International Electronic Journal of Mathematics Education
2022, 17(4), em0700
e-ISSN: 1306-3030
https://www.iejme.com

MODESTUM

Research Article

OPEN ACCESS

Exploring Calculus I students' performance between varying course times among other predictive variables

Zachariah Benton Hurdle [1*], Enes Akbuga [2], Paul Schrader [3]

[1] Department of Mathematics, Utah Valley University, Orem, UT, USA
[2] Department of Mathematics & Computer Science, Drake University, Des Moines, IA, USA
[3] Information Directorate, Air Force Research Laboratory, Rome, NY, USA [†]
[†]Additional note from the institution: The views expressed are those of the author and do not necessarily represent the official policy or position of the Department of Defense, Department of the Air Force, or the U.S. Government (See DoDI 5230.29 and 5 CFR 2635.807 for this language).
***Corresponding Author:** ZHurdle@uvu.edu

Citation: Hurdle, Z. B., Akbuga, E., & Schrader, P. (2022). Exploring Calculus I students' performance between varying course times among other predictive variables. *International Electronic Journal of Mathematics Education, 17*(4), em0700. https://doi.org/10.29333/iejme/12234

ARTICLE INFO

Received: 25 Mar. 2022

Accepted: 29 Jun. 2022

ABSTRACT

This study focuses on the analysis of certain performance predictors for calculus I. We collected data from 717 students from 2013 through 2018 at a southeastern university in the United States to explore any correlation between course times (particularly very early versus the rest) and student performance in this specific course, along with a handful of other variables. This represented all calculus I students over this time period. A two-proportion test confirmed that time was a significant variable in performance. We then used regression to determine similar impacts of gender, major, instructor, and term on student performance. Initial findings portrayed statistical differences between terms and course times; other findings included the significance of major and instructor in different contexts. Interaction effects were used with time to complete our analysis of its impact, and controls were later used accordingly. We also display appropriate models for comparing categories. We conclude with some basic assertions and argue some departmental recommendations on how to use these findings in undergraduate mathematics education.

Keywords: calculus, course times, college mathematics, mathematics performance

INTRODUCTION

Calculus I, the first course in the calculus sequence, is a critical mathematics course that provides a turning point for student STEM success. Because of this, it is imperative that faculty keenly observe the factors that provide insight into success rates of students working through a challenging undergraduate course which ranks historically as a major blockage to the STEM (science, technology, engineering, and mathematics) career pipeline. In becoming more familiar with data on passing rates in a course like calculus I, departments are equipped to make improved and more informed decisions supporting student success in future mathematics courses and careers. Therefore, the purpose of our study is to determine whether any controllable predictors exist. Much of calculus education research is devoted to solving the student issues in an early STEM course like calclus I (Bressoud et al., 2015). Departments can use this information when setting expectations for students as they navigate through the appropriate calculus sequence in their degree plan. The literature provides one example of potential influencers such as *when* students decide to take a calculus I course. Beyond the course time, there is also timing, which we define as either spring semester (January through May) or fall semester (August through December). Fall is the first semester of the academic year. This leads to a more specific purpose of this paper: determine any factors such as course times or scheduled terms that impact calculus I performance. We look to determine information that could help dictate students' more successful paths through calculus I and, as a result, their STEM degree through completion. Thus, our overall research questions are as follows:

1. How does student performance in calculus I compare between 8:00am and other course start times? Similarly, how does student performance in calculus I compare between spring and fall terms?
2. How does gender, major, instructor, start time, and term predict student performance in calculus I when considered altogether in a multivariable model?

LITERATURE REVIEW

Student retention in STEM has emerged as an important problem in higher education, particularly in the U.S. Calculus I is critical when determining success in any given STEM field, with the more immediate concern in graduation rates. This course is considered a threshold to a successful STEM degree through university (Hall et al., 2003; Wu et al., 2018). Bressoud et al. (2013) even state that calculus I is

> "a filter, discouraging all but the very strongest students from pursuing a career in science or engineering" (p.685).

