The AIR Professional File

Fall 2019 Volume

Supporting quality data and decisions for higher education.

Association for Institutional Research

LETTER FROM THE EDITOR

This volume of the *AIR Professional File* illustrates creative approaches to data analysis in two familiar areas of institutional research – class size and enrollment projection.

In *Perception Isn't Everything: The Reality of Class Size*, Umbricht and Stange challenge traditional measures of class size found in institutional fact books, used by ranking publications, and computed for the Common Data Set. What if, instead, we evaluated class size through the lens of the student experience? The authors explain how to compute a student-centric metric of class size and offer advice for the effective use of this more sophisticated and nuanced measure.

Who among us hasn't been pressed to find a magic formula to accurately forecasts student enrollment? In *Enrollment Projection Using Markov Chains*, Gandy, Crosby, Luna, Kasper and Kendrick show how an underutilized forecasting tool can be enlisted to better understand and predict student flow, even in subgroups of students. With extensive description of the methodology and presentation of results, the authors provide a blueprint for replication at your institution.

Do you have a creative approach to data analysis? Your colleagues are waiting for you to share it through the *AIR Professional File*!

Sincerely,
Sharron Ronco

ISSN 2155-7535

Table of Contents

Perception Isn't Everything: The Reality of Class Size

Mark Umbricht and
Kevin Stange

Acknowledgments

We would like to thank Paul Courant, Ben Koester, Steve Lonn, Tim McCay, and numerous participants in seminars at the University of Michigan for helpful comments. This research was conducted as part of the University of Michigan Institutional Learning Analytics Committee. The authors are solely responsible for the conclusions and findings.

About the Authors

The authors are with the University of Michigan. Kevin Stange is also with the National Bureau of Economic Research.

Abstract

Every term, institutions of higher education must make decisions about the class size for each class they offer, which can have implications for student outcomes, satisfaction, and cost. These decisions must be made within the current higher education landscape of tightening budgets and calls for increased productivity. Beyond institution decision making, prospective students and their families may use class size as one factor in deciding whether an institution might be a good fit for them. The current measure of class size found in university fact books, and subsequently sent to numerous ranking groups such as U.S. News & World Report (hereafter U.S. News), is an inadequate gauge of the student experience in the classroom, as measured by the percent of time students spend in classes of varying sizes. The current measure does not weight for enrollment, credits, or multiple components of a class, which results in a misleading representation of the student experience of class size. This paper will discuss these issues in depth, explain how class size varies across institutions, and offer recommendations on how to reweight class size in the Common Data Set to accurately describe it from the student's perspective. Institutions could use this new metric to better understand class size, and subsequently to understand the student experience and cost of a class, while prospective students and their families could use the metric to gain a clearer picture of the class sizes they are likely to experience on campus.

Keywords: class size, productivity, student outcomes

The AIR Professional File, Fall 2019
Article 146

https://doi.org/10.34315/apf1462019

INTRODUCTION

Organizing class delivery is a key operational decision for institutions of higher education. Each term these institutions must decide how many students will be taught within each section given the classes they offer, the faculty and instructors they have available to teach, and the confines of physical spaces they have on campus. Within these constraints, institutions must decide how to deliver classes. Consider a popular class taken by nearly every first-year student: Should this class be taught as one large lecture by a famous professor, many small sections taught by graduate students, or a combination of the two? Should small classes target freshmen who are acclimating to college or seniors as they specialize in their field?

Institutions, particularly public institutions, make these decisions within the current context of increased accountability and decreased resources. Traditional wisdom argues that smaller classes increase engagement, facilitate student-faculty interactions, and improve student success. The opportunity to learn from prominent scholars in the field is also considered a strength of undergraduate education at research universities. However, smaller classes and senior faculty are costlier and their use comes at the expense of other ways of enriching and supporting the undergraduate student experience. As Courant and Turner (forthcoming) argue, institutions have an interest in providing curriculum efficiently, meaning they must strike a balance between quality, costs, and tuition revenue. If an institution or department has an influx of students, it must decide whether it will increase the size of its faculty or the class size of its courses. Therefore, decisions about class size have first-order influence on student success and institutional costs. From the student perspective, class size could be influential in the college choice process, with some students seeking intimate class settings with small class sizes, and others preferring to blend in to a large classroom. Students and their families rely on institutional websites and rankings, such as Princeton Review or U.S. News, and other publicly available data for information about class size. These data are typically drawn from the Common Data Set (CDS), which is a collaborative effort among data providers and publishers to improve the quality and accuracy of information provided to prospective students, and to reduce the reporting burden on data providers (CDS Initiative, 2018). While it is helpful to have a measure that can be reported across multiple campuses, the class size metric used by the CDS is measured at the classroom level rather than at the student level. This difference in measurement leads to a disconnect between the metric and the phenomena it is trying to describe, as described in the following example.

Imagine a high school student researching her nearby public, research university as a prospective student. She sees that, according to U.S. News in 2018, only 17% of classes have more than 50 students, and 57% of classes have fewer than 20 students. The student thinks, What luck! She thinks she can attend a high-quality research institution while spending most of her time in small classes. After graduation from that college, the same student looks back and sees that she spent more than 41% of her time in classes with more than 50 students, and only 20% of her time in classes with fewer than 20 students. These differences in the perception versus reality are not exaggerated, but rather are many students' average experience. This paper will show that the measure of class size calculated for the CDS, and subsequently used by many other sources, does not provide an accurate approximation of the true class size experienced by students at the University of Michigan (U-M), a large, research university in the Midwest. By the term "student experience," we mean the percent of time or credits spent in classes of varying sizes. The results of this case study could be replicated at any institution, with varying degrees of departure from the true student experience depending on the institution type and size. Specifically, this paper will argue for a new class size metric to be used in the CDS and will address the following questions:

Framing Questions

1| How does the standard definition of class size vary from the student experience?

2| How does a student-centric version of class size vary across an institution?

3| How can institutional researchers practically recalculate class size to better approximate the student experience without significantly increasing the burden of data providers?

Importance of the Topic and Literature Review

This topic is important for students, institutions, and the field of institutional research. From the student perspective, students and those assisting in their decisions need accurate and meaningful information to make the best decision about which college to attend. A small number of studies have shown that class size is an important factor for students as they select an institution (Drewes & Michael, 2006; Espinoza, Bradshaw, & Hausman, 2002). This makes sense since lower class size is perceived to be linked to gains in student outcomes. Literature in the secondary setting is clear that lower class size is associated with gains across multiple areas, including test scores, noncognitive skills, college enrollment, and other outcomes (Angrist & Lavy, 1999; Chetty et al., 2010; Dee & West, 2011; Dynarski, Hyman, & Schanzenbach, 2013; Hoxby, 2000; Krueger, 1999). However, in higher education the relationship between class size and outcomes is not well established, with studies finding either negligible association (Bettinger, Doss, Loeb, Rogers, & Taylor, 2017; Lande, Wright, & Bartholomew, 2016; Stange & Umbricht, 2018; Wright, Bergom, & Lande, 2015) or a negative relationship between class size and outcomes (Bettinger & Long, 2018; De Giorgi, Pellizzari, & Woolston, 2012; Kokkelenberg, Dillon, & Christy, 2008). Institutions that gain a more accurate and more nuanced version of class size from the student experience perspective could aid prospective students in their decision-making process.

Class size is also important to institutions for planning purposes. Courant and Turner (forthcoming) argue that institutions must strike a balance between quality, costs, and tuition revenue. In recent years, institutions have been asked to cut back and do more with fewer resources, which would imply that increasing class size would be an appropriate strategy. In fact, class size is one of the most important drivers of instructional costs (Hemelt, Stange, Furquim, Simon, & Sawyer, 2018). However, lower class size is perceived to lead to better student outcomes and is subsequently tied to rankings such as those at U.S. News. This common perception pulls institutions to keep class size lower, putting institutions in a situation where a logical solution is to hire cheaper instructors, such as noncontingent faculty. The ultimate decision on how to strike this balance is not traditionally made at the institution level, but rather at the department level. Cross and Goldenberg (2009) found that the number of noncontingent faculty at elite research institutions rose significantly in the 1990s, which was due to micro- (department-)level decisions. Departments (or colleges) that are particularly concerned with the quality (or perceived quality) of small class sizes would find it difficult to adequately assess how much time their students spend in classes of a given size with the current metric, which is at the class level. A new student experience version of class size would allow departments to compare the class size experience across multiple majors or between departments, which could assist in balancing the class size constraints for long-term planning. Having an accurate understanding of imbalances by class size across colleges, departments, or majors could help institutions pinpoint areas that need improvement. In addition, institutions could examine whether access to smaller classes is inequitable across certain student groups, such as among minority, first-generation, or first-year students.

We will argue that the current definition of class size used in the CDS Initiative (2018) is insufficient for both internal planning and external consumption. As institutional researchers, it is our job to provide meaningful and accurate information to both internal and external parties. While the traditional measure of class size may be accurate, this paper describes ways in which we could provide data that are more meaningful. Institutional researchers and higher education professionals have an obligation to update this metric to reflect the actual student experience.

DEFINING CLASS SIZE

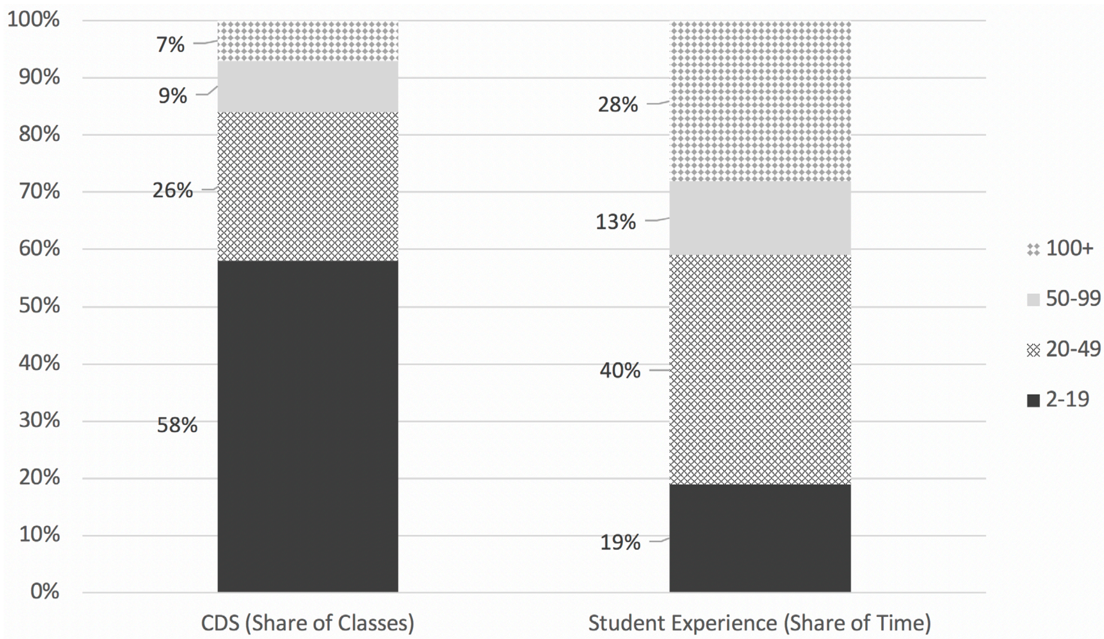
Based on the conventions of the CDS, undergraduate class size is calculated based on the number of classes with a given class size range. Classes are divided into sections and subsections. A class section is an organized class that is offered by credit, is identified by discipline and number, meets at one or more stated times in a classroom or setting, and is not a subsection such as a laboratory or discussion section. A class subsection is a part of a class that is supplementary and meets separate from the lecture, such as laboratory, recitation, and discussions sections. In calculations of class size, we count only the sections of a class and discard the subsections. The CDS conventions consider any section or subsection with at least one degree-seeking undergraduate student enrolled for credit to be an undergraduate class section, but exclude distance learning, noncredit, and specialized one-on-one classes such as dissertation or thesis, music instruction, one-to-one readings, independent study, internships, and so on. If multiple classes are cross-listed, then the set of classes are listed only once to avoid duplication (CDS Initiative, 2018). This means that we count stand-alone classes, defined as having only one component, once per section in the class section portion.

