JOURNAL OF
ECONOMICS and FINANCE EDUCATION

JEFE

editor
E. F. Stephenson
MANAGING EDITOR

AEF
ACADEMY of ECONOMICS and FINANCE

Volume 19 FALL 2020 Number 2

Academy of Economics and Finance

Using Cluster Analysis in Financial Services: A Teaching Case

Satish Nargundkar and Rasha Ashraf[1]

ABSTRACT

A teaching case appropriate for students in finance, analytics, or related areas is presented for applying cluster analysis to the financial services industry. Based on a real case (modified for confidentiality), it presents students with a real-life scenario where cluster analysis is the appropriate tool to use for addressing a business problem. A credit marketing problem is presented, and incorporates the idea of predictive modeling within clusters. A top-down approach is taken to teach cluster analysis, with the business application and interpretation of results presented first, followed by the algorithm and other details.

Introduction

The Green City Bank's Vice President for marketing, Lisa Carson, was unhappy with the current efforts at attracting and retaining profitable customers. The models currently in place to predict the likelihood of customer response to offers, customers' risk of default, and customers' overall profitability were not accurate enough in their predictions. She tasked her new analytics team, led by Greg Williams, to rebuild the models to show improved performance.

Greg realized that the models that Green City Bank has been using were built on the entire population of customer data available, and that this is generally not a good strategy. From his experience, he knew that the population is rarely homogenous, and so attributes that are predictive of certain behavior for one segment may not be as predictive (or have a different relationship) in another segment. Using the entire dataset as one large group could also systematically eliminate certain segments from consideration. For instance, young people (in the, say, 18–25 year range) have little or no credit history, so many of the variables that indicate low risk, such as a long credit history (measured by the attribute "Age of Oldest Trade") would automatically work against young people. Separate models should therefore be built for "young" and "mature" groups.

Greg wanted to explore the data a bit more to see if he could find segments of the population that were different from each other, based on a mix of demographic and credit related attributes considered simultaneously, which would better aid the marketing effort. He began by considering only those customers who were over 30 years of age, and then used some key attributes to segment them further. Similar analysis would be done later for the group below 30 years of age.

The Clusters and Their Interpretations

Greg used readily available statistical software to perform K-means cluster analysis, a method of clustering recommended for use with large datasets. The method requires the analyst to segment the data into clusters or subgroups. The clustering technique partitions the dataset into distinct subgroups or clusters, such that observations within each group are similar to each other and observations across groups or clusters are different from each other based on some chosen attributes. The number of clusters (K) is specified by the analyst. Given that the number of clusters in the data is not known in advance, the typical approach is to try various values of K and then pick the best solution. The methodology for choosing the best value of K is discussed below.

[1] Nargundkar: Clinical Professor of Management and Assistant Dean for Professional and Flex MBA, Robinson College of Business, Georgia State University, Atlanta GA 30302. Ashraf: Clinical Associate Professor of Finance, Robinson College of Business, Georgia State University, Atlanta GA 30302.

Greg chose seven different attributes for clustering. He wanted a profile of customers based on a mix of demographic data (age, income, and length of residence) and credit related data (number of cards, balance, utilization, and recency of new credit). The demographic data would give him an idea of the stability and income levels, while the credit data would help paint a picture of the customers' proclivity towards using credit. Based on these variables, he started with a two-cluster solution, then tried a three-cluster solution, and so on, all the way up to 10 clusters. He finally settled on a five-cluster solution as the best one. This decision was based on some statistical indicators (discussed below) as well as the meaningfulness of the interpretations that this solution afforded him. The solution he chose is presented in Table 1.

Table 1: The Five-Cluster Solution

	Cluster					
Attribute	A	B	C	D	E	All Data
Customer Age	38	62	48	44	52	43
Income ($1000s)	48	45	88	73	71	62
Length of Residence (years)	9	12	9	6	9	9
Number of Bankcards	24	2	10	15	8	7
Number of Bankcards with a balance > 0	13	1	3	6	2	3
Bankcard Utilization %	69	6	29	51	7	30
Months since most recent bankcard acquired	11	75	21	15	32	35

For each of the seven attributes chosen, the solution shows the average values of the attributes of all the customers that were assigned to that particular cluster. Thus, for customers in cluster A, the average age was 38, the average income was $48,000, and so forth. The entire column of seven numbers for each cluster (called the *cluster centroid*) was essentially the average profile of people in that cluster, based on the chosen attributes. Greg proceeded to summarize each cluster as follows:

Cluster A–Lower income, hungry for credit, high risk!
Cluster B–Retirees, not interested in more credit.
Cluster C–Country Club set–the richest, with a fair amount of credit usage.
Cluster D–Well-to-do, use credit frequently, likely most profitable.
Cluster E–Well-to-do, less credit usage and likely payoff immediately, not very profitable.

Digging Deeper into the Clusters

Analysis of the customers within each cluster separately showed Greg some interesting insights into their behaviors. As a preliminary step, he noted the number of people who fell into each cluster. He examined past responses to marketing campaigns to see how likely people in each segment were to respond to an offer, and created a "Response Index," defined as the probability of response multiplied by 100. An index of 250 thus means that the probability of response to a marketing campaign is 2.50%. From his experience with banks, Greg knows that the average response to a new product offering is typically around 1.00% (in this dataset, the average for all the clusters combined was roughly 1.27%). Thus a 2.50% response rate is a rather high number. It makes sense in this case, since cluster A represents credit hungry people. He also examined the profitability of the customers over a period. Profitability for each customer over a month was determined by taking the interest and fee income generated by each customer, and subtracting the amount lost due to default, as well as other costs allocated per customer. Table 2 summarizes what he found.

Table 2: Analysis of the Cluster Segments

	Cluster					
	A	B	C	D	E	All Data
Volume (% of population)	18	27	20	25	10	100.00
Response Index (Probability*100)	250	30	110	175	80	127.00
Average Profitability Per Person ($)	-75	5	48	86	-5	18.45

This analysis immediately helped answer some questions about the marketing efforts. It was clear to Greg that people in cluster A were not profitable. Their high credit dependence meant that they were much more likely than others to respond to an offer of credit, but their default rate was high, and they would end up being a losing proposition for the bank. Thus, it would be best to eliminate them from campaigns for new products. Cluster B represented the near polar opposite - these were people of very low risk, but not interested in any more credit. So once again, it would be a wasted effort to market new products to them. Cluster E included people known as "Transactors" – those that pay off their bill in full every month, and are not profitable to the bank, despite posing little risk. Hence, this table indicated to Greg that people in clusters C and D were the best potential targets for new credit products. This information alone could save the bank a significant amount in marketing costs, by targeting the right groups. It also helped tailor the product to the right group. As an illustration, assume a total population of 100 customers. Mailing to all of them with an average profitability of $18.45 per customer during the time period analyzed would yield $1845. On the other hand, mailing to only clusters C and D (45% of the population) would yield a profit of $3110 (weighted average computed as 48*20 + 86*25). This is nearly 70% more profit by targeting the best 45% of the population.

In addition, Greg could now build predictive models within each cluster to predict delinquency risk, and the risk levels could be managed with appropriate interest rates or security deposits, or by withholding credit altogether. Such models within each cluster were likely to do a better job of predicting behavior than a single model across the entire population.

Procedural Details

Several questions arise when working on a cluster analysis procedure. Greg made a note of a few frequently asked questions and his responses to them, to aid new analysts and the management team in better understanding the process.

1. Is there a best way to choose the variables based on which the clustering is done?

This is a business decision. One chooses variables (attributes) based on business domain knowledge and experience. There are some basic rules to follow to ensure that the procedure works well:

a. The variables must be numeric. Categorical variables (gender, for instance) are already segments, so they do not need to be considered as part of this procedure. That kind of segmentation is self-evident, and needs no mathematics.

b. The variables must not be strongly correlated with one another, so that the segments are being created with the help of different attributes, not similar ones with different names. If there are several variables in the dataset with many possible correlations, it may be useful to first perform Principal Components Analysis to identify the underlying factors and use them in the clustering process.

c. The variables must be standardized so that each variable has a mean of 0 and a standard deviation of 1 before performing cluster analysis. This is because the procedure relies on measuring the distances between points – the closest ones to each other across all dimensions end up in the same cluster. If variables are on widely differing scales (age and income, for instance), the one with the larger numbers will dominate unless standardized.

d. Data must be cleaned. It is important to remove outliers from the data, or to cap the values. For instance, for the income variable, it may be reasonable in some cases to take all values above $200,000 and simply recode them as $200,000, thus eliminating the extremely large values that might exist, while still preserving the idea that they are high-income individuals. Also, as with many statistical procedures, missing values must be appropriately handled.

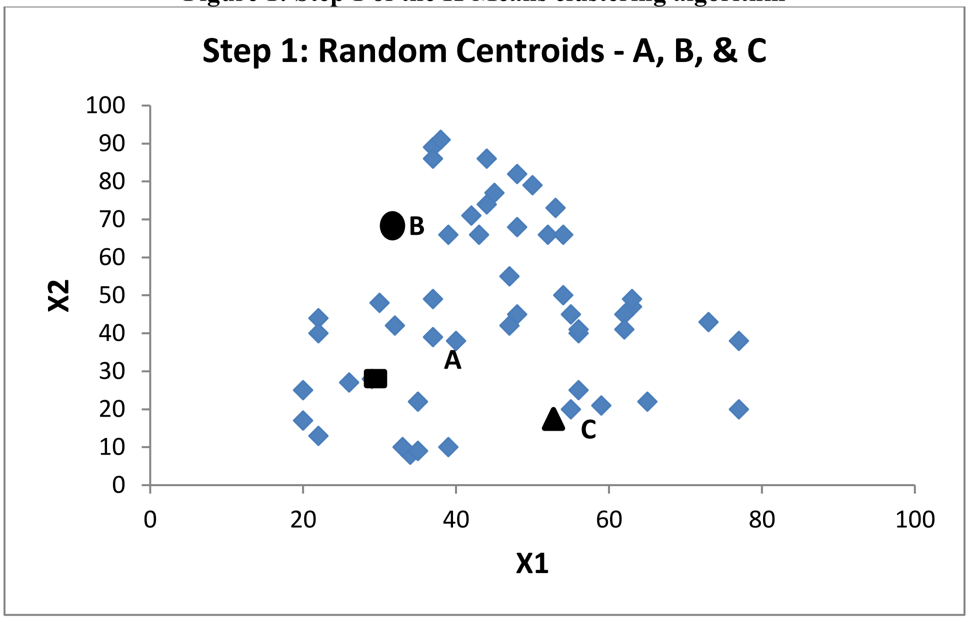
e. Aside from the above basics, however, the choice of the variables themselves is up to the analyst, and depends on the task at hand. Apart from cluster analysis, a group of people, for instance, can be segmented by gender or by age or both. What is the right choice? It depends on what one wants to do with the segments.

2. How does the software determine how to split up the points into the requisite number of clusters? What is the algorithm?

For simplicity, assume there are only two variables, and a three-cluster solution is desired. The cluster procedure uses the following simple algorithm based on Euclidean distances to figure out the cluster memberships:

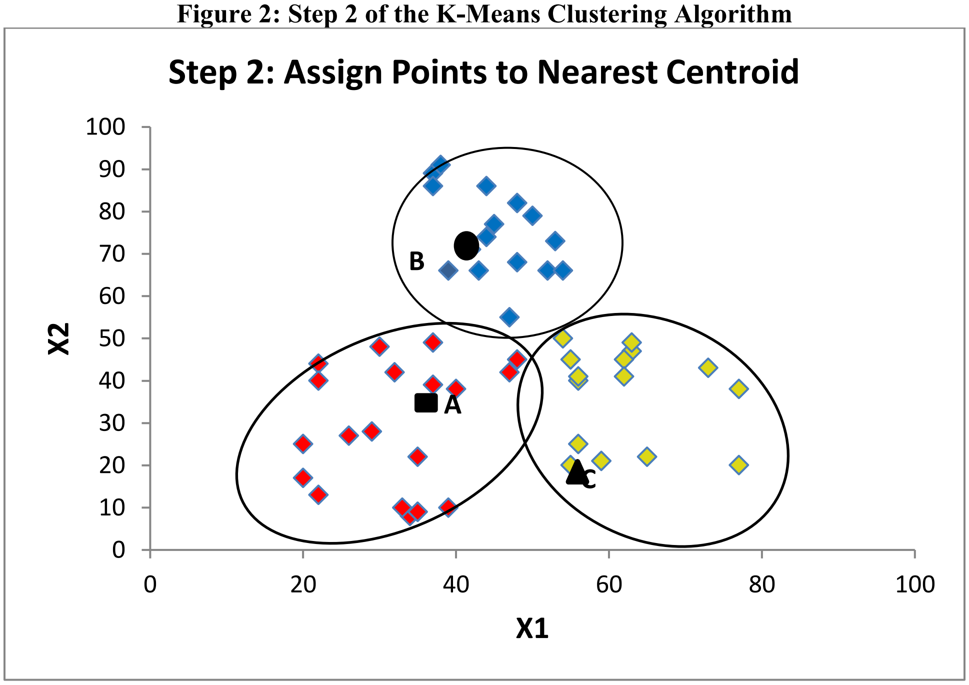
a. Assign any three random points (A, B, C) in the data as the cluster centroids as shown in Figure 1.

Figure 1: Step 1 of the K-Means clustering algorithm



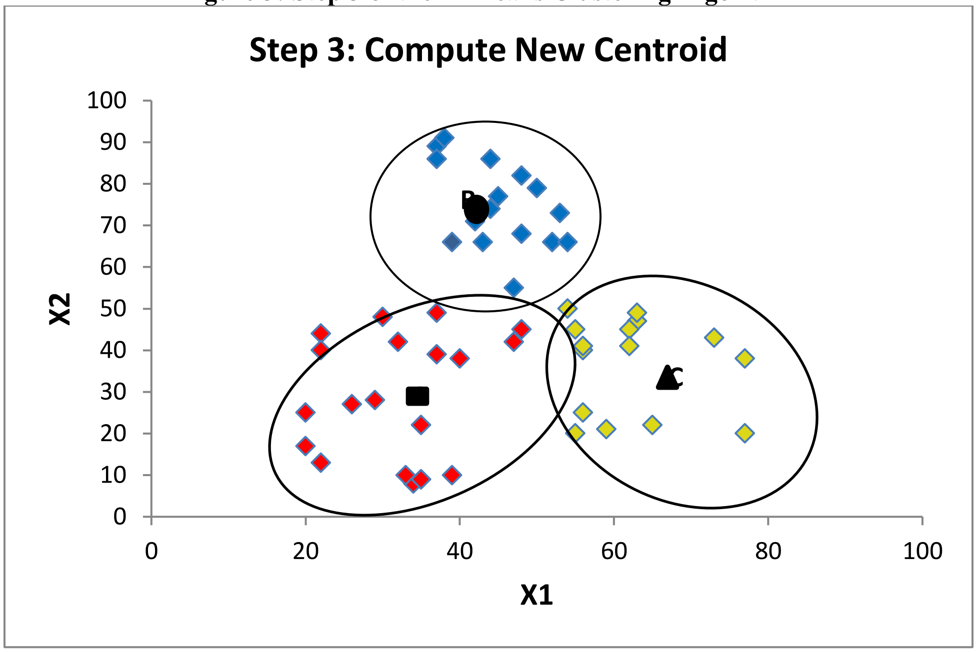
b. Measure the distance from each point to A, B, and C. Assign the point to A, B, or C based on whichever is closest. Figure 2 shows the points color coded based on the centroid to which they are assigned.

