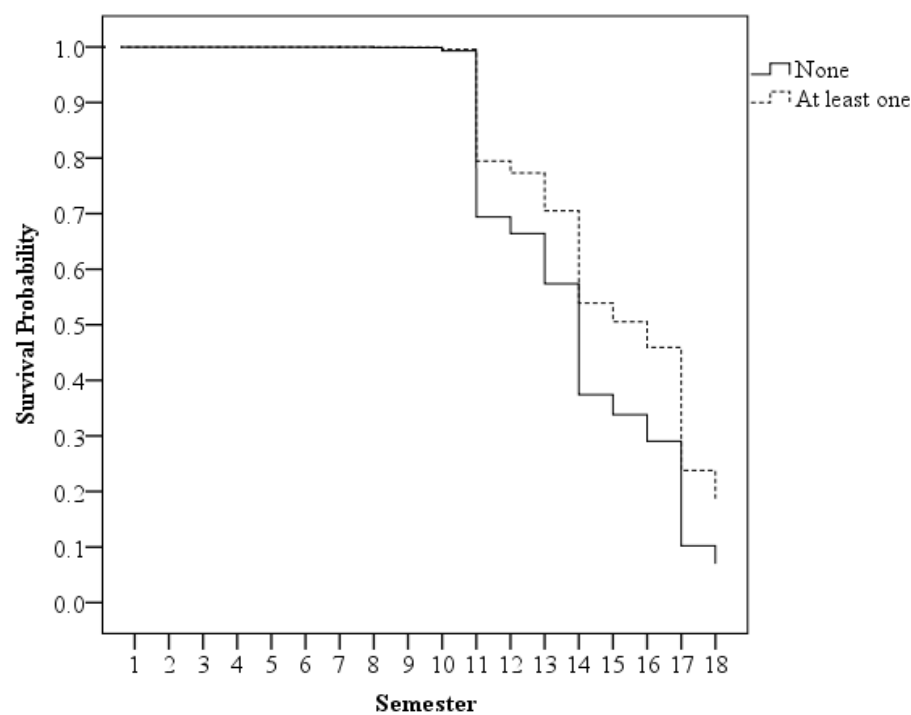


Figure 4.8. Sample Survivor Function for Time to Degree Completion by School-to-School Transfers. Survivor function shown is unconditional (i.e., not based on regressors).

Credits attempted and the percentage of credits earned are significant variables in the model,  $\chi^2(1, 124,397) = 116.754, p < .001$  and  $\chi^2(1, 124,397) = 162.341, p < .001$  respectively. A unit increase in the number of credits attempted results in a 5.4 percent decrease in the hazard of graduating, whereas a one unit increase in the percentage of credits earned results in a 1.7 percent increase in the hazard of graduating. As graduation requirements require a student to complete a specified number of credit hours (usually 120), it makes sense that the number of credits a student attempts each semester and the percentage of the credits they earn have a significant effect on graduating.

Cumulative GPA is also significant,  $\chi^2(1, 124,397) = 2,273.566, p < .001$ . Based on the interpretation of the odds ratio for continuous variables, a one unit increase in cumulative GPA results in a 5 percent increase in the hazard of graduating. As can be seen in Table 4.6, the covariate cumulative GPA provides the single largest contribution to the time to degree completion model of all the enrollment variables.

According to Table 4.6, all of the financial aid variables except work study earnings are significant in the model. Grants, scholarships, and student loans are significant predictors in time to degree completion,  $\chi^2(1, 124,397) = 144.61, p < .001$ ,  $\chi^2(1, 124,397) = 5.16, p = .023$ , and  $\chi^2(1, 124,397) = 168.27, p < .001$  respectively. As the beta coefficients and odds ratios for these variables are 0.00 and 1.00 respectively, the effects of grants, scholarships, and student loans on graduation appear to be indirect.