For classes with multiple components, such as a lecture section combined with a lab or discussion section, we count each lecture section once in the class section portion while we count each associated lab or discussion section once in the class subsection table. In traditional class size metrics, the CDS counts only the class section portion of the class while the CDS discards the subsection from the calculation. This metric is relatively easy to compute and is comparable across campuses, but it may not be representative of the student experience.

We define "student experience class size" as the percent of time spent by a student in classes of various sizes, using credits as a proxy for time.[1] Calculations for this metric will be discussed later in

1. We assume that each credit associated with a class is approximately 50 minutes of class time. While this pattern is not universal across U-M, the calculation is easy to make and should be readily accessible in any institution's data warehouse. A more complicated, but 100% accurate approach, would be to use the day and time location to derive the true number of minutes spent in each section. These data might not be accessible and would require substantial coding to calculate. We tested both approaches and our simplified approach did not meaningfully differ.

Figure 1. Class Size by Various Sources



this paper. Figure 1 shows the difference between the CDS method and our new student experience method of computing class size. Sources that are often used as references for prospective students and institutions, such as U.S. News and institutional websites, draw data from the CDS. The CDS metric describes the share of classes in a given range rather than the share of time spent in classes of varying sizes. According to the U.S. News and the U-M websites, 84% of classes at U-M have fewer than 50 students, and 57% have fewer than 20 students. However, using our student experience class size metric, only 19% of a student's classroom time is spent in classes with fewer than 20 students and nearly 30% of their time is spent in classes with at least 100 other students. Why do these metrics differ so drastically?

There are three primary reasons driving these differences. First, the traditional measure for class size is not weighted by the number of students enrolled. A 500-student section and a 5-student section both count as one class, even though many more students experience the larger section. Second, classes are not weighted by the number of credits associated with the class. A class worth five credits counts for the same as a class worth one credit, even though students likely spend five times as much time in the first class. Finally, the traditional measure does not incorporate subsections. It is typical for large lecture classes to have multiple components, such as a large lecture of 200 that meets for 2 hours per week and 10 associated small discussion groups of 20 students each that meet for 1 hour per week. Students spend 67% of their classroom time in a large lecture and 33% of their time in a small discussion, but the traditional metric counts only the lecture portion. This means the 200-person lecture counts as one class, ignoring the subsections. Our new student experience class size metric accounts for these three factors, as will be explained in detail in the methods section. This new metric provides a more accurate representation of the student experience within the classroom.

Given the immense difference between the traditional class size measure and our student experience version, it is only natural to question why institutions have not moved to a different calculation of class size. To be clear, it is not the authors' belief that institutions are purposely trying to push an inaccurate measure of class size. The traditional class size measure does hold value in describing the number of classes available to students of various class sizes. U-M students can choose from many small classes and could theoretically construct a set of classes to minimize the amount of time spent in large classes. In reality, though, several forces could make it difficult for institutions to switch to a student experience version of class size.

First, there are serious consequences in rankings and optics for many institutions, particularly larger ones. U.S. News currently provides points to institutions based on the share of small sections (19 and fewer students), partial points for the share of medium sections (20–49 students), and no points for large sections (50 or more students).[2] Universities that use larger class sizes to teach a large number of students will see their rankings negatively impacted if classes were weighted by the number of students taking the class. The shift in class size would also create poor optics for prospective students and may impact whether they choose one institution over another. The same would be true for a student experience measurement of instructor type, where larger research institutions are much more likely to use graduate students as instructors. In addition, having non-tenure track instructors primarily responsible for teaching large classes, and therefore many students, could have bad optics for institutions. Therefore, there is a disincentive for a single college, or a small group of colleges, to recalculate their class size based on the student experience. The exception would be institutions that uniformly have very small classes, such as small liberal arts colleges, that would see little or no change in their calculation of class size.

A second difficulty is measuring class size from the student perspective. Leaders of the CDS Initiative already consider the measurement of class size to be the second-most difficult part of the CDS, with only calculations of financial aid deemed more difficult (Bernstein, Sauermelch, Morse, & Lebo, 2015). At U-M, measures required to recalculate class size to the student perspective are readily available and clean, and require little manipulation to combine. Institutions vary significantly in their data capacity and availability of staff to adjust the CDS measures. Given these challenges, the authors of this study still believe that shifting to a student experience version of class size would provide many benefits, including an accurate representation of the amount of time students spend in classes of varying sizes and with various instructor types. This shift would be beneficial for institutions for planning purposes as well as for prospective students as they weigh various institutions during the selection process.

2. This recently changed from a system that provided points for small courses and penalties for large courses in an effort to minimize gaming of the system (Supiano, 2018). However, there is no evidence provided to back up this claim. Regardless of whether the new or old system is used, having a larger number of small classes is rewarded, and the metric is at the course level, which does not appropriately describe the student experience.

Table 1. Example of Distribution of Class Credits in Our Framework

Number of Students	Class 1	Class 2	Class 3	Class 4	Total Credits	% of Total
2–9						
10–19						
20–29	1 (Lab)	1 (Discussion)			2	17%
30–39						
40–49			3 (Lecture)		3	25%
50–99				2 (Lecture)	2	17%
100–199	2 (Lecture)				2	17%
200+		3 (Lecture)			3	25%
Total Credits	3	4	3	2	12	100%

METHODS

The purpose of this study is to create a student experience version of class size. We drew data from the U-M data warehouse, specifically from the Learning Analytics Data Architecture (LARC) and College Resource Analysis System (CRAS).[3] LARC is a flattened, research-friendly version of the raw data warehouse that houses data about students, their background, their progress, and their coursework. CRAS is a data warehouse system that houses data about classes and the instructors that teach them. The sample included first-time freshman students in cohorts between 2001 and 2012, examining classes taken within 4 years of entry. Freshman cohorts are between approximately 5,500 and 6,500 students during this period. Individual study classes and one-on-one classes were removed from the sample, as were classes with no CRAS information, including subjects such as medicine, dentistry, armed forces, study abroad, and classes through the Committee on Institutional Cooperation program. The final sample included 70,426 first-time students and 3,398,320 class sections taken in these students' first 4 years of study, between 2001 and 2016.

Calculating Class Size

As previously noted, we made three adjustments to the traditional measure of class size: (1) weighting for number of students, (2) weighting for credits associated with the section, and (3) incorporating subsections. Table 1 provides an example of how

3. LARC is unique to U-M, although some institutions have developed a similar research-friendly database. CRAS is also unique to U-M because there are some calculations made with institution-specific formulas. However, these databases comprise data that are regularly available (although not always clean) at all institutions, such as student and class information, classes taken, and the number of students in a given class.

we accounted for credits and subsections in our student experience framework.

This example shows a hypothetical student's coursework for one term. This student took four classes in this term for a total of 12 credits. Classes 1 and 2 had multiple components, with a lecture and either a discussion or a lab. We divided credits for these classes among the section and subsection based on the amount of time or credits associated with each component. Then we put each component into the appropriate class size bin, shown in the left column. Classes 3 and 4 were stand-alone classes that contained only a lecture component, so there was no need to divide their credits. Once we had distributed all the credits for each class to the appropriate class size bin, we totaled credits across each class size row. In doing so, we accounted for credits and the multicomponent nature of classes, fixing two of the issues with the standard definition of class size. The third piece relates to weighting class size by the number of students in the class. In this framework, we theoretically create a table like this for each student and each term the student attends. We then summed across every student and term. Since the level of observation is a class enrollment, we naturally weight by the number of students in the class because there will be 50 observations if there are 50 students in a class, or 5 observations for a class with 5 students. A basic assumption made by this framework is that one credit is equal to approximately 1 hour of class time. While there are a small number of classes that violate this assumption, we do not believe it would impact our overall results in a meaningful way. It is also important to note that enrollments for cross-listed classes were combined into the home class. At U-M, if there are multiple cross-listed sections, one is considered the home class and the rest are considered away classes. This means that if there are three cross-listed sections with 12, 15, and 18 students, our data would show one class with 45 students.

Once we calculated the percent of time (using credits as a proxy) that a student earned in various class sizes across his first four years, we calculated percentiles for each enrollment group across the entire university by college, by major, and by year in school.[4] College and major are determined by the last college or major associated with a student. If a student graduated with a bachelor's degree, we used his graduating college or major. If a student departed prior to completing his degree, we used the last known college and major associated with him.[5] Rather than showing just the median or mean, we chose to use five percentiles (10th, 25th, 50th, 75th, and 90th) to show the distribution of student experiences. Finally, we mapped these percentiles into figures that show the range of student experiences for a given class size.

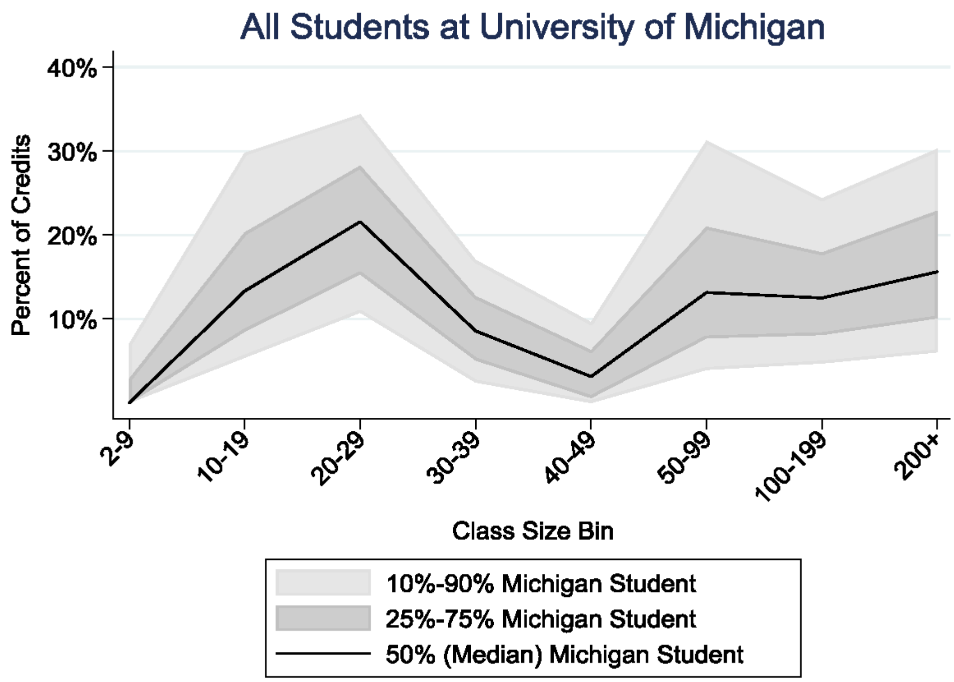
CASE STUDY

This section will examine how class size varies across U-M. The figures in this section represent the distribution of time that students spend in classrooms of varying sizes. Class size was grouped into eight bins of varying sizes to create a smooth figure and to mimic the traditional measure of class size. Figure 2 shows the distribution of class size across the entire university. The black line represents the median student, the dark gray shaded area represents the 25th to 75th percentile, and the light gray shaded area represents the 10th

4. Results show the 10th, 25th, 50th, 75th, and 90th percentile for each enrollment group. The term "college" refers to an academic college, such as engineering, liberal arts and science, or education.