Figure 2: Step 2 of the K-Means Clustering Algorithm



c. When all points are thus assigned, take all points assigned to A and compute the average for both variables. The resulting point is the new centroid for cluster A. Do the same for the other two clusters. Figure 3 illustrates the new clusters.

Figure 3: Step 3 of the K-Means Clustering Algorithm

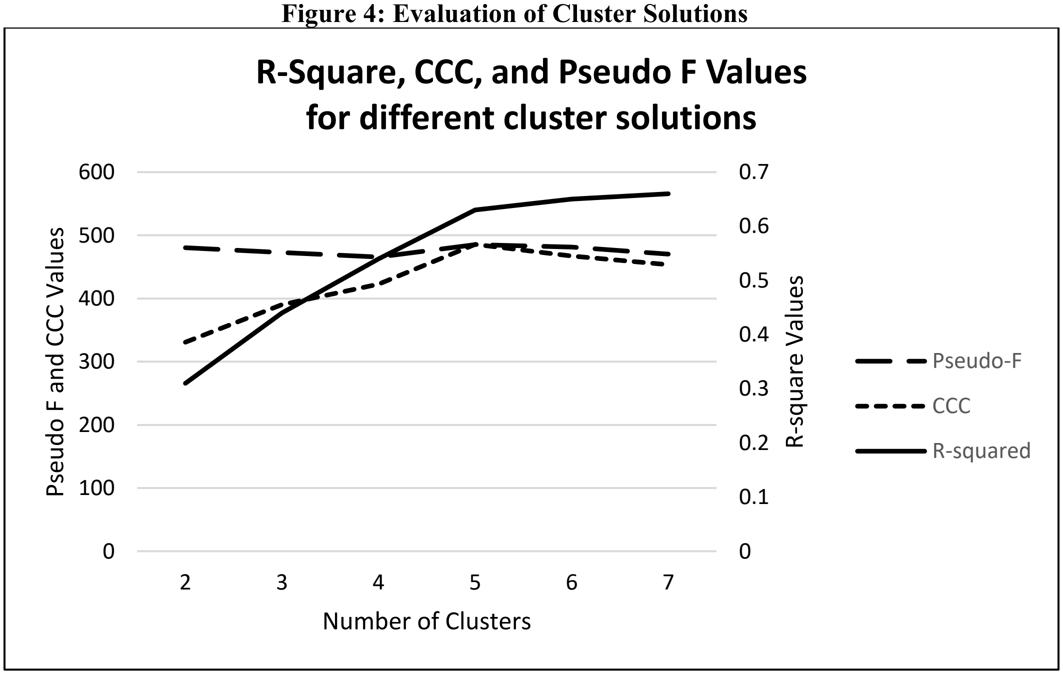


d. If the newly computed centroids in step c) shifted sufficiently from their original positions, go back to step b) and repeat until the centroids stop moving, or you have performed a certain number of iterations.

3. How do you determine the best number of clusters?

There are statistical indicators like R-squared, Pseudo F, and the Cubic Clustering Criterion (CCC) (Milligan and Cooper 1985). As with regression analysis, R-square keeps increasing with each additional cluster, and there is usually a point of diminishing rate of increase. Pseudo F and CCC can fluctuate. Usually, the best solution has relatively high values of all three indicators. Figure 4 below shows the values for these indicators as Greg tried solutions with varying numbers of clusters.

Figure 4: Evaluation of Cluster Solutions



Based on Figure 4, Greg decided that the 5-cluster solution was the best choice. It had the highest values for CCC and Pseudo F, and the R-squared values showed diminishing improvement beyond five clusters.

One should note that unlike regression analysis, none of these measures guarantees an optimal solution. In fact, clustering is a heuristic technique, and optimality is never guaranteed. As an exploratory technique, that is less important than simply being able to interpret the clusters meaningfully. The within-cluster variations should be relatively small (a large variation in a cluster may indicate the need to split further). If one can interpret the clusters in a way that makes sense from a business standpoint, and these interpretations help drive business decisions, you have a good solution.

Instructor's Teaching Notes

Case Synopsis

The case is suitable for an undergraduate or graduate class on data analytics and/or machine learning. The case outlines how the K-means clustering method can be applied to identify subgroups or clusters within a dataset. The clustering method has a wide variety of applications in different business domains. It may be applied as a marketing strategy to segment customers based on their purchasing habits, demographics, and risk profiles. An organization can target marketing efforts based on intended goals for each segment. The case highlights an example of how banks can use clustering to segment customers in order to better target customers and improve overall profitability. Through the case, students get to see how clustering helps in decision making and building the right models for business decisions.

Learning Objectives

The key objectives of this case are to help the students understand how clustering can be applied in the business world to aid decision making in a way that clearly has financial consequences. Specifically, students should be able to:

1. Analyze a business problem and see the need for segmentation.
2. Determine what attributes make sense to use in clustering.
3. Perform cluster analysis with software like SAS or R (we can provide code for either one to interested instructors).
4. Interpret the results.
5. Choose the best cluster solution.
6. Design a business strategy for each cluster, after further modeling within clusters if necessary.

Case Data Source

This case is based on a real consulting project with a bank. However, to protect confidentiality, the names of the bank as well as the people involved have been altered. Likewise, the results table shows realistic but fictitious numbers.

Theoretical Linkages

K-means clustering is an unsupervised learning technique used widely in modern machine learning applications. The data is segmented into *K* distinct non-overlapping groups based on some selected attributes. The data is partitioned in such a way that observations within the same cluster are similar to each other and observations across different clusters are different from each other.

We present the K-means clustering method following James et al. (2013) and it is briefly described below. For a more detailed understanding, see James et al. (2013). The objective of clustering is that within-cluster variation, which is the sum of all pairwise squared Euclidean distances between observations, is as small as possible. If $W(C_k)$ measures the amount by which observations within a cluster differ from each other, then the *K*-means clustering solves the following minimization problem:

$$\underset{C_1,\ldots,C_K}{Min}\left\{\sum_{k=1}^{K} W(C_k)\right\}$$

This means the observations are portioned into K clusters such that the total within-cluster variation, summed over all K clusters, is as small as possible. Here $C_1, \ldots\ldots C_K$ denote cluster sets containing the indices of the observations within each cluster.

Within-cluster variation is the sum of all pairwise squared Euclidean distances between observations and is explained by the following equation:

$$W(C_k) = \frac{1}{N_k} \sum_{i,i' \in C_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2$$

Here N_k denotes the number of observations in the kth cluster, p denotes the number of attributes the clustering is based on, and x_{ij} implies the measure of the j^{th} attribute in the i^{th} observation. The first (inside) summation adds the Euclidean distance between two data points for all feature vectors $j=1, \ldots p$ and the second (outside) summation adds the distances measured in first summation for all pairwise data points (i, i') within the cluster.

The above optimization problem is a very difficult problem, as there are K^n ways to segment the data with n observations into K clusters. The problem becomes computationally challenging if K and n are large. Therefore, K-means clustering method is solved based on a heuristic algorithm, explained in the case, which provides a good solution.

Suggested Teaching Approaches

Students have generally learned better when first presented with the case and the application area and reasons for using cluster analysis are discussed. Following this motivational introduction, they are shown the results and asked to interpret the clusters and come up with business strategies for each cluster. Once they understand the business implications, they can be introduced to the details of variable selection, the clustering procedure, and the selection of the best solution.

The clustering algorithm can be illustrated in class with three tennis balls. Select three students at random, hand them a ball each, and label them A, B, and C. Ask them raise their arms holding the ball above their heads. Ask each of the other students to look at those three balls. Those who are closest to ball A should all raise their hands. You can then visually look at the center of that group and pass ball A to the person at the center. Do the same with balls B and C. A couple of iterations generally suffice to see that the centroids do not move any more, and a final cluster solution is obtained with the students divided into three clusters based on proximity of seating.

Assignment/Discussion Questions

We can provide a data set with 14,042 observations on customers who were approved for auto loans. The dataset has 55 variables that include some demographic information, credit bureau variables, information on the car purchased, and the loan terms. (Alternatively, students may use other datasets that are available online, for instance at *kaggle.com*.) We can also provide the R or SAS code to construct the clusters. R is open-source statistical software widely used in the data analytics and machine learning community. The following is a typical assignment:

1. Use the auto loan dataset provided. Pick any 7 to 10 numeric variables from the data. Justify your selection.
2. Clean the data.
    a. Run correlations to make sure the variables are not too correlated with each other (correlations > .30 are not desirable). If some are, then drop one or more and consider replacing it with a different variable.
    b. Deal with missing data appropriately.

c. Eliminate outliers.
d. Standardize the variables.

3. Perform K-means cluster analysis with 2,3,4, …10 clusters, and then pick the solution that makes the most sense to you, based on our discussions in class. Interpret the solution. What are the characteristics of each cluster? If you were marketing to them, what sorts of offers would you make to each segment?
4. Optional assignment (advanced): Develop separate predictive models for each cluster, and compare the performance of those within the respective cluster with the predictive performance of a single model for the entire dataset.

Solution to the Assignment

The nature of clustering is such that there is no single right answer. What we are looking for is that the student has a rationale for picking the variables, the final solution, and can interpret the clusters meaningfully.

Evidence of Effectiveness

The use of the case and the teaching method outlined above showed an improvement in student learning in a couple of ways. Examination questions related to this topic include asking students to interpret clustering output, discuss the algorithm for K-means clustering, as well as answer questions regarding the general need for clustering and applications in business. Student scores on this material increased from an average of about 79% (n=85) to about 87% (n=63), showing statistically significant improvement over a traditional lecture that began with data collection, then discussed the algorithm and then an application. In addition, more teams (three teams upon using the case and assignment, compared to one team before) included segmentation in their final projects as a result, even though the project requirement was to build predictive models, and segmentation was optional. This was encouraging, even though there was not enough data on the team projects for statistical comparison. Subjectively, student response and participation in the class also showed a marked increase.

References

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning, with Applications in R*. New York: Springer.

Milligan, Glenn W., and Martha C. Cooper. 1985. "An Examination of Procedures for Determining the Number of Clusters in a Data Set." *Psychometrika* 50: 159-179.

Traditional, Online, or Flipped Classes–Which Do Students Prefer?

Anil K. Lal [1]

ABSTRACT

This paper describes the creation of a flipped classroom based on experimentation, student feedback, and an instructor's impressions about learning outcomes. Based on 601 student responses to a survey administered between 2014 and 2018 to lower and upper level economics classes at a four-year public university, this paper finds that the majority of students preferred flipped to traditional as well as to 100% online classes. Other results indicate that students preferred watching lecture videos to reading textbooks or attending in-class lectures, and preferred weekly-structured assignments over daily-structured assignments. However, findings indicate mixed student preferences with respect to flipped mastery.

Introduction

Technological developments have wide implications for various sectors of the economy–how and what we produce and consume. Technical change has not spared even the education sector. As Bishop and Verleger (2013) noted, technical innovations, among other things, have made possible amplification and duplication of information at an extremely low cost. For example, the introduction of the printing press over 500 years ago meant that books or printed material could be produced at substantially lower costs. This development, among other things, led to a system of teaching, where students listened to the lectures in class and worked on problems (or application of their class learning) outside the classroom, called the "traditional" classroom.

Though the basic format of teaching has not changed in centuries, several attempts have been made to make educational opportunities available to those who could not attend classes at universities and colleges. For example, attempts were made to reach out to a larger body of students via printed material through correspondence courses in 1728, through audio lectures broadcast over radios in 1920s, via television in 1950s, and telecourses in 1970s (see, for example, Bower and Hardy 2004). For the most part, these programs were kept separate from programs offered by brick and mortar schools, which required the physical presence of students in the classroom.

The advent of the world wide web and the availability of a variety of devices (particularly, portable devices) with which one can access the internet, is transforming the way we teach and students learn. Specifically, this has been made possible for the following reasons: first, the availability and subsequent development of learning management systems (LMS) such as Blackboard, Canvas, etc., since 1997 (see, for example, Chaubey and Bhattacharya 2015) provided instructors with a portal through which they can disseminate information (post announcements, place files of teaching materials, post grades, send messages to students) and also test students' learning online (through assignments and tests). This led to the creation of "online" classes and, since the 2000s, there has been an explosion of online classes offered by traditional (brick-mortar) universities. These online classes, which overcame physical barriers to learning, are different from earlier distance learning programs. Now students have a choice to complete a degree program through a mix of online and traditional courses to pursue their educational goals. The decision to shift to online was often made by the administration (see, for example, Allgood et al. 2015), ostensibly to work around students' schedules and to increase revenue by charging students a small additional fee for use of online courses. Since creation of an online class requires effort on the part of the instructors, the universities opened centers to train

[1] Professor of Economics, Kelce Undergraduate School of Business, Pittsburg State University, Pittsburg, KS 66762 USA. I am grateful to an anonymous referee, Paul Grimes, June Freund, Mike McKinnis for their suggestions, and to Irene Robinson for editorial assistance.

instructors to create online classes using LMS and gave them financial incentives to do so. Recognizing this trend, the publishing world innovated and made course cartridges (containing PowerPoint slides, quizzes, ebooks, test banks, videos, etc.) available that can seamlessly be uploaded by instructors into their LMS. At the same time, open educational resources that are free and easily accessible have provided a wealth of educational material to create online classes.[2] Thus, universities compete with one another in terms of how many online courses and degrees are offered, and it has become possible for students to complete courses and even degrees without stepping into an actual classroom.

The second development in advancing higher education has been due to the availability of technology to create and share audio and video files. Readily accessible tools enabled Wall Street analyst Sal Khan in 2004 to create and share his lectures via videos, and this showed the power of lecture videos (a recording of drawings on an electronic blackboard, which is similar to the style of teaching in a traditional classroom) to provide alternative means of learning. It is interesting to note that the popularity of Khan's formal lecture videos, though anecdotal, reflected enhanced student interest and learning, based on voluntary student viewing.[3] Motivated by the viewing success of his initial lecture videos, Khan started the Khan Academy (a not for profit organization, www.khanacademy.org) in 2008 to make learning tools such as lecture videos and online assignments on a variety of subjects available for free. Inspired by Khan, various universities have made a massive amount of free and quality educational material available online. Stanford professors Sebastian Thurn and Peter Norvig opened access to their online course Introduction to Artificial Intelligence for a fee in fall 2011 and had an enrollment of over 160,000 students worldwide. By 2012-13, this had resulted in the popular emergence of MOOCs (Massive Open Online Courses, available to all those who want to learn and with no participation restrictions) via online initiatives such as Udacity, Coursera, and EdX (see, for example, Baturay 2015). MOOCs, among other features, use short lecture videos to teach and embed questions in the videos to test student learning and understanding.

This phenomenon of creating and sharing lecture videos has also led to the creation of another method of teaching called the "flipped" classroom.[4] This teaching format contains two parts–one, students watch lecture videos and initially assess their understanding before coming to class; and two, students participate in various active learning exercises in the classroom. Thus, in a flipped classroom, the lowest level of cognitive domain (remembering and understanding) is accomplished outside the classroom via the use of video lectures.[5] Higher levels of cognitive domain (applying, analyzing, evaluating, and creating) are completed in the classroom by engaging students through group discussions, problem solving exercises, etc.[6] Further, since the flipped classroom is student-centered, students learn valuable life skills of time management, learning through discussions, working with other students, and seeking help when needed.

The research literature acknowledges various taxonomies to classify classroom pedagogy. For the purposes of this paper, I classify available methods of teaching as traditional, online, and flipped, though

---

[2] This term was first coined at UNESCO's 2002 forum on Open Education Resources (OER). Hewlett Foundation defines OER as "teaching, learning, and research materials in any medium–digital or otherwise–that reside in the public domain or have been released under an open license that permits no cost access, use, adaptation and redistribution by others with no or limited restrictions." OERs include full courses, course materials, modules, textbooks, streaming videos, tests, software, and any other tools, materials, or techniques used to support access to knowledge. The United Nations, philanthropic foundations, and universities have funded the growth of OERs and this has become a global phenomenon.

