Leveraging Data-Driven Decisions:
A Predictive Model to Enhance Student Success in Higher Education

Ming Chen
California State University Long Beach

Jasmine Yur-Austin
California State University Long Beach

Higher education institutions often face challenges in achieving key success metrics, such as on-time graduation rates, particularly when supporting students with diverse backgrounds, needs, and barriers. This study examines a detailed dataset containing student demographics and academic records from College of Business at a large public university spanning ten years. Our results highlight significant achievement gaps based on factors such as gender, age, race, family income, and parents' education levels. These gaps persist even among students with similar academic histories. We develop prediction models that enable institutions to identify high-risk students early, facilitating timely intervention and providing a framework for comparing institutional efforts to improve student success. This study highlights how data-driven approaches can enhance institutional management by enabling proactive identification of challenges, optimizing resource allocation, and supporting strategies that improve student outcomes.

Keywords: higher education management, student success, data-driven, prediction

INTRODUCTION

Student success, particularly measured by on-time graduation, remains a persistent challenge for many higher education institutions. Despite ongoing efforts to improve retention and completion rates, significant disparities persist across various student populations. These disparities often manifest as higher failure rates, lower retention, and delayed graduation, even among students who enter college with comparable academic preparation. Higher education institutions are increasingly aware that addressing such challenges requires a strategic management approach that integrates policy reform, targeted support services, and the deployment of data-driven decision-making tools. These tools enable early identification of at-risk students, thereby enhancing resource allocation, optimizing student retention and graduation and supporting evidence-based organizational practices aimed at improving academic outcomes.

Our study is grounded in the context of the California State University (CSU) system, the largest public university system in the United States. Comprising 23 campuses and serving more than 480,000 students, CSU is home to one of the most diverse student populations in the country. Nearly half of CSU students are from underrepresented minority groups, approximately one-third are first-generation college students, and 54% receive Pell Grants, indicating low-income backgrounds. These characteristics make CSU an ideal setting for examining student outcomes across a wide range of demographic and socioeconomic

backgrounds. Within this system, our research is situated at the College of Business (COB), which is committed to offering accessible, affordable, and high-quality business education tailored to its ethnically, economically, and academically diverse student body. In pursuit of its mission, COB actively works to dismantle academic barriers and implement systemic strategies to engage students from disadvantaged backgrounds. COB aims to prepare all students with the knowledge and skills necessary to thrive in their professional careers, regardless of their background. As such, the development of a predictive model that enables early identification of students at risk of delayed graduation aligns directly with the institution's goals and commitment to student success.

While existing literature has extensively documented the predictive power of factors like high school GPA, standardized test scores, and socioeconomic status on college persistence and graduation outcomes (Tinto, 1993; Galla et al., 2019; Tucker and McKnight, 2019; Barbera et al., 2020), fewer studies have offered institution-ready models that operationalize these findings into real-time interventions. Moreover, our focus is pragmatic and outcome-driven: to demonstrate how a robust, data-informed model can help any institution improve graduation rates by providing targeted and timely support to all students, regardless of their background. To evaluate academic achievement, we adopt two key metrics: (1) graduation within standard time windows, i.e., 4-year and 6-year benchmarks for first-time freshmen, and 2-year and 4-year benchmarks for transfer students; and (2) cumulative GPA at graduation. These indicators not only reflect student performance but also institutional effectiveness in guiding students to successful completion. Using a robust dataset spanning over ten years of academic records and demographic data, we investigate how variables available at the point of entry, such as gender, age, standardized test scores (SAT or ACT), ethnicity, Pell Grant eligibility, parental education level, and transfer status, can predict student academic achievement. We build on prior efforts that emphasize the importance of early warning systems (Arnold & Pistilli, 2012; Jayaprakash et al., 2014), but extend these approaches by validating our predictive model across a large and diverse dataset covering multiple academic cohorts. Our analysis underscores that achievement gaps—often assumed to emerge only later in college—are, in fact, predictable based on information available at matriculation. More importantly, our model can be adapted and scaled across institutions, providing a flexible and evidence-based framework to guide resource allocation, advising strategies, and support services.

The rest of the paper is structured as follows. In Section 2, we review relevant literature. In Section 3, we analyse data and measure academic achievement gaps across various dimensions. In Section 4, we develop several predictive models and demonstrate how these models can be used to identify high-risk students based on both demographic information and prior academic performance. Additionally, we illustrate how our prediction models can provide a common framework for comparing different institutions with varying demographic compositions and admission criteria concerning student success. We conclude our paper in Section 5.

LITERATURE REVIEW

The Obama and Trump administrations both implemented various initiatives to enhance the College Scorecard, including updates to graduation rates, college costs, and student loan information, to help prospective students make informed decisions about their college choices. In support of federal initiatives to increase degree attainment, 33 states have implemented revised funding mechanisms for public colleges to boost graduation rates (Dougherty et al. 2016). Our study, based on over a decade students' data since the introduction of the College Scorecard, aims to evaluate whether college degree attainment has improved across various demographic groups. Our research is anchored in the rich literature that collectively supports the notion that a persistent academic achievement gap is observed within and across higher education institutions.

Previous research on divergent degree attainment outcomes has underscored the importance of identifying factors that enhance college readiness, provide financial support, improve retention rates, and ultimately contribute to higher graduation rates. These prior studies consistently highlight student-related factors, including pre-college academic performance (high school GPAs), national standardized test scores

(SAT/ACT), socioeconomic status (Pell Grant recipients), first-generation college status, gender, and ethnicity (racially minoritized groups), which significantly impact students' core course completion rates, retention rates, and graduation rates.

Building on this foundation, numerous empirical studies have examined how student demographic and academic characteristics influence the likelihood of on-time graduation. Toutkoushian (2025) constructs a panel dataset spanning student data from 2003 to 2020. Employing a two-way fixed-effects model, he investigated whether changes in compositional diversity—specifically in race, gender, major, and ACT scores—were linked to shifts in retention and graduation rates. The results suggest that increasing diversity in terms of race/ethnicity and academic ability may, if anything, hinder institutional efforts to enhance retention and graduation outcomes. These findings are unexpected and concerning, as prior studies have documented the psychosocial and attitudinal benefits of racial diversity for all students.

Demeter et al. (2022) utilize machine learning algorithms, particularly the Random Forest model, to analyse key predictors of graduation outcomes. Their study identifies credit hours completed, college and high school GPA, estimated family financial support, and performance in essential gateway courses within the student's field of study as significant factors. Together, these indicators achieved an impressive overall prediction accuracy of 79% for on-time graduation outcomes. The relationship between high school GPAs and national standardized test scores (SAT/ACT) and college graduation rates has been extensively examined in existing research. Some studies suggest that high school GPAs are stronger predictors of college readiness and degree attainment outcomes (Allensworth and Clark, 2020; Bowen et al., 2009; Geiser & Santelices, 2007; Hiss & Franks, 2014; Jackson & Kurlaender, 2014; Kobrin et al.,2008; Tierney et al., 2009; Tucker & McKnight, 2019) than the conventional views that national standardized test scores are a more robust yardstick for assessing the student readiness and degree attainments. Although high school GPAs have been proven to be a reliable predictor of student college readiness, Klasik and Strayhorn (2018) report that the lack of consideration for ethnicity leads to an overestimation of college readiness for racially and ethnically minoritized students. Hall et al. (2017), in their study of 1,300 college students, found that a combined measure of ethnic-racial discrimination from professors and peers was associated with reduced academic confidence, but no notable differences were observed between the various ethnic-racial groups. Toro and Hughes (2020) demonstrate that timely graduation is equally negatively impacted by both peer and professor discrimination. Further, peer discrimination has a stronger effect on worsening students' mental health and health issues during their first year of college. While experiences of discrimination varied by ethnicity and race, its effects on student outcomes were consistent across all groups. In their study of all 4-year institutions for the 2011 cohort, Ginder et al. (2018) found that graduation rates were lower for Native American, Black, Native Islanders, and Latinx students. Furthermore, the graduation rate for recipients of the Pell Grant was 48%, which is notably lower than the overall graduation rate of 60.8% among all students. The National Center for Education Statistics (NCES) reported similar trends in 6-year graduation rate for first-time, full-time undergraduate students in the 2016 cohort. The lowest graduation rates were observed for Black students (45.2%) and Latinx students (48.7%), compared to Asian students (77.8%) and White students (68.1%).

