5-2024

A Mixed Methods Evaluation of Three Quantitative Approaches for Examining Persistence

Angela J. Rockwell
University of Northern Colorado

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

A MIXED METHODS EVALUATION OF THREE QUANTITATIVE APPROACHES FOR EXAMINING PERSISTENCE

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

Angela J. Rockwell

College of Education and Behavioral Sciences
Department of Applied Statistics and Research Methods

May 2024

This Dissertation by: Angela J. Rockwell

Entitled: *A Mixed Methods Evaluation of Three Quantitative Approaches for Examining Persistence*
has been approved as meeting the requirement for the Degree of Doctor of Philosophy in College of Education and Behavioral Sciences in Department of Applied Statistics and Research Methods, Program of Applied Statistics and Research Methods
Accepted by the Doctoral Committee

____
William Merchant, Ph.D., Research Advisor

____
Chai-Lin Tsai, Ph.D., Committee Member

____
Randy Larkins, Ph.D., Committee Member

____
Sue Hyeon Paek, Ph.D., Faculty Representative

Date of Dissertation Defense _________________________

Accepted by the Graduate School

____
Jeri-Anne Lyons, PhD
Dean of the Graduate School
Associate Vice President for Research

ABSTRACT

Rockwell, Angela J. *A mixed methods evaluation of three quantitative approaches for examining persistence.* Published Doctor of Philosophy dissertation, University of Northern Colorado, 2024.

Institutional researchers use skills from their diverse backgrounds to collect, analyze and report data about their institutions to stakeholders representing various interests and levels of data literacy. However, there is little research into how these professionals process data and none into what aspects are important to institutional researchers when planning and executing these analyses. Given the importance of student persistence in higher education, this study used mixed methods to examine the factors institutional researchers consider when selecting a quantitative approach to exploring persistence and how they apply to selecting a quantitative approach. Three approaches (proportions, logistic regression, and discrete time survival analysis) were used to analyze student persistence at a bachelor's degree granting college and the findings from each approach were compared. The differences among these findings were shared in four focus groups of institutional researchers and administrators where participants discussed their experiences in collecting and analyzing persistence data and communicating the findings. Participants shared that the operational limitations of the approaches, the need for methodological rigor, their institutional data culture, and the way persistence was operationalized in the request are driving factors in which approaches they use. These findings present a different perspective on how to select a quantitative analysis approach than the purely methodological process taught in academia suggesting that practical learning would improve future institutional researchers' preparation for the profession.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

CHAPTER I

INTRODUCTION TO THE STUDY

Academic programs train researchers in quantitative, qualitative, and mixed methods under controlled conditions. The data sets are familiar, and the research questions are chosen to highlight aspects of the day's lesson. However, when students begin to apply these skills, it is quickly apparent that the preferred methods of academia find little quarter in applied research. Many research methods programs are adapting to include experiential learning and opportunities to work with clients in addressing actual research needs, but this change is slow and in applied research fields like institutional research, the gap between the analytical approaches that should be used from a methodological perspective and the ones that are used endures.

Institutional research professionals have many options of quantitative approaches when exploring persistence, some methodologically robust others arithmetically simple. Although there are so many to choose from, persistence rates (i.e., proportions) are the most used. This study used mixed methods to examine the factors influencing institutional researchers' decisions of which approach to use, beginning with the application of three quantitative approaches of increasing methodological complexity and appropriateness, followed by focus groups with institutional research professionals and higher education administrators at colleges in the United States to explore their experiences related to persistence modeling and applied quantitative research in higher education.

Institutional researchers have multiple stakeholders: prospective students and their families, current students, alumni, faculty and staff of their institutions, their communities, accreditors, legislative bodies, and boards of trustees or regents to name a few. Of these stakeholders, institutional researchers interact most closely with the administrators at their institutions who are their primary clients. Administrators use information to decide on budgets and for curriculum, program planning, and review, so information produced by institutional researchers touches every aspect of higher education. With so much responsibility, it is essential that institutional researchers use the best tools available to meet their institutions' information and communication needs.

Why Persistence Matters in Higher Education

The United States Department of Education's National Center for Education Statistics (NCES), through the amended Higher Education Act of 1965, requires that institutions that participate in federal student financial aid programs provide regular information about their students and institutions through the Integrated Postsecondary Education Data System (IPEDS) survey components (National Center for Education Statistics, n.d.-a). In this survey, institutions provide aggregated information about their students' retention and degree completion. Along with being publicly available through the IPEDS Data Center website, this information is also accessible through the College Navigator. Some institutions that do not participate in the federal student financial aid programs also participate in IPEDS because of its utility in advertising the institution to prospective students.

There are many incentives beyond the institutions' fiduciary responsibility to the federal taxpayers of the United States to retain students until they complete their credential. Leading proprietary college ranking algorithms heavily weight institutions' retention and graduation rates (Morse & Brooks, 2021). Some states have tied their financial support of public higher education to performance metrics like retention and completion, with at best, mixed benefit to the institutions and communities they serve (Hillman et al., 2015). There is big money in tracking and improving retention and completion rates and although higher education is a social business, it is short sighted to ignore the business aspect. Most importantly, students and their families make sacrifices to pursue higher education. Students delay fully entering the workforce, incur debt to finance their education, and forgo spending time on more enjoyable endeavors, all in the belief that they will earn a credential that will benefit them in the future.

Statement of the Problem

Studying retention and completion is natural for institutions, but much of the research is based on the definitions established for responding to the IPEDS survey components (Davidson & Blankenship, 2017; National Center for Educational Statistics, 2021a, 2021b). The omnipresence of these definitions makes comparison across institutions easy but may delude decision makers into thinking that persistence can be adequately measured with such a basic measuring stick. IPEDS's relative silence on retention and completion in students' later terms likely extends to institutional researchers' and administrators' failing to explore this period as well. For so many institutions the depth of published persistence analyses ends at the level of proportions (Purdue University, n.d.; University of Michigan, Office of Budget and Planning, 2021). A further concern with relying on these primitive approaches is that the metrics they produce do not account for students' intersectional identities or how variation in the student body composition affect persistence rates over time; they fulfill the mandated reporting requirements but are too coarse to support decision making.

Methodologically, proportions are an inadequate tool to understand persistence, but it is naïve to think that methodological concerns are the only ones that direct which approach researchers use. However, there is no research into the barriers institutional researchers face to using more appropriate analyses. The overwhelming variety of resources available on the methodological advantages of candidate approaches compound the difficulty given the competing demands on researchers' time and diverse backgrounds.

Purpose of the Study

The purpose of this study is to understand what the primary considerations are for institutional researchers when selecting a quantitative approach to exploring persistence. I examined persistence at an institution, here called Regional College, using three candidate approaches: proportions, logistic regression, and discrete time survival analysis. According to the National Center for Education Statistics, the fall-to-fall retention rate of students who entered higher education at a public 4-year institution in fall of 2022 was 81.2%, meaning that about one in five students did not return for their second year (National Center for Educational Statistics, 2024).

I posed the following questions:

Q1 What factors do institutional researchers consider when selecting a quantitative approach to exploring persistence?

Q2 How do the factors apply to selecting a quantitative approach to exploring persistence?

Design of the Study

To answer this study's research questions, I used an embedded mixed method design with the quantitative strand embedded in the qualitative strand (Teddlie & Tashakkori, 2009). The quantitative analysis was first, and its findings informed the qualitative strand in which I followed the constructivist paradigm as described by Creswell and Poth (2018) to explore the meanings and implications of the researchers' and administrators' realities and to focus on the complexities and interactions that surround persistence modeling. I integrated the two strands in both the data collection and analysis phases. In the data collection phase, I documented my experiences in performing the quantitative analyses in the researcher journal, which I used along with the findings from the quantitative analysis to inform the focus groups. In the analysis phase, I used the findings from both strands to identify the factors considered by institutional researchers and to examine how those factors interplay in practice.

This study required a mixed methods design. There was little research into the application or interpretation of quantitative research methods within institutional research, so qualitative methods were necessary to explore the topic and lay a foundation. Because the topic centers on quantitative methods, this study also needed to incorporate quantitative research methods. After receiving the attached Institutional Review Board approval (Appendix A) to access and analyze the student data from Regional College in the quantitative strand, I applied each of the three candidate approaches to explore persistence at that institution. By sharing my experience using the three approaches along with the meta-findings from those analyses in the later focus groups, I focused the discussion, provided new discussion points, and prepared myself to better appreciate the challenges and opportunities mentioned by the participants.

Adhering to the procedure outlined in the Institutional Review Board approved protocol for the qualitative strand in Appendix B I held four focus groups of collectively 14 participants via Zoom®. I chose to hold the focus groups via the popular online videoconferencing software to minimize disruption to the participants' schedules and to encourage participation by people with a variety of personal backgrounds and from a variety of institutions. I used a brief presentation of my experience and meta-findings from the quantitative strand to split the discussion into general aspects of persistence modeling followed by a discussion of persistence research within the context of the three approaches. After transcribing the discussions, I used the classic analysis strategy as outlined in Krueger and Casey (2009) to analyze the transcripts, an application of constant comparison similar to the structural coding described by (Saldaña, 2016, p. 98).

Definition of Terms

In this section are explanations for the key phrases essential to this study.

Degree Completion. Institutions each have their own requirements and processes for awarding degrees, but for this study, institutions award degrees in the term in which the student has completed all the requirements of their program including all administrative tasks. At Regional College, degrees were awarded in the term that the Registrar has verified that all requirements are met, which is generally the same term that the student completes the required coursework for their program. Although it is an administrative delay, failing to apply for graduation may result in the degree being awarded in a term after the last of the coursework was completed. Degree completions are reported in the Completions component and as graduation rates in the Graduation Rates and Graduation Rates 200% components of the IPEDS surveys (National Center for Educational Statistics, 2021a).

Enrollment. Students were enrolled at Regional College only in those terms in which they are registered for at least one credit bearing section as of the official freeze date for the term. Any changes made to the students' registration statuses after that date are not considered for this study. Students may make changes to their course registrations until the add and drop deadlines for each section. They are registered in all sections for which their latest registration activity is one of adding or withdrawing from the section and are not registered in all sections for which they have no registration activity or for which their latest registration activity is one of dropping the section. Under rare circumstances, Registrar's office staff may change registration activity after the add or drop deadlines. All activity occurring before the official freeze date for the term is reflected in this study. Institutional enrollment is reported in aggregate in the 12-Month Enrollment and Fall Enrollment components of the IPEDS surveys (National Center for Educational Statistics, 2021a).

New First-Time Student. Degree-seeking students whose first college level coursework after high school graduation was at Regional College in a fall term or was in a summer term and they re-enrolled at Regional College in the immediately following fall term are first-time students (National Center for Educational Statistics, 2021a). Students regularly enter higher education having already earned college credits, so not all enter as freshmen. Graduates from high schools with early college programs complete an associate degree along with their high school diploma so they enter higher education as juniors. Fall starting new first-time students compose the fall cohorts used for multiple calculations in IPEDS including those in the Outcome Measures component (National Center for Educational Statistics, 2021a).

Retention. Students enrolled in the initial term who re-enrolled in the follow up term retained to the follow up term. Retention is reported in the Fall Enrollment component of the IPEDS surveys (National Center for Educational Statistics, 2021b). IPEDS allows institutions to adjust their cohorts by excluding students who passed away, were deployed as an active military member, or were engaged on a religious mission before the follow up term (National Center for Educational Statistics, 2021a).

Persistence. Persistence is a composite of degree completion and retention and is not a term used in IPEDS. A student retained to the follow up term or who completed a degree before the follow up term has persisted from the initial term until the follow up term. As a composite metric, persistence does not penalize institutions for students who do not re-enroll in the follow up term because the student had completed their degree and so would not be expected to re-enroll.

Significance of the Study

This study contributes to the literature by exploring and documenting institutional researchers' considerations when selecting a quantitative approach to modeling persistence. Methodological literature is clear on which approaches researchers should use. However, this is based solely on their methodological benefits and ignores other constraints in the research environment. The researchers who use those approaches and decision makers who digest the findings may not be aware of the implications of using a methodologically inferior approach. Before any advocacy for the appropriate approaches can be effective, we need to first identify and understand the barriers to better professional practice. Other researchers have investigated critical quantitative research within institutional research, the factors that influence institutional researchers use of qualitative methods, and the role of institutional research in supporting decision making, but this is the first study to investigate factors that influence which quantitative approaches institutional researchers use. This study ties these themes together as critical theory and decision support motivate approach selection and institutional research professional's considerations around qualitative methods parallel those they have around the more complex quantitative methods.

The second contribution of this study is the findings from my analyses that I provided Regional College to deepen the understanding of its students' persistence. Like at many institutions, the research into persistence at Regional College was previously limited to proportions and rudimentary logistic regression analysis. The findings contributed to persistence modeling by considering more complex data structures, highlighting longitudinal trends, and confirming that widely observed relationships are present in their community.

Delimitations and Limitations of the Study

This study examined the considerations of institutional researchers and used data from higher education institutions in the United States that reported to IPEDS in the 2020-21 survey cycle, although the findings may apply to other applied science fields and to institutional research at other institutions. I only explored the perspectives and experiences of senior administrators and institutional research professionals, which may differ from others' in higher education such as lower and middle administrators, faculty, and students. I limited the variables in the quantitative analyses to those that were readily available. This study may have had different findings had I included non-cognitive variables which are known to have strong correlations with persistence. Lastly, my dual role as a researcher and an institutional researcher impacted my selection of study design and interpretation of the findings.

Researcher's Background

I have almost a decade of experience working in institutional research, the field in which professionals "identify information needs…collect, analyze, interpret, and report data and information…plan and evaluate…serve as stewards of data and information…[and] educate information producers, users, and consumers" in higher education (Association for Institutional Research, 2022). Like the academic backgrounds of many institutional researchers, mine is eclectic, including chemistry, German, public health, computer science, and statistics. Having lived in the worlds of the physical sciences, the last harbor of positivism, as well as the humanities, I appreciate the challenge of applying the academic knowledge of the classroom to the situations outside those walls. As an institutional researcher, I found that the considerations dictating what procedures I used and what questions I asked extended beyond the methodological. Often, I would use a procedure because it was the best for a situation although methodologically inferior; my experience could not be unique. Unfortunately, many institutional researchers lack the statistical skills necessary to safely navigate the hazards of violating assumptions. Beginning from the naïve position of methodological superiority, my need to understand the differences between statistical analysis as taught in academia and the application of those skills led me to investigate the applied environment of institutional researchers.

CHAPTER II

REVIEW OF THE LITERATURE

This study continued the work began by other researchers in the fields of higher education student success, quantitative analysis, and institutional research. This chapter begins by outlining applications of the three chosen analytical approaches in exploring student persistence and continues with presenting primers on the three approaches with special attention paid to potential challenges. In the later part, it presents literature on virtual focus groups for qualitative data collection.

Exploring Persistence in Higher Education

Before diving into persistence metrics, it is important to discuss what exactly it is that they explain or predict. There are many intangible benefits to the student and their community for having attended higher education, but the most discussed and most easily quantified are the economic benefits: the earnings premium of the different degree levels and of selected fields of study over others (Kim et al., 2015; Shafiq et al., 2019). With this mindset, graduation and continues enrollment until graduation are the primary measures of student success. However, this reduces the benefits of higher education to earning credentials. Certificates and degrees represent students' growth and effort in convenient parcels, but they are not the only way. The benefits of continued education are not held in waiting until a student earns their credential. Instead, they accrue along the way to graduation with a cumulative benefit when the credential is awarded. There is growing literature on the career trajectories of students who start but do not complete a bachelor's degree, highlighting the diverse paths students take through higher education to reach their goals (Luckman & Harvey, 2019).

Theoretical Models

From the mid-1970's on, researchers developed theories to examine the causes of attrition and persistence. (Tinto, 1975) developed a conceptual framework connecting the academic system, academic integration, social system, social integration, commitments, and student level factors to the student's decision to leave an institution called the Student Integration Model. This model emphasizes institutional commitment, focusing on the student's attachment to a particular institution as a means to attain their goal. By 1993, Tinto had made only minor changes to the Student Integration Model, including interactions between the student and their peer group and between the student and faculty and staff, again within the context of institutional commitment (Tinto, 1993, p. 114).

In 1982, Bean proposed his Model of Student Departure that offered one explanation of the causal connections between student, institutional attributes, and student attrition. In a later study, Bean (1983) used an ordinary least squares path model based on ten factors relating to 820 female freshmen students' experiences at the institution, their opportunities to transfer to a different institution, and their intentions to marry or to leave. Not surprisingly, the largest explanatory power came from the student's intention to leave. Cabrera et al. (1992) explored the convergent validity of Tinto's Student Integration Model and Bean's Model of Student Departure concluding that researchers gain a deeper understanding of persistence when both models are used. Their findings confirmed Bean's (1983) that intent to persist has the highest correlation with actual persistence.

With their 2008 study, Kuh, Cruce, Shoup, Kinzie and Gonyea furthered our understanding of persistence and demonstrated how large-scale student surveys could be leveraged to understand persistence at the institutional level. They used a logistic regression to estimate the relationships between student's ethnicity, selected measures of college preparedness and expectations, financial need, high school grades, academic outcomes, and other variables with the students' persistence to their second year. Although this model was not based on a theory that proposed causal relationships between the variables, such atheoretical models give a snapshot of the health of the student body, indicating where researchers should focus their attention. Bean's (1982) objection that there were already too many atheoretical models in research four decades ago is not without merit, but researchers need to create and use far more atheoretical models in practice than we do now. Studies using these models may not contribute to the body of literature, but in institutional research theory is useful only as far as it serves a practical purpose.

Atheoretical Models

Atheoretical models continue to thrive in practice because even after decades of theory, the mechanisms of persistence remain largely obscure. Traits related to the student, their environment, and their communities help explain the likelihood of a student to persist. Gansemer-Topf et al. (2017) found that the likelihood of persisting varies across major or field of study and Allen and Robbins (2010) found that alignment between a student's major and their interests is associated with timely degree completion and better first year academic performance. Persistence varies across ethnicities, citizenship status, high school grade point average, aptitude test scores, parental education, and residency (Gansemer-Topf et al., 2017; Zhang et al., 2004). Student's non-cognitive traits, like their sense of belonging, degree commitment, overall evaluation of the institution, and academic self-efficacy and adjustment may also play a role in persistence (Gansemer-Topf et al., 2017). Even administrative aspects of the students' academic environment are correlated with persistence, like whether the student successfully completes an online course and what proportion of the student's course load is delivered online (Jaggars & Xu, 2010; Shea & Bidjerano, 2019).

Institutional Research Practice

Institutional research is an applied social sciences research field, but one heavily influenced by quantitative epistemology. The largest professional organization for institutional researchers in North America, the Association for Institutional Research (AIR), offers resources and professional development opportunities primarily on quantitative topics or on IPEDS reporting (Nelson, A., 2022; Deom, Talley, Sauer, & Fiorini, 2024; Association for Institutional Research, n.d.). Although AIR regularly surveys its members, there is little research on institutional researchers either by the organization or by other entities or individual researchers. Relevant topics in the existing research are the role of institutional research in the institutional decision-making processes, the emerging influence of critical quantitative research in the profession, and the role of data sense-making in institutional research.

Rouse (2018) found that "institutional researcher's work is guided by a calendar of deadlines, often external deadlines, which includes deadlines for IPEDS, mandated reporting and information for ranking publications." Alongside these obligations, institutional researchers were also found to initiate research for meeting institutional needs such as attracting more applicants and supporting additional grant submissions. Further, some institutional researchers believed they could influence policy and internal decision making while others had hope that institutional research's change agency was increasing.

Along with increased awareness of inequities in higher education comes an increase in the attention paid to the structures that perpetuate them. One approach, critical quantitative research applies critical theory to quantitative analysis. Long's (2020) dissertation found that institutional researchers less frequently use the methods and perspectives of critical theory when examining persistence data broken out by race. Senior institutional researcher participants tended to use a color-blind framing, ignoring influences of systemic racism and privilege.

In Villalobos Meléndez's (2023) action research project exploring the experiences of institutional researchers in data sensemaking of student equity data, participants shared how their personal and professional identities were lenses through which they interpreted data and how their social environments further influenced their interpretations. Their process of sensemaking was continual as they drew connections between sources and prior knowledge as well as reiteratively revisiting their earlier sensemaking efforts to reprocess and identify deeper meanings. Importantly, Villalobos Meléndez was unable to observe the influence of power structures on the institutional researchers' sensemaking, which she acknowledged is an important aspect of the process.

Analytical Approaches

Quantitative models of persistence impose a structure on the relationship between the variables assumed to contribute to persistence and the persistence outcome measure of interest. Persistence can be measured as a dichotomous variable, as the likelihood of a student persisting, or as the count of students who persisted. Each of these outcomes is associated with an analytical process: proportions, logistic regression and discrete time survival analysis as a special case of logistic regression, and Poisson regression. In institutional research, proportions are more often called rates to reflect that the number of students who experienced the event is over a period of time. In this section, I provide a brief overview of how the first two processes are used in exploring persistence.

Proportions

The most widely used process for exploring persistence is the proportion of counts made popular through the IPEDS survey components. With the information collected through the Fall Enrollment (EF), Graduation Rates (GR), Graduation Rates 200 (GR200), and Outcome Measures (OM) components, retention and graduation rates are calculated for cohorts of students (National Center for Education Statistics, 2021b, 2021a, 2021c, 2021d). An overwhelming majority of colleges report to IPEDS making its definitions standard throughout higher education. These persistence rates are shared with the public through the IPEDS Data Center and College Navigator, and many colleges also publish these rates on their websites and include them in marketing communications.

Logistic Regression

Application in Persistence Research

Building on persistence rates, which have the proportion of students who experience the persistence event as their outcome, are those models which incorporate statistical error and instead estimate the likelihood of observing the persistence event as their outcome. Ordinary least squares (OLS) regression can estimate a dichotomous outcome where the mean of the outcome variable represents the probability of an event occurring. When OLS is used in such a situation, the coefficients are unbiased, but the expected probability can fall outside the range of [0, 1] and the OLS assumption of a linear relationship between the covariates and the outcome variable is violated (Menard, 1995, p. 7).