## CHAPTER V. SUMMARY AND CONCLUSIONS

In an era of increasing demand for college, declining fiscal resources, and the rising costs of undergraduate education, the study of college student retention and graduation, especially timely graduation, has been of great importance to the higher education community for decades. Research has shown that there are societal, individual, and institutional costs associated with students who leave college before degree completion (Pascarella & Terenzini, 1991). A college educated workforce benefits society by boosting economic development, increasing tax revenues, increasing civic engagement, and decreasing government spending on healthcare, criminal justice, and welfare. Postsecondary degree attainment benefits individuals through increased employment opportunities, higher earnings, employer-provided health and pension benefits, and improved quality of life. Graduation, especially timely graduation, is essential to institutions of higher education in order to manage enrollment numbers for economic stability as funding for higher education has declined dramatically over the last twenty years. Furthermore, retention and graduation rates are increasingly dominating higher education policy debates in the United States as national interest in performance-based accountability grows.

The purpose of the present study was to develop a greater understanding of graduation behavior, particularly at the study institution. Recent studies on college student retention and degree completion advance event history analysis as a more appropriate technique for studying a longitudinal process such as graduation behavior. Although the application of event history modeling to educational research is relatively new, researchers have long recognized the shortcomings of using cross-sectional data and

traditional static methodologies when studying longitudinal processes in higher education. The application of event history analysis to the study of college student retention and degree completion provides researchers and administrators with much more information about the temporal nature of factors over time.

The present study used event history analysis to model time to degree completion at Rutgers University for the fall 2001, 2002, and 2003 entering cohorts of first-time, full-time, degree-seeking undergraduates enrolled at the flagship campus. Consistent with other studies employing event history analysis to student retention and degree completion, adding a time dimension improves our understanding of event occurrence. Although many of the pre-enrollment characteristics were found to be statistically significant, their effect on graduation was generally very small. This finding is consistent with previous studies that noted the diminishing effect of pre-enrollment variables when enrollment and financial aid variables were included in time to degree completion models. An interesting finding is that first generation college students were not found to be statistically different in regards to graduation behavior as their peers who have at least one parent with a 4-year college degree as exhibited in other studies. This finding suggests that the Rutgers serves the first generation college student population quite well, and should be examined further to identify practices and policies that diminish the academic, social, and financial barriers this group typically experience during postsecondary education.

Consistent with previous studies, the present study provides support for the strong relationship between the longitudinal effects of academic performance while in college (as measured by cumulative GPA) and graduation. As cumulative GPA increases, the

likelihood of graduation increases. Intuitively this makes sense as students who are performing poorly do not experience the reinforcing effect of good grades and depending on the degree of their poor performance may also be academically dismissed from the institution for not making satisfactory academic progress. Furthermore, recent changes to the federal satisfactory academic progress regulations within the Office of Financial Aid may result in the suspension of federal aid for poor academic performance until the student regains eligibility. The recent changes in the satisfactory academic progress policies at Rutgers will impact these variables for those students who do not meet the new regulations. In particular, the new policy requires financial aid recipients to earn a cumulative completion rate of 50 percent and have a cumulative GPA of 1.50 for the first 30 credits, 60 percent and 1.80 for 31 to 59 credits, 70 percent and 2.00 for 60 to 89 credits, and 75 percent and 2.00 for 90 or more degree credits. If a student fails to meet at least one of the two standards, then the student will not be eligible for federal aid until he or she has met both standards. This can have a significant impact on student retention, and ultimately graduation, for low-income and middle-class students who do not have the financial means to finance their college education.

For those students who rely heavily on federal aid to finance their college education, poor academic performance may cause a systematic institutional departure before being formally dismissed by the institution for students who wish to continue but do not have the financial means to pay out of pocket or borrow the entire cost of attendance through high interest student loans. As Rutgers remains committed to access and opportunity, university administrators should make an active effort to inform all current and incoming students about the changes in the satisfactory academic progress

guidelines and how these changes can potentially impact their eligibility for federal financial aid. University administrators should also communicate to students the importance of seeking out help early in the semester when experiencing academic difficulties and talking to an academic advisor before dropping or withdrawing from courses as cumulative GPA and the percentage of credits completed are the measures by which financial aid eligibility will be determined.