In addition, parental education level, a common indicator of first-generation college status, has been consistently associated with college outcomes. Previous empirical evidence consistently shows that, on average, students whose parents did not attend college are less likely to enroll in college (Choy, 2001; Ward et al., 2012). Even if they do enroll, they are less likely to graduate (Choy, 2001; Engle & Tinto, 2008). Corroborating these prior findings, Toutkoushian et al. (2021) confirm a direct and significant relationship between parental education level and higher college completion rates. Additionally, their research delves into more detailed classifications of FGCS (first-generation college students), revealing that compared to their peers with two college-educated parents, students with only one college-educated parent face a greater risk of not graduating from college. Research indicates that FGCSs are more likely to come from low socioeconomic status families, belong to ethnic or racial minorities, and tend to be female and older (Chen & Carroll, 2005; Choy, 2001; Pascarella et al., 2004; Toutkoushian et al.). Well-documented findings have shown how financial aid influences student success (Braunstein et al., 2000) and how students burdened by financial concerns or high levels of debt are also more likely to take more than four years to attain a college

degree (Letkiewicz et al., 2014). Targeting timely degree completion for first-generation and low-income transfer students, Sumihig (2016) identifies significant factors, including pre-transfer units, term 1 GPA, and term 1 enrolled units, that impact their degree attainment within two years of transferring.

Gender has been identified as a consistent predictor, with females generally exhibiting higher persistence and graduation rates than males across most higher education contexts (Goldrick-Rab, 2006). Since 1982, female college graduates have consistently outpaced their male counterparts in degree attainment rates across most ethnic groups and socioeconomic backgrounds (Buchman & DiPrete, 2006; Goldin et al., 2006). These disparities are often attributed to differences in academic engagement, social integration, and non-cognitive factors, such as motivation and time management skills. For example, Buchmann and DiPrete (2006) note that women tend to complete college at higher rates than men, largely due to variations in academic habits and perspectives on education. Bastedo et al. (2023) create contextualized measures of math, science, English, high school GPA, and ACT scores alongside traditional raw scores in their logistic regression model. Their findings indicate that both raw and contextualized ACT scores are associated with first-year GPA for Pell Grant recipients and women. For example, women outperform their male counterparts in both retention and four-year graduation rates, though the difference in retention rates between genders is relatively small. The disparity in college degree attainment rates between females and males has widened. According to DeAngelo et al. (2011), 43.8% of females achieved college degrees, compared to 32.9% of males. This difference is larger than the 7.1% gap (32.6% for males vs. 39.7% for females) reported by Astin and Oseguera (2005). Keels (2013) demonstrates that the influence of gender on college degree attainment varies to some extent depending on the moderating effect of socioeconomic status. After accounting for sociodemographic factors, the gender gap in degree attainment widened among Black students, while it narrowed and became statistically insignificant for Latinx students.

Collectively, these studies reinforce the idea that demographic and academic inputs available at the time of college entry are powerful predictors of students' ability to persist and graduate on time. Our study extends this body of research by applying these factors to a large dataset from a diverse public university system, providing insights that can inform early intervention strategies to support timely degree attainment.

DATA AND ANALYSIS

Description of Dataset

The dataset for this project includes student demographic information and academic performance data for all COB students in incoming cohorts from Fall 2011 to Fall 2021. Our study utilizes a decade-long dataset of over 15,000 students to predict time to degree completion and extract insights for both strategic and operational decision-making in higher education. More specifically, the dataset contains student information including age, term of entry, student type (freshmen or transfer), major at entry, gender, race and ethnicity, minority status, first generation, Pell Grant eligibility, high school GPA, transfer GPA, SAT and ACT scores, term and major when degree was awarded, and cumulative GPA upon graduation.

Summary Statistics

The dataset contains information for a total of 15,983 students, and among them 6,962 are first-time freshmen students and 9,021 are transfer students. The student body is comprised of 46% female and 54% male students. Table 1 presents the age distribution for both freshmen group and transfer group. As indicated in the table, the freshman group is quite homogeneous in terms of age, with most students starting their freshman year at 18 or 19. Transfer group, on the other hand, exhibits large variations in terms of age distribution. There are over 30% of the students who started their junior year at the age of 25 or older. The oldest student observed in the dataset is 71.

TABLE 1
STUDENT AGE DISTRIBUTION

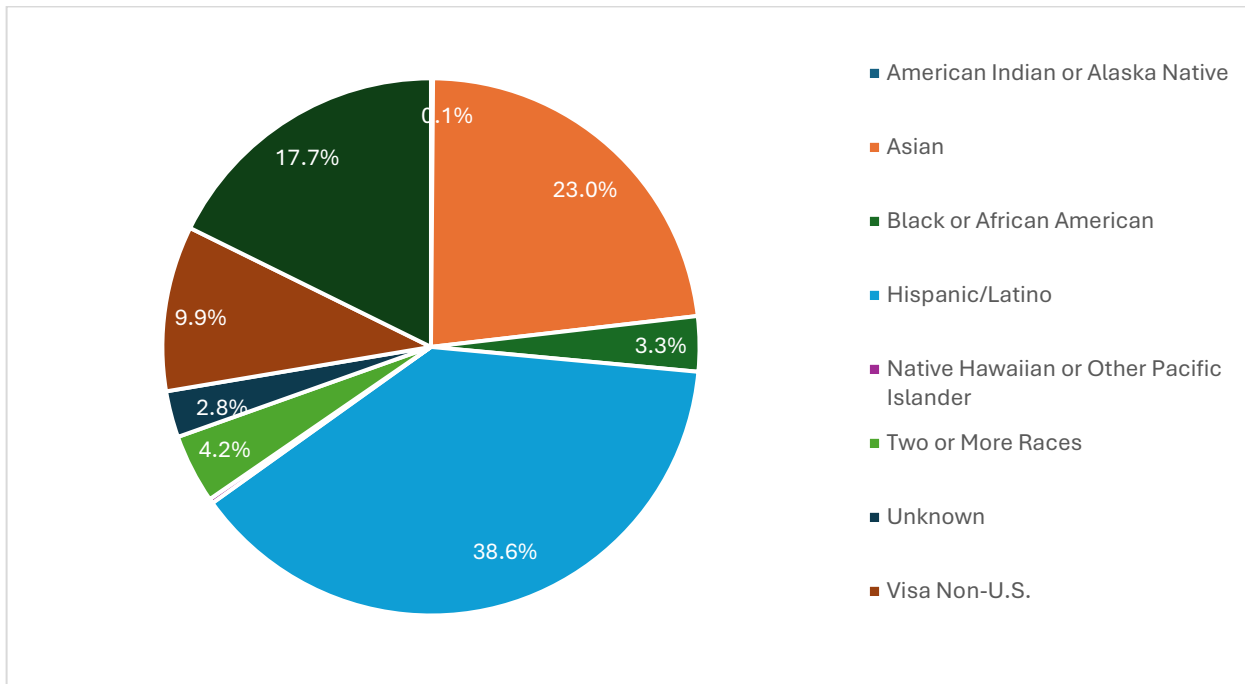| Age Group	FRESHMEN	TRANSFER	All
19 and younger	95.8%	0.4%	42.0%
20-24	3.9%	66.1%	39.0%
25-29	0.2%	21.3%	12.1%
30-39	0.0%	9.8%	5.5%
40-49	0.0%	1.9%	1.0%
50-59	0.0%	0.5%	0.3%
60 and older	0.0%	0.1%	0.1%
Sum	**100.0%**	**100.0%**	**100.0%**

FIGURE 1
STUDENT RACE DISTRIBUTION

Figure 1 presents the composition of races among all students. The largest demographic is Hispanic and Latino students, comprising 38.6% of the total student population, followed by Asian students at 23%, and white students at 17.7%. Note in U.S. higher education, Hispanic/Latino, Black, and American Indian or Alaska Native are classified as underrepresented minorities. Therefore, in our dataset, these three groups collectively account for more than 40% of the total student population.

