The When and Who of Graduation and Dropout Predictors: A Moderated Hazard Analysis

Sara R. Berzenski

Abstract
This study examined graduation and persistence among social and behavioral science students at a regional comprehensive university. Hazard analyses identified predictors of student trajectories, times at which predictors were more or less impactful, and interactions between predictors such that particular risk factors were more detrimental for certain groups. Results revealed that first-time freshmen and first-time transfer students operate under very different risk models. Total units enrolled and cumulative grade point average emerged as the most salient predictors across all models; however, numerous predictors varied in their salience across time and student subgroup. Differential predictors functioned as risks early versus late in student trajectories. Underrepresented minority status emerged as a risk factor in interaction with other predictors, such that it amplified other risks. These findings suggest that examinations of graduation and persistence must highlight the complex ways in which traditional risk factors interact with time and context to impact student success.

Keywords
graduation, persistence, retention, hazard analysis, timing

1Department of Psychology, California State University, Northridge, CA, USA

Corresponding Author:
Sara R. Berzenski, Department of Psychology, California State University, Northridge, CA 91330, USA. Email: sara.berzenski@csun.edu
Supporting and enhancing college student success is a critical goal for scholars and practitioners of higher education. This success, at the most basic level, is defined by the extent to which students persist toward and ultimately achieve graduation with a college degree. Indeed, a mountain of research has been dedicated to this pursuit over the past several decades. Although a set of common factors has been identified that contribute to student success, neither the one size fits all approach, nor the identification of individualized predictive factors, has yielded tremendous gains in student success across the board (Tinto, 2006). Thus, increasingly nuanced approaches to studying these phenomena are warranted to identify the specific ways in which predictors of student success operate and can be most effectively translated into practical gains. Specifically, the goals of this study were to examine commonly accessed student success predictive factors for their differential influence across time and for the ways in which the effects of some predictors may depend on others such that risk is magnified for particular subgroups. Thus, this study sets itself apart from the existing literature in its ability to identify when in a student’s college career, and for which type of student, each predictive factor is most salient. In so doing, it is poised to provide crucial information to practitioners regarding targets of first year versus later year student interventions, and which interventions may be best aimed toward vulnerable students.

A recent national report on student persistence (i.e., the extent to which students continue on in higher education from one year to the next, at any institution) revealed that the 2016–2017 first year persistence rate was 73.9%, indicating that more than a quarter of students did not return to college after their first year. Furthermore, Black students had the lowest persistence rate, at 67.0%, followed by Hispanic students at 70.7% (National Student Clearinghouse Research Center, 2018). Thus, it is of supreme importance to continue investigating pathways to supporting student persistence and graduation (i.e., eventual degree completion), and in particular, to focus on the ways in which traditional supports are more or less effective for underrepresented students.

Predictors of graduation and persistence can be identified according to demographic characteristics, institutional characteristics, and those that pertain to engagement (the transactional relationship between the student and the institution; Wolf-Wendel, Ward, & Kinzie, 2007). Chen (2012) suggest that theories on student persistence can likewise be organized along organizational, interactionalist, or structural-demographic domains. Effects of demographic factors abound in organized summaries of the graduation and persistence literature, with consistent effects identified of lower socioeconomic status (SES), male gender, racial/ethnic minority status, and lower parental educational attainment/first-generation student status (Chen, 2012; Ishitani, 2003; Kuh, Cruce, Shoup, Kinzie, & Gonyea, 2008; Kuh, Kinzie, Buckley, Bridges, & Hayek, 2006; Munro, 1981). However, it is likely these predictors may be
better explained by (Millea, Wills, Elder, & Molina, 2018; Pike & Kuh, 2005), or operate through their indirect effects on (Kuh et al., 2006; Munro, 1981; Sass, Castro-Villarreal, Wilkerson, Guerra, & Sullivan, 2018), predictors in other domains, such as engagement and connectedness. Sense of belonging, representing how students feel about their connections to the university and the extent to which they feel welcome, and that their community is included in the larger campus community, has been a robust predictor of graduation and persistence outcomes, along with similar precursor constructs such as engagement (Wolf-Wendel et al., 2007). Early academic performance is also a robust predictor of graduation and potential mediator (Gershenfeld, Ward Hood, & Zhan, 2016; Raju & Schumacker, 2015). However, key scholars note that given these established relations, it is critical to move toward considering whether and how established predictors and theories apply to underrepresented students and identifying potential effects on subgroups of students (Wolf-Wendel et al., 2007).

Evidence establishing specific conditions under which predictors of graduation and persistence operate is sparse, despite ample evidence that underrepresented students enter higher education with differential levels of risk. For example, first-generation students experience greater financial insecurity, lower levels of self-perceived academic competence, and increased difficulty engaging with peers and campus activities than continuing generation students (Pratt, Harwood, Cavazos, & Ditzfeld, 2019). Minority students often experience financial pressures, discrimination, exclusion, increased life stressors, and other structural barriers (Carter, 2006; Sólorzano, Villalpando, & Oseguera, 2005). However, it is critical to not only identify these factors but also extend this work to examine whether traditional risk factors actually predict graduation and persistence to different degrees, that is, are more or less risky for these students, given their differential experiences. When examining interactions between subgroups of students and predictors, some work has failed to find differences in the predictive capabilities of institutional characteristics (Chen, 2012) and noncognitive factors (e.g., academic mindset) for students of different ethnic groups (Farruggia, Han, Watson, Moss, & Bottoms, 2018). However, other work indicates that minority and nonminority students may experience predictors differently. Studies suggest that support from campus organizations and family members (Kuh et al., 2006) along with engagement (Kuh et al., 2008), may be particularly important for ethnic minority students compared with White students. More specifically, Xu and Webber (2018) find that psychosocial engagement may be more important for minority students, whereas academic engagement may be more important for White students. Furthermore, they find that goal commitment is important for Black students, but not White students, in predicting plans to drop out. Effects of grade point average (GPA) on persistence are inconsistent in terms of variation by ethnic group. Allen (1999) found that parent education predicted GPA and persistence only for nonminority students, and that GPA predicted persistence only for minority
students, whereas more recent work found that GPA is related to persistence only for White students, and not for Black students or other minorities (Farruggia et al., 2018; Xu & Webber, 2018). Taken together, this body of work suggests that varying levels of specific risk factors may mitigate typically expected effects of demographic factors and may be more or less salient for particular groups. Furthermore, although the limited and inconsistent state of this literature precludes specific expectations about interaction effects, these findings clearly highlight the necessity of further investigating traditionally expected effects for their differential relevance in particular subgroups of students.