5. We considered removing these students but decided that doing so could introduce some selection bias. For example, students that drop out may select larger courses; if we remove them, we may distort the student experience.

Figure 2. Percent of Time Spent by Size Across the University of Michigan

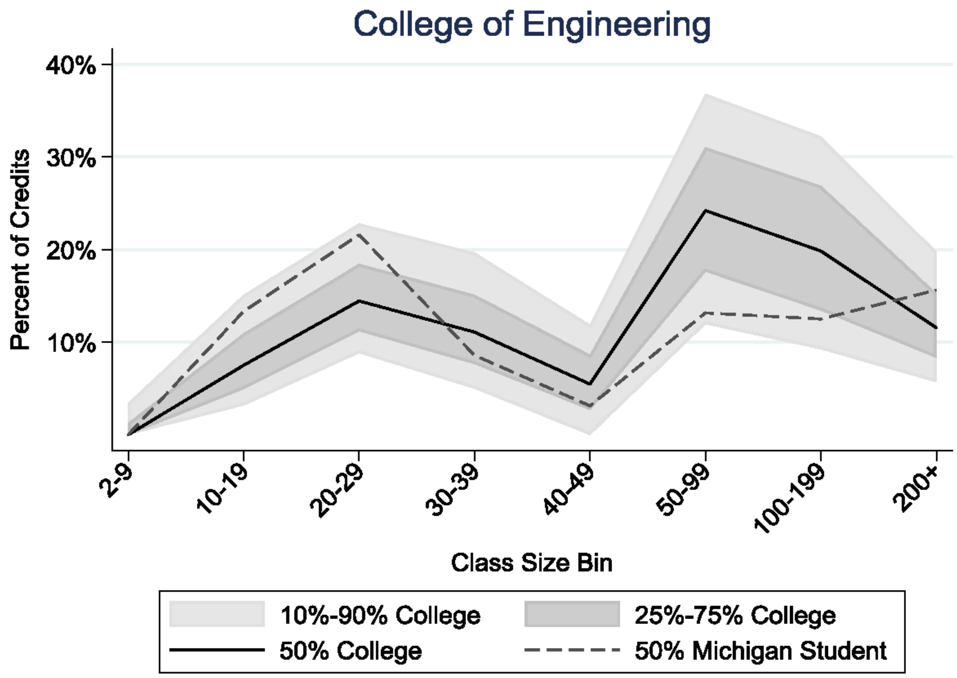


to 90th percentiles.

To interpret this figure, consider the enrollment group of 200+ students. The black line indicates that the median U-M student spent about 15% of her time in classes with 200 or more students. Moving above the black line to the edge of the dark gray area, we see that about a quarter of students spent about 23%–24% of their time in classes with 200 or more students. The outer edge of the light gray area indicates that 10% of students at U-M spent more than 30% of their time in these very large classes. If we move to the 20–29 enrollment group, we see that the median student spent about 20% of her time in classes with 20–29 students. There is a very wide range of experiences (25%) between the 10th and 90th percentiles, indicating that students may spend vastly different amounts of time in classes with 20–29 students. The spread is only about 10% wide for enrollment groups between 30 and 49 students, indicating there is less variability in the percent of time spent in medium-size classes. Overall, it is clear that students' time is more heavily weighted in both large (50+ students) and small (10–29 students) classes. This means that students spend their time in classes of varying sizes, but classes at U-M tend to favor high or low enrollments on average. This also implies that very large classes likely have a smaller enrollment component tied to them, such as a discussion or lab section.

At a university the size of U-M, with an undergraduate population of almost 30,000, it is natural to assume that student experiences may vary greatly across the institution, such as by college, major, or year in school. Figure 3 shows the distribution of class size in the College of Engineering, which we chose because it is the

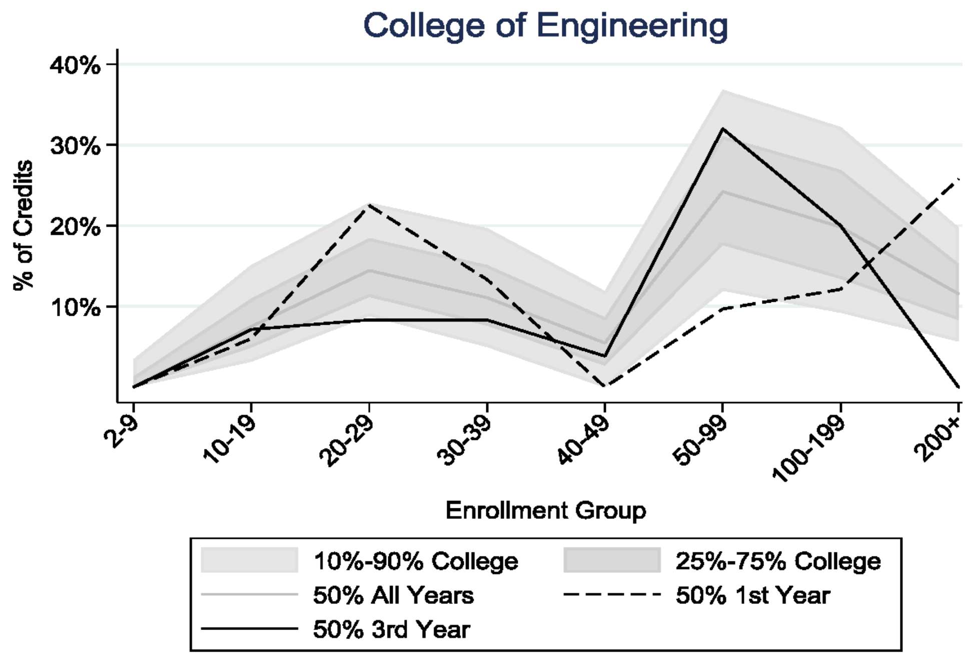
Figure 3. Percent of Time Spent by Enrollment Group in the College of Engineering



second-largest college on campus and differs significantly from the trends for the median student. Once again, the black line and gray shaded areas represent the percentiles for students in a given college. We added the gray dotted line here to show the median U-M student from Figure 2, allowing us to examine how the college differs from the overall pattern at the university. In the College of Engineering students tend to take coursework with very different class sizes compared to the median U-M student. In particular, engineering students take fewer small classes (10–29 students) and very large classes (200+ students), and instead take more classes in the middle range (30–199 students). Part of the difference could be attributed to deliberate planning by the College of Engineering and part could be related to the size of classrooms in the engineering buildings. Classroom caps, and subsequently enrollment caps, for classes in engineering tend to lie in the middle of the enrollment group distribution.

While not shown, we created figures for every college on campus. The trends of these figures show that class size differs significantly across colleges. The College of Literature, Science, and the Arts (LSA) has very similar trends to the median U-M student, in part because it is the largest college on campus. Given that more than 60% of students are in LSA, this college drives much of the median class size. As one would expect, smaller and more narrowly focused colleges such as the College of Architecture and Urban Planning, the College of Art and Design, and the College of Music, Theatre, and Dance had significantly higher levels of small classes (2–29 students) and fewer very large classes (200+ students). The College of Public Policy tends to mimic the trends of LSA, in part because students

Figure 4. Class Size Within the College of Engineering by Academic Year



spend their first two years in LSA before declaring their major. The College of Business had very high levels of medium-size classes (40–99 students) because many of its core classes have enrollments of 40 to 80 students.

Class size also varies in systematic ways by major, even within a college. Comparing the major of Arts and Ideas in the Humanities, a small, multidisciplinary major, to the major of Biopsychology, Cognition, and Neuroscience (BCN), a large, premed major, yielded large differences in class size. The median Arts and Ideas major spent about 50% of her time in classes with 20 or fewer students, which is twice the 25% of the median LSA student. Students in the BCN major, on the other hand, took a much larger share of large lectures, rising out of the 90th percentile for LSA students. They spent nearly 33% of their time in classes with more than 200 students, compared to only 20% of the median LSA student's time. While we observed systematic differences between majors, we also found that there were some majors that had very similar class size structures.

A final way to observe how class size varies across an institution is by comparing class size by academic level. Rather than aggregating across a college or major, we aggregated by a student's year in school (e.g., freshman, sophomore, junior, senior). Students in their first year typically fulfill their general education requirements, which tend to be classes taught to many students at once. By their junior or senior year, students tend to take many classes within their major of increasing depth and specialization, characterized by smaller class sizes. Splitting the data in this manner did yield interesting differences over time. As shown by the dotted line in

Figure 4, students in the College of Engineering took similar coursework to LSA students in their first year compared to the median U-M student from Figure 2, taking mainly classes consisting of large lecture and small discussion groups. This is likely because students are fulfilling their general education requirements. By their third year, the solid black line indicates the median engineering student deviates from the pattern at the college level and takes a much higher proportion of classes with 50–100 other students; these classes comprise classes in the students' major.

Overall, we created figures for every major, college, academic level, and student across campus to examine how class size varied across U-M. The median student within some majors had class size distributions that were very similar to the median student in their college and the university, but many majors differed significantly from the general trends. Similarly, there were some students within majors that varied significantly from the median student in their major. These figures show that class size can vary significantly across an institution. A department, college, or institution can use this information to gain a more nuanced understanding of the experiences of their students. For example, if students that tend to take only large lectures or small discussion classes perform worse, then an institution could adjust its advising to promote students to take classes of varying sizes. Similarly, an institution could identify whether certain student groups, such as first-generation students, may benefit from a more intimate classroom environment where they receive more attention from instructors.

HOW SHOULD WE MEASURE CLASS SIZE?

This paper has shown that the traditional measure of class size is not sufficient if it is meant to provide information about a student's actual experience in the classroom. Previous research has shown that increasing class size has a mixed but generally negative impact on student learning and satisfaction. Accurately measuring class size is also important to institutions because it impacts productivity. For example, increasing or decreasing the class size of introductory calculus, a class that most students on campus take, can have vast implications for the cost of the class. If an institution wanted to recalculate class size with the student experience at the core, what would it look like?