An option for addressing this nonlinearity between covariates and outcome variables is to transform the outcome variable which is the foundation for logistic regression. Logistic regression increased in popularity in published higher education research between 1988 to 1999 and the models and techniques became increasingly complex (Peng, So, et al., 2002). Peng, So, et al. (2002) attributed this trend to improved computing power and statistical analysis software necessary for the procedures. More than half of the articles they reviewed used logistic regression to study university enrollment and retention. Logistic regression can also be extended to model polytomous ordered outcomes (Peng, Lee, et al., 2002). This feature allows for modeling the levels of an outcome such as differentiating the probability of retaining to the college with the initial major and retaining to the college but changing majors.

Multivariate analysis is also straightforward with logistic regression. In 2017, Gansemer-Topf et al. used principal axis factoring to select informative factors from a 150 item transition survey administered to a group of students at a large 4-year research institution. They combined these with institutional data in logistic regressions to separately explain fall to fall persistence of students in science technology engineering and math (STEM) majors and those in non-STEM majors. They included interactions of gender, ethnicity, and residency which were not significant in the STEM model but were so in the non-STEM model. Non-STEM in-state students were more likely to be retained than their out of state counterparts and within the in-state group, non-White students were more likely to be retained than White students. They also found that ACT scores, being a member of a learning community, and academic self-efficacy and adjustment were only significant in the STEM model.

In a study like logistic regression approach in this study, Zhang et al. (2004) used a dataset of students pursuing engineering to understand the likelihood of graduating within six years of matriculation. They included ethnicity, gender, high school GPA, SAT math score, SAT verbal score, and citizenship status as covariates. The logistic regression approach in this study differs from their analysis because they selected the six-year mark for their follow up time because that was the point when graduation rates plateaued, whereas here it was selected without such a consideration.

Logistic regression pairs well with many other regression approaches. When used with multilevel modeling, a logistic regression can account for the nesting inherent in educational data. Arreola and Wilson (2020) used instructors and majors as random effects in a multilevel logistic regression to explain student academic success. The student's term GPAs for three sequential semesters were nested within students and students were nested within multiple instructors, but only one major. Logistic regression also works well with hierarchical regression in modeling persistence outcomes as demonstrated by Perry et al., (1999), and Allen and Robbins (2010) developed a logistic path model with high school achievement, first generation status, ethnicity, and gender as exogenous variables to explain first year academic performance and timely degree completion.

Reporting

The reporting requirements for a logistic regression are also a step up from those required for an OLS regression. Beyond the standard components of OLS reporting including the full model used, an evaluation of the model and individual predictors, and the sample size, researchers should also report the link function and an evaluation of the predicted probabilities for a logistic regression (Peng, Lee, et al., 2002). Additionally, Peng, So, et al. (2002) recommend reporting the change in probabilities for continuous covariates and odds ratios for categorical covariates. Researchers most commonly report the statistics provided as a default by the software package, specifically, the odds ratio, confidence interval, and p-value, but rarely report alternatives such as marginal effects and predicted probabilities (Niu, 2020). Peng and So (2002) also recommend reporting the statistical package used, because after reviewing 6 popular statistical packages for logistic regression, none were error-free.

The change in probability attributable to a change in the covariates is an appealing measure to describe the relationship between the covariates and probability of the outcome, but it is prone to misuse and misinterpretation. Readers may assume that the change in probability depends only on the amount of change in the variable, X_i, and the regression coefficient, β_i, as in OLS, but in logistic regression, the change also depends on the initial value of X_i. Because the change in probability, ΔP, is not linear with the change in X_i, the interpretation of ΔP, differs from the interpretation of the change in the linear outcome, ΔY, in OLS regression; ΔY is constant throughout the range of X_i, but ΔP is not (Petersen, 1985). ΔP's dependence on the values of all other covariates in the model can be a strength when used to highlight specific scenarios. However, communicating marginal effects and predicted probabilities appropriately means providing the values of all covariates and explaining why those values were selected (Niu, 2020). For example, Kuh et al. (2008) included the predicted probabilities for each level of each variable with all other variables held at their means. However, they did not explain why they chose to hold the other values at the means. It may seem obvious to choose the means, but if no one in their population resembles that profile, using the means is a mathematical convenience that poorly describes the relationships in the population.

There are four options for reporting the communicating strength of the relationship, also known as the effect size, for continuous variables. The first is the change in the event's probability, ΔP mentioned earlier. Another option is the event's odds or the probability the event occurs divided by the probability that it does not. The odds ratio, which I discuss later is based on the odds but communicates different information and is easily confused with odds. Related to ΔP is marginal probability, the instantaneous change in the probability holding all but one covariate constant. As a third option, it is less intuitive and less effective than either ΔPs or odds ratios (Peng, So, et al., 2002). The limitation with marginal probability though, is that the probability curve is not linear and marginal probability describes the line tangent to the probability curve at the selected value. It is difficult to translate a marginal probability into the expected change in probability given a discrete change in the covariate of interest. This is particularly problematic when the covariate is categorical. The fourth option is the log of the event's odds, called the logit or log odds. Given that these metrics can all take on values between 0 and 1 exclusive, researchers should indicate which they have provided and not just the calculated value itself to avoid misinterpretation (Peng, So, et al., 2002).

Categorical variables, like full-time/part-time status, major, or whether the student received financial aid, have three additional options for communicating effect size. The first is relative risk, which is calculated as the proportion of the group coded 1 that had the event divided by the proportion of the reference group (the group coded 0) that had the event. Using relative risk makes it more explicit that the comparison is between the groups, although calculating ΔP for categorical variables makes the same comparison (Tabachnick & Fidell, 2007, p. 463). The second is the odds ratio, which is most analogous to a correlation coefficient in ordinary least squares regression. Unfortunately, correct interpretation of odds ratios is not intuitive. An odds ratio is the odds of an event in one group divided by the odds of the event in another group and can communicate the odds of a first-generation student retaining to the next term compared to the odds of a non-first-generation student. However, some researchers incorrectly interpret the odds ratio as the relative risk (Niu, 2020). Others describe odds ratios accurately but use phrasing that is easily misunderstood by readers as if the odds ratio were a comparison of probabilities. When possible, a better alternative is the risk ratio. Risk ratios are intuitive, but their calculations must be based on the entire population with an implicit assumption of causation and should be presented alongside odds ratios (George et al., 2020). Unfortunately, their assumption of causation means they are not an option for communicating the findings of observational studies. When the outcome is very rare, odds ratios may be used to approximate risk ratios, but in persistence research it is rare for the outcome to be rare. Interpreting odds ratios as risk ratios when the outcome is not rare exaggerates the effect of the covariate (George et al., 2020).

The best study is only as good as how well its findings are communicated with stakeholders. In addition to avoiding misleading statistics and selecting the most effective metrics to communicate the findings, researchers must consider how the findings themselves are presented. Peng, Lee, et al. (2002) recommend that the results of logistic regressions be presented in both tabular and graphical formats because readers may otherwise find it difficult to correctly interpret the findings. Charts are particularly useful to communicate the interaction between categorical and continuous variables. A chart will show the relationship between the continuous variables and the predicted probability much more clearly than a table with the same information.

Survival Analysis

Application in Persistence Research

Survival analysis allows researchers to answer questions about the time until an event happens, like how long until a toaster breaks to determine an appropriate warranty period.

Although originally developed for continuous-time outcome variables, it was later adapted for use when time is measured discretely (Singer & Willett, 1993). Discrete time survival analysis is based on logistic regression with the major differences being the nature and structure of the data used and the addition of covariates representing time and a proportional hazards assumption. The concepts are the same whether the outcome is continuous or discrete. In this study, retention, graduation and withdrawal events were only observed once per term.

Survival analysis is a valuable tool for understanding how the relationships between student traits and aspects of their environment change over time. Gury (2011) used a discrete time event-history analysis to examine the higher education dropout rates in France of the 5,383 students who entered middle school in 1989 and eventually entered higher education. His survival analysis explored the effect of time on the hazard of dropping out of higher education, filling the gap of many earlier studies that used simpler approaches like logistic regression.

In 2018, Shea and Bidjerano (2018) used a logistic regression to explore the relationship between the proportion of courses community college students took online (vs traditional/face to face) and the students' academic outcomes. Then in 2019, they reexamined the relationship using the more complex survival analysis which allowed them to explore not only whether but when a student completed their associate degree, transferred to a 4-year institution, or left higher education before completing a degree (Shea & Bidjerano, 2019). They accounted for there being multiple mutually exclusive outcomes by using a competing risks model. Students who completed their associates' before transferring to a 4-year institution were coded as degree completers because that event happened first. Those who did not complete before transferring were coded as transfer students. If they had not considered these two outcomes jointly and focused only on whether the student completed a degree, the transfer student would look the same as a student who left higher education all together although these are very different outcomes.

In a further extension of discrete time survival analysis, Muthén and Masyn (2005) used a discrete-time mixture model survival analysis to identify latent groups within an elementary and middle school retention data set. With one of their models they identified long term survivors, who are all of participants who do not experience the event during the observation period. The goal of this analysis was to identify the traits of these long-term survivors as a group. A similar analysis in higher education could be used to identify the traits of undergraduate students who continue on to graduate school at the same institution, i.e. continue to re-enroll after their observation as an undergraduate student ends.

As an alternative to survival analysis for modeling longitudinal data with a categorical outcome, Hu et al. (1998) propose using random effects models to estimate subject-specific rates and generalized estimating equations to estimate population-averaged effects. These options handle time dependency differently than survival analysis though. Examples of population-averaged effects models are the Mantel-Haenszel method (Mantel & Haenszel, 1959) and the standard logistic models for independent binary outcomes. Assuming that the observations are independent though, yields biased (exaggerated) standard error estimates when multiple observations are made on the same participant. The random effects modeling treats the repeated observations as nested within subjects in a multilevel structure and estimates a random effect that is similar to frailty, which explains the variation in hazard functions among participants with the same levels of the covariates.

Reporting

The recommendations for reporting logistic regression findings also apply to discrete time survival analysis with a few more specific to survival analysis. Risk ratios, odds ratios, and hazard ratios should be reported along their confidence intervals and p-values along with non-ratio statistics for clarity (George et al., 2020). Ratio statistics describe how two groups compare with each other, but all ratios need references of scale. For example, knowing that an animal is strong enough to support 3,400 times its body weight with its neck is impressive, but knowing that the animal is an ant, and therefore has a much smaller body weight than a human puts that ratio into proper perspective (Nguyen et al., 2014). Further, many survival analysis models have the proportional hazards assumption. When reporting on these models, the median time to event should be reported along with the hazard ratio for each subgroup as supporting evidence to show the proportional hazards assumption was not violated (George et al., 2020). Singer and Willett (1993) suggest also using the median lifetime to communicate differences among groups, even if the data makes exact estimation of one of the median lifetimes impossible. Further, they suggest presenting the fitted hazard and survivor functions using parameter values important to the audience as a more tangible method of communicating results than only interpreting odds ratios. This recommendation is similar to that of Niu (2020) who suggested using meaningful values for the other covariates when communicating marginal effects.

Researchers have many options from basic proportions of students who retain or graduate, to models that estimate the likelihood of persisting and those that estimate how these probabilities change over time. Each approach has found favor in publications and has strengths that make it useful for exploring persistence. However, the more complex the approach becomes, the more onerous the reporting requirements to steer readers away from incorrect conclusions.

Qualitative Data Collection

Focus Groups

In response concerns about the data quality produced by the structured and restrictive interviewing techniques that predominated before the 1950's, researchers explored data collection techniques that matured into the focus groups of today. Krueger & Casey (2009, p. 2) explain that focus groups are a data collection method using discussions on defined topics of between within small groups who are selected because of their similarities. The discussions are guided by a skilled moderator and are held in a relaxed and non-threatening environment where participants are encouraged to share their opinions and experiences without pressure to come to consensus or agree.

They advise that focus groups are best used in research intended to understand diverse perspectives or to identify what influences opinions, behaviors, or motivations (Krueger and Casey, 2009, pp. 19-20). Focus groups are useful for soliciting input from individual participants as well as the perspectives that result from the dynamic interactions of the participants and effectively capture the common vocabulary and options participants consider when making decisions. However, Krueger and Casey (2009) advise against using focus groups when participants must come to a consensus or when the flow of information is one way. Focus groups are also ill suited to collecting quality data on sensitive or emotionally charged topics and researchers should only use them when their high cost and resource demands are likely to be offset by collecting higher quality data.

Virtual versus Face-to-Face Focus Groups

Although researchers have largely embraced online mediated focus groups, some have expressed concerns about focus groups mediated online since the adoption of virtual modalities in 1998 (James & Busher, 2009). Specific disadvantages that should be addressed in deciding to hold the focus group virtually include distractions from the technology, less moderator control, more chaotic transcripts, potentially excluding participants without internet access, and topics being covered with less depth (Abrams et al., 2015). Methodological researchers have warned of relatively shallow discussion in online focus groups, citing the fewer spontaneous verbal and nonverbal cues and shorter response length (Larkins, 2015; Abrams et al., 2015; Davies, et al., 2020).

However, the longer responses in face-to-face focus groups may not be wholly an asset. Davies, et al., (2020) found that the shorter responses in online formats were generally more on topic and agreed with Dodds and Hess (2021) that participants were more likely to share sensitive information in the online format. Participants in the study by Dodds and Hess (2021) also appreciated the increased privacy of not having the interviewer physically present in their personal environment. Virtual focus groups have fewer dominant participants, a reduced risk of group conformity, and the option for the moderator to less obtrusively follow a script or protocol (Abrams et al., 2015; Dodds & Hess, 2021; Lobe & Morgan, 2021). The mass adoption of video conferencing technologies in March 2020 may have changed how participants interact in the format.

Importantly, when both face to face and virtual focus groups are viable options, researchers must balance the benefits to participants and researchers of avoiding travel related fatigue with the costs of contributing to “Zoom fatigue”, the emotional tole of prolonged communication via video conferencing software (Dodds & Hess, 2021; Lee, 2020). The budgetary implications of deciding between a face to face and an online focus group are also nuanced. Recruitment may be more expensive for a virtual format and while some software options for virtual focus groups are free, others cost hundreds of dollars per license (Wilkerson et al., 2014). However, these additional costs may be offset by avoiding space rental and refreshment costs. Some virtual focus group options have the added benefit of reducing the time and financial cost of transcription.

The concerns around virtual focus groups extend beyond the operational. Larkins (2015) highlighted ethical issues that may arise in an online focus group, finding that privacy was not a concern for his young adult participants, who were desensitized to the surveillance of a web cam and freely shared protected academic information about themselves. Online data collection is by nature more vulnerable to snooping, so much so that researchers must consider the encryption options available and storage policies that affect their data (Marhefka et al., 2020).

The concern that virtual focus groups exclude participants who lack the resources to participate (Abrams et al., 2015; Davies, et al., 2020; Krueger & Casey, 2009) is largely a function of the target population. There remain many communities where financial and infrastructure limitations would prevent or curtail participation in a virtual focus group, but researchers must also recognize that face-to-face focus groups have their own resource-related challenges. Participants must have not only the time available to participate in face-to-face focus groups themselves, but also the time and resources to travel to and from the location. Participation in face-to-face focus groups also requires proximity to the meeting location, making a regionally diverse focus group cost- and time-prohibitive and excludes participants who have transportation or mobility constraints (Davies, et al., 2020). Participants who also manage professional careers appreciate the convenience of participation via a videoconferencing app when logistics make in person participation prohibitively inconvenient (Archibald et al., 2019). Additionally, people within regional communities tend to be more similar to each other than to people in other communities, limiting the diversity within a single face-to-face focus group.

The social distancing guidelines and mandates effected in response to the Corona virus pandemic forced many researchers to change data collection modalities; those who had been interacting with participants face to face needed to find alternatives or forego further collection after early spring 2020. To help these researchers navigate the distance data collection options, Marhefka et al. (2020) developed a series of decision checklists, outlined best practices, and described protocols for selecting the modality and platform for continuing data collection.

This chapter reviewed how researchers have explored persistence using theoretical and atheoretical models as well as the limited relevant research on institutional researchers. Then it discussed the three quantitative approaches used in this study, focusing on how each approach has been used in persistence research and concerns related to interpreting and reporting the findings. This chapter concluded with a discussion of focus group data collection and the use cases of virtual focus groups in preference to in-person focus groups.

CHAPTER III

METHODOLOGY

Chapter three begins by presenting the research questions and epistemology that guided this study then describes the methodology used to collect, analyze, and interpret data using qualitative and quantitative methods. Here, I explain my rationale for using mixed methods followed by a description of my processes in the qualitative and quantitative strands. In the quantitative section, I describe the methods used for the three quantitative approaches. Chapter three concludes with an explanation of how I joined the qualitative and quantitative strands.

Introduction

Institutional research is an applied field staffed by people from diverse professional and academic backgrounds with varying levels of research autonomy and authority who work at institutions with wide varying funding levels, organizational structures, and missions. This means that institutional researchers as a profession have many options in how we address research interests, ranging from descriptive statistics to machine learning and from surveys to panel focus groups. Many researchers have previously evaluated the quantitative techniques' theoretical and mathematical fitness, but there is no research into selecting an appropriate technique accounting for institutional researchers' operational environments, specifically around evaluating student persistence. This study aims to evaluate three quantitative approaches to evaluating persistence as a binary outcome through the lens of institutional research.

To guide this research, I identified the following research questions:

Q1 What factors do institutional researchers consider when selecting a quantitative approach to exploring persistence?

Q2 How do the factors apply to selecting a quantitative approach to exploring persistence?

I approached this question using both quantitative and qualitative research methods. I "tested" the three approaches on a real data set from Regional College to identify the strengths and weaknesses of each and to experience for myself the different findings available by using each of them. Then I used these findings and experiences as the scaffold of four focus groups to which I invited institutional research professionals and higher education administrators. In these focus groups we discussed their experiences around institutional research, focusing on how persistence is explored at their institutions.

Epistemology

Epistemology is how we know what we know about our realities and what evidence we accept to judge new information as valid and true. In this study, I used two divergent stances on knowledge. In my qualitative evaluation, I operated through the constructivist epistemology, accepting that there were multiple realities and that in exploring these realities with others, they and I together created a more complex and nuanced understanding of our shared reality. In this aspect of the study, knowledge was not fixed, and my role was to record the knowledge from multiple similar perspectives, accepting that each perspective provides a unique meaning. I captured as much of the detail as possible, the evolution of processes, and the biases that my participants and I had that distorted the picture. When applying the three quantitative techniques in the quantitative strand, I used the post-positivist epistemology; knowledge was quantifiable and measurable within some level of uncertainty.

Persistence is foundational to student success in higher education, but there is no universal understanding of persistence within institutional research or higher education. Individual approaches to evaluating persistence are paired with a distinct conceptual model of what persistence is and how it happens. By using a particular approach, researchers make force assumptions about the nature of persistence. Although academic training suggests otherwise, choosing an approach is more complex than picking the statistical model that most closely aligns with the researcher's conceptual model. The varying approaches allow researchers to answer different questions, and vary in complexity, ease of use and interpretation, data requirements, and the validity of the conclusions researchers can draw for specific scenarios. To illuminate the considerations that researchers use to select the best approach for modeling persistence, I executed a mixed methods design which allowed me to leverage the strengths of qualitative and quantitative research methodologies.

Choice of Methodology

By the definition of Teddlie and Tashakkori (2009), this study was an embedded integrated mixed methods design with two phases. The first phase had a quantitative strand in which I used three nested statistical approaches to examine persistence using data from a real institution and a qualitative strand in which I used my researcher journal maintained during the quantitative analysis, as a resource of the experiences of an institutional researcher. I selected the three approaches using personal experience and the prevalence of approaches cited in literature and evaluated each approach's ability to explain persistence, ease of use, operational considerations, and assumption violations in this application. I also examined how the conclusions differed based on the approach chosen.

I used my researcher journal to inform and supplement the qualitative strand in the second phase, which consisted of a set of focus groups with fellow institutional researchers and institutional administrators. I mirrored my quantitative strand off of the analyses institutional researchers routinely perform, so that the focus group participants would share experiences, attitudes, and beliefs related to the same phenomena. Administrators often initiate research interests and also rely on the information produced by institutional researchers in their decision-making, making the relationship between institutional researchers and administrators cyclical.

The qualitative and quantitative strands came together in the data collection stage when I used my experiences collected in the researcher journal to guide my focus group questions and in the interpretation stage when I examined the experiences and perceptions of the institutional researchers and administrators alongside my own. The findings from the qualitative and quantitative strands will complement one another by more fully describing the decision environments that researchers encounter and outlining why particular approaches are used. The findings from each strand describe part of institutional researchers' justification for selecting a particular approach. I recognized that institutional research is made of people from diverse backgrounds and that their roles vary among institutions. I attributed differences in the findings between strands as the natural sequelae of these differences in researcher experiences and professional environments.

Data Collection

For the qualitative strand, I conducted three focus groups with institutional researchers and one with college administrators and analyzed the researcher journal I maintained while pursuing the quantitative arm. Collectively, the qualitative strand was structured to identify institutional researchers' considerations related to quantitative analysis and the three selected approaches in particular. The data for the quantitative strand was student demographic, enrollment, and graduation data that I obtained from Regional College following IRB approval. The enrollment data included all enrollments from fall 2014 to fall 2022 or until the students completed their first bachelor's degree at Regional College which ever came first. The degree completion data was the term of the students' first bachelor's degree completion. The students included were all new first-time full-time degree seeking fall starting students who started at Regional College in fall 2014 to fall 2021.

Focus Groups

Collecting data through focus groups was a natural choice for my qualitative data collection because I wanted to uncover systematic themes in the institutional researchers' and administrators' experiences, and I especially wanted to tap the rich data that comes from participant interactions. The group interactions allowed me to explore my relative inexperience in IR and the emphasis on statistics-driven analysis in my academic background so that I better understood my position in the research. The participants and I interacted identifying nuances in their individual experiences that I did not plan for.

I received Institutional Review Board approval from the University of Northern Colorado under protocol number 2302048368 for the qualitative portion of this study.