Table 2 presents the information on parental education level across different groups of students. Overall, 31.3% of the students have parents who have graduated from college, while slightly more (32.9%) are first-generation college students. Among minority students, only 18.3% have parents who graduated from college, while 47.5% are first-generation students. In contrast, 43.5% of the non-minority students have parents who graduated from college and only 21.1% are first-generation students.

TABLE 2
PARENTS' EDUCATION LEVEL

Parent Education Level	Minority	Non-Minority	Unknown	Visa Non-U.S.	All
Parent Attended Some College	25.4%	28.8%	26.2%	8.6%	25.3%
Parent Graduated College	18.3%	43.5%	35.0%	29.5%	31.3%
Student is First Generation to Attend College	47.5%	21.1%	22.9%	27.5%	32.9%
UNKNOWN	8.8%	6.6%	15.9%	34.4%	10.5%
Grand Total	**6727**	**7220**	**446**	**1590**	**15983**

Lastly, we examine family income and students' financial needs. Table 3 presents Pell Grant eligibility for various student groups. Overall, 47.4% of the students are Pell eligible while 44.3% are not Pell eligible. For minority students, the percentage of Pell eligible students rises to 61.8%, with only 29.3% not eligible. These numbers illustrate the financial challenges that many minority students face while attending college.

TABLE 3
PELL GRANT ELIGIBILITY

Family Income	Minority	Non-Minority	Unknown	Visa Non-U.S.	All
NON TRADITIONAL	8.9%	8.9%	10.5%	2.8%	8.3%
NOT PELL ELIGIBLE	29.3%	47.0%	45.3%	94.8%	44.3%
TRADITIONAL	61.8%	44.0%	44.2%	2.4%	47.4%
Grand Total	**6727**	**7220**	**446**	**1590**	**15983**

In this study, one of the key metrics used to measure student success is the time it takes for a student to graduate, if they ever do. Among the freshmen group, overall, 26.0% graduated on time or within four years, 58.3% graduated within six years, and 60.3% managed to graduate including those who graduated within an extended period of time longer than six years. For transfer group, 33.3% graduated on time or within two years, 72.0% graduated within four years, and 73.9% managed to graduate including those who graduated within an extended period of time longer than four years.

Measure of Academic Achievement Gap

In this section, we analyze and measure academic achievement gaps based on gender, minority status, and parents' educational level. The results, based on age, family income, and prior academic performance, are presented in the appendix. We use two metrics to measure academic achievements, graduation rates within a given time window (4-year and 6-year for freshmen students; and 2-year and 4-year for transfer students), and cumulative GPA at graduation. In Section 3.3.4, we also compare academic achievements for different groups of students with similar level of prior academic performance.

Academic Achievement Gaps by Gender

Table 4 shows academic achievement gaps by gender. For transfer group, female students clearly outperform male students in terms of GPA (3.15 as compared to 3.09). However, female students only slightly outperform male students in terms of graduation rates. In contrast, female freshmen students significantly outperform their male counterparts in both dimensions. For instance, the 4-year graduation rate for female students is 30.4%, compared to 21.6% for male students. The GPA for female students is 3.17 as compared to 3.04 for male students.

TABLE 4
GRADUATION RATES AND GPA BY GENDER

Gender	Student Count	Graduation Rates				GPA
		2-Year	4-Year	6-Year	Overall	
FRESHMEN	6962		26.0%	58.3%	60.3%	3.11
FEMALE	3497		30.4%	60.8%	62.1%	3.17
MALE	3465		21.6%	55.7%	58.5%	3.04
TRANSFER	9021	33.3%	72.0%		73.9%	3.12
FEMALE	3863	33.8%	72.8%		74.9%	3.15
MALE	5157	33.0%	71.4%		73.1%	3.09
UNKNOWN	1	0.0%	0.0%		0.0%	
Grand Total	**15983**				**68.0%**	**3.11**

Academic Achievement Gaps by Minority Status

Table 5 presents academic achievement gaps by minority status. Minority students exhibit significantly lower graduation rates and GPAs compared to their peers, a trend observed among both freshmen and transfer students. For instance, for the freshman group, the 4-year/6-year graduate rates for minority students are only 20.7%/52.8%, compared to 30.3%/64.6% for non-minority students. Minority students' GPA at graduation is 3.03 as compared to 3.15 for non-minority students.

TABLE 5
GRADUATION RATES AND GPA BY MINORITY STATUS

Minority Status	Student Count	Graduation Rates				GPA
		2-Year	4-Year	6-Year	Overall	
FRESHMEN	6962		26.0%	58.3%	60.3%	3.11
Minority	3097		20.7%	52.8%	55.1%	3.03
Non-Minority	3066		30.3%	64.6%	66.4%	3.15
Unknown	133		29.3%	57.1%	57.9%	3.16
Visa Non-U.S.	666		30.3%	55.1%	56.9%	3.17
TRANSFER	9021	33.3%	72.0%		73.9%	3.12
Minority	3630	30.6%	68.3%		70.5%	3.06
Non-Minority	4154	35.4%	74.6%		76.2%	3.14
Unknown	313	34.2%	72.8%		74.4%	3.13
Visa Non-U.S.	924	34.7%	74.4%		76.6%	3.20
Grand Total	**15983**				**68.0%**	**3.11**

Academic Achievement Gaps by Parents' Education Level

Table 6 indicates that first-generation students significantly underperform in both graduation rates and GPA compared to their peers, among both freshmen and transfer students. For instance, the 4-year graduate rate for first-generation students in the freshman group is only 20.3%, compared to 30.7% for those whose parents graduated from college. Additionally, first-generation students have an average graduation GPA of 3.06, while those with college-educated parents achieve an average GPA of 3.14.

TABLE 6
GRADUATION RATES AND GPA BY PARENTS' EDUCATION LEVEL

Parent Education Level	Student Count	Graduation Rates				GPA
		2-Year	4-Year	6-Year	Overall	
FRESHMEN	6962		26.0%	58.3%	60.3%	3.11
Parent Attended Some College	1780		26.6%	59.5%	61.7%	3.10
Parent Graduated College	2463		30.7%	61.3%	63.4%	3.14
Student is First Generation	2073		20.3%	56.9%	58.9%	3.06
UNKNOWN	646		25.1%	48.0%	49.1%	3.12
TRANSFER	9021	33.3%	72.0%		73.9%	3.12
Parent Attended Some College	2262	34.8%	72.9%		74.8%	3.11
Parent Graduated College	2539	37.7%	74.7%		76.4%	3.14
Student is First Generation	3185	28.3%	69.4%		71.4%	3.11
UNKNOWN	1035	35.0%	71.3%		73.2%	3.12
Grand Total	**15983**				**68.0%**	**3.11**

Academic Achievement Gaps by Prior Academic Performance and Minority Status

In this section, we present the first main findings of this study, investigating whether the earlier results hold true when comparing students with similar prior academic performance. To streamline the discussion, we focus on the results for freshmen group categorized by minority status. Table 7 presents the results across different high-school GPA ranges. The results indicate that, even among those with comparable prior academic performance, minority students continue to underperform their peers in terms of both graduation rates and GPA upon graduation. Similar patterns are observed for age, gender, parents' education level, and family income. In summary, all findings identified earlier hold true even when comparing students with similar prior academic performance. These findings underscore the persistent influence of demographic and socioeconomic factors, highlighting critical areas for further investigation and intervention.

