ASSOCIATION FOR INSTITUTIONAL RESEARCH
Data and Decisions for Higher Education

Back to All Proposals

PROPOSAL DETAILS

Personal Information	
Name	Dr. Jeffrey Christopher Sklar
Informal Name	Jeff
Affiliation	Cal Poly Corporation on behalf of Cal Poly
Unit/Department	Statistics Department
Title	Associate Professor
Year began this position	2005
Email	jsklar@calpoly.edu
Alternate Email	jcsklar@gmail.com
Cell Phone	805-252-6012
Preferred Mailing Address	California Polytechnic State University 1 Grand Avenue San Luis Obispo, California 93410-9001 United States Phone: 805-756-6353
Secondary Address	188 Miramar Lane Pismo Beach, California 93449 United States

Financial Representative	
Name	Melissa R Mullen
Informal Name	
Institution / Affiliation	Cal Poly Corporation on behalf of Cal Poly
Unit / Department	Sponsored Programs, Cal Poly Corporation
Title	Sponsored Programs Manager
Preferred Mailing Address	One Grand Avenue California Polytechnic State University
City	San Luis Obispo
Country	United States
State	California
Zip/Postal Code	93407-0830
Email	sponprog@calpoly.edu
Phone	805-756-1123
Fax	805-756-5588

Secondary Investigator(s)

Project Description I

Title:

The Impact of Change of Major on Time to Bachelor's Degree Completion with Special Emphasis on STEM Disciplines: A Multilevel Discrete -Time Hazard Modeling Approach

Statement of the research problem and national importance:

Although the current literature addresses an array of student and institutional factors associated with college dropout and degree completion (DesJardins, Ahlburg, and McCall, 1999; 2002; Titus, 2004; 2006(a); 2006(b); Osegura and Rhee, 2009; Chen and DesJardins, 2010; Chen, 2012), very little research has directly examined the impact of change of major behavior on postsecondary outcomes, particularly at the national level. As a potential explanation for delaying graduation and lowering 4-year and 6-year graduation rates, it would be in the best interests of students, as well as campus administrators, to understand the implications that changing major has on persistence and graduation rates. Figures for the proportion of all first-time full-time degree seeking students at four-year institutions who change their major at least once vary by data used, but the percentage has been reported as high as 75% (Kramer, Higley, and Olsen, 1994). According to more recent single-institution reports (Micceri, 2001; Filce, 2010; Foraker, 2012; Sklar, 2013), the proportion of students who change majors at least once (within 6 years) ranges between 21% and 50%.

Yet even with only a few studies completed, there are still conflicting conclusions as to whether changing major increases the likelihood of graduating college and/or increases the time to graduate (Micceri, 2001; Foraker, 2012; Sklar, 2013). Single-institution studies typically use simple descriptive measures and non-model-based methods that ignore student background characteristics, as well as the longitudinal nature of college enrollment data (the study by Sklar (2013) is an exception). It is likely that the association between changing major and graduation outcomes depends on a variety of student and institutional characteristics. Needless to say, although change of major activity is prevalent throughout postsecondary institutions, the impact on graduation outcomes is not completely understood. The proposed research will attempt to fill an apparent void in the literature by examining the association between changing major and graduation outcomes, specifically risk of graduation and time-to-graduation, at the national level. In addition, due to the current push to retain and graduate students in science, technology, engineering, and mathematics (STEM) fields (NSB, 2007; PCAST, 2012), primary (student-level) control variables to be used in the proposed analysis include STEM status of the first declared major and the STEM-status of the new major for students who switch. Recent figures indicate that the proportion of bachelor's degree seeking students beginning as STEM majors who eventually changed to a non-STEM major was 28.1% (Chen & Ho, 2012), suggesting that the technical classification of the major may be an important consideration when deciding to switch.

Identifying important factors that impact graduation is crucial at the national level, considering the high financial costs associated with attending college, as well as low overall 6-year graduation rates. Annual costs to attend public and private 4-year institutions averaged $20,100 and $39,800, respectively, for the 2010-2011 academic year, while only about 58 percent of first-time, full-time students who began seeking a bachelor's degree at a 4-year institution in fall 2004 completed a bachelor's degree at that institution within 6 years (Aud et al., 2012). The results of a rigorous examination of change of major behavior may help students make more informed decisions regarding a change of major if they know the additional time, risk, and financial costs that could be potentially incurred if they switch. In addition, because a primary goal of campus administrations is to graduate students in a timely manner, institutions might consider revisions to their change of major policies to facilitate (or restrict) the transition to a new major if the effect of switching majors on graduation outcomes is substantial.

The primary aims of the proposed research study are to: 1) Investigate the impact of change of major behavior on student persistence and graduation rates over time at the national level; 2) Examine whether the impact depends on a variety of student and institutional-level characteristics including STEM classification of first declared major, gender, ethnicity, and public/private status of the institution; and 3) Estimate the expected additional time to graduate for students who switch major for designated subgroups of students. The study will also contribute to the STEM literature by investigating and comparing the persistence and graduation profiles over time for students who change from STEM to non-STEM majors, non-STEM to STEM majors, and those who persist in STEM majors. These objectives will be addressed by developing an appropriate multilevel discrete-time hazard model using restricted-use data from the 2004/2009 Beginning Postsecondary Students Longitudinal Study (BPS:04/09)

Review the literature and establish a theoretical grounding for the research:

Prior Research

There is a large body of literature focused on examining student and/or institutional factors related to postsecondary outcomes. Studies investigating student-level characteristics have revealed that high school GPA, socioeconomic status, gender, SAT and ACT scores, and social and academic integration are associated with persistence outcomes (Astin, 1993; Tinto, 1993; Cabrera, Nora, and Castañeda, 1993). Studies examining the effects of institutional factors have found that size, control, selectivity, expenditures and financial aid awarded, and retention climate are significantly associated with persistence and degree attainment (Berger & Milem, 2000; Titus, 2004; 2006(a); 2006(b); Osegura and Rhee, 2009; Chen, 2012). In addition, various researchers have focused on factors associated with the selection of a college major or persistence in a college major (Porter and Umbach, 2006; Pike, 2006; Allen and Robbins, 2008) using component's of Holland's theory of careers (Holland, 1997; Feldman, Smart, and Ethington, 1999). Although there is an abundance of work on student retention and degree completion, as well as college major choice and persistence in major, there are very few studies that specifically address how persistence (or non-persistence) in a major contributes to the risk of graduating. It has been suggested that decreasing the number of students who change major may improve university graduation rates (Allen & Robbins, 2008), but no substantial evidence has been provided to support this claim.

Various single institution studies have been conducted investigating the impact of changing major, but results tend to conflict. Micceri (2001) reported that students at one university who had changed their major at least once had a higher overall graduation rate than those students who had not changed majors. Only raw percentages were provided and results were not adjusted for background student characteristics. Foraker (2012) found that changing major after the second year was associated with lower graduation rates and longer time-to-graduation; however, the methods used were also descriptive and ignored any background student characteristics. In one of the few single-institution studies that included change of major as a predictor of graduation, Kreysa (2007) found that for students requiring remedial courses, changing majors increased the likelihood of graduation, but for non-remedial students, changing majors decreased the likelihood of graduating. Although enlightening, the study did not focus solely on the impact of changing major, and was limited in the sense that the classification of the new and previous declared major (e.g. STEM versus non-STEM) was not investigated. In studies that have addressed persistence in STEM majors, researchers have focused primarily on student and institutional characteristics associated with STEM degree completion or leaving a STEM major (Crisp, Nora, & Taggart, 2009; Eagan, Hurtado, & Chang, 2010; Chang, Sharkness, Newman, & Hurtado, 2010). In contrast to these studies where STEM degree completion or leaving a STEM major have served as the outcomes of interest, the proposed study would examine persistence in a STEM major as an explanatory variable of interest, and time-to-graduation as the specific outcome variable.

A more thorough examination of the effects of changing major at a single institution has been conducted by Sklar (2013). By applying discrete-time survival analysis (event history) techniques to 6-year enrollment data on first-time freshmen at a large public mostly undergraduate university, Sklar (2013) found that after controlling for several background student variables, the effect of changing major on the risk of graduating significantly depended on the categorization of the first declared major, quarterly GPA, and year that the switch occurred, but did not significantly depend on ethnicity or gender of the student. Furthermore, switching to a STEM major tended to decrease the risk of graduating at the end of four years (compared to students who persisted in their major); however, switching to a non-STEM major in the second year tended to increase the risk of graduating at the end of four years. In addition, students who switched to a STEM major during any year of enrollment tended to take longer to graduate, while students who switched to a non-STEM major tended to take less time to graduate than students who persisted in their major. Although illuminating, the results of this investigation are only applicable to other colleges and universities with student-level and institutional-level characteristics similar to those of the study institution. If results are to be extended to a nationally representative population of first time freshmen pursuing a bachelor's degree, as well as college administrators, then analysis of a national level data set needs to be performed.

Conceptual Framework

The conceptual framework of the proposed research study is adapted from various theories and empirical models of student persistence/dropout. Traditional theories of student persistence have emphasized academic integration (Tinto, 1975; 1993; Bean, 1980; 1990; Berger & Milem, 2000) as an important predictor of student retention, and over the years various studies have confirmed its significance (Pascarella & Chapman, 1983; Cabrera, Nora, & Castañeda, 1993; Chen, 2012). A good match between institutional environment and student commitment results in better integration into the social and academic domains of college life, increasing the likelihood of persisting through college until degree completion. Furthermore, academic integration and commitment to remain at the institution are associated closely with satisfaction with academic environment (Tinto, 1993). It is arguably justified and noted by Allen and Robbins (2008; 2010) that major persistence and changing major are important indicators of satisfaction (or dissatisfaction) with academic environment.

In addition, recent conceptual models have emphasized the integration of institutional characteristics with student level factors when examining persistence (Titus, 2004; Oseguera & Rhee, 2009); however, studies by these authors focused on the role of institution level variables, and have not integrated the longitudinal nature of college enrollment data. The framework for this study incorporates institutional-level factors, but focuses specifically on student behavior, i.e. change of major. Hence, the final conceptual frame for this study posits that change of major behavior (student-level) is directly associated with the risk of graduation over time, but that the association will depend to an extent on student demographic and academic characteristics, as well as institutional-level variables.

Describe the research method that will be used:

Using multilevel statistical models, the proposed study will investigate the impact of change of major behavior on persistence and graduation over time by addressing the following research questions:

1) Adjusting for student-level and institutional-level variables, does the risk of completing a bachelor’s degree depend on change of major behavior, i.e. does the risk significantly vary between students who persist in their first declared major, switch to a STEM major, or switch to a non-STEM major?

2) Adjusting for student-level and institutional-level variables, does the association between changing major and risk of degree completion depend on student-level and institutional-level characteristics including type of first declared major (STEM versus non-STEM), gender, ethnicity, and public/private status of the institution?