Modality

After deciding on focus groups for data collection, I invited participants from institutions spread across the United States, spanning six time zones and thousands of miles. Because of the challenges involved in meeting for focus groups in-person, I elected to hold the focus groups via video conferencing software because it came into widespread use in response to the social distancing need of the early Corona virus pandemic in 2020. Ultimately, my participants were comfortable with the modality. The responses were not noticeably truncated, and the topics generally received adequate discussion. My experience supports Dodds' and Hess' (2021) findings that the pandemic response measures increased participants' comfort and familiarity with the video conferencing software.

Video Conferencing Logistics

Following the key considerations Marhefka et al. (2020) outlined for platform selection, I chose Zoom® because it had the features necessary for my data collection; it was available for computers and mobile devices, was free to download and use, offered encrypted recording, and features that prevent uninvited people from joining the focus group (Zoom Video Communications, 2022; Archibald et al., 2019).

Participants joined the focus groups through meeting invitations I had emailed to them, so they had full control over where they joined from and independent control over the privacy of their environments. Nearly all participants chose to join from a dedicated personal office space, most often at their respective institutions. To respect participants' privacy and to acknowledge the mental toll of "Zoom fatigue", I supported camera-off participation and encouraged participants to mute themselves when not contributing. Interestingly, most participants chose to have cameras on for the full focus group. Two of the three participants who had their cameras off had them off intermittently, one of them because of bandwidth and connectivity issues. The third participant remained camera off for their entire focus group. After introducing the focus group and reminding participants to change their screen names to their chosen pseudonyms, I began the recordings. Although Zoom® offered simultaneous transcription, I did not use it because I was not familiar with its relative strengths and weaknesses.

Quality of Data

One group had a dominant participant whose experiences were in opposition to those of another participant. The infrequency of me needing to manage overly active exchanges or to manage multiple participants speaking at once aligns with Lobe's & Morgan's (2021) similar findings from interviewing smaller groups in a virtual format. Participants appeared to feel comfortable with internal disagreement and there were multiple instances of a participant offering a divergent experience or perspective. More generally as well, the convenience of participating virtually allowed for more diverse participation and allowed me to show respect for the participants' time as professionals.

Participants

I aimed for focus groups of between four and six participants as recommended by Krueger and Casey (2009). Professional obligations of my prospective participants meant the final groups had between 2-5 members. The smaller size allowed me to better ensure that all members could share their thoughts and minimized redundancy. Each group allowed few perspectives, so I held multiple focus groups to ensure I reached saturation. People in institutional research roles made up three of the focus groups. The fourth group was made up of administrators in academic affairs or a similar division. I separated the participants by role to maintain perceived similarity among participants within each focus group and to target discussion questions specific to each groups' roles.

I recruited participants representing diverse professional career paths and institutions in what Teddlie and Tashakkori (2009) called maximum variation sampling. I sought to include institutional researchers with varying experience with statistical analyses and from institutions that varied on:

- Size of the institution by undergraduate headcount
- Sector
- Level
- Size of the institutional research office
- Main function of the institutional research office
- Data use culture of the institution-what information is used, how is it used, and how often

Although I could not determine before the focus groups where my participants and their institutional fell along these traits, my participants represented sufficient diversity as shown in Table 1.

Most participants represented institutions that awarded degrees of four or more years, including the community college that was represented. There was equal representation between public and private not-for-profit institutions. However, none of the prospective participants from private for-profit institutions chose to participate. It is unlikely that including participants from this sector would have altered my findings because participants from the private not-for-profit sector comprised half of the participant pool. Interestingly, most institutional research offices represented had a primarily institutional research function, although institutional effectiveness offices were well represented as well.

Table 1

Representation of Institutional Traits Among Focus Group Participants

Institutional Trait	Count	Percent
Institution Size: Fall 2021 UG Enrollment Reported to IPEDS (National Center for Educational Statistics, n. d.-m)		
<3,000	1	7
3000-9000	8	57
9,000-20,000	3	21
>20,000	1	7
Consortium	1	7
Sector		
Public	6	43
Private not-for-profit	7	50
Consortium	1	7
Level		
At least two but less than four years	1	7
Four or more years	12	86
Consortium	1	7
IR Office Size		
One person	2	14
Two to five people	9	64
Six to twelve people	2	14
More than twelve people	1	7
Main Function of IR Office		
Institutional Research	7	50
Institutional Research/Analytics	1	7
Institutional Research/Analytics/Effectiveness	1	7
Institutional Research/Information Technology	1	7
Institutional Analysis	1	7
Institutional Effectiveness	3	21

Note: One participant had previously worked in institutional research at an institution but worked at in institutional research at a consortium at the time of the focus groups.

Recruitment

I recruited using contact information published on the websites of institutions where at least one member of AIR worked and through referrals from professionals I had invited to participate. AIR is the main professional organization for institutional research professionals and membership demonstrates an institution's commitment to support institutional researchers in their role as student success data analysist. Among its mission objectives is that "AIR will educate and support higher education professionals in: Contextualizing data across campus and throughout higher education…Learning methods and tools of the institutional research profession", which closely align with the purpose of this study.

My exclusion criteria for the participants were that participants in both groups must have at least one year of experience in their role, be comfortable communicating in English, and comfortable using Zoom® videoconferencing software. They must also have had access to an appropriate electronic device. Participants in the institutional researcher group must have had a role in which they performed institutional research functions regardless of their position title as position titles in institutional research are eclectic. Participants in the administrator group must have had an upper-administrative role in academic affairs at their institution such as college dean, associate provost, or associate vice president/chancellor for academic affairs.

Incentives

For their participation, each participant had the choice of either receiving a $20 electronic Amazon gift card or I made a $20 donation to the general scholarship fund at their institution. I made the scholarship fund donation alternative available because the policies at some institutions preclude employees from accepting gifts or additional compensation related to their role at the institution.

Focus Group Protocol

Both sets of focus groups followed the same format. I used Zoom® to host and record the focus groups with consent of all participants. Before joining the meeting, I informed each participant that the meeting would be recorded, that they were to change their screen name to a pseudonym upon entering the meeting room, that they were welcome to have their camera of whenever they chose. I preferred it when they kept their camera on, but I also respected their interest in privacy. They were allowed to turn their mics off as well when they were not contributing.

Each focus group began with my introduction and the agenda. The discussion then moved from descriptions of the institutional research function at their institutions to the meanings they and their colleagues assigned to persistence and how they researched persistence. The protocols of questions for the first discussions are listed in Appendix D. I let the natural flow of the discussion guide which questions I asked; I did not ask each group all the questions and I asked incipient questions in response to the participants' contributions. After this part of the discussion, I gave a brief 15 minute slide presentation on my experiences and the meta-findings from running the three approaches that was tailored to the participants' roles and to their level of comfort with statistical analyses. The participants had greater familiarity with statistics than I had expected, but a few asked for further technical explanation after the presentation.

The second discussion focused on the benefits and limitations of the three approaches I presented within the context of their own institutions. The protocols of questions for the second discussions are listed in Appendix E. I concluded the focus groups by asking the participants whether they would be willing to review the list of themes from coding.

Transcription

I used the Otter® app version 3.25.0-6044 for Android for the initial transcription and corrected mistakes made by the software using the video component to supplement the audio and provide additional detail about the emotion of the speaker. I watched each recording three times before accepting the transcriptions as correct. I took the opportunity of first run, when Otter® transcribed the audio to watch the meeting and observe participants non-verbal interactions. During the second and third run throughs, I split my attention between the visual recording and verifying that the transcript matched the audio recording. As Pocock et al. (2021) warned, a few minutes of recording were incomplete or difficult to understand because of technical issues including bad signal, malfunctioning devices, and participants forgetting to turn microphones back on. For the participants who had their cameras off for part or all of their groups, I corrected Otter's® transcription to the best of my ability without the supplement of the visual data. The feature in Zoom® of highlighting the speaker's image was particularly helpful with these participants attributing the words to the speaker. I sent the transcript to each participant of the focus group giving them at least two weeks for review and made all changes they requested. Most participants did not request any changes and the changes that were requested were minor.

Researcher Journal

The second source of qualitative information I explored was my own experiences in running the quantitative analyses that I documented in a researcher journal. I maintained this journal through the qualitative phase as the focus shifts from my own experiences to the experiences of others. During the quantitative strand, I documented challenges I encountered in the analyses and my thoughts about challenges people with different backgrounds may have when running the analyses. I used the journal along with my quantitative output for the short presentation, so I specifically recorded how the quantitative findings vary across the three approaches and my perceptions on relative costs and benefits of each approach. In the qualitative strand, I documented my reactions and interpretations to what I experience in the focus groups in the journal to tease out connections between the focus group participants' experiences.

Data Analysis

I used an analysis strategy similar to the processes Krueger and Casey (2009) and Saldaña (2016, p. 98) described. I copied the text form all four transcripts and my researcher journal into an Excel® workbook for coding. In Excel®, I created a table of the qualitative data loosely following the layout common for databases with a column for the speaker, the comment number, the source of the text, and the comment. Initially, each comment was a new record and was identified by an integer unique and sequential within the source and the source name. However, some comments were quite long and included too many themes to code efficiently. I broke these comments into separate comments and updated the comment number to include a decimal value that kept the comments sequential and unique so that they were easier to code while maintaining context. I split the text from my researcher journal into one comment per paragraph and copied the journal entry date to the beginning of each applicable comment.

With the text in this table format, I read through the comments in order, adding columns to the table with themes as I identified them and marking the presence of that theme within the comment by entering an “x” into the coordinate cell of the theme column and the comment row. This process allowed me to code a comment with multiple themes and recode comments and themes evolved. This was particularly efficient when deciding whether to merge codes or to create new codes for more detail; I was able to filter the table and quickly identify comments with specific codes, seeing who the speaker was and where the comment was sourced from. I used in vivo coding to identify themes (Saldaña, 2016, p. 105). In my first read through the table, I created all relevant codes and split the comments as previously described. In my second read through the table, I assigned codes that I had created because of later comments to the earlier comments they applied to as well. From participating in the focus groups, watching the recordings and reviewing the transcripts three times, and reading the table twice for coding, I gained a deep understanding of the information that I used to create meta-themes to group the individual themes.

Inference Quality

Quantitative and qualitative research methods have each specific standards for ensuring inference quality. Whereas quantitative research focuses on the appropriateness of the analytical technique, respecting assumptions, and sufficient sample size, qualitative research focuses on trustworthiness. Lincoln and Guba (1985) gave four criteria for trustworthiness in qualitative research: credibility, dependability, confirmability, and transferability. My years of professional experience as an institutional researcher, working alongside institutional researcher and administrative members of the populations of interest for this study provide credibility to my qualitative findings. However, that same experience and familiarity made it more likely that I would lose my researcher role and excessively empathize with my participants. I minimized the distortions resulting from professional lens by documenting my experiences and thoughts in the researcher journal and triangulating the participant responses across focus groups and with my researcher journal. Triangulation provided the feedback I needed to correctly identify what my participants' lived experiences were, while also filtering out what they weren't (Schwandt, 2007).

The researcher journal also supports dependability, how well the research could be repeated and arrive at the same findings, because it is documentation of my processes. The quantitative findings documented in the researcher journal were the basis of the focus group presentations. These presentations structured peer debriefing during the focus groups as the participants, my peers, discussed my quantitative findings and how I presented them. Each focus group provided valuable feedback based on these presentations, in effect validating whether my findings were supported by their context.

Participants reviewed the transcripts as member checking, supporting confirmability, the criteria that the findings of the research were founded in the data and not the spontaneous creation of the researcher (Schwandt, 2007). Sharing comments among focus groups allowed members of later groups to vet the preliminary themes I identified from earlier ones and led to a richer discussion. I addressed the last criteria, transferability, by documenting my processes throughout and by only working with comments in their larger context during interpretation and when communicating the findings. I expect other researchers to use this surrounding information, the “thick description” described by Lincoln and Guba (1985, p. 316), to decide whether and how findings from this study apply to their situation of interest. Thick description is more than compiling descriptors of the study’s scenario, context, or happenings, but also includes initial interpretations of meanings and motivations as well (Schwandt, 2007). The findings from this study are only intended to generalize to other similar applications of the approaches described in the next section.

Quantitative Strand

The quantitative portion of this study was the application of three quantitative approaches to modeling persistence to a real data set. The three approaches I evaluated were proportions, logistic regression, and survival analysis. The most common approach in practice is univariate proportions, which have the greatest limitations. They are only descriptive and do not allow researchers to account for student attributes, nor for the nested data structure encountered in educational settings or the longitudinal nature of persistence. Logistic regression allows researchers to account for student covariates but ignores the impacts of time, the presence of competing outcomes, and the nested data structure. Survival analysis allows researchers to model time and account for competing outcomes, but it also ignores the nested data structure.

I received Institutional Review Board approval from the University of Northern Colorado under protocol number 2204037984 for the quantitative portion of this study.

Data

For the quantitative strand, I used the frozen student enrollment and graduation data from fall 2014 to fall 2022 stored in the student management system at Regional College for the college's new first-time full time fall cohorts for IPEDS, i.e. the eight cohorts: fall 2014 to fall 2021. To be included as enrolled in a term, the student must have been a degree seeking undergraduate student registered for at least one credit as of the date when the official data used for mandated reporting was frozen, generally one week after the end of the term. Due to the timing of my data request, the enrollment data for fall 2022 is as of the college's census date also used for mandated reporting, two weeks into the term. At Regional College, students may drop a course before a set deadline so that the registration doesn't show on their transcript, and they are refunded the tuition and fees for the course. If the student leaves the course after the drop deadline, the attempt shows on their transcript with the grade of "withdrawn" and the institution keeps the tuition and fees for that course. Depending on extenuating circumstances, students are occasionally granted late drops in which the student leaves a course after the drop deadline, but their attempt is removed from their transcript, and they may receive a refund of the tuition and fees. When these late drops are processed after the data is frozen, the student is still included as enrolled for mandated reporting, so for this analysis, I included them as enrolled as well. I included students who Regional College may exclude from their first-time student cohorts because of death, permanent disability, active military deployment, or religious missions according to the NCES definition. I did exclude students who are later removed from the official cohort because they had been incorrectly included (National Center for Educational Statistics, 2021a).

I evaluated the status of each included student from their first enrolled term until they left the dataset by either failing to re-enroll (left the college) or completing their degree (graduation). The outcomes of interest for the proportions and logistic regression were continued enrollment and completion, and for the survival analysis they were completion and leaving the institution. In all three analyses, I excluded students once they completed their degrees even though some re-enrolled at the undergraduate level. I included students who left the college but had re-enrolled by the term of interest as enrolled in that term, aligning with the definition of retention set by IPEDS (National Center for Educational Statistics, 2021b). I included gaps in enrollment as a set of indicator variables in the survival analysis as I explain in the variables section.

Five metrics are common in institutional research for assessing persistence: fall to spring retention, fall to fall retention and completion within 100%, 150%, and 200% of degree time. Term to term retention measures what proportion of students in the first term re-enrolled in the second term. The only undergraduate degrees that Regional College confers are 4-year bachelors, so the 100%-, 150%-, and 200%-time frames are the 4 year, 6 year, and 8 year graduation rates. The data set covers five cohorts to complete within 100%, three to complete within 150%, and one to complete within 200%. Following IPEDS reporting definitions (National Center for Educational Statistics, 2021a), I only considered fall starting terms and students who began in a summer semester were only included if they also enrolled in the immediately following fall semester. Students must have been new to the college as a new first-time student who had not attended any other higher education institution after completing secondary education.

Variables

To focus on the experience of modeling undergraduate persistence, I examined the following covariates. I selected them based on personal experience and on them regularly being used to explain differences in persistence.

High school GPA. This is the high school grade point average used in the student's admissions decision, whether that is the weighted or unweighted version. I truncated those GPAs with a value above 4.0 to 4.0 and I treated students who did not report a high school GPA, who completed a GED, who were home schooled, or who graduated from an institution that does not report grade point averages on a 4.0 scale as having a missing value for high school GPA.

Admissions score. This is the highest equivalent score reported by the student on the SAT® or ACT® college aptitude tests. For most students, the institution has either an ACT® score or an SAT® so that there have been excessive missingness if I were to use the highest score from each test for a student. I used the 2018 Concordance Tables to convert the SAT® Total to an ACT® Composite score. Then, I used the highest score among the ACT® Composite scores and the ACT® Composite score equivalent of the SAT® Total (ACT, 2022). A few years before this study, Regional College opted to no longer require these test scores for admissions decisions. A smaller, but notable proportion of applicants continued to submit them.

First-generation identity. Neither parent or guardian of a first-generation student has earned a bachelor's degree. Students at Regional College provided their first generation identity via the Free Application for Federal Student Aid® (FAFSA®) and via Regional College's application for admission. They generally complete the application for admission only once, but complete the FAFSA® annually, so it is common for a student to have indicated being both first-generation and non-first generation or for missing information to be later completed. For this variable, I followed Regional College's policy which was a student who has ever indicated that they are a first-generation student, regardless of the source, is a first-generation student. To capture the most complete data available, I used the students' first-generation identities as known to the institution as of fall 2022 instead of the identities as recorded in the official enrollment files.

Under-represented minority identity. In the application for admission and on the student portal, students are asked to provide their ethnic identities using the definitions established by the Office of Management and Budget in 1997 (National Center for Educational Statistics, n.d.-b). The college then categorizes students as belonging to only one of the following ethnic categories: Hispanic or Latino, American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Two or more races, Nonresident alien, and Race and ethnicity unknown. Students holding Hispanic or Latino, American Indian or Alaska Native, Black or African American, or Native Hawaiian or Other Pacific Islander ethnic identities are under-represented in higher education. To monitor progress toward enrollment and achievement equity while accounting for the small number of students who hold some of these identities, Regional College uses the meta ethnicity-classification of under-represented minority (URM). As with first generation status, the quality of the ethnicity information captured increases over time. So, I used the students' ethnic identities recorded as of fall

2022. Although I expected this variable to have substantial explanatory capabilities, included it with great caution. This variable combined the experiences of many unique communities, likely disguising important differences. Further, researchers harm students by misinterpreting parameter estimates, mistaking causality or causal order, and failing to understand the limitations inherent in ethnicity data.

Temporary United States resident status. Temporary U.S. residents, commonly called international students, are those who are present in the United States lawfully on a temporary basis under a visa. This does not include international students studying remotely from abroad, U.S. citizens or permanent residents, or students who have dual citizenship between the United States and another country. Temporary U.S. resident students experience challenges similar to domestic students but have additional financial and immigration challenges too. For example, they are ineligible for most federal and state financial aid and may be restricted in how many hours per week they can work, and they pay out of state tuition at Regional College even if they graduated from an in-state high school. Students' immigration status rarely changes, so I used the status as of their first fall for this variable.

Number of credit hours earned as of the beginning of the term. Increasingly, students enter higher education with advanced standing (college credit earned while in high school). By extension, the proportion of students who begin their higher education careers as a sophomore is also increasing and the number of credits students need to complete their bachelor's is decreasing. Theoretically, students with advanced standing are closer to earning their degrees making this an important covariate for understanding persistence, but previous credits may not apply efficiently to students' degrees. I used the student's number of credits earned as of the beginning of the term as a term dependent variable ranging from 0 to over 200.

Credits attempted in current term. Students are expected to take 15 credits in each fall and spring term and 0 credits over summer to complete their bachelor's degree within four years. Students who attempt fewer than these 15 credits per fall and spring term are at increased risk of not completing within the four years, and students who take at least 15 credits in their first semester have better outcomes (Davidson & Blankenship, 2017). In the proportion and logistic regression analyses, I used the student's credit load in their first term. In the survival analysis, each record included a variable of the student's credit load in that term.

Female. Students may provide their sex on the admissions application in a voluntary question. Although an option, very few students at Regional College have identified as other than as male or female. The small count and proportion of students who select this third option is so small that including it would have caused computational issues in the logistic and survival analysis models. As a balance of minimizing these issues while partially respecting the student's autonomy, I included sex as an indicator for whether the student identified as female; students who did not identify as female were estimated in the baseline rates and the variation associated with identifying as female was estimated by the parameter. Students may update their identified sex, but for few students does the sex information vary over time. I used the most recent sex identity information as of fall 2022 and treated sex as a constant throughout the study window.

Pell eligibility. This college measured financial need using the federal guidelines for Pell Grant eligibility. Eligibility was based on calculations within the FAFSA® of personal and family assets, so there were limitations to using it to identify students with financial need. Students who are not eligible for federal financial aid, such as temporary U.S. residents and other students without U.S. citizenship or permanent resident status, rarely complete the FAFSA®. There were also concerns that the application itself was onerous and confusing to some applicants and their families. Accordingly, not all students with financial need are identified as such in this dataset. I sorted students into those with known Pell grant eligibility and those with unknown eligibility or known ineligibility. Students must reapply for federal financial aid each year, so their Pell eligibility is coded by term.

Enrollment. Students must meet the definition of enrolment used for reporting to IPEDS as of the final freeze date for terms before fall 2022 and the census freeze date for fall 2022 (National Center for Educational Statistics, 2021a). Each fall and spring term a student is enrolled was indicated with a record in the dataset. In the survival analysis, two additional variables captured enrollment information. A cumulative variable indicated the number of fall and spring terms the currently enrolled student had not been enrolled since their initial fall term. Another variable indicated the length in fall and spring terms of the student's most recent gap in enrollment.

Degree Completion. Students must have fulfilled all of the academic and non-academic requirements for their degree program and have applied for graduation to count as having completed their degree and the graduation term is the one entered on the student's degree record even if the student was not enrolled that term. In the rare situations when a student is not enrolled in their graduation term, it is most often due to non-coursework requirements being incomplete as of the date degrees were awarded. An example of missing requirements is if the student neglects to drop a minor or second major that they did not intend to complete. Because these issues are not related to the student's enrollment I coded the last enrolled term for these students as the term in which they completed their degree.

I used the college's frozen data for mandated reporting for these analyses. Before being frozen, they were used in daily operations which highlighted most of the errors and necessitated their correction. The offices responsible for the data maintained data quality by regularly running data integrity and validation reports which alerted them to less-routine errors in the data. The exceptions to these validation checks are the sex, under-represented minority, and first-generation status fields, which are students' self-reported information; these fields could have only been validated by the students themselves. By using this data, I insured consistency between my calculations and those produced internally by the college but at the cost of accepting that errors discovered after the data was frozen remained errors. The consequences to the college of inaccurate reporting are dire, so it is unlikely that substantial errors were present. By dint of using this dataset and the previously listed definitions, admissions score is the only field with missing data.