Although the effect of the financial aid variables – particularly grants, scholarships, and student loans – do not have a direct impact on degree completion at Rutgers, it is important to include these aid-related variables in the model for time to degree completion because previous studies suggest financial aid is helpful for reducing student departure. In order to understand the impact the changing satisfactory academic policy will have on student retention, and ultimately graduation, the types of financial aid should be disaggregated further (e.g., breakdown the loans variable by unsubsidized, subsidized, and private loans) and combined with the remaining data in the present study in order to develop a model of student departure. Rutgers administrators would then be able to use the model of student departure to simulate the implication of the new satisfactory academic policies on student retention, particularly for those students who are on the border. That is, they met the old guidelines but under the proposed changes would not meet the guidelines.

Rutgers should also take a more proactive approach to student retention by developing and implementing an early warning system that would provide a systematic way for academic and residence life support staff to identify students who could benefit from someone reaching out and helping them connect with the appropriate resources at

the university. For example, students experiencing academic difficulties might be referred to one of the Rutgers Learning Centers. Located on each of the New Brunswick campuses, the Rutgers Learning Centers offer a comprehensive range of free academic support programs to promote student achievement. In addition to offering both individual and group tutoring services, the Rutgers Learning Centers employ a number of Academic Coaches to assist students with better time management skills, reading and test-preparation strategies, and public speaking techniques. The earlier Rutgers can identify and connect students who are struggling with the academic and/or social demands of college with the appropriate university resources, the more likely the issue can be resolved before it manifests into a much larger problem. For some students, this could mean having to spend additional semesters in college in order to meet all degree requirements. For others, what started out as a minor issue could contribute to or become a major driving force of a permanent departure from the institution either voluntarily or involuntarily.

### **Policy Implication**

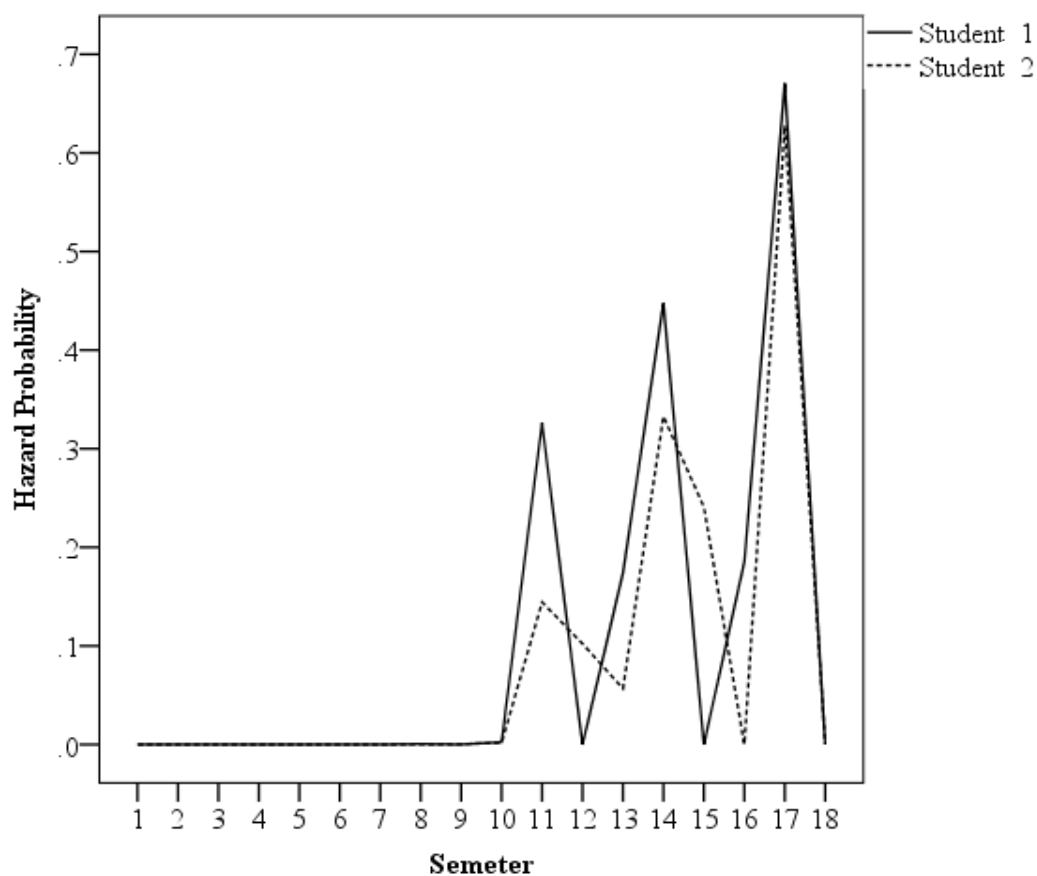
The event history model in the present study can be used to illustrate graduation behavior to provide university administrators at Rutgers with information to identify students who could benefit from targeted interventions. For example, let us assume two students who enrolled at Rutgers with different characteristics. Student 1 is an 18-year-old, White male from New Jersey who enrolled in the School of Environmental and Biological Sciences in fall 2001. He comes from a family with at least one parent with a 4-year college degree, and to save money commuted to campus his entire time at Rutgers. He scored about average on the SAT verbal section with a 580 and slightly below