When we look at the current status of calculus I courses nationally, Bressoud et al. (2015) found that only 50% of 14,000 students from 160 institutions were able to earn an A or a B, while 27% fell into the DFW (i.e., a conventionally failing letter grade of "D", "F", or withdrawal from the course) category (as cited in Akbuga, 2018). This means that the national passing rate (letter grades A, B, and C) for calculus I is roughly 73%. Research continually shows that calculus is consistently a difficult subject for students, mainly because of the abstract concepts required with many of the topics (White & Mitchelmore, 1996).

> "Since calculus is often seen to be highly symbolic in nature, students often try to get through a first course in calculus by manipulating the symbols without developing a real understanding of the meaning of the symbols" (Berry & Nyman, 2003).

Courses that contribute to student fail and withdrawal rates, a pressing issue, are often called "gateway courses". Gateway courses affix students to their degree plan and prohibit progress toward graduation (Bloemer et al., 2017; Koch & Prestilli, 2014). Indeed, calculus I has been argued to be one of these gateway courses for STEM-field majors (Akbuga et al., 2019). Being dissatisfied with their grade is the top reason many undergraduates drop a course (Hall et al., 2003). More alarmingly, Ellis et al. (2016) found that women are 1.5 times more likely than men to switch out of the calculus sequence, ending their path as a STEM major. Motivated by these prior studies, we are convinced that in order to address equity in calculus completion rates, we need a focused investigation on calculus I to identify impactful factors of student success that aligns well with current research studies on the national scale.

Our interests, both generally and in this study, lie in more controllable factors at the departmental or institutional levels. Historically, there are many student-based factors such as anxiety, motivation, and other affective variables. For instance, research shows that self-efficacy and students' personal study skills can be a predictor of GPA (Young-Jones et al., 2013), as can the prerequisite path that students take before calculus I (Hurdle & Mogilski, 2022). However, departmental intervention can and cannot control some of these factors, particularly depending on university policy around the globe. Since course start time is a variable that departments can generally control, we wanted to investigate this as a potential contributor to student success. We also found that the existing literature discussing the impact of course start times on student academic performance, particularly regarding mathematics, portrays conflicting results. For example, Pope (2016) found that their

> "results tend to show that students are more productive earlier in the school day, especially in math" (p. 10),

as it pertains to middle and high school students. Yet, Carrell et al. (2011) found that "early school start times negatively affect student achievement," in a post-secondary setting. However, they admitted that the study was not specific to mathematics.

Overall, while studies have firmly concluded that later class start times have plenty of benefits (such as improving alertness, attendance, tardiness, etc.), the results are mixed on whether academic performance is improved based on class times (Wheaton et al., 2016). For example, one study found that moving courses to later times provided students more time to stay up and abuse substances like alcohol, thus impeding academic performance further (Onyper et al., 2012). There have also been some studies linking sleep quality, duration, and times with effectiveness in the course, one of which found graduation rates improved when school schedules started later in high school (Wahistrom, 2002). Limited research exists regarding term-wise (e.g. fall vs spring) performance comparisons over various course subjects, and the results differ significantly (Eskew, 2013; Reardo et al., 2007). Additionally, early studies indicate that despite gender stereotypes, deeper data analysis shows that males and females perform similarly for precalculus and calculus topics (Bridgeman & Lewis, 1996). Furthermore, Webb (2016) found that math instructor impact varies not by instructional quality, but by grading standards. Many of the findings in the existing research and the study described in this paper can and should be used in departmental and institutional policy implementation, because "academic advisors are uniquely positioned to both affect, and be affected by, important aspects of educational research" (Aiken-Wisniewski et al., 2010). As advisors and professional academics, we are strongly convinced of and affirm the importance of being a one-on-one contact point for undergraduate students equipped with state-of-the-art educational research and pedagogical strategies. Thus, our study pertains to external framing, defined as "the influence of agents outside of the classroom that administration and policy-makers can control" (Hagman et al., 2017).