3) What is the expected additional amount added to graduation time for students who switch major, and how does this amount differ between students who switch to STEM and those who switch to non-STEM majors?

Since the research questions focus on the longitudinal process of persistence and degree completion, survival analysis (also called event history analysis) techniques are appropriate. Survival analysis techniques incorporate “censoring,” a particular feature common to time-to-event data that occurs when the event of interest occurs after observation on the subject ends. Censoring is especially relevant to college enrollment data because students may drop out during their college career or may graduate after the data collection period ends. Furthermore, since the metric used to measure time until graduation will be measured in years, it is more appropriate to use *discrete-time* survival analysis methods (Singer and Willett, 2003). For the proposed study, time until degree completion is measured in academic years from fall 2003. Students who graduated by the end of 2009 will have complete event times, i.e. the time until graduation will be known exactly. Those students who were still enrolled in the first institution at the end of 2009, or dropped out or stopped out of the first institution without re-enrolling and graduating by 2009 will have censored event times, i.e. their times to graduate are known only up to the year of last enrollment. Finally, due to the nested structure of the BPS data (students within institution), multilevel discrete-time hazard models will be used. Multilevel models will account for the possibility that students within a particular institution may behave more similarly than students attending other institutions. Multilevel discrete-time survival methods have not been used extensively in the persistence literature; however, some recent examples include works by Wao (2010), who used multilevel discrete-time hazard models to investigate the time until doctorate degree completion, and by Chen (2012) who examined institutional characteristics associated with dropout. In addition, the proposed methods have been explored in sociology (Barber, Murphy, Axinn, and Maples, 2000) and demography (Browning, Leventhal, and Brooks-Gunn, 2004).

The final sample to be analyzed will be limited to full-time, first-time, bachelor's degree seeking students who enrolled in a four-year institution during the 2003-2004 academic year, and who had a declared major at the time of fall 2003 enrollment. Since this study focuses on change of major and its impact at the first institution of attendance, students who transfer from their first institution will be removed from the analyses. The majority of the predictor variables that will be included in the hazard models are time-invariant, i.e. their values do not change over time (e.g. gender and ethnicity); however, change of major can occur at different times in the study. Change of major information can be extracted from the MAJ06CHG and MAJ09CHG variables in BPS:04/09, which provide the number of times students had changed their major as of 2006 and 2009. For those who had changed majors, the variables MAJ06A and MAJ09A provide the new major fields in 2006 and 2009. These majors can then be further classified as STEM or non-STEM. Using change of major indicators from follow-up stages of the BPS, a time-varying predictor to indicate whether a change of major occurred will be derived.

To implement the appropriate hazard models, data must be converted from the person-level format, i.e. one set of measurements for each individual, to a person-period format, i.e. each individual has a record of measurements for each time period (Singer and Willett, 2003). Once the person-period data is arranged, a hierarchical logistic regression model will be fit to the data to obtain the estimated coefficients for the multilevel discrete-time hazard model. Although standard software for implementing multilevel models such as HLM software (Raudenbush, Bryk, Cheong, and Congdon, 2004) is available, the free open-source R software package (R Development Core Team, 2011) may also be used to fit the multilevel discrete-time hazard models. In particular, the R library HGLMMM (Molas and Lesaffre, 2011) contains a suite of functions that can be used to fit a variety of hierarchical generalized linear models. To account for the unequal probability of selection into BPS:04/09, appropriate weighting of the individuals in the final sample will be performed.

The statistical models used to specifically address Research Questions (1)-(3) are provided in the attached "Statistical Models AIR2013" document. The predictor variables that will be used in the models to address Research Questions (1)-(3) are presented in the attached "List of Variables AIR2013" document and are separated by Student Level (Level 1) and Institutional Level (Level 2).

Uploaded Appendix Document(s):

- List of Variables AIR2013
- Statistical Models AIR2013

Project Description II

Will you use NCES target dataset? Yes

Please check all NCES datasets that apply
- Beginning Postsecondary Student (BPS) Longitudinal Study and Transcript Data
- IPEDS Fall Enrollment (EF)
- National Postsecondary Student Aid Study (NPSAS)

Explain why each dataset best serves this research. Include a variable list for each dataset used.

The primary data that will be used to address the proposed research questions and hypotheses come from the 2004/2009 Beginning Postsecondary Students Longitudinal Study (BPS:04/09) and the Integrated Postsecondary Education Data System (IPEDS) from 2003 and 2004. Additional information from the 2004 National Postsecondary Student Aid Study (NPSAS:2004) and College Board included on the BPS:04/09 data file (CD) will also be used. Data from BPS:04/09 as well IPEDS and NPSAS data provide the opportunity to explore the effects of change of major controlling for a variety of student level variables and how these effects may vary by institutional level. The appendix titled "List of Variables AIR2013" provides the names and brief descriptions of the variables, as well as the corresponding source dataset in parentheses, that will be used as Level 1 and Level 2 predictors in the multilevel discrete-time hazard model. All predictor variables have been found to be of significant importance in the student persistence literature.