Finally, it is critical to examine traditional predictors of retention and persistence with respect to their differential importance over time. Tinto’s (1988) theoretical model specifies that incorporation into college is the third stage in a multistep process of integration for new students. Thus, this model would suggest that some factors identified to predict retention and persistence may be differentially important earlier or later in a student’s college experience. Indeed, one study examining percent of variance explained by clusters of predictors in varying years found that socialpsychological variables (e.g., goals, faculty contact) did gain importance across time in college (John, 1985).

Given the limited resources universities have to deploy interventions targeting particular risk factors, knowledge of the times when each risk is most salient would serve a critical function. Yet, systematic longitudinal examinations of these predictors using event/hazard analysis remain extremely rare. Hazard analysis uses logistic regression to model the likelihood of an event or “hazard” (e.g., college student dropout) occurring, with time as a predictor. In this way, not only can the event’s likelihood be modeled over time, but interactions between time and other predictors can be tested to evaluate whether the effect of a given factor (e.g., gender) on dropout likelihood depends on year in school. Thus, the results from this type of analysis can qualify the effects of predictors that have previously been established for which it has not yet been determined whether effects are felt earlier, later, or do not vary over time.

Although scarce and inconsistent, previous time-sensitive analyses of various types have identified a few results of this nature. First, it has been demonstrated that the overall risk of dropout is higher in the first year than in other years (Chen, 2012). Furthermore, although one study found that risk of dropout for ethnic minority students is consistent across time (Chen, 2012), two others found that minority students were at greater risk earlier than later (Ishitani, 2003, 2016), and a third found that Black students were more likely to drop out only in Year 3, and found no differences for Hispanic students (DesJardins, Ahlburg, & McCall, 1999). In two studies, female students were found to be at greater risk of dropout later, compared with males (DesJardins et al., 1999; Ishitani, 2003), and another study found males more at risk earlier than later (Ishitani, 2016). First-generation student status and lower income
(Ishitani, 2003) were found to be greater risks in the first year in one study but more salient later in another (Ishitani, 2016). In addition, lower ACT scores were found to predict drop out only in the second year (DesJardins et al., 1999), and lower GPA was found to decline in importance as years increased (DesJardins et al., 1999; Ishitani, 2003). Finally, a recent study found that financial and academic pressures varied in their influence over time as well (Villano, Harrison, Lynch, & Chen, 2018). Of course, these studies varied in their recency and populations examined, so there are many possible reasons for these discrepancies. Thus, these extant studies suggest it is crucial to examine these predictors over time, but that their definitive importance at particular time points has not yet been established.

Tinto (2006) suggests that one crucial obstacle to the success of student retention research lies in the research-practice gap, and the extent to which interventions actually translate these identified factors into solutions and implement them. Thus, one potential strategy toward making tangible gains may be to focus on factors that are easily measured by institutions and therefore more easily targeted by interventions. Therefore, this study seeks to explore particularly accessible factors (e.g., first-generation student status, underrepresented minority status, and SES) and identify in depth characteristics of when and on whom they exert their effects so as to inform intervention efforts that target these accessible characteristics. In contrast to more nuanced concepts such as student engagement and sense of belonging, these factors are easily measured and available at even the most basic level of institutional data. However, at the same time, although these factors have demonstrated influence on student success, it is important to avoid concluding that these factors put all identified students at equal risk without adopting an individualized approach to examining them. In addition, these factors are not easily modifiable, and so the specific ways in which their identification translates into intervention centers around isolating the times and conditions under which particular groups are at greater risk and most vulnerable to a targeted intervention.

Therefore, the goals of this study were to (a) plot competing risks trajectories of graduation and dropout over a period of 10 years, (b) identify predictors of increased risk of graduation and dropout across time, (c) explore how the influence of these predictors might be differentially salient across time, and (d) explore potential interactions between these predictors.

**Method**

**Participants and Procedures**

Participants were students in the College of Social and Behavioral Sciences at a large west coast regional comprehensive university, who were either first-time freshmen \( (N=1,202) \) or first-time transfer students \( (N=1,700) \) who first
enrolled between Fall 2007 and Fall 2010, and were followed through Fall 2016. First-time freshmen and first-time transfer students were examined in separate models because evidence suggests that these students may face different sets of retention and graduation challenges and pathways (Kuh et al., 2006). First-time freshmen were 72% females, 50.4% Latinx, 19.0% African-American, 16.8% White, 6.5% Asian American, .3% Native American, .2% Pacific Islander, .5% International, and 6.2% Unknown. Furthermore, 41.6% were first-generation students and 51.9% received Pell Grants (i.e., federal tuition assistance). First-time transfers were 64.4% female, 35.3% Latinx, 5% African-American, 36.8% White, 5.7% Asian American, 3.0% Multiracial, .3% Native American, 1.7% International, and 12.1% Unknown. Moreover, 35.4% were first-generation students and 48.3% received Pell Grants.

