Simulating Student Flow Through a University's General Education Curriculum

Robert Saltzman*
San Francisco State University, San Francisco, California, USA

Stewart Liu
San Francisco State University, San Francisco, California, USA

Theresa Roeder
San Francisco State University, San Francisco, California, USA

Many public universities require undergraduate students to take a substantial number of general education courses distributed among various areas of knowledge. This article describes a computer model that simulates the flow of undergraduates through the lower division portion of the general education curriculum at San Francisco State University, one of the 23 campuses of the California State University system. The model permits various changes in the general education curriculum, such as course sequencing, pass rates, retention rates, capacities, and enrollment, to be tested for their potential impact on students. Key outcome measures are the elapsed time taken to complete all requirements and the percentage of these requirements that students complete within three years of starting this curriculum. Experiments conducted show that time to complete requirements and percent complete behave as expected to changes in pass rates and capacities. Increased retention rates pose challenges unless they are met with increased capacities.

* Corresponding Author. E-mail address: saltzman@sfsu.edu

I. INTRODUCTION

The California State University (CSU) system educates more than 480,000 students every year and, with 23 campuses and eight off-campus centers, is the largest four-year public university system in the United States (California State University, 2018). Like many other public university systems, the CSU requires undergraduate students to take a substantial number of general education (GE) courses. While a student's major provides in-depth study in one subject area, GE imposes a breadth of study that aims to instill knowledge in students that will serve them for a range of future experiences and provide them with the intellectual agility to move from one career to another (San Francisco State University, 2018). The GE curriculum at all CSU campuses, including San Francisco State University (SFSU), requires students to complete 48 units in 16 courses distributed across various areas of knowledge, such as English language communication, science and quantitative reasoning, arts and humanities, and social sciences (White, 2017).

In September 2016, CSU launched its Graduation Initiative 2025 in order to increase completion rates for all students while closing achievement gaps among low-income and underserved students (California State University, 2018). The initiative's ambitious goals challenge the system to more than double its current 4-year graduation rate for first-time freshmen in nine years, as well as graduate 500,000 additional students by 2025 – meeting the CSU's share of the state's projected degree shortfall. For SFSU, in particular, the goal is to raise its 4-year graduation rate from 22% for the 2013 cohort to 33% for the 2021 cohort (San Francisco State University, 2017). Consequently, there is keen interest on campus to analyze our curriculum (and the student experience more broadly) and identify barriers to student success (Altura et al., 2018).

However, before possibly altering policies and allocating resources to try and achieve the target 4-year graduation rate, it would be helpful to first understand more clearly how any actions taken might affect student flow through the curriculum. To do so, we developed a discrete-event computer simulation model that mimics how undergraduates move through the lower division portion of GE, which consists of 13 requirements spread across 5 major areas (see Fig. 1). We chose to focus on the GE curriculum as it is the major, common focus of incoming freshmen and can have a significant effect on student success. Further, it is something that is more easily controlled at the university level than are curricula in the major. Our modeling effort was limited to the lower division portion of the GE curriculum because upper division requirements may be satisfied with a very large number of courses and do not seem to be constraining student progress toward graduation.

By design, SFSU's GE curriculum is broad, requiring students to take courses in a wide variety of subject areas, with few embedded prerequisites; *e.g.,* while all Area E courses require students to have first completed a Written English course in subarea A2, only a small fraction of humanities courses in subarea C2 have a similar requirement. Meanwhile, Lab Science courses in subarea B3 may be taken concurrently with a Physical or Life Science class. A more thorough description of SFSU's GE program can be found in the university's on-line bulletin (San Francisco State University, 2018). The purpose of our modeling effort is to provide an objective framework for assessing the impact of various changes to the GE program. In particular, this paper describes using the model to investigate how performance measures, such as the elapsed time to complete GE requirements, might change if course pass rates and year-to-year retention rates were improved. It also reports on the potential impact of increasing the GE program's capacity to seat more students and of changes in the number of newly enrolled students.