[3] Sharing of video files became much easier with the creation of YouTube.com in 2005, which allows users to upload, view, and share videos free. The popularity of videos on YouTube attests, among other things, to the easiness of understanding an activity when it is viewed rather than read. Stelzer et al. (2009) found that the students receiving the multimedia learning modules performed significantly better on tests than the students experiencing the text-based presentations in an introductory-level physics course. Vazquez and Chiang (2016) found similar results in a Principles of Microeconomics class.

[4] Initially, a classroom was considered "flipped" if there was a simple re-arrangement of learning activities–the events that took place in classroom were moved to outside the classroom and vice-versa (see, for example, Lage et al. 2000). Based on research on how students learn best and recent developments in technology, a flipped classroom now represents an *expansion* of curriculum rather than a simple re-arrangement of activities (see, for example, Bishop and Verleger 2013).

[5] Students can get first exposure to the course material through assigned readings and/or lecture videos. In this paper, I have used a restrictive definition of flipped classroom, where students receive initial exposure to a course topic through lecture videos. The main reason for using this restrictive definition is that a more general definition of flipped classroom makes it difficult to assess student preferences or effectiveness of the flipped classroom.

[6] These active learning exercises, which engage students to play an active role in applying their knowledge to different contexts in creative ways, have been found to result in higher learning gains. Freeman et al. (2014), based on meta-analysis of 225 studies, found that devoting class time to active learning resulted in higher learning gains in STEM fields.

fully recognizing that there are variants within each teaching method. For example, blended or hybrid courses combine traditional classroom with online activities. I consider a flipped classroom to be an example of blended or hybrid learning, where students gain initial knowledge about the course material via recorded lecture videos and class time is devoted to active learning exercises. Thus, for a class to be considered flipped, first exposure to the course material cannot be primarily readings (textbook or others) and lectures cannot be a primary activity during class meeting. This definition of flipped classroom is based on decades of research on how students learn best (see, for example, Bishop and Verleger 2013, Goffe and Kauper 2014, Vazquez and Chiang 2016, Balaban et al. 2016). Furthermore, this teaching format encourages continual self-assessment of their understanding by students leading to the development of meta-cognitive skills.

The outline of the paper is as follows: the next section reviews relevant literature on students' preferences and learning outcomes under alternative methods of teaching, the following section outlines the evolution of my flipped classroom and its current structure, the subsequent section summarizes and discusses students' preferences, and the last section concludes the research.

Literature Review

From the perspective of research regarding different methods of teaching (traditional, online, or flipped), two questions are relevant: first, which method of teaching results in better student learning outcomes; and second, which method of teaching do students prefer. Research on learning outcomes compares the online or flipped classroom to the traditional classroom. Research studies that controlled for student selection reported that students in online economics classes attained significantly lower test scores than students in traditional classes (Allgood et al. 2015). An interesting and comprehensive study by Xu and Jaggars (2014), using a dataset of nearly 500,000 courses taken by over 40,000 community and technical college students in Washington state, found a much lower performance in online classes relative to traditional classes. Their research showed that this result was consistent across different demographic characteristics (blacks, younger, lower previous academic performance), though the performance gap was larger for some demographic characteristics and smaller for some. They suggested that to improve student performance in online courses, colleges could take at least three distinct approaches: screening, early warning, and scaffolding (learning self-directed work skills). Further, Pyne (2007) found that taking an online introductory microeconomics course had a negative impact on learning outcomes in subsequent economics courses. Thus, most research that compared learning outcomes between online and traditional classroom found student performance in online classes was lower than student performance in traditional classes.

Recent research that compares flipped classroom to traditional classroom finds higher learning gains in the former. For example, based on a 729 student sample in Introduction to Microeconomics classes at University of North Carolina-Chapel Hill, Balaban et al. (2016) found that the flipped classroom format improved students' effort and increased students' final exam performance relative to traditional classroom. Their study also evaluated the impact of flipped classroom on student learning in different cognitive levels (based on Bloom's 1956 taxonomy). Their findings showed that the flipped classroom positively impacts students' comprehension, application, and analytical capabilities and that there was no difference in student knowledge between flipped and traditional classroom frameworks. Using online lecture videos available from Khan Academy and online assignments created by Aplia, Caviglia-Harris (2016) compared student achievements in traditional versus two versions of flipped classroom (fully and partially flipped) and found that students in both versions of flipped classroom performed better than students in traditional classroom.[7]

Thus, available research shows higher learning outcomes in traditional classrooms relative to 100% online classes and higher learning outcomes in a flipped classroom relative to a traditional classroom. Though there is no research that comprehensively compares learning outcomes in the three methods of teaching, based on existing evidence, one might infer that alternative teaching methods can be ranked as flipped, traditional, and then online.

Various studies have shown that students' preferences and learning outcomes are related and are not independent (see, for example, Ainley 1998 and Renninger 2000). It is likely that a useful learning experience may lead students to greater interest in the course material and thus result in better learning outcomes. All this research in economics started with the pioneering work by Lage et al. (2000), who discussed the benefits

[7] Researchers have found positive learning outcomes associated with variants of flipped classroom (blended or inverted) particularly in STEM (science, technology, engineering, and math) fields. See, for example, Balaban et al. 2016.

of an "inverted" classroom and compared the preferences of students and instructors between an "inverted" classroom and a "chalk and talk" classroom. Based on a sample of about 189 students in a Principles of Microeconomics class in fall 1996, their results showed that students generally preferred the "inverted" classroom over a "chalk and talk" classroom. It is interesting to note that their experiment used reading the textbook and watching lecture videos before the class meeting. These lecture videos could only be played on multimedia computers and/or VCRs, and their "inverted" classroom predates the technical developments relating to the world wide web that have led to relative ease of creating, distributing, and viewing lecture videos using a variety of portable and non-portable devices.[8]

Roach (2014) presented results about students' preferences based on his experience in a "partially" flipped classroom in Principles of Microeconomics. In this framework, the students were required to watch one video (a lecture video created by Khan Academy or some application-based video), followed by two days of in class lecturing and one day of active learning exercises. Based on 96 responses, he found that students responded positively to flipped learning. Joyce et al. (2014) assessed students' preferences about meeting once a week versus twice a week, when students had access to online lecture videos, PowerPoints slides, and pre- and post-lecture online quizzes in both classes, and compared the learning outcomes between once a week class meetings and twice a week class meetings. Their results showed that students preferred once a week meetings to twice a week meetings by 23% at the beginning of the semester. However, by the end of the semester (closer to final exams), they found that students preferences had reversed (students preferred twice a week class meetings to once a week meetings by 11%).

The existing research on students' preferences over alternative teaching methods suffers from the following. First, both studies (Roach 2014 and Joyce et al. 2014) on students' preferences do not qualify as flipped classroom in the sense that outside classroom lecture videos were accompanied by full-fledged in class lectures. Second, findings of both studies are based on the one semester experience of one class (Roach 2014) or the one semester experience of four classes under two formats viz., two hybrids and two traditional classes (Joyce et al. 2014). Third, both studies are based on students' preferences in Principles of Microeconomics. Fourth, the student response rate was 54% in case of Roach (2014), though much higher and representative (student response rate of 96%) in the case of Joyce et al. (2014). Fifth, as far as I am aware, there is an absence of research that compares online with flipped classroom.

Given the limitations of earlier work, this paper compares students' preferences over alternative methods of teaching (traditional, online, or flipped) and bases its results on data collected for: one, a fairly large sample of students (601 student responses) with a high response rate (88.6%); two, a four-year time period; three, students in both lower and upper level economics classes. In addition, this paper summarizes students' preferences about use of lecture videos over textbook reading, lecture videos over in-class lectures, and the structure of work outside the classroom.

Evolution and Current Structure of My Flipped Classroom

Designing a Flipped Classroom

Inspired by the learning outcomes associated with online lecture videos, I decided to create my own lecture videos for Intermediate Microeconomics class in 2012. About a year was spent on searching for an app that was available free, easy to use, and interactive enough to provide a lecture experience close to face-to-face lectures. In July 2013, I stumbled upon an app called KNOWMIA and started to create my own lecture videos. During fall 2013, I created about 20 lecture videos, each no more than 15 minutes in duration, on four to five topics in Intermediate Microeconomics that would have been covered in about five weeks during the semester. My classes were scheduled to meet for 75 minutes each on Tuesdays and Thursdays. I had an agreement with students that they watch three to five lecture videos before coming to a class meeting on Tuesdays. The Tuesday class time was reserved for Q & A session, discussions, individual or group problem solving exercises, etc. If all parties (the students and the teacher) were satisfied by the extent of learning in

[8] Note that the lecture videos initially started with using tapes played on VCRs and video was projected on TV screens. However, though useful, they never had the convenience that is afforded by portable devices that are currently available. Their use was somewhat limited for various reasons such as the difficulty with availability, viewing time, etc. Since most of the videos were made by professionals, they appeared more movie like and did not truly replace the classroom lecture experience. All this is now possible and, as Joyce et al. (2014, p. 21) noted, "students appear to be drawn to videos in which their professor appears rather than an unknown 'talking head.'"

one class meeting, the class meeting on Thursday was cancelled. The students enjoyed getting time off and I, as an instructor, was impressed by the quality of their questions, by their responses to my questions, and their performance on class assignments.[9] Further, I observed that during these class meetings, their attendance was close to perfect.

Encouraged by my own impressions and students' preferences of this format, I created more lecture videos and online assignments (five multiple-choice questions) associated with each lecture video for my Intermediate Microeconomics class. By the end of spring 2014, I had created a complete set of lecture videos and online assignments associated with each lecture video for my Intermediate Microeconomics class. During spring and fall 2014, I had the same arrangement with students–Thursdays off, if they watched the lecture videos at home, completed the assignments associated with those lecture videos, and all of us were satisfied with their learning during class meetings on Tuesdays. This worked well and the class did not meet on Thursdays. In spring 2015, after three semesters of experimentation, Intermediate Microeconomics was formally offered as a hybrid that met only on Tuesdays. Freed from my responsibility of creating lecture videos and online assignments for Intermediate Microeconomics, I devoted my efforts to creating lecture videos and associated online assignments for my Principles of Microeconomics and International Trade classes. Once again, I followed the same steps that I had followed for my Intermediate Microeconomics class–agreement with students to meet only once a week that could easily be extended to twice a week meeting, if either any student or I was not satisfied with learning outcomes associated with this method of teaching. In fall 2016, both Principles of Microeconomics and International Trade classes were formally offered as hybrid classes that met only once a week. During summer 2016, a weekly discussion forum was added to provide a platform for students to ask questions and respond to questions asked by others before the class meeting. Participation in the discussion forum was for extra credit and I intervened only when I felt necessary. Thus, though not my original intention, I had created a full-fledged flipped classroom.

Structure of My Current Flipped Classroom

The entire course for each class is organized on a weekly basis. Prior to coming to class, students watch three to five lecture videos (each less than 15 minutes in duration) and/or read the textbook,[10] ask questions and respond to questions in the discussion forum (a platform for and by the students in which I respond only on a need-to-respond basis), and complete online assignments associated with each lecture video. Each online assignment contains five multiple-choice questions, each student gets two attempts, and only the latest (not the highest) score is recorded. Further, the correct answers for online assignments are posted on the day of the class meeting, when the online assignments are not available for submission. Further, the students cannot discuss answers to any online assignment questions in the discussion forum. All this is organized in weekly modules and students have the flexibility during a given week to complete all pre-class activity.

The students are required to come prepared to ask specific questions about the topic(s) covered during the week and actively participate in various learning exercises during the class meeting. Before each class meeting, I spend about 1-3 hours reviewing information available on Canvas about each student (course analytics, scores on online assignments, Q & A on the discussion forum, etc.) and designing each classroom meeting with the twin objectives of removing misunderstandings and enabling a higher level of learning. Given the size of my classes (typically 10 to over 100 students), it is easier for me to personalize the classroom experience by actively engaging each student in learning through Q & A sessions, discussions, and problem-solving assignments. Lastly, all exams are held in class.

Since the summer of 2014, I have administered an end-of-semester survey in all of my classes. The survey includes ten questions on a five-point Likert scale along with a section for written comments.[11] Student preferences are measured on a scale of 1 to 5, with 1 representing "strongly disagree," 2 representing "disagree," 3 representing "neither agree nor disagree," 4 representing "agree," and 5 representing "strongly

[9] Joyce et al. (2014) compared student performance in classes that met once a week with the classes that met twice a week and found no significant difference in student performance. Thus, they recommended the use of hybrid classes, with little or no impact on student performance. This, inter-alia, should lead to higher productivity of college and university resources.

[10] Note that the lecture videos are a substitute for in-class lectures and not a substitute for textbook reading. Irrespective of the teaching format, textbooks are recommended but not required in any of my classes.

[11] Summary tables of students' responses to the survey as well as detailed students' comments on different aspects of flipped classroom for each class are available upon request from the author.

agree." For the purposes of discussion, scores of 1 and 2 are combined to represent "disagree," score of 3 represents "neither agree nor disagree," and scores of 4 and 5 are combined to represent "agree." The current design of my flipped classes is based on student feedback and my personal impressions regarding the level of learning.

Students' Preferences

The purpose of the survey was to assess student preferences with respect to alternative teaching methods and the structure of the course to enable better course design. Improved design accommodates student preferences and, based on my impressions, has led to better learning outcomes. Participation in the survey was voluntary, although students received extra credit points as an incentive to respond. To ensure student anonymity and reduce reporting bias, the survey did not include questions regarding demographic background or other identifying information. This was important as my class sizes varied from 10 to over 100 students. These data were collected from summer 2014 to spring 2018 and included all of my classes that were completely taught using lecture videos at a four-year public university.[12] As stated earlier, 601 students responded to the survey, with a response rate of about 88.6%. Table 1 summarizes combined responses of all students in Principles of Microeconomics (Econ 200), Intermediate Microeconomics (Econ 318), and International Trade (Econ 440) classes.[13]

Table 1: Summary of Student Responses to Survey Questions

Survey Statement	Percentage Agree	Percentage Disagree	Percentage Neither Agree nor Disagree	Mean (Standard Deviation)
Prefer recommended to required textbook	85.9	7.8	6.3	4.4 (1.07)
Prefer lecture videos to textbook reading	87.5	4.7	7.8	4.5 (0.88)
Prefer short online assignment to no online assignment	78.4	8.2	13.5	4.1 (1.04)
Prefer short online assignment to one longer online assignment	85.7	7.0	7.3	4.4 (1.01)
Prefer weekly-structured assignments to daily-structured assignments	86.5	4.3	9.0	4.5 (0.88)
Prefer flipped mastery to no flipped mastery	23.6	43.3	33.1	2.67 (1.26)
Prefer lecture videos to in-class lectures	66.2	14.5	19.3	3.79 (1.14)
Prefer flipped to 100% online class based on power points and readings	74.2	10.3	15.5	4.1 (1.10)
Prefer flipped to 100% online class based on lecture videos and reading	72.1	10.1	17.8	4.0 (1.11)
Prefer flipped to traditional class	62.7	17.5	19.8	3.72 (1.21)

Notes: Sample size (N) = 601.

[12] This is a fully accredited regional university that offers bachelor's and master's degrees, with an overall enrollment of about 7,000 students. About 6.8% of the student body is comprised of international students and about 8.3% of the students are transfer students. The average age of undergraduate students is 21.7 years and of graduate students is 32.4 years. Average ACT scores of undergraduate students is 21.8 and average high school GPA is 3.36. It is interesting to note that between academic years 2012-13 and 2016-17, the number of online classes offered by the university increased by 50% and the number of hybrid classes almost tripled during the same time. Out of the 148 classes offered in the College of Business during fall 2016, 14 classes were offered online and seven were offered as hybrid classes. The overall enrollment in the College of Business is about 1,100 students.