average on the SAT math section with a 530. Not qualifying for a scholarship, his financial aid package consisted of grants and student loans. He does not transfer out of the School of Environmental and Biological Sciences, and only takes classes in the fall and spring semesters. Student 2 is an 18-year-old, Black/African American male from New Jersey who enrolled in the School of Arts and Sciences in fall 2001. From the data reported by the student, university administrators do not know if the student is a first generation college student. His SAT scores are lower than average with a 430 on the math section and 420 on the verbal section. Not qualifying for a merit scholarships, his financial aid package consists of grants, student loans, and work study. Living on campus, he takes classes every semester, including summer sessions, except during semester 16. Both students enroll in a credit load of 12-15 credits during the fall and spring semesters, and Student 2 enrolls in 4-6 credits in the summer semesters. Student 1 tends to earn a larger percentage of the credits he attempted, and has a cumulative GPA as low as 1.900 and as high as 2.689. Student 2 does slightly worse academically, with a cumulative GPA that as low as 1.400 and as high as 2.546.

Using the results from the model (Table 4.6) in this study, we can graphically display the longitudinal effects of graduation risks for Students 1 and 2 (see Figure 5.1). Overall, the two students have different hazards for graduating during the time periods based on their characteristics. Although both students have similar hazard at the end of the 6-year observation period, Student 1 has a higher risk of graduating at the 4-year (Semester 11) and 5-year (Semester 14) mark. The illustrated longitudinal graduation behavior can enable university administrators to identify subgroups of students that could benefit from targeted interventions based on their characteristics. To reduce the time to

degree at the study institution, university administrators could use the simulated students illustrated in Figure 5.1 to identify Student 2 as a candidate for special programming in his first three years to help increase his likelihood for graduating in Semesters 11 and 14. The application of the time-specific graduation risks of students, as measured by hazard, would help university administrators to strengthen their outreach to subgroups of students who could benefit from strategic interventions.



*Figure 5.1. Simulated Longitudinal Graduation Hazard*

### **Limitations and Implications for Future Research**

The present study has several limitations that merit discussion. First, analyses conducted on a single institution are often criticized for the lack of generalizability of the results beyond the study institution. When research is done across multiple institutions,

the degree to which findings are applicable to any one institution is diminished by aggregation and the utility for effecting change at the campus level is diminished because the aggregate across multiple institutions does not fit any one institution well. With the disparities in graduation rates from one institution to another and the push for colleges and universities to do a better job of graduating their students, research on the time to degree completion should be conducted at the institutional level where the information has the greatest likelihood of effecting change in graduation behaviors. Furthermore, as New Jersey policy makers place more emphasis on time to degree completion to control the growing costs of higher education in the state, event history models generated at the institutional level will provide important empirical evidence about why students are taking more than four years to graduate. As the largest postsecondary institution in the state and the designated state university of New Jersey, Rutgers has the capacity to take the lead in reducing the time to degree completion for New Jersey's students.