METHODOLOGY

Our sample consisted of 717 calculus I students, from fall 2013 through spring 2018, at a small university in the southeastern United States, blinded and coded to maintain confidentiality. Data is collected behind the scenes to track institutional effectiveness at this school, and confidentiality was maintained through this office. **Figure 1** shows how the sample was distributed by term. **Figure 2** shows how the sample was distributed by start time. **Figure 3** shows how the sample was distributed by gender. These three graphs demonstrate the undergraduate major selections are dispersed under these categories. In our

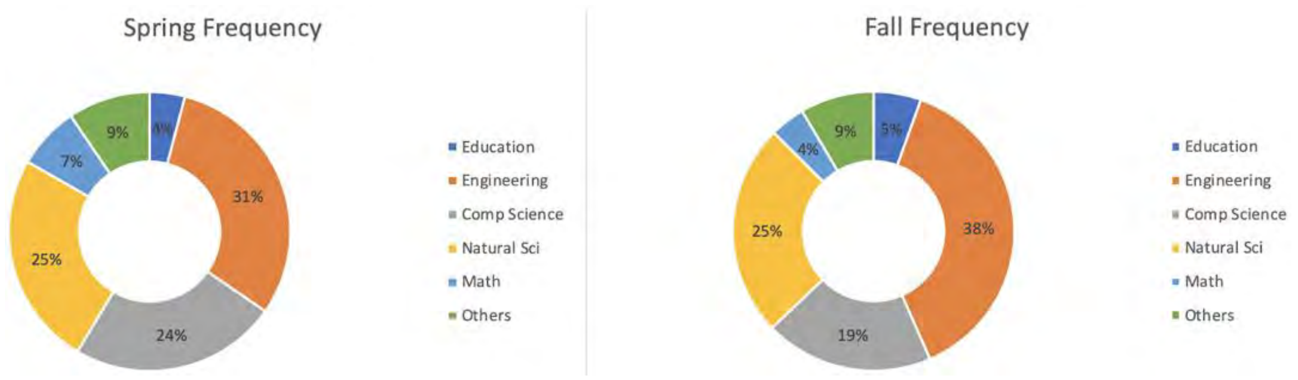

Figure 1. Major distribution by term

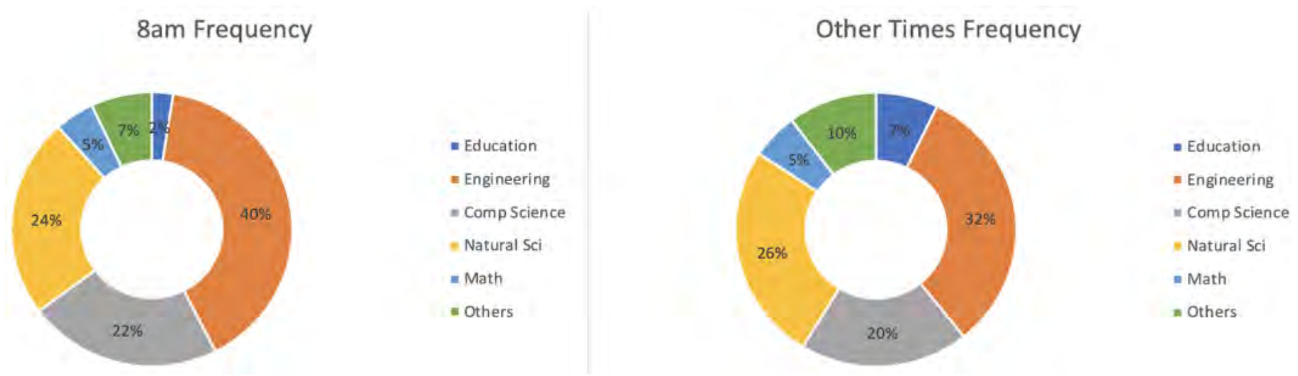

Figure 2. Major distribution by start time

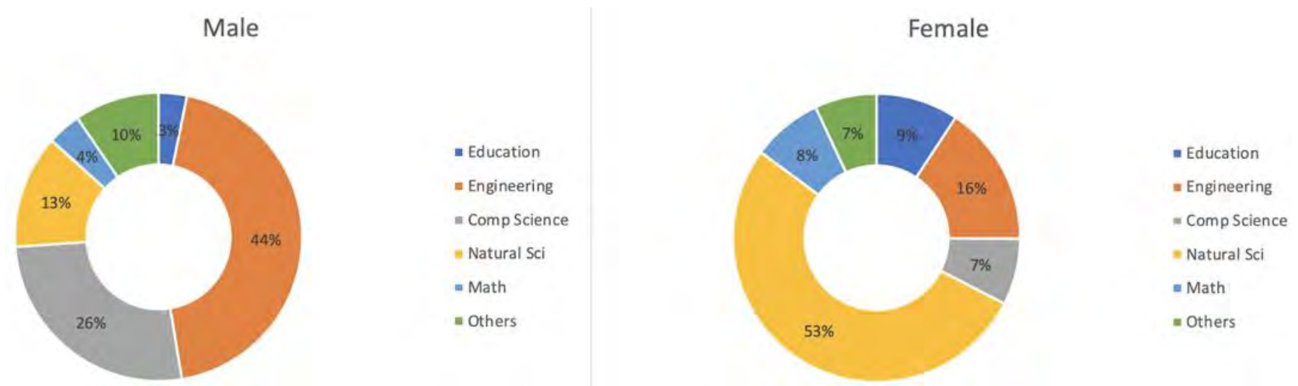

Figure 3. Major distribution by gender

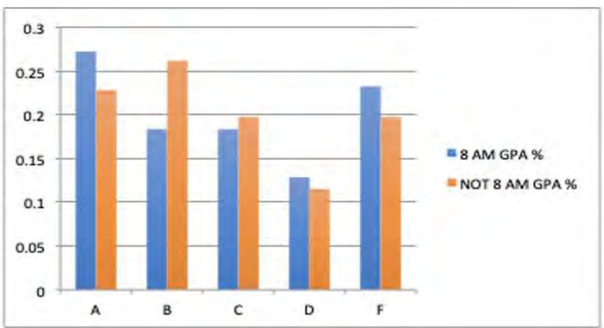

Figure 4. Course time distribution of calculus I GPA

study, we defined "passing" or "passing rate" using the letter grades A, B, or C in the course, and the letter grades D, F, and W rates as not passing grades (the widely-accepted DFW rates). Recall that A is 4.0, B is 3.0, C is 2.0, D is 1.0, and F is 0.0 credits. This grade schema requiring a C or higher is typically an important evaluation of any course, instructor, or program across the globe, particularly in certain degree programs with rigorous requirements at the college level.

Figure 4 shows what percentage of the two-time categories achieved each letter grade in a side-by-side comparison. The percentage of each grade achieved out of all calculus I 8:00am courses is shown in blue, and the percentage of each grade achieved

Table 1. Earned grade rates per time and gender

Grade	8am	Rate	other	Rate	M	F
A	89	.2722	89	.2282	106	72
B	60	.1835	102	.2615	114	48