Data were archival—gathered, compiled, and managed by the institutional research office at the university. Specifically, demographic and family history information were gathered from admissions information provided by the students on their university applications. Information provided by students regarding high school GPA and SAT scores were verified by official transcripts and ETS records by the admissions and records office. Remediation eligibility was also determined by the admissions and records office based on placement tests and high school records. Information on Pell Grant eligibility were provided by the financial aid office, and information on educational opportunity program (EOP) status was gathered by the EOP office. Data regarding college GPA and enrollment on a continuing basis was received from the registrar’s office each semester.

The enrollment information for each semester, combined with information on degree receipt, was used to create variables indicating each participant’s time to graduation and/or dropout from the first year in which they were enrolled and recoded so that the first data point was the participant’s first enrolled semester (rather than the first year of data collection). These data do not take into account gaps in enrollment but rather simply models time to eventual graduation or dropout (indicated by the final semester of enrollment). If participants were still enrolled within a year of the end of data collection (Fall 2016), they were considered still enrolled (not yet graduated or dropped out) for the purposes of this study. The predictors were as follows: gender (coded 0 = female, 1 = male), ethnicity (also recoded as underrepresented minority status [URM; coded 0 = White, Asian American, 1 = African-American, Native American, Latinx, Pacific Islander] in some analyses), Pell Grant status (0 = not received, 1 = received), Math remediation needed upon enrollment for first-time freshmen (0 = no, 1 = yes), English remediation needed upon enrollment for first-time freshmen (0 = no, 1 = yes), EOP status for first-time freshmen (0 = did not participate, 1 = participated), first-generation student status (0 = no, 1 = yes), high school GPA for first-time freshmen and transfer GPA for first-time transfers, SAT scores for first-time freshmen (highest math + highest verbal recorded), GPA per term, cumulative campus GPA, total units enrolled per term, and total units earned per term.
Analytic Plan

To address the study goals, hazard analyses were conducted to predict graduation and dropout. Hazard analysis plots conditional risk across time by examining the proportion of students who experience each “hazard” (i.e., graduation or dropout) at each time point among those who remain in the data set at that time point (i.e., have not previously graduated or dropped out, or aged out of the data set due to the end of data collection). These dichotomous outcomes then can be predicted by other variables across time points using logistic regression models, from which coefficients are interpreted as odds ratios (OR). Predictor values can be constant across time (e.g., gender, first-generation student status) or can vary at each time point (e.g., cumulative GPA, units enrolled), and differences in their effects can be examined as varying across time by evaluating interactions between predictors and time. In these models, time was modeled as discrete in terms of the baseline model (i.e., different hazard values at each time point), and as continuous for the purpose of examining interactions with predictors (such that predictor effects could be stronger earlier vs. later across continuous time, rather than display separate effects at each time point). The decision to model time interactions as linear was made to preserve power (i.e., including interaction terms with each discrete time point would involve including 8 extra variables for each possible interaction). Furthermore, discrete time was modeled across years, rather than semesters, (a) to preserve power by including 10 time intercepts rather than 19 and (b) because graduation in fall is by nature more rare than in spring, and thus hazard models that plot risk by semester would present a misleading trajectory.

Results

Hazard Trajectory Plots

The conditional probability of graduating each year (i.e., hazard rate), among students who remained enrolled, was calculated for first-time freshmen. For freshmen, the hazard rate peaked at year 5 (52.85% of enrolled students graduated that year), dropped but remained fairly constant across years 6 (43.81%) and 7 (42.31%), and then continued to decline, indicating that the probability of graduating was reduced each year the longer students remain enrolled past Year 5 (and particularly last Year 7). However, in examining the conditional probability of dropping out each year, among students who remained enrolled, there was a peak at Year 1 (24.92%), but then a relatively constant risk each year after that such that the dropout hazard was between 4.18% and 11.76% for each subsequent year (through year 8, after which no dropouts were recorded among the very small number of still-enrolled students). Thus, it is noteworthy that the graduation and dropout hazard plots are not simply mirror images of each
other, such that the reduced probability of graduating during later years reflects an only slightly increased probability of dropout, but also potentially a lengthening of graduation trajectories (see Figure 1).

The conditional probability of graduating each year, among students who remained enrolled, was calculated for first-time transfers. For transfers, the hazard rate peaked at Year 3 (68.64%) but was fairly consistently high across Years 2 (51.48%) through 5 (56.67%). There was a decline in Year 6 (39.13%) and then a sharp drop off in Year 7 (7.69%); however, only a handful of students remained enrolled at this point ($N = 13$) so that rate should be interpreted with caution. In examining the conditional probability of dropping out each year, among students who remained enrolled, the rate peaked at Year 1 (9.76%) and then remained fairly constant (between 5.00% and 7.69%) between Years 2 through 7, with the exception of 0% drop out in Year 6. Again, the sample size of students still enrolled in Years 6 and 7 is likely too small to fully interpret these rates. Overall, it appears the graduation and dropout hazards may be more constant across time for first-time transfers than they are for first-time freshmen, although some features were consistent across the groups such as a peak for dropout in the first year (see Figure 2).

**Figure 1.** Competing risks (graduation/dropout) hazard plot for first-time freshmen.

**Individual Predictor Effects**

Next, each potential predictor was examined for its effects on graduation hazard rate and dropout hazard rate both for first-time freshmen and first-time transfers. Separate models were analyzed for each competing risk, as it is necessary
that separate dependent variables be predicted in separate logistic regression models. Competing risks hazard analysis relies on an assumption of independence of the two outcomes; however, in practice this is frequently an untenable assumption at the level of the baseline models. Therefore, the assumption is maintained by including the same predictors in each model, and theoretically assuring that the resulting graduation and dropout rates are independent after controlling for all predictors. First, each predictor was analyzed in an independent regression model. This was done to preserve power in the models and to examine the contribution of each predictor at face value. Although examining the predictors as competing simultaneously in models has obvious utility for prioritizing intervention efforts (and was examined in the second section of these results), given the exploratory perspective with which several of the variables were examined, information about each separate predictor’s salience (or lack thereof) was also uniquely valuable. Also, it is important to note that the effects of each predictor on the hazard are held constant across time, although it may not appear this way in the figures because of the proportional nature of the metric. In this section, main effects of predictors as well as interactions with time and interactions with other predictors will be discussed.