Variables used from NPSAS, IPEDS, College Board and BPS include: GENDER (NPSAS), RACE (NPSAS), C04027 (NPSAS), C04028 (NPSAS), PELL04 (NPSAS), TUITION2 (IPEDS), ACAINX04 (NPSAS), SOCINX04 (NPSAS), HCGPAREP (College Board), TESATMDE (College Board), TESATVDE (College Board), HCYSENGL (College Board), HCYSMATH (College Board), HCYSSCIE (College Board), MAJ04A (NPSAS), MAJ06A (BPS), MAJ09A (BPS), MAJ06CHG (BPS), MAJ09CHG (BPS), UNITID (BPS), LOCALE (IPEDS), ENRLSIZE (IPEDS), PCT_MIN (IPEDS), FGRNT_P (IPEDS), SELECTV2 (IPEDS), CNTLAFFI (IPEDS), DEG1 (BPS), DEGDT1 (BPS). Derived Variables include: DATEBEG, DATEEND, STEM04, CHANGE06, CHANGE09, and CHMAJ(tij).

Will you use NSF target dataset? No

Explain why each dataset best serves this research. Include a variable list for each dataset used.

Will you address the NPEC focus topic? Yes

If yes, please briefly describe:
Although the current research proposal does not directly address what types of data students and their families might utilize in the college search process, the results of the study will provide information regarding graduation rates and typical graduation times that is not typically available to students which could be used in the selection process. One of the most important pieces of information that prospective students want to know about a specific institution is the percent of students who graduate; however, conventional 4-year or 6-year graduation rates published by individual institutions do not directly apply to students who eventually change major. As an example, the single-institution study conducted by Sklar (2013) found that of the students who had changed their major at least once, roughly 88% had graduated by the end of six years, while only about 74% of students who persisted in their first declared major graduated within six years. If change of major is as routine as the literature suggests, and time to graduation and likelihood of graduation are significantly affected by changing major, then students should consider this information not only if/when they are deciding if they should change major, but also when they are considering a major to declare when applying to college.

Project Description III

Provide a timeline of key project activities:

May-July 2013: Organize Data and Compute Descriptive Statistics

- Appropriate samples and variables will be extracted from the BPS:04/09 data file during the spring and early summer of 2013.
- Create new variables to be used in addition to those provided in the BPS:04/09 data file.
- Prepare data to be analyzed by conversion to person-period form (Singer & Willet, 2003).
- Compute descriptive statistics for the student-level quantitative variables of, such as SAT score and high school GPA, as well as the institutional-level variables of interest such as enrollment size and percent minority enrollment. In addition, contingency tables will be constructed for student and institutional-level variables to display the relationships between pairs of important categorical variables, for example STEM classification of first declared major and ethnic category.

August-September 2013: Explore hazard models and summarize results

- Multilevel discrete-time hazard models will be explored, developed, and tested to address the primary research questions.
- Summarize preliminary model results, interpret estimated model coefficients, and prepare estimated hazard and survival profiles for several groups of interest.

October-December 2013: Prepare AIR materials

- Summarize responses to the three primary research questions of the study.
- Submit proposal to present at 2014 AIR Annual Forum.
- Prepare and submit mid-year progress report.

January-March 2014: Manuscript preparation

- Begin preparation of publishable manuscript and presentation for AIR Forum

April-May 2014: Finalize reports and disseminate results

- Present at AIR Forum.
- Complete and submit final report to AIR.
- Complete and submit manuscript for publication.

List deliverables such as research reports, books, and presentations that will be developed from this research initiative:

Based on the results of the proposed research study, mid-year reports will be submitted to AIR, the AIR Forum presentation is planned, and at least one manuscript will be submitted to a peer-reviewed journal. Possible journals include *Research in Higher Education*, *Journal of Higher Education*, *Journal of College Student Retention*, *American Educational Research Journal,* or *Educational Assessment, Evaluation and Accountability*.

Describe how you will disseminate the results of this research:

The impact of changing major on risk of graduation and time-to-graduation is potentially of interest to diverse groups of individuals, including academic scholars, campus administrators, academic advisors, and students. Presenting at various conferences, including the 2014 AIR Forum and possibly the 2014 AERA Conference, will be the most effective method for disseminating the outcomes of the investigation to scholars. Because significant results of the study may have different policy implications at individual institutions, a potential avenue to reach a large audience of advisors and administrators is through a presentation at the 2014 National Academic Advising Association (NACADA) Annual Conference.

Should the national-level findings prove to be significant, then several research offices at local colleges and universities (e.g. the Office of Institutional Planning and Analysis at Cal Poly) will be contacted to determine the level of interest in change of major impact and for possible distribution of the key findings of the proposed study. Institutions could communicate and publicize some of the national-level findings through their research and planning offices and corresponding websites to prospective and current students.

Finally, a manuscript presenting the results and policy implications of the study will be submitted to one (or more) suitable journals (specific titles provided in the List of Deliverables section) after the research has been revised based on peer-review by experts in the areas of postsecondary persistence and retention.

Provide a reference list of sources cited:

Allen, J. and Robbins, S.B. (2008). Prediction of college major persistence based on vocational interests, academic preparation, and first-year academic performance. *Research in Higher Education, 49*(1)*,* 62-79.

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.

Astin, A.W. (1993). What Matters in College: Four Critical Years Revisited. Jossey-Bass, San Francisco, CA.

Aud, S., Hussar, W., Johnson, F., Kena, G., Roth, E., Manning, E., Wang, X., and Zhang, J. (2012). *The Condition of Education 2012* (NCES 2012-045). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved 12-12-12 from http://nces.ed.gov/pubsearch.

Barber, Jennifer S., Susan A. Murphy, William G. Axinn and Jerry Maples. (2000). Discrete-time multilevel hazard analysis. *Sociological Methodology*, *30*, 201-235.

Berger, J. B. & Milem, J. F. (2000). Organizational behavior in higher education and student outcomes. In J. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 15, pp. 268–338). New York: Agathon Press.