We will first consider two simple adjustments: weighting for credits and weighting for students. Table 2 shows the distribution of class size given four different calculations and Figure 5 provides a visual representation of the table. The first calculation is the traditional measure of class size, with no adjustments. This means there is one observation per lecture. The second column weights the traditional measure by number of credits. A class of four credits is now worth twice as much as a class of two credits. This slightly shifts the class size distribution down because large class sections have corresponding subsections. Consider a class worth three credits, two earned in a lecture and one earned in a lab. Since the lab, or subsection, is not counted in the traditional measure, one of the three credits is discarded, deflating the value of a large class. The third column accounts for subsections and credits, appropriately distributing all the credits associated with each class. A class worth three credits including a lecture and a lab, as described above, is now fully included in the metric for class

Table 2. Class Size by Three Different Calculations

Class Size	Traditional Measures (1)	Weight for Credits (2)	Account for Subsections (3)	Weight for Students (4)	Weight for All Changes (5)
2–9	13.9%	14.1%	11.4%	2.0%	2.5%
10–19	32.1%	32.5%	31.1%	13.0%	15.4%
20–29	23.0%	23.6%	34.0%	14.2%	21.9%
30–39	8.2%	8.5%	9.4%	6.8%	9.3%
40–49	4.4%	4.3%	3.2%	4.6%	4.4%
50–99	10.5%	10.2%	6.4%	18.2%	15.8%
100–199	4.9%	4.5%	2.9%	17.2%	13.8%
200+	2.9%	2.4%	1.7%	24.0%	16.8%

size (Column 3). This lowers class size considerably because we are not removing the subsection but instead including it in our calculation. Subsections tend to be smaller, which drives the distribution down.

The fourth column shows class size distribution if we weighted only for the number of students in a class. Using this calculation, a large lecture of 500 students is counted 500 times, while a small class of 5 students is counted only 5 times. As expected, this dramatically shifts the distribution of class size upward, nearly flipping the distribution of class size from the traditional measure. Column 5 accounts for weighting by credits and students, and includes subsections simultaneously. This gets closer to the perfect measure of the student experience described earlier in this paper and vastly improves the understanding of class size for institutions and potential students. Figure 5 shows a visual representation of each strategy for calculating class size. We can see that the traditional measure (nonstudent experience), weighting for credits, and weighting for subsections are very similar, with slight increasing low enrollments (fewer than 30) and decreasing other enrollments (30–200+) after accounting for subsections. Once we weight for students there is an immediate shift, nearly flipping the distribution. However, this shift goes too far; accounting for credits and subsections provides a balance between the three strategies.

It is important to note that our calculation of class size used student-level micro data to account for these changes, in part due to other related research. However, this calculation may be burdensome for institutions. A simpler way to achieve the same results would be to take the class-level data used for

Figure 5. Class Meeting Size Distribution by Five Different Methods

the CDS and simply weight each class section and subsection by the number of students and credits. This simplifies the process and limits the resources required to pull and process the data.

LIMITATIONS OF NEW CLASS SIZE MEASUREMENT

While this new class size metric may provide a clearer picture of the student experience at an institution, it does not come without limitations and challenges. The first challenge is the technical barriers of calculating the new class size measurement. For example, institutions may house the required variables in different data systems, but replacing or adding a new metric could be difficult to implement and could increase the reporting burden on institutions. Most of what is required to adjust to this new metric is already required for the CDS version of class size. Institutions must calculate the number of sections and subsections of a given class size. This means they must know the exact class size for every section and subsection. The only missing piece is to allocate time between the section and the subsection. At U-M, a field for distributed hours is contained in the data warehouse, but that may not be the case for all institutions, some of which may not even use the Carnegie credit system. For these institutions, a calculation could be made to allocate credits based on the amount of classroom time for each section and subsection. For example, a three-credit class that meets for 2 hours in a lecture section and 1 hour in a discussion section could split the class into two credits for the lecture and one credit for the discussion. This would require some up-front work to create these calculations based on day/time information, but is not unmanageable.

The second set of challenges relate to the nuance that is inherent in the new metric. Presenting students with a single metric for an entire university is simple and easy to explain but leaves out a lot of nuance. If an institution were to recreate the figures in this paper, it would introduce some challenges for interpretation. As shown in Figure 3, there are stark differences in class size between engineering and LSA students; the same differences can be shown between majors or year in school. The large number combinations and comparisons by college, major, and year in school present challenges when trying to show or explain the data to prospective or current students. While a single, institution-wide metric leaves out nuance, it is a vast improvement over the current metric used by institutions and rankings. If an institution wanted to provide more nuance, it could create an interactive dashboard for prospective and current students to view the nuanced version of class size presented in the figures of this paper. Students could select a college or major from a list to see what the distribution of class size looks like for students in that major. Institutions could pare down these figures to show only the median for simplicity or to provide a detailed explanation of how the percentiles work when students first use the new dashboard. However, with a nuanced view an institution could also show students that class size may be lower for a given major, or how class size may lower as students progress toward their degrees.

DISCUSSION AND FUTURE WORK

As institutional researchers it is imperative for us to provide data that are both accurate and meaningful. This paper argues that the traditional measure of class size is not a meaningful representation of what students experience in the classroom. The traditional measure of class size illustrates only the proportion of classes that are small, not the amount of time that students spend in small classes. This is problematic for prospective students, who could use meaningful class size data to determine where they want to attend college. It is also problematic for institutions, which may not understand the extent of large or small classroom experiences on their campuses. While limited, previous research in higher education suggests that class size matters for student outcomes and satisfaction in classes.

This paper suggests that the measurement of class size could be altered by weighting for the number of students and credits associated with the section, and accounting for subsections. We suggest that institutional researchers consider revamping their class size metric to reflect the student experience in the classroom more accurately. Nearly all the required components (number of students in each class and section/subsection, and number of classes) for this calculation are used by the current CDS metric. We believe that distributed credits, the potential missing component, is likely captured and readily available at many institutions, which would make this adjustment relatively easy. While this could require an investment of time on behalf of some institutions, we believe the potential benefits will outweigh the investment. At a minimum, we suggest that institutional research professionals consider reweighting their current metric of class size by the number of students in each lecture section. This new metric provides a more accurate description of class size from the student experience.

Future work on this topic could yield more improvements in the description of the student experience at institutions of higher education. First, the range of class size also varies significantly across colleges and departments, indicating that a simple institution-wide metric masks important differences for students who plan to major in different fields. Institutions could take this idea one step further to create an interactive dashboard that allows prospective and current students the opportunity to see the range of class size experiences for majors in which they have interest. A second improvement institutions could make would be to pair data about class size and instructor type (e.g., IPEDS instructor type). By combining the percent of time spent in varying class sizes and varying instructor types, students will gain a clearer picture of what their classroom experience would be at a particular institution.

REFERENCES

Angrist, J., & Lavy, V. (1999). Using Maimonides Rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics, 114*(2), 553–575. https://www.jstor.org/stable/2587016?seq=1#page_scan_tab_contents

Bernstein, S., Sauermelch, S., Morse, R., & Lebo, C. (2015). Fundamentals and best practices for reporting Common Data Set (CDS) data. Presented at the Association for Institutional Research Annual Conference, Denver, CO, May 27.

Bettinger, E., Doss, C., Loeb, S., Rogers, A., & Taylor, E. (2017). The effects of class size in online college courses: Experimental evidence. *Economics of Education Review, 58*, 68–85. https://doi.org/10.1016/j.econedurev.2017.03.006

Bettinger, E. P., & Long, B. T. (2018). Mass instruction or higher learning? The impact of college class size on student retention and graduation. *Education Finance and Policy, 13*(1), 97–118. https://doi.org/10.1162/edfp_a_00221

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Quarterly Journal of Economics, 126*(4), 1593–1660. https://doi.org/10.1093/qje/qjr041

Common Data Set (CDS) Initiative (2018). 2018–2019 Common Data Set Initiative. http://www.commondataset.org/

Courant, P. N., & Turner, S. (forthcoming). Faculty deployment in research universities. In Hoxby, C. M. & Stange, K. (Eds.) *Productivity in higher education.* Chicago: University of Chicago Press.

Cross, J. G., & Goldberg, E. N. (2009). *Off-track profs: Nontenured teachers in higher education.* Cambridge, MA: MIT Press.

De Giorgi, G., Pellizzari, M., & Woolston, W. G. (2012). Class size and class heterogeneity. *Journal of the European Economic Association*, 10(4), 795–830. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1542-4774.2012.01073.x

Dee, T. S., & West, M. R. (2011). The non-cognitive returns to class size. *Educational Evaluation and Policy Analysis, 33*(1), 23–46. https://doi.org/10.3102/0162373710392370

Drewes, T., & Michael, C. (2006). How do students choose a university?: An analysis of applications to universities in Ontario, Canada. *Research in Higher Education, 47*(7), 781–800. https://doi.org/10.1007/s11162-006-9015-6

Dynarski, S., Hyman, J., & Schanzenbach, D. (2013). Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management, 32*(4), 692–717. https://doi.org/10.1002/pam.21715

Espinoza, S., Bradshaw, G., & Hausman, C. (2002). The importance of college choice factors from the perspective of high school counselors. *College and University, 77*(4), 19–23.

Hemelt, S. W., Stange, K. M., Furquim, F., Simon, A., & Sawyer, J. E. (2018). Why is math cheaper than English? Understanding cost differences in higher education. Working Paper 25314, National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w25314

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of Economics, 115*(4), 1239–1285. https://doi.org/10.1162/003355300555060

Kokkelenberg, E. C., Dillon, M., & Christy, S. M. (2008). The effects of class size on student grades at a public university. *Economics of Education Review, 27*(2), 221–233. https://doi.org/10.1016/j.econedurev.2006.09.011

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics, 114*(2), 497–532. https://doi.org/10.1162/003355399556052

Lande, E., Wright, M., & Bartholomew, T. (2016, October). Impact of smaller class sizes in U-M mathematics courses. Unpublished internal report, Center for Research on Learning and Teaching, University of Michigan, Ann Arbor.

Stange, K., & M. Umbricht, 2018. Undergraduate class size reduction: Effects and costs. Unpublished working paper, University of Michigan, Ann Arbor.

Supiano, B. (2018, March 21). Are small classes best? It's complicated. *Chronicle of Higher Education.* https://www.chronicle.com/article/Are-Small-Classes-Best-It-s/242878

Wright, M., Bergom, I., & Lande, E. (2015, September). Impact of smaller class sizes in U-M Spanish and German courses. Internal report, Center for Research on Learning and Teaching, University of Michigan, Ann Arbor. Executive summary at http://crlt.umich.edu/sites/default/files/SmallClassExecSum.pdf

Enrollment Projection Using Markov Chains: Detecting Leaky Pipes and the Bulge in the Boa

Rex Gandy, Lynne Crosby, Andrew Luna, Daniel Kasper, and Sherry Kendrick

About the Authors

The authors are with Austin Peay State University.

Abstract

While Markov chains are widely used in business and industry, they are used within higher education only sporadically. Furthermore, when used to predict enrollment progression, most of these models use student level as the classification variable. This study uses grouped earned student credit hours to track the movement of students from one academic term to the other to better identify where students enter or leave the institution. Results from this study indicate a high level of predictability from one year to the next. In addition, the use of the credit hour flow matrix can aid administrators in identifying trends and anomalies within the institution's enrollment management process.

Keywords: Markov chains, enrollment projections, enrollment management, enrollment trends, enrollment

The AIR Professional File, Fall 2019

Article 147

https://doi.org/10.34315/apf1472019

INTRODUCTION

The current challenges facing higher education administrators create myriad reasons to find a crystal ball of sorts to effectively forecast enrollments, predict how many current students will stay at the institution, forecast new students, and adequately estimate revenues. These challenges have become only more pressing in recent years.