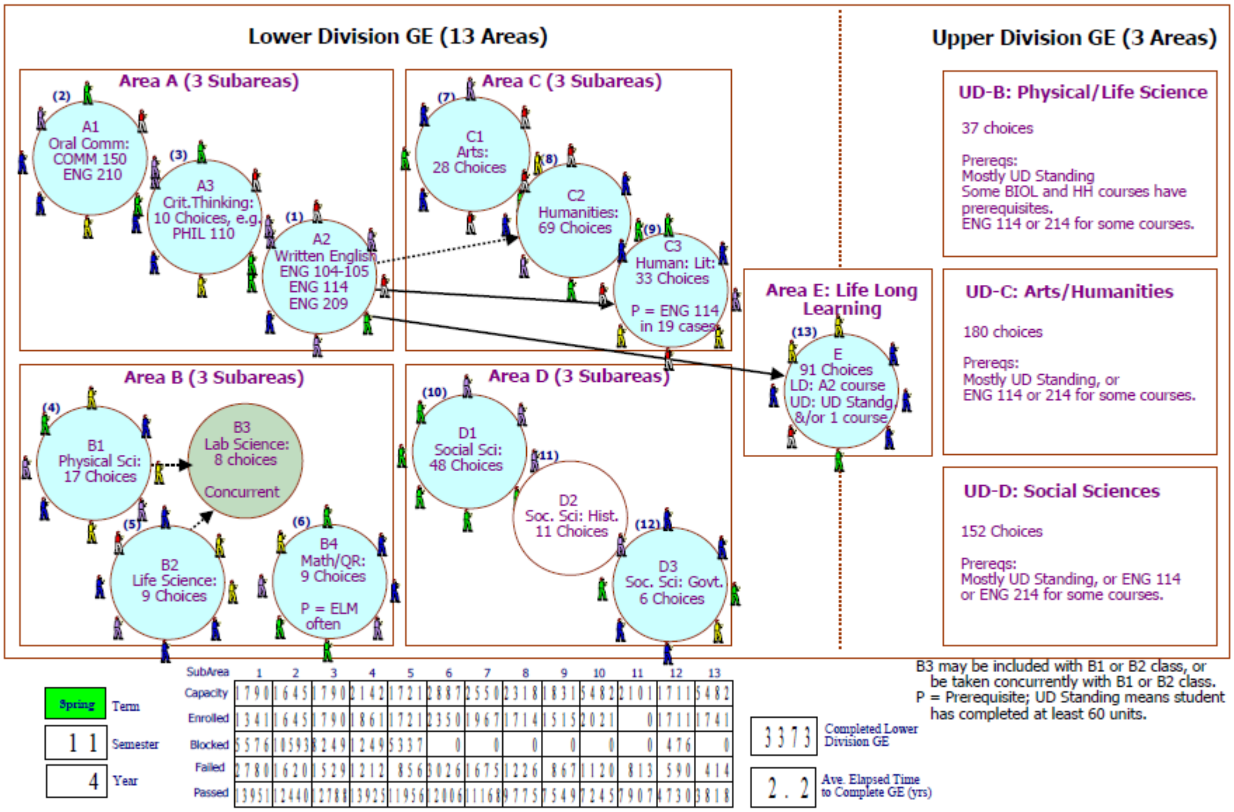

FIGURE 1. SNAPSHOT OF THE ANIMATED PORTION OF THE LODGE SIMULATION MODEL.

II. LITERATURE REVIEW

Although operations research techniques have been applied to educational issues for more than 50 years, this review will limit itself to approaches taken to address planning issues in higher education. These techniques generally fall into four major categories: statistical methods, mathematical programming, Markov chain models, and computer simulation models. Statistical methods, including machine learning algorithms, have been applied to student experiences in many settings, including student flow, retention rates, and graduation rates (Willett, 1982; Simpson, 1987; Knight, 2002). Both descriptive and explanatory models have been created, and can be used to evaluate the impact of changes in the independent variables. For example, Knight (2004) developed a multiple regression model to gauge the impact of dozens of student-behavioral factors (*e.g.,* academic preparation, college grade point average (GPA), financial need) on elapsed time to degree. One difficulty in doing so is the collinearity of these predictors. Because student behavior is complex, many factors (nonlinearly) interact and affect the outcomes being studied. Goenner and Snaith (2004) also developed a multiple regression model to predict graduation rates at R1 institutions, including institutional characteristics (*e.g.*, student-faculty ratio, weighted tuition rates) in their model. They found that student and institutional factors both play a role in graduation rates, but, as in Knight (2004), that it is challenging to explicitly characterize the different aspects of these factors. In this sense, simulation modeling is better suited for prediction because it captures these behaviors more readily.

Arsad, Buniyamin, and Ab Manan (2014) compared the performance of neural network and linear regression models in predicting the final cumulative GPAs of electrical engineering students at a Malaysian university. They found that neural networks have higher correlations and lower MSE values than regression models, and that student grades in foundational engineering courses are strong predictors for final GPAs. Slim et al. (2014) used a Bayesian Belief Network to accurately predict student GPA starting early in a student's academic career, given the student's performance-to-date. The networks in this paper represent the curriculum map of required courses and their prerequisite structure. By contrast, Slim et al. (2016) used Sequential Pattern Mapping to identify the sequence of courses taken by "successful" (*i.e.*, high-GPA) students to compare it to the order in which less successful students progress through the curriculum. Their analysis showed a dramatic difference between the sequences and even the courses taken by the two groups. Ojha (2017) and Ojha et al. (2017) used a variety of statistical learning models to predict delay in graduation based on an increasing number of input variables. Their initial models incorporated pre-university and demographic information on students, while subsequent models added data on student performance up to their second and third years. The accuracies of the models were in the 67-76% range.

Mathematical programming models have also been widely applied to analyze enrollment, student success, and teacher-student ratios. Correa (1967) surveyed numerous models that form the basis of these analyses. The number of enrolling students is typically calculated based on factors such as the number of new enrollments, repeats, dropouts, re-enrollments, graduates, and deaths. In particular, Gani (1963) showed that a strong correlation exists between student enrollment and student pass and repeat rates, emphasizing the importance of collecting student data. However, mathematical models are typically isolated from other social aspects that influence the overall supply of students. Stone (1965) used an economic input-output model to analyze demographic changes over time as an indicator for the supply of future students. The model used pre-determined parameters for population growth based on birth rates but omitted factors such as immigration. It also made other simplifying assumptions to ensure linearity, including a pre-defined constant student pass rate and limiting the educational curriculum to one sequence. Using another economic input-output model, Stone (1966) found that student pass rates are often not constant or monotonic over time. In addition, the level of educational penetration varies widely across different social classes.