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education*, *12*(2), 155–187.

Bean, J. P. (1990). Why students leave: Insights from research. In D. Hossler & J. P. Bean (Eds.), *The strategic management of college enrollments*. San Francisco: Jossey-Bass.

Browning, C., Leventhal, T., & Brooks-Gunn, J. (2004). Neighborhood structural disadvantage, collective efficacy, and the timing of first intercourse. *Demography, 41(4),* 697-720.

Cabrera, A.F., Nora, A., & Castañeda, M.B. (1993). College persistence: Structural equations modeling test of an integrated model of student retention. *The Journal of Higher Education*, *64*(2), 123-139.

Chang, M.J., Sharkness, J., Newman, C., & Hurtado, S. (2010). *What matters in college for retaining aspiring scientists and engineers*. Paper presented at the annual meeting of the American Educational Research Association, Denver, CO.

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

Chen, R. and DesJardins, S.L. (2010). Investigating the impact of financial aid on student dropout risks: Racial and ethnic differences. *The Journal of Higher Education*, *81*(2), 179-208.

Chen, X. & Ho, P. (2012). STEM in postsecondary education: Entrance, attrition, and coursetaking among 2003-04 beginning postsecondary students (NCES 2013-152). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved 12-12-12 from http://nces.ed.gov/pubsearch.

Chen, X., & Weko, T. (2009). Students who study science, technology, engineering, and mathematics (STEM) in postsecondary education: NCES 2009-161. Washington, DC: National Center for Education Statistics.

Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Educational Research Journal*, *46*(4), 924-942.

DesJardins, S.L., Ahlburg, D.A., & McCall, B.P. (1999). An event history model of student departure. *Economics of Education Review, 18*, 375–390.

DesJardins, S.L., Ahlburg, D.A., & McCall, B.P. (2002). A temporal investigation of factors related to timely degree completion. *The Journal of Higher Education, 73*(5), 555-581.

Eagan, M.K., Hurtado, S., and Chang, M.J. (2010). What Matters in STEM: Institutional Contexts That Influence STEM Bachelor's Degree Completion Rates. Paper presented at the 2010 Annual Meeting of the Association for the Study of Higher Education Indianapolis, IN.

Feldman, K.A., Smart, J.C., & Ethington, C. (1999). Major field and person-environment fit: Using Holland's theory to study change and stability in college students. *The Journal of Higher Education, 70*(6), 642-669.

Filce, J. (2010). Meandering Pathways –Understanding the Effects of Changing Majors. Presentation prepared for the Institutional Research and Planning, Humboldt State University, CAIR -November 18, 2010.

Foraker, M.J. (2012). Does Changing Majors Really Affect the Time to Graduate? The Impact of Changing Majors on Student Retention, Graduation, and Time to Graduate. Paper presented at the Annual Meeting of the Association for Institutional Research (New Orleans, LA, June 2-6, 2012).

Holland, J.L. (1997). *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments (3rd Ed.)*. Odessa, FL: Psychological Assessment Resources.

Kramer, G.L., Higley, H., & Olsen, D. (1994). Changes in academic major among undergraduate students. *College & University*, 69, 88–90.

Kreysa, P.G. (2007). The impact of remediation on persistence of under-prepared college students. *Journal of College Student Retention, 8*(2), 251-270.

Micceri, T. (2001). Change Your Major and Double Your Graduation Chances. Paper presented at the Annual Meeting of the Association for Institutional Research (Long Beach, CA, June 3-6, 2001).

Molas, M. and Lesaffre, E. (2011). Hierarchical Generalized Linear Models: The R Package HGLMMM. *Journal of Statistical Software*, *39* (13), 1-20.

National Science Board (NSB). (2007). *A National Action Plan for Addressing the Critical Needs of the U.S. Science, Technology, Engineering, and Mathematics Education System.* Arlington, VA: National Science Foundation.

Oseguera, L. & Rhee, B. (2009). The Influence of institutional retention climates on student persistence to degree completion: A multilevel approach. *Research in Higher Education*, *50*(6), 546-569.

Pascarella, E.T., & Chapman, D.W. (1983). A multi-institutional, path analytic validation of Tinto's model of college withdrawal. *American Educational Research Journal*, *20*(1), 87-102.

Pike, G.R. (2006). Vocational preferences and college expectations: An extension of Holland's principle of self-selection. *Research in Higher Education*, *47*(5), 591-612.

Porter, S.R. & Umbach, P.D. (2006). College major choice: An analysis of person-environment fit. *Research in Higher Education*, *47*(4), 429 -449.

President's Council of Advisors on Science and Technology (U.S.), & United States. Executive Office of the President. (2012). *Report to the president, engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, D.C.: Executive Office of the President, President's Council of Advisors on Science and Technology.

R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Raudenbush, S. W., Bryk, A. S., Cheong, Y., & Congdon, R. T. (2004). *HLM 6: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International.

Singer, J., & Willett, J. (1993). It's about time: Using discrete time survival analysis to study the duration and the timing of events. *The Journal of Educational Statistics, 18*(2), 155-195.

Singer, J., & Willett, J. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*, Oxford University Press, New York, NY.

Sklar, J. (2013). Exploring the impact of changing major on time to graduation at a primarily undergraduate university. A report in progress based on work completed for Extramural Funding Initiative project "Assessing the Impact of Change of Major on College Persistence and Bachelor's Degree Attainment Using Discrete-Time Hazard Models".