More than 20 years ago, when public college and university revenues were ample, administrators were not readily concerned about the future of college enrollments or student persistence. State appropriations were healthy and usually made up more than half of an institution's revenue source. Moreover, with lower tuition more students could afford to obtain a degree without going into significant financial debt (Coomes, 2000).

The costs to run higher education have skyrocketed, however, causing today's institutions to seek scarce resources within an ever-diminishing financial pool. As states tackle other pressing issues such as infrastructure, entitlements, and prisons, the amount they give to higher education naturally wanes. Decreased state revenue, therefore, compels institutions to increase tuition to make up the difference. According to Seltzer (2017), for every $1,000 cut from per student state and local appropriations, the average student can be expected to pay $257 more per year in tuition and fees. He further notes that this rate is rising.

In addition to decreases in state revenues, higher education administrators are under increasing pressure to be accountable to federal and state governments as well as to regional and discipline-based accreditors. This accountability is increasingly seen in tougher reporting standards, outcomes-based funding formulae, and mandated student achievement thresholds.

The closest resource to a crystal ball available to administrators is a set of mathematical prediction tools. These prediction tools range from simple formulae contained in spreadsheets to much more complicated regression, autoregressive integrated moving average (ARIMA), and econometric time series models.

According to Day (1997), current predictive tools that are statistically based rely on the institution's ability to access and manipulate large datasets and individual student-record data. While more-complicated statistical models incorporate variables such as tuition cost, high school graduate numbers, economic factors, and labor-market demand, other models look more specifically at institutional indicators such as high school grade point averages of entering freshmen, as well as the retention, progression, and graduation rates of students.

One such model, the Markov chain, has been relatively underutilized as an enrollment projection tool in higher education. When used properly, however, it can aid institutions in determining progression of students. Specifically, Markov chains are unique from more-traditional ARIMA and regression prediction tools in that the following is true:

1| Markov chains can give accurate enrollment predictions with only the previous year's data. These predictions can be helpful when large longitudinal databases are not available.

2| They can generate predictions on segments of a group of students rather than on the entire population. Other models often require the use of the entire population.

3| The almost intuitive nature of the Markov chain lends well to changes in student flow characteristics that often cannot be explained by a complex statistical formula.

Moreover, Markov chains might be particularly helpful in determining progression of students during benchmark years when enrollments vary significantly due to state mandates and policies, or due to institutional changes in admission standards. The purpose of this study is to show how a Southeastern, masters-level (Larger Programs) public institution utilized the unique properties of this model to create a tool to better understand credit hour flow and student persistence.

Enrollment Management's Problem with Leaky Pipes and the Bulge in the Boa

While enrollment management has clearly evolved since the inception of the field of enrollment management in the 1970s, some fundamental processes have essentially stayed the same. Institutions have always wanted to attract the right students who fit well within the institution's role, scope, and mission. Once matriculated into the institution, there is also a strong desire for students to adequately progress through their program and graduate within a reasonable amount of time (Hossler, 1984). As enrollment management developed through time, however, administrators became increasingly aware that college-age students were more difficult to enroll, higher tuition was causing some students to forgo their degree, and institutional loyalty was waning as students transferred to similar or different institutions. Furthermore, institutions have seen an increasing number of students who are not fully prepared for the rigors of college work, putting greater enrollment strain on institutions (Johnson, 2000).

After more than 40 years of enrollment management within higher education, it is not surprising that metaphorical associations have entered the lexicon of the profession as administrators try to better understand and predict student matriculation, persistence, and graduation. For instance, Ewell (1985), referred to students progressing and moving throughout their program as student flow, while Clagett (1991) discussed following the flow of student cohorts through to graduation. Luna (1999) used the concept of student flow to explain the various pathways by which the institution may retain students, and Torraco and Hamilton (2013) discussed the student flow of selected groups of minority students. Furthermore, many software companies have exploited the student flow metaphor to describe use of data to identify areas where leakage is present in student flow pipelines. It is easy, then, to see how the management of student retention can be associated with a pipeline and how administrators are busy trying to plug the leaks.

Markov chains are uniquely suited to identifying these leaks because they can model student flow as a set of transitions between several states, much like a set of pipes with various inflows, outflows, and interconnections. In addition to using the model to project enrollments, it is also possible to observe from year to year where students enter the absorbed state (i.e., do not return to the institution). Leakage within the student credit hour (SCH) flow pipeline occurs when students withdraw or stop out due to reasons that are academic, nonacademic, or both. If the model can isolate where the major leaks occur, the institution can identify causes and work to retain and maintain the flow of students within the pipeline. These leaks in the student flow pipeline can be detected and monitored from term to term so that the institution can develop strategies to maintain a healthier flow.

Another colorful bit of jargon among enrollment management professionals is the idea of bulging enrollments. For example, Fallows and Ganeshananthan (2004) use the term "bulging of enrollments" to describe a significantly larger share of students needing financial aid or when, due to rising tuition costs, students bulge into less-expensive 2-year colleges. Herron (1988) uses the term "bulge in the boa" to define instances of oversupply in student populations quickly entering the student flow pipeline, much as a boa constrictor swallows a large meal. Liljegren and Saks (2017) added that these bulges can significantly affect higher education and its future. These bulges occur when large groups of students suddenly enter higher education, putting a strain on the student flow pipeline. As the bulge dissipates, its effects may remain, and it may redefine student flow for the future. With Markov chain models, institutions can monitor these bulges in the system so that they can address issues such as course offerings and instructor availability.

Markov Chains and Higher Education

A Markov chain is a type of projection model created by Russian mathematician Andrey Markov around 1906. It uses a stochastic (random) process to describe a sequence of events in which the probability of each event depends only on the state attained in the previous event.

The Markov chain is a stochastic rather than a deterministic model. Unlike a deterministic process where the output of the model is fully determined by the parameter values and by sets of previous states of these values, a stochastic process possesses inherent randomness: the same set of parameter values and initial conditions can lead to different outputs.

Take, for example, the scenario of an individual returning home from work. In a deterministic process, there is only one route (Route A) from work to home, and the amount of time to get home depends only on the variable speed of the driver. In a stochastic process, the individual will have multiple routes (Routes A, B, and C) from which to choose, and each of the routes intersects the other routes at various points. The randomness of the process occurs when the individual combines routes to go home, if she makes the choices at each intersection randomly. For example, the driver may take Route A part of the time, followed by Route C, then Route B, and back to Route A again, or take some completely different path. There are many random possibilities the individual may take to get home, leading to a variety of possible driving times.

Markov chains utilize transition matrices that represent the probabilities of transitioning from each possible state to each other possible state. These states can be absorbing or nonabsorbing: nonabsorbing states allow future transitions to other states while absorbing states do not.

Markov chains have been widely and successfully used in business applications, from predicting sales and stock prices to personnel planning and running machines. Markov chains also have been used in higher education, albeit with much less frequency.

In most studies where Markov chains were used in enrollment management, the various transitional states were categorized either by student classification or by other simpler dichotomous measures. Given the strength of the Markovian stochastic process in generating student flow probabilities using data only from the previous year, the process of classifying students into other kinds of states could be appealing. Such states could include SCHs, student debt, and (on a more systemwide level) the transitioning from one institution or program to another. The possibilities are diverse.

One of the first to use Markov chains in determining enrollment projections was Oliver (1968) when he compared Markov chains to the much more established use (at that time) of grade progression ratios to predict enrollments at the University of California. According to Oliver's study, enrollment forecasting made a prediction on the basis of historical information on past enrollment and admission trends. In determining a stochastic process, Oliver demonstrated that the fraction of students who leave one grade level (class status) *i* and progress to class status *j* is a fraction p_{ij}; that progress could also be time dependent. These fractions p_{ij} can also be interpreted as random transition probabilities. He determined that the process allowed for contributions in one grade level that were identified by their origins, such as prior grade level, returning to the same grade level, and new admissions (Oliver, 1968).

According to Hopkins and Massy (1981), the use of Markov chains allows the researcher to observe the flow of students from one classification level (i.e., freshman, sophomore, junior, senior) to the next class level. The chain also incorporates students who stay at the same class level from one year to the next. Therefore, the Markov chain for class level, as studied by Hopkins and Massy, can be described as follows:

1| The number of students in class level *i* who progress to class level *j*

2| The number of students in class level *i* who stay in the same level

3| The number of students who leave the institution (drop out, stop out, or graduate)

Similarly, Borden and Dalphin (1998) used Markov chains to develop a 1-year enrollment transition matrix to track how students of each class level progressed. The authors found that unique Markov chain models were valuable in measuring student progression without having to rely on 6-year graduation rate models, which could be ineffective due to the large time lags. Specifically, the model was built around a transition matrix where student flow was tracked from one year to the next, and the rates of transition from four nonabsorption states (i.e., freshman to sophomore) were placed into a matrix that was separate from the two absorption states (i.e., drop out, graduate).

Using the percentages in the two matrices, those students who continue in nonabsorption states were processed through the matrix using the established rates of transition until, asymptotically, all students reach the final absorption state.

Additionally, Borden and Dalphin (1998) developed discrete Markov chain processes to simulate the effect of changes in student body profile on graduation rates. In these models, the authors incorporated credit-load and grade performance categories. Their results indicated that, while there was a strong association between grade performance and persistence, it took very large changes in levels of student performance to impact retention and graduation rates modestly.

In a more narrowly focused study, Gagne (2015) used Markov chains to predict how English Language Institute (ELI) students progressed through science, technology, engineering, and math (STEM) programs. Specifically, the model created transitional (nonabsorbing) states based on classification level and three absorbing states to include those students who left the institution, those who graduated from a STEM program, or those who graduated from a non-STEM program. Findings from their study indicated that the ELI students tended to progress at a higher rate than non-ELI students in STEM programs, and that ELI students who repeated the freshman year were more likely to repeat again than they were to transition to the sophomore year.

Correspondingly, a study by Pierre and Silver (2016) used Markov chain models to determine the length of time it took students to graduate from their institutions. As with previous studies, students were divided into nonabsorbing transitional states (i.e., freshman, sophomore, junior, senior) and absorbing states (i.e., graduate, nonreturning). Using the Markovian property, the future probability of transitioning from one state to another depended only on the present state of the process and was not influenced by its history. The study found that it took 5.9 years for a freshman to graduate and 4.5 years for a sophomore to graduate from the institution.

Brezavšček, Bach, and Baggia (2017) successfully used Markov chain models to investigate the pattern of students' enrollment and academic performance at a Slovenian institution of higher education. The model contained five transient or nonabsorbing states and two absorbing states. The authors used student records for a total of eight consecutive academic seasons, and estimated the students' progression toward the next stage of the program. From those transition percentages they were able to obtain progression, graduation, and withdrawal probabilities.

As mentioned earlier, most Markov chain models involving enrollment management and prediction use student classification to create the various states of the model. Using student classification in model specification, however, could create states that are overly broad in nature since, at most semester-based colleges and universities, student classification varies by 30 hours.