Non-linear programming models have also been applied to analyze the relationship between student enrollment, student success, and teaching staffs. Correa (1967) presented a model that assumes a positive correlation between successful students and the supply of good teachers. This model shows the positive evolutionary cycle of the educational system and demonstrates the importance of improving educational systems in general. Oliver and Hopkins (1972) presented a network flow model at equilibrium for estimating student enrollments of different groups. Their model assumed a linear relationship between student enrollment and the number of various types of teaching staff (tenure track faculty, part time faculty, and teaching assistants) but did not consider changes in the size of the teaching staff. It accurately predicted student enrollment at UC Berkeley.

Educational systems have been viewed by some as Markov processes in which students move through an ordered set of states based on a transition matrix holding the probabilities of moving from state to state (Johnes 2015). Markov models can be straightforward to set up and solve, allowing the steady-state probability of being in any state after many transitions to be found. One such study by Bessent and Bessent (1980) analyzed the progression of doctoral students through their degree program. Another by Kwak et al. (1986) extended this approach to a trimester-based institution and was able to fairly accurately forecast departmental enrollments and graduation rates. Shah and Burke (1999) used a Markov model to estimate the average time needed by undergraduates to complete their degrees while accounting for each student's course of study, gender, and age when starting the degree program. Nicholls (2009) built a small, 10-state Markov model of students moving through a part-time doctoral program in order to forecast the number of students in each year of the program; these forecasts, in turn, were used to estimate the program's expected revenue and supervisor workload.

It should be noted that Markov chain models require the system under study to possess both the Markovian property and stationary transition probabilities (Hillier and Lieberman 1990). Many systems, unfortunately, lack one or both of these conditions. At universities, transition probabilities are not stationary because both the number of seats made available and the number of students competing for these seats vary from term to term. Simulation models, by contrast, require neither of these conditions and can handle subtle aspects of a process, such as the prioritization of various groups of students when they register for courses. Because of its flexibility, simulation's advantage over other analytical methods tends to increase as system complexity increases.

One effort applying simulation to an educational planning setting was reported by Plotnicki and Garfinkel (1986) who used simulation to create course schedules that might allow the largest number of students to flow smoothly through their university's MBA curriculum while maintaining feasibility for the departments who offer the courses. Although our goal is also to understand and facilitate student flow, we take the university's extensive GE class schedule (built from myriad departments' independently developed schedules) as a fixed input to the model. Webster (1997) describes a spreadsheet simulation model used to analyze costs in Papua New Guinea's educational system, and make funding decisions for the whole country.

Mansmann and Scholl (2007) built a decision support system for a flexible curriculum in Germany that allows users to input modifications to the curriculum and then simulate the impact on resource requirements. Simulation experiments undertaken by Schellekens et al. (2010) tested the effectiveness of a redesigned program at a Dutch university that allows students to take any course at virtually any time. Saltzman and Roeder (2012) constructed a discrete event simulation model of undergraduate student flow that was used to evaluate potential changes in curriculum policy, prerequisite structure, and staffing capacity decisions at a large public business college facing budget cuts. In a similar vein, Weber (2013) simulated students moving through an undergraduate engineering program in order to test the impact of possible enrollment increases, capacity reductions, and grading option changes. Hickman (2017) developed an open-source software library flexible enough to input any curriculum, compute metrics of its structural complexity, and simulate its impact on students' ability to successfully move through it. Experiments with the library indicated that more complex curricula are negatively correlated with student success.

As described in the next section, the present article adapts the focused business core curriculum framework found in Saltzman and Roeder (2012) to the broader GE curriculum at SFSU. We hope that the results from this effort can be used to inform decision making and resource allocation at the university level regarding GE courses required of all incoming students. To the best of our knowledge, this is the first effort to simulate the flow of thousands of students through a university's GE program.

III. THE LODGE SIMULATION MODEL

SFSU's GE curriculum is comprised of courses in eight areas spanning lower and upper division classes (see Fig. 1). Because there are so many course choices for students to fulfill their upper division requirements in areas UD-B, UD-C, and UD-D, these areas do not create bottlenecks. To focus on what matters most, our model tracks student progress only in (lower-division) areas A–E; consequently, it is called the Lower Division General Education (LODGE) model. The LODGE model was built with Arena 15, a process-oriented simulation package (Kelton, Sadowski and Zupick, 2015). Students are the model's main entities moving through the process. GE (sub)areas are modeled as resources with large but finite capacities, so students may be blocked from taking courses in a particular subarea in a given term. When the model runs, students can be seen moving through the various subareas, as in Fig. 1. However, since thousands of students move through the system, only a small fraction of them are shown on screen. Meanwhile, counters dynamically update the number of students enrolled in each subarea, as well as the cumulative number who have passed, failed, and been blocked in each subarea. The lower right part of Fig. 1 also displays the current number of students who have completed all lower division GE courses and the average elapsed time needed to do so.