Tinto, V. (1975). Dropouts from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, *45*(1), 89–125.

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

Titus, M. A. (2004). Examining the influence of institutional context on persistence at four-year colleges and universities: A multilevel approach. *Research in Higher Education, 45(7)*, 673 - 699.

Titus, M. A. (2006a). Understanding the influence of the financial context of institutions on student persistence at four-year colleges and universities, *The Journal of Higher Education*, *77*(2), 353-375. *Research in Higher Education, 47(4)*, 371-398.

Titus, M. A. (2006b). Understanding college degree completion of students with low socioeconomic status: The influence of the institutional financial context.

Wao, H.O. (2010). Time to the doctorate: Multilevel discrete-time hazard analysis. *Educational Assessment, Evaluation and Accountability, 22* (3), 227-247.

IRB Statement

Statement of Institutional Review Board approval or exemption:

Prior to award acceptance and beginning data analysis (April 2013), a Human Subjects Protocol Approval Form will be submitted to the IRB with the expectation that the proposed research will be exempt from further review. Cal Poly's Institutional Review Board complies with Federal requirements. The Federal Wide Assurance number is 00000342 and the Human Subjects Review Committee Registration number is 000001118.

Statement of Use of Restricted Datasets

The proposed research requires the use of the restricted data from the 2004/2009 Beginning Postsecondary Students Longitudinal Study (BPS:04/09). The BPS:04/09 restricted data is the most recent large scale longitudinal data available on the attributes and outcomes of postsecondary students and is not available on a public-use CD. In addition, online NCES statistical tools such as PowerStats and the Data Analysis System (DAS) cannot perform the appropriate analysis methods on the BPS data to address the research questions.

Biographical Sketch

Jeffrey Sklar's Biography Sketch

Jeffrey Sklar is an Associate Professor of Statistics at California Polytechnic State University, San Luis Obispo. Prior to his current position, he was a Postdoctoral Researcher at the Gevirtz Graduate School of Education (GGSE) at the University of California, Santa Barbara, as well as a Lecturer of statistics courses in the GGSE and the Department of Statistics and Applied Probability at UC Santa Barbara. Dr. Sklar received his PhD in Statistics from UC Santa Barbara in 2003.

He currently teaches a variety of introductory statistics courses, as well as upper-level applied courses such as multivariate statistics, regression analysis, and survival analysis. In addition, Dr. Sklar has served as the Statistics Department consultant, addressing research questions posed by university faculty, students and staff using a variety of methods, including fixed and mixed effects models (hierarchical linear models), log-linear models, regression models, and multivariate and nonparametric techniques. He has extensive knowledge of various statistical software programs including R, SPSS, and Minitab.

Although Dr. Sklar's formal training is in theoretical and applied statistics (specifically nonparametric regression), he has strong interests in educational policy, practices, and methods that address and promote persistence and graduation, especially among students majoring in STEM fields. In varying capacities, he has worked on several educational research projects throughout his academic career. His investigation of the impact of changing major on graduation for students at California Polytechnic State University was funded by an Extramural Funding Initiative grant ($2000 and 8 units of release time). Dr. Sklar is currently collaborating with faculty from the Biological Sciences Department and the Women's and Gender Studies Department examining the attitudes of freshmen and sophomores majoring in STEM fields to increase understanding of the factors contributing to problems in retention of science and mathematics majors, and to identify factors which may contribute to the gaps in retention for first-generation students, low-income students and underrepresented minorities. In another ongoing project, Dr. Sklar is working with administrators from the San Luis Obispo County Office of Education to create logistic regression models to predict whether ninth grade students will be credit deficient (fail at least one class) based on eighth grade characteristics such as GPA, attendance, and a variety of other background predictors.

Dr. Sklar has also served as a research assistant for an American Educational Research Association (AERA) funded project (NSF grant #RED-9452861) and as a Postdoctoral Researcher (and later Senior Researcher) for another project funded by the National Science Foundation (NSF) (NSF#0352519). The AERA project investigated the extent to which the SAT predicts college grade point average for language minorities and the extent to which the SAT predicts college graduation for language minorities and native English speakers using regression and discrete-time survival analysis. This research required analysis of restricted-use NCES data from the High School and Beyond survey. For the NSF project, he assisted in the development and evaluation of web-based instructional modules in educational measurement and statistics to help school personnel acquire the "assessment literacy" required for their respective positions. Results from the projects eventually lead to several peer-reviewed publications in various educational journals.

Select Publications and Technical Reports:

Sklar, J. (2013). Exploring the impact of changing major on time to graduation at a primarily undergraduate university. A report based on work completed for Extramural Funding Initiative project "Assessing the Impact of Change of Major on College Persistence and Bachelor's Degree Attainment Using Discrete-Time Hazard Models".

Sklar, J., Wu, J., Meiring W., & Wang Y. (2013). Non-parametric regression with basis selection from multiple libraries. To appear in *Technometrics*. Online version available at: http://www.tandfonline.com/doi/abs/10.1080/00401706.2012.739104

Sklar, J. & Zwick, R. (2009). Multimedia presentations in educational measurement and statistics: Design considerations and instructional approaches. *Journal of Statistics Education*, *17*(3).

Zwick, R., Sklar, J., Wakefield, G., Hamilton, C., Norman, A., & Folsom, D. (2008). Instructional tools in educational measurement and statistics (ITEMS) for school personnel: Evaluation of three web-based training modules. *Educational Measurement: Issues and Practice*, *27* (2), 14-27.