Ewell (1985), who also used Markov chains to predict college enrollments, noted two limitations of the models. First, because the estimation of the probabilities rests on historical data, Markov chains may be sensitive to when the data were collected. This could be especially true with significant enrollment gains or declines from one year to the next. Second, according to Ewell, different subpopulations may behave in different ways, thus necessitating the need to disaggregate into smaller groupings.

However, the Markov chain's attributes may allow a unique ability to detect the leaks and bulges. Because this type of projection model uses the stochastic process to describe a sequence of events in which the probability of each event depends only on the state attained in the previous event, changes to student flow are immediate and are not subject to potentially skewed results of the past. In short, the limitations mentioned by Ewell (1985) can be utilized when building the student flow matrices to detect significant shifts in enrollment and to determine which groups of students are leaving the institution at a higher rate.

METHODOLOGY

The current study used Markov chains to predict Fall enrollment at a Southeastern, masters-level (Larger Programs) public institution based on annual Fall semester enrollment for degree-seeking undergraduates. The process involved obtaining data from the institution's student information system and separating students into groupings based on their cumulative SCHs earned. Student flow was measured from Fall of year *i* to Fall of year *i+1* based on whether students stayed within their credit hour category, moved into another credit hour category, or did not enroll at the institution. These student flow changes for each category were then summed and applied to year *i+2* as a prediction of enrollment.

Within the model, at a given point in time each student has a particular state, and each student is treated as having a particular probability of transitioning to each other state or staying within the same state. Most of these states are based on the number of SCH the student has accumulated (i.e., the SCH category). Because the SCH category of a student was determined by the number of cumulative SCHs a student earned, most of the credit hour flow scenarios included students advancing to a higher credit hour category or students withdrawing or graduating. While it is rare for a student to move from a particular credit hour category to a lower category, it can happen through the transfer process when, after the student has enrolled, the current institution does not accept certain SCHs from the former institution.

The characteristic that makes this model a Markov chain is the fact that a given student's transition probabilities between states are assumed to depend only on that student's current state and not on any of the student's previous states. This is a simplifying assumption that allows all students within a given state to be treated similarly regardless of their histories. Otherwise, the model would become much more complicated and difficult to apply.

The main parameters of the model are estimates of these transition probabilities. These transition probabilities are estimated by calculating the fractions of students that transitioned from each state to each other state relative to the number of students initially in that state in past years' enrollment data. The other parameters of the model are the fractions of new incoming students by credit hour category. The total number of new incoming students is assumed to be fixed, thus the estimated number of incoming students by credit hour category follows from these fractions.

The model process is recursive in that predictions for Fall X are produced from the enrollment data from Fall X-2 and Fall X-1 and the subsequent flow rates from Fall X-2 to Fall X-1.

We can now describe the basic assumptions that we used to construct the predictive models:

1| Each model models flow from one year to the next and is named accordingly. For example, Fall 2013 to Fall 2014 is known as the 13_14 Model and is based on the starting data for Fall 2013 and the new student data from Fall 2014.

2| As the model is applied, the output headcount by SCH level for the *(i+1)*th year becomes the input headcount for the next iteration of the model.

3| When the model is applied to a future year, the total number of new students is assumed to be constant and the same as the number of new students for the *(i+1)*th year. The distribution of new students by SCH level is also assumed to be constant.

4| When the model is applied to a future year, it is assumed that the fractional student loss and fractional student continuation ratios are fixed by SCH level.

5| When the model is applied to a future year it is assumed that the fractional flow from SCH level to SCH level is the same as for the year used to construct the model.

The model divides the undergraduates into 24 6-SCH groupings. This method uses historic ratios of SCH student subsets gathered from the student information system to predict future Fall headcounts.

The 6-SCH groupings used in this model are individually less broad than the more familiar student classification levels. However, it is possible to aggregate the 6-SCH bins into a version of these student levels, which we define as

- Freshmen: ≤30 SCH
- Sophomore: >30 SCH and ≤60 SCH
- Junior: >60 SCH and ≤90 SCH
- Senior: >90 SCH

Note that these classification-level definitions do not exactly match the institution's definitions. In using SCH groupings, the enrollment pipeline may be much more finely observed and enrollment patterns among students may be more precisely distinguished. While it is the goal of this study to develop a model to predict the coming Fall enrollment once the previous Fall enrollment is known, the model will not address enrollment by major, academic department, or college.

MODEL DESCRIPTION

The student information system parsed out students into the various SCH categories based on the predetermined groupings. These students were then tracked during the following Fall semester to determine student flow percentages. Within this study, student flow states are defined as:

1| students in credit hour group *j* who stayed within that group,

2| students in credit hour group *j* who moved to a different credit hour group,

3| students in other credit hour groups who moved to group *j*, and

4| students who were no longer enrolled at the institution.

Within this model, the following terms and symbols are used:

1| n is the number of SCH levels in the model (n = 24 for the 6-SCH groupings).

2| h_{ij} is the ith Fall semester headcount for the jth SCH level.

3| H_i is the total undergraduate headcount for the ith semester.

4| l_{ij} is the number of the h_{ij} subset students not enrolled the next Fall semester.

5| L_i is the total number of undergraduates enrolled in the ith Fall semester that are not enrolled in the $(i+1)$th Fall semester.

6| $c_{ij} = h_{ij} - l_{ij}$ is number of continuing students in the jth SCH level.

7| C_i is the total number of undergraduates that enrolled in the ith Fall semester that are also enrolled in the $(i+1)$th Fall semester.

8| d_{ijk} is the number of the continuing c_{ij} subset students that move from SCH level j to SCH level k from the ith Fall to the $(i+1)$th Fall.

9| w_{ij} is the number of the C_i subset students that flow from all other levels into level j.

10| o_{ij} is the number of the c_{ij} subset students that flow out of level j into all other levels.

11| $s_{(i+1)j}$ is the number of the new incoming students for the $(i+1)$th Fall semester where j is the SCH level.

12| $N_{(i+1)}$ is the total number of incoming new undergraduate students for the $(i+1)$th semester.

With this terminology in place, the previously stated assumptions of the models can now be described algebraically:

1| When applying a model to a future period from Fall ($i+1$) to Fall ($i+2$), the total number of incoming students is assumed to be the same as it was for the period used to build the model, so it is assumed to have the value N_{i+1}. The fraction of new students by SCH level for that upcoming year is also assumed to be the same as it was in the period used to train the models, so each is assumed to be $s_{(i+1)j}/N_{i+1}$. Therefore, the estimated number of new students for a particular SCH level in that future year can be obtained by multiplying the value of this fraction by the estimated total number of students in the current year. That is, the estimate for the number of new students in the future year for that particular SCH level is given by $s_{(i+1)j}/N_{i+1} \times N_{i+1} = s_{(i+1)j}$.

2| The fractional loss and fractional continuation ratios are also assumed to be fixed by SCH level. In other words, for a future year these ratios are assumed to be l_{ij}/h_{ij} and c_{ij}/h_{ij}, the same as they were in the year used to build the model. Therefore, for the upcoming future period from Fall $(i+1)$ to Fall $(i+2)$, the estimated number of lost and continuing students for the jth SCH level are obtained by multiplying these ratios by the number of students $h_{(i+1)j}$ in that SCH level in the current Fall $(i+1)$. This multiplication is $l_{ij}/h_{ij} \times h_{(i+1)j}$ to estimate lost students in the jth SCH level and $c_{ij}/h_{ij} \times h_{(i+1)j}$ to estimate continuing students in the jth SCH level.

3| Finally, the fractional flow from a particular SCH level to another SCH level is assumed to be fixed. In other words, for a future year these ratios are assumed to be d_{ijk}/c_{ij}, the same as they were in the year used to build the model. Therefore, for the upcoming future period from Fall $(i+1)$ to Fall $(i+2)$, the estimated number of students transitioning from SCH level j to SCH level k is given by the value of this ratio d_{ijk}/c_{ij} multiplied by the estimated number of continuing students in the jth SCH level.

The processes described above can be applied iteratively to obtain estimates for years even farther into the future by using the estimated values from one iteration as inputs into the next iteration.

Using the terms and formulae, we created a spreadsheet matrix (Table 1) that includes the various credit hour classifications as well as the nonabsorbed transient student states and the absorbed state of no longer enrolled.

Table 1. Basic Structure Matrix of the Markov Chain Model

SCH Levels	Fall *i* Headcount	Lost	Continuing			Movement from SCH Level 1	2	.	.	*n*	Static	Inflow	Outflow	Fall *i*+1, New Students	Fall *i*+1 Headcount
1	h_{i1}	l_{i1}	c_{i1}	Movement to SCH Level	1	d_{i11}					d_{i11}	w_{i1}	o_{i1}	$s_{(i+1)1}$	$h_{(i+1)1}$
2	h_{i2}	l_{i2}	c_{i2}		2	d_{i12}	d_{i22}				d_{i22}	w_{i2}	o_{i2}	$s_{(i+1)2}$	$h_{(i+1)2}$
.	.	.	.		.	.	.	.	.		.	.	.	.	.
.	.	.	.		.	.	.	.	.		.	.	.	.	.
n	h_{in}	l_{in}	c_{in}		*n*	d_{i1n}	.	.	.	d_{inn}	d_{inn}	w_{in}	o_{in}	$s_{(i+1)n}$	$h_{(i+1)n}$

Note: This table shows the basic structure matrix of the headcount SCH flow associated with the Markov chain model that connects the undergraduate headcount in the ith Fall to the headcount in the *(i+1)*th Fall.

From this SCH flow structure, we can observe the relationships of credit hour flow between and among the various states, including flow into nonabsorbing states (staying or moving into another credit hour state) or into absorbing states (not enrolling at the institution). The relationships among the variables are as follows:

1| $c_{ij} = \sum_{k=1}^{n} d_{ijk}$ represents those current students who were in SCH level *j* who stayed at the institution.

2| $o_{ij} = \sum_{\substack{k=1 \\ k\neq j}}^{n} d_{ijk}$ represents those current students who were in SCH level *j* who moved to all other SCH levels.

3| $w_{ik} = \sum_{\substack{j=1 \\ j\neq k}}^{n} d_{ijk}$ represents those current students who were in SCH levels other than *k* who moved to SCH level *k*.

4| $H_i = \sum_{j=1}^{n} h_{ij}$ represents semester headcount at Fall semester *i*.

5| $L_i = \sum_{j=1}^{n} l_{ij}$ represents those students at Fall semester *i* who did not reenroll.

6| $C_i = \sum_{j=1}^{n} c_{ij}$ represents those students at Fall semester *i* who did reenroll.

The following relationship,

$$\sum_{k=1}^{n} w_{ik} = \sum_{j=1}^{n} o_{ij}$$

shows two equivalent ways of expressing the collection of students who remain at the institution and move from any SCH level to a different SCH level during the year. Conservation of student flow is obtained only when students from level *j* stay in SCH

Table 2. Annual Enrollment Data, Fall 2010–Fall 2017

Fall *i*	Fall *i* Headcount	Lost	Continuing	New	Fall (*i*+1) Headcount
Fall 2010	9,652	3,773	5,879	3,957	9,836
Fall 2011	9,836	4,082	5,754	3,721	9,475
Fall 2012	9,475	3,965	5,510	3,761	9,271
Fall 2013	9,271	3,843	5,428	3,574	9,002
Fall 2014	9,002	3,685	5,317	3,598	8,915
Fall 2015	8,915	3,792	5,123	3,993	9,116
Fall 2016	9,116	3,945	5,171	3,919	9,090
Fall 2017	9,090	not known	not known	not known	not known

level j or move to other SCH levels, or when students from other SCH levels move into SCH level j.