C	60	.1835	77	.1974	100	37
D	42	.1284	45	.1154	64	23
F	76	.2324	77	.1974	118	35
Total	**327**		**390**		**502**	**215**

Table 2. Grade distributions per major

Grade	Education	Engineering	Computer science	Natural sciences	Mathematics	Other
A	5	59	35	57	15	7
B	8	56	28	48	14	8
C	7	54	25	34	10	7
D	6	30	20	14	9	8
F	10	57	41	24	15	6
ABC Rate	**0.55**	**0.66**	**0.59**	**0.79**	**0.62**	**0.61**

Table 3. Two-proportion, one-tailed testing values

Name	$\hat{p}_1$	$\hat{p}_2$	z-score	α	p-value
8am vs. Others	0.639	0.687	-1.3576	0.1	.08692
Fall vs. Spring	0.663	0.671	-5.5172	0.01	<.00001
$\hat{p} = .665$					

out of all other calculus I courses is shown in orange. Using the value-based schema described, we calculated the mean GPA for 8:00am calculus I as 2.135, with a standard deviation of 1.521, and the mean grade point score for all other calculus I times, classified as “not 8:00am” as 2.208, with a standard deviation of 1.427.

Note that 8:00am courses contained much more extreme values (As and Fs) than the other course times, among a few other more obvious visual patterns. To succinctly show the rates certain grades were earned, **Table 1** is organized by gender and **Table 2** is organized by selected undergraduate major, both over the five years of data collection.

Performance score is a categorical variable. To compare the difference in means between the populations, we decided to begin by performing a two-proportion hypothesis test to compare student performance between 8:00am and all other classes (combined), as well as spring and fall enrollments. To explore the first research question, a two-proportion, one-tailed z-test model was used (meant for independent populations). Our alpha value to determine our confidence interval was adjusted based on appropriate calculations and will be discussed in the results section.