Second, the present study does not incorporate the many changes to the academic engagement programming that came out of the Transforming Undergraduate Education (TUE) initiative at Rutgers. In addition to the proposed reorganization of the liberal arts colleges into a single liberal arts school, the TUE initiative included a comprehensive series of recommendations covering all aspects of the undergraduate experience. As a direct result of the task force's final recommendations, Rutgers' administration committed a great deal of time, energy, and resources to improve the first-year experience. For example, Rutgers established Byrne First-Year Seminars, which are elective, one-credit courses designed to engage students academically and socially with faculty members during their first year in college in a small group setting. Rutgers also

greatly expanded the number and variety of learning communities it offered to first-year students, upperclassmen, and transfer students. Although these programs will undoubtedly vary across colleges and universities, the growing popularity of academic engagement programming in higher education and their integration into the undergraduate experience provide an opportunity for researchers and university administrators to examine the longitudinal effects such programs have on their institution's student retention and graduation behaviors. However, in order for researchers to include this type of information in event history models, colleges and universities need to collect and store participation data within a student's record in the university's administrative student information system so that six years later the data is available to examine their longitudinal effects over time when modeling time to degree completion.

Third, as with all statistical models, there is always a potential for bias in the parameter estimates. Unobserved heterogeneity can cause parameter estimation bias when using logistic regression on the students' event history data if important explanatory variables are not included in the model. In particular, the presence of unobserved heterogeneity may seriously bias the effects of time-varying covariates. In the case of the present study, some of the unobserved heterogeneity is locked up in endogeneity. That is, it is determined within the system. For example, transferring between academic schools is modeled as a time-independent variable. Although the time at which the student makes the transfer (e.g., first year vs. third year) certainly has an effect on the time to degree completion, the event history model is limited by its inability to tease out this effect because the data cannot be stretched that far. Other times, the

unobserved heterogeneity is endogenous, and cannot be controlled in the model. For example, continuously enrolling between the fall and spring semesters certainly has an effect on graduation as students who experience a stopout have an increased likelihood of subsequent stopouts, which can eventually lead to a permanent departure from an institution. However, continuous enrollment is part of the longitudinal process being explained and it would be inappropriate to include such a variable in the model. Future research should examine the possibility of bias in the present model caused by unobserved heterogeneity and ways to empirically address unobserved heterogeneity through the use of more advanced statistical software (e.g., STATA).

## **Conclusion**

As graduation rates are increasingly tied to institutional performance and resource allocation, college and university administrators need to have a better understanding of graduation behaviors at their institution to develop effective interventions that help students not only complete their degree, but complete their degree in a timely manner. As advocated in recent research on college student retention and graduation, this study used event history analysis to examine the temporal nature of graduation behavior. By focusing on the institutional level, this study produced fine-grained results that are more useful to university administrators at the study institution than those produced by analyses conducted at the national level. Although the present study was driven by a need to better inform the university administrators at Rutgers University, this study also benefits the higher education system by advancing event history analysis as a methodology for studying graduation behaviors.

## REFERENCES

- Adelman, C. (1999). Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment. Jessup, MD: U.S. Department of Education.
- Alexander, F. K. (2000). The changing face of accountability: Monitoring and assessing institutional performance in higher education. *Journal of Higher Education*, 71, 411-431.
- Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. *American Sociological Association*, 13, 61-98.
- Allison, P. D. (1984). *Event history analysis: Regression for longitudinal event data*. Newbury Park, CA: SAGE Publications.
- Allison, P. D. (1995). *Survival Analysis Using the SAS System: A practical guide*. Cary, NC: SAS Institute Inc.
- Altbach, P. G., Reisberg, L., & Rumbley, L. E. (2010). Tracking a global academic revolution. *Change: The Magazine of Higher Learning*, 42(2), 30-39.
- American Association of State Colleges and Universities. (2006). *Graduation rates and student success: Squaring means and ends*. Washington, DC: AASCU.
- Barger, B., & Hall, E. (1965). Time of dropout as a variable in the study of college attrition. *College and University*, 41, 84-88.
- Baum, S., Ma, J., & Payea, K. (2010). *Education Pays 2010: The Benefits of Higher Education for Individuals and Society*. Washington, DC: College Board. Retrieved from [http://trends.collegeboard.org/downloads/Education\\_Pays\\_2010.pdf](http://trends.collegeboard.org/downloads/Education_Pays_2010.pdf).
- Baum, S., Payea, K., & Cardenas-Elliott, D. (2010). Trends in student aid. *Trends in Higher Education Series, 2010*. Washington, DC: College Board. Retrieved from [http://trends.collegeboard.org/downloads/Student\\_Aid\\_2010.pdf](http://trends.collegeboard.org/downloads/Student_Aid_2010.pdf).
- Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education*, 12, 155-187.
- Bean, J. P. (1982). Student attrition, intentions, and confidence: Interaction effects in a path model. *Research in Higher Education*, 17, 291-319.
- Bean, J. P. (1983). The application of a model of turnover in work organizations to the student attrition process. *Review of Higher Education*, 6, 129-148.