Given these relationships, the number of undergraduates by level in the second Fall semester can be calculated using the following formula:

$$h_{(i+1)j} = h_{ij} - \ell_{ij} - o_{ij} + w_{ij} + s_{(i+1)j}$$

This is the number of total transient students in one of the SCH levels after 1 year who were not absorbed by withdrawing or graduating. Therefore, the total number of students in the *(i+1)*th Fall semester is simply given by

$$H_{i+1} = H_i - L_i - N_{i+1}$$

since the inflow and outflow terms cancel upon summation.

RESULTS

The model used actual data from a Southeastern, masters-level (Large Programs) public institution for Fall 2010 through Fall 2017. The enrollments for these 8 years are displayed in Table 2.

In developing the Markov chain matrix for each year, the total number of students within each category were noted and tracked to the following year. Within this matrix, one can observe the various student states by each category to determine who is moving into transitional (nonabsorbing) states and who is graduating or not returning. These more-granular data within the matrix offer clues as to when students may be leaving the institution and where there are potential bulges in the system coming from new or transfer students.

Table 3. Fall 2016 to Fall 2017 6-SCH Matrix

SCH Level Groupings	SCH Level Definition	HC1 (Fall 2016)	Lost	Continuing	GradA16	GradA16E	GradB16	GradB16E
1	(0–6)	1589	597	992	0	0	0	0
2	(7–12)	264	81	183	0	0	0	0
3	(13–18)	245	110	135	1	0	0	0
4	(19–24)	322	141	181	1	1	0	0
5	(25–30)	501	147	354	11	9	0	0
6	(31–36)	386	121	265	17	13	0	0
7	(37–42)	332	108	224	13	8	0	0
8	(43–48)	316	94	222	16	13	0	0
9	(49–54)	291	92	199	19	17	0	0
10	(55–60)	399	114	285	31	21	0	0
11	(61–66)	446	105	341	37	27	0	0
12	(67–72)	369	85	284	16	13	1	0
13	(73–78)	346	84	262	18	12	14	1
14	(79–84)	347	121	226	13	12	34	1
15	(85–90)	347	165	182	9	5	100	0
16	(91–96)	400	248	152	11	6	175	0
17	(97–102)	341	204	137	8	6	147	1
18	(103–108)	323	230	93	5	1	188	5
19	(109–114)	271	188	83	3	3	159	2
20	(115–120)	237	171	66	5	5	135	3
21	(121–126)	179	131	48	3	2	98	3
22	(127–132)	154	116	38	1	1	90	2
23	(133–138)	135	89	46	1	1	56	3
24	(139–144)	113	79	34	0	0	48	1
25	(145–150)	96	78	18	2	0	56	0
26	(151–156)	61	45	16	4	2	32	0
27	(157–162)	49	25	24	1	1	21	1
28	(>162)	257	176	81	4	4	124	2

Movement to SCH Level Number	Movement from SCH Level Number																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1																												
2	14																											
3	60	5																										
4	158	6	5																									
5	408	10	9	7		1																						
6	258	42	18	13	11																							
7	60	85	35	31	19	5																						
8	15	21	45	58	57	19	9																					
9	3	5	13	46	103	31	17	11		1																		
10		3	4	16	115	82	33	17	10																			
11	1	1		4	38	89	59	39	13	9																		
12				1	5	24	70	43	27	21	9																	
13					3	6	19	69	49	26	23	12																
14						1	10	29	56	58	28	13	7															
15						1	1	8	31	84	79	25	18	8														
16								3	7	71	123	79	34	17	2													
17							1			9	53	85	64	17	10	5				1								
18									1	5	20	46	79	84	28	13	5											
19											4	18	45	62	52	23	16	8										
20									1			4	5	28	60	49	19	10	6									
21													5	8	18	38	44	17	11	7								
22													2	1	8	16	27	18	11	12	3							
23															1	1	16	25	23	8	6	4						
24													1			2	4	12	22	21	6	4	5					
25													1				2		4	10	14	4	6	5				
26																			1	4	10	10	9	5	3			
27																			2	1	6	8	6	3	1			
28																	1				1	4	18	20	13	16	24	

SCH Level Groupings	Static	Inflow to	Outflow from	New	NewUnder30Hrs	Transfer	HC2 (Fall 2017)
1	15	0	977	1606	1606	32	1621
2	5	14	178	239	239	23	258
3	6	65	129	195	195	46	266
4	5	169	176	161	161	37	335
5	3	435	351	130	113	60	568
6	6	342	259	116	0	57	464
7	5	235	219	103	0	45	343
8	3	224	219	79	0	32	306
9	4	230	195	85	0	38	319
10	1	280	284	116	0	74	397
11	2	253	339	141	0	103	396
12	2	200	282	105	0	63	307
13	1	207	261	107	0	56	315
14	1	202	225	96	0	45	299
15	3	255	179	77	0	25	335
16	5	336	147	87	0	24	428
17	3	245	134	62	0	21	310
18	3	281	90	46	0	19	330
19	3	228	80	25	0	7	256
20	2	182	64	50	0	19	234
21	2	148	46	45	0	17	195
22	4	98	34	41	0	12	143
23	2	84	44	32	0	12	118
24	1	77	33	20	0	2	98
25	1	46	17	22	0	5	69
26	0	42	16	14	0	2	56
27	0	27	24	10	0	4	37
28	81	97	0	109	0	34	287

Table 3 represents one such matrix, the 6-SCH matrix from Fall 2016 to Fall 2017. The 28 6-SCH groupings are labeled down the left with the same corresponding 28 groupings across the center of the matrix. This table also contains headcount by groupings, how many within each grouping did not return, how many graduated, and how many new students enrolled in Fall 2017 but not Fall 2016. Matrices such as this one can be examined to identify the aforementioned leaks and bulges in the enrollment pipeline.

The following labels are used in Table 3:

1| HC1 (Fall 2016): Fall 2016 census undergraduate enrollment excluding special groups.

2| Lost: Enrolled Fall 2016 but not in Fall 2017. This includes students that graduated without reenrolling, as a subset. When determining if the student returned in Fall 2017, only undergraduate students, excluding special groups, were considered.

3| Continuing: Enrolled in Fall 2016 and Fall 2017.

4| GradA16: Awarded an associate degree in Fall, Spring, or Summer of Academic Year 2016–17. Note that only one degree is counted per student to avoid double-counting, with bachelor's degrees given precedence over associate's degrees.

5| GradA16E: Awarded an associate's degree and enrolled in next Fall term in another degree program. These students are a subset of GradA16.

6| GradB16: Awarded a bachelor's degree in Fall, Spring, or Summer of Academic Year 2016–17.

7| GradB16E: Awarded a bachelor's degree and enrolled in next Fall term in another degree program. These students are a subset of GradB16.

8| Columns in the center indicate movement of continuing students from the Fall 2016 SCH categories to the Fall 2017 SCH categories. Note that the central portion of Table 3 does not include counts for students who enrolled both semesters but remained in the same SCH level; these counts are instead separately labeled Static.

9| Static: Enrolled in Fall 2016 and Fall 2017 and stayed in the same SCH level.

10| Inflow to: Enrolled in Fall 2016 within a different SCH level but moved to the current SCH level in Fall 2017.

11| Outflow from: Enrolled in the SCH level during Fall 2016 but moved to another SCH level in Fall 2017.

12| New: Enrolled in Fall 2017 but did not enroll in Fall 2016. (NewUnder30Hrs and Transfer are subsets of New.)

13| NewUnder30Hrs: New students with fewer than 30 hours.

14| Transfer: Transfer students.

15| HC2: Fall 2017 census undergraduate enrollment excluding special groups.

According to the table, in Fall 2016 there were 1,589 students in the (0–6) SCH group. Out of these, 597 did not return the next Fall semester. A total of 408 of these students transitioned into the (25–30) SCH group, indicating that they were progressing normally, while 232 transitioned into groups of 24 or fewer SCH. With a quick examination of the flow, it is easy to see that the majority of students are not returning within the SCH groupings that make up the freshman and sophomore years as denoted in the Lost column. In the (85–90) SCH grouping, 109 students graduated, and 5 of the students who graduated reenrolled in Fall 2017, meaning

Table 4. Actual Enrollment and Predictions, Fall 2012–Fall 2017

Model		Fall 12	Fall 13	Fall 14	Fall 15	Fall 16	Fall 17
Reality	Actual Headcounts	9,475	9,271	9,002	8,915	9,116	9,090
Model 10_11 6SCH	Predicted Headcounts % Diff. from Actual	9,948 4.99%	9,999 7.85%	10,002 11.11%			
Model 11_12 6SCH	Predicted Headcounts % Diff. from Actual		9,244 –0.29%	9,076 0.82%	8,958 0.48%		
Model 12_13 6SCH	Predicted Headcounts % Diff. from Actual			9,105 1.14%	8,980 0.73%	8,903 –2.34%	
Model 13_14 6SCH	Predicted Headcounts % Diff. from Actual				8,839 –0.85%	8,745 –4.07%	8,694 –4.36%
Model 14_15 6SCH	Predicted Headcounts % Diff. from Actual					8,874 –2.65%	8,865 –2.48%
Model 15_16 6SCH	Predicted Headcounts % Diff. from Actual						9,258 1.85%

Note: The model creates predictions for the next 3 years (when actual data are available for comparison) for each of the models using the 6-SCH methods.

that 104 of the students who graduated did not reenroll. A total of 165 students in the (85–90) SCH grouping were lost (did not reenroll); subtracting the aforementioned 104 students leaves 61 students who neither graduated nor reenrolled.

A total of 914 new transfer students entered for Fall 2017, indicating a significant number of students who took some type of transfer credit. Many of these new transfers could constitute dual-enrolled students who took both high school and college classes. The bulk of the new transfer students, however, are entering with more than 54 and fewer than 84 SCHs.

In observing the higher groupings, the table indicates that 865 students had accumulated more than 126 SCH and 448 (52%) graduated. Of the students who earned more than 126 SCHs, 608 did not reenroll in the institution.

While this table represents only one of the six matrices created for this study, the possibilities of tracking student flow by groupings, classifications, or years are numerous. Moreover, it can be argued that the process of tracking student flow through transitional states within the Markov process is somewhat intuitive and indicative of the strong predictive properties of the model.

Table 4 shows the predictions for the next 3 years, along with the actual data. The model was built using the flow of students over a particular academic year. There were six such academic years used for construction of the models. The columns of Table 4 show the years for which an enrollment prediction was generated. As can be seen in the table, predictions for the 10_11 Model for both methods were overspecified by about 5% for Fall 2012 and about 11% for Fall 2014. The 11_12 Models produced better projections, coming within less than 1% of the actual values for all 3 years. The prediction of the 12_13 Model differed from the actual enrollment by an average of –0.2%. Results from the 13_14 Model indicate that the prediction differed by an average of 3.1%. In most cases, predictions farther into the future from the years used to train the models have greater residuals, which is to be expected in any forecasting problem.