Next, we utilized a multivariable linear regression model to determine predictor variables toward students’ calculus I performance. Unfortunately, the inclusion of time dominated the model, and significance of any variables in terms of direct effect on the outcome was not apparent at our desired levels with its inclusion. This prompted us to look at the interaction effect of time on the different variables through logistical regression. Finally, we later controlled for time so that the underlying significance could be more evident. This information, including the detailed tests and resulting data, will be described in the next section.

DATA ANALYSIS

Two-Proportion Testing

With the sample size n=717, the most appropriate split in the population was a sample for 8:00am class start times (n_1=327) and all other start times, including 11:00am, 12:40pm, 1:10pm, and 2:10pm (n_2=390). These populations are independent among the five years of data collected. The proportions are as follows: p_1=209/327=0.639, p_2=268/390=0.687, p=477/717=0.665, where p_1 is 8:00am passing rate, p_2 was all other times combined into one passing rate, and p was overall five-year passing rate (all values which are necessary for effective two-proportion testing). Our hypothesis (H_A), based on prior experience and these initial calculations, was that the 8:00am passing rate would be lower than the rest ($p_1<p_2$). With a 10% significance level (α=.10, CI=90%), we were able to *reject* the null that there was no significant difference ($p_1=p_2$) between passing rates between 8:00am and all other course start times, concluding that 8:00am passing rates were statistically significantly worse than these other course start times. **Table 3** shows summarizes this information concisely.

When using a similar statistical model for fall calculus I courses (n_1=495) and spring calculus I courses (n_2=222), also independent populations, our proportions were as follows: p_1=328/495=0.663, p_2=149/222=0.671, p=477/717=0.665, where p_1 was fall passing rate, p_2 was spring passing rate, and p was overall five-year ABC rate (obviously, the same proportional value as earlier). Our hypothesis (H_A), based on experience and these initial calculations, was that the fall pass rate would be consistently lower than the spring passing rate ($p_1<p_2$). We were able to *reject* the null, this time at a 1% significance level (α=.01, CI=99%) that there

Table 4. Multivariable regression, all variables included

Coefficients	Estimate	Standard error	z value	Pr(>|z|)
(Intercept)	-0.20288	0.34962	-0.580	0.561724
Time	0.01332	0.08854	0.150	0.880443
Major	0.18075	0.05087	3.553	**0.000381**
Gender	-0.32095	0.18660	-1.720	**0.085436**
Instructor	0.21910	0.04448	4.926	**8.39e-07**
Term	-0.12876	0.18196	-0.708	0.479181

Table 5. Interaction effect with time and major

Coefficients	Estimate	Standard error	z value	Pr(>|z|)
(Intercept)	0.138003	0.233517	0.591	0.5545
Time	0.008878	0.171363	0.052	0.9587
Major	0.134000	0.065486	2.046	**0.0407**
Time:Major	0.046777	0.049147	0.952	0.3412

was no significant difference between passing rates between spring and fall passing rates, concluding that fall term passing rates were statistically significantly worse than spring term passing rates. Again, **Table 3** summarizes these results.

Main Effects from the Linear Regression Analysis

While two-proportion tests revealed the importance of time (and also term) in course performance, we wanted to model the impact of predictor variables, particularly the variable time, on passing rates in calculus I at this institution, as per our second research question. The data was coded by increasing units representing later times in the day, and our codes for some other variables needed a ranking system as well: we decided on decreasing DFW rates for major and instructor. For context, we ordered majors as education, computer science, mathematics, other, engineering, and natural sciences, with the highest calculus I DFW rates for the education majors, and lowest for natural science majors; this information is again available in **Table 2**.

We also wanted to return to the idea of multivariable regression models but knew that time was distorting the significance of any other values. To control for time, we separated the data into 8:00am courses and all other start times before then running the regression model (separated as **Table 1** had described). Additionally, our proportion tests showed there was significance when comparing times in this way with our sample sizes closer in size under this categorization. This required 15 logistic regression models for 8:00am, and 15 more for all other times combined. The 15 models were all possible combinations of variables from the full list of n=4 (major, gender, instructor, term), with k=1, 2, 3, and 4, where k is the number of explanatory variables in each regression experiment; this accounted for 30 logistic models in total. When controlling for time and running a regression model, 8:00am course times showed that major and instructor were significant in every test that included those variables (each present eight times in the models), while gender and term were never significant (each present eight times in the models). Alternatively, non-8:00am course times showed that major, instructor and gender were all significant in every test that included those variables (each present eight times in the models), while term was never significant (present eight times in the models).