[13] Note that there was statistically no significant difference in students' preferences between lower and upper level classes. Hence, only combined results for all classes are discussed.

The survey contained two questions relating to the use of textbooks–one, do students prefer that textbooks be recommended or required, and two, do students prefer watching lecture videos to reading a textbook. As shown in Table 1, a majority of the students (85.9%) preferred that textbooks be recommended rather than required, and 7.8% students preferred that the textbook be required rather than recommended. Also, 87.5% of the students preferred watching lecture videos to reading a textbook (only 4.7% preferred reading a textbook to watching video lectures). This is consistent with earlier research–for example, Vazquez and Chiang (2015) found that 92% of students preferred multimedia pre-lectures to textbook reading. Caviglia-Harris (2016) noted that Khan Academy videos are white board scribe animation (i.e., do not include images of people) and are known to be conducive to learning economics or any other complex material that can be presented in steps (such as graphs and equations). Further, Vazquez and Chiang (2016) found that students with access to visual and audio presentation performed better than those with access only to text-based presentation in Principles of Microeconomics course.[14] They also suggest, from a practical perspective, nearly all students today have grown up in an era of the internet and portable devices, and therefore are more comfortable learning using these tools. They also stated that some instructors create time for active learning by simply replacing textbook reading before class. Their findings suggested that this strategy might be counterproductive, because it denies the student an effective way of receiving basic information for the first time: the spoken lecture. They recommended that instructors move the traditional lecture outside the classroom by assigning pre-recorded multimedia lectures to the students that better replicate the chalk-and-talk experience, thereby preparing students for a more effective active learning experience in class.

The survey contained four questions regarding the structure of pre-class activities. The first two questions were related to the need for online assignments and to the extent of topics covered in online assignments. A majority of the students (78.4%) preferred a short online assignment associated with each lecture video as compared to no online assignments, and only 8.2% preferred no online assignment to a short online assignment. Further, 85.7% of the students preferred a short online assignment associated with each lecture video as compared to one long online assignment for the entire topic (based on several lecture videos), and only 7.0% preferred a long online assignment to a short online assignment associated with each video lecture. The next question related to the structure of the weekly work. The students preferred that they have the independence to watch lecture videos and complete online assignments anytime during the week (weekly-structured assignments) as compared to having to watch lecture videos and complete online assignments on a daily basis (daily-structured assignments). As the table shows, 86.5% of the students preferred weekly-structured assignments over a daily-structured assignment and only 4.3% preferred daily-structured assignments to weekly-structured assignments. This is in line with our desire that students learn the course material at a time most convenient and productive for them and, as a step towards maturity, students learn how to allocate their time efficiently. An interesting result is with respect to the question of flipped mastery (i.e., the students cannot move on to the next topic, unless they have a certain level of competency in the current topic in terms of their score on an online assignment) versus no flipped mastery (the students can move on to the next topic, irrespective of their performance on an online assignment). The survey results show that 23.6% of students preferred flipped mastery over no flipped mastery, 43.3% preferred no flipped mastery over flipped mastery, and 33.1% were indifferent between the two**.** This mixed result relating to students' preferences about flipped mastery, perhaps reflects the fact that the students were provided with several opportunities to remedy their lack of understanding via alternative means–(i) before class meeting: re-watching lecture videos, reading textbook, using online discussion forum to ask questions or see someone else's response to the same question asked by another student, and (ii) during class meeting: asking specific question(s) in class, participating in classroom discussion (directed or otherwise), and working on problem-solving exercises. This design encourages students to wrestle with content and discover answers on their own.

The last four questions relate to alternative methods of teaching. The first question asks if students prefer lecture videos to in-class lectures. The survey results show that 66.2% of students preferred lecture videos to in-class lectures and only 14.5% of the students preferred in-class lectures to lecture videos. This result is in contrast to the earlier research that showed that students tend to prefer in-class lectures to lecture videos (see, for example, Bishop and Verleger 2013). Compared to classroom lectures, the lecture videos have distinct

[14] Note that Stelzer et al. (2009) found similar results in an introductory physics course. All these empirical results confirm decades of theoretical research on neuro and cognitive sciences on how brain functions when stimulated by information delivered through alternative means. See Vazquez and Chiang (2016) for a summary.

advantages for students. Lecture videos enable students to watch multiple times if needed, print slides from the lecture videos, take notes only when they feel comfortable in terms of their understanding (this is important as in classroom lecture, students are under pressure to understand and take notes at the same time). In addition, they can clarify doubts or lack of understanding by watching the videos again and, if that does not help, by reading the textbook. The next two questions compare student preferences of a flipped classroom to two formats of online classes (online classes based on PowerPoint slides and readings and online classes based on lecture videos and readings). Table 1 shows that 74.2% of the students preferred the flipped class structure over a class that is completely online using PowerPoint slides and reading, and 10.3% preferred this structure over the flipped class structure. Approximately 72% of students preferred the flipped class to 100% online class based on lecture videos and readings, and 10.1% of students preferred a 100% online class based on video lectures and readings over a flipped classroom.

The last question compared students' preferences between flipped and traditional class structures. Based on the survey, 62.7% of students preferred a flipped class to a traditional class and 17.5% preferred a traditional class to a flipped class. This finding (preference of flipped classroom over traditional classroom) is consistent with students' preference of lecture videos over in-class lectures and their preference of lecture videos over textbook reading. The former also attests to students' preference for receiving their first exposure to course material through lecture videos rather than in-class lectures. Further, their preference of a flipped class over an online class (particularly, an online class based on lecture videos and readings) highlights the importance of class meetings.

Conclusion and Suggestions

This paper describes the creation of a classroom based on experimentation, students' feedback, and instructor's impressions about learning outcomes. Findings of this paper, based on students' preferences, highlight the importance of lecture videos over textbook reading and also over in-class lectures; flexibility to work whenever most convenient and productive for students within a given time period (a week); the need to test learning based on each sub topic or the topic covered by each lecture video rather than one longer weekly assignment based on topic(s) covered in three to five lecture videos; the need for a weekly class meeting–based on their preference of flipped over 100% online classes based on lecture videos or reading; and the need for devoting classroom time on various active learning exercises. Though not measured in terms of students' preferences, written comments about the discussion forum[15] suggest that the ability to ask and answer questions by students is a useful learning tool and, from my perspective, another way to assess students' learning (based on students' effort and quality of their posts, etc.).

Though it is hard for any single teaching method to appeal to the learning preference of every student, it is possible for the instructors to engage a group of learners who vary in terms of level of preparedness, interest, and effort.[16] For example, course analytics on Canvas provide instructors with a wealth of information about each student's study habits (when they study and for how long), the areas in which they are having difficulty, etc. The online weekly discussion forum also provides excellent information about students' understanding or lack thereof. Thus, the information available prior to the class meeting is very useful for instructors to prepare material for the class meeting to comprehensively address problems faced by different students and enhance their understanding. The role of the instructor, then, can shift away from teaching/lecturing to a typical or an average student to where the instructor becomes a facilitator enabling students to transition their own learning of course material to acquiring skills and competencies.

---

[15] My flipped classes have a weekly discussion forum–mainly by and for the students. Students can ask any question(s) about the topics covered during the given week and respond to each other's questions. I respond to any unanswered questions by 6:00 PM the day before the class meeting. Participation in the discussion forum is voluntary and students receive extra credit points, based on the quality of Q & A. These discussion forums have become extremely useful as: (1) the students are able to see the benefits of academic discussion in terms of learning; (2) they make everyone aware of alternative resources (lecture videos created by others, interesting articles, etc.) that can enable basic learning or take discussion to a higher level; and (3) they may eliminate the need for some students to hire a tutor. The participation rate in these discussion forums is about 40 to 70% and every week I am impressed by the quality of questions and responses of students.

[16] For example, research on low completion rates of MOOCs and also on online courses offered at community colleges shows that it is difficult for these online courses to cater to a diverse set of students who vary in terms of background skills, preparedness, and effort (see, for example, Khalil and Ebner 2014 and also Xu and Jaggars 2014). Until now, it was not possible to customize teaching according to students' preparedness to do well in a given course. With a flipped classroom, it is possible for an instructor to create interactive active learning exercises based on students' preparedness and motivation.

References

Ainley, Mary D. 1998. "Interest in Learning in the Disposition of Curiosity in Secondary Students: Investigating Process and Context." In *Interest and Learning: Proceedings of the Seeon Conference on Interest and Gender*, edited by Lore Hoffman, Andreas Krapp, K. Ann Renninger, and Jürgen Baumert. Kiel, Germany: IPN.

Allgood, Sam, William B. Walstad, and John J. Siegfried. 2015. "Research on Teaching Economics to Undergraduates." *Journal of Economic Literature* 53: 285-325.

Baturay, Meltem H. 2015. "An Overview of the World of MOOCs." *Procedia-Social and Behavioral Sciences* 174: 427-433.

Balaban, Rita A., Donna B. Gilleskie, and Uyen Tran. 2016. "A Quantitative Evaluation of the Flipped Classroom in a Large Lecture Principles of Economics Course." *The Journal of Economic Education* 47: 269-287.

Bishop, Jacob L., and Matthew A. Verleger. 2013. "The Flipped Classroom: A Survey of Research." In *2013 ASEE Annual Conference & Exposition*.

Bower, Beverly L., and Kimberly P. Hardy. 2004. "From Correspondence to Cyberspace: Changes and Challenges in Distance Education." *New Directions for Community Colleges* 128: 5-12.

Caviglia-Harris, Jill 2016. "Flipping the Undergraduate Economics Classroom: Using Online Videos to Enhance Teaching and Learning." *Southern Economic Journal* 83: 321-331.

Chaubey, Aabha, and Bani Bhattacharya. 2015. "Learning Management System in Higher Education." *International Journal of Science Technology & Engineering* 2: 158-162.

Freeman, Scott, Sarah L. Eddy, Miles McDonough, Michelle K. Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary P. Wenderoth. 2014. "Active Learning Increases Student Performance in Science, Engineering, and Mathematics." *Proceedings of the National Academy of Sciences* 111: 8410-15.

Goffe, William L., and David Kauper. 2014. "A Survey of Principles Instructors: Why Lecture Prevails." *The Journal of Economic Education* 45: 360-375.

Hewlett Foundation. "Open Educational Resources." *https://hewlett.org/strategy/open-educational-resources/*.

Joyce, Theodore J., Sean Crockett, David A. Jaeger, Onur Altindag, Stephen D. O'Connell. 2014. "Does Classroom Time Matter: A Randomized Field Experiment of Hybrid and Traditional Lecture Formats in Economics." National Bureau of Economic Research Working Paper No. 20006.

Khalil, Hanan, and Martin Ebner. 2014. "MOOCs Completion Rates and Possible Methods to Improve Retention–A Literature Review." *Proceedings of World Conferences on Educational Multimedia, Hypermedia, and Telecommunication* 1236-1244.

Lage, Maureen J., Glenn J. Platt, and Michael Treglia. 2000. "Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment." *The Journal of Economic Education* 31: 30-43.

Pyne, Derek. 2007. "Does the Choice of Introductory Microeconomics Textbook Matter?" *The Journal of Economic Education* 38: 279-96.

Renninger, K. Ann. 2000. "Individual Interest and its Implications for Understanding Intrinsic Motivation." In *Intrinsic and Extrinsic Motivation: The Search for Optimum Motivation and Performance*, edited by Carol Sansone and Judith M. Harackiewicz. New York: Academic Press.

Roach, Travis. 2014. "Student Perceptions toward Flipped Learning: New Methods to Increase Interaction and Active Learning in Economics." *International Review of Economics Education* 17: 74-84.

Stelzer, Timothy, Gary Gladding, José P. Mestre, and David T. Brookes. 2009. "Comparing the Efficacy of Multimedia Modules with Traditional Textbooks for Learning Introductory Physics Content." *American Journal of Physics* 77: 184-90.

Vazquez, Jose J., and Eric P. Chiang. 2015. "Flipping Out! A Case Study on How to Flip the Principles of Economics Classroom." *International Advances in Economics Research* 21: 379-390.

Vazquez, Jose J., and Eric P. Chiang. 2016. "Preparing Students for Class: A Clinical Trial Testing the Efficacy between Multimedia Pre-lectures and Textbooks in an Economics Course." *Journal of College Teaching & Learning* 13: 37-46.

Xu, Di, and Shania S. Jaggars. 2014. "Performance Gaps between Online and Face-to-Face Courses: Differences across Types of Students and Academic Subject Areas." *The Journal of Higher Education* 85: 633-659.

Priming Effects of Instructor-Expressed Judgments of the Difficulty of Quantitative Material on Student Performance and Perceptions

Kevin S. S. Henning and Natalie D. Hegwood[1]

ABSTRACT

We investigate whether instructors teaching quantitative material can influence student performance through priming, a psychological phenomenon wherein exposure to a stimulus can influence memory and behavior. We first examine whether instructors can influence performance by priming students with statements stating that a quantitative problem has "Advanced" or "Basic" difficulty. We then investigate effects of priming on perceived difficulty. Priming does not affect performance but does affect perceptions. Compared to no priming, Basic priming reduces perceived difficulty, while Advanced priming increases female students' perceived difficulty. Male students view Basic problems to be easier, but do not see Advanced problems as harder.

Introduction

Every college course presents its own unique pedagogical challenges, but instructors of quantitative courses frequently face the additional hurdle of counteracting years of inculcated fear and ambivalence toward all things mathematical that students have brought with them from past experiences. The challenges of teaching mathematical content while overcoming fear and negative attitudes in students have been discussed for decades (Betz 1978; Dillon 1982; Gal and Ginsburg 1994; Hembree 1990; Zanakis and Valenzi 1997). Estimates vary based on measurement instruments, but it has been estimated that one quarter of students in four-year undergraduate programs and up to 80% of students in community colleges report having experienced math anxiety, a condition brought on by situations involving quantitative reasoning that is characterized by feelings of panic, helplessness, dread, and even physical illness such as nausea and heart palpitations (Chang and Beilock 2016; Henrich and Lee 2011). Given that economics and statistics courses are staples of the modern undergraduate experience (Allgood et al. 2015), instructors are increasingly under pressure to mitigate the effects of math anxiety because of the significant relationship between math anxiety and performance (Benedict and Hoag 2002; Hoag and Benedict 2010).

Common advice for the instructor interested in striking a balance among student engagement, accessibility, and rigor is to revamp his or her existing pedagogy through innovations such as flipped classrooms, competitive learning games, and case- and problem-based learning (e.g., Aguilar and Soques 2015; Chulkov and Nizovtsev 2015; Vazquez and Chiang 2015). Our interest in this paper is in exploring the potential of instructors of math-intensive courses to influence student performance and perceptions of difficulty in a subtle way by communicating judgments of how difficult the material is. One of the distinguishing features of quantitative subjects is the propensity for the authors of textbooks and many instructors to editorialize the material as it is presented.[2] Most students taking quantitative courses will, sooner or later, encounter a problem that is pitched to them as being "an obvious application" of some theorem or process that requires such "an easy calculation" that the student cannot help but "clearly see" that "the result follows directly." Depending on the instructor and circumstances, such statements could serve as

[1] Henning (corresponding author): Department of Economics and International Business, Sam Houston State University, Huntsville, TX, 77341, henning@shsu.edu. Hegwood: Department of Economics and International Business, Sam Houston State University, Huntsville, TX, 77341, eco_ndh@shsu.edu.