First-time freshmen. Ethnicity significantly predicted first-time freshmen graduation and dropout hazard rates. Specifically, White students were 2.23 times more likely than African American students \((p < .001)\) and 1.88 times more likely than Latinx students \((p < .001)\) to graduate at any given point. Conversely, African American students were 2.07 times more likely \((p < .001)\), and Latinx...
students 1.44 more likely ($p < .001$) than White students to drop out at any given point.

There were no effects of gender on first-time freshmen graduation or drop-out rates.

Students who did not receive Pell Grants were 1.33 times more likely to graduate at any given point than Pell Grant recipients ($p = .007$; see Figure 3), who were 1.25 times more likely to drop out at any given point ($p = .019$).

Students not needing Math remediation were 1.52 times more likely to graduate at any given point than students needing Math remediation ($p < .001$); however, this effect was qualified by an interaction with time, such that it was more salient in earlier years than later years ($p = .009$, $OR = 1.28$). Furthermore, students who needed Math remediation were 1.42 times more likely to drop out at any given point ($p < .001$), and this effect did not vary across time.

Students not needing English remediation were 1.66 times more likely to graduate at any point than students needing English remediation ($p < .001$);
however, this effect was qualified by an interaction with time, such that it was more salient in earlier years than later years ($p = .014$, $OR = 1.26$). Furthermore, students who needed English remediation were 1.45 times more likely to drop out at any given point ($p < .001$), and this effect did not vary across time.

Students not involved in EOP were 1.33 times more likely to graduate at any given point than those who were not involved ($p = .034$). Although there was no main effect of EOP on likelihood of dropping out, there was an interaction with time ($p = .002$, $OR = 1.216$) such that involvement with EOP increased the risk of dropout later more so than earlier.

Continuing generation students were 1.38 times more likely to graduate at any given point than first-generation students ($p = .005$), who were 1.42 times more likely to drop out ($p < .001$).

Higher high school GPA predicted higher likelihood of graduation at any given point ($p < .001$, $OR = 2.17$, indicating that a one full point increase in GPA predicted a 2.17 times greater likelihood of graduating) and lower likelihood of dropout ($p < .001$, $OR$ for lower GPA = 2.35).

Higher SAT scores predicted higher likelihood of graduation at any given point ($p = .001$, $OR = 1.002$, indicating that a 1 point increase in SAT predicted a 1.002 times greater likelihood of graduating). This effect was qualified by an interaction with time, such that it was more salient earlier than later ($p = .041$, $OR = .999$). Furthermore, higher SAT score predicted lower likelihood of dropout ($p = .002$, $OR$ for lower SAT = 1.001), and this effect did not vary over time.

Higher GPA within a given year ($p < .001$, $OR = 3.63$) and higher cumulative campus GPA ($p < .001$, $OR = 7.45$) predicted increased likelihood of graduation in that year, and decreased likelihood of dropout (lower term GPA $p < .001$, $OR = 4.90$, indicating a 4.90 times greater dropout risk compared with a GPA one full point higher; lower cumulative GPA $p < .001$, $OR = 7.04$, indicating a 7.04 times greater dropout risk compared with a cumulative GPA one point higher). The effect of cumulative campus GPA on dropout was qualified by an interaction with time, such that it was more salient later than earlier ($p < .001$, $OR = .62$).

Higher total units enrolled within a given year ($p < .001$, $OR = 1.28$) and higher total units earned within a given year ($p < .001$, $OR = 1.37$) predicted increased likelihood of graduation in that year and decreased likelihood of dropout (fewer units enrolled $p < .001$, $OR = 1.22$; fewer units earned $p < .001$, $OR = 1.42$). These effects were qualified by interactions with time, such that they were more salient earlier than later ($p_{\text{graduation}} < .001$, $ORs = 1.10$ and 1.08, respectively, $p_{\text{dropout}} < .001$, $ORs = 1.05$ and 1.04 respectively).

In addition, a number of interaction effects were observed. Underrepresented minority status significantly interacted with first-generation student status ($p = .038$, $OR = .41$) in predicting graduation, and with English remediation status ($p < .001$, $OR = 2.53$) and SAT score ($p = .011$, $OR = .998$) in predicting dropout, such being an underrepresented minority student was only a risk factor
among students already at risk via these other factors. For example, for first-generation students, underrepresented minority status posed a much greater risk to graduation than for nonfirst-generation students (see Figure 4).

Similarly, underrepresented minority status posed a risk for dropout only among students in need of English remediation and not among those who did not need English remediation (see Figure 5).

Gender significantly interacted with total units enrolled ($p = .018, OR = 1.10$) in predicting dropout, such that this risk was more salient for females. Pell Grant status significantly interacted with total units earned ($p = .015, OR = 1.11$) in predicting graduation, and with English remediation ($p = .036, OR = 1.57$), SAT score ($p = .027, OR = .998$), term GPA ($p = .015, OR = .73$), and cumulative campus GPA ($p = .014, OR = .66$) in predicting dropout, such that these risks were more salient for Pell Grant recipients. Pell Grant status interacted with Math remediation ($p = .024, OR = 1.66$), such that non-Pell Grant recipients only had increased graduation rates among those who did
not need Math remediation. Finally, first-generation student status significantly interacted with cumulative campus GPA in predicting graduation ($p = .019$, $OR = 1.90$) such that this risk was more salient for first-generation students.