The model's structure is essentially the same as that of the business undergraduate flow (BUF) model of Saltzman and Roeder (2012); a detailed description of the BUF model can be found in that article, but its modified operation to fit the LODGE model can be summarized as follows. At the start of each term, new First-Time Freshman (FTF) students are created and subarea capacities are reset. New and continuing students are assigned updated priority numbers for the upcoming course registration based on how long they have been at SFSU. In the registration process, the model first determines how many GE courses the student wishes to register for by sampling from the appropriate GE course load distribution (Table 2); it then tries to enroll students in those GE courses. During registration, students may be blocked from enrolling in a subarea that is already filled, in which case the student attempts to enroll in a course in the next available subarea. Students then take all the courses in which they are enrolled. At the end of the term, students either pass or fail each course taken (based on input subarea pass rates). A large array is employed to track every student's success or failure in each subarea.

At the end of each term, the model records information such as the time to complete GE requirements for students who have passed courses in all GE subareas. At the end of each academic year, some students randomly drop out of the university based on input retention rates. They are counted and then removed from the model. Continuing students return to the registration station for the next term. After 45 terms (*i.e.,* 15 years) have been simulated, the model reaches the end of the run and writes out key summary statistics.

From a process perspective, the BUF and LODGE models are quite similar, *e.g.,* both simulate courses being taken by students in three terms per academic year (fall, spring, and summer), and neither tracks detailed student characteristics such as GPA. However, some changes had to be made to the BUF model as it was adapted to the flow of students through the lower division requirements of GE, as delineated in Table 1. Perhaps the most important change is that the LODGE model simulates activity not at the individual course level but at the subarea level, which is comprised of multiple courses. We felt that explicitly representing each of the hundreds of GE courses in the simulation model would prove too difficult and time consuming, without adding significant insights.

The LODGE model makes three other important structural assumptions. First, it only tracks students who arrive as First Time Freshmen, and assumes they have not completed any GE requirements prior to arrival. While some upper division transfer (UDT) students may still need a few GE courses, this number is assumed to be negligible. Second, the model does not represent subarea B3 because many lab science courses are actually part of a 4-unit science course in subareas B1 and B2, while others exist as stand-alone 1-unit courses. Third, when registering for classes, freshmen are actually given highest priority (once special case students have been accommodated), followed by seniors, juniors, and sophomores. The model approximates this priority scheme by assigning each student a priority number at the start of each term equal to $1 \cdot (t \leq 3) + 2 \cdot (t \geq 10) + 3 \cdot (7 \leq t \leq 9) + 4 \cdot (4 \leq t \leq 6)$, where t indicates how many terms have started since the student arrived on campus, and the mutually exclusive logical expressions in parentheses evaluate to either one (if true) or zero (if false). The registration queue processes students with lower priority numbers before those with higher numbers.

TABLE 1. RELATIONSHIP BETWEEN THE BUF AND LODGE MODELS.

	BUF Model	LODGE Model
Entities	FTF and UDT students	FTF students only
Resources	19 business core courses	13 GE areas and subareas, each comprised of multiple courses
Prerequisite Structure	Extensive, with up to six prerequisites per course	Minimal, with only area E and subarea C3 having one prerequisite each
Key Input Data	Core course load distribution, by term; Capacity by course, year & term; Incoming FTF, UDT, by year & term; Pass rates, by course; and Retention rates, by year	GE course load dist., by year & term; Capacity by subarea, year & term; Incoming FTF, by year & term; Pass rates, by subarea; and Retention rates, by year
Performance Measures	ETD: Elapsed time to degree for FTF and UDT; and 6YGR, 4YGR: 6-year and 4-year graduation rates for FTF and UDT	ETC: Elapsed time to complete lower division GE requirements; and PctComp: Percent of GE requirements completed by students within 3 years

Gathering input data for the LODGE model was challenging because SFSU does not report data aggregated by GE area. The following describes the model's five major inputs and assumptions related to them.

1. The number of *incoming FTF* for terms from Fall 2014–Fall 2018 come from the

actual number of incoming FTF found on SFSU's Office of Institutional Research website (ir.sfsu.edu). For each term after Fall 2018, the number of incoming FTF was taken to be the moving average of the values for that term in the two preceding years.

2. *Subarea capacities* for terms from Fall 2016–Fall 2018 were found by aggregating actual enrollment data (found at ir.sfsu.edu) from all courses within each subarea. However, summer enrollments by subarea are not available, so they were estimated by multiplying the fall plus spring enrollments in each subarea by the summer's total enrollment relative to that of the preceding fall and spring semesters' enrollments. Capacities for terms prior to Fall 2016 are assumed to be the same as those of the corresponding term during AY 2016-17; those in terms after Fall 2018 were taken to be the moving average of the two preceding years' values. Enrollments (seats occupied) are used for capacities rather than seats offered because we want model behavior to be consistent with what actually happened in the past.

3. Student *GE course loads* were estimated by examining the transcripts of a random sample of 100 students who began as FTF. The number of GE classes taken per term in each of their first three years was recorded and analyzed to find a GE course load distribution by student year and term (see Table 2). The distributions in the fall and spring of each year were similar to one another, and very different from those in the summer. Students in the system after three years are assumed to take on a GE course load from the third year distributions.