[2] The authors of the present paper count ourselves among these "many instructors."

mere rhetorical devices, or verbal "filler," or they could be uttered by well-intentioned instructors to boost the confidence of students before moving on to a problem that appears challenging. Whatever the motivation is for instructors to comment on difficulty, the effects of doing so on the experience of students in the classroom are less well known compared to those of newer and more dynamic teaching interventions.

In this paper, we investigate the role of instructor-expressed judgments of difficulty on student performance as well as perceptions through the theoretical lens of psychological priming. A "priming effect" is said to occur when exposing a person to subtle stimuli (words, colors, sounds, etc.) alters that person's cognitive processing of, or response to, subsequent stimuli through implicit, or unconscious, memory. Priming effects can manifest in various ways, from simple decreases in the time it takes for individuals to identify words or pictures that have been previously presented to them, to changes in complex behavior and judgments (Bargh et al. 2001; Dijksterhuis and van Knippenberg 1998; Tulving and Schacter 1990). A simple example of the priming effect is the increased likelihood that a person who is exposed to information related to food and eating will think of "soup" when shown the word stem SO__P compared to a person not exposed to such information (Kahneman 2011). Because by definition, priming involves a subtle stimulus that can be activated in numerous ways, the instructor has practically limitless ways to prime students for behavior. Examining the nature and degree of priming effects in a classroom context (in particular, the communication of the judgments of problem difficulty) could help inform best practices of instructors within the classroom.

In this paper, we add to the literature on priming and education by first investigating whether an instructor's expressed judgment of the difficulty of a quantitative problem can significantly alter student performance, as measured by grades on a quiz with mathematical content. We then examine whether the mere perception of difficulty can be altered through priming. Overall, we find the effects of priming to be mixed. While we do not find that priming significantly affects actual student performance, we do find evidence that priming influences perceptions of difficulty. In particular, priming students to expect that a problem will have an advanced difficulty level ("Advanced") tends to increase perceptions of difficulty compared to when students are primed to expect a basic difficulty level ("Basic"). Moreover, Basic priming results in a general decrease in perceived difficulty relative to no priming ("None"), but there is no significant change in perceived difficulty when Advanced priming is given over none at all. We find significant gender differences related to priming and perception, however. Advanced priming increases female students' perceptions of difficulty compared to no priming but does not do so for males. Further, while female students tend to view Advanced-primed problems to be more difficult than Basic-primed problems, they do not view Basic-primed problems to be significantly easier than problems without priming. Males, however, do perceive Basic-primed material to be easier.

In the next section, we provide additional background on the nature of priming, and more specifically, priming in an educational context. We then discuss our methodology, which involves a quasi-experiment using data from quizzes given to two sections of an undergraduate business statistics course as part of the normal assessment process. After presenting the results, we conclude with a discussion of the main findings and provide some directions for future research.

Background and Literature Review

Priming is a phenomenon studied in psychology, education and, more recently, behavioral economics (e.g., Kamenica 2012) in which a person's exposure to a subtle or even subliminal stimulus can influence his or her judgments, goals, and behavior when exposed to a seemingly unrelated stimulus (Bargh et al. 2001; Cohn and Maréchal 2016; Elliot et al. 2007; Molden 2014). The idea of priming in a psychological sense was first outlined in the early 1950s (Bargh 2014), and over the subsequent six decades, the phenomenon has been expanded to include effects on behaviors, such as the achievement of goals (Chartrand and Bargh 1996; Dijksterhuis and van Knippenberg 1998; Förster et al. 2009).

Priming effects on college students have been documented in specific subjects. In an economics classroom setting, Jackson (2016) found a significant positive difference in the performance of students on a quiz between groups of students who were shown images of well-known "high performing" persons with others shown images of well-known "low performance" individuals. Priming has also been studied as a mechanism that could cause student performance to decline due to "stereotype threat" activation, in which individuals who are reminded of their membership in a social group with certain stereotypes begin to exhibit behaviors that are stereotypical of that group. Stone et al. (2012) found that priming African American student athletes by having them note their identity as either "athlete" or "scholar-athlete" on a mathematics performance test resulted in poorer performance relative to assessments involving no priming, even when they are otherwise "academically engaged," that is, regarding academic performance as a key aspect of their overall self-worth. This work extended Harrison et al.'s (2009) finding that priming athlete identity and gender identity negatively affected the performance of female students on a verbal analogy test, but priming athlete identity resulted in better performance from males, even when faced with more difficult questions.

Prior research on the effects of priming in the classroom has also examined the influence of the instructor on perceptions, attitudes, beliefs, and actions of students. In statistics classrooms, an instructor who endeavors to be approachable by using welcoming verbal and nonverbal cues (such as smiling, incorporating humor into the classroom, etc.) may significantly lower student anxiety levels (Williams 2010). Jackson and Leffingwell (1999) demonstrated the importance that students place on their instructors' judgments, finding in a survey of elementary education students that 27% of the respondents traced the origin of their mathematics anxiety to experiences in college classrooms. Some specific anxiety-inducing behaviors of instructors that the students reported included being told they were "dumb" or "slow" or that "[They] should know this." Offhand pronouncements of intellectual ability or learning speed using such emotionally laden words are clearly inappropriate, but occasionally even well-meaning instructors trying to demonstrate caring attitudes can nevertheless negatively affect students' perceptions of their mathematical ability and their motivation to learn. A key determinant appears to be the degree to which an instructor believes that intelligence is an innate, immutable attribute (the "entity view" of intelligence) versus being an incremental attribute, that is, one that can be honed over time (Blackwell et al. 2007). For example, Rattan et al. (2012) found that instructors with an entity view of intelligence tended to support giving struggling students comforting statements to help them accept their limitations. Students perceived such statements as communicating their instructor's low expectations, which tended to lower motivation. Some work closely related to our notion of priming comes from the literature on "metacognitive prompting," also called "metacognitive cueing" (Hoffman and Spatariu 2008; Veenman et al. 2000). This type of intervention, however, is designed to be conspicuous to the learner and trigger reflective cognitive processes consciously in learners to enhance performance.

Given the subtle nature of priming, it is difficult to imagine how an instructor could *not* exert some influence over how students perceive the material. In the traditionally lecture- and problem-driven economics or statistics classroom, the opportunities to shape students' perceptions of the topic, and the discipline itself, abound. The literature in statistics education is particularly rich with suggestions, best practices, and guidelines on content, delivery, and assessment (Garfield et al. 2005; Wood et al. 2018). Less well known is how the presence of psychological primes related to quantitative material, specifically, an instructor's expressed judgment of how difficult that material is, might influence how students perform on assessments, and whether their performance is consistent with their perceptions of how difficult the material is. In a classroom setting, these judgments could arise in a segue from discussing theory to working a problem, to bookend a set of worked problems, or as a well-intentioned reminder of learning outcomes. In this paper, we first investigate the role of priming on actual student performance as measured by quiz grades, which is a measure of how hard a problem actually is. We next investigate the role of priming on perceptions of problem difficulty, that is, on how hard students *think* a problem is.

Methodology and Data

We performed a quasi-experiment using three priming scenarios: no priming statement (i.e., no communicated judgment from the instructor regarding difficulty of the question), which we refer to as the "None" condition; a statement communicating to the student that a question was of "basic difficulty," which we call the "Basic" condition; and a statement communicating to the student that question was of "advanced difficulty," which we call the "Advanced" condition. Data were collected from students enrolled in a required junior-level business statistics course at Sam Houston State University's College of Business Administration. The course covers the standard topics of inferential statistics, namely, confidence intervals and hypothesis tests for means and proportions, ANOVA, chi-squared tests, and regression. As the faculty members in charge of assessing learning outcomes related to this course, we have been interested in gathering data that might provide insight into how students are thinking about the course as they are completing the work rather than merely post-hoc measures of performance.

Motivated by this desire to understand the learning process more deeply in this course as well as examine the central question of our paper, we asked students in two sections of the course—one on a

Monday/Wednesday (MW) schedule and the other on a Tuesday/Thursday (TTH) schedule—to complete a survey about their impressions of the difficulty and familiarity of the material after completing a graded in-class quiz. Because the quiz was required—in the sense that a student's score on it would affect his or her final grade—we wanted to mitigate potential bias that could arise from the tendency of some students to answer the survey according to what they believed the instructor wanted to see. Were the quiz not required (e.g., offered as an extra credit assignment), it would be justifiable to link the survey and the quiz to a particular student.[3] Therefore, the survey was contained on a sheet of paper separate from the quiz so that students could turn in the quiz and survey separately, with the survey responses being anonymous. The mandatory nature of the quiz also led us to design the priming manipulation at the quiz-by-class level: that is, every student in a particular section (MW or TTH) would take the same quiz, and every student would receive the same question and priming: either None, Basic, or Advanced. The assignment of priming conditions to classes and quizzes is as shown in Table 1.

Table 1: Assignment of Priming Conditions to Course Sections

Class	Quiz Number	Priming Condition
MW	2	None
MW	3	None
TTH	2	Basic
TTH	3	Advanced

Note: "MW" indicates a class meeting on a Monday/Wednesday schedule, and "TTH" indicates a class meeting on a Tuesday/Thursday schedule.

The MW (non-primed) class met from 12:30 p.m. to 1:50 p.m., and the TTH (primed) class met from 11 a.m. to 12:20 p.m. By examining courses that met close to the same time, we were able to partially control for time-of-day effects that could influence students' performance and perceptions of difficulty. Research suggests that it is important to consider such effects. For example, Carrell et al. (2011) found a general increase in academic performance in adolescents when the school day started even just 50 minutes later. Class start times have also been associated with math performance in particular, at least for middle school and high school students (Edwards 2012; Fike and Fike 2013; Pope 2015).

To reduce the effect of instructor bias in the selection of questions for the Advanced and Basic conditions, rather than creating and then rating our own questions, we chose to use problems from a test bank that accompanies a popular business statistics textbook that were rated as either Advanced or Basic. Because the priming conditions were assigned at the class level, we can investigate the relationship between priming and performance using grades on the quizzes. However, because the surveys were anonymous and submitted separately from the quiz, we cannot match difficulty perceptions with grades at the student level.

The exact wording of the Advanced priming statement appeared at the top of the student's quiz as follows:

> *This problem was given a difficulty rating of Advanced (on a scale of basic, intermediate, advanced) by the source from which it was taken.*

A similar statement was provided for the Basic priming question. In the non-primed class, there were no statements about the difficulty of the problem, but otherwise the quizzes were identical in appearance to those in the primed class. Because priming can arise from both written and auditory stimuli, and students do not always read directions closely, to ensure that the students in the priming conditions actually did receive a priming stimulus, the instructor read the priming statement and the directions printed on the survey aloud to the entire class; for students in the non-primed class, the instructor read just the directions.

Again, due to the mandatory nature of the quiz, in the interest of fairness, we chose to "never lie to the students;" that is, if a student saw an Advanced (or Basic) priming statement, the problem was indeed Advanced (or Basic) according to the rating system for questions in the test bank. Students in both classes took the first quiz of the semester, Quiz 1, with no priming interventions. This allowed us to establish a baseline for student performance before the priming interventions. The priming interventions began with

[3] A study with such a direct link to priming, difficulty perceptions, and performance would be a natural next step.

Quiz 2 in the TTH class. A student in the MW section taking Quiz 2 received a similar—but not identical—Basic-rated question as a student in the TTH class, but without a priming statement. Similarly, for Quiz 3, a student in the MW section received a similar Advanced-rated question as a student in the TTH section, but, again, without the priming statement. By giving students similar rather than identical questions, we were able to partially account for the fact that a MW student could share quiz information with a TTH student, giving the latter an advantage regardless of priming. By using similar questions, even if this information sharing did occur, the TTH student could not score highly on the quiz solely because someone in the MW class had shared information with him or her.[4]

Each quiz was scored out of 20 points and was administered and graded by the same instructor. This quiz grade ("Grade") is the dependent variable that we use to investigate the main research question, namely, what effect priming has on student performance. To avoid introducing an extraneous "stereotype threat" effect, wherein female students could be primed to artificially underperform relative to male students due simply to being reminded of the negative stereotype that women are weaker in math (e.g., Bench et al. 2015; Spencer et al. 1999; Steele and Ambady 2006), we did not ask students to report their gender on the quizzes; rather, we recorded their genders using student registration information.

Our secondary objective was to examine what, if any, effect priming may have on how difficult students perceive a problem to be, irrespective of how they actually performed. Implications of being able to alter students' impressions of how difficult the material is, even if the effect does not translate into improved actual performance, could nevertheless help instructors assuage their students' math anxiety by creating a more positive classroom environment, which could aid in retaining students within courses and, more broadly, encourage persistence within quantitative majors. To measure perceived difficulty (which we will, at times, refer to as simply "difficulty" for brevity when there is no risk for confusion), we adapted a 0-100 Likert-style scale that Bandura (2006) discusses for measuring self-efficacy, a person's belief in his/her ability to accomplish a goal or task (Bandura 1977, 1982, 1989, 1997). Self-efficacy has been linked to student performance in mathematical problem solving in general (e.g., Pajares and Miller 1994, 1995) and in statistics specifically (Finney and Schraw 2003). Of course, difficulty itself is not equivalent to performance, although it has been found that students' self-efficacy to solve math problems is greater at higher levels of difficulty (Hoffman and Schraw 2009). Using a scale with steps from 0 to 100 (which is equivalent to an 11-point Likert scale) can result in a measurement item that results in observed data closer to behaving as if normally distributed and more sensitive than the usual four- or six-level scales (Leung 2011). Further, in the context of an education study, a scale ranging from 0 to 100 will be consistent with how many students conceptualize grading (e.g., Pajares et al. 2001). We asked students the following question to measure perceived difficulty:

> *How would you rate the level of difficulty of this quiz problem on a scale of 0-100? Define "0" as "not difficult at all", 50 as "moderately difficult", and 100 as "very difficult."*
> *0 10 20 30 40 50 60 70 80 90 100*

One factor that would obviously influence a student's perception of difficulty is his or her knowledge of the material. Because the quizzes required students to provide a worked-out solution to a question, no amount of priming could lead a student to a fully correct solution if he or she lacked knowledge of the concepts. As we discussed earlier, our design did not link grades to the responses to the anonymous survey. Thus, as a proxy variable for content knowledge, we included a question asking students to report their level of familiarity with the material. In our analyses, we used this self-reported familiarity measure ("familiarity") as a control variable to account partially for the differences in difficulty that can arise from content knowledge, irrespective of priming. We asked the following question to measure familiarity, using a five-point Likert scale:

> *How would you rate your level of familiarity with the material on this question? In other words, did it seem similar to ones worked in class and in homework problems?*
> *1: not at all familiar*
> *2: slightly familiar*
> *3: somewhat familiar*
> *4: moderately familiar*
> *5: extremely familiar*

[4] In our empirical models, we formally check for a class effect on our outcome variables.

Additionally, we asked students to report their gender on the surveys to give us a means of examining possible differential effects of priming based on gender identity. All surveys and quizzes were administered to classes taught by the same instructor, which allowed us to control for instructor effects.

From administering the survey to the two sections across two quizzes, we were able to gather N_{survey} = 269 usable responses from the original 270 responses; one observation was dropped because the respondent did not indicate a gender. Because the same classes took Quiz 2 and Quiz 3 in succession, the structure of the data for the survey $N_{\text{survey}} = 269$ can be regarded as a pooled cross-section where the time dimension is captured by the quiz number. The data from the quiz itself, however, can be treated as a panel because we are able to link quiz grades to the students over a time dimension that is captured by the quiz number.