**First-time transfers.** Ethnicity significantly predicted first-time transfer dropout rates ($p = .005$). Specifically, African American students were 2.05 times more likely ($p = .006$), and Multiracial students were 2.22 times more likely ($p = .012$) than White transfer students to drop out at any given point. There was no significant effect of ethnicity on graduation hazard rates for transfer students.

Female transfer students were 1.34 times more likely to graduate at any given point than male transfer students ($p = .001$), who were 1.48 times more likely than female students to drop out ($p = .001$).

There were no significant effects of Pell Grant status or first-generation student status on graduation or dropout rates for transfer students.
Higher transfer student GPA predicted higher likelihood of graduation at any given point \((p < .001, \text{OR} = 1.48)\), indicating that a 1 full point increase in GPA predicted a 1.48 times greater likelihood of graduating and lower likelihood of dropout \((p < .001, \text{OR for lower GPA} = 2.43)\).

Higher GPA within a given year \((p < .001, \text{OR} = 2.68)\) and higher cumulative campus GPA \((p < .001, \text{OR} = 3.30)\) predicted increased likelihood of first-time transfer graduation in that year and decreased likelihood of dropout \((\text{lower term GPA} p < .001, \text{OR} = 4.17; \text{lower cumulative GPA} p < .001, \text{OR} = 4.48)\). The effect of cumulative campus GPA on dropout was qualified by an interaction with time, such that it was more salient later than earlier \((p < .001, \text{OR} = .61)\).

Higher total units enrolled within a given year \((p < .001, \text{OR} = 1.33)\) and higher total units earned within a given year \((p < .001, \text{OR} = 1.41)\) predicted increased likelihood of graduation in that year, and decreased likelihood of dropout \((\text{fewer units enrolled} p < .001, \text{OR} = 1.22; \text{fewer units earned} p < .001, \text{OR} = 1.45)\). The effects on graduation were qualified by interactions with time, such that they were more salient earlier than later \((ps < .001, \text{ORs} = 1.16\text{ and }1.15\text{ respectively})\).

In addition, two interaction effects were observed. Pell Grant status significantly interacted with term GPA \((p = .011, \text{OR} = .69)\) in predicting dropout, such that this risk was more salient for non-Pell Grant recipients. Furthermore, first-generation student status significantly interacted with total units earned \((p = .044, \text{OR} = 1.07)\) in predicting transfer student graduation, such that low units earned was a more salient risk factor for first-generation students.

**Final Models**

Finally, multiple predictor logistic regression models were analyzed for graduation and dropout rates for first-time freshmen and first-time transfers, using the significant predictors identified earlier. As noted earlier, models for graduation and dropout were required to contain the same predictors as each other (but predictors varied between first-time freshmen and first-time transfer models). Only one of the campus GPA variables (cumulative instead of term) and units variables (total enrolled instead of total earned) was included in each model, due to the very high correlations between these variables.

In the final model predicting first-time freshmen graduation, several predictors emerged as significant when simultaneously competing with other potential predictors. More total enrolled units predicted higher likelihood of graduation \((p < .001)\) and Pell Grant recipient status predicted lower likelihood of graduation \((p = .026)\). An interaction between time and units enrolled was significant \((p = .001)\), such that more units enrolled predicted greater likelihood of graduation, but only earlier not later. Finally, Pell Grant status significantly interacted
with Math remediation ($p = .018$) such that this risk factor was particularly salient for non-Pell Grant recipients.

In the final model predicting first-time freshmen dropout, several predictors emerged as significant when simultaneously competing with other potential predictors. Pell Grant recipient status ($p = .003$) predicted higher likelihood of dropout, whereas more total units enrolled ($p = .025$), higher campus GPA ($p < .001$), and needing English remediation ($p = .041$) predicted lower likelihood of dropout. There was a significant interaction effect between time and cumulative campus GPA ($p < .001$) such that higher campus GPA predicted lower likelihood of dropout later but not earlier. Finally, there were significant interactions between underrepresented minority status and English remediation ($p = .006$), such that this risk was more significant for underrepresented minorities, and between Pell Grant status and SAT scores ($p = .003$), such that SAT scores predicted lower likelihood of dropout more strongly for non-Pell Grant recipients.

In the final model predicting first-time transfer graduation, several predictors emerged as significant when simultaneously competing with other potential predictors. More total enrolled units ($p < .001$) and higher cumulative campus GPA ($p < .001$) predicted higher likelihood of graduation, whereas higher transfer GPA ($p = .005$) predicted lower likelihood of graduation. It is important to keep in mind that the value of these predictors controls for all other predictors, so we interpret this result to say that higher transfer GPA is a risk factor for lower graduation, at constant levels of the other factors (e.g., including cumulative campus GPA). There was a significant interaction between time and total units enrolled ($p < .001$) such that total units enrolled predicted higher likelihood of graduation later but not earlier.

In the final model predicting first-time transfer dropout, several predictors emerged as significant when simultaneously competing with other potential predictors. Male gender ($p = .010$) predicted higher likelihood of dropout, whereas more total enrolled units ($p < .001$) and higher cumulative campus GPA ($p < .001$) predicted lower likelihood of dropout. Finally, there was a significant interaction between time and cumulative campus GPA ($p < .001$) such that lower cumulative campus GPA predicted increased risk of dropout later but not earlier. See Table 1 for a summary of significant individual and final predictor main effects in all models. Interaction effects are not included in the table, for clarity.