TABLE 7
FRESHMEN GRADUATION RATES AND GPA BY HIGH SCHOOL GPA AND
MINORITY STATUS

High School GPA	Student Count	Graduation Rate			GPA
		4-Year	6-Year	Overall	
2.00-2.49	**16**	**6.3%**	**25.0%**	**31.3%**	**2.67**
Minority	3	0.0%	33.3%	66.7%	2.43
Non-Minority	8	12.5%	25.0%	25.0%	2.86
Visa Non-U.S.	5	0.0%	20.0%	20.0%	2.78
2.50-2.99	**590**	**12.5%**	**47.3%**	**52.4%**	**2.83**
Minority	279	7.9%	41.6%	48.0%	2.78
Non-Minority	243	14.0%	52.7%	56.8%	2.83
Unknown	14	7.1%	42.9%	42.9%	2.84
Visa Non-U.S.	54	31.5%	53.7%	57.4%	3.03

3.00-3.49	**2472**	**20.2%**	**58.2%**	**60.7%**	**2.96**
Minority	1186	16.2%	52.6%	54.8%	2.90
Non-Minority	1031	23.8%	66.0%	69.0%	3.00
Unknown	45	15.6%	48.9%	51.1%	2.93
Visa Non-U.S.	210	26.2%	53.3%	55.2%	3.11
3.50-3.99	**3034**	**31.7%**	**62.5%**	**63.9%**	**3.19**
Minority	1307	25.9%	57.2%	59.1%	3.11
Non-Minority	1463	36.6%	68.0%	69.0%	3.24
Unknown	58	43.1%	67.2%	67.2%	3.27
Visa Non-U.S.	206	30.6%	55.3%	57.3%	3.19
4.00-4.49	**712**	**31.6%**	**51.0%**	**51.3%**	**3.47**
Minority	321	27.4%	45.2%	45.2%	3.46
Non-Minority	315	35.2%	54.6%	54.9%	3.50
Unknown	16	37.5%	56.3%	56.3%	3.48
Visa Non-U.S.	60	33.3%	61.7%	63.3%	3.34
4.50-4.99	**1**	**100.0%**	**100.0%**	**100.0%**	**2.99**
Minority	1	100.0%	100.0%	100.0%	2.99
NO DATA	**137**	**36.5%**	**56.2%**	**56.9%**	**3.22**
Non-Minority	6	50.0%	50.0%	50.0%	3.20
Visa Non-U.S.	131	35.9%	56.5%	57.3%	3.22
Grand Total	**6962**	**26.0%**	**58.3%**	**60.3%**	**3.11**

PREDITION MODEL

This section constructs several logistic regression models to estimate the probability of a specific student graduating within a certain timeframe, incorporating students' demographic information and prior academic performance variables outlined in Section 3.3. These prediction models serve two main purposes. First, they assess individual students' probability of academic success, enabling campus administrators and advising office to identify high-risk students early, even with limited information. This allows campus to use limited resources more effectively to support targeted student groups. Second, the models enable administrators to predict the overall success rate for incoming cohorts with varying demographics and admission standards, which results in two key benefits. (1) helping administrators better understand the resource needed to support student success when establishing admission criteria, and (2) providing a common framework for comparing student success across multiple campuses with different demographics and admission criteria. We develop the prediction models for both freshmen (presented in Section 4.1) and transfer students (presented in the appendix). We apply the prediction models to estimate the probability of academic success for individual students in Section 4.2 and to evaluate the graduation rates for any incoming cohorts with different demographics and admission criteria in Section 4.3.

Prediction Model for Freshmen

In this section, we develop two prediction models for freshmen students (named Models FM1 and FM2, respectively) with different demographics and prior academic performance to assess their chances of graduating on time. The demographic factors are gender, minority status, Pell Grant eligibility, and parents' education level. Since there is little variation in age among freshmen, we excluded age as an explanatory variable. FM1 uses high-school GPA as an indicator of prior academic performance while FM2 uses SAT score. We define response variable as 1 if a student graduated within four years and 0 otherwise.

Table 8 presents the results for model FM1, showing all explanatory variables significantly impact on-time graduation probability. The estimated parameter values indicate that "Female," "Parent Graduated College," "Not Pell Eligible," and high school GPA have a positive impact on the probability of on-time graduation, while "minority" has a negative impact on this probability. Notably, high-school GPA has the most impactful effect among all explanatory variables, followed by gender and minority status.

TABLE 8
MODEL FM1 OUTPUT

Effect Summary		
Source	**Logworth**	**PValue**
ESA_HS_GPA_FIN	27.857	0.00000
DAE_SEX_CODE_FIN	8.426	0.00000
DAE_MIN_STATUS	6.008	0.00000
PELLTOT_ELIG_TYPE_FIN	4.147	0.00007
DAE_FIRST_GEN_FIN	3.180	0.00066

Whole Model Test				
Model	**-**	**DF**	**ChiSquare**	**Prob>ChiSq**
Difference	160.2547	8	320.5093	<.0001*
Full	3343.3534			
Reduced	3503.6080			
RSquare (U)	0.0457			
AICc	6704.74			
BIC	6765.19			
Observations (or Sum Wgts)	6123			
AUC	0.64909			

Parameter Estimates				
Term	**Estimate**	**Std Error**	**ChiSquare**	**Prob>ChiSq**
Intercept	-4.3511883	0.312101	194.37	<.0001*
DAE_SEX_CODE_FIN[FEMALE]	0.18200939	0.0309779	34.52	<.0001*
DAE_MIN_STATUS[Minority]	-0.2584354	0.0559016	21.37	<.0001*
DAE_MIN_STATUS[Non-Minority]	0.06775302	0.0533762	1.61	0.2043
DAE_FIRST_GEN_FIN[Parent Attended Some College]	0.04144231	0.0447661	0.86	0.3546
DAE_FIRST_GEN_FIN[Parent Graduated College]	0.13712005	0.0445931	9.46	0.0021*
PELLTOT_ELIG_TYPE_FIN[NON TRADITIONAL]	-0.0191902	0.0834597	0.05	0.8181
PELLTOT_ELIG_TYPE_FIN[NOT PELL ELIGIBLE]	0.16351046	0.0540566	9.15	0.0025*
ESA_HS_GPA_FIN	0.93655357	0.0861677	118.13	<.0001*

FM2, using the same set of explanatory variables, found only SAT score, gender, and Pell eligibility to be significant. Consequently, the model was rerun with only these three explanatory variables, as presented in Table 9. As shown, SAT score has the strongest impact among all explanatory variables, followed by

gender and Pell Grant Eligibility. The estimated parameter values indicate that "Female", "Not Pell Eligible", and SAT score all positively impact the probability of on-time graduation.

TABLE 9
MODEL FM2 OUTPUT

Effect Summary

Source	Logworth	PValue
ESA_SAT_COMP_SCORE	43.420	0.00000
DAE_SEX_CODE_FIN	20.023	0.00000
PELLTOT_ELIG_TYPE_FIN	7.040	0.00000

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	160.6073	4	321.2146	<.0001*
Full	1876.7044			
Reduced	2037.3117			
RSquare (U)	0.0788			
AICc	3763.43			
BIC	3794.09			
Observations (or Sum	3419			
AUC	0.69341			

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-5.3885103	0.3371519	255.44	<.0001*
DAE_SEX_CODE_FIN[FEMALE]	0.37687742	0.0409959	84.51	<.0001*
PELLTOT_ELIG_TYPE_FIN[NON TRADITIONAL]	-0.1533274	0.1084451	2.00	0.1574
PELLTOT_ELIG_TYPE_FIN[NOT PELL ELIGIBLE]	0.30741966	0.0675774	20.69	<.0001*
ESA_SAT_COMP_SCORE	0.00411791	0.0003083	178.43	<.0001*

Predicting Individual Student Success

Applying the prediction models from Section 4.1, we estimate the on-time graduation probabilities for a particular type of student, highlighting their correlation with prior academic performance, such as high school GPA or SAT scores. Figure 2 presents the probabilities of on-time graduation for six representative types of students with varying high-school GPA based on the results of model FM1. The details of these six types of students are described as follows.