Zwick, R. and Sklar, J. (2005). Predicting college grades and degree completion using high school grades and SAT scores: The role of student ethnicity and first language. *American Educational Research Journal, 42*(3), 439-464.

Zwick, R. and Sklar, J. (2005). A note on the standard errors for survival curves in discrete time survival analysis. *Journal of Educational and Behavioral Statistics, 30*(1), 75-92.

Zwick, R., Brown, T., & Sklar, J. (2004). California and the SAT: A reanalysis of University of California admissions data. Center for Studies in Higher Education, UC Berkeley, Research and Occasional Papers Series, July 2004, http://ishi.lib.berkeley.edu/cshe/publications/papers/papers.html

Budget Requirements

Jeffrey Sklar' Budget

Personnel-Time on Project
%(FTE) Academic Year: 35.56
%(FTE) Summer: 0.00

Personnel-Salary & Benefits
Academic Year: $ 107199.00
Summer: $ 0.00

Graduate Research Assistant's Budget

Personnel-Time on Project
%(FTE) Academic Year: 0.00
%(FTE) Summer: 0.00

Personnel-Salary & Benefits
Academic Year: $ 0.00
Summer: $ 0.00

Total Salary and Wages: $38119.96

Travel: $1500.00
Other travel related expenses: $0.00
Other research expenses: $380.00
Total Request: $39999.96

Funding History

This project has not received funding, and no funding is pending for the proposed research. Funding is being sought to be used in conjunction with a one quarter sabbatical that has been planned for Fall 2013. The proposed budget is being requested to relieve Dr. Sklar from a substantial portion of his teaching responsibilities in the winter and spring quarters of 2014, reducing his typical load in those quarters from 3 courses to 1 course per quarter.

<u>Research Question 1</u>

To address Research Question 1, the following multilevel discrete-time hazard model is proposed:

Student-Level (Level 1) Model:

$$\ln\left(\frac{h_{ijt}}{1-h_{ijt}}\right)=\sum_{k=1}^{6}\alpha_{jk}D_{ijk}+\sum_{p=1}^{P}\beta_p X_{pij}+\sum_{c=1}^{2}\delta_c (CHMAJ_{ijt})_c , \quad (1)$$

$$i=1,2,\ldots,I,\ \ t=1,\ 2,\ \ldots,6,\ \ j=1,\ 2,\ \ldots,J,$$

where:

- h_{ijt} is the hazard (or risk) of graduating for person *i* at institution *j* during year *t* , i.e. it is the probability that student *i* graduates from institution *j* during year *t* given that the student had not previously graduated. Note that the final number of students and institutions to be used is not exactly known, so *I* and *J* are purposely left undesignated in the model.

- α_{jk} is the (baseline) logit hazard of graduating at institution *j* in year *k* (i.e., when Level 1 predictors are set to 0), and depends on the set of Level 2 predictors (see Level 2 and description below).

- D_{ijk} is a dummy variable taking value 0 for student *i* at institution *j* for time periods when the student has not graduated, and the value 1 during the time period when the student has graduated.

- X_{pij} , $p=1,\ldots,P$, is the *p*th student-level (time-invariant) predictor for person *i* at institution *j*. The student-level predictors include background/demographic, financial aid and expenses, pre-college preparation, and measures of social and academic integration provided in the “List of Variables AIR2013” document. Note that since the final number of predictor terms is not precisely known (due to possible exclusions of certain groups) *P* is purposely not defined.

- $CHMAJ_{ijt}$ is a time varying categorical predictor indicating whether the student has changed majors (at least once) and taking the following values:

$$CHMAJ_{ijt}=\begin{cases}0 \text{ if student i at institution j has not changed major by year } t\\ 1 \text{ if student i at institution j has changed to a STEM major by year } t\\ 2 \text{ if student i at institution j has changed to a non}-\text{STEM major by year } t\end{cases}$$

  Note that since $CHMAJ_{ijt}$ is a categorical predictor that can take on 3 values, only the two dummy variables $(CHMAJ_{ijt})_1$ and $(CHMAJ_{ijt})_2$ are included in the model.

- β_p , $p=1,\ldots,P$, and δ_1 and δ_2 are the Level 1 parameters measuring the (partial) effects of the student-level predictors on logit hazard.

Institution-Level (Level 2) Model:

$$\alpha_{jk}=\gamma_{0k}+\sum_{q=1}^{Q}\gamma_{qk}W_{qj}+\varepsilon_{jk},\ k=1,\ldots,6\ , \quad (2)$$

where:

- W_{qj} is the *q*th institutional-level predictor for person *i* at institution *j*. The institution-level variables are provided in the Variable List supporting document.

- γ_{qk} , $q=1,\ldots,Q$ are the Level 2 parameters representing the (partial) effects of the institution-level predictors on logit hazard of graduating at time *k*.

- ε_{jk} is the random error associated with institution *j*, and it is assumed (for simplicity) that the vector of error terms $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots \varepsilon_6)^T$ follows a multivariate normal distribution with zero mean vector, and covariance matrix $\boldsymbol{\Sigma}$.