**Discussion**

**Summary and Interpretations**

This study examined several prominent predictors of graduation and dropout rates among first-time freshmen and first-time transfer students in the college of
Table 1. Main Effects of Individual and Final Model Predictors.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>First-time freshmen</th>
<th>First-time transfers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Graduation</td>
<td>Dropout</td>
</tr>
<tr>
<td></td>
<td>Time variation OR p</td>
<td>Time variation OR p</td>
</tr>
<tr>
<td>White vs. African American</td>
<td>2.23 .001</td>
<td>2.07 .001</td>
</tr>
<tr>
<td>White vs. Latinx</td>
<td>1.88 &lt;.001</td>
<td>1.44 &lt;.001</td>
</tr>
<tr>
<td>White vs. Multiracial</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Gender</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Pell Grant Status</td>
<td>1.33 .007</td>
<td>1.25 .019</td>
</tr>
<tr>
<td>Math remediation needed</td>
<td>1.52 &lt;.001 Everyone</td>
<td>1.42 &lt;.001</td>
</tr>
<tr>
<td>English remediation needed</td>
<td>1.66 &lt;.001 Everyone</td>
<td>1.45 &lt;.001</td>
</tr>
<tr>
<td>EOP status</td>
<td>1.33 .034</td>
<td>ns</td>
</tr>
<tr>
<td>First-generation student status</td>
<td>1.38 .005</td>
<td>1.42 &lt;.001</td>
</tr>
<tr>
<td>High school GPA</td>
<td>2.17 &lt;.001</td>
<td>2.35 &lt;.001</td>
</tr>
<tr>
<td>Transfer GPA</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>SAT scores</td>
<td>1.00 .001</td>
<td>1.00 .002</td>
</tr>
<tr>
<td>GPA: term&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.63 &lt;.001</td>
<td>4.90 &lt;.001</td>
</tr>
<tr>
<td>GPA: cumulative&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7.45 &lt;.001</td>
<td>7.04 &lt;.001</td>
</tr>
<tr>
<td>Units enrolled</td>
<td>1.28 &lt;.001</td>
<td>1.22 &lt;.001</td>
</tr>
<tr>
<td>Units earned&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.37 &lt;.001</td>
<td>1.42 &lt;.001</td>
</tr>
</tbody>
</table>

All odds ratios (OR) are presented in the positive direction for ease of interpretation. However, as noted in text, these effects represent oppositely coded predictors, for example, a positive effect of Pell Grant Status on graduation indicates individuals without Pell Grants were more likely to graduate, whereas a positive effect of Pell Grant Status on dropout indicates individuals with Pell Grants were more likely to drop out. Refer to the text to confirm the direction of each effect as necessary. OR should not be directly compared between variables with different units (e.g., SAT vs. GPA). Bold text denotes predictors that remained significant in the final combined models (significant interaction terms from the final models are not included in this table). ns = nonsignificant.

<sup>a</sup>GPA: Term and Units Earned were not included in the final models due to overlap with the other GPA and unit predictors.

<sup>b</sup>GPA: Cumulative only interacted with time in predicting transfer dropout in the final model.

<sup>c</sup>Units enrolled predicted graduation for first-time transfers earlier in the individual models but later in the final model.
social and behavioral sciences at a west coast regional comprehensive university. Critically, our findings extend the existing literature by clarifying which factors are salient only for particular subgroups of students and which are more predictive of success earlier versus later. By the nature of examining four separate models, along with a number of individual predictors and potential interaction effects, individual findings were numerous. Indeed, separating graduation from dropout and separating freshmen from transfer students in these models represent key innovations of this study. However, a few important take-home messages will be offered to help synthesize this collection of findings.

First, it is clear that first-time freshmen and first-time transfer students operate under very different models of graduation and persistence. This is apparent from the shape of the underlying baseline hazard profiles and also in terms of specific salient predictors. Transfer students had much less variability in conditional probabilities of graduation and dropout in any given year than did first-time freshmen. In terms of predictors, male gender was a significant risk factor for dropout and lack of graduation among transfer students, but not first-time freshmen. On the other hand, first-generation student status was a significant risk factor for first-time freshmen but not for transfer students. This may be because transfer students have already taken some time to figure out how to navigate the college experience before arriving at the university and as such are not as vulnerable to the challenges this presents to first-generation student freshmen. Pell Grant status was also a risk factor for first-time freshmen but not for transfer students. Second, graduation and dropout are complementary but not completely overlapping converse outcomes. These results suggest that certain factors may delay graduation without necessarily predicting increased dropout rates. For example, ethnicity was a significant predictor of transfer student dropout rates but not graduation, whereas EOP involvement was a predictor of first-time freshmen graduation rates but not dropout. In addition, some of the time varying effects suggested that particular factors (e.g., remediation need) may delay graduation early on but not have differential effects with regard to the timing of dropout.

Third, in terms of the differential importance of specific predictors, as expected based on the existing literature, each separate predictor examined had unique effects on graduation and persistence outcomes for at least one group or the other. However, some specific predictors emerged as particularly strong when competing in the simultaneous predictor models. Specifically, total units enrolled and cumulative campus GPA were arguably the most salient predictors across all the models. These predictors are notable for their broad applicability to both transfer students and first-time freshmen, to predicting graduation and dropout rates, and for surfacing in the large models as the most robust predictors when competing with other factors. These factors were more salient across models than underrepresented minority status, first-generation student status, remediation needs, and high school/transfer GPA.
However, it is extremely important to remember that these variables may be proxies for a number of other measured and unmeasured factors. For example, given the effects of first-generation student status at the single predictor level, it may be that the effect of total enrolled units in the larger models captures with it some of the first-generation student variance. Furthermore, Pell Grant status likely is a higher level construct that could be unpacked into effects of things like commuting distance, hours worked off campus, child care needs, and so forth. Thus, although these predictors may have accounted for these overlapping factors in the parsimonious models, if we intervene at the broad level and ignore the underlying mechanisms we are at a great danger of misallocating efforts and resources. The importance of this point and careful interpretation of the implications of these findings cannot be overstated. That said, these findings do reveal robust potential targets for further study and targeted attempts to understand and intervene on the underlying mechanisms of these particularly salient effects.