TABLE 2. GE COURSE LOAD DISTRIBUTIONS AT SFSU BY STUDENT YEAR AND TERM.

Student Year and Term	Number of GE courses taken by a student						Mean
	0	*1*	*2*	*3*	*4*	*5*	
Year 1: Fall & Spring (%)	0.5	3.5	15.0	34.5	31.5	15.0	3.38
Year 1: Summer (%)	82.8	13.1	3.0	1.0			0.22
Year 2: Fall & Spring (%)	5.7	7.7	29.4	29.9	21.1	6.2	2.72
Year 2: Summer (%)	95.5	2.2	1.1	1.1			0.08
Year 3+: Fall & Spring (%)	21.6	34.6	22.8	16.0	4.3	0.6	1.49
Year 3+: Summer (%)	94.3	5.7					0.06

4. *Subarea pass rates* are based on aggregated numbers of students who passed courses within each subarea during Fall 2016, the most recent term for which data are available at ir.sfsu.edu.
5. *Student retention rates* applied in the LODGE model at the end of each of the first three years come from actual first-, second-, and third-year retention rates (80.1%, 87.1%, and 94.7%, respectively) for the Fall 2011 entering cohort, the most recent cohort for which data are available at ir.sfsu.edu. Retention rates applied after the third year are assumed to remain constant at 94.7%.

IV. EXPERIMENTS WITH THE LODGE MODEL

Each replication of the LODGE model simulates 15 academic years, from Fall 2014 through Summer 2029, with historical input data driving the model for the first four and a half years of each replication.

To avoid start-up bias as well as end-of-run effects (from students who enter SFSU but do not finish the GE curriculum by the end of year 15), output below reflects system performance from years 8-10 only. Key performance measures are the percentage of GE requirements that students complete in three years (PctComp) and the elapsed time (in years) taken by students to complete all GE requirements (ETC). Model results given below represent means across five replications.

Unfortunately, SFSU does not report any performance data about the GE curriculum, so the only model validation we could conduct was to compare model output to statistics from the 100 sampled transcripts used to estimate the model's GE course load distributions. For each of the two performance measures of interest, Table 3 indicates that there is overlap in the 95% confidence intervals for the population mean generated by the model and by the sample data, implying that model behavior is consistent with data from the real system.

TABLE 3. COMPARING LODGE MODEL PERFORMANCE TO ACTUAL DATA.

	100 Sampled Transcripts		LODGE Model	
Performance Measure	mean	95% CI hw	mean	95% CI hw
Elapsed time to complete GE (years)	2.69	0.19	2.85	0.03
Pct. of GE completed in 3 years (%)	93.3	2.4	91.7	0.3

In the experiments described below, proposed changes to the system were all assumed to start at the beginning of the sixth academic year (Fall 2019). For ease of presentation, model output from all experiments run are given in Table 4 and then discussed in turn in the following subsections. The input parameters PassRateΔ and RetRateΔ shown in Table 4 refer to additive changes in base case area pass rates and year-to-year retention rates, respectively; CapMult and DmdMult to multiplicative changes in base case area capacities and incoming numbers of FTF students, respectively.

4.1. Higher Area Pass Rates

The first experiment examined the potential impact from higher pass rates of students taking courses in all 13 GE subareas. In particular, pass rates in all subareas (which averaged 88.5% across all GE areas in Fall 2016) were increased by one to five percentage points. As seen in Figure 2, both performance measures steadily improve with students passing their courses at higher rates: for each percentage point increase in pass rates, PctComp rises by about 0.62 percentage points, while ETC drops by about 0.06 years. A five percentage point increase in pass rates, for example, could decrease the elapsed time to complete GE by nearly a third of a year (*i.e.,* one term).

TABLE 4. RESULTS FROM ALL LODGE MODEL EXPERIMENTS.

	INPUT PARAMETER				PERFORMANCE MEASURE	
	PassRateΔ	RetRateΔ	CapMult	DmdMult	ETC (yrs)	PctComp
BASE CASE	0	0	1	1	2.85	91.7
EXPMT. 1 (FIG. 2)	1	0	1	1	2.77	92.4
	2	0	1	1	2.70	93.1
	3	0	1	1	2.65	93.8
	4	0	1	1	2.61	94.1
	5	0	1	1	2.54	94.9
EXPMT. 2 (FIG. 3)	0	1	1	1	2.88	91.6
	0	2	1	1	2.89	91.6
	0	3	1	1	2.97	90.9
	0	4	1	1	2.98	91.2
	0	5	1	1	3.10	89.9
EXPMT. 3	0	1	1.02	1	2.82	91.9
	0	2	1.04	1	2.80	92.5
	0	3	1.06	1	2.80	92.6
	0	4	1.08	1	2.82	92.6
	0	5	1.10	1	2.81	93.1
EXPMT. 4 (FIG. 4)	0	0	0	0.90	2.73	93.5
	0	0	0	0.94	2.74	93.0
	0	0	0	0.97	2.78	92.3
	0	0	0	1.00	2.85	91.7
	0	0	0	1.03	2.91	91.1
	0	0	0	1.07	2.99	89.9
	0	0	0	1.10	3.24	87.1

4.2. Higher Retention Rates

The second experiment considered the possible effects of students being retained at higher rates from year to year than is currently the case. Retention rates in the model determine whether or not each student remains in school for another year; they were increased in this experiment by one to five percentage points (see Figure 3). Perhaps surprisingly, both performance measures are slightly degraded by higher retention rates. While retaining a larger proportion of students each year is desirable in many respects, doing so effectively creates more demand for classes and, without additional resources, actually impedes the overall ability of students to progress through the curriculum. Higher retention rates need to be accompanied by greater course capacities in order to maintain the status quo. This can be seen in the third section of Table 4 where a set of five scenarios paired each percentage point increase in the retention rate with a two percentage point increase in the capacity of all GE areas. Doing so enables ETC to remain at its base case value while PctComp inches up modestly.