Preliminary Diagnostics

I used R version 4.2.2 through R Studio version 2022.07 for all analyses, which I began with diagnostic tests on the data. For all statistical testing in this study, I used the $\alpha = 0.05$ significance level. When making multiple comparisons within the same model, I used the Bonferroni adjustment of $\leq \frac{0.05}{n}$, where n is the number of comparisons within that model. Typically, the Bonferroni adjustment accounts for all comparisons, not just those within a particular model. To simulate how these models are ran in practice, I only corrected for the inflated type I error rate within each model.

I verified that, in accordance with the literature and my experience, each independent variable had a sufficiently strong relationship with the outcome variables to warrant inclusion. The outcome variables I used were the students' enrollment status in their first spring and second fall term and whether they had completed their degree within four or six years. I treated the admissions scores, credits attempted, and the number of credit hours earned as of the beginning of the term as continuous variables, so I measured the association between these and the outcome variables with point biserial correlations. The remaining independent variables were dichotomous, so I used tetrachoric correlations to measure their associations with the four outcome variables. I retained all variables that had an association larger than .7 with any of the four outcome variables, which resulted in retaining all of the variables.

To check for multicollinearity, I estimated a simultaneous entry logistic regression model for each of the outcome variables using all of the covariate variables. Logistic regression (with the logit link function) was developed to analyze data with categorical outcomes because these data violate assumptions of traditional ordinary least squares regression (Peng, Lee, et al., 2002). It replaces the assumption of constant error variance and normally distributed errors with the assumption that the errors (observed value minus the predicted value) and that the outcome variables follow the binomial distribution. Parameter estimates that are abnormally high or low are suggestive of multicollinearity.

R provides the raw parameter estimates and odds ratios for easier interpretation. Odds and probability are sigmoidally related; probability is the likelihood of a specific event occurring divided by the likelihood of any event occurring and odds are the ratio of the probability of an event happening and the probability of the event not happening. Odds ratios are symmetric, unlike a ratio of probabilities, meaning that one can work backward from a known outcome to interpret correlated variables without having to include an entire population. The explanation given by George et al. (2020) can be adapted using a persistence odds ratio of 1.5 for students who were Pell eligible. This odds ratio means that students who were Pell eligible had 50% higher odds of persisting and that students who persisted had a 50% higher chance of being Pell eligible.

I used the Wald test, calculated as $W_j = \frac{\widehat{\beta_j}}{\widehat{SE(\beta_j)}}$ and following the X^2 distribution, to determine the variables' contribution to the model (Tabachnick & Fidell, 2007, p. 445). Using previous experience and findings reported by other researchers, I reviewed the parameter estimates for reasonableness and the standard errors of the estimates for their stability. Using a plot the residuals against the predicted probabilities, I identified a few outliers which did not appear to be the result of a data entry error, so I retained them in the dataset. After completing these preliminary diagnostics, I ran three sets of analyses on the dataset: proportions analysis, logistic regression, and survival analysis.

Proportions

In this proportions analysis, I calculated 37 sets of proportions for each level of the discrete covariate variables: first-generation identity, under-represented minority identity, temporary US resident status, female, and Pell eligibility to identify across subgroups. 16 sets measured one of the retention rates. Nine sets measured one of the freshmen completion rates, and twelve sets measured one of the junior completion rates. The initial and follow up terms for the freshmen completion rate models are listed in Table 2.

Table 2

Initial and Follow-Up Terms for the Freshmen Completion Rates.

New first-time student in term	Follow up: 100% (4 years)	Follow up 2: 150% (6 years)	Follow up 3: 200% (8 years)
Fall 2014	Fall 2018	Fall 2020	Fall 2022
Fall 2015	Fall 2019	Fall 2021	
Fall 2016	Fall 2020	Fall 2022	
Fall 2017	Fall 2021		
Fall 2018	Fall 2022		

Junior completion rates are a modification of the more commonly used completion rates calculated within IPEDS (freshmen completion rates) (National Center for Educational Statistics, 2021a). They are the proportion of students from an initial cohort who reached junior standing (60 earned credit hours) by their third fall, who then complete their degrees within a set time frame. Nearly all of Regional College's bachelor's degrees are 120 credits long, so the students in the base population of the junior completion rates had earned at least half of their required credits within two years. With just half of their credits remaining, the times allowed to complete within 100%, 150%, and 200% are also halved to 2, 3, and 4 years. These junior completion rate models highlight effects on students' timelines when they're most vulnerable to disruption; a freshmen's timeline is less sensitive to a required class being cancelled than a senior's timeline. Junior completion rates also estimate the variables' relationships based on a shorter and more actionable time frame, allowing information users to account for or respond to cross-sectional effects such as pandemic disruptions or policy changes. The initial and follow up terms for the twelve junior completion rate models are listed in Table 3.

Table 3

Initial and Follow-Up Terms for the Junior Completion Rates.

New first-time 2 years prior	Follow up 1: 100% (4 years)	Follow up 2: 150% (6 years)	Follow up 3: 200% (8 years)
Fall 2016	Fall 2018	Fall 2019	Fall 2020
Fall 2017	Fall 2019	Fall 2020	Fall 2021
Fall 2018	Fall 2020	Fall 2021	Fall 2022
Fall 2019	Fall 2021	Fall 2022	
Fall 2020	Fall 2022		

I calculated retention as the student re-enrolling in a following term. Following convention, I used only the spring and second fall as destination terms and ignored enrollment behavior in later terms. Because of the timing of my study, the final official enrollment data for fall 2022 was not yet available, so I used the enrollment data as of Regional College's official census reporting date instead. The initial and follow up terms for the retention models are listed in Table 4.

Table 4

Initial and Follow-Up Terms for the Retention Rates.

New first-time student in term	Follow up 1: First spring	Follow up 1: Second fall
Fall 2014	Spring 2015	Fall 2015
Fall 2015	Spring 2016	Fall 2016
Fall 2016	Spring 2017	Fall 2017
Fall 2017	Spring 2018	Fall 2018
Fall 2018	Spring 2019	Fall 2019
Fall 2019	Spring 2020	Fall 2020
Fall 2020	Spring 2021	Fall 2021
Fall 2021	Spring 2022	Fall 2022

Survival Analysis

The third approach I used was a discrete time competing risks survival analysis. Survivor analysis was developed to model mortality from disease over time, so the associated terminology has a negative connotation. I modeled student's continued enrollment at the institution and my outcomes were of the student departing from the institution before completing their degree or completing their degree. Departure is commonly perceived as negative, like mortality, but degree completion is consistently a positive outcome. Many modern applications of survival analysis are similarly faced with using terminology with a negative connotation when modeling a positive outcome.

The strength of survival analysis is that it allows researchers to follow a group of students until each has experienced an event or the follow up period ends. In my analysis, I followed the fall starting new first-time students from terms fall 2014 through fall 2021 until fall 2022, so that I observed the students who started in fall 2014 for up to eight years and the fall 2021 starting students for up to one year. Over 80% of students experienced one of the two events, with more students from the later cohorts still enrolled as of fall 2022. I measured time as the number of terms the student was enrolled, so that I was able to calculate the risks according to the student's time at the institution, in line with how retention and graduation rates are conceptualized from the proportions section. Just as with the logistic regression analyses, survival analysis adjusted my expectation for each student as I included additional information; a student who entered with advanced standing was expected to have a higher likelihood of retaining than one who did not.

Looking at the events longitudinally, I calculated their hazards and hazard ratios. Hazard is the rate the event occurs based on the number of remaining students who have not experienced an even before that time. In continuous time analyses, the hazard is the instantaneous rate of occurrence, but in discrete time analyses like mine, the hazard is a ratio of the number of occurrences per time unit i.e. completions per term (Cox, 1972; Singer & Willett, 1993). The completion hazard is the proportion of students who were still enrolled (i.e. had not yet withdrawn nor completed before that term) who completed in that term. The relationship between a variable, like advanced standing and the completion hazard is measured with a hazard ratio. For dichotomous variables, hazard ratios are the ratio of the hazard for members of one group divided by the hazard for the members of the other group. For continuous variables, hazard ratios measure the increase in hazard associated with a one unit increase in the variable. Interpreting hazard ratios requires the assumption that the two hazards are constant over the time segment, the proportional hazard assumption.

This can get confusing because the proportional hazards assumption for the Cox model only applies to the hazards associated with the covariates. The baseline hazard function does not need to have a fixed slope and the researcher can choose from a variety of distributions to represent the baseline hazard function (Jóźwiak & Moerbeek, 2012). This distinction means that having substantially more completions per term after the eighth term is irrelevant to the proportional hazards assumption, because this pattern is the baseline hazard function for all students. Applying a model with the proportional hazards assumption to data that violate this assumption complicates model interpretation, results in a substantial loss of power, and easily leads to incorrect conclusions (Hess, 1995) (Singer & Willett, 1993). At the recommendation of Singer and Willett (1993), I first estimated an intercept only set of discrete time models that included the indicator variables for each term to isolate the baseline risk of withdrawing or completing over time.

The inverse of the hazard function is the survival function, which describes the likelihood over time that the event will not happen given that the event had not yet happened. The height of the survival function is the survival probability and the time that passes before the event is the survival time. If the median survival times are not equal across groups, the proportional hazards assumption has been violated (George et al., 2020). In my analysis, the risk set is all students still enrolled at the beginning of the term. These students are still eligible to experience an event. Once a student experienced an event, they were removed from the risk set.

Sample Size

The minimum sample size requirements for the survival analysis are the same as those for the logistic regressions apply because discrete time survival analysis is based on logistic regression, differing only in structure and interpretations. However, the longitudinal component of survival analysis leads to more missingness than conventional logistic regressions, which may affect same size needs. Unlike logistic regression, survival analysis has the assumption that the outcome it is assumed to happen after the study ended if it is not yet observed by the last follow up time point, allowing a portion of the process to happen outside of the study window. This unobserved portion of the process is censored (missing) from the analysis. For example, my analysis did not capture the last years of a student who started as a new first-time freshmen in fall 2021; their academic career is right censored. Some students were simultaneously registered at more than one institution causing a form of interval censoring in my data set.

Peduzzi et al. (1995) explored the impact of sample size and number of observed events on power in continuous time survival analyses and found that a minimum of 10 events must be observed for each covariate to minimize bias in the parameter estimates. Sample size relates to the power of the analysis in two ways: the number of participants and the number of time points the participants are followed for. Increasing the number of participants or the number of time periods both increase the power. Power is also sensitive to the difference in the proportion of participants who experience the event across groups and how close the overall proportion is to either 1 or 0. (Jóźwiak & Moerbeek, 2012). About 35% of the students in this sample of over

14,000 students followed for up to seventeen timepoints completed and about 45% withdrew, so the power of each model was lower because of the baseline proportions but was likely still adequate. In their simulations, Jóźwiak and Moerbeek (2012) achieved a power of .8 from a dataset of 400 participants with an overall proportion near .5, just four time points, and a difference between control and treatment group proportions of .15.

Analysis Functions

I used the same software for the survival analysis that I had used for the logistic regressions, that being the glm function in the stats package in base R version 4.2.2 to estimate the models and the log likelihood test and McFadden's R2 from the pR2 function in the pscl package version 1.5.5.1 and the Hosmer Lemeshow from the hoslem.test function in the ResourceSelection package version 0.3-5 to assess model fit (R Core Team, 2022; Jackman, S., 2020; Lele, S.R., Keim, J.L., Solymos P., 2019).

Assumptions

As with conventional logistic models, discrete time survival analysis assumes linearity between continuous variables and the logit transform of the probability, here hazard. I checked this assumption visually by breaking the continuous variables into subgroups and fitting models to each of the groups separately as recommended in Singer and Willett (1993). If the proportional hazards assumption had been violated, the respective hazard functions would not have been roughly parallel and would not have been spaced roughly proportional to the width of the grouping.

Both discrete time survival analysis and logistic regression assume that the continuous covariates are linearly related to the link function, and here again I used the logit link. To evaluate this assumption and to check for outliers, I again used smoothed scatterplot approach and residual graphs (Hosmer et al., 2008, 2013).

Modeling censored data requires extending the independence of observations assumption to include independence of censoring, also called non-informative censoring. Non-informative censoring when modeling completion means assuming that all students will eventually complete a bachelor's degree, which is a faulty assumption. In a model with only the completion outcome, the censoring event would provide information about the student's completion risk. Specifically, withdrawing from the institution would result in censoring and would also influence the student's completion risk. I addressed this issue by using the competing risks approach, ensuring that the parameter estimates were more accurate (Prentice et al., 1978).

Model Fit

I used the log likelihood test of fit applied to the survival analyses as I did with the logistic regressions, and it was particularly useful because it is blind to whether the added parameters were time varying or invariant (Singer & Willett, 1993). To check the fit of the time component, I followed the recommendations of Singer and Willett (1993) and interpreted the fitted hazard functions. I also compared the fitted hazard curves to the observed hazards of the students when sorted by the subgroupings listed in Table 5.

Table 5

Subgroupings for Model Classification Rate Analysis

Subgrouping	Members of Part A:	Members of Part B:
A	under-represented minority ethnicity students	non-minority ethnicity students, well-represented minority ethnicity students, students who did not provide an ethnicity, and student who were a temporary United States resident
B	female students	male students and other non-female students
C	students whose high school grade point average (GPA) was less than a 3.0 on a 4.0 scale	students who did not provide a high school GPA or whose high school GPA was at least 3.0 on a 4.0 scale
D	students who received Pell grant financial aid	students who were known to be ineligible for Pell grant financial aid and students whose Pell grand eligibility was unknown

Joining Quantitative and Qualitative

As Teddlie and Tashakkori (2009) explain, mixed methods research faces the challenge of joining disparate paradigms. I explored the meanings and implications of the participants' realities, focusing on the complexities and interactions of using the three quantitative approaches, in line with Lincoln and Guba's explanation of constructivism during the qualitative strand (1985), and I used the post-positivist paradigm during the quantitative strand. To develop a cohesive and contiguous whole, I integrated the strands using the pragmatic paradigm. Using the pragmatic paradigm adds value to this research because it acknowledges that reality is too complex to be captured by one paradigm. The quantitative strand of applying the three approaches was nested within the overall qualitative strand of describing their value. The findings from the quantitative strand informed the focus group questions, were presented in the focus groups, and contributed to the experiences I brought to the discussions.

In chapter III I described the methods I used to collect and analyze the qualitative and quantitative data for this study. I used three quantitative approaches to model persistence using a real dataset, then used my findings from that experience as recorded in my researcher journal to inform the four focus groups I held over Zoom®. In chapter IV I summarize my findings from the two strands.

CHAPTER IV

FINDINGS

Through the contributions of fourteen focus group participants and my own experiences as recorded in the researcher journal, this mixed methods study used qualitative methods to examine the value of three quantitative approaches to modeling binary outcomes in institutional research. The findings from the quantitative phase fed into the qualitative phase by directing the focus groups discussion and informing my researcher perspective. I compared the three quantitative approaches on their model fit, classification accuracy, the classification accuracy for subgroups of interest, utility of the findings, and the complexity of data cleaning, analysis, and interpretation. Following the order of data collection and analysis, the quantitative findings are presented first.

Quantitative

Proportions

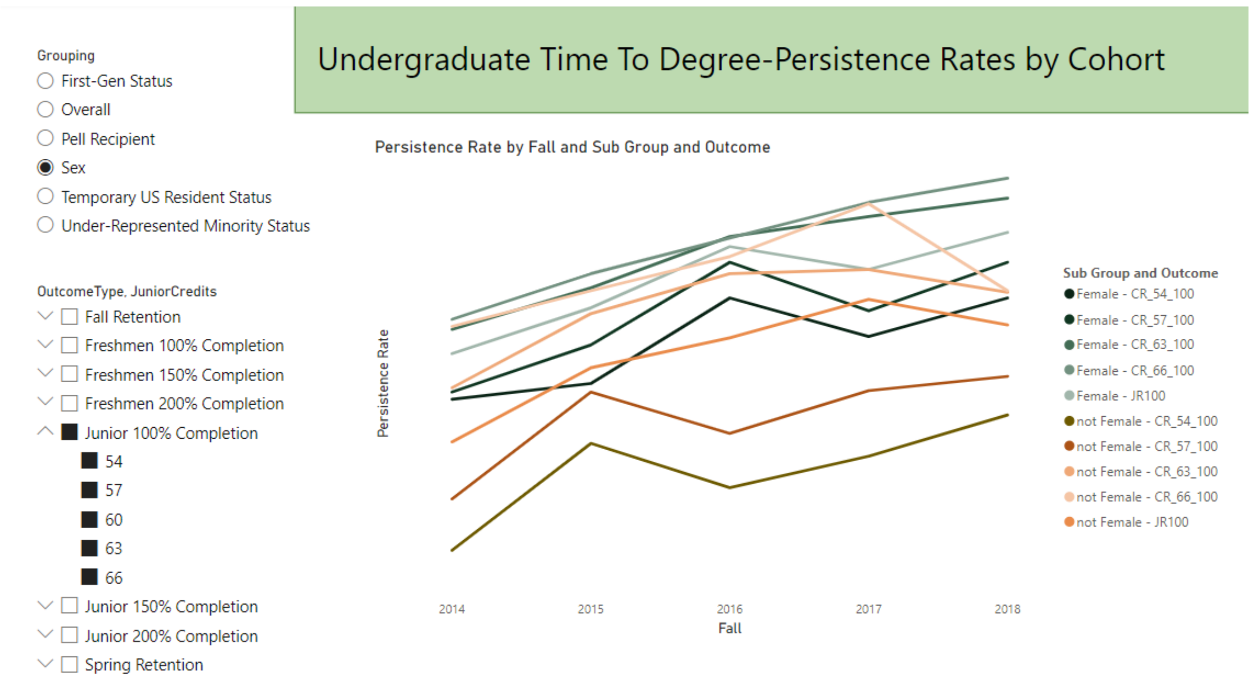
I split the dataset using the two values of each dichotomous variable and calculated the persistence rate for each group. For example, I calculated the retention rate for female students and the retention rate for students of any sex other than female; these two retention rates constituted a pair that I interpreted together as a means of quantifying the differences related to sex. There were five pairs: female students contrasted against students of any other sex, first generation students contrasted against students whose first generation status is unknown and non-first generation students, students of an under-represented ethnic minority contrasted against students of a well-represented ethnic minority or the ethnic majority, students who received Pell grant aid contrasted against those who did not, and students whose high school GPA was below 3.0 on a 4.0 scale contrasted against those whose high school GPA was higher or who had no high school GPA. I calculated the outcome rates for each persistence outcome over the eight cohorts, which in total yielded 37 sets for each of the five pairs. In total, I calculated 370 persistence rates. Many of the findings were expected; in all cohorts, first generation students had lower graduation and retention rates than students who were not first generation and female students most often had higher graduation rates than students of any other sex. The junior graduation rates were higher than the freshmen graduation rates; this held for students regardless of their group membership across the five binary variables, suggesting a survivorship bias in the junior graduation rates. The students still enrolled as of their junior fall were likely demonstrably different than those who were not so that whatever enabled these students to persist to their junior fall was also likely to influence their likelihood of graduating--a missing covariate.

There were some unexpected findings though. Although the Temporary U.S. Resident group was very small, it tended to perform better than the group of U.S. residents and those present in the U.S. under another standing. The four-year rate gaps shrunk by up to 4 percentage points between the freshmen and junior rates for the First Generation, Female, and Pell Eligible variables. However, the 6-year graduation rate gaps shrunk by up to 9 percentage points for the First Generation, Pell Eligible, and the Underrepresented Minority variables. These narrowing gaps suggest that the elapse of time effects the disadvantaged and advantaged groups differently. A smaller proportion of students in the disadvantaged group of each variable were still enrolled, a 6 to 9 percentage point shift in demographics, but these shifts do not fully account for the closing gaps.

These patterns warranted further investigation within the proportion approach. I used Power BI® to visualize the trends of five additional grad rate base populations: students enrolled in their third fall term who had earned at least 54, 57, 60, 63, or 66 credits. Figure 1 shows the visualization in which I selected the grouping, outcome of interest, and level of credits earned using the parameters on the left and the legend on the right.

Figure 1

Power BI® Visualization of Persistence Rates from the Proportions Approach.



Not surprisingly, overall and within each subgroup, students who earned fewer credits by their third fall had lower graduation rates. What was surprising was the overlap of graduation rates within the variables. For example, from the 2014 cohort to the 2017 cohort, non-first gen students who earned at least 54 credits by their third fall had the same graduation rates as first-generation students who had earned at least 66 credits. Also, the graduation rates for non-female students had a larger range among the credits earned levels, but there was negligible difference between the sex subgroups for those who earned at least 66 credits. Female students had higher achievement in each level of credits earned, but the non-female student's junior graduation rates were more sensitive to credits earned. Further inquiry is needed to determine whether programming targeted toward non-female students enrolled in their third fall with at least 54, but fewer than 60 credits earned would impact these graduation rates. This is a small group of students, but there is a noteworthy gap between their graduation rates and the rates of similar students who had earned 60 credits.

Logistic Regression

Sample Size

The next step in modeling binary persistence outcomes was to estimate the 37 logistic regression models. The new first-time cohort (including both full- and part-time students) for Regional College ranged from 1,200 to 2,100 students for the falls included in the data set. Using the minimum sample size recommendations by Tabachnick and Fidell (2007) of a ratio of at least 10 observations for each variable and at least 100 observations, this sample allowed for 120 variables in each model. This far exceeds the 10 variables I was interested in, so I determined that my data set was large enough for the logistic regression analyses.

Using another recommendation for calculating the necessary sample size for a multivariate logistic regression, that of having at least 500 observations and no more than i number of observations based on n number of observations where $n = 100 + 50i$, the enrollments supported estimating 22 parameters (Bujang et al., 2018). Peng, So, et al., (2002) recommends the sample size guidelines from Lawley and Maxwell (1971), explaining that the sample size requirements of the estimating procedure, maximum likelihood, are the limiting factors. Specifically, they used the rule of no fewer than 100 cases, with at least 500 as ideal, and at least 10 observations per parameter so that there are at least 51 more cases than the number of estimated variables, resulting in at least 50 degrees of freedom. As yet other, similar sample size guidelines, Thorndike (1978) suggested two other thresholds of at least 50 cases plus 10 times the number of estimated parameters and at least 50 times the square of the number of estimated parameters. By these three guidelines as well, there were sufficient enrollments to support estimating all the relationships I was interested in.