The fact that some students in our sample did not take all three quizzes, either due to missing class on the day that the quiz was given or to dropping the course, could lead to attrition bias. We decided to perform our analysis of the grade effects of priming by creating a balanced panel consisting only of data from students who took all three quizzes. This choice also afforded us a way to control for a practice effect, whereby a student's performance might change, not due to priming, but rather due to having experience with the structure of the instructor's quizzes. Based on $n = 124$ unique students across two classes, with $q = 3$ quizzes per student, our balanced panel contained $N_{\text{quiz}} = n \times q = 124 \times 3 = 372$ responses.

Table 2 gives the summary statistics for these two samples. The average of Difficulty was 66.41 on the 0 to 100 scale, indicating that students overall tended to view the quiz problems as more than moderately difficult. The average Familiarity level on a 1 to 5 scale was 3.42, indicating that students were slightly more than moderately familiar with the material.

Table 2: Means and Standard Deviations for Variables in the Survey and Quiz

	Data Source			
	Survey (*N* = 269)		Quiz (*N* = 372)	
Variable	*M*	*SD*	*M*	*SD*
Difficulty	66.41	20.51	-	-
Familiarity	3.42	0.95	-	-
Female	0.52	0.50	0.53	0.50
Grade	-	-	15.15	4.51

Note: M and *SD* are the mean and standard deviation, respectively. Female = 1 if a student is female, and 0 otherwise. The survey and quiz sample sizes differ because the survey was anonymous and was only given for Quiz 2 and 3, while we have grade data for all three quizzes.

Both samples were almost evenly split by gender, with female students making up 52% of the survey sample and 53% of the quiz sample. The average quiz grade across all priming conditions (including the None condition) was 15.15 out of a maximum of 20 points.

A typical student in the panel would have had three grades: one for Quiz 1 (which contained no priming statement for either class), one for Quiz 2, and one for Quiz 3. Because of our study design, a student (i) and quiz (j) combination defines a priming condition. Therefore, a basic random effects panel model relating the grade on quiz j for student i is given is as follows:

$$Grade_{ij} = \beta_0 + \beta_1 Adv_{ij} + \beta_2 None_{ij} + v_i + \varepsilon_{ij} \tag{1}$$

where Adv_{ij} and $None_{ij}$ are indicator variables for the Advanced and None priming conditions, respectively, $v_i \sim N(0, \tau)$ is a student random effect, and $\varepsilon_{ij} \sim N(0, \sigma)$ is the individual student-quiz random error. The random effects v_i and the errors ε_{ij} are assumed mutually independent. Given evidence in the literature that priming can have differential effects based on gender, our next empirical model allows for estimation of an interaction effect of gender on priming through the indicator variable Fem_i, which equals 1 if student i is female and 0 otherwise. This gives the following model:

$$\begin{aligned} Grade_{ij} = \beta_0 + \beta_1 Adv_{ij} + \beta_2 None_{ij} + \beta_3 Fem_i + \beta_4 (Fem \times Adv)_{ij} \\ + \beta_5 (Fem \times None)_{ij} + v_i + \varepsilon_{ij} \end{aligned} \tag{2}$$

where v_i and ε_{ij} are defined as for Model 1. We also consider a model with a quiz order effect because students might find later material to be more (or less) difficult than earlier material regardless of the type of priming. Model 3 incorporates these quiz ordering effects as the dummy variables $Quiz1_j$ and $Quiz3_j$:

$$\begin{aligned} Grade_{ij} = \beta_0 + \beta_1 Adv_{ij} + \beta_2 None_{ij} + \beta_3 Fem_i + \beta_4 (Fem \times Adv)_{ij} \\ + \beta_5 (Fem \times None)_{ij} + \beta_6 Quiz1_j + \beta_7 Quiz3_j + v_i + \varepsilon_{ij} \end{aligned} \tag{3}$$

Finally, we consider a model with a class effect to account, in part, for differences in performance that might arise due to the possibility of students in the MW class sharing information about the quiz to students in the TTH class (or vice versa). The class effect would also account for the possibility that the day of the week or even the time of day could impact performance (e.g., Carrell, et al. 2011). This effect is added to Model 3 to produce Model 4:

$$\begin{aligned} Grade_{ij} = \beta_0 + \beta_1 Adv_{ij} + \beta_2 None_{ij} + \beta_3 Fem_i + \beta_4 (Fem \times Adv)_{ij} \\ + \beta_5 (Fem \times None)_{ij} + \beta_6 Quiz1_j + \beta_7 Quiz3_j + \beta_8 TTH_i \\ + v_i + \varepsilon_{ij} \end{aligned} \tag{4}$$

where, again, v_i and ε_{ij} are defined as for Model 1. In all of our models, the Basic priming condition is the reference category.

To address our secondary research question of what effect, if any, priming has on perceptions of difficulty, we examine three empirical models of increasing complexity. We code the priming and gender conditions in the same manner as for the main analysis: the dummy variables Adv and $None$ equal 1 for their respective priming conditions, and 0 otherwise, and $Fem = 1$ if the student is female, and 0 otherwise. Each of the models for Difficulty also includes the control variable Fam, the student's self-reported familiarity with the material. Basic priming is again the omitted category.

Model 5, the simplest of these, examines first-order priming effects on perceived difficulty, and is given as follows:

$$Diff_i = \beta_0 + \beta_1 Adv_i + \beta_2 None_i + \beta_3 Fam_i + \varepsilon_i \tag{5}$$

To capture the possibility of a gender effect on difficulty perceptions, we expand Model 5 using the gender dummy variable Fem, resulting in Model 6:

$$Diff_i = \beta_0 + \beta_1 Adv_i + \beta_2 None_i + \beta_3 Fem_i + \beta_4 Fam_i + \varepsilon_i \tag{6}$$

Finally, we examine an empirical model that allows for differential effects of priming based on gender by interacting Fem with the priming conditions; this gives Model 7:

$$\begin{aligned} Diff_i = \beta_0 + \beta_1 Adv_i + \beta_2 None_i + \beta_3 Fem_i + \beta_4 (Fem \times Adv)_i \\ + \beta_5 (Fem \times None)_i + \beta_6 Fam_i + \varepsilon_i \end{aligned} \tag{7}$$

For all models, we assume the random error term $\varepsilon_i \sim N(0, \sigma)$.

In the next section, we report descriptive statistics and the results of the main analysis on priming and performance based on Models 1 through 4. We then address the secondary research question of priming and perceptions of difficulty based on Models 5 through 7.

Results

Priming and Performance

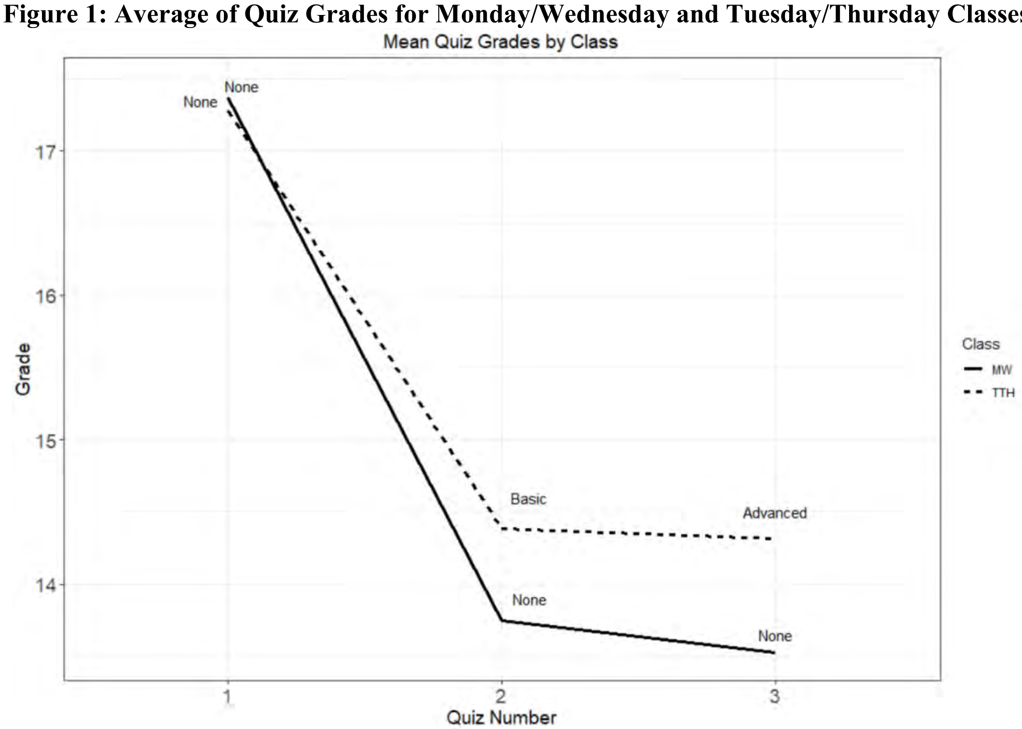
Table 3 gives the means and standard deviations for grades for all three quizzes by class and priming condition (which is tied to the class in our design). This information is depicted in Figure 1 as a time series average of the three quiz scores between the two classes as well as the priming conditions associated with each quiz. Students in both classes did the best on Quiz 1, which contained no priming statement for either class. Neither class significantly outperformed the other on this first quiz ($t(88.8) = -0.384, p = 0.702$). Differences in the sample means for performance appear for Basic and Advanced priming, but in what seems like a counterintuitive pattern both Basic and Advanced priming are associated with higher quiz performance compared to no priming. Because the quizzes were given over the course of the semester, we must consider that performance could be higher on Quiz 1 than Quiz 3 simply due to the earlier course content being easier for the students than the later quiz content, irrespective of priming. This is accounted for in Model 3 with the Quiz 1 and 3 fixed effects. Because only one class received the treatments, there is, as we have mentioned, a reason to investigate class effects, which are captured by Model 4.

Table 3: Quiz Grade Summary Statistics by Class and Priming Condition

	Quiz Number (Priming Condition)							
	1 (None)		2 (None or Basic)		3 (None or Advanced)		Marginal	
Class	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
MW	17.37	2.36	13.75	5.04	13.53	4.58	14.88	4.50
TTH	17.59	3.76	14.43	4.66	14.39	4.41	15.47	4.53
Marginal	17.47	3.06	14.06	4.86	13.92	4.51		

Note: Quiz 1 for both classes contained no priming statement. The MW class had no priming ("None") for all three quizzes. The TTH class had "Basic" priming for Quiz 2 and "Advanced" priming for Quiz 3. Overall means for classes and priming conditions are given under the "Marginal" headings.

Figure 1: Average of Quiz Grades for Monday/Wednesday and Tuesday/Thursday Classes



Note: There was no priming for Quiz 1 in either class. In the Tuesday/Thursday class, Quiz 2 used Basic priming and Quiz 3 used Advanced priming.

Table 4 contains the results of fitting Models 1 through 4. Model 1 suggests that students who receive no priming (None) tend to perform better on average than students receiving Basic priming, while Model 2 suggests that males fare better with no priming versus basic priming ($b_2 = 2.564, z = 2.961, p = 0.003$), with no significant effect for females. Both models have small adjusted r-squared values, however, and do not account for changes in performance over time. Models 3 and 4, which add the quiz number and class effects, perform much better, with adjusted r-squared values of 0.199 and 0.197, respectively. In both of these models, the None priming condition is no longer significant; only the dummy for Quiz 1 is significant and positive, which reflects the significant drop in the mean grade after Quiz 1 that we see in Figure 1. Model 4, which includes the Tuesday/Thursday class dummy variable TTH, performs slightly worse than Model 3. Moreover, we do not find a significant class effect. Because Model 3 is the more parsimonious of the two in addition to having the highest adjusted r-squared, we prefer it. Overall, however, we find no significant effect of priming on performance after controlling for quiz fixed effects. The significant dummy variable coefficient for Quiz 1 ($b_6 = 3.676$, $z = 7.124$, $p < 0.001$) indicates that the mean grade for students in both classes was significantly higher on the first quiz, which contained no priming statements and no manipulation of difficulty, than on the second quiz, which contained Basic priming. This is reflected in Figure 1.

Table 4: Random Effects Panel Regression Results for Models 1 through 4

	Dependent variable:			
	Grade			
	(1)	(2)	(3)	(4)
Advanced	-0.036	0.536	0.756	0.756
	(0.692)	(0.979)	(1.025)	(1.026)
None	1.929^{***}	2.564^{***}	0.005	0.131
	(0.608)	(0.866)	(0.889)	(0.990)
Female		1.112	1.172	1.176
		(1.156)	(1.063)	(1.065)
Female.Advanced		-1.143	-1.143	-1.143
		(1.384)	(1.220)	(1.221)
Female.None		-1.243	-1.182	-1.176
		(1.218)	(1.096)	(1.097)
Quiz1			3.676^{***}	3.618^{***}
			(0.516)	(0.554)
Quiz3			-0.221	-0.221
			(0.553)	(0.554)
TTH				0.222
				(0.766)
Constant	13.805^{***}	13.244^{***}	13.792^{***}	13.619^{***}
	(0.576)	(0.819)	(0.758)	(0.966)
Observations	372	372	372	372
Adjusted R^2	0.034	0.030	0.199	0.197

Note: Standard errors in parentheses. "Basic" priming is the omitted category. "[Variable 1].[Variable 2]" indicates the interaction coefficient for Variable 1 and Variable 2. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

Thus, the addition of priming appears to not have caused any significant change, for better or for worse, in the actual performance of students. This could be due to the changes in the difficulty of the material on the quizzes not being sufficiently large for priming to have a noticeable impact. Based on their performance, non-primed students (MW class) did not find the material on Quiz 3 to be significantly harder compared to Quiz 2 ($b_7 = -0.221$, $z = -0.399$, $p = 0.690$), although it was intended to be based on our choice of problems according to the test bank difficulty rating system. The question remains, however, whether priming can influence the *perception* of difficulty. We present results for this question in the next section.

Priming and Perceived Difficulty

Table 5 presents the summary statistics for perceived difficulty ("Difficulty") tabulated by priming condition and gender. Overall, male students tended to rate the difficulty of the problems slightly higher than females at 68.1 versus 64.86, respectively. The means for the priming conditions had the expected ordering, with Basic priming resulting in the lowest difficulty (M = 60.95, SD = 20.38), Advanced priming resulting in the highest difficulty (M = 70.63, SD = 18.04), and no priming (None) giving results approximately halfway between Advanced and Basic (M = 66.96, SD = 21.20).

Table 5: Perceived Difficulty Summary Statistics by Gender and Priming Condition

	Priming Condition (Class)							
	Advanced (TTH)		Basic (TTH)		None (MW)		Marginal	
Gender	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Female	73.00	19.32	57.58	19.85	64.81	21.86	64.86	21.37
Male	68.48	16.79	64.67	20.63	69.47	20.27	68.10	19.48
Marginal	70.63	18.04	60.95	20.38	66.96	21.20		

Note: "Marginal" row and column values summarize Difficulty by gender and priming condition, respectively. N = 269. Advanced and Basic priming was only given to the TTH class. The MW class got no priming.