Another important take-home message from this set of findings is that the factors that emerged as most salient are potentially modifiable. This is encouraging, compared with an alternative situation in which, for example, underrepresented minority status would have predicted graduation and persistence above and beyond cumulative GPA. The fact that it did not suggests that although underrepresented minorities are at risk for lower GPAs (given the established opportunity gap), being an underrepresented minority does not predict graduation and persistence obstacles above and beyond these potentially modifiable targets.

Fourth, related to the earlier point, being an underrepresented minority student was only a risk factor for delayed graduation and increased dropout among those already at risk due to other factors. However, underrepresented minority students without additional risk factors (e.g., who did not need remediation, were not first-generation students) were not particularly at risk. That said, it is worth noting that underrepresented minority students may be especially likely to experience these other risk factors, so they still may be at increased indirect risk for that reason.

Finally, specific predictors emerged as varying in their effects over time. For example, remediation necessity, SAT scores, and total units enrolled predicted both lower likelihood of graduation and higher rates of dropout; however, this was particularly the case for graduation rates in early years. Conversely, cumulative campus GPA was a more salient predictor of dropout in later years.

**Implications for Practice**

This set of results suggests numerous exciting intervention recommendations. First, two general and very clear recommendations for practitioners stand out above and beyond any specific risk factor targets. Specifically, the study revealed very different models for student success for first-time freshmen versus transfer
students and for graduation versus dropout. Therefore, student success interventions should differentially target the needs of first-time freshmen and first-time transfer students, whether it be in line with these specific predictors or other salient factors not assessed in this study. The results also suggest that first-time freshmen may be more mutable targets for intervention. In addition, the differential importance of persistence versus timely graduation as targets of a given student success intervention may dictate attention to particular factors over others. For example, for a campus most concerned with dropout, if receiving remediation has an indirect effect on cumulative campus GPA, then its negative impact in delaying graduation timing may be a small consequence for reducing overall dropout rates. To be clear, this study did not assess the effects of receiving remediation, simply whether incoming students needed remediation, so testing this specific hypothetical comparison is beyond the scope of the study, but worth examining as a follow-up. Prioritizing dropout avoidance versus time to graduation may be a tall order to consider when designing interventions, as often shorter time to graduation and reduced dropout are seen as complementary goals. However, from these findings we can infer that some interventions may indirectly extend time to graduation while reducing dropout, so it may be important to disaggregate these goals when possible, and design more narrowly focused interventions.

In terms of recommendations based on specific emergent predictors from this research, total units enrolled and cumulative GPA were the most robust factors identified. Thus, one implication of this work may be to pay particular attention to these factors in intervention efforts (this could be through increased attention to students’ financial obstacles as well as tutoring and study skills training designed to improve grades and allow students to successfully complete more units). However, as noted earlier, these variables likely stand in for other more specific factors that actually drive their effects, and these underlying mechanisms are likely even more important targets for intervention. In other words, if we encourage students to enroll in more units without addressing the underlying factors that may define students who are able to enroll in more units comfortably already (and thus are at lower risk for dropout and delayed graduation), our intervention will likely be unsuccessful. Therefore, more work is needed to uncover these mechanisms before interventions should be deployed on these targets.

In addition, it is encouraging that there were very few direct effects of dispositional (i.e., unmodifiable) factors such as gender, underrepresented minority status, and first-generation student status. However, the finding that these factors only become salient in the context of other risks means these still may be important markers of groups of students more at risk, to whom we may benefit from targeting interventions. For example, among those needing English remediation, underrepresented minority students were particularly at risk; thus, the mechanisms by which this risk factor operates may be tied into this status, and
these students may be particularly vulnerable to a targeted intervention on this factor.

Finally, and critically, the novel feature of this study in examining these factors for differential effects over time suggests specific recommendations for practitioners in terms of when particular interventions are deployed. For example, interventions focused on increasing enrolled units should likely be targeted at incoming students but not pushed on continuing students, whereas interventions targeting cumulative GPA may be more effective later.

**Limitations and Considerations for Interpretation**

It is especially important to consider these findings in light of several significant limitations. First, this study examined only students in the college of social and behavioral sciences. It is by no means a foregone conclusion that these same predictors would be differentially salient or exert influence at the same times across other colleges (e.g., college of science and math). Furthermore, the study can only be generalized to students from this university or those from very similar types of regional comprehensive universities with a comparable student body. It is imperative that the study be replicated in other colleges and at other universities.

Critically, not enough information was available to discern among students’ specific ethnic backgrounds. For example, the Asian American student category included both Southeast Asian Americans and East Asian Americans. Although Asian American students are not typically considered underrepresented minorities, these subgroups of students are distinct, particularly given the most recent trend in U.S. immigration policies which hyperselect for highly educated East Asian immigrants (Zhou & Lee, 2017). Furthermore, recent evidence has emerged regarding an opportunity gap between White students and Southeast Asian students. This gap stems from specific barriers Southeast Asian students face such as comparatively high rates of poverty and low rates of parent engagement and English language proficiency (Southeast Asia Resource Action Center, 2013). Thus, it is important to disaggregate this group so that more meaningful conclusions can be drawn regarding underrepresented minority status.