After running these tests and observing the results, we decided to run the full test as shown, where p/f represents our pass/fail as the dependent variable (output) for all tested variables in the model, and t=time, m=major, g=gender, s=term, and i=instructor:

$$p/f \sim t+m+g+s+i.$$

Table 4 summarizes the results of this model, with significance (p-values) bolded in the last column, where major (p=0.000), gender (p=0.085), and instructor (p=0.000) were found to be significant. Values were found through inputting interaction effects and the results will be shown in the next few tables. This supported some trends viewed in earlier tests we performed with combinations of variables while controlling for term and finding that these three variables still showed significance under different circumstances. The interpretation of the "estimate" values from **Table 4** are as follows, using our model: for every one unit change in major (that is, a shift in decreasing ranking of DFW rate), the probability of passing the course increases by 18%, at a very high significance level; for every one unit change in gender (that is, a shift from female to male), the probability of passing the course decreases by 32%, with a fairly high significance level; for every one unit change in instructor (that is, a shift in decreasing ranking of DFW rate), the probability of passing the course increases by 22%, at a very high significance level. These significance levels are bolded in **Table 4**.

Interaction Effects from the Logistic Regression Analysis

Even though these findings were important, time was still at the forefront of our study. With practically no probability change in the time variable, and no significance on that variable itself, the overall impact of time on passing rates in calculus I was still unclear. However, our prior knowledge from the two-proportion tests suggested that course start time did indeed have value (showing the significance 8:00am vs not-8:00am). Thus, we decided to next look for an interaction effect through

$$p/f \sim t^*x_n,$$

where x_n is the inclusion of another distinct independent variable. This required four trials to run; time interacting with major, gender, instructor, and term. We summarize the outputs and results of interaction effect (with significance bolded), keeping in mind that each row represents the coefficients in the following equation:

$$\hat{y}=b_0+b_1x_1+b_2x_2+b_3x_1x_2.$$

Table 6. Interaction effect with time and gender

Coefficients	Estimate	Standard error	z value	Pr(>|z|)
(Intercept)	0.7498	0.1965	3.815	**0.000136**
Time	0.3028	0.1643	1.842	**0.065407**
Gender	-0.2583	0.2314	-1.116	0.264323
Time:Gender	-0.2098	0.1883	-1.114	0.265195

Table 7. Interaction effect with time and instructor

Coefficients	Estimate	Standard error	z value	Pr(>|z|)
(Intercept)	0.052922	0.190600	0.278	0.781274
Time	-0.023891	0.193839	-0.123	0.901906
Instructor	0.195530	0.053833	3.632	**0.000281**
Time:Instructor	0.007895	0.040710	0.194	0.846223

Table 8. Interaction effect with time and term

Coefficients	Estimate	Standard error	z value	Pr(>|z|)
(Intercept)	0.584134	0.319249	1.830	**0.0673**
Time	0.157838	0.265338	0.595	0.5519
Term	-0.013788	0.224476	-0.061	0.9510
Time:Term	-0.007667	0.161738	-0.047	0.9622

None of the models showed a statistical significance in the interaction term (final row in each table, labeled as "time:variable"), but the time and gender isolated interaction were the most significant of the four cases. However, isolating each variable, paired with time, did yield some interesting results. According to **Table 5**, based on the estimate column, at a 5% significance level, we can use the coefficient estimate of 0.138003 to describe an increase in probability of passing the course by 13% as major changes, when the model only involved major and time.

According to **Table 6**, based on the estimate column, at a 10% significance level, we can use the coefficient estimate of 0.3028 to describe an increase in probability of passing the course by 30% as time changes, when the model only involved gender and time.

According to **Table 7**, based on the estimate column, at a 1% significance level, we can use the coefficient estimate of 0.195530 to describe an increase in probability of passing the course by 20% (rounded) as instructor changes, when the model only involved instructor and time.

According to **Table 8**, there were no significant interactions when the model only involved term and time. Note that the models for time vs gender, time vs major, and time vs instructor showed as most significant.