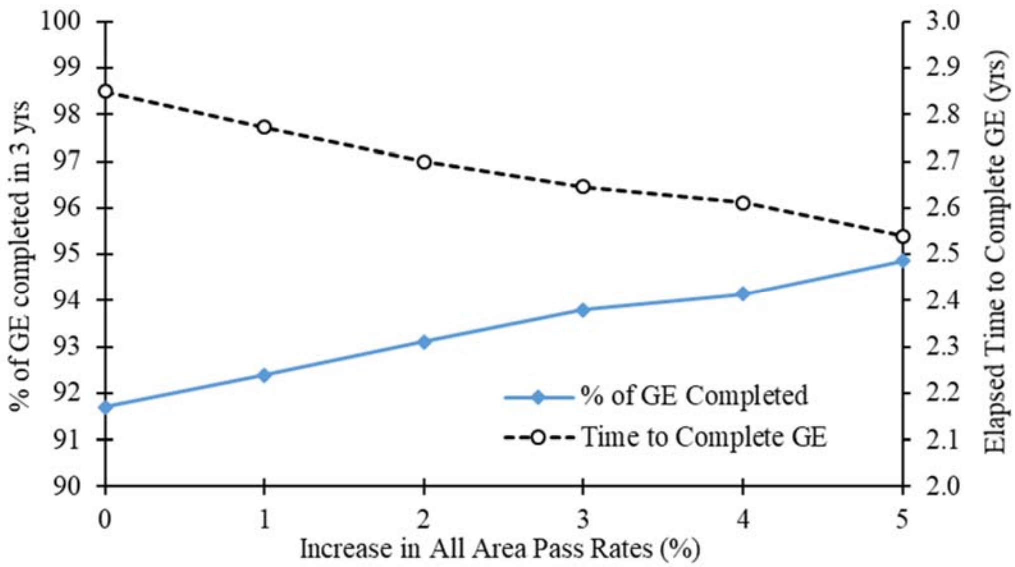

FIGURE 2. IMPACT OF HIGHER PASS RATES.

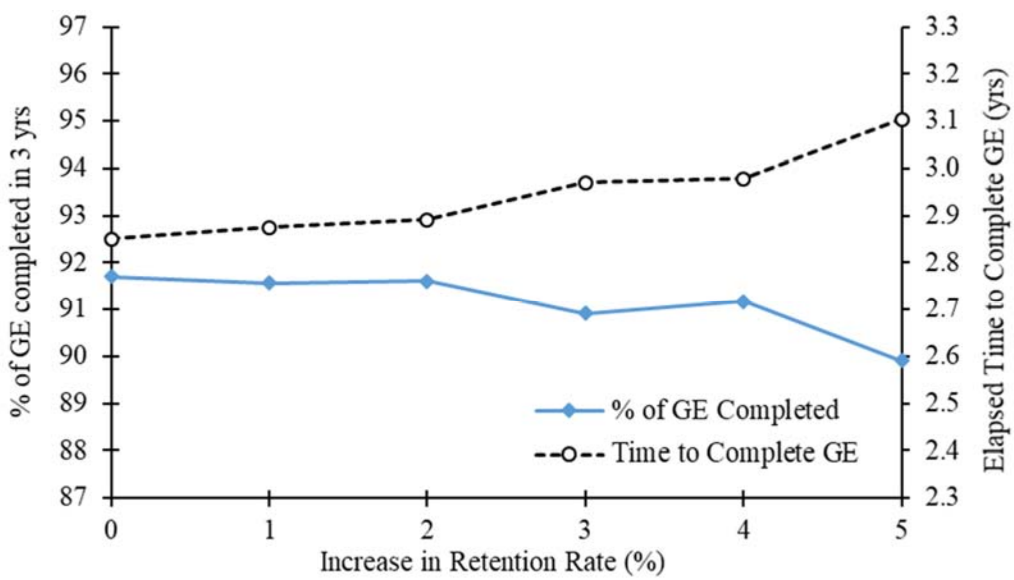

FIGURE 3. IMPACT OF HIGHER RETENTION RATES.

4.3. Changes in Incoming FTF Demand

Finally, the last set of experiments performed with the LODGE model involved altering the demand, *i.e.*, the incoming number of FTF students. The demand multiplier parameter (DmdMult) was changed from 0.90 to 1.10 so that the system was subjected to both decreases from and increases to base case enrollment. Here, the direction of the results was as expected: as demand dropped performance generally improved. More specifically, each percentage point drop in base case demand caused ETC to drop by about four-tenths of a percent and PctComp to rise by two-tenths of a percent. However, as demand increased above its base case value, especially beyond seven percent, performance deteriorated more rapidly, *e.g.,* a 10% increase in demand increased ETC by almost 0.4 years (14%) and reduced PctComp by 4.6 percentage points from their respective base case values.