- Bean, J. P. (1985). Interaction effects based on class level in an exploratory model of college student dropout syndrome. *American Educational Research Journal*, 22, 35-64.
- Belfield, C. R., & Levin, H. M. (2007). *The price we pay: The economic and social costs of inadequate education*. Brookings Institution: Washington, DC.
- Berger, J. B., & Lyon, S. C. (2005). Past to present: A historical look at retention. In A. Siedman (Ed.), *College Student Retention: Formula for Student Success* (pp. 1-29). Westport, CT: Praeger Publishers.
- Braxton, J. M. (2000). Introduction: Reworking the student departure puzzle. In J. M. Braxton (Ed.), *Reworking the student departure puzzle* (pp. 1-8). Nashville, TN: Vanderbilt University Press.
- Braxton, J. M., & Hirschy, A. S. (2005). Theoretical developments in the study of college student departure. In A. Seidman (Ed.), *College Student Retention: Formula for Student Success* (pp. 61-87). Westport, CT: Praeger Publishers.
- Braxton, J. M., Hirschy, A. S., & McClendon, S.A. (2004). *Understanding and reducing college student departure*. ASHE-ERIC Higher Education Report No. 3. San Francisco, CA: Jossey-Bass.
- Braxton, J. M., & Lee, S. D. (2005). Theoretical developments in the study of college student departure. In A. Seidman, *College Student Retention: Formula for Student Success* (pp. 61-87). Westport, CT: Praeger Publishers.
- Braxton, J. M., & Lien, L. A. (2000). The viability of academic integration as a central construct in Tinto's interactionalist theory of college student departure. In J. Braxton (Ed.), *Reworking the Student Departure Puzzle* (pp. 11-28). Nashville, TN: Vanderbilt University Press.
- Braxton, J. M., Sullivan, A. S., & Johnson, R. M. (1997). Appraising Tinto's theory of college student departure. In J. Smart (Ed.), *Higher Education: Handbook of Theory and Research, Vol. 12* (pp. 107-164). Bronx, NY: Agathon Press.
- Bureau of Labor Statistics. (2010a). *Employment status by educational attainment, sex, race, and Hispanic ethnicity*. Retrieved from <http://www.bls.gov/cps/demographics.htm#education>.
- Bureau of Labor Statistics. (2010b). *Volunteering in the United States – 2009*. Retrieved from <http://www.bls.gov/news.release/pdf/volun.pdf>.
- Burke, J. C. (2005). *Achieving accountability in higher education: Balancing public, academic, and market demand*. San Francisco, CA: Jossey-Bass.