<u>Research Question 2</u>

To address Research Question 2, multilevel discrete-time hazard models with interaction terms between student-level variables and $CHMAJ_{ijt}$ and between institution-level variables and $CHMAJ_{ijt}$ are proposed. For example, to investigate whether the association between changing major and risk of graduation depends on STEM status of first declared major, the following model is proposed:

Student-Level (Level 1) Model:

$$\ln\left(\frac{h_{ijt}}{1-h_{ijt}}\right) = \sum_{k=1}^{6} \alpha_{jk} D_{ijk} + \sum_{p=1}^{P} \beta_p X_{pij} + \sum_{c=1}^{2} \delta_c (CHMAJ_{ijt})_c + \sum_{c=1}^{5} \kappa_c [(STEM04)(CHMAJ_{tij})]_c \ , \qquad (3)$$

where each κ_c measures how the effect of changing major on (logit) hazard of graduation varies by STEM status of first declared major (STEM04). The interaction between CHMAJ and STEM04 consists of the sum of the products between each $CHMAJ_{ijt}$ dummy variable and a dummy variable representing the levels of STEM04. The remaining terms in the model are defined as in (1).

Institution-Level (Level 2) Model:

$$\alpha_{jk} = \gamma_{0k} + \sum_{q=1}^{Q} \gamma_{qk} W_{qj} + \varepsilon_{jk}, \ k = 1, \ldots, 6, \qquad (4)$$

where all terms are defined as in (2).

Equations (1) and (3) will be compared using change in deviance tests (Singer and Willett, 2003) to determine whether the interactions are significant.

<u>Research Question 3</u>

To address Research Question 3, median times to graduate for students who persist in their majors, switch to a STEM major, and switch to a non-STEM major will be predicted based on the fitted hazard probabilities of the final model. Let $\hat{h}_t$ represent the estimated hazard probability of graduating at time *t* (for sake of brevity, the *i* and *j* subscripts have been dropped), and $\hat{S}_j = \prod_{t=1}^{j}\left(1-\hat{h}_t\right)$ be the estimated probability of not having graduated (surviving) beyond time *j*. Then the estimated median time to graduate is given by:

$$\text{Estimated Median} = m + \left[\frac{\hat{S}_m - .5}{\hat{S}_m - \hat{S}_{m+1}}\right], \qquad (5)$$

where m is the first time when the estimated survival function $\hat{S}_t$ is above .5 (see Singer and Willett (2003)).

Student-Level (Level 1)

- GENDER: gender of student (NPSAS)
- RACE: racial category of student (NPSAS)
- C04027: Father's highest grade of education (NPSAS)
- C04028 Mother's highest grade of education (NPSAS)
- PELL04: Amount of Pell grant funds the respondent received in 2003-04 (NPSAS)
- TUITION2: Amount of tuition and fees in 2003-04 (IPEDS)
- ACAINX04: measure of academic integration of the student at first institution in 2004 (NPSAS)
- SOCINX04: measure of social integration of the student at first institution in 2004 (NPSAS)
- HCGPAREP: cumulative high school grade point average (College Board)
- TESATMDE, TESATVDE: Derived SAT math and verbal scores (College Board)
- HCYSENGL, HCYSMATH, HCYSSCIE: Years of English, math, and science taken in high school (College Board)
- MAJ04A: Student's primary major in 2004 (NPSAS)
- MAJ06A, MAJ09A: Student's primary major in 2006 and 2009 (BPS)
- MAJ06CHG and MAJ09CHG: number of times student changed major from 2004-2006 and 2006-2009, respectively. These variables will be used in conjunction with MAJ04A, MAJ06A, and MAJ09A to create a derived time-varying predictor to indicate whether and when the student changes to a STEM or non-STEM major. (BPS)

Institution-Level (Level 2) Predictors

- UNITID: School IPEDS identification number needed to perform the hierarchical analysis (BPS)
- LOCALE: Measure of degree of urbanization of institution in 2003-2004 (IPEDS)
- ENRLSIZE: Enrollment size of institution in 2003-2004 (IPEDS)
- PCT_MIN: Percent minority enrollment of institution in 2003-2004 (IPEDS)
- FGRNT_P: Percent of students at institution receiving federal grants (IPEDS)
- SELECTV2: Selectivity level of first institution in 2003-2004 (IPEDS)
- CNTLAFFI: Institution control and affiliation in 2003-2004 (IPEDS)

The following list contains the names and brief descriptions of the primary derived variables (not in BPS:04/09) that will be used as student-level (Level 1) predictors in the multilevel discrete-time hazard model:

- STEM04: indicates whether first declared major in 2003-2004 academic year was in a STEM field. This will be determined by classifying major provided in MAJ04A using the most recent NSF STEM guidelines.

- CHANGE06: indicates whether the student is persisting in first declared major, has changed to a STEM major, or has changed to a non-STEM major as of 2006. MAJ06CHG and MAJ06A will be used to determine whether a change occurred, and whether the change was to a STEM or non-STEM major.

- CHANGE09: indicates whether student is persisting in first declared major, has changed to a STEM major, or has changed to a non-STEM major as of 2009. MAJ09CHG and MAJ09A will be used to determine whether a change occurred, and whether the change was to a STEM or non-STEM major.

- CHMAJ(tij): a time varying categorical predictor indicating whether the student has changed majors and takes the following values: 0 if student *i* at institution *j* has not changed major by time *t*, 1 if student *i* at institution *j* has changed to a STEM major by time *t*, and 2 if student *i* at institution *j* has changed to a STEM major by time *t*. Here, time *t* represents number of years after first enrollment.

The following list contains the names and brief descriptions of the BPS:04/09 variables that will be used to create the time to graduation and censoring variables for the multilevel hazard model:

- DATEBEG: Enrollment starting date of student (Derived variable)
- DATEEND: Enrollment ending date of student (Derived variable)
- DEG1: Indicates whether the bachelor's degree was earned (BPS)
- DEGDT1: Date the bachelor's degree was earned (BPS)