To estimate the logistic regression models, I used the glm function in the stats package in base R version 4.2.2 (R Core Team, 2022). I did not encounter convergence issues, so there was no need to use the glm2 function instead of glm (Marschner, 2018). I used powerLogisticBin to run an a priori analysis of 1,800 new first-time students who had a fall to fall retention rate of 70% and found that it had the power of .830 to detect an odds ratio of 1.10 or a retention rate of .748 for members of one group and .68 for the other (Qiu, n.d).

There were few temporary U.S. resident students in some cohorts, making the estimates for that parameter unstable. These small cell sizes only affected the one variable and did not extend to more than 20% of the cells so I did not need to collapse categories to increase cell sizes (Hosmer et al., 2013, p. 146) (Tabachnick & Fidell, 2007, p. 442). Complete separation happens when the model has overfit the data and results from insufficient sample size or poor sample selection; it was not present in these data (Hosmer et al., 2013, pp. 147-9).

Assumptions

Logistic regression assumes that the continuous covariates are linearly related to a transformed version of the likelihood of the outcome occurring where the nature of the transformation is determined by the link function used. The decision between using the probit link function and the logit link function is driven by the assumed distribution of the event and the resulting errors from the model (Peng, So, et al., 2002). I chose the logit link function because it assumes that the estimated outcomes and errors follow the standard logistic distribution, while the probit and tobit link functions assumes that they instead follow the standard normal distribution. The link function chosen also affects the interpretation of the parameters so that parameter coefficients in a model with a logit link function are the change in the log-odds of the event, while in a probit model, it is equated to a change in the z-score.

Dependent variables in models with the logit link function are assumed to be ordinal categorical, aligning with definitions I used for retention, completion, and persistence. It is convention to assume that the observations in datasets like the one I used are independent. However, in practice there are subgroupings, like major, that segment students who have more in common with others in their group than with students in other groups. These group differences may have affected variables I included in the model, such as the under-representation of female and URM students in science, technology, engineering, and math majors. In this study though, I followed convention and assumed independence of observations.

By reviewing smoothed scatterplots as explained in Hosmer et al., (2013) I assessed whether my continuous covariates were linearly related to the logit transforms of my outcome variables and identified an issue with the high school GPA variable. To obtain a linear relationship, the high school GPA variable had to be raised to the fifth power. The relationships between the other independent and dependent variables were approximately linear, so I did not transform any other variables.

I also used scatterplots to identify outliers which may have unduly influenced the models. Pearson residuals are more intuitive, but the distribution is skewed in many applications, so I used the deviance residuals to identify outliers (Davison, 1989). Following the advice of Stage

(1990), I used a criterion of 2.5 for minimum deviance ratio to indicate good fit, although as of 2002, this had only been validated for applications using structural equation modeling (Peng, So, et al., 2002).

I examined the same measures of association I outlined in the preliminary diagnostics section to check for multicollinearity among the variables. These could only identify bivariate collinearity, so I also examined the parameter estimates, the β's, in the models. In the presence of multicollinearity, parameter estimates often take on counterintuitive values, being smaller or larger than expected. I elected to examine parameter estimates instead of odds ratios because odds ratios are not scalar.

There is no way to test the assumption of no unobserved heterogeneity--the assumption that all variation is due exclusively to the variables included in the model. Although other research has identified explanatory variables that I omitted, I proceeded as if all important variables were included. If I did omit an important variable, the analysis would have pooled the profiles resulting from that variable, yielding profiles substantially different than the ones I estimated. More concerning is that if the omitted variable had no cases at its mean, the pooled function would represent an "average" that applies to no one in practice (Singer & Willet, 1993).

Model Fit

Researchers have many options for assessing model fit, which Peng, Lee, et al. (2002) and Peng, So, et al. (2002) group into four categories:

- compare the accuracy of predicted outcomes and simply assigning all observations to the outcome category with the highest frequency
- assess whether individual predictors statistically significantly contribute to the effectiveness of the model

- compare how the predicted outcomes coincide with observed outcomes
- assess whether participants who experienced the event have higher expected probabilities than participants who did not experience the event

In ordinary least squares (OLS) linear regression, the R^2 statistic measures the proportion of variation accounted for by the variation in the model's covariates and is directly tied to the model's overall significance. However, logistic regression does not have such a statistic because of how the parameters are estimated. Researchers have proposed many statistics similar to the R^2 coefficient of determination from OLS regression, which are also called R^2. These pseudo R^2 statistics describe the overall fit of the model, but none directly measure the proportion of variation explained by the model (Menard, 2000). The pseud R^2 I used was derived by McFadden (1974) and is calculated as $\frac{(-2\ log\ likelihood_{initial} - -2\ log\ likelihood_{model})}{-2\ log\ likelihood_{initial}}$. McFadden's index has the advantage over other R^2 alternatives that it can be used to compare non-nested models as long as the comparison models have the same outcome variable and are based on the same dataset (Menard, 2000).

After verifying that the models were overall statistically significant and that each of the retained parameters contributes significantly to the model, I evaluated the models' goodness of fit and correct classification rates. I used the log likelihood test and McFadden's R^2 which I obtained from the pR2 function in the pscl package version 1.5.5.1 along with the Hosmer Lemeshow test that I obtained from the hoslem.test function in the ResourceSelection package version 0.3-5 to assess model fit (Jackman, 2020; Lele, Keim, & Solymo, 2019). The log likelihood gains between the null and full models were negligible. McFadden's R^2 values were uniformly low. None of the Hosmer Lemeshow tests were statistically significant. All of these indicate that the models had poor fit. However, persistence models are also used to classify students, so I evaluated classification accuracy.

Classification accuracy measures how well predicted and observed outcomes coincide overall. At the individual level, the outcome predicted by the model can match the observed outcome: the event was predicted and occurred (true positive) or the event was not predicted and did not occur (true negative). Alternatively, the predicted and observed outcomes can be different: the event is predicted but did not occur (false positive) or the outcome was not predicted but did occur (true negative) as shown in Table 6.

Table 6

Relationships Between Expected and Observed Outcomes.

Outcome Prediction	Observed Outcome	
	Outcome observed (Positive cases)	Outcome not observed (Negative cases)
Predicted to happened	True Positive	False Positive
Not predicted to happen	False Negative	True Negative

However, the dependent variable of logistic regression with the logit link function is a continuous value, the estimated likelihood of the event occurring, not a binary value of whether the event will occur. Researchers must select a threshold value, dividing the likelihood estimates into predictions that the event will occur or not. Sensitivity is the proportion of positive cases that are true positives and specificity is the proportion of negative cases that are true negatives. The sensitivity and specificity of the model depend on the threshold the researcher selected. A threshold for a graduation model that is too low will correctly predict more graduated students in the “did graduate” group, but it will also incorrectly predict many students who did not graduate in the “did graduate” group. Conversely, a threshold that is too high will predict more students not to graduate, including more of those who did graduate.

Peng, Lee, et al. (2002) found that the algorithm used to sort observations for a contingency table varies by software package, leading to differences when the same model is fit in different packages. Following the advice of Hosmer et al. (2013, p. 175) I visually identified the optimal threshold value for each model by comparing the sensitivity vs. threshold and specificity vs. threshold plots, selecting the threshold where the two lines intersect. Using these thresholds, I also compared the classification accuracy for the subgroups listed above in Table 5 to the overall classification accuracy to identify whether these models might have perpetuated inequalities among student identities.

The parts of each subgrouping that I examined were those that tended to have poorer persistence rates and tended to be the focus of more administrative attention at Regional College. Because of this additional attention, the college may feasibly have been more sensitive to changes in persistence rates for these groups, so it would be more important that any changes in persistence rates be properly attributed.

I then examined each model's receiver operating characteristic curve (ROC curve), a plot of the probability of detecting true signal or a false signal. Plots of sensitivity and specificity against threshold values identify the best threshold, but the ROC curve, and the area under the curve (AUC) show how well the sensitivity and specificity compare to the ideal scenario in which false positive rate of near zero occurs at the same low threshold as a true positive rate of near one. The baseline for the ROC curve is a straight line from the lower left corner to the upper right indicating that the model is no better than random chance at predicting the outcome. The ROC curve for models with better predictive ability deviates further toward the upper left corner. The area under the ROC curve (AUC) ranges from .5 to 1, increasing as the ROC curve deviates further from baseline, indicating better ability to distinguish true from false positives.

Student Persistence

The primary motivation for escalating to this approach is to learn whether the differences I observed in the persistence rates were likely due to random chance. Where I found evidence that there was a statistically significant relationship between the variables of interest and the odds of persisting, I examined the coefficients to determine the nature of that relationship. By initially modeling the outcomes with individual variables such as the odds of retaining to the next fall depending on students' first-generation status, I found that temporary U.S resident status did not explain the odds of retaining or completing their degree. Similarly, most variables I had found to be significant in the univariate models were not so in the multivariate models. Under-represented minority status was not significant in any of the multivariate models, and Pell eligibility was only significant in eight of the graduation models. Interestingly, these models were the four freshmen graduation models and their corresponding junior graduation models.

The log likelihood gains between the null and full models were negligible in all 37 multivariate models and McFadden's R^2 values were uniformly low. However, only two of the models had statistically significant Hosmer-Lemeshow tests; those groups with higher predicted probabilities also had higher observed probabilities. These findings suggest that the models provided little additional explanatory power over their null models although the predicted probabilities do somewhat align the observed probabilities. Interestingly, the classification accuracy of the models on the full testing datasets was relatively consistent within each outcome. As shown in Table 8, the classification accuracy ranges are similar for the overall testing dataset and the under-represented minority ethnicity, all sexes other than female, and Pell eligible subsets of the testing data.

Table 7

Accuracy Ranges of Logistic Models by Outcome and Testing Data Subset

	Overall Percent	Under-Represented Minority Ethnicity Percent	All Sexes Other than Female Percent	High School GPA below 3.0 Percent	Pell Eligible Percent
Spring Retention	88-89	85-88	87-88	79-81	88-89
Fall Retention	70-71	66-69	71-74	55-63	71-74
Freshmen 100%	72-73	70-71	70-72	89-89	72-75
Freshmen 150%	68-69	69-73	69-70	74-78	69-72
Freshmen 200%	67	70	69	71	69
Junior 100%	72-73	70-73	69-71	89-89	70-75
Junior 150%	68-70	66-72	69-71	76-79	66-71
Junior 200%	68-69	71-73	69-71	76-79	69-72

The subset with the biggest difference in classification accuracy from the overall testing dataset was the group of students with a high school GPA of less than 3.0, subgroup C. The models with the two retention outcomes were markedly less accurate in classifying the low high school GPA students and two models performed little better than random chance. The graduation outcome models correctly classified the low high school GPA subset much better than the other testing datasets. This may be because the misclassifications in the retention models were most often false positives and the misclassifications in the graduation models were most often false negatives. However, the high school GPA variable, transformed for linearity, was nearly always statistically significant and was associated with moderately higher odds of retaining and even higher odds of graduating.

In the proportions approach I had found evidence suggesting that the relationship between first generation status and graduation likelihood changed over time. While it was significant in univariate models of 36 of the 37 outcomes, it was significant in only twelve of the multivariate models. As with under-represented minority status, first generation status was only significant in a few freshmen models and their corresponding junior models, but it was also significant in a few retention models. In the graduation models where it showed significance, there was no clear pattern of the coefficient increasing (becoming less negative) between the freshmen-junior model pairs.

Survival Analysis

The survival analysis modeled two outcomes, graduation and withdrawal separately. This allowed me to separately explore trends over time in the likelihood of the two events. As with the logistic regression approach, I built the two final multivariate models based on the significant variables from univariate models. The null models included the baseline hazard captured by a set of spline functions. The baseline hazard for graduation was captured with three splines modeled with three variables to reflect the change in baseline hazard between terms six and seven, between terms eight and nine, between terms ten and eleven, and between terms thirteen and fourteen. Spline 1 included terms one through six and terms fourteen and later. Spline 2 included terms seven and eight. Spline 3 included terms nine and ten. Spline 4 included terms eleven through thirteen. As the terms progressed, the likelihood of graduating increased dramatically for those students still enrolled.

Graduation

The first order relationships of graduation likelihood with admissions scores and high school GPA, both transformed for linearity, were positive. However, the relationship between graduation odds and admissions scores changed velocity so that by terms 11-13 the relationship was negative. Pell eligibility, under-represented minority status, and first-generation status did not contribute significantly after accounting for the baseline graduation hazard--the overall trend in graduation likelihood, akin to the intercept in a linear regression. Two variables closely tied to time were significant as first order time dependent variables: an indicator for whether the enrolled term was spring 2020 or later and the length of the student's most recent gap in enrollment. These variables were more challenging to interpret, but very telling. The net relationship of the enrolled term being post-pandemic and graduation odds varied by where the student was in their academic career. After the pandemic's onset, the pattern of increases in graduation odds was amplified; the odds of graduating was higher for all students post-pandemic, but increased the most over terms seven and eight, a little less over terms nine and ten, and the least over terms eleven through thirteen.

Withdrawal

The baseline withdrawal hazard had two time periods of increased risk. The first was between terms two and three, roughly between first spring and second fall. This captures the relatively high proportion of students who do not return for their second fall, which impacts the fall-to-fall retention rate. The second increase in withdrawal hazard was between the second spring and third fall, a pattern that cannot be captured by any common retention or completion metrics. Examining the baseline hazard curves by cohort showed that the second period of increased withdrawal risk has become more pronounced in recent cohorts.

In the logistic regression models, I transformed high school GPA by raising it to the fifth power and centering it within each year so that the relationships of it with retention and graduation likelihood were approximately linear. I also used this transformed high school GPA variable in the survival analysis where it was significant in explaining withdrawal odds; students with higher high school GPAs had a lower odds of withdrawing across all terms, which was expected. First generation status was significant in the withdrawal model, and contrary to what I had expected to find from the proportions analysis, the withdrawal risk was constant across all terms. The COVID parameter was significant, but none of its interactions with the splines were, so although the overall withdrawal risk increased post-pandemic, the change in risk was constant for all students, regardless of their term at the time. The only time dependent variable was the number of credits students took in the term which had an overall negative relationship with withdrawal risk that was even more strongly negative in the first two terms.

The most intriguing finding was how gaps in enrollment were associated with overall withdrawal risks. Having more gaps and having a longer most recent gap were associated with a higher withdrawal risk. However, the coefficient for the cumulative gap length nearly cancelled out that for the most recent gap length. Most students with gaps had only one gap, so for them, the two gap length parameters nulled each other out, leaving only the number of gaps to contribute to their withdrawal odds. These variables get more interesting for the handful of students who had more than one gap because their net relationship was a decrease in overall withdrawal risk.

Summary of Quantitative Analyses

The results from each analysis scaffolded on the earlier ones, providing more context and even answering questions I posited based on the findings from the proportions analysis. The majority of the findings from the analyses were unremarkable in that they reflected well-established relationships. While it is good practice to verify that these patterns were also present at Regional College, they would not justify the time expense of running these analyses.

The proportions provided a general impression of the retention and graduation trends at Regional College and suggested that the difference in graduation likelihood between first-generation students and their non-first-generation peers decreases over time. Breaking out the graduation rates for those students enrolled as of their third fall showed that the graduation rate gaps vary by the number of credits earned as of that term. Under-represented minority status and temporary United States (US) resident status were not significant in the logistic regressions, but for different reasons. There were too few temporary US resident students to provide stable estimates. The non-significance of under-represented minority status was due to measures of pre-collegiate academic success also being included in the models. Importantly, the classification accuracy was not consistent across all the subgroups, with the models fitting poorer for the group of students whose high school GPA was below a 3.0. In the survival analyses, a second period of increased withdrawal risk was apparent and the narrowing of graduation rate gap by first-generation status that was suggested by the proportions analysis was dismissed. The survival analysis also provided one of now many estimates of the COVID pandemic's impact on student outcomes.

Qualitative

Four focus groups with a total of 15 participants and my researcher journal were the source materials for the qualitative data collection. The participants in the first three focus groups had institutional researcher roles and the participants in the fourth were administrators. From these sources, I identified four meta-themes relevant to the utility and value of quantitative measures of persistence in general and the three approaches I chose in particular. The first three meta-themes, the role of institutional research at the institution, data governance, and institutional data use culture, give context for the decisions institutional researchers make when exploring persistence. The last meta-theme is about persistence itself and how it is explored.

Role of Institutional Research

The roles of IR among the represented institutions vary outside a set of core responsibilities. As Rouse (2018) found, these core responsibilities include mandated reporting at the federal, state, and regional levels.

> Right now, we're doing our strategic plan, our five year strategic plan. I've had to provide a lot of information for decision making there. Also, we are on a productivity funding model for the state. So, in order to earn points that equal our funding, that's very data driven, and the information we give the state, it’s very important that it's correct and that it's entered in the state system correctly. So twofold there and then of course, IPEDS reporting is also extremely important. (Jen, Institutional Researcher Group 3)

Participants spoke of institutional researchers as facilitators of using data for decision making by providing data for grant administration, measuring progress along a strategic plan, and improving financial aid efficiency. They collaborate with the many other units on campus who own the data IR offices work with and by analyzing these data, institutional researchers affect policy.

> But yeah, we've had a lot of changes to our academic policies in the last five years or so. We've changed our add-drop deadlines. And then more recently, we changed our academic calendar, so our fall semester mirrors our spring semester. They took a week out of the semester and that created all kinds of work. We had to survey students and faculty. We had to do focus groups. We had to do all kinds of things during the pilot to see how it was affecting student performance or outcomes or perceptions of their workload. (Carlos, Institutional Researcher Group 3)

Institutional researchers rely on relationships that span the breadth of their institutions from athletics to academics and from housing and dining to advising. They must be renaissance people who leverage eclectic skill sets including both technical and interpersonal skills. Most successful IR professionals wander into the profession though, bringing with them skills developed in neighboring disciplines.

> I don't think there's anyone that ever grew up saying “I want to be an institutional researcher.” I'm just guessing here. … you don't ever take off your statistics hat, but you're going to decorate it. And you're going to put a big flower right here and this is going to be the, you know, the person that does the narrative, and you're going to have a little button right here and this button is the part of your job that asks all the questions while you're actively listening because really, they don't know what they want. So, you almost have to become a counselor, and be like, “Well tell me how you feel” and “What will this help you do?” And so now you've got a button for the counselor. Then you've got your data, you know, that your data visualization button that you've added to your statistics. So, you’ve got this beautifully decorated hat, but at the base of it, you're always going to be a statistician, right? And that's really what's going to drive your data to be so valid and meaningful. (Jenn, Institutional Researcher Group 3)

The findings from Ducharme’s 2014 study also highlighted the diverse backgrounds of IR professionals, with 13% of the survey respondents having an undergraduate degree in arts and humanities and about one third having a highest degree in each education or social sciences. With all their skills though, institutional researchers are limited in their scope of practice. Analyses typically end at descriptive statistics and frequencies and IR professionals rarely get to do more complex analyses. These are instead contracted out to external vendors who do the analysis and report the findings back to the institution.

> We also have an external vendor right now, who's doing a retention prediction model for us. It creates a likelihood score of how likely they are to be re-enrolled the next fall. And then we're looking at the students who have those lowest likelihood scores, reviewing them in case management meetings with our associate deans of the college. (Carlos, Institutional Researcher Group 3)

Phillip J Fry described how his office guides stakeholders through interpreting results, but at other institutions, IR offices only provide the results and the interpretation and implementation of findings is left to the requester. It is their loss if stakeholders choose to disregard the institutional knowledge that lives in IR offices, but it is even worse when IR professionals have information that they do not actively share.

> Or at least that closing the circle as you mentioned, Carlos, it’s not just you, you're not isolated in that experience. It seems to be just the nature of IR for some reason. We're really meticulous about collecting data. We're really good about keeping it clean, organizing it, analyzing it, and then we have all this knowledge that doesn't get put back out. (Mano, Institutional Researcher Group 3)

The knowledge within IR offices is hard earned. Institutional researchers are driven to find and solve problems, compelled to understand how things work. After the presentations, the IR professionals asked questions and shared comments that showcased their natural curiosity, impulse to assign meaning to numbers, and need to form connections.

> It would be interesting to explore the qualitative on the second-year bump: withdrawals. We see it here as well. And similar to yours, it's not as pronounced as the first year but it

> does exist. And I think if, if I look into some of the user experience design training that I'm doing right now, in a separate course, it's… you're reaching that sunk cost threshold. (Katherine Danziger, Institutional Researcher Group 2)

However, many challenges prevent them from devoting their time to "playing with the data" and force them to triage requests. Multiple participants lamented not having more personnel resources, both in the size of the office and in specific skill sets, agreeing with Rouse's (2018) and Brown's (2008) findings that institutional researchers feel the need for additional technical skills and training. Similarly, Ducharme (2014) found that inadequate time, resources, and relevant skills are major barriers to IR's use of qualitative methods.

> This is the type of stuff that I wish I could do more of because, you know, you're getting I don't know, it's just a lot more interesting. But, you know, unfortunately, with all the reporting needs that we have at the top and we're very small office. (Hendrix 64 IR2)
> Well, I'm an office of one. And as much as I would love to play around with all the models and ways that we can show information, I simply just don't have the time for the work demand. I have to stick with what I know is the most receptive and communicates information, the most accurate. (Jenn, Institutional Researcher Group 3)

Because of their diverse backgrounds, not all IR professionals have formal research training; many are self-taught or learned on the job. Institutional researchers need to know more than just statistics to thrive. Coding, problem solving, and storytelling skills are also essential. Much of IR professionals' time is spent in one reporting tool or another, creating and maintaining reports and dashboards in reporting libraries that often become overgrown and unwieldy. Stakeholders find that having too much information to wade through is the same as having no information at all.