- Cabrera, A. F., Castaneda, M. B., Nora, A., & Hengstler, D. (1992). The convergence between two theories of college persistence. *Journal of Higher Education*, 63, 143-164.
- Cabrera, A. F., Nora, A., & Castaneda, M. B. (1993). College persistence: Structural equations modeling test of an integrated model of student retention. *Journal of Higher Education*, 64, 123-139.
- Carnevale, A. P., Smith, N., & Strohl, J. (2010). *Help wanted: Projections of jobs and education requirements through 2018*. Washington, DC: Georgetown University Center on Education and the Workforce. Retrieved from <http://cew.georgetown.edu/jobs2018>.
- Chen, R., & DesJardins, S. L. (2008). Exploring the effects of financial aid on the gap in student dropout risks by income level. *Research in Higher Education*, 49, 1-18.
- Cheng, D., & Reed, M. (2010). *Student debt and the class of 2009*. Oakland, CA: Project on Student Debt. Retrieved from <http://projectonstudentdebt.org/files/pub/classof2009.pdf>.
- Cook, B., & Pullaro, N. (2010). *College graduation rates: Behind the numbers*. Washington, DC: American Council on Education. Retrieved from [http://www.acenet.edu/links/pdfs/cpa/2010\\_09\\_30\\_collegegraduationrates.html](http://www.acenet.edu/links/pdfs/cpa/2010_09_30_collegegraduationrates.html).
- DesJardins, S. L. (2003). Event history methods: Conceptual issues and an application to student departure from college. In J. Smart (Ed.), *Higher Education: Handbook of Theory and Research*, Vol. 18 (pp. 421-472). London: Kluwer Academic Publishers.
- DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (1999). An event history model of student departure. *Economics of Education Review*, 18, 375-390.
- DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (2002a). A temporal investigation of factors related to timely degree completion. *Journal of Higher Education*, 73, 555-581.
- DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (2002b). Simulating the longitudinal effects of changes in financial aid on student departure from college. *Journal of Human Resources*, 37, 653-679.
- DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (2006). The effects of interrupted enrollment on graduation from college: Racial, income, and ability differences. *Economics of Education Review*, 25, 575-590.
- DesJardins, S. L., McCall, B. P., Ahlburg, D. A., & Moye, M. J. (2002). Adding a timing light to the "tool box". *Research in Higher Education*, 43, 83-114.



- Desrochers, D. M., Lenihan, C. M., & Wellman, J. V. (2010). *Trends in college spending 1998-2008*. Washington, DC: Delta Cost Project. Retrieved from [http://www.deltacostproject.org/resources/pdf/trends\\_in\\_spending-report.pdf](http://www.deltacostproject.org/resources/pdf/trends_in_spending-report.pdf).
- Duderstadt, J. J., & Womack, F. W. (2003). *The future of the public university in America: Beyond the crossroads*. Baltimore, MD: The John Hopkins University Press.
- Durkheim, E. (1951). *Suicide: A study of sociology*. (J. A. Spaulding & G. Simpson, Trans.). Glencoe, IL: Free Press. (Original work published 1897).
- Eckland, B. K. (1964). College dropouts who came back. *Harvard Educational Review*, 34, 402-420.
- Gekowski, N., & Schwartz, S. (1961). Student mortality and related factors. *Journal of Educational Research*, 54, 192-194.
- Gladieux, L., & Perna, L. (2005). *Borrowers who drop out: A neglected aspect of the college student loan trend*. San Jose, CA: The National Center for Public Policy and Higher Education. Retrieved from <http://highereducation.org/reports/borrowing/borrowers.pdf>.
- Hossler, D. (1984). *Enrollment Management: An Integrated Approach*. New York: College Entrance Examination Board.
- Hurley, D. J. (2009). *Considerations for state colleges and universities in a post-recession America* [White paper]. Retrieved from <http://www.aascu.org/media/publications/09hurleywhpaper.pdf>.
- Ishitani, T. T., & DesJardins, S. L. (2002). A longitudinal investigation of dropout from colleges in the United States. *Journal of College Student Retention*, 4, 173-201.
- Ishitani, T. T. (2003). A longitudinal approach to assessing attrition behavior among first-generation students: Time-varying effects of pre-college characteristics. *Research in Higher Education*, 44, 433-449.
- Ishitani, T. T. (2006). Studying attrition and degree completion behavior among first-generation college students in the United States. *Journal of Higher Education*, 77, 861-885.
- Jex, F. B., & Merrill, R. M. (1962). A study in persistence. *Personnel and Guidance Journal*, 40, 762-769.
- Johansson, C. B., & Rossmann, J. E. (1973). Persistence at a liberal arts college: A replicated, five-year longitudinal study. *Journal of Counseling Psychology*, 20, 1-9.