CONCLUSIONS, LIMITATIONS, AND THE FUTURE

Limitations and the Future

This study is unique through its context as a college calculus-specific multi-term examination of data. This can be considered as a first step toward more specific research within undergraduate mathematics and more careful consideration toward predictor variables for student success. While we uncovered important conclusions from this data analysis, we cannot claim that these results would work in other content courses, mathematics courses, or other universities/colleges. These are simply findings at a particular four-year school for calculus I, broken down in order to analyze a variety of possible predictors of passing rates more deeply. Additionally, this is only data from a university in the United States, and results may vary in other portions of the country or the world. There could be other underlying influences beneath these factors including student course backgrounds. Further data for student profiles, including previous mathematical preparation, academic motivation, and college experience was unavailable due to university policy and Family Educational Rights and Privacy Act (FERPA) anonymity concerns. Furthermore, under the assumption that 8:00am start times are unpopular with most students, there could be a complication in determining which students are registering for the variety of course times. We also do not consider student performance in major-related courses outside of calculus I when ranking majors by passing rate. Similarly, we cannot claim to know the causes of fall term calculus I courses showing lower performance than spring term calculus I courses, instead we argue that a correlation is apparent.

We advise faculty from other institutions to replicate these strategies in their own departments. Further studies could be done to determine outside potential factors, such as attendance, homework habits, number of first-year students present in the classes, adequate time to adjust to the college lifestyle, and the expectations of students who have possibly not completed enough prerequisites to fit in the calculus I course. Also, we are only including students who finished calculus I (i.e., the Withdrawals in standard DFW rates), not those that withdrew, which admittedly includes students that may stubbornly push to the end of the term despite tremendously low odds of passing the course. Additionally, a future extension of this study could test the pairs of variables with time, but without implied interaction effects. These methods could also be used for other "gateway courses" considered in other undergraduate fields of study to discover similar trends and conclusions.

Overall Conclusions

Our results show that in early 8:00am calculus I courses, students perform worse than in later calculus I courses (defined as lower passing rates). This supports Carrell et al. (2011) in that post-secondary early course start times impact performance negatively and can reinforce the findings through a mathematics-specific context. While the term (spring or fall) was not a significant predictor in our projection models, through two-proportion testing we validated the claim that spring students perform better (defined as higher passing rates) in calculus I than fall students. This supports a clear indication that *when* a student takes calculus I (course time and academic term) can affect student performance, which differs from the mixed conclusions that the literature provided. Perhaps by taking calculus I in the spring, these students were given time to prepare more adequately with appropriate prerequisites. Time and gender appeared related: at first, all genders found more difficulty in early 8:00am calculus I courses, but as the classes began later in the day, the gap widened between females and males, favoring females as course times progressed. While our sample numbers for male and female varied, our evidence was very clear and supported Bridgeman and Lewis (1996). Instructor pass rates varied considerably, complementing the study from Webb (2016), and thus showed up as a very significant factor in calculating passing probabilities. Interaction effects between time and each variable were not significant on their own, but our additional analysis showed that when these variables were isolated, more impact was visible through statistical analysis.

Author contributions: All authors have sufficiently contributed to the study, and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Aiken-Wisniewski, S. A., Smith, J. S., & Troxel, W. G. (2010). Expanding research in academic advising: Methodological strategies to engage advisors in research. *NACADA Journal, 30*(1), 4-13. https://doi.org/10.12930/0271-9517-30.1.4

Akbuga, E. (2018). *Motivation intervention through calculus tasks with science and engineering applications.* https://digital.library.txstate.edu/handle/10877/7413

Akbuga, E., Hurdle, Z., Daniel, S., & Laffey, R. (2019). Using calculus writing assignments to foster student motivation. *Mathematics in School, 47*(4), 37-39.