Furthermore, it was not possible to draw specific conclusions about multiracial students. For example, national trends in academic performance and employment among multiracial individuals suggest that for this group, the picture is complex (Musu-Gillette et al., 2016). Multiracial students tend to complete high school and college at comparable rates to White students, but they also are more likely to be from poor families, less likely to be employed after college, and lag behind White counterparts in median yearly income. Research on social identity and social perceptions suggest that multiracial individuals’ identities, experiences, and outcomes may depend on phenotype, such that
a White-passing biracial person may be treated differently by teachers, administrators, and peers than one who is not passing. Furthermore, regardless of phenotype, perceivers tend to label multiracial individuals as belonging to one racial/ethnic group, most often assigning the lower status parent racial group to multiracial individuals (e.g., a Black-White biracial person like Barack Obama is often thought to be Black, rather than biracial or White; Ho, Sidanius, Levin, & Banaji, 2011). Therefore, it is imperative that research on graduation and persistence move toward targeted investigations of the complex ways risk and protective factors function for specific groups of students, with particular attention paid to groups such as Southeast Asian Americans, Pacific Islanders, and multiracial students, whose profiles are complex and are often neglected in research.

Second, these data are subject to significant selection effects, given that students were not (obviously) randomly assigned to particular numbers of units to enroll in or particular GPAs, for example. Thus, it is impossible to determine the underlying causal factors behind these effects. For example, as noted earlier, the finding that total units enrolled robustly related to graduation and persistence rates does not signify that enrolling in more units will cause graduation to increase and dropout to decrease. On the contrary, there is a large selection effect at play in which certain factors dictate which students are enrolling in more units than others naturally, and it is likely this set of factors that is the underlying causal mechanism. This means that it would be inappropriate to intervene directly on the factors identified here without understanding these mechanisms and the potential selection effect inherent in the data. However, these findings still represent an important step forward, in that they point us in the direction of this underlying mechanism. Total units enrolled may be an important factor signifying a latent group of individuals who are uniquely suited to persist and to graduate in a timely fashion. The present findings reveal that the next important step may be to examine these individuals and determine what that underlying mechanism is, rather than to intervene directly on units enrolled.

In addition, although the sample size is fairly large, when examined over time, the sample by nature gets progressively smaller. Thus, the study lacks power to examine meaningful effects at later time points and to some degree may lack power to examine the number of predictors evaluated here overall, once the sample size is reduced. Furthermore, given differentially missing data across predictors, it is difficult to explicitly compare the single predictor models to each other and to the larger competing model since they likely contain slightly different sets of participants. In addition, potential power issues precluded doing a more detailed evaluation of time-varying effects, such that specific differences at each time point could be examined (rather than looking at a single linear interaction). Future targeted analyses on larger samples could seek to evaluate this question for theoretically specified periods of time.
Relatedly, this study did not examine less objective predictors, such as student engagement, in this study. As noted earlier, the goal of the study was to focus on the most accessible data, such that intervention efforts could be targeted to widely available targets across universities. However, it is likely that these types of predictors also play an important role in graduation and persistence, and at this time, we are unable to make comparative statements regarding their differential importance.

In addition, this study does not address the fact that there were gaps in enrollment for many students. In terms of predicting graduation timing, there may be differences between students who were enrolled continuously and those who were not, for which future analyses should attempt to account. This fact also makes it difficult to accurately predict dropout date, as some students return after long absences.

Next, the exploratory nature of the study and the focus on internal assessment for this project dictated that an extremely large number of tests be conducted. Thus, the study suffers from a potentially inflated Type I error rate. We must therefore be especially prudent in our conclusions, particularly about smaller effects, as some of the effects observed here may be emerging spuriously. Replication of this study in independent samples will be critical in evaluating which of these predictors and effects are robust to this potential limitation.

Furthermore, the models examined here rely on the competing risks model assumption described earlier, that graduation and dropout can be examined in separate models if we assume they are independent outcomes. This assumption can be met if sufficient predictors are included in the models; however, we cannot truly evaluate whether we have met this assumption; thus, these results once again need to be interpreted with caution.

Finally, though not a limitation of this study, it is important to remember the context in which these results should be interpreted. The probabilities described here represent conditional probabilities, risks given the set of students continuing to be enrolled at each time point. Thus, the conclusions reached here regard risk at each particular given point in time and are different in nature than conclusions about overall graduation and dropout rates among all students (beginning from enrollment). In addition, it is important to remember that the OR reported here are contingent on units of measurement, and thus should not be directly compared with each other in size (e.g., there is a large difference between a one unit increase in GPA and a one unit increase in SAT scores).

**Conclusion**

Despite these limitations, this study extends the current literature by revealing very important findings about the nature of differential risk profiles for first-time freshmen and transfer students and identifying several robust predictors that are
consistent over time and several that vary over time. These findings have clear implications for both the timing and targets of interventions pending future examination of the underlying mechanisms of the effects brought into relief by this study.

**Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was supported by resources and funding provided by California State University, Northridge’s 2016-17 Data Champions program, under the direction of Dr. Janet Oh, Senior Director of Institutional Research, and Dr. Kristy Michaud, Director of Student Success Innovations and Professor of Political Science. Opinions in this article are solely those of the author and do not necessarily represent those of California State University, Northridge, or its Data Champions program.

**ORCID iD**

Sara R. Berzenski  
https://orcid.org/0000-0003-2435-0411

**Note**

1. Multiracial students were excluded from the analyses presented here due to the complex nature of their categorization. However, all study analyses were also conducted including multiracial students as underrepresented minorities, and all findings remained consistent with those presented here.

**References**


Retention: Research, Theory & Practice, 17(4), 469–488. doi:10.1177/1521025115579251


**Author Biography**

**Sara R. Berzenski**, PhD, is an associate professor of quantitative and developmental psychology at California State University, Northridge. Her major research interests center on examining emotional development in contexts of adversity, measurement, and using quantitative methods to highlight individual differences across a broad range of applications.