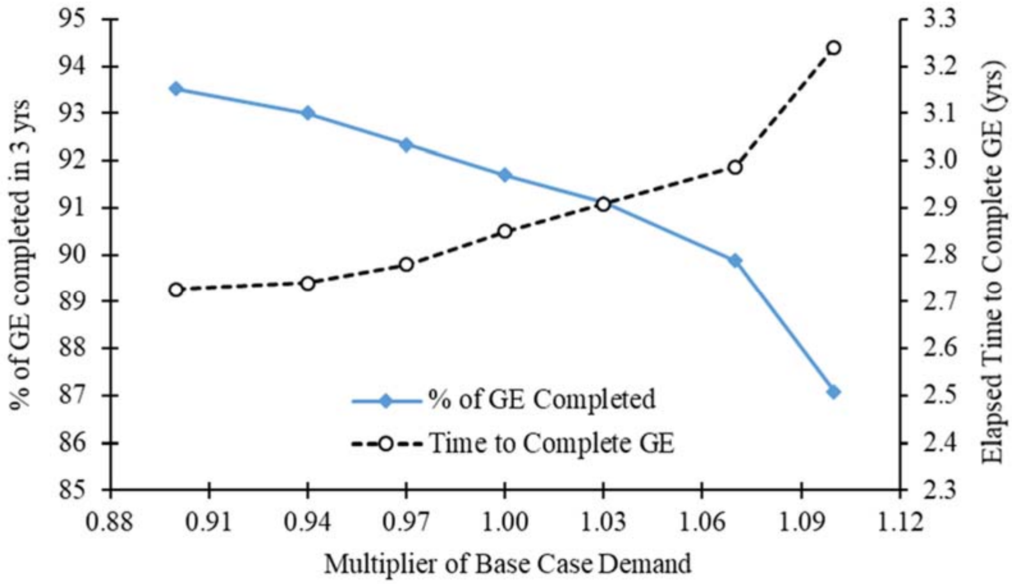

FIGURE 4. IMPACT OF CHANGES IN DEMAND.

V. CONCLUSION

This article has described how a model previously built to analyze the flow of students within a college of business's core curriculum was adapted to simulate the flow of students through a large public university's lower division general education curriculum. The resulting LODGE model was used to assess the potential impact on key performance measures due to changes in pass rates, retention rates, curriculum capacity, and incoming student enrollment. Insights gained from such experiments could prove helpful in deciding how the university should alter its policies and allocate resources. For example, model output indicates that higher course pass rates would reduce the elapsed time to complete the GE program, and thus help lower the university's average elapsed time to degree. Achieving higher course pass rates might require the campus to invest more in tutoring resources for difficult subjects, intervening when students are doing poorly in courses, developing mentoring programs, and so on.

Our experimentation was somewhat limited by access to data that would have made more detailed research possible. For example, it would have been useful to know summer enrollment rates for GE courses. We also would have preferred population data on enrollment patterns to manually-created sample data, which are more prone to error and bias.

The LODGE model could also be used to run other experiments, such as testing the impact of changes in the student GE load distribution (*e.g.,* what if students were advised to take more GE courses in their first two years than they currently do?) and in the university's registration priority scheme. It could even be used to assess the potential impact of changes in the structure of the GE program itself, such as that made recently at SFSU when it removed Written English II (formerly subarea A4) from area A in order to comply with CSU Executive Order 1100 (White, 2017). In the wake of this change, the GE program at SFSU is somewhat in a state of flux, with the restructuring having repercussions on courses in other areas of GE and beyond. Simulating the changes in advance of their implementation might have provided some reassurance that their impact would be positive for students.

REFERENCES

Altura, K., et al., "Student Academic Pathways Project: What Roadblocks and Challenges do Our Students Face?" *San Francisco State University*, http://studentsuccess.sfsu.edu/content/student-academic-pathways-project-sapp (accessed November 14, 2018).

Arsad, P.M., Buniyamin, N. and Ab Manan, J.-L., "Neural Network and Linear Regression Methods for Prediction of Students' Academic Achievement," *2014 IEEE Global Engineering Education Conference (EDUCON),* April 2014, 916–921.

Bessent, E.W. and Bessent A.M., "Student Flow in a University Department: Results of a Markov Analysis," *Interfaces*, 10(2), 1980, 52-59.

California State University, "About the CSU," *California State University*, https://www2.calstate.edu/csu-system/about-the-csu (accessed October 7, 2018).

Correa, H., "A Survey of Mathematical Models in Educational Planning," *Mathematical Models in Educational Planning*. Organisation for Economic Cooperation & Development (ed), Paris, OECD**,** 1967, 21-93.

Gani, J., "Formulae for Projecting Enrolments and Degrees Awarded in Universities," *Journal of the Royal Statistical Society Series A,* 126(3), 1963, 400-409.

Goenner, C.F. and Snaith, S.M., "Predicting Graduation Rates: An Analysis of Student and Institutional Factors at Doctoral Universities," *Journal of College Student Retention: Research, Theory & Practice*, 5(4), 2004, 409–420.