> We have a very extensive collection of analytics/visualization reporting tools that my office has developed, and actually it was my predecessor to be fair, but because the volume of data is so vast, the utility of the data has dropped significantly. (Seudo Nim, Institutional Researcher Group 1)

Institutional researchers use data to support institutional goals in resource-constrained environments with limited personnel. They're good at converting data into information but sometimes struggle to effectively communicate that information to stakeholders. They spend substantial time on low-level descriptive analyses and on maintaining reports and dashboards while the "fun" analyses get completed by contracted vendors. Statistical and research skills are not every IR professional's strengths, but they use a variety of other analysis skills and continue to grow.

Data Quality, Availability, and Access

Quantitative analyses require varying amounts of data, but challenges related to data were mentioned in all four focus groups. Higher education generates a lot of data, most of which is low-impact quantitative data collected and stored for operational use and not in a format conducive to research. The skill for IR professionals is to find the meaningful nuggets in with all the straw. They translate requesters' end goals into studies where the accessibility, availability, and quality of data are key factors.

> They want to replicate something that they saw like in the Chronicle or Inside Higher Ed for our university or for their department. And sometimes you can, but other times, you know, their department might only have 200 students in it. So I can't do a multivariate

> logistic regression with 20 different predictors on 200 students, so it's always balancing those different needs and trying to figure out what are they really asking and what information do they really need to make a decision. (Carlos, Institutional Researcher Group 3)

Unfortunately, IR professionals work with data stored across multiple data platforms which increases the potential for issues in accessing the data. When the same data are stored in multiple sources, the values in the data may not match. Formatting in the respective sources can also be a barrier, sometimes so bad that researchers must use an intermediary software or data format. The primary purpose of the software is operational, like facilitating course registration or managing admissions, and reporting needs are secondary. Some software "plays nice" with each other, but more often each was created as an island unto itself.

> We've had some discussions about trying to merge in our online learning platform, Canvas… The way Canvas has structured data is complicated, and it has absolutely nothing in common with our student information system, with Colleague. So like, which is where basically all our data comes from. So that hasn't gone anywhere. (Seudo Nim, Institutional Researcher Group 1)

In addition to technical issues, interpersonal barriers may prevent access. IR professionals rarely own the data they work with. Sources may fall under the purview of different people, often called data stewards, who among other responsibilities, grant access and work to ensure data integrity. Very rarely do data stewards oversee their data as small fiefdoms, but they do have the final say in who is granted access and when. Researchers may never have direct access to the data, instead relying on data pulled by others, trusting that their parameters were correctly understood in a game of "data request telephone".

> I came in right as we were switching to a new ERP. They're switching from a very antiquated mainframe system to Banner, and so I didn't even… I never got access to the data. I didn't get trained. They're like, wait 'till the transition is done, and then we'll give you access… they didn't even have a reporting tool for another year and a half. So, I never got my hands on the raw data until five years into my tenure there. (Outside LV, Institutional Researcher Group 2)

If working with multiple platforms were not frustrating enough, there are also multiple definitions in use for the most basic terms. In exploring persistence, higher education is naturally interested in how many students are enrolled, but there are multiple definitions to choose from for even something as simple as "enrolled student". Information technology may call anyone with an active account to register for classes enrolled for the sake of determining how many software licenses to buy. Marketing may call anyone who has registered for classes in a future term enrolled because they'll need to receive communications about upcoming events. The definitions also proliferate because external entities, like the Department of Education and the various accrediting bodies, each have their own definitions.

> A lot of times the accreditor has really specific definitions that they have to adhere to. And then accreditors, obviously, they're not… the nursing accreditor and the welding accreditor are not talking to each other when they determine how to calculate persistence. And the first thing you learn in this job is like 90% of the work is coming up with the right definition. (Seudo Nim, Institutional Researcher Group 1)

Often the difference in definitions is a matter of timing. Because much of institutional data is transactional, the data are nearly constantly in flux. Each change yields a new dataset to apply the definition to. Students add and drop courses throughout the term so someone who dropped their full schedule would change from enrolled to not-enrolled overnight. A few stragglers finally complete financial aid forms at the end of the academic year and are not identified as low income until after the fall term has ended. To dampen this noise in the data, business intelligence and institutional research offices take snapshots of the data at predetermined dates to serve as the official record. Data as of those dates are the only versions of the data to be used for official reporting. However, not all requests need official data, and some are better served by live data. Data definitions are just algorithms, instructions on how and what to count. A definition gives different values depending on the timing of the data it is applied to.

> So for external reporting, whether it's to IPEDS or Common Data Set or US News or wherever it's going to be, we want them to be frozen and snapshotted and consistent and always able to replicate the exact same information. For other things like room scheduling for our registrar's office … Doing some more of those kind of real time analysis direct drawing directly from the databases has been a growth area in our office and something that we seem to be doing more and more of over the last year or two. (Carlos, Institutional Researcher Group 3)

Persistence measures need to be accurate and valid, but having multiple values circulating, all called the same thing, weakens confidence in the values themselves. Stakeholders regularly question the numbers. Institutional researchers are not perfect. They make mistakes, but more often discrepancies are the result of the multiple definitions and the timing of the data. However, soft ledgers and shadow databases haunt colleges and units calculate their own version of persistence, adding to the cacophony.

> I had a mid-level administrator, dean/director kinda level, argue with me and said, “This isn't the right data.” So I said “Well. Okay, well, perhaps something's not in the central system. Perhaps something got overlooked. Do you have numbers?” “Yes, I always keep a tally: Physically count every student.” “Okay? Could you send me a copy of that?” I'm not kidding. I got a photograph of their retention rates written on a sticky note taped to the side of their monitor and I got a picture of all that. That was their official data. (Person1, Institutional Researcher Group 1)

Sometimes though, stakeholders’ attitudes and biases impede rational information interpretation; they do not trust the information that comes from the institutional research office. This is debilitating for IR professionals. Most stakeholders have no means of independently validating the numbers that come from institutional research, so although stakeholders’ mistrust likely has very little basis it leaves them stranded without reliable data for decision making. Too often the mistrust stems from stakeholders’ unfamiliarity with the data and business practices at the institution, but other times their weak analytical skills are at fault.

> But in general, right, the way our data culture is right now is, anytime anything like that is presented, it's ripped apart. And it's, “Hey, you're showing me some indicators at the student level that this student isn't likely to succeed. But I know that student. I know all these other variables that are not captured in here that show that student is likely to succeed. I don't trust this, therefore, I'm not going to use it.” (Phillip J Fry, Institutional Researcher Group 1)

With multiple definitions available, it is a challenge to balance the need for specific cases with standardization and comparability. IR professionals know that the standard definitions will not work for everyone, so they try to minimize the use of non-standard definitions unless absolutely necessary.

> Our standard definitions often don't work. And so how do we make it to where it's still, there's still some… You rein it in or you have checks to make sure it's, you know… you don't just do everything for everybody. Because then nobody agrees to anything and there's no comparability and no way of making sense of it at an aggregate level, but still be responsive to those specific needs. (Phillip J Fry, Institutional Researcher Group 1)

For all the heart ache that data integrity issues cause, costing time and resources to identify and explain or correct, they are unavoidable. As an industry, higher education generates a lot of data and institutional researchers are tasked with finding the information therein. They navigate the challenges of using data they do not own that is stored across multiple platforms not designed for reporting all in a swarm of definitions. Collectively, higher education professionals work to ensure consistency and comparability in values by adhering to standard definitions, but individualism within institutions frustrates those efforts with shadow databases and data hoarding that erode confidence in the information produced by institutional research.

Institutional Data Use

Many of the challenges institutional researchers face around data use stem from data culture and the role data plays at their respective institutions. Data literacy is foundational to appropriate data use. It gets a lot of attention in discussions, but stakeholders vary widely in their ability to interpret information and their familiarity with how data are collected and analyzed. Similarly, one of Ducharme's (2014) survey respondents highlighted the importance of being able to communicate findings to nonresearchers. Some stakeholders are happily ignorant because they lack the time to understand data fully; they are interested in the output and only the essential context.

> Equally important is how you package and present the data. You're not going to put anything out there that's not like wrapped up and you know, with a bow on top, because you don't want any loose threads to be taken out of context. Not everyone understands the data like you do. And not everyone understands data in general. You want to make sure that if a question was asked, you give the answer to the point concisely, neatly, with no loose threads. (Mano, Institutional Researcher Group 3)

Focus group participants mentioned repeatedly that the value of data and information is in its actionability. Data for data's sake is a waste because information's value is determined by what questions it can answer, nothing more. However, actionability does not always align with statistical significance. Because institutional researchers work primarily with secondary data, they cannot go back in time to collect information on confounding variables nor increase their sample sizes. The validity of statistical significance is based on the methods' assumptions being met, which is not the case when relevant variables are excluded or the analysis is under powered. The likelihood of statistical error is high, but that may be beside the point. If there are decisions to be made with less-than-ideal information, the available information is the best information available.

> Data as having operational-actionable value independent of its predictive sort of mathematical realities has served us well as an institution and kind of like getting people away from just you know, "what is statistically significant" to "what is operationalizable" if that's the word. (Diana, Administrator Focus Group)

Ideally, decisions are made following a scientific method where data evidence is used to select the best course of action, but reality is not the ideal. Information is regularly politicized, cherry picked, and used to support a decision that was already made. Even if statistical techniques are used, these decisions are made without the benefit of inferential statistics.

> Our leadership uses the data or asks for data in a reactionary sense, rather than a proactive sense. 30% is to justify decisions that have already been made or have been decided and we're now looking for some way to shape the narrative around that. (Seudo Nim, Institutional Researcher Group 1)

Thankfully this is changing. As in Rouse’s (2018) study, participants in this study believed that institutional research had the ability to influence internal decisions or were optimistic of increasing agency. As Phillip J Fry observed, data informed decision making is gaining traction as data evidence is intentionally integrated into the decision-making process.

> When it's clear that decisions are being made is there's a process that that data [analysis] gets embedded into. That's what we've been doing a lot more lately. There's an initiative going on, and data gets embedded into where, where the expectation is that data is being used for decisions in these ways. (Phillip J Fry, Institutional Researcher Group 1)

Participants mentioned experiencing both situations at their institutions; sometimes data informed the decision and others the decision determined what data would be used. With so much information available for stakeholders on demand though, it is difficult to predict which information is used when and for what.

> We too have dashboards that break down retention and graduation rate by academic unit, all the way down to the program level. I also don't know the extent to which decisions are

> made based on those things. I do know some departments find them useful. (Hendrix, Institutional Researcher Group 2)

Reports and dashboards are quick and efficient tools to communicate information to stakeholders, but they only communicate one direction, from the report writer to the data user. There is little opportunity to provide additional context or to get feedback and the data users are left to come to their own conclusions. Too often though, stakeholders misinterpret or misunderstand data and the error is not caught until it leads to a larger issue.

Persistence Measures

Persistence calculations have many applications in higher education, although some applications raise questions of equity. The measures are tied to funding through IPEDS and state mandated reporting and to faculty and staff jobs through academic program review. They are used to evaluate equity within an institution and by accreditors to evaluate institutional health. They also have operational value for improving institutional policy.

> … our first to second year retention was very good, but we'd see after they're in their third or fourth year, students started to leave and looking at the Clearinghouse data, we can see a lot of them were transferring to other universities and you know, part of it was their gap… So we changed our financial aid policies. We've seen an increase in our graduation rates over time, largely because we're now meeting the full need of our students. (Carlos, Institutional Researcher Group 3)

Because colleges are resource constrained, relatively small changes in persistence rates can have large impacts on programs and institutions. In this application, persistence is a measure of academic quality; programs with declining enrollments or low persistence rates face budget cuts that have real human impacts for their faculty and staff.

And then we take that information as an institution to look at what academic programs may need review. As far as may need strengthening or may need expansion or may even need to be sunset. You know, we look at persistence: we see students it started with 100 when they were freshmen. Now they're down to five students… is that a program that is viable? (Administrator X, Administrator Focus Group)

Persistence measures are used to identify struggling students. Propensity scores based on persistence measures help identify students who would most benefit from additional supports. These likelihood scores condense multiple attributes down to a single easy-to-interpret measure, ideal for front line student support workers. In this application, likelihood of persisting is an inverse measure of students' need.

So, what do you do calculate a persistence, likelihood and frame that sort of in in broad bands for the staff so that they have some sense of "this student maybe someone that needs more time, more attention, more care. (Diana, Administrator Focus Group)

Retention and graduation are usually discussed as the only positive outcomes of persistence, but transferring and taking time off are not necessarily failures by the student or the institution. Students may need to step away from their studies for a time; their college can support their return. Students may need to change institutions to complete their degree; colleges can support their transfer. One of community colleges' missions is to prepare students for transferring to a 4-year institution where they will ultimately finish their degree. Conventional graduation and retention rates treat these transfers as failures for the community college, discounting their institutional missions and the successes students achieved along the way.

Transferring is part of our mission. We in fact, have many 2 + 2 agreements with four-year institutions that allow them to go, allows a student to go to that that 4-year school for

> the tuition rate that we offer, as a 2-year institution. And so we don't consider a student is transferred out to be a bad part of our persistence. (Jenn, Institutional Researcher Group 3)

Graduation rates also impose specific deadlines that are not universally appropriate and meaningful. It is some students need longer to complete their degrees than the conventional timelines, while those who began college with advanced standing should complete sooner. Using the conventional timelines in these situations becomes an issue of equity. When interpreting graduation rates broken out by race, Long (2020) found that institutional researchers tend to use a color-blind framing, ignoring influences of systemic racism and privilege.

> For some students, like getting through one college class is a success. Getting you know, half of your degree: that is a success. That will never be included in national statistics, unfortunately... For tribal colleges, [persistence rates] mean a little bit less in terms of success for students and kind of how students move through higher education. (JMH, Institutional Researcher Group 2)

Regardless of the many functions persistence measures serve, the values are primarily interpreted through the positivist epistemology, that there is one objective truth, that “numbers don’t lie” (Administrator X, 90 Admin). However, some applications impose a different understanding. In the quantitative strand of this study, I influenced the reality I observed through the variables and the data I chose to used, which aligns with the post positivist epistemology, acknowledging that findings always depend on how one looks at the data. The participants in Ducharme’s (2014) study also shared that the information preferences of their stakeholders drive which approaches institutional researchers choose to use.

> But I think the overall narrative that leaders tell themselves about persistence has a bigger influence on the kinds of analyses that we're asked to do. And let me give an example. For a long time, we had an Enrollment Management leader who was very convinced that the single biggest factor in persistence was the ability of the student. And so then it became, all the analyses became about, you know, segmenting persistence by some incoming measure of the ability of the student and analyzing it through that lens. Whereas, you know, we had a Provost come in who has a background in curriculum and instruction and education, and believed very strongly that persistence was a function of the quality of the instruction that takes place at an institution. And then the analyses that you conduct become more about well, where are the where the systemic bottlenecks in the students' experience going through, you know, through the system. (Horace, Institutional Researcher Group 3)

Horace further described how this researcher influence plays out in their own analyses and describing how they might counter the biases that enter the analyses because of these preferences.

> There's also been a trend in circles that I've been hearing too about, like, not just measuring someone's demographic variables, but trying to measure the effects of those variables. So, like if you suspect that there's like some racism endemic to your system, rather than try to like measure the race or ethnicity like try to quantify like, exactly what kinds of racism that person might experience and that could be a potentially interesting way to go. (Horace, Institutional Researcher Group 3)

This parallels Long's (2020) findings that institutional researchers infrequently provide context along with persistence data broken out by race and that the additional information is generally operational and only rarely addresses race-based inequities. Without caution and concern, institutional researchers perpetuate racial inequality through their analyses.

Although represented with numbers, persistence is interwoven with stories about the institution and about its students that influence the meanings they assign to the values. There are alternatives to quantitative approaches that better capture and acknowledge these stories like the focus groups, case study, and mixed methods mentioned by the participants. These methods are ill suited to the current institutional research environment, beginning with their assumption that knowledge is subjective.

> [If I get] person A, B, and C to tell me the same story, but I still get something different each time I hear the story from A, B, and C. Even though they witnessed the same situation. They were a part of the same circumstance. The outcome was the same for all of them. The reasoning the rationale behind is not it's not necessarily the same. (Administrator X, Administrator Focus Group)

On a more practical level, these methods are less common because few IR professionals have qualitative research skills. The participants unequivocally valued quantitative skills higher, but they lamented not having qualitative skills within their offices, going as far as making light of the paucity of qualitative skills within the profession. Ducharme (2014) also found that institutional researchers reported using quantitative methods in preference to qualitative methods for all nine of the IR functions examined. These functions included enrollment management, accreditation, outcomes assessment, program review and strategic planning, all areas in which persistence is key.

> The running joke is you put 10 researchers in a room and based on the publication volume in the world… you might get less than one person in the room who actually

> knows what qualitative epistemology is. The other nine are very well versed in every statistical method that they like to run, and they continuously run, because it's the one they want to use. (Person1, Institutional Researcher Group 1)

In Ducharme's (2014) survey, nearly all respondents agreed that qualitative methods complement quantitative methods. Similarly, the focus group participants for this study recognized the strengths of the two methods. The value of qualitative research to participants was in being able to explain the "why" of observed patterns, but it was more important to them that quantitative approaches are less human resource intensive. Quantitative methods leverage secondary data sources, greatly reducing data collection time, and the calculations and data presentation can be automated through reports and dashboards.

> So being able to pull it up in real time and consider the question right away… I built some Tableau dashboards that let me do that. I guess they're available, anyone on our campus can see them but I'm probably the person who's in there the most because I can pull them up when asked a quick question and answer pretty quickly. (Carlos, Institutional Researcher Group 3)

Persistence measures have a variety of internal and external uses and are most often discussed in terms of a positivist epistemology--as being the only truth. However, their assumptions greatly reduce their utility for some populations, mainly because of systematic differences between these subpopulations and the majority of college students. Institutional researchers have used qualitative methods to assess persistence to offset these issues, but they are resource intensive and qualitative analysis skills are not as common in institutional research as quantitative skills are. It is also more challenging to communicate the nuances that make qualitative methods so valuable to diverse stakeholder audiences.

Many factors of institutional researchers' environments influence how they approach analyses in general and exploring persistence specifically. They work in complex environments where insufficient resources and high demands are the norm. Their problem-solving skills and curiosity help them navigate these challenges, but because of other demands on their time, they are limited in the depth and breath of analyses they can perform. Persistence has many definitions and applications, so institutional researchers pair the definition that best fits with the application for each analysis. Sometimes, the definitions are such a poor fit though, that they fail to meet the needs of the institutions. In these instances, IR professionals use other methods to explore persistence, but these tend to be more resource intensive than the standard methods, thereby limiting their wide-spread adoption.

CHAPTER V

DISCUSSIONS AND CONCLUSIONS

Persistence is a many-headed beast that colleges and institutional researchers encounter daily. This study evaluated three quantitative approaches for exploring persistence and in this final chapter I review its purpose and background followed by a discussion of my findings and an interpretation of those findings. After explaining the limitations of this study, this study is concluded.

Purpose of the Study

Persistence measures have many purposes, ranging from targeting supports for individual students to meeting the requirements for awarding federal financial aid. The measures are most often calculated as proportions although other approaches with their own assumptions and limitations, and benefits. Understanding why institutional researchers use one approach over others can highlight areas for professional development within IR offices and other college units as well as guiding discussions of data use within higher education. The purpose of this study was to evaluate three quantitative approaches for exploring persistence using the lens of institutional research and to contribute to the minimal existing research on the methods used by institutional researchers. The three approaches I evaluated were proportions of counts, logistic regression, and survival analysis. I selected these three approaches because they were present in mandated reporting requirements or were present in persistence literature.

Summary of Results

This study used both qualitative and quantitative methods to evaluate the three approaches. In the quantitative strand, I applied the three approaches to a dataset of new first-time full-time fall starting degree seeking students who started at Regional College between fall 2014 and fall 2022. For the qualitative strand, I analyzed the transcripts from four focus groups of institutional researchers and college administrators and my researcher journal to bridge the gap between methodological rigor and robustness and the utility required by institutional research.

Qualitative

In the qualitative strand of this study, the theme "multiplicity" came up across many aspects of institutional research and persistence modeling. Persistence itself has multiple definitions and its measures fulfil multiple functions. Thankfully, the definitions are often paired with an expected function of that measure, such as the definitions set by NCES for IPEDS data collection being paired with benchmarking across institutions and with meeting mandated reporting requirements. When new needs arise, researchers re-use existing definitions of persistence to ensure comparability, mitigate misinterpretation, streamline and automate calculations, and reduce the reporting burden. Most definitions specify the time components of the request; for those that do not, the institutional researcher negotiates them with the requester or uses their best judgement to assign them based on what the persistence metric is needed for.

The raw data are stored across multiple platforms which complicates data extraction and cleaning. The data sources are not designed with reporting in mind and require substantial effort to integrate. Sometimes, the data in a particularly obstinate platform is orphaned until it can be hopefully accessed at a future date. The platforms themselves are not the only barrier to accessing useful data; the data culture within an institution can pose a challenge when it is the norm for other units to share extracted data with institutional researchers instead of them pulling the data themselves. Even when IR professionals have direct access to fully connected data platforms, simply having the same information stored in multiple locations increases the likelihood of errors.

In addition to research, IR offices have administrative and operational responsibilities. The predictable, mandated requests are completed alongside ad hoc requests from stakeholders ranging from faculty to boards of regents and the National Center for Education Statistics. Because they are subject to so many stakeholders, IR professionals work on multiple concurrent projects with little time to devote to each one. Analyses must be fast and efficient and the findings easily communicated to people with minimal understanding of the research process or business practices. Further, some IR professionals lack formal research and statistics training. Instead, they entered their roles through necessity or because they have other advanced skills necessary for the position such as database administration, data visualization, or coding.

Quantitative

The quantitative strand demonstrated the relative strengths and weaknesses of the three approaches. Interestingly, the findings from the proportions analysis suggested that the graduation rate gap between first-generation students and their non-first-generation peers may close over time. If this were the case, Regional College would have evidence that its existing supports are at least partially successful. However, this finding did not withstand the higher rigor of the logistic regression or survival analysis approaches.