- S1: Female, non-minority, with parents graduated from College, and Not Pell eligible
- S2: Male, non-minority, with parents graduated from College, and Not Pell eligible
- S3: Female, non-minority, first generation, and Pell eligible
- S4: Male, non-minority, first generation, and Pell eligible
- S5: Female, minority, first generation, and Pell eligible
- S6: Male, minority, first generation, and Pell eligible

FIGURE 2
PROBABILITY OF ON-TIME GRADUATION AND HIGH-SCHOOL GPA
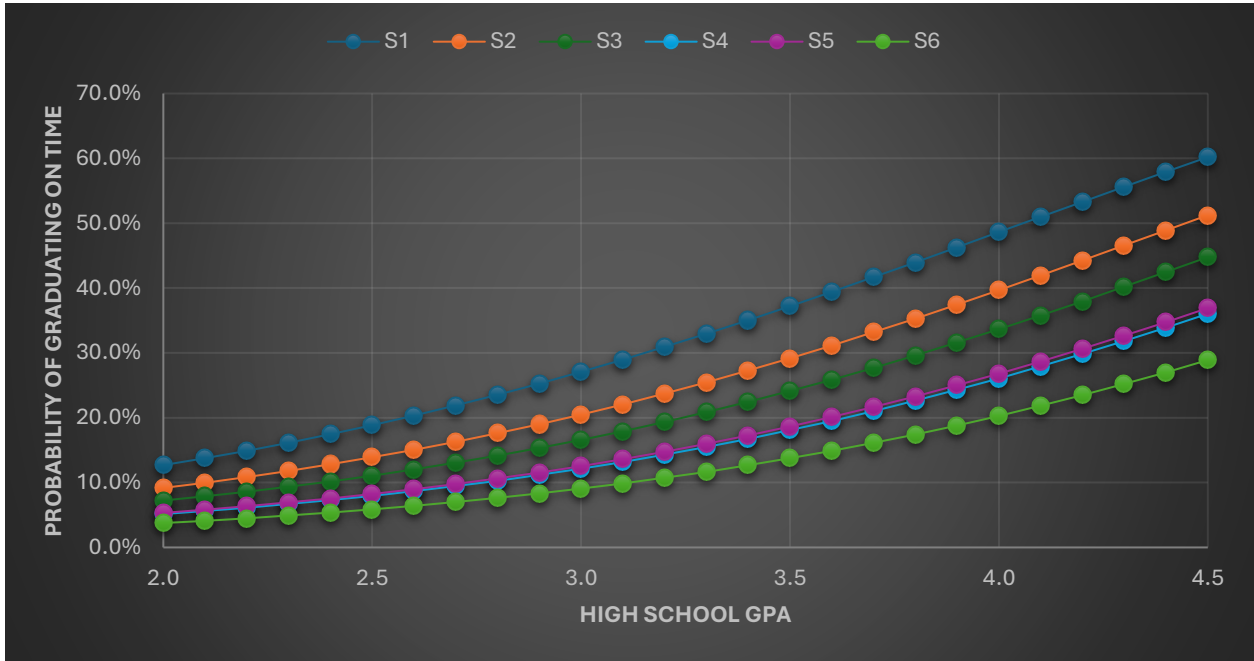

Type S1 represents a female, non-minority student with college-educated parents and no Pell Grant eligibility, while Type S6 represents a male, minority, first-generation college student eligible for a Pell Grant. Among all types of students, S1 students have the highest probability of on-time graduation, while S6 students have the lowest. Surprisingly, an S1 student with high-school GPA 2.5 has a slightly higher chance of graduating on time than an S6 student with high-school GPA 3.9 (18.9% compared to 18.8%). This highlights the need for extra academic support for S6 students, even those with higher GPAs, if on-time graduation is the measure of success.

Figure 3 presents the probabilities of on-time graduation for four representative types of students with varying SAT scores based on the results of model FM2. The details of these four types of students are described as follows.

- T1: Female, and Not Pell eligible
- T2: Male, and Not Pell eligible
- T3: Female, and Pell eligible
- T4: Male, and Pell eligible

FIGURE 3
PROBABILITY OF ON-TIME GRADUATION AND SAT SCORES
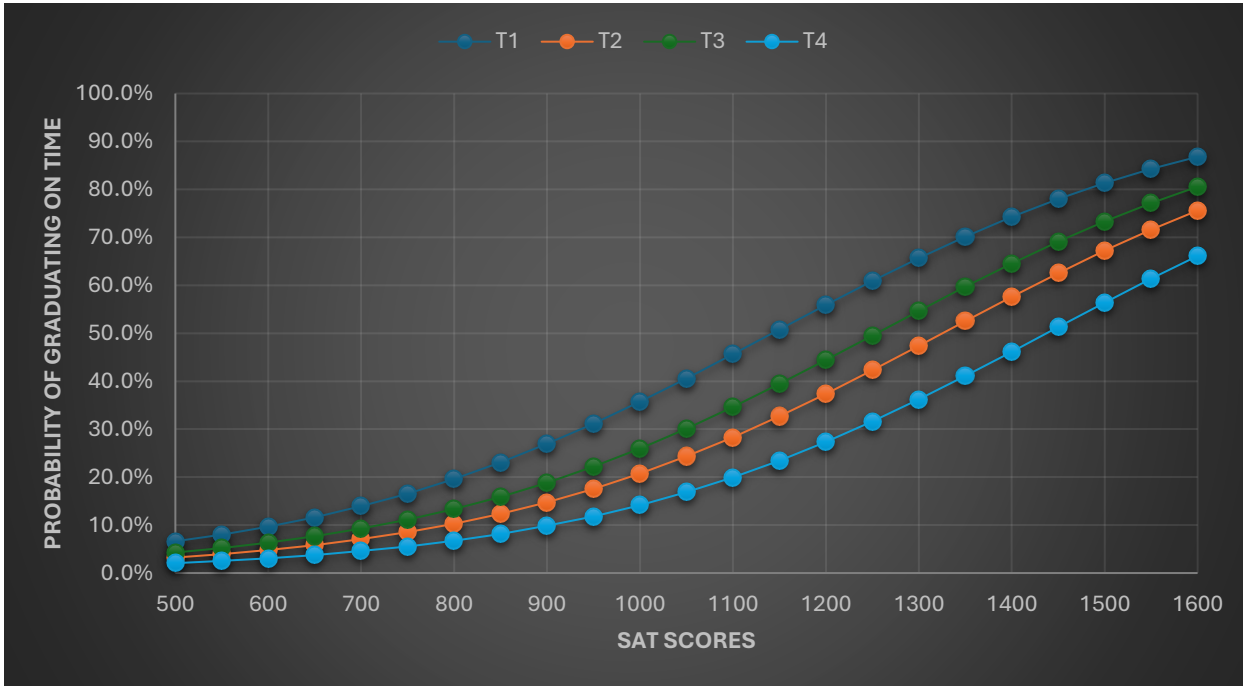
T1 represents a female student ineligible for Pell Grant, while T4 represents a male student eligible for Pell Grant. T1 has the highest on-time graduation probability, and T4 has the lowest. A T1 student with an SAT score of 800 has almost the same graduation chance as a T4 student with an SAT score of 1100 (19.6% vs. 19.9%).

Figures 2 and 3 illustrate how the prediction models enable university administrators to identify high-risk students, allowing them to allocate limited resources effectively to support student success.

Evaluating Campus-Wide Academic Success

This section applies the prediction models from Section 4.1 to estimate on-time graduation rates for campuses with different student composition and admission criteria. We conduct two sets of analysis. First, model FM1 is applied to analyse campuses by different racial breakdowns and entry requirements on high-school GPA (Table 10). Second, model FM2 is applied to examine campuses with different compositions of family income levels and entry requirements based on SAT scores (Table 11).