We calculated averages of the absolute values of the percentage differences between the actual and predicted values for enrollment using the actual and predicted enrollment from Table 4. The percentage difference between the predicted and actual value is defined as

$$\textit{\% difference} = \frac{\textit{predicted value - actual value}}{\textit{actual value}} \times 100\%$$

We can examine the predictive ability of the models by using the average value of the absolute values of these percentage differences, because these values show on average how far off the models were, regardless of sign. In a mathematical sense, the absolute value between two numbers is known as the standard Euclidean distance between two points and indicates the real distance between two numbers (Bartle & Sherbert, 2011). The results as shown in Table 5 clearly indicate that the predictive ability of the model decreases as number of years out from the years used to build the model increases, which is expected, similar to how weather forecasts become less accurate the farther they go into the future.

Table 5. Mean Absolute Value of Percent Differences by Years Out for 6-SCH Models

Prediction Time Frame	Mean Absolute Value of Percent Difference
1 year out	1.96%
2 years out	3.19%
3 years out	4.57%

Based on the results from Table 5, the study will examine only 1-year-out predictions, because these were the most accurate. The actual values are compared with those 1-year-out predictions in Table 6. The predicted enrollment for Fall X in Table 6 is produced from the enrollment data from Fall X-2 and Fall X-1 and subsequent flow rates from Fall X-2 to Fall X-1.

Note that the 6-year average of the absolute values of the percentage differences by class range from 2.8% to 4.7%. The 2016 freshman percent difference of –12.9% represents an outlier due to a major university initiative to increase new freshmen enrollment. This influx of new freshmen was significantly different from past years and clearly signals the bulge in the student flow pipeline as mentioned above. By utilizing the iterative process of producing Fall X projections from the enrollment data from Fall X-2 and subsequent flow rates from Fall X-2 to Fall X-1, the effect of this bulge in the system can be tracked into the future to plan upcoming course offerings.

Table 6. The 6-SCH Models' 1-Year-Out Predictions Compared to Actual Enrollment, 2012–17

		Freshman	Sophomore	Juniors	Seniors	All Levels	Mean Absolute % Difference of Class Levels
2012	Actual	2,876	2,035	1,871	2,693	9,475	
	Predicted	3,114	2,090	1,966	2,778	9,948	
	% Difference	8.28%	2.68%	5.10%	3.17%	5.00%	4.81%
2013	Actual	2,729	1,890	1,870	2,782	9,271	
	Predicted	2,817	1,875	1,834	2,718	9,244	
	% Difference	3.23%	–0.79%	–1.92%	–2.30%	–0.29%	2.06%
2014	Actual	2,644	1,803	1,870	2,685	9,002	
	Predicted	2,709	1,800	1,789	2,807	9,105	
	% Difference	2.47%	–0.16%	–4.36%	4.53%	1.14%	2.88%
2015	Actual	2,533	1,944	1,738	2,700	8,915	
	Predicted	2,574	1,809	1,816	2,640	8,839	
	% Difference	1.60%	–6.93%	4.51%	–2.24%	–0.85%	3.82%
2016	Actual	2,921	1,724	1,855	2,616	9,116	
	Predicted	2,543	1,885	1,801	2,644	8,874	
	% Difference	–12.93%	9.36%	–2.89%	1.07%	–2.65%	6.56%
2017	Actual	3,048	1,829	1,652	2,561	9,090	
	Predicted	3,053	1,788	1,766	2,651	9,258	
	% Difference	0.17%	–2.24%	6.91%	3.51%	1.85%	3.21%
	Mean Absolute % Difference	4.78%	3.69%	4.28%	2.80%	1.96%	

By observing the predictive capabilities of the model, it is easy to see how administrators and enrollment managers can use these results to plan for classes and instructional personnel. Here, both annual projections and classification average projections for the 5-year period were off by no more than 6.6%, which should fall within the margin of error for most larger institutions.

Furthermore, Monte Carlo simulation could be used to obtain enrollment predictions that give a range of plausible values instead of a single point estimate for a future year's enrollment. Monte Carlo simulations have been used in the context of higher education by Torres, Crichigno, and Sanchez (2018) to examine degree plans for potential bottlenecks. In applying these methods to this enrollment model, the fractions of students transitioning between specific levels would be treated more like the result of many coin flips than as fixed fractional values, and the ranges of predicted values could be obtained by repeated random simulation. This level of simulation was not performed in this study.

CONCLUSIONS

The use of Markov chains in projecting enrollment and the management thereof has gained popularity among professionals in higher education. The short-term projections created by this stochastic process are unique to other time-tested forecasting tools used in enrollment management. When used properly, Markov chains can aid institutions in determining progression of students that are different from more-traditional ARIMA and regression prediction tools in that:

1| they can give accurate enrollment predictions with only 2 previous years' data, which can be helpful when large longitudinal databases are not available;

2| they can be used to generate predictions on segments of a group of students rather than the entire population, which may be required for other models; and

3| the almost intuitive nature of the Markov chain lends well to changes in student flow characteristics, which often cannot be explained by a complex statistical formula.

By creating groupings and tracking students within those groupings by the state they transition into, the researcher can also get a better picture of what type of students are leaving and when they are leaving.

As shown in this study, the strong predictability of Markov chains allows administrators to better plan course scheduling and instructor demand while managing tight budgets. In this study, several predictive headcount models were developed using SCH flow as the annual driver. Eight years of Fall enrollment data from the institution were used to develop the models. When applied to historical data each gives 1-year-out predictions within a calculated level of uncertainty. The models can easily be modified to change the new student input data, the continuation rates, and the interlevel flow rates, should that be desired. Furthermore, similar models could be used to track Fall to Spring retention as well as Spring to Fall retention.

REFERENCES

Bartle, R. G., & Sherbert, D. R. (2011). *Introduction to real analysis.* Urbana-Champaign: University of Illinois.

Borden, V. M. H., & Dalphin J. F. (1998). Simulating the effect of student profile changes on retention and graduation rates: A Markov chain analysis. Paper presented at the Annual Forum of the Association for Institutional Research, Jacksonville, FL, May 19.

Brezavšček, A., Bach, M., & Baggia, A. (2017). Markov analysis of students' performance and academic progress in higher education. *Organizacija*, 50(2), 83–95.

Clagett, C. A. (1991). *Institutional research: The key to successful enrollment management.* Office of Institutional Research. Largo, MD: Prince George's Community College.

Coomes, M. D. (2000). The historical roots of enrollment management. In M. D. Coomes (Ed.), *The role student aid plays in enrollment management: New directions for student services*, 89 (5–18). San Francisco: Jossey-Bass.

Day, J. H. (1997). Enrollment forecasting and revenue implications for private colleges and universities. In D. T. Layzell (Ed.), *Forecasting and managing enrollment and revenue: An overview of current trends, issues, and methods: New directions for institutional research*, 93 (51–65). San Francisco: Jossey-Bass.

Ewell, P. T. (1985). Recruitment, retention and student flow: A comprehensive approach to enrollment management. National Center for Higher Education Management Systems, Monograph 7, Boulder, CO.

Fallows, J., & Ganeshananthan, V. (2004, October). The big picture. *The Atlantic.* https://www.theatlantic.com/magazine/archive/2004/10/the-big-picture/303520/

Gagne, L. (2015). Modeling the progress and retention of international students using Markov chains. Honors Research Projects 3, University of Akron, Akron, OH.

Herron, J. (1988). Universities and the myth of cultural decline. Detroit, MI: Wayne State University Press.

Hopkins, D. S. P., & Massy, W. F. (1981). *Planning models for higher education.* Stanford, CA: Stanford University Press.

Hossler, D. (1984). *Enrollment management: An integrated approach.* New York: College Board.

Johnson, A. L. (2000). The evolution of enrollment management: A historical perspective. *Journal of College Admission, 166*, 4–11.

Liljegren, A., & Saks, M. (Eds.). (2017). *Professions and metaphors: Understanding professions in society.* New York: Routledge, Taylor & Francis Group.

Luna, A. (1999). Using a matrix model for enrollment management. *Planning for Higher Education*, 27(3), 19–31.

Oliver, R. M. (1968). Models for predicting gross enrollments at the University of California. Ford Foundation Program for Research in University Administration, Berkeley, CA.

Pierre, C., & Silver, C. (2016). Using a Markov chain model to understand the behavior of student retention. In N. Callaos, H. Chu, J. Ferrer, S. Fernandes, & B. Belkis Sánchez (Eds.), *The 7th International Multi-Conference on Complexity, Informatics, and Cybernetics/The 7th International Conference on Society and Information Technologies: Proceedings* (pp. 248–251). Winter Garden, FL: International Institute of Informatics and Systemics.

Seltzer, R. (2017, July 24). State funding cuts matter. *Inside Higher Ed.* https://www.insidehighered.com/news/2017/07/24/new-study-attempts-show-how-much-state-funding-cuts-push-tuition

Torraco, R. J., & Hamilton, D. W. (2013). The leaking U.S. educational pipeline and its implications for the future. *Community College Journal of Research and Practice, 37*(3), 237–241.

Torres, D., Crichigno, J., & Sanchez, C. (2018). Assessing curriculum efficiency through Monte Carlo simulation. *Journal of College Student Retention: Research, Theory & Practice.* https://doi.org/10.1177/1521025118776618

THANK YOU

AIR expresses sincere appreciation for the members who serve as peer reviewers. Authors who submit materials to AIR Publications receive feedback from AIR members. The following individuals reviewed manuscripts submitted to this volume of *The AIR Professional File*.

Sesime Adanu

David Allen

Rebecca Bell

Timothy Chow

Michael Duggan

Sarah Fitzgerald

Maxwell Kwenda

Ming Li

Mary Milikin

Gordon Mills

Joseph Roy

Steve Simpson

Logan Tice

ABOUT *THE AIR PROFESSIONAL FILE*

The AIR Professional File features journal-length articles grounded in relevant literature that synthesize current issues, present new processes or models, or share practical applications related to institutional research. All submissions are peer-reviewed. For more information about AIR Publications, including *The AIR Professional File* and instructions for authors, visit www.airweb.org/collaborate-learn/reports-publications.

Association for Institutional Research
Christine M. Keller, Executive Director and CEO
www.airweb.org

ISSN 2155-7535

MORE ABOUT AIR

The Association for Institutional Research (AIR) represents a worldwide community of institutional research, institutional effectiveness, and other decision-support professionals. Incorporated in 1966, AIR empowers higher education professionals to effectively and ethically use data for better decisions through education and professional development, research, and collective action.

A HOLISTIC APPROACH TO IR

This course provides a foundation for participants to meet and navigate the ever-growing demands for data and information in the current higher education landscape.

IPEDS KEYHOLDER ESSENTIALS: A BEGINNER'S GUIDE

Created for data providers with less than 9 months of experience as keyholders, this course covers basic concepts and definitions.

IPEDS KEYHOLDER EFFICIENCIES: REDUCING THE REPORTING BURDEN

Created for data providers with 10-24 months of experience, this course is designed to enhance your knowledge as an IPEDS Keyholder.

WEBINARS

Subject matter experts provide insight into various aspects of IR, IE, Assessment, and related fields.

AIR FORUM

Learn, connect, and share at the largest gathering of IR professionals in the world, and take advantage of recorded sessions, presentation slides, handouts, and scholarly papers available after the conference.

AIR

Association for
Institutional Research

www.airweb.org