Proportions

The proportions were by the fastest to calculate and required the least amount of data. They were easiest to communicate and enabled direct comparisons between Regional College and other institutions. However, I ended up with an unwieldy number of statistics as I broke students out by their intersectional identities. Further, I had no way of judging whether the differences I observed were likely due to random chance, the variable of interest, or to a moderately correlated covariate such as any of the other variables I examined in the multivariate analyses. Although a test for the difference of proportions exists, it is sensitive to large sample sizes and assumes that the difference between the two proportions is due entirely to the variable of interest. This assumption that the two groups were in all other ways equal extended to the application of the findings as well. Being able to compare the persistence rates of Regional College with other colleges implies that the institutions are equal in all other relevant aspects, that the two institutions have similar admissions criteria, financial resources, and mission so that the rates have the same meanings. A community college and an ivy league university are intuitively different. The graduation rate for the community college will look dismal next to that of the ivy league university without the context that one has no admissions requirement while the other is highly selective; therefore the two institutions are incomparable.

Logistic Regression

My findings from the proportions of counts approach showed that under-represented minority students perform differently from their well-represented minority and non-minority peers, but this finding was not borne out in the logistic regression approach. After adding in measures of academic preparedness along with the other variables, the under-represented minority variable lost significance. This aligns with a participant's suggestion of trying to measure the effects of racism (i.e. poor academic preparedness resulting from racism in k-12 schooling) instead of trying to measure students' ethnicities. Aside from this finding, there were few new insights from this second set of analyses; most findings paralleled the patterns common across higher education. It was useful to learn that the students of Regional College was not unique, so that presumably programs that work elsewhere might be effective here as well. Unfortunately, this also limits the actionability of the findings because they give no new avenues to pursue. Thankfully there were fewer statistics to consider, and each came with an estimate of how likely the association is real. Because of the shift from proportions to odds ratios though, the statistics were more difficult to communicate and more prone to misinterpretation.

Survival Analysis

The final approach was the most informative, but also the most complex. Because I used snapshotted official data, no cleaning was necessary for any of the three approaches beyond reformatting and transforming the variables. If I had not had access to these clean data, the survival analysis approach may have been prohibitively time consuming because it required so much more data. Learning that the association of first-generation status and graduation likelihood was time-invariant (as opposed to the time varying association suggested by the proportions of counts approach) suggests that the college can further support first generation students by continuing programming through senior year. The most actionable finding of all three approaches is the second period of increased withdrawal risk year that I found through the survival analysis. Focus group participants had observed a similar pattern at their institutions, but this was yet unexplored at Regional College.

Interpretation of the Findings

All three approaches are used in practice, with the proportions of counts being most widely used although it is the least methodologically robust approach. According to the focus group participants, none of the approaches can be used all the time; the approach used is determined by the situation and methodological robustness is only one consideration of many.

Research Question One

> What factors do institutional researchers consider when selecting a quantitative approach to exploring persistence?

The findings from this study suggest four criteria guiding which approach is best suited for a situation: implementation of the approach, quality of the findings, institutional data culture, operationalization of persistence.

Implementation of the Approach

The three approaches share a need for clean data, although as the complexity of the analysis increases, so does the quantity of data needed. Without access to sufficient clean data, proportions of counts may be the only option available for exploring persistence. Logistic regression and survival analysis allow researchers to explore student's intersectional identities in relation to persistence outcomes, but the college must collect this information about its students to experience the methodological benefit. If the information is collected, but the college has poor data governance, cleaning the data to the point of usability may be prohibitively time consuming, especially if the bad data practices are not corrected. The time available to complete an analysis may also determine which approach is used. Institutional researchers consider resource requirements when selecting an approach.

Proportions of counts are obtained through simple arithmetic, while logistic regression and survival analysis require statistical software and familiarity with statistical thinking. IR professionals have diverse backgrounds that do not always include research methodology or statistical analysis, so while some statistical software is available without cost, IR professionals may face other barriers to its use. Those professionals without these skills tend to rely on proportions for data storytelling because they are easy, familiar, and get the job done. Proportions of counts are the most comfortable approach and although institutional researchers move between positivist, post-positivist, and interpretivist epistemologies, proportions and their positivist epistemology reinforce the feeling of certainty in the findings.

Quality of the Findings

The institutional researcher participants of the focus groups recognized that they make trade-offs so that the methodological quality of the findings is rarely the primary driver for selecting an analytical approach. The quality of findings is determined by their actionability, including how well their motivations can be identified and how well they align with prevailing theories. College administrations tend to be fiscally conservative and hesitant to divert limited resources toward new projects; convincing them of a new trend or relationship requires more compelling evidence. Generally, there are insufficient resources to cross validate findings through a secondary analysis as I did in the quantitative strand of this study. I was able to discredit the suggestion from the proportions approach that the graduation rate gap by first-generation status closed over time using the survival analysis approach by finding that there were no significant interaction effects between the first-generation variable and the variables for the terms. Institutional researchers also triangulate their findings with existing resources like knowledge of institutional policies’ effects and trends in population demographics. If, as with my logistic regression approach, the findings from one approach are as expected or are uninteresting, other approaches will gain favor for contributing new information to the conversation or for more efficiently delivering similar findings.

Institutional Data Culture

The stakeholders that institutional researchers prepare information for most frequently are close to them in the organizational structure of their institutions, as close as their direct superiors. This familiarity gives IR professionals a deep understanding of their data consumers' data use abilities and preferences. Collaborating with more distant stakeholders within their institutions and across higher education exposes them to the abilities and preferences of a broad range of people. Through this depth and breadth of experiences, IR professionals learn of potential barriers to effective data use, like data hoarding, the existence of shadow databases, and low trust in data. These, along with stakeholders' analytical skills and fluency with business practices, are aspects of data culture that influence which analytical approaches institutional researchers use. When communicating to groups with higher data literacy that are less likely to push back on the information, institutional researchers can use automated mechanisms like dashboards and reports. Otherwise, information communication is more bespoke, possibly even a dialogue, giving IR professionals the opportunity to explain more nuance and to clarify misconceptions. When they must also counter data mistrust, institutional researchers prefer approaches that are easy to explain so that the analysis does not contribute to the misunderstanding.

Operationalization of Persistence

The final criteria influencing which approach IR professionals select is how persistence is operationalized. There are multiple definitions of persistence to choose from, so thankfully the definition used is often indicated in the request such as the definitions provided in the instructions for IPEDS reporting. Regardless of the definition used though, the underlying data are unavoidably plagued with issues and Long's (2020) lament at the lack of context in interpreting the information presented publicly through the IPEDS website is not unique. Some students meet their educational goals without earning an award while others never intended to earn an award and only declared as degree seeking to receive financial aid. Students may mis-identify as transfer students or may be transfer students and not be recorded as such. Definitions of persistence ignore the progress and successes of transfer students, even if that is core to the mission of the institution. Naturally, institutional researchers are disinclined to invest time and effort in robust persistence modeling analyses when the resulting metrics unavoidably misrepresent the community of interest.

Stakeholders presume a process to persistence that also determines what approach Is taken. Identities that researchers expect to differ on graduation likelihood enter models as variables, and intersectional identities are entered as interactions. When constructs are expected to mediate other effects, a more complex approach is required. Models are built and approaches are selected based on pre-existing narratives coming from peer-reviewed published literature and from administrator's personal beliefs. However, even the most robust experimental design cannot address why persistence values are what they are.

Research Question Two

How do the factors apply to selecting a quantitative approach to exploring persistence?

Implementation of the Approach

Proportions are best suited to applications where standardization is more important than capturing nuance such as when the IPEDS components mandate the use of proportions. Small or understaffed IR offices can produce them quickly and consistently and they are the best option for audiences with low data literacy and analytical skills. The strength of proportions is their simplicity and minimal data requirements. They require minimal data pre-processing and are also effectively and accurately communicated through automated methods like dashboards and reports which adds additional efficiency by not requiring personalized communication to disseminate the findings. Because of their use in IPEDS reporting, they are the lengua Franca for mandated reporting and external surveys as other organizations work to reduce reporting burden through standardized definitions.

Logistic regression is more resource intensive than proportions and survival analysis, as a specific type of logistic regression, is even more so. It is a poor choice when researchers have insufficient resources to run and interpret an analysis. Many institutional research offices have only one employee and others are so small that all institutional research tasks are completed by someone whose primary role is outside institutional research; some institutions have no office of institutional research. For these institutions, proportions are the best option because they leverage the effort already required for mandated reporting to meet internal needs. Another implementation-related barrier to using logistic regression is having insufficient data. Even medium sized institutions may encounter this in the form of small class sizes, few students with the relevant identities, changes in data practice, or legislation preventing certain uses of data. Also, data may be missing or incomplete or direct access may be restricted. inadequate quantitative experience or training in specific approaches may also limit institutional researchers' use of logistic regression and survival analysis.

Quality of the Findings

The simplicity of proportions is also their weakness; they imply a comparability among colleges and communities that may not exist. The fall-to-fall retention of students who began in fall of 2018 is qualitatively different than the fall-to-fall retention of students who began just one year later. The mechanics are the same, but comparing the two proportions ignores the impacts of a global pandemic. Compared with proportions, logistic regression is well suited to exploring the relationships between intersectional identities and persistence and yields results with higher methodological quality. Survival analysis yields the most methodologically appropriate findings, accounting for the longitudinal nature of persistence. They both give researchers the statistical significance of variables and overall model as another tool for determining importance and actionability, but the findings are also more information dense than those from a proportions analysis. The quality of findings in institutional research describes their actionability first and foremost with the methodological rigor relegated to just one aspect of quality. Findings that are not actionable, either because they are too complex for stakeholders to apply or because they impart no new knowledge, are poor quality even if they stem from a methodologically rigorous analysis.

Institutional Data Culture

Institutions dealing with information mistrust may find the simplicity of proportions useful as they build or rebuild trust in the more complex analyses. After an initial explanation, stakeholders are less likely to misinterpret or misconstrue proportions which allow readers and researchers to stay in a positivistic epistemology, representing a “single truth”, streamlining communication and decision making. When considering a more nuanced truth matters, proportions are a poor choice. Researchers can use logistic regression as a first step toward exploring those nuances, but the increased information density of the findings may confound stakeholders with low data literacy. Logistic regression and survival analysis are poor choices when data analysis comes after the decision has been made because there is no return on the additional time and effort they require over proportions. Survival analysis findings are the most complex out of the three approaches included in this study, and the approach puts more onus on researchers to process the findings to effectively communicate them to stakeholders.

Operationalization of Persistence

In the unavoidable situations when institutional researchers must use persistence definitions that misrepresent their students, proportions are the preferred approach because they minimize how much time and effort is wasted. Propensity scores from logistic regression models allow front line student support staff to target efforts toward students who are most likely to benefit and allow researchers to estimate program effects in quasi-experimental research designs. Logistic regression and survival analysis provide researchers with a more holistic understanding of persistence. Logistic regression and survival analysis can both use the IPEDS definitions of persistence as starting points, but their increased robustness allows researchers to simultaneously estimate the relationships between multiple variables and persistence and to examine persistence behavior outside of the few timepoints set in the IPEDS definitions.

In academic programs, students are taught the approaches to use for specific needs, but these approaches often do not coincide with the ones used in practice. When institutional researchers examine persistence, they consider other factors beyond the statistical basis for selecting an approach. Available resources, specifically technical skill, time, and quality data, relevance of the persistence definitions to their community, stakeholders' ability to digest and apply findings, actionability of findings, whether an approach is mandated, and ease of communicating findings all contribute to which approach they select. Most of these factors are outside their control, but institutional researchers do the best they can in a less than ideal analysis environment.

Recommendations for Practice

From a methodological perspective, methodological rigor is the primary, if not only, driver of approach selection. In practice, though, institutional researchers and other applied researchers must consider other factors, often at the expense of methodological rigor. From this study I have five recommendations for practitioners.

First, I recommend reviewing current practice using these four factors to identify why the most methodologically rigorous approach is not used. This review should also consider the possibility that the barriers to the more rigorous approaches may be insurmountable, at least in the immediate future. However, colleges should be open to opportunities to overcome these barriers. Improving current practice must be an institutional priority to be effective, and as the people most attuned to persistence research, institutional research professionals should be vocal advocates for themselves and the quality of the information they provide. The second period of increased withdrawal risk at Regional College was a previously unknown opportunity to target programming and resources for ultimately increasing graduation rates. Given the increasingly competitive higher education environment, institutions are well served by tools that help them retain and graduate more students.

Second, colleges and institutional researchers all benefit from robust data governance, so every institution should have a sufficiently empowered data governance initiative. Through data governance, institutional researchers can ensure data hygiene, increase data literacy, and advocate for necessary data collection. Data siloing and shadow databases are prime targets of data governance efforts because of their toxic effects on data hygiene and on appropriate data use. Adequate governance also reduces the mandated reporting burden, freeing capacity for more actionable projects. Importantly, it does not simply move the responsibility of managing bad data from institutional research to another unit; instead, it targets the etiology of bad data, improving business practices and providing efficiencies to units across the college.

Third, I recommend that colleges and professional organizations provide sufficient resources for professional development. The use of survival analysis to examine persistence was novel to a few focus group participants, and others had never even heard of the approach. As much as institutional researchers want to provide the most actionable and accurate information, it is too easy to neglect professional development without outside support. Offerings should extend beyond the already incredibly valuable IPEDS keyholder trainings to include free or low cost resources on quantitative and qualitative research methods. Even a compilation of resources already provided by other organizations would increase the accessibility of these skills.

Fourth, participants also acknowledged that their inadequate communication skills were a limitation. The baseline hazard curves from the survival analysis were evocative as was the Power BI® report I used for exploring the proportions, demonstrating the effectiveness of automated communication for simplistic information. For more complex findings, I agree with one of the participants that having a representative stakeholder review communications before wider dissemination helps the institutional researcher anticipate questions. Using the researcher journal allowed me to process the quantitative findings multiple times, honing my understanding each time and reducing what information would have needed to communicate to just the essentials. Institutional researchers can ease communications with lay stakeholders by referencing simpler approaches, reserving any mention of methods until being asked for them, and by focusing tightly on only the immediately relevant findings. In the interest of open communication, offices of institutional research should then also follow the example of industry and prepare internal research briefs or technical papers following complex analyses so prevent their own information hoarding.

Fifth, colleges should work to support a culture of data-informed decision-making, the foundation of which is both quantitative and qualitative data literacy. Many focus group participants shared their stakeholders' limited quantitative skill as a barrier to using methodologically rigorous approaches, even as far as needing to avoid charts and tables in presentations. It is important to meet stakeholders where they are, but there must also be an expectation that key stakeholders will remediate their data literacy.

When choosing an approach, mythologists would prefer that practitioners used the most rigorous approach possible, but that is not always the case. Before jumping to a new approach though, practitioners should evaluate why that approach was not already in use. If the barriers can be eliminated or reduced, they should be. Although, the limitations of practice should be recognized as a difference between academic and applied research.

Suggestions for Additional Research

As a mixed methods evaluation of quantitative approaches to exploring persistence, this study is the first of its kind. The three approaches are all commonly used, so this study helps to bridge the gap between how people are taught to use the approaches and how they are used in practice. Academic faculty could use further research in this vein to support students entering careers in applied sciences as they transition from the academic research environment to one that is much more convoluted. The three approaches I selected are not exhaustive of the approaches available to explore persistence, so future research could consider additional approaches and other applied research professions to identify additional factors that influence researchers' methodological decisions.

One barrier mentioned by multiple participants is the difficulty in communicating statistical findings, especially ones from complex analyses, to lay stakeholders. Given the value of the more methodologically robust approaches, future research should investigate best practices in communicating statistical findings. General recommendations for these communications already exist, so the future research should include assessing the effectiveness of the recommendations, the frequency of their implementation, and barriers to their use.

Limitations of the Study

This study evaluated three quantitative approaches for exploring persistence through the lens of institutional research in the United States. Although some alternative approaches were mentioned, this study did not evaluate them thoroughly. The focus group participants and the quantitative data were all selected because of their attachment to a United States higher education institution all of which reported to IPEDS in the 2020-21 survey cycle. Most institutions in the United States report to IPEDS either by mandate or voluntarily, so this study cannot represent the approaches of persistence modeling used at those US institutions that do not report or institutions outside of the US.

The focus group participants included institutional researchers and upper administrators, so I did not capture the perspectives of lower and middle academic administrators like department chairs and deans. Because of their different vantage points, they may have had different experiences related to persistence modeling at their institutions. My role as both the researcher for this study and an institutional researcher professionally gave me insight into the experiences of the participants but influenced my selection of study design and interpretation of the findings.

I selected variables used in the quantitative strand because they are common in persistence research and were readily available at Regional College. The findings from that very section indicate that had I included other variables my evaluation of the approaches may have been different. This study should not be taken as an authority on impacts of the COVID pandemic in higher education because only two years of enrollment and graduation data after spring 2020 were analyzed.

Conclusion

Institutional researchers consider many factors when deciding which quantitative approach to use for exploring persistence at their institutions. In many situations, the approach is mandated by the requester so even when that approach has noteworthy methodological limitations, that is the approach they use. Other times limited resources within the IR office and poor data literacy and analytical skills among stakeholders play a larger role in approach selection than methodological considerations like assumptions and sample size. Regardless of the causes though, institutional research professionals often defer to less methodologically robust approaches to explore persistence. This creates a disconnect between applied research practice and the approach selection methods taught in academic programs. Through understanding and appreciation of this disconnect, faculty can better prepare students for the transition to practice. Training resources introducing the many institutional researchers enter the profession without academic statistical training to more methodologically appropriate approaches should address these non-statistical considerations. Higher education institutions could also provide professional development opportunities around increasing data literacy, research skills, and data security so that institutional researchers have an audience prepared for findings from more robust analyses.

REFERENCES

Abrams, K. M., Wang, Z., Song, Y. J., & Galindo-Gonzalez, S. (2015). Data richness trade-offs between face-to-face, online audiovisual, and online text-only focus groups. *Social Science Computer Review*, *33*(1), 80-96. https://doi.org/10.1177/0894439313519733

ACT. (2022). *ACT / SAT Concordance*. https://www.act.org/content/act/en/products-and-services/the-act/scores/act-sat-concordance.html

Allen, J., & Robbins, S. (2010). Effects of interest—major congruence, motivation, and academic performance on timely degree attainment. *Journal of Counseling Psychology*, *57*(1), 23-35. https://doi-org.unco.idm.oclc.org/10.1037/a0017267

Archibald, M. M., Ambagtsheer, R. C., Casey, M. G., & Lawless, M. (2019). Using Zoom videoconferencing for qualitative data collection: Perceptions and experiences of researchers and participants. *International Journal of Qualitative Methods*, *18,* 1-8. https://doi.org/10.1177/1609406919874596

Arreola, E. V., & Wilson, J. R. (2020). Bayesian multiple membership multiple classification logistic regression model on student performance with random effects in university instructors and majors. *PLoS ONE*, *15*(1), https://doi.org/10.1371/journal.pone.0227343

Association for Institutional Research. (n.d.). *IPEDS training center*. https://www.airweb.org/collaborate-learn/ipeds-training-center

Association for Institutional Research. (2022). *Duties and functions of institutional research*. https://www.airweb.org/ir-data-professional-overview/duties-and-functions-of-institutional-research

Bean, J. P. (1982). Conceptual models of student attrition: How theory can help the institutional researcher. *New Directions for Institutional Research*, *1982*(36), 17-33. https://doi.org/10.1002/ir.37019823604

Bean, J. P. (1983). The application of a model of turnover in work organizations to the student attrition process. *The Review of Higher Education*, *6*(2), 129-148. https://doi.org/10.1353/rhe.1983.0026

Brown, J. A. (2008). Institutional research and external reporting: The major data seekers. *New Directions for Higher Education*, *2008*(141), 87-96. https://doi.org/10.1002/he.296

Bujang, M. A., Sa'at, N., Tg Abu Bakar Sidik, T., & Chien Joo, L. (2018). Sample size guidelines for logistic regression from observational studies with large population: Emphasis on the accuracy between statistics and parameters based on real life clinical data. *The Malaysian Journal of Medical Sciences*, *25*(4), 122-130. https://doi.org/10.21315/mjms2018.25.4.12

Cabrera, A. F., Castañeda, M. B., Nora, A., & Hengstler, D. (1992). The convergence between two theories of college persistence. *The Journal of Higher Education*, *63*(2), 143-164. https://doi.org/10.1080/00221546.1992.11778347

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B, Methodological*, *34*(2), 187-220. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Los Angeles: Sage Publications, Inc.

Davidson, J. C., & Blankenship, P. (2017). Initial academic momentum and student success: Comparing 4- and 2-Year students. *Community College Journal of Research and Practice*, *41*(8), 467-480. https://doi.org/10.1080/10668926.2016.1202158

Davies, L., LeClair, K. L., Bagley, P., Blunt, H., Hinton, L., Ryan, S., & Ziebland, S. (2020). Face-to-face compared with online collected accounts of health and illness experiences: A scoping review. *Qualitative Health Research*, *30*(13), 2092-2102. https://doi.org/10.1177/1049732320935835

Davison, A. C. (1989). Model-checking I: General regression models. *Brazilian Journal of Probability and Statistics*, *3*(2), 77-86. http://www.jstor.org/stable/43601341

Deom, G., Talley, J., Sauer, M., & Fiorini, S. (2024, January 1). *A Three-part Series on the Application and Production of Shiny Applications for IR: Part 3*. Tech Tips.