Hickman, M.S., "Development of a Curriculum Analysis and Simulation Library with Applications in Curricular Analytics," Unpublished Master's Thesis, *University of New Mexico*, 2017, http://digitalrepository.unm.edu/ece_etds/388 (accessed October 9, 2018).

Hillier, F.S. and Lieberman, G.J., *Introduction to Stochastics Models in Operations Research*, McGraw-Hill, New York, 1990.

Johnes, J., "Operational Research in Education," *European Journal of Operational Research*, 243(3), 2015, 683-696.

Kelton, W.D., Sadowski, R.P. and Zupick, N.B., *Simulation with Arena 6th ed.*, McGraw-Hill, New York, 2015.

Knight, W.E., "Toward a Comprehensive Model of Influences Upon Time to Bachelor's Degree Attainment," *AIR Professional File* 85, 2002.

Knight, W.E., "Time to Bachelor's Degree Attainment: An Application of Descriptive, Bivariate, and Multiple Regression Techniques," *AIR IR Applications* 2, 2004.

Kwak, N.K., Brown, R. and Schiederjans, M.J., "A Markov Analysis of Estimating Student Enrollment Transition in a Trimester Institution," *Socio-Economic Planning Sciences,* 20(5), 1986, 311-318.

Mansmann, S. and Scholl, M.H., "Decision Support System for Managing Educational Capacity Utilization," *IEEE Transactions on Education*, 50(2), 2007, 143-150.

Nicholls, M., "The Use of Markov Models as an Aid to the Evaluation, Planning and Benchmarking of Doctoral Programs," *Journal of the Operational Research Society* 60(9), 2009, 1183-1190.

Ojha, T., Heileman, G., Martinez-Ramon, M. and Slim, A., "Prediction of Graduation Delay Based on Student Performance," *2017 International Joint Conference on Neural Networks*, 2017, 3454-3460.

Ojha, T., "Prediction of Graduation Delay Based on Student Characteristics and Performance," Unpublished Master's Thesis, *University of New Mexico*, 2017, https://digitalrepository.unm.edu/ece_etds/355 (accessed November 25, 2018).

Oliver, R.M. and Hopkins, D.S.P., "An Equilibrium Flow Model of a University Campus," *Operations Research* 20(2), 1972, 249-264.

Plotnicki, W.J. and Garfinkel, R.S., "Scheduling Academic Courses to Maximize Student Flow: A Simulation Approach," *Socio-Economic Planning Sciences*, 20(4), 1986, 193-199.

Saltzman, R. and Roeder, T., "Simulating Student Flow Through a College of Business for Policy and Structural Change Analysis," *Journal of the Operational Research Society*, 63(4), 2012, 511-523.

San Francisco State University, "Graduation Initiative 2025 Preliminary Progress Update," *San Francisco State University*, https://air.sfsu.edu/sites/default/files/GI2025-Progress-Update-San-Francisco-10-5-2017.pdf (accessed December 19, 2017).

San Francisco State University, "Introduction to General Education," *San Francisco State University*, http://bulletin.sfsu.edu/undergraduate-education/#IGE (accessed October 7, 2018).

Schellekens, A., Paas, F., Verbraeck, A. and van Merriënboer, J.J.G., "Designing a Flexible Approach for Higher Professional Education by Means of Simulation Modelling," *Journal of the Operational Research Society*, 61, 2010, 202-210.

Shah, C. and Burke, G., "An Undergraduate Student Flow Model: Australian Higher Education," *Higher Education*, 37, 1999, 359-375.

Slim, A., Heileman, G.L., Kozlick, J. and Abdallah, C.T., "Predicting Student Success Based on Prior Performance," *2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM),* Dec. 2014, 410–415.

Slim, A., Heileman, G.L., Al-Doroubi, W. and Abdallah, C.T., "The Impact of Course Enrollment Sequences on Student Success," *2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA),* Crans-Montana, 2016, 59-65.

Simpson, W.A., "Tracking Students Through Majors: Methodology and Applications," *The Journal of Higher Education*, 58(3), 1987, 323-343.

Stone, R., "Input-Output and Demographic Accounting: A Tool for Educational Planning," *Minerva* 3(2), 1965, 172-186.

Stone, R., "A Model of the Educational System," *Minerva* 4(3), 1966, 365-380.

Weber, A.C., "Simulating the Flow of Students Through Cal Poly's Undergraduate Industrial Engineering Program for Policy Analysis," Unpublished Senior Project, *California*

Polytechnic State University-San Luis Obispo, 2013, https://digitalcommons.calpoly.edu/imesp/140/ (accessed October 10, 2018).

Webster, T. "Cost Analysis and Its Use in Simulation of Policy Options: The Papua New Guinea Education Finance Model," *International Review of Education*, 43(1), 1997, 5-23.

White, T.P., "General Education Breadth Requirements-Executive Order 1100-Revised August 23, 2017," https://www.calstate.edu/eo/EO-1100-rev-8-23-17.html (accessed November 28, 2018).

Willett, L.H., "Continuing Education Student Flow Analysis," *Research in Higher Education*, 17(2), 1982, 155-164.