TABLE 10
**ESTIMATED ON-TIME GRADUATION RATES FOR CAMPUSES WITH DIFFERENT
RACIAL BREAKDOWNS AND HIGH-SCHOOL GPA REQUIREMENTS**

	High School GPA Requirements					
Student Composition Minority, non-minority, VISA	2.0	2.5	2.8	3.0	3.2	3.5
Campus A: 40%, 40%, 20%	25.2%	25.5%	26.2%	27.0%	28.6%	31.1%
Campus B: 20%, 60%, 20%	26.8%	27.1%	27.9%	28.7%	30.4%	33.0%
Campus C: 50%, 40%, 10%	24.0%	24.2%	24.9%	25.7%	27.2%	29.7%
Campus D: 40%, 50%, 10%	24.8%	25.1%	25.8%	26.6%	28.1%	30.7%
Campus E: 30%, 60%, 10%	25.6%	25.9%	26.6%	27.5%	29.0%	31.6%
Campus F: 20%, 70%, 10%	26.4%	26.7%	27.4%	28.3%	29.9%	32.6%
Campus G: 40%, 55%, 5%	24.6%	24.9%	25.6%	26.4%	27.9%	30.4%
Campus H: 20%, 75%, 5%	26.2%	26.5%	27.2%	28.1%	29.7%	32.3%

TABLE 11
**ESTIMATED ON-TIME GRADUATION RATES FOR CAMPUSES WITH DIFFERENT
FAMILY INCOME COMPOSITION AND SAT SCORES REQUIREMENTS**

	SAT Score Requirements				
Student Composition Non-traditional, not Pell-eligible, traditional	600	800	1000	1200	1400
Campus I: 8%, 82%, 10%	29.9%	31.2%	37.9%	51.5%	67.5%
Campus J: 8%, 72%, 20%	29.1%	30.4%	36.9%	50.4%	66.5%
Campus K: 8%, 62%, 30%	28.3%	29.5%	35.9%	49.3%	65.4%
Campus L: 8%, 52%, 40%	27.4%	28.7%	35.0%	48.2%	64.4%
Campus M: 8%, 42%, 50%	26.6%	27.8%	34.0%	47.1%	63.4%
Campus N: 8%, 32%, 60%	25.8%	26.9%	33.0%	46.0%	62.4%

Higher education institutions serving diverse student populations often encounter unique challenges in achieving success metrics, such as on-time graduation rates, particularly when these metrics overlook differences in student demographics and socioeconomic contexts. This analysis is based on the critical assumption that institutional efforts and resource allocations remain constant. Under this assumption, we run empirical models incorporating various combinations of student demographics and academic metrics, generating the results presented in Tables 10 and 11. This assumption is essential, as it allows us to argue

that on-time graduation rates can be considered equally comparable across different spectrums of student demographics and academic metrics when analysed through this lens. Ultimately, the findings in Tables 10 and 11 provide a framework for a more equitable evaluations of institutional efforts to promote student success. For example, as shown in Table 10, Campus C, with 50% of its students from underrepresented groups and a minimum GPA requirement of 2.5, achieving a 24.2% on-time graduation rate, demonstrates comparable performance to Campus F, which has 20% underrepresented students, a minimum GPA requirement of 2.5, and a 26.7% on-time graduation rate. Similarly, the results in Table 11 show that Campus M, where 58% of students are Pell Grant recipients and the minimum SAT requirement is 800, achieving a 27.8% on-time graduation rate, performs similarly to Campus I, where only 18% of students are Pell eligible, the minimum SAT requirement is 800, and the on-time graduation rate is 31.2%. These examples illustrate how our models can serve as a standardized framework for objectively evaluating institutional efforts to foster student success, considering differences in student composition and admission criteria.

CONCLUDING REMARKS

When evaluating academic progress, one of the key measures of student success is the time it takes to graduate. Previous studies have shown that factors such as student demographics, socioeconomic status, and prior academic achievement significantly affect the time required to attain a degree. The COB's diverse student population reflects these dynamics. This study cohort comprises both first-time freshmen and transfer students, with women accounting for 46% of the student body. Notably, minority students (Hispanic/Latino, Black, and American Indian or Alaska Native) constitute over 40% of the total population. Within this minority group, 47.5% are first-generation students, in contrast to 21.1% among non-minority students. Additionally, 61.8% of minority students qualify for Pell Grants, underscoring the financial challenges they face in pursuing higher education, which often exacerbates other barriers to degree completion.

Our prediction models provide a robust, data-driven framework for analyzing graduation rates at the university level, accommodating the complexities of diverse student populations, including both incoming freshmen and transfer students. By integrating prior academic records-such as high school GPA, SAT scores, and transfer GPA-with demographic factors like gender, minority status, Pell Grant eligibility, and parents' education level, these models offer nuanced insights into the factors influencing graduation outcomes. This approach ensures a comprehensive understanding of the diverse challenges faced by higher education institutions serving varied student demographics.

Higher education institutions serving diverse student populations often face unique challenges in achieving success benchmarks, such as on-time graduation rates. Evaluating these metrics without factoring in differences in student demographics or socioeconomic backgrounds can obscure the significant efforts of campuses to support student success. Our findings address this limitation by providing a standardized framework for equitable comparisons across institutions with different student compositions and admission criteria. As demonstrated in our empirical findings, our models account for these disparities, enabling a fair assessment of institutional performance. For instance, campuses with higher proportions of underrepresented or Pell-eligible students can achieve graduation outcomes that are contextually comparable to those of campuses with less diverse student populations, even when their on-time graduation rates differ. This underscores the importance of considering student demographics and baseline admission criteria when evaluating institutional performance, rather than relying solely on graduation rate comparisons.

The flexibility of our proposed prediction models allows for targeted analyses of specific student groups, further highlighting academic achievement gaps and informing strategies to address them. For incoming freshmen, SAT scores emerge as the strongest predictors of four-year graduation success, followed by gender and Pell Grant eligibility. For transfer students, Pell Grant eligibility and transfer GPA have the most significant influence, along with age, first-generation status, and minority status. These insights reveal not only the predictors of success but also the systemic barriers that may disproportionately

affect certain groups. For example, while female, non-minority, non-Pell-eligible students with college-graduate parents consistently show the highest likelihood of on-time graduation, male, minority, first-generation, Pell-eligible students face persistent challenges, even with strong academic credentials. Such findings underscore the need for tailored support strategies to ensure that high academic performance translates into equitable graduation outcomes for all students.

In addition to pinpointing high-risk students for early intervention, our proposed predictive models serve as a strategic tool for institutional planning. By providing accurate estimates of students' graduation probabilities, administrators can allocate resources more efficiently, refine admission standards, and benchmark institutional efforts across campuses. Furthermore, the models enable proactive advising strategies that identify students most in need of support, ensuring that interventions are targeted and impactful. This approach not only optimizes resource allocation but also enhances institutions' ability to promote student success effectively.

Ultimately, this study highlights the effectiveness of a data-driven approach in enhancing organizational outcomes within the context of higher education. By analyzing a range of academic and demographic factors that influence graduation outcomes, our prediction models empower higher education institutions to proactively identify and support students who may be at risk of delayed graduation or attrition. Implementing these models enables the design of targeted, evidence-based interventions aimed at enhancing timely graduation rates and overall student academic achievement. These findings underscore the importance of utilizing predictive analytics as a strategic tool for fostering student achievement and optimizing institutional resources. As such, this framework provides a practical and scalable approach to enhancing educational outcomes across diverse student populations.

REFERENCES

Allensworth, E.M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher*, *49*(3), 198–211.

Arnold, K.E., & Pistilli, M.D. (2012). Course signals at Purdue: Using learning analytics to increase student success. Proceedings of the *2nd International Conference on Learning Analytics and Knowledge*, pp. 267–270. https://doi.org/10.1145/2330601.2330666

Austin, A.W., & Oseguera, L. (2005). *Degree attainment rates at American colleges and universities* (Revised ed.). Loas Angeles: Higher Education Research Institute, UCLA.