Dodds, S., & Hess, A. C. (2021). Adapting research methodology during COVID-19: Lessons for transformative service research. *Journal of Service Management*, *32*(2), 203-217. https://doi.org/10.1108/JOSM-05-2020-0153

Ducharme, D. R. (2014). *Factors that influence the use of qualitative methods by institutional researchers*

Gansemer-Topf, A. M., Kollasch, A., & Sun, J. (2017). A house divided? Examining persistence for on-campus STEM and Non-STEM students. *Journal of College Student Retention: Research, Theory & Practice*, *19*(2), 199-223. https://doi.org/10.1177/1521025115611671

George, A., Stead, T. S., & Ganti, L. (2020). What's the risk: Differentiating risk ratios, odds ratios, and hazard ratios? *Cureus*, *12*(8), e10047. https://doi.org/10.7759/cureus.10047

Gury, N. (2011). Dropping out of higher education in France: a micro-economic approach using survival analysis. *Education Economics*, *19*(1), 51-64. https://doi.org/10.1080/09645290902796357

Hess, K. R. (1995). Graphical methods for assessing violations of the proportional hazards assumption in Cox regression. *Statistics in Medicine*, *14*(15), 1707-1723. https://doi.org/10.1002/sim.4780141510

Hillman, N. W., Tandberg, D. A., & Fryar, A. H. (2015). Evaluating the impacts of "new" performance funding in higher education. *Educational Evaluation and Policy Analysis*, *37*(4), 501-519. https://doi.org/10.3102/0162373714560224

Hosmer, D. W., Lemeshow, S., & May, S. (2008). *Applied survival analysis: Regression modeling of time-to-event data* (2nd ed.). Wiley.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). John Wiley & Sons, Incorporated. https://ebookcentral.proquest.com/lib/unco/detail.action?docID=1138225.

Hu, F. B., Goldberg, J., Hedeker, D., Flay, B. R., & Pentz, M. A. (1998). Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. *American Journal of Epidemiology*, *147*(7), 694-703. https://doi.org/10.1093/oxfordjournals.aje.a009511

Jaggars, S. S., & Xu, D. (2010). *Online learning in the Virginia Community College System*. New York: Community College Research Center. https://ccrc.tc.columbia.edu/publications/online-learning-virginia.html

James, N., & Busher, H. (2009). Developing online research strategies: A methodological discussion. In N. James, & H. Busher, *Online Interviewing* (pp. 30-41). London: SAGE Publications Ltd .

Jóźwiak, K., & Moerbeek, M. (2012). Power analysis for trials with discrete-time survival endpoints. *Journal of Educational and Behavioral Statistics*, *37*(5), 630-654. https://doi.org/10.3102/1076998611424876

Kim, C., Tamborini, C. R., & Sakamoto, A. (2015). Field of study in college and lifetime earnings in the United States. *Sociology of Education*, *88*(4), 320-339. https://doi.org/10.1177/0038040715602132

Krueger, R. A., & Casey, M. (2009). *Focus groups: A practical guide for applied research* (4th ed.). London: SAGE Publications Ltd.

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education*, *79*(5), 540-563. https://doi.org/10.1353/jhe.0.0019

Larkins, R. J. (2015). *Project success: A methodological and evaluative case study of the early alert program Interventions*..

Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method* (2nd ed.). American Elsevier Publishing Company.

Lee, J. (2020, March 17). *A neuropsychological exploration of Zoom fatigue*. Psychiatric Times. https://www.psychiatrictimes.com/view/psychological-exploration-zoom-fatigue

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry.* Beverly Hills, California: Sage Publications.

Lobe, B., & Morgan, D. L. (2021). Assessing the effectiveness of video-based interviewing: a systematic comparison of video-conferencing based dyadic interviews and focus groups. *International Journal of Social Research Methodology*, *24*(3), 301-312. https://doi.org/10.1080/13645579.2020.1785763

Long, K. B. (2020). *How senior institutional research leaders interpret graduation outcomes split by race category*

Luckman, M., & Harvey, A. (2019). The financial and educational outcomes of bachelor degree non-completers. *Journal of Higher Education Policy and Management*, *41*(1), 3-17. https://doi.org/10.1080/1360080X.2018.1553106

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, *22*(4), 719-748. https://doi.org/10.1093/jnci/22.4.719

Marhefka, S., Lockhart, E., & Turner, D. (2020). Achieve research continuity during social distancing by rapidly implementing individual and group videoconferencing with participants: Key considerations, best practices, and protocols. *AIDS and Behavior*, *24*(7), 1983-1989. https://doi.org/10.1007/s10461-020-02837-x

Marschner, I. C. (2018). *glm2: Fitting generalized linear models*. [R package] https://www.rdocumentation.org/packages/glm2/versions/1.2.1/topics/glm2

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka, *Frontiers in Econometrics* (pp. 105-142). New York: Academic Press.

Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *The American Statistician*, *54*(1), 17-24. https://doi.org/10.2307/2685605

Menard, S. W. (1995). Applied logistic regression analysis. In M. S. Lewis-Beck, (Ed.), *Sage University Paper series on Quantitative Applications in the Social Sciences*. London: Sage Publications, Inc.

Morse, R., & Brooks, E. (2021, Sep 12). *How U.S. News Calculated the 2022 Best Colleges Rankings*. U.S. News & World Report, Education. https://www.usnews.com/education/best-colleges/articles/how-us-news-calculated-the-rankings

Muthén, B., & Masyn, K. (2005). Discrete-time survival mixture analysis. *Journal of Educational and Behavioral Statistics*, *30*(1), 27-58. https://doi.org/10.3102/10769986030001027

National Center for Education Statistics. (n.d.-a). *About IPEDS*. U.S. Department of Education. https://nces.ed.gov/ipeds/use-the-data/annual-survey-forms-packages-archived/2021

National Center for Education Statistics. (n.d.-b). *Definitions for New Race and Ethnicity Categories*. U.S. Department of Education. Retrieved April 4, 2022, from https://nces.ed.gov/ipeds/report-your-data/race-ethnicity-definitions

National Center for Education Statistics. (2021a). *Graduation Rates for 4-year institutions reporting on a fall cohort (academic reporters)* (OMB No. 1850-0582 v. 29). U.S. Department of Education.. https://nces.ed.gov/ipeds/use-the-data/annual-survey-forms-packages-archived/2021

National Center for Education Statistics. (2021b). *Fall Enrollment for 4-year degree-granting institutions* (OMB No. 1850-0582 v. 29). U.S. Department of Education. https://nces.ed.gov/ipeds/use-the-data/annual-survey-forms-packages-archived/2021

National Center for Education Statistics. (2021c). *200% Graduation Rates for 4-year institutions* (OMB No. 1850-0582 v. 29). U.S. Department of Education. https://nces.ed.gov/ipeds/use-the-data/annual-survey-forms-packages-archived/2021

National Center for Education Statistics. (2021d). *Outcome Measures* (OMB No. 1850-0582 v. 29). U.S. Department of Education. https://nces.ed.gov/ipeds/use-the-data/annual-survey-forms-packages-archived/2021

National Center for Education Statistics. (2024). *Graduation and Retention Rates: What is the full-time retention rate in postsecondary institutions?* [IPEDS Fall Enrollment component final data (2006-2021) and provisional data (2022)] [Table]. U.S. Department of Education. https://nces.ed.gov/ipeds/TrendGenerator/app/answer/7/32?f=1%3D1%3B4%3D1

Nelson, A. R. (2022, October 27). *Building higher education comparison groups: A Data Science informed approach*. Tech Tips.

Nguyen, V., Lilly, B., & Castro, C. (2014). The exoskeletal structure and tensile loading behavior of an ant neck joint. *Journal of Biomechanics*, *47*(2), 497-504. https://doi.org/10.1016/j.jbiomech.2013.10.053

Niu, L. (2020). A review of the application of logistic regression in educational research: Common issues, implications, and suggestions. *Educational Review*, *72*(1), 41-67. https://doi.org/10.1080/00131911.2018.1483892

Peduzzi, P., Concato, J., Feinstein, A. R., & Holford, T. R. (1995). Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates. *Journal of Clinical Epidemiology, 48*(12), 1503-1510. https://doi.org/10.1016/0895-4356(95)00048-8

Peng, C. Y. , & So, T. S. (2002). Modeling strategies in logistic regression with SAS, SPSS, Systat, BMDP, Minitab, and STATA. *Journal of Modern Applied Statistical Methods*, *1*(1), 147-156. https://doi.org/10.22237/jmasm/1020255720

Peng, C. Y., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, *96*(1), 3-14. https://doi.org/10.1080/00220670209598786

Peng, C. Y., So, T. S., Stage, F. K., & St. John, E. P. (2002). The use and interpretation of logistic regression in higher education journals: 1988-1999. *Research in Higher Education*, *43*(3), 259-293. https://doi.org/10.1023/A:1014858517172

Perry, S. R., Cabrera, A. F., & Vogt, W. P. (1999). Career maturity and college student persistence. *Journal of College Student Retention: Research, Theory & Practice*, *1*(1), 41-58.

Petersen, T. (1985). A comment on presenting results from logit and probit models. *American Sociological Review*, *50*(1), 130-131. https://www.jstor.org/stable/2095348

Pocock, T., Smith, M., & Wiles, J. (2021). Recommendations for virtual qualitative health research during a pandemic. *Qualitative Health Research*, *31*(13), 2403-2413. https://doi.org/10.1177/10497323211036891

Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Flournoy, N., Farewell, V. T., & Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, *34*(4), 541-554. https://doi.org/10.2307/2530374

Purdue University. (n.d.). *Graduation and Retention Rates.* Data & Reports. https://www.purdue.edu/enrollmentmanagement/data-reports/grad-retention-rates.php

Qiu, W. (n.d.). *powerLogisticBin: Calculating power for simple logistic regression with binary predictor*. (Version 0.3.4) [R function] The Comprehensive R Archive Network. https://www.rdocumentation.org/packages/powerMediation/versions/0.3.4/topics/powerLogisticBin

Rouse, D. L. (2018). *Institutional researchers: Perceptions of data influence in internal decision making and policy formulation*

Saldaña, J. (2016). *The Coding Manual for Qualitative Researchers* (3rd ed.). London: SAGE Publications Ltd.

Schwandt, T. (2007). *The SAGE dictionary of qualitative inquiry* (3rd ed.). SAGE Publishing Inc. https://doi.org/10.4135/97814129862681.n354

Shafiq, M. N., Toutkoushian, R. K., & Valerio, A. (2019). Who benefits from higher education in low- and middle-income countries? *The Journal of Development Studies*, *55*(11), 2403-2423. https://doi.org/10.1080/00220388.2018.1528351

Shea, P., & Bidjerano, T. (2018). Online course enrollment in community college and degree completion: The tipping point. *International Review of Research in Open and Distributed Learning*, 19(2), 282-293.

Shea, P., & Bidjerano, T. (2019). Effects of online course load on degree completion, transfer, and dropout among community college students of the state university of New York. *Online Learning*, *23*(4), 6-22. https://doi.org/10.24059/olj.v23i4.1364

Singer, J. D., & Willett, J. B. (1993). It's about time: Using discrete-time survival analysis to study duration and timing of events. *Journal of Educational Statistics*, *18*(2), 155-195. https://doi.org/10.3102/10769986018002155

Stage, F. K. (1990). LISREL: An introduction and applications in higher education. In J. C. Smart (Ed.), *Higher Education: Handbook of Theory and Research* (Vol. VI, pp. 427-466). New York: Agathon Press.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Boston: Pearson

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. New York: Sage Publications, Inc.

Thorndike, R. M. (1978). *Correlational Procedures for Research*. New York: Gardner Press.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89-125. https://doi.org/10.1080/14783363.2013.807677

Tinto, V. (1993). *Leaving College* (2nd ed.). Chicago: University of Chicago Press.

University of Michigan, Office of Budget and Planning. (2021). *The Michigan Almanac, March 2021*. https://obp.umich.edu/campus-statistics/michigan-almanac/

Villalobos Meléndez, A. (2023). *Leveraging institutional research to scaffold data sensemaking and understanding of student equity among community college administrators*

Wilkerson, J. M., Iantaffi, A., Grey, J. A., Bockting, W. O., & Rosser, B. R. S. (2014). Recommendations for internet-based qualitative health research with hard-to-reach populations. *Qualitative Health Research*, 24(4), 561-574. https://doi.org/10.1177/1049732314524635

Zhang, G., Anderson, T. J., Ohland, M. W., & Thorndyke, B. R. (2004). Identifying factors influencing engineering student graduation: A longitudinal and cross-institutional study. *Journal of Engineering Education*, *93*(4), 313-320. https://doi.org/10.1002/j.2168-9830.2004.tb00820.x

Zoom Video Communications. (2022). *Video Conferencing, Cloud Phone, Webinars, Chats, Virtual Events*. Retrieved from Zoom: https://zoom.us/

APPENDIX A

INSTITUTIONAL REVIEW BOARD APPROVAL
FOR THE QUANTITATIVE PHASE

UNIVERSITY OF
NORTHERN COLORADO

Institutional Review Board

Date:	04/15/2022
Principal Investigator:	Angela Rockwell
Committee Action:	**IRB EXEMPT DETERMINATION – New Protocol**
Action Date:	04/15/2022
Protocol Number:	2204037984
Protocol Title:	A mixed methods exploration of the technical and operational considerations influencing institutional researchers' selection of more methodologically appropriate persistence modeling techniques
Expiration Date:	

The University of Northern Colorado Institutional Review Board has reviewed your protocol and determined your project to be exempt under 45 CFR 46.104(d)(704) for research involving

Category 4 (2018): SECONDARY RESEARCH USING IDENTIFIABLE DATA OR SPECIMENS. Secondary research for which consent is not required: Secondary research uses of identifiable private information or identifiable biospecimens, if at least one of the following criteria is met: (i) The identifiable private information or identifiable biospecimens are publicly available; (ii) Information, which may include information about biospecimens, is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained directly or through identifiers linked to the subjects, the investigator does not contact the subjects, and the investigator will not re-identify subjects; (iii) The research involves only information collection and analysis involving the investigator's use of identifiable health information when that use is regulated under 45 CFR parts 160 and 164, subparts A and E, for the purposes of "health care operations" or "research" as those terms are defined at 45 CFR 164.501 or for "public health activities and purposes" as described under 45 CFR 164.512(b); or (iv) The research is conducted by, or on behalf of, a Federal department or agency using government-generated or government-collected information obtained for nonresearch activities, if the research generates identifiable private information that is or will be maintained on information technology that is subject to and in compliance with section 208(b) of the E-Government Act of 2002, 44 U.S.C. 3501 note, if all of the identifiable private information collected, used, or generated as part of the activity will be maintained in systems of records subject to the Privacy Act of 1974, 5 U.S.C. 552a, and, if applicable, the information used in the research was collected subject to the Paperwork Reduction Act of 1995, 44 U.S.C. 3501 et seq.

Carter Hall 2008 | Campus Box 143 | Greeley, CO 80639 | Office 970-351-1910

UNIVERSITY OF
NORTHERN COLORADO

Institutional Review Board

You may begin conducting your research as outlined in your protocol. Your study does not require further review from the IRB, unless changes need to be made to your approved protocol.

As the Principal Investigator (PI), you are still responsible for contacting the UNC IRB office if and when:

- You wish to deviate from the described protocol and would like to formally submit a modification request. Prior IRB approval must be obtained before any changes can be implemented (except to eliminate an immediate hazard to research participants).
- You make changes to the research personnel working on this study (add or drop research staff on this protocol).
- At the end of the study or before you leave The University of Northern Colorado and are no longer a student or employee, to request your protocol be closed. *You cannot continue to reference UNC on any documents (including the informed consent form) or conduct the study under the auspices of UNC if you are no longer a student/employee of this university.
- You have received or have been made aware of any complaints, problems, or adverse events that are related or possibly related to participation in the research.

If you have any questions, please contact the Research Compliance Manager, Nicole Morse, at 970-351-1910 or via e-mail at nicole.morse@unco.edu. Additional information concerning the requirements for the protection of human subjects may be found at the Office of Human Research Protection website - http://hhs.gov/ohrp/ and https://www.unco.edu/research/research-integrity-and-compliance/institutional-review-board/.

Sincerely,

Nicole Morse
Research Compliance Manager

University of Northern Colorado: FWA00000784

APPENDIX B

INSTITUTIONAL REVIEW BOARD APPROVAL FOR THE QUALITATIVE PHASE

UNIVERSITY OF
NORTHERN COLORADO

Institutional Review Board

Date:	03/29/2023
Principal Investigator:	Angela Rockwell
Committee Action:	**IRB EXEMPT DETERMINATION – New Protocol**
Action Date:	03/29/2023
Protocol Number:	2302048368
Protocol Title:	A mixed methods exploration of the technical and operational considerations influencing institutional researchers' selection of more methodologically appropriate persistence modeling techniques
Expiration Date:	

The University of Northern Colorado Institutional Review Board has reviewed your protocol and determined your project to be exempt under 45 CFR 46.104(d)(702) for research involving

Category 2 (2018): EDUCATIONAL TESTS, SURVEYS, INTERVIEWS, OR OBSERVATIONS OF PUBLIC BEHAVIOR. Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met: (i) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects; (ii) Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation; or (iii) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination required by 45 CFR 46.111(a)(7).

You may begin conducting your research as outlined in your protocol. Your study does not require further review from the IRB, unless changes need to be made to your approved protocol.

As the Principal Investigator (PI), you are still responsible for contacting the UNC IRB office if and when:

UNIVERSITY OF
NORTHERN COLORADO

Institutional Review Board

- You wish to deviate from the described protocol and would like to formally submit a modification request. Prior IRB approval must be obtained before any changes can be implemented (except to eliminate an immediate hazard to research participants).
- You make changes to the research personnel working on this study (add or drop research staff on this protocol).
- At the end of the study or before you leave The University of Northern Colorado and are no longer a student or employee, to request your protocol be closed. *You cannot continue to reference UNC on any documents (including the informed consent form) or conduct the study under the auspices of UNC if you are no longer a student/employee of this university.
- You have received or have been made aware of any complaints, problems, or adverse events that are related or possibly related to participation in the research.

If you have any questions, please contact the Research Compliance Manager, Nicole Morse, at 970-351-1910 or via e-mail at nicole.morse@unco.edu. Additional information concerning the requirements for the protection of human subjects may be found at the Office of Human Research Protection website - http://hhs.gov/ohrp/ and https://www.unco.edu/research/research-integrity-and-compliance/institutional-review-board/.

Sincerely,

Nicole Morse

Nicole Morse
Research Compliance Manager

University of Northern Colorado: FWA00000784

APPENDIX C

FOCUS GROUP CONSENT FORM

UNIVERSITY OF NORTHERN COLORADO

Informed Consent Form for Participation in Research

Title of Research Study: A mixed methods exploration of the technical and operational considerations influencing institutional researchers' selection of more methodologically appropriate persistence modeling techniques

Researcher(s): Angela Rockwell
Phone Number: (970) 313-7176 email: rock2487@bears.unco.edu
Research Advisor: William Merchant
Phone Number: (970) 351-2015 email: william.merchant@unco.edu

Procedures: This is a mixed methods study including the analysis of undergraduate retention and completion at a United States regional university and focus groups with institutional research professionals and academic leadership from a variety of institutions across the United States. At this time, we are asking members of academic leadership and institutional research professionals from a diverse selection of higher education institutions within the United States, regardless of job title, to participate in focus groups held in Zoom. In these focus groups, we will discuss the role of institutional research, the value and analysis needs of persistence information at the represented institutions, and gaps in current practice. To focus the discussion, there will be a brief presentation outlining three analysis methods, then the group will discuss perceptions and responses to those approaches.

In total, the focus groups may take about 2 hours to complete. The focus groups will be recorded and you will select a pseudonym to be used in the transcriptions and reporting the findings. Your name and institution will not be associated with your responses and your participation is completely voluntary. For participating, you will receive either a $25 Amazon digital gift card or a $25 contribution to your institution's foundation, at your discretion, at the completion of the focus group. Please note that it is your responsibility to be aware of and follow any gift or compensation policy at your institution that is relevant to receiving compensation for participating in this study. If you elect to receive compensation in the form of the contribution to your institution's foundation, you will receive confirmation at your preferred email address of that contribution no later than two weeks after the focus group is held. If you elect to receive compensation in the form of the Amazon digital gift card, you will receive the link at your preferred email address no later than two business days after the focus group is held.

Questions: If you have any questions about this research project, please feel free to contact Angela Rockwell at rock2487@bears.unco.edu or William Merchant at william.merchant@unco.edu. If you have any concerns about your selection or treatment as a research participant, please contact the University of Northern Colorado IRB at irb@unco.edu or 970-351-1910.

Voluntary Participation: Please understand that your participation is voluntary. You may decide not to participate in this study and if you begin participation you may still decide to stop and withdraw at any time. Your decision will be respected and will not result in loss of benefits to which you are otherwise entitled.

Please take all the time you need to read through this document and decide whether you would like to participate in this research study.

If you agree to participate in this research study, please sign below. You will be given a copy of this form for your records.

______________________________	__________
Participant Signature	Date
______________________________	__________
Investigator Signature	Date

APPENDIX D

PRE-PRESENTATION FOCUS GROUP QUESTIONS

Table 8

Institutional Researcher Focus Group Pre-Presentation Discussion Questions

1. How does institutional research support decision making at your institutions?
2. How is persistence information used at your institution?
3. How sensitive is your institution to changes in persistence rates?
4. How does your stakeholder’s ability to digest and interpret data affect what kinds of analyses you choose to explore persistence?
5. What analytical techniques do you use to explore persistence?
6. How do you handle conflicts between the findings of different persistence analyses?
7. What questions about persistence would you like to answer that you can't answer right now?

Table 9

Administrator Focus Group Pre-Presentation Discussion Questions

1. How does institutional research support decision making at your institution?
2. What is the dynamic between institutional research and your specific offices?
3. Are administrators at your institutions data literate?
4. Who decides what the research question is?
5. How is persistence information used at your institutions?
6. Which subgroups do you typically track your persistence on?
7. How well does the persistence information you receive meet your needs in understanding student persistence?
8. Do you have ancillary needs, where qualitative information or more of a mixed methods approach would be more useful for understanding student persistence?
9. What sources do you use to frame the information you get internally with student persistence?
10. How is information gathered by front line workers integrated into persistence models?

APPENDIX E

POST-PRESENTATION FOCUS GROUP QUESTIONS

Table 10

Institutional Researcher Focus Group Post-Presentation Discussion Questions

1.	What questions you think the approaches would be best at answering?
2.	What limitations you see to using them at your institutions?
3.	As experienced institutional researchers, what advice would you have for somebody in my position on how to use these tools?

Table 11

Administrator Focus Group Post-Presentation Discussion Questions

1.	What limitations do you see to using these quantitative approaches at your institution?
2.	What benefits do you see to using these quantitative approaches at your institution?
3.	What questions do you think they could answer?
4.	How, what barriers do you see to your institutions having to using a mixed methods or a qualitative approach?