- Kennedy, P., & Sheckley, B. (1999). *Attrition and persistence in higher education: A review of the literature*. College Park, MD: Institute for Research on Adults in Higher Education, University of Maryland University College.
- Lumina Foundation for Education. (2010). *A stronger nation through higher education: How and why Americans must achieve a "big goal" for college attainment*. Retrieved from [http://www.luminafoundation.org/publications/A\\_stronger\\_nation.pdf](http://www.luminafoundation.org/publications/A_stronger_nation.pdf).
- Marks, E. (1967). Student perceptions of college persistence, and their intellectual, personality, and performance correlates. *Journal of Educational Psychology*, 58, 210-221.
- Murtaugh, P. A., Burns, L. D., & Schuster, J. (1999). Predicting the retention of university students. *Research in Higher Education*, 40, 355-371.
- Nora, A., & Cabrera, A. F. (1996). The role of perceptions of prejudice and discrimination on the adjustment of minority students to college. *Journal of Higher Education*, 67, 119-148.
- Obama, B. (2010, August). Remarks by the president on higher education and the economy at the University of Texas at Austin. Retrieved from, <http://www.whitehouse.gov/the-press-office/2010/08/09/remarks-president-higher-education-and-economy-university-texas-austin>.
- Panos, R. J., & Astin, A. W. (1968). Attrition among college students. *American Educational Research Journal*, 5, 57-72.
- Pantages, T. J., & Creedon, C. F. (1978). Studies of college attrition: 1950-1975. *Review of Educational Research*, 48, 49-101.
- Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco, CA: Jossey-Bass.
- Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research* (Vol. 2). San Francisco, CA: Jossey-Bass.
- Price, J. L. (1977). *The Study of Turnover*. Ames, IA: Iowa State University Press.
- Price, J. L., & Mueller, C. W. (1981). A causal model of turnover for nurses. *Academy of Management Journal*, 24, 543-565.
- Rossmann, J. E., & Kirk, B. A. (1970). Factors related to persistence and withdrawal among university students. *Journal of Counseling Psychology*, 17, 56-62.

- Schneider, M. (2010). *Finishing the first lap: The cost of first-year student attrition in America's four-year colleges and universities*. Washington, D.C.: American Institutes for Research.
- Singer, J. D., & Willett, J. B. (1991). Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psychological Bulletin*, 110, 268-290.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.
- Spady, W. (1970). Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange*, 1, 64-85.
- Spady, W. (1971). Dropout from higher education: Toward an empirical model. *Interchange*, 2, 38-62.
- Starr, A., Betz, E. L., & Menne, J. (1972). Differences in college student satisfaction: Academic dropout, non-academic dropouts, and non-dropouts. *Journal of Counseling Psychology*, 19, 318-322.
- Summerskill, J. (1962). Dropout from college. In N. Sanford (Ed.), *The American College* (pp. 627-655). New York: Wiley.
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45, 89-125.
- Tinto, V. (1982). Defining dropout: A matter of perspective. In E. Pascarella (Ed.), *New Directions for Institutional Research: Studying Student Retention*, No. 36 (pp. 3-15). San Francisco: Jossey-Bass.
- Tinto, V. (1987). *Leaving college: Rethinking the causes and cures of student attrition*. Chicago: The University of Chicago Press.
- Tinto, V. (1988). Stages of student departure: Reflections on the longitudinal character of student leaving. *Journal of Higher Education*, 59, 438-455.
- Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: The University of Chicago Press.
- Transforming Undergraduate Education. (2005). *Report of the task force on undergraduate education*. Retrieved from [http://urwebsrv.rutgers.edu/transform\\_ru/pdf/fullreport.pdf](http://urwebsrv.rutgers.edu/transform_ru/pdf/fullreport.pdf).

- U.S. Census Bureau. (2010a). *Current population survey, 2009 annual social and \ economic supplement*. Retrieved from [http://www.census.gov/hhes/www/cpstables/032010/perinc/new03\\_010.htm](http://www.census.gov/hhes/www/cpstables/032010/perinc/new03_010.htm).
- U.S. Census Bureau. (2010b). *Voting and registration in the election of November 2008*. Retrieved from <http://www.census.gov/prod/2010pubs/p20-562.pdf>.
- Willett, J. B., & Singer, J. D. (1991). From whether to when: New methods for studying Student dropout and teacher attrition. *Review of Educational Research*, 61, 407-450.
- Willett, J. B., & Singer, J. D. (1993). Investigating onset, cessation, relapse, and recovery: Why you should, and how you can, use discrete-time survival analysis to examine event occurrence. *Journal of Consulting and Clinical Psychology*, 61, 952-965.