Barbera, S.A., Berkshire, S.D., Boronat, C.B., & Kennedy, M.H. (n.d.). Review of undergraduate student retention and graduation since 2010: Patterns, predictions, and recommendations for 2020. *Journal of College Student Retention: Research, Theory & Practice*, *22*(2), 227–250.

Bastedo, M.N., Umbricht, M., Bausch, E., Byun, B.-K., & Bai, Y. (2023). How well do contextualized admissions measures predict success for low-income students, women and unrepresented students of color? *Journal of Diversity in Higher Education*. https://dx.doi.org/10.1037/dhe0000499

Braunstein, A., McGrath, M., & Pescatrice, D. (2000). Measuring the impact of financial factors on college persistence. *Journal of College Student Retention: Research, Theory & Practice*, *2*(3), 191–203.

Bowen, W.G., Chingos, M.M., & McPherson, M.S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton University Press. 1–416.

Buchman, C., & DiPrete, T. (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review*, *71*, 515–541.

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

Chen, X., & Caroll, C.D. (2005). *First-generation students in postsecondary education: A look at their college transcripts. Postsecondary education descriptive analysis report.* National Center for Education Statistics, pp. 1–103.

Choy, S. (2001). Students whose parents did not go to college: Postsecondary access, persistence, and attainment. *The Condition of Education*, pp. 18–43.

DeAngelo, L., Franke, R., Hurtado, S., Pryor, J.H., & Tran, S. (2011). *Completing college: Assessing graduation rates at four-year institutions.* Loas Angeles: Higher Education Research Institute, UCLA.

Demeter, E., Dorodchi, M., Al-Hossami, E., Benedict, A., Walker L.S., & Smail, J. (2022). Predicting first-time-in-college students' degree completion outcomes. *Higher Education*, *84*, 589–609. https://doi.org/10.1007/s10734-021-00790-9

Dougherty, K.J., Jones, S.M., Lahr, H., Natow, R.S., Pheatt, L., & Reddy, V. (2016). *Performance funding for higher education.* Baltimore: Johns Hopkins University Press.

Engle, J., & Tinto, V. (2008). *Moving beyond access: College success for low-income, first-generation students.* Pell Institute for the Study of Opportunity in Higher Education. Retrieved from http://www.pellinstitute.org/publications-Moving_Beyond_Access_2008.shtml

Galla, B.M., Shulman, E.P., Plummer, B.D., Gardner, M., Hutt, S.J., Goyer, J.P., . . . Duckworth, A.L. (2019). Why high school grades are better predictors of on-time college graduation than are admissions test scores: The roles of self-regulation and cognitive ability. *American Educational Research Journal*, *56*(6), 2077–2155. https://doi.org/10.3102/0002831219843292

Geiser, S., & Santelices, M.V. (2007). Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes. *Center for Studies in Higher Education*, pp. 1–35.

Ginder, S.A., Kelly-reid, J.E. & Mann, F.B. (2018). *Graduation rates for selected cohorts, 2009-14; outcome measures for cohort year 2009-10; student financial aids, academic year 2016-17; and admissions in postsecondary institutions, Fall 2017.* National Center for Education Statistics, pp. 1–32.

Goldin, C., Katz, L., & Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, *20*, 133–156.

Hiss, W.C., & Franks, V.W. (2014). Defining promise: Optional standardized testing policies in American college and university admissions. *The National Association for College Admission Counselling*, pp. 1–70.

Jackson, J., & Kurlaender, M. (2014). College readiness and college completion at broad access four-year institutions. *American Behavioural Scientist*, *58*(8), 947–971. https://doi.org/10.1177/0002764213515229

Jayaprakash, S.M., Moody, E.W., Lauría, E.J.M., Regan, J.R., & Baron, J.D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, *1*(1), 6–47. https://doi.org/10.18608/jla.2014.11.3

Keels, M. (2013). Getting them enrolled is only half the battle: College success as a function of race or ethnicity, gender, and class. *American Journal of Orthopsychiatry*, *83* (2, 3), 310–332.

Klasik, D., & Strayhorn, T.L. (2018). The complexity of college readiness: Differences by race and college selectivity. *Educational Researcher*, *47*(6), 334–351. https://doi.org/10.3102/0013189X18778598

Kobrin, J.L., Patterson, B.F., Shaw, E.J., Mattern, K.D., & Barbuti, S.M. (2008). Validity of the SAT® for predicting first-year college grade point average. *Research Report No. 2008-5, College Entrance Examination Board*, pp. 1–15.

Letkiewicz, J., Lim, H., Heckman, S., Bartholomae, S., Fox, J.J., & Montalto, C.P. (2014). The path to graduation: Factors predicting on-time graduation rates. *Journal of College Student Retention: Research*, *Theory & Practice*, *16*(3), 351–371. https://doi.org/10.2190/CS.16.3.c

Perna, L.W., & Finney, J.E. (2014). The attainment agenda: State policy leadership in higher education. Baltimore, MD: Johns Hopkins University Press.

Sumihig, J.A. (2016). *Timely degree completion: the success and setbacks of first-generation and low-income transfer students.* Thesis and Dissertation: California State University, Sacramento.

Tierney, W.G., Bailey, T., Constantine, J., Finkelstein, N., & Hurd, N.F. (2009). *Helping students navigate the path to college: What high schools can do: A practice guide* (NCEE #2009-4066). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/

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

Toro, J.D., & Hughes, D. (2020). Trajectories of discrimination across the college of years: Associations with academics, psychological, and physical adjustment outcomes. *Journal of Youth and Adolescence*, *49*, 772–789. https://doi.org/10.1007/s10964-019-01147-3

Toutkoushian, R.K. (2025). Student compositional diversity and college retention and graduation rates. *Journal of Diversity in Higher Education*, *18*(1), 97–109. https://doi.org/10.1037/dhe0000491

Toutkoushian, R.K., May-Trifiletti, J.A., & Clayton, A.B. (2021). From "first in family" to "first to finish": Does college graduation vary how first-generation college status is defined? *Educational Policy*, *35*(3), 481–521. https://doi.org/10.1177/0895904818823753

Tucker, L., & McKnight, O. (2019). Assessing the validity of college success indicators for the at-risk student: Toward developing a best-practice model, *Journal of College Student Retention: Research, Theory & Practice*, *2*(12), 166–183.

Ward, L., Siegel, M., & Davenport, Z. (2012). *First-generation college students: Understanding and improving the experience from recruitment to commencement*. San Francisco, CA: John Wiley, pp. 1–176.

Witham, K., Malcom-Piqueux, L., Dowd, A.C. & Bensimon, E.M. (2015). *Americas unmet promise: The imperative for equity in higher education.* Washington DC: Association of American Colleges and Universities.

APPENDIX

In Section A.1, we analyse and measure academic achievement gaps based on age, family income, and prior academic performance. These findings complement those presented in the main text, which address achievement gaps related to gender, minority status, and parents' education levels.

Section A.1 Measure of Academic Achievement Gap
Section A.1.1 Academic Achievement Gaps by Age

We begin by examining academic achievement gaps across age groups, focusing on transfer students due to the homogeneity of freshmen age distribution. The results in Table A1 show that graduation rates generally decline with age, for 2-year, 4-year, and overall graduation rates. While older students often take longer to graduate, their GPAs may not be lower than younger students. In fact, students aged 40-49 and 50-59 outperform their younger peers in GPA at graduation.