Table 6: Regression Results for Models 5, 6, and 7

	Dependent variable:		
	Difficulty		
	(5)	(6)	(7)
Advanced	11.036***	10.901***	7.909*
	(3.329)	(3.329)	(4.733)
None	7.370***	7.405***	8.894**
	(2.828)	(2.825)	(4.141)
Female		-2.737	-2.742
		(2.278)	(4.738)
Female.Advanced			6.280
			(6.686)
Female.None			-2.768
			(5.689)
Familiarity	-8.528***	-8.498***	-8.491***
	(1.203)	(1.202)	(1.213)
Constant	89.110***	90.442***	90.423***
	(4.615)	(4.743)	(5.009)
Observations	269	269	269
Adjusted R^2	0.173	0.174	0.176

Note: Standard errors in parentheses. "Basic" priming is the omitted category. "[Variable 1].[Variable 2]" indicates the interaction coefficient for Variable 1 and Variable 2. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

The results from fitting Models 5 through 7 are shown in Table 6. The models perform similarly, but by the criterion of adjusted r-squared, Model 7 is preferred. The Familiarity control variable is highly significant in the expected direction, indicating that a one-level increase in a student's self-reported familiarity with the quiz content tends to reduce perceived difficulty by about 8.5 points on average, irrespective of any priming effect. The coefficient for the None priming condition, $b_2 = 8.894$, is significant ($t(262) = 2.148$, $p = 0.033$), and interpretable as the estimated difference in mean perceived difficulty between males given no priming and males given Basic priming; males who were not primed tended to perceive the problem to be more difficult than males who were given Basic priming. The coefficient for Advanced priming, highly significant in Models 5 and 6 ($p < 0.01$ for both), becomes marginally significant in Model 7 ($b_1 = 7.909$,

$t(262) = 1.671$, $p = 0.096$). This coefficient is interpretable as the estimated difference in mean perceived difficulty between the Advanced and Basic priming conditions for male students, controlling for familiarity.

Model 7 also permits us to examine linear restriction hypotheses comparing the various priming conditions to one another. The results of seven such hypotheses are shown in Table 7. We find overall, using full/restricted model F tests, significant positive differences in mean perceived difficulty between the Advanced and Basic conditions, and the Basic and None conditions, but not between the Advanced and None conditions. Advanced priming increases perceived difficulty over Basic priming by about 11 points on average ($p = 0.001$), and Basic priming decreases it by about 7.5 points ($p = 0.008$), controlling for familiarity. For female students, the priming effect is stronger, at about 14 points, between the Advanced and Basic conditions, while there is no significant effect of priming in female students between the Basic and None conditions ($p = 0.115$). The equivalent comparisons for male students are calculated using the coefficient estimates, except for the comparison of Advanced versus no priming for male students, which is the final linear restriction hypothesis in Table 7. We find no significant effect of Advanced priming for males ($p = 0.804$).

Table 7: Linear Hypotheses Comparing Perceived Difficulty Between Priming Conditions

Comparison	Null Hypothesis	Estimate	F Statistic	P Value
Advanced Vs. Basic†	$\beta_1 + \frac{\beta_4}{2} = 0$	11.049 (3.328)	11.025	0.001
Advanced Vs. None	$\beta_1 - \beta_2 + \frac{\beta_4 - \beta_5}{2} = 0$	3.539 (2.820)	1.574	0.211
Basic Vs. None†	$-\beta_2 - \frac{\beta_5}{2} = 0$	-7.51 (2.828)	7.05	0.008
Female: Advanced Vs. Basic†	$\beta_1 + \beta_4 = 0$	14.189 (4.700)	9.114	0.003
Female: Advanced Vs. None†	$\beta_1 + \beta_4 - \beta_2 - \beta_5 = 0$	8.062 (4.007)	4.048	0.045
Female: Basic Vs. None	$-\beta_2 - \beta_5 = 0$	-6.127 (3.877)	2.497	0.115
Male: Advanced Vs. None	$\beta_1 - \beta_2 = 0$	-0.985 (3.970)	0.062	0.804

Note: A comparison significant at the 0.05 level is indicated by a †. Standard errors of the estimated linear restrictions are given in parentheses in the "Estimate" column. The F statistics are the squares of the t statistics obtainable by dividing the estimates by their respective standard errors.

Because of the nature of the data-gathering process, there is the possibility that the effects we have so far attributed to priming could be caused by other factors such as the difference in material covered on Quiz 2 versus Quiz 3 and class effects. In the next section, we discuss some robustness checks we employed to address these issues.

Robustness Checks for Perceived Difficulty

In our design, the material covered on Quiz 2 was the same for both the MW and TTH classes (although with non-identical questions) and similarly for Quiz 3. Therefore, if we have evidence that students, on average, did not perceive a difference in difficulty between these quizzes, our findings in the previous section are more robust, and we have more reason to believe that priming is the factor driving the differences. To investigate this possibility, we used the fact that the class receiving the priming treatment was always the TTH section to fit the following interaction model:

$$Diff_i = \gamma_0 + \gamma_1 Q3_i + \gamma_2 TTH_i + \gamma_3 Fem_i + \gamma_4 (Q3 \times TTH)_i + \gamma_5 (Q3 \times Fem)_i + \gamma_6 (TTH \times Fem)_i + \gamma_7 (Q3 \times TTH \times Fem)_i + \gamma_8 Fam_i + \varepsilon_i \quad (8)$$

where $\varepsilon_i \sim N(0, \sigma)$ is again the random error term and $Q3$, TTH, and Fem are dummy variables: $Q3 = 1$ if the student took Quiz 3, and 0 if the student took Quiz 2; $TTH = 1$ if the student was in the TTH class, and 0 if the student was in the MW class; and $Fem = 1$ if the student was female, and 0 otherwise. The results of fitting Model 8 are shown in Table 8.

Table 8: Model for the Robustness Checks for Quiz Content Effect

	Dependent variable:
	Difficulty
Q3	-4.029
	(4.588)
TTH	-10.809**
	(4.655)
Female	-10.350**
	(4.332)
Q3.TTH	11.960*
	(6.574)
Q3.FEM	9.993
	(6.245)
TTH.FEM	7.631
	(6.429)
Q3.TTH.FEM	-3.742
	(9.146)
Familiarity	-8.538***
	(1.213)
Constant	101.373***
	(5.240)
Observations	269
Adjusted R^2	0.178

Note: Standard errors in parentheses. "[Variable 1].[Variable 2]" indicates the interaction coefficient for Variable 1 and Variable 2. *p < .1; **p < .05; ***p< .01

The relevant linear hypothesis that allows for an examination of a possible effect of quiz material on mean perceived difficulty is, conceptually, the difference between the average of the Quiz 3 and Quiz 2 difficulties for male and female students in the MW class. Using cell means notation (Kutner et al. 2005, chap. 16) to aid in exposition, we can write this null hypothesis as

$$H_{01}: \frac{\mu_{M,Q3,MW} + \mu_{F,Q3,MW}}{2} - \frac{\mu_{M,Q2,MW} + \mu_{F,Q2,MW}}{2} = 0,$$

where the subscripts M and F denote "male" and "female," respectively, and $Q3$ and $Q2$ denote Quiz 3 and Quiz 2, respectively. This can be expressed in terms of the coefficients of Model 8 as follows:

$$H_{01}: \gamma_1 + \frac{\gamma_5}{2} = 0$$

The point estimate using the output in Table 8 is $\hat{\gamma}_1 + \frac{\hat{\gamma}_5}{2} = 0.968$; however, it does not reach statistical significance ($F(1, 260) = 0.096$, $p = 0.757$). We therefore cannot conclude that students in the MW section perceived a difference in difficulty, on average, between quizzes 2 and 3 that could be attributed to changes in the difficulty of the material itself. This is consistent with the appearance of the "None" line in Figure 1 and the findings in the previous section.

However, when looking at just Quiz 2, we are able to conclude that female students in the MW class found Quiz 2 to be less difficult than did males in the MW class ($\hat{\gamma}_3 = -10.35$, $t(260) = -2.389$, $p = 0.018$).

We also find evidence that males in the TTH class (the group that received Basic priming) perceived the difficulty of Quiz 2 to be less on average than males in the MW (non-primed) class taking the quiz covering the same topics ($\hat{\gamma}_2 = -10.809$, $t(260) = -2.322$, $p = 0.021$). Thus, Basic priming reduces difficulty for males by about 11 points on average, controlling for familiarity.

We would like to determine whether there is evidence of a priming effect when considering differences in content between the quizzes. We can test this hypothesis by examining the difference in mean difficulty between quizzes with Advanced priming versus quizzes with Basic priming, net of any difficulty difference that could be attributed to the material on the quiz and familiarity and averaged over gender. Again, using cell means notation, we can express this hypothesis as follows:

$$H_{02}: \left(\frac{\mu_{M,Q3,TTH} + \mu_{F,Q3,TTH}}{2} - \frac{\mu_{M,Q2,TTH} + \mu_{F,Q2,TTH}}{2}\right) - \left(\frac{\mu_{M,Q3,MW} + \mu_{F,Q3,MW}}{2} - \frac{\mu_{M,Q2,MW} + \mu_{F,Q2,MW}}{2}\right) = 0$$

Conceptually, H_{02} compares the difference in the change in mean difficulty between the primed Quiz 2 and Quiz 3 and the non-primed Quiz 2 and Quiz 3. This hypothesis can be expressed in terms of a linear restriction on the coefficients of Model 8 as follows:

$$H_{02}: \gamma_4 + \frac{\gamma_7}{2} = 0$$

The point estimate $\hat{\gamma}_4 + \frac{\hat{\gamma}_7}{2} = 10.089$ indicates that students perceive quizzes given with Advanced priming to be more difficult than those given with Basic priming ($F(1{,}260) = 4.906$, $p = 0.028$), and this difference is not attributable simply to differences in quiz material. Thus, our finding in the previous section, that priming a problem to be Advanced has a significant effect on perceived difficulty versus priming a problem to be Basic, is supported.

Discussion and Suggestions for Future Research

In this paper, we examined the effect that priming statements communicating a judgment about the difficulty of quiz material have on actual student performance as well as their perception of difficulty, irrespective of their actual performance. Students in quantitative courses frequently encounter such statements when reading textbooks, for example, in the form of transitional statements in the vein of "It is easy to see that…" or "Obviously, we have…" or from a live lecture peppered with statements that "it is trivial" to see some fact, or how "a basic calculation" will show some result. Our first research question was whether statements priming students with regard to difficulty could influence actual difficulty as measured by quiz grades. The secondary question was whether students' perceptions of difficulty could be similarly influenced. We examined both research questions using survey and quiz data from an undergraduate business statistics course.

Overall, we find the effects of priming to be mixed. We do not find that priming statements significantly affect *actual* student performance, but we do find evidence that priming influences *perceptions* of difficulty. We find that a statement priming students to perceive a problem as having Advanced difficulty results in a higher average perception of difficulty compared to a statement priming for Basic difficulty. A statement suggesting a problem is of Basic difficulty results in a decrease in perceived difficulty relative to no priming. On the other hand, we cannot conclude that students given an Advanced priming statement will tend to perceive a problem to be more difficult than students given no priming at all. Thus, if an instructor wishes to reduce perceived difficulty (for example, to mitigate feelings of anxiety among students), a Basic prime may help. However, if the goal is to build students' confidence by suggesting that they are solving challenging problems, our findings indicate that such an intervention may not help or hurt compared with telling the students nothing.

Through subgroup analyses, we find some significant gender differences in the effects of priming on difficulty perception. Advanced priming increases perceived difficulty for female students over no priming, but it does not have a significant effect on male students over no priming. Further, female students perceive Advanced-primed material to be significantly more difficult than Basic-primed material, but do not appear to be helped by Basic priming over no priming at all. These results are consistent with other literature that has found that some priming effects can be moderated by gender (e.g., Harrison et al. 2009; Steele and Ambady 2006; Thompson and Musket 2005).

Our finding of differential priming effects based on gender has important implications. While programs in science, technology, engineering, and mathematics (STEM) are gaining popularity, a gender gap persists. Compared to men, women tend to possess a smaller percentage of awarded STEM degrees (Beede et al. 2011; Stoet and Geary 2018), report reduced intentions to pursue math as a field of study (Bench et al. 2015), and tend to have lower self-efficacy and self-concept (Robnett 2016). These gender gaps in STEM are particularly interesting considering research that finds female students tend to have higher overall achievement than male students as measured by grades, but with a gap that narrows in courses with quantitative content (Voyer and Voyer 2014). Therefore, one future avenue of research would be an investigation of the potential for priming to activate gender-based differences in attitudes toward mathematics through gender-based learning styles (Kulturel-Konak et al. 2011; Orhun 2007).

The fact that Basic priming has the ability to lower perceived difficulty, but does not appear to affect the actual difficulty students experience, suggests that including priming statements could be a way to address the issues of retention (that is, keeping students enrolled in the course) and persistence (i.e., encouraging students to take additional courses within a given discipline). Priming, operating unconsciously by definition, can motivate individuals to achieve goals with the same level of perseverance and performance as goals that are set consciously (Bargh 2014; Custers and Aarts 2010). Students' quantitative skills are a strong predictor of performance in introductory economics classes (Ballard and Johnson 2004; Schuhmann et al. 2005). Moreover, the introductory economics class is one that has a pivotal role in the pathway to the economics major. Emerson and McGoldrick (2019) found, using data on nearly 97,000 students over 23 years, that about 45% of students declaring an economics major at the time of taking their first principles course switched to other majors at the conclusion of the course. The ability of priming to alter perceptions of difficulty could help encourage more students to stay in the major. Our present study highlights the short-term effects of priming on perceptions of difficulty. A future study could examine the longer-term effects of priming on student retention and persistence within the major by using longitudinal data on students gathered from their first declaration of a major through graduation.

Another important question that remains to be answered is how robust priming effects are. Our study methodology involved no deception: if students were told a problem was Advanced (or Basic), it was indeed rated as being Advanced (or Basic) according to the test bank from which the question came. If students were given problems actually rated as Basic but were told the problems were Advanced, or if they were given Advanced-rated problems and told they were Basic, would there be an observable priming effect on perceived difficulty and/or performance? A true factorial experiment crossing priming conditions with problem content could be used. In addition to examining effects by gender, such an analysis could be made richer by incorporating additional student attributes beyond self-reported familiarity with the material, such as a measure of self-efficacy, GPA, major, age, baseline math aptitude, and other factors.

In addition to examining the question of the effectiveness of priming using different outcome measures and experimental designs, it would be useful to determine what form of priming has the greatest effect on student performance. Many forms and modalities of priming exist. In our study, the prime came in the form of a printed statement on the quiz that the instructor also read aloud word for word. However, instructor judgments of difficulty can also come in much more informal ways, for example, as offhand in-class comments such as "Okay, everyone. Let's work an easy one." It would be informative to examine whether the form of the prime matters in how students perceive the problem.

Regardless of its purpose, form, or origin, priming has clear relevance for instructors of required quantitative courses who face pressure to make material accessible. Radel et al. (2009) report that students who find it difficult to be in "mindfulness," a state in which a person focuses on thoughts and sensations of the present, respond positively to priming through enhanced academic performance. The effect of priming, moreover, can be greater for high-value goals (Förster et al. 2009; Weingarten et al. 2016), and few investments have higher lifetime value than a college degree. Given the complex links among perceptions of difficulty, attitude, and achievement, particularly in statistics courses (Chiesi and Caterina 2010), priming has the potential to serve as a "low-cost and effective tool…to promote desirable behavior" (Cohn and Maréchal 2016, p. 20). If nothing else, instructors can take heart that, at least sometimes, our students listen to us.

References

Aguilar, Mike, and Daniel Soques. 2015. "Fiscal Challenge: An Experiential Exercise in Policy Making."

The Journal of Economic Education 46: 285-299.

Allgood, Sam, William B. Walstad, and John J. Siegfried. 2015. "Research on Teaching Economics to Undergraduates." *Journal of Economic Literature* 53: 285-325.

Ballard, Charles L., and Marianne F. Johnson. 2004. "Basic Math Skills and Performance in an Introductory Economics Class." *The Journal of Economic Education* 35: 3-23.

Bandura, Albert. 1977. "Self-Efficacy: Toward a Unifying Theory of Behavioral Change." *Psychological Review* 84: 191-215.