Berry, J. S., & Nyman, M. A. (2003). Promoting students' graphical understanding of the calculus. *The Journal of Mathematical Behavior, 22*(4), 479-495. https://doi.org/10.1016/j.jmathb.2003.09.006

Bloemer, W., Day, S., & Swan, K. (2017). Gap analysis: An innovative look at gateway courses and student retention. *Online Learning, 21*(3), 5-14. https://doi.org/10.24059/olj.v21i3.1233

Bressoud, D. M., Carlson, M. P., Mesa, V., & Rasmussen, C. (2013). The calculus student: Insights from the Mathematical Association of America national study. *International Journal of Mathematical Education in Science and Technology, 44*(5), 685-698. https://doi.org/10.1080/0020739X.2013.798874

Bressoud, D. M., Mesa, V., & Rasmussen, C. L. (2015). *Insights and recommendations from the MAA national study of college calculus.* https://doi.org/10.5951/mathteacher.109.3.0178

Bridgeman, B., & Lewis, C. (1996). Gender differences in college mathematics grades and SAT-M scores: A reanalysis of Weiner and Steinberg. *Journal of Educational Measurement*, *33*(3), 257-270. https://doi.org/10.1111/j.1745-3984.1996.tb00492.x

Carrell, S. E., Maghakian, T., & West, J. E. (2011). A's from Zzzz's? The causal effect of school start time on the academic achievement of adolescents. *American Economic Journal: Economic Policy, 3*(3), 62-81. https://doi.org/10.1257/pol.3.3.62

Ellis, J., Fosdick, B. K., & Rasmussen, C. (2016). Women 1.5 more likely to leave STEM pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit. *PLOS One, 11*(7), 1-14. https://doi.org/10.1371/journal.pone.0157447

Eskew, R. W. (2013). *Trends in grade distributions at Hilbert College: Report to the Provost's Council*. https://www.hilbert.edu/docs/cdefault-source/Academics/Institutional-Research-Assessment/oira_report-on-grade-distributions_2009_12.pdf?sfvrsn=8e29bc3f_0

Hagman, J. E., Johnson, E., & Fosdick, B. K. (2017). Factors contributing to students and instructors experiencing a lack of time in college calculus. *International Journal of STEM Education, 4*(12), 1-15. https://doi.org/10.1186/s40594-017-0070-7

Hall, M., Smith, K., Boeckman, D., Ramachandra, V., & Jasin, J. (2003). Why do students withdraw from courses? In *Proceedings of Southern Association for Institutional Research Conference*. San Antonio, TX, USA.

Hurdle, Z. & Mogilski, W. (2022). The impact of prerequisites for undergraduate Calculus I performance. *International Electronic Journal of Mathematics Education, 17*(3), em0696. https://doi.org/10.29333/iejme/12146

Koch, A. K., & Pristilli, M. D. (2014). *Analystics and gateway courses: Understanding and overcoming roadblocks to college completion*. John Gardner Institute.

Onyper, S. V., Thacher, P. V., Gilbert, J. W., & Gradess, S. G. (2012). Class start times, sleep, and academic performance in college: A path analysis. *Chronobiology International, 29*(3), 318-335. https://doi.org/10.3109/07420528.2012.655868

Pope, N. G. (2016). How the time of day affects productivity: Evidence from school schedules. *The Review of Economics and Statistics, 98*(1), 1-11. https://doi.org/10.1162/REST_a_00525

Wahistrom, K. (2002). Changing times: Findings from the first longitudinal study of later high school start times. *NASSP Bulletin, 86*(633), 3-21. https://doi.org/10.1177/019263650208663302

Webb, D. (2016). Appling principles for active learning to promote student engagement in undergraduate calculus. In *Proceedings of the 13th International Congress on Mathematical Education*.Hamburg, Germany.

Wheaton, A. G., Chapman, D. P., & Croft, J. B. (2016). School start times, sleep, behavioral, health, and academic outcomes: A review of the literature. *The Journal of School Health, 86*(5), 363-381. https://doi.org/10.1111/josh.12388

White, P., & Mitchelmore, M. (1996). Conceptual knowledge in introductory calculus. *Journal for Research in Mathematics Education, 27*(1), 79-95. https://doi.org/10.2307/749199

Wu, X., Deshler, J., & Fuller, E. (2018). The effects of different versions of a gateway STEM course on student attitudes and beliefs. *International Journal of STEM Education, 5*(1), 44. https://doi.org/10.1186/s40594-018-0141-4

Young-Jones, A. D., Burt, T. D., Dixon, S., & Hawthorne, M. J. (2013). Academic advising: Does it really impact student success? *Quality Assurance in Education, 21*(1), 7-19. https://doi.org/10.1108/09684881311293034