TABLE A1
GRADUATION RATES AND GPA BY AGE GROUP

| Age Group | Student Count | Graduation Rates | | | GPA |
		2-Year	4-Year	Overall	
19 and younger	34	29.4%	85.3%	88.2%	3.19
20-24	5964	34.9%	74.2%	75.6%	3.12
25-29	1918	31.8%	68.9%	71.6%	3.08
30-39	882	30.5%	66.4%	68.9%	3.17
40-49	167	19.8%	61.7%	68.3%	3.26
50-59	46	15.2%	54.3%	58.7%	3.32
60 and older	10	0.0%	50.0%	60.0%	3.05
Grand Total	**9021**	**33.3%**	**72.0%**	**73.9%**	**3.12**

Section A.1.2 Academic Achievement Gaps by Family Income

Next, we examine academic achievement gaps by family income, using Pell Grant eligibility as a proxy. Table A2 presents the findings for three groups of students, where "traditional" is the group of students who are eligible for Pell Grants and hence are from low-income families. The "Non-Traditional" is the group of students who is Pell-eligible after the first semester, which represents a relatively small percentage of the total population. The results show that low-income students significantly underperform, especially in timely graduation rates. For example, the freshmen 4-year graduation rate for Pell-eligible students is 21.2%, compared to 30.5% for non-eligible students. The transfer 2-year graduation rate for Pell-eligible students is 30.0%, compared to 39.1% for non-eligible students. Additionally, low-income students earned a lower GPA, with an average of 3.06 compared to 3.15 for freshmen, and 3.09 compared to 3.16 for transfers.

TABLE A2
GRADUATION RATES AND GPA BY PELL ELIGIBILITY

| Pell Eligibility | Student Count | Graduation Rates | | | | GPA |
		2-Year	4-Year	6-Year	Overall	
FRESHMEN	6962		26.0%	58.3%	60.3%	3.11
NON TRADITIONAL	402		25.4%	62.9%	70.4%	3.06
NOT PELL ELIGIBLE	3401		30.5%	59.6%	60.9%	3.15
TRADITIONAL	3159		21.2%	56.3%	58.3%	3.06
TRANSFER	9021	33.3%	72.0%		73.9%	3.12
NON TRADITIONAL	930	26.3%	76.7%		80.2%	3.07
NOT PELL ELIGIBLE	3679	39.1%	72.3%		74.0%	3.16
TRADITIONAL	4412	30.0%	70.7%		72.4%	3.09
Grand Total	**15983**				**68.0%**	**3.11**

Section A.1.3 Academic Achievement Gaps by Prior Academic Performance

In this section, we explore the correlation between students' prior academic performance and college success, using pre-college GPA and SAT scores as indicators. Tables A3 and A4 present results for freshmen with different high school GPAs and SAT scores, while Table A5 shows results for transfer students with varying transfer GPAs. The findings clearly demonstrate that prior academic performance

strongly predicts college success, including graduation rates and GPA. For example, freshmen with a high school GPA below 2.5 and an SAT score below 800 tend to struggle and have a much lower chance of academic success. A similar pattern is observed for transfer students.

TABLE A3
FRESHMEN GRADUATION RATES AND GPA BY HIGH-SCHOOL GPA

HS GPA	Student Count	Graduation Rates			GPA
		4-Year	6-Year	Overall	
2.00-2.49	16	6.3%	25.0%	31.3%	2.67
2.50-2.99	590	12.5%	47.3%	52.4%	2.83
3.00-3.49	2472	20.2%	58.2%	60.7%	2.96
3.50-3.99	3034	31.7%	62.5%	63.9%	3.19
4.00-4.49	712	31.6%	51.0%	51.3%	3.47
4.50-4.99	1	100.0%	100.0%	100.0%	2.99
NO DATA	137	36.5%	56.2%	56.9%	3.22
Grand Total	**6962**	**26.0%**	**58.3%**	**60.3%**	**3.11**

TABLE A4
FRESHMEN GRADUATION RATES AND GPA BY SAT SCORES

SAT Score	Student Count	Graduation Rates			GPA
		4-Year	6-Year	Overall	
500-599	3	0.0%	0.0%	33.3%	2.55
600-699	25	8.0%	60.0%	60.0%	2.95
700-799	127	5.5%	59.1%	64.6%	2.89
800-899	410	11.7%	74.1%	79.8%	2.92
900-999	817	18.0%	80.5%	84.2%	2.96
1000-1099	1009	32.5%	84.1%	88.3%	3.08
1100-1199	781	39.1%	85.1%	86.9%	3.16
1200-1299	388	39.7%	82.5%	85.6%	3.16
1300-1399	120	45.8%	82.5%	85.8%	3.36
1400-1499	22	31.8%	59.1%	63.6%	3.29
NO DATA	3260	23.3%	32.5%	32.7%	3.22
Grand Total	**6962**	**26.0%**	**58.3%**	**60.3%**	**3.11**

TABLE A5
TRANSFER GRADUATION RATES AND GPA BY TRANSFER GPA

		Graduation Rates			
Transfer GPA	Student Count	2_Year	4_Year	Overall	GPA
1.50-1.99	2	0.0%	50.0%	50.0%	3.12
2.00-2.49	146	21.2%	52.1%	56.2%	2.53
2.50-2.99	2406	29.4%	67.4%	70.1%	2.78
3.00-3.49	4268	33.3%	73.4%	75.2%	3.10
3.50-3.99	2058	38.2%	76.2%	77.2%	3.50
4.00-4.49	139	41.0%	66.9%	67.6%	3.85
NO DATA	2	100.0%	100.0%	100.0%	3.03
Grand Total	**9021**	**33.3%**	**72.0%**	**73.9%**	**3.12**

Section A.2 Prediction Model for Transfer

In this section, we develop a prediction model for transfer students (Model TF) to assess their chance of graduating on time based on demographics and prior academic performance. The demographic factors included are age, minority status, Pell Grant eligibility, and parents' education level. Transfer GPA is used as an indicator of prior academic performance. The response variable is defined as 1 if a student graduates within two years, and 0 otherwise.

The results presented in Table A6 suggest that all explanatory variables significantly impact a student's probability of on-time graduation. The estimated parameter values indicate that "Parent Graduated College", "Not Pell Eligible" and Transfer GPA positively impact the on-time graduation probability, while "age", "Student is First Generation to Attend College" and "Non-traditional" negatively impact the probability. In particular, Pell-eligibility and transfer GPA have the strongest effect among all explanatory variables, followed by age, first-generation, and minority status.

TABLE A6
MODEL TF OUTPUT

Effect Summary		
Source	Logworth	PValue
PELLTOT_ELIG_TYPE_FIN	13.147	0.00000
TRANSFER_GPA	10.966	0.00000
DAE_AGE	6.725	0.00000
DAE_FIRST_GEN_FIN	4.210	0.00006
DAE_MIN_STATUS	1.848	0.01418
Whole Model Test		

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	108.6855	10	217.3711	<.0001*
Full	5631.1111			
Reduced	5739.7966			
RSquare (U)	0.0189			
AICc	11284.3			
BIC	11362.4			
Observations (or Sum	9019			

Parameter Estimates				
Term	**Estimate**	**Std Error**	**ChiSquare**	**Prob>ChiSq**
Intercept	-1.5309946	0.2352473	42.35	<.0001*
DAE_AGE	-0.0241202	0.0047511	25.77	<.0001*
DAE_MIN_STATUS[Minority]	-0.0003155	0.0471555	0.00	0.9947
DAE_MIN_STATUS[Non-Minority]	0.08155293	0.0449499	3.29	0.0696
DAE_MIN_STATUS[Unknown]	0.09464637	0.0933894	1.03	0.3108
DAE_FIRST_GEN_FIN[Parent Attended Some College]	0.02508484	0.0406449	0.38	0.5371
DAE_FIRST_GEN_FIN[Parent Graduated College]	0.08146445	0.0399558	4.16	0.0415*
DAE_FIRST_GEN_FIN[Student is First Generation to Attend College]	-0.1781918	0.0390087	20.87	<.0001*
PELLTOT_ELIG_TYPE_FIN[NON TRADITIONAL]	-0.2488991	0.0526049	22.39	<.0001*
PELLTOT_ELIG_TYPE_FIN[NOT PELL ELIGIBLE]	0.29057018	0.0378473	58.94	<.0001*
TRANSFER_GPA	0.41699002	0.0615547	45.89	<.0001*