Bandura, Albert. 1982. "Self-Efficacy Mechanism in Human Agency." *American Psychologist* 37: 122-147.

Bandura, Albert. 1989. "Regulation of Cognitive Processes through Perceived Self-Efficacy." *Developmental Psychology* 25: 729-735.

Bandura, Albert. 1997. *Self-Efficacy: The Exercise of Control*. New York, NY: W.H. Freeman and Company.

Bandura, Albert. 2006. "Guide for Constructing Self-Efficacy Scales." In *Self Efficacy Beliefs of Adolescents*, edited by Frank Pajares and Tim Urdan. Greenwich, CT: IAP.

Bargh, John A. 2014. "The Historical Origins of Priming as the Preparation of Behavioral Responses: Unconscious Carryover and Contextual Influences of Real-World Importance." *Social Cognition* 32 (Supplement): 209-224.

Bargh, John A., Annette Lee-Chai, Kimberly Barndollar, Peter M. Gollwitzer, and Roman Trötschel. 2001. "The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals." *Journal of Personality and Social Psychology* 81: 1014-1027.

Beede, David N., Tiffany A. Julian, David Langdon, George McKittrick, Beethika Khan, and Mark E. Doms. 2011. "Women in STEM: A Gender Gap to Innovation." SSRN Scholarly Paper ID 1964782. Rochester, NY: Social Science Research Network.

Bench, Shane W., Heather C. Lench, Jeffrey Liew, Kathi Miner, and Sarah A. Flores. 2015. "Gender Gaps in Overestimation of Math Performance." *Sex Roles* 72: 536-546.

Benedict, Mary Ellen, and John Hoag. 2002. "Who's Afraid of Their Economics Classes? Why Are Students Apprehensive about Introductory Economics Courses? An Empirical Investigation." *The American Economist* 46: 31-44.

Betz, Nancy E. 1978. "Prevalence, Distribution, and Correlates of Math Anxiety in College Students." *Journal of Counseling Psychology* 25: 441-448.

Blackwell, Lisa S., Kali H. Trzesniewski, and Carol Sorich Dweck. 2007. "Implicit Theories of Intelligence Predict Achievement across an Adolescent Transition: A Longitudinal Study and an Intervention." *Child Development* 78: 246-263.

Carrell, Scott E., Teny Maghakian, and James E. West. 2011. "A's from Zzzz's? The Causal Effect of School Start Time on the Academic Achievement of Adolescents." *American Economic Journal: Economic Policy* 3: 62-81.

Chang, Hyesang, and Sian L Beilock. 2016. "The Math Anxiety-Math Performance Link and Its Relation to Individual and Environmental Factors: A Review of Current Behavioral and Psychophysiological Research." *Current Opinion in Behavioral Sciences* 10: 33-38.

Chartrand, Tanya L., and John A. Bargh. 1996. "Automatic Activation of Impression Formation and

Memorization Goals: Nonconscious Goal Priming Reproduces Effects of Explicit Task Instructions." *Journal of Personality and Social Psychology* 71: 464-478.

Chiesi, Francesca, and Primi Caterina. 2010. "Cognitive and Non-Cognitive Factors Related to Students' Statistics Achievement." *Statistics Education Research Journal* 9: 6-26.

Chulkov, Dmitriy, and Dmitri Nizovtsev. 2015. "Problem-Based Learning in Managerial Economics with an Integrated Case Study." *Journal of Economics and Economic Education Research* 16: 188-197.

Cohn, Alain, and Michel André Maréchal. 2016. "Priming in Economics." *Current Opinion in Psychology* 12: 17-21.

Custers, Ruud, and Henk Aarts. 2010. "The Unconscious Will: How the Pursuit of Goals Operates Outside of Conscious Awareness." *Science* 329: 47-50.

Dijksterhuis, Ap, and Ad van Knippenberg. 1998. "The Relation between Perception and Behavior, or How to Win a Game of Trivial Pursuit." *Journal of Personality and Social Psychology* 74: 865-877.

Dillon, Kathleen M. 1982. "Statisticophobia." *Teaching of Psychology* 9: 117.

Edwards, Finley. 2012. "Early to Rise? The Effect of Daily Start Times on Academic Performance." *Economics of Education Review* 31: 970-983.

Elliot, Andrew J., Markus A. Maier, Arlen C. Moller, Ron Friedman, and Jörg Meinhardt. 2007. "Color and Psychological Functioning: The Effect of Red on Performance Attainment." *Journal of Experimental Psychology: General* 136: 154-168.

Emerson, Tisha L. N., and KimMarie McGoldrick. 2019. "Switching Majors – into and out of Economics." *The Journal of Economic Education* 50: 321-332.

Fike, David S., and Renea Fike. 2013. "A Multilevel Analysis of the Association of Class Schedule with Student Outcomes in Community College Developmental Math." *Community College Journal of Research and Practice* 37: 816-827.

Finney, Sara J., and Gregory Schraw. 2003. "Self-Efficacy Beliefs in College Statistics Courses." *Contemporary Educational Psychology* 28: 161-186.

Förster, Jens, Nira Liberman, and Ronald S. Friedman. 2009. "What Do We Prime? On Distinguishing between Semantic Priming, Procedural Priming, and Goal Priming." In *Oxford Handbook of Human Action*, edited by Ezequiel Morsella, John A. Bargh, and Peter M. Gollwitzer. New York, NY: Oxford University Press.

Gal, Iddo, and Lynda Ginsburg. 1994. "The Role of Beliefs and Attitudes in Learning Statistics: Towards an Assessment Framework." *Journal of Statistics Education* 2: 1-15.

Garfield, Joan, Martha Aliaga, George W. Cobb, Carolyn Cuff, Rob Gould, Robin Lock, Tom Moore, Allan Rossman, Bob Stephenson, Jessica Utts, Paul Velleman, and Jeff Witmer. 2005. *Guidelines for Assessment and Instruction in Statistics Education (GAISE): College Report*. Alexandria, VA: American Statistical Association.

Harrison, C. Keith, Jeff Stone, Jenessa Shapiro, Sharon Yee, Jean A. Boyd, and Vashti Rullan. 2009. "The Role of Gender Identities and Stereotype Salience with the Academic Performance of Male and Female College Athletes." *Journal of Sport and Social Issues* 33: 78-96.

Hembree, Ray. 1990. "The Nature, Effects, and Relief of Mathematics Anxiety." *Journal for Research in Mathematics Education* 21: 33-46.

Henrich, Allison, and Kristi Lee. 2011. "Reducing Math Anxiety: Findings from Incorporating Service Learning into a Quantitative Reasoning Course at Seattle University." *Numeracy* 4: Article 9.

Hoag, John, and Mary Ellen Benedict. 2010. "What Influence Does Mathematics Preparation and Performance Have on Performance in First Economics Classes?" *Journal of Economics and Economic Education Research* 11: 19-42.

Hoffman, Bobby, and Gregory Schraw. 2009. "The Influence of Self-Efficacy and Working Memory Capacity on Problem-Solving Efficiency." *Learning and Individual Differences* 19: 91-100.

Hoffman, Bobby, and Alexandru Spatariu. 2008. "The Influence of Self-Efficacy and Metacognitive Prompting on Math Problem-Solving Efficiency." *Contemporary Educational Psychology* 33: 875-893.

Jackson, Carol D., and R. Jon Leffingwell. 1999. "The Role of Instructors in Creating Math Anxiety in Students from Kindergarten through College." *The Mathematics Teacher* 92: 583-86.

Jackson, Thad. 2016. "Priming Performance of Economics Students: Even More Evidence." SSRN Scholarly Paper ID 2877005. Rochester, NY: Social Science Research Network.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York, NY: Farrar, Straus, and Giroux.

Kamenica, Emir. 2012. "Behavioral Economics and Psychology of Incentives." *Annual Review of Economics* 4: 427-452.

Kulturel-Konak, Sadan, Mary Lou D'Allegro, and Sarah Dickinson. 2011. "Review of Gender Differences in Learning Styles: Suggestions for STEM Education." *Contemporary Issues in Education Research* 4: 9-18.

Kutner, Michael H., Christopher J. Nachtsheim, John Neter, and William Li. 2005. *Applied Linear Statistical Models* (5th ed.). New York, NY: McGraw-Hill Irwin.

Leung, Shing On. 2011. "A Comparison of Psychometric Properties and Normality in 4-, 5-, 6-, and 11-Point Likert Scales." *Journal of Social Service Research* 37: 412-421.

Molden, Daniel C. 2014. "Understanding Priming Effects in Social Psychology: What Is 'Social Priming' and How Does It Occur?" *Social Cognition* 32 (Supplement): 1-11.

Orhun, Nevin. 2007. "An Investigation into the Mathematics Achievement and Attitude towards Mathematics with Respect to Learning Style According to Gender." *International Journal of Mathematical Education in Science and Technology* 38: 321-333.

Pajares, Frank, James Hartley, and Giovanni Valiante. 2001. "Response Format in Writing Self-Efficacy Assessment: Greater Discrimination Increases Prediction." *Measurement and Evaluation in Counseling Development* 33: 214-221.

Pajares, Frank, and M. David Miller. 1994. "Role of Self-Efficacy and Self-Concept Beliefs in Mathematical Problem Solving: A Path Analysis." *Journal of Educational Psychology* 86: 193-203.

Pajares, Frank, and M. David Miller. 1995. "Mathematics Self-Efficacy and Mathematics Performances: The Need for Specificity of Assessment." *Journal of Counseling Psychology* 42: 190-198.

Pope, Nolan G. 2015. "How the Time of Day Affects Productivity: Evidence from School Schedules." *The Review of Economics and Statistics* 98: 1-11.

Radel, Rémi, Philippe Sarrazin, Pascal Legrain, and Lucie Gobancé. 2009. "Subliminal Priming of Motivational Orientation in Educational Settings: Effect on Academic Performance Moderated by Mindfulness." *Journal of Research in Personality* 43: 695-698.

Rattan, Aneeta, Catherine Good, and Carol S. Dweck. 2012. "'It's Ok — Not Everyone Can Be Good at Math': Instructors with an Entity Theory Comfort (and Demotivate) Students." *Journal of Experimental Social Psychology* 48: 731-737.

Robnett, Rachael D. 2016. "Gender Bias in STEM Fields: Variation in Prevalence and Links to STEM Self-Concept." *Psychology of Women Quarterly* 40: 65-79.

Schuhmann, Peter W., KimMarie McGoldrick, and Robert T. Burrus. 2005. "Student Quantitative Literacy: Importance, Measurement, and Correlation with Economic Literacy." *The American Economist* 49: 49-65.

Spencer, Steven J., Claude M. Steele, and Diane M. Quinn. 1999. "Stereotype Threat and Women's Math Performance." *Journal of Experimental Social Psychology* 35: 4-28.

Steele, Jennifer R., and Nalini Ambady. 2006. "'Math Is Hard!' The Effect of Gender Priming on Women's Attitudes." *Journal of Experimental Social Psychology* 42: 428-436.

Stoet, Gijsbert, and David C. Geary. 2018. "The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education." *Psychological Science* 29: 581-593.

Stone, Jeff, C. Keith Harrison, and JaVonte Mottley. 2012. "'Don't Call Me a Student-Athlete': The Effect of Identity Priming on Stereotype Threat for Academically Engaged African American College Athletes." *Basic and Applied Social Psychology* 34: 99-106.

Thompson, Ted, and Sarah Musket. 2005. "Does Priming for Mastery Goals Improve the Performance of Students with an Entity View of Ability?" *British Journal of Educational Psychology* 75: 391-409.

Tulving, Endel, and Daniel. L. Schacter. 1990. "Priming and Human Memory Systems." *Science* 247: 301-306.

Vazquez, Jose J., and Eric P. Chiang. 2015. "Flipping Out! A Case Study on How to Flip the Principles of Economics Classroom." *International Advances in Economic Research* 21: 379-390.

Veenman, Marcel V. J., Lieneke Kerseboom, and Cornelie Imthorn. 2000. "Test Anxiety and Metacognitive Skillfulness: Availability Versus Production Deficiencies." *Anxiety, Stress, and Coping* 13: 391-412.

Voyer, Daniel, and Susan D. Voyer. 2014. "Gender Differences in Scholastic Achievement: A Meta-Analysis." *Psychological Bulletin* 140: 1174-1204.

Weingarten, Evan, Qijia Chen, Maxwell McAdams, Jessica Yi, Justin Hepler, and Dolores Albarracín. 2016. "From Primed Concepts to Action: A Meta-Analysis of the Behavioral Effects of Incidentally Presented Words." *Psychological Bulletin* 142: 472-497.

Williams, Amanda S. 2010. "Statistics Anxiety and Instructor Immediacy." *Journal of Statistics Education* 18: 1-18.

Wood, Beverly L., Megan Mocko, Michelle Everson, Nicholas J. Horton, and Paul Velleman. 2018. "Updated Guidelines, Updated Curriculum: The GAISE College Report and Introductory Statistics for the Modern Student." *CHANCE* 31: 53-59.

Zanakis, Stelios H., and Enzo R. Valenzi. 1997. "Student Anxiety and Attitudes in Business Statistics." *Journal of Education for Business* 73: 10-16.

Dave Ramsey's Personal Finance: A Primer and Critique

***William C. Wood and M. Scott Niederjohn*[1]**

ABSTRACT

Author and radio host Dave Ramsey advises millions of people on personal finance. Starting from the assumption that people are not good at maximizing their own utility, he recommends a rules-based approach that often directly contradicts conventional instruction in personal finance. This article shows how Ramsey's recommendations follow from his behavioral assumptions and highlights some of his specific recommendations about financial products and services that are at odds with traditional personal finance. After making these contrasts, the article outlines implications for educators who, whether they agree with Ramsey's approach or not, should understand the strengths and limitations of his approach.

Introduction

Personal finance instructors may be asked what they think of Dave Ramsey, a popular talk radio show host and author with a distinctive approach to personal finance. While many instructors have probably heard of Ramsey and might have a general sense that his approach places a heavy emphasis on debt avoidance, they are likely unfamiliar with all of Ramsey's approach. Hence, this paper outlines the key themes of Ramsey's approach and relates them to more conventional economic and personal finance education.

Before turning to fully explaining and critiquing Ramsey's approach, we begin with some information on Ramsey's background and enterprise. Among the many popular offerings in personal finance (Faulkner 2017), Dave Ramsey stands out. His syndicated radio show is the highest-rated personal finance program nationally, with ratings in recent years surpassed only by two political talk radio shows (Salmon and Poppick 2013). Ramsey runs a multiproduct firm, Ramsey Solutions, Inc., but the products are all built on core teachings in personal finance from his books (Ramsey 2011 and 2013, Ramsey and Cruze 2014). Ramsey's signature "Financial Peace University" is a nine-week class mediated by local proctors, often meeting at churches. Ramsey (2011) is the official handbook for "Financial Peace University" and contains essentially all the course's analytical content and advice.

For a widely followed financial guru, Dave Ramsey had a distinctly volatile start. After becoming a millionaire in his 20s, Ramsey was dragged down by leveraged holdings in a real estate crash when he was 26. He lost everything and filed for bankruptcy, an event that affected him deeply. By his own account, he was "totally broke and completely broken" (Ramsey 2019a). The experience also left him with a strong aversion to debt and a tendency to counsel others not to file an avoidable bankruptcy (Ramsey 2019a).

We see multiple points of interest in Ramsey's published and broadcast work. First, his rule-based approach challenges economic optimizing models, both as descriptions of what people do and of what they should do. Second, his approach of telling people what to do challenges the pedagogy of traditional personal finance, which focuses instead on providing information and encouraging learners to find their own solutions. Third, Ramsey (2013, 20) claims his approach dominates optimizing models. In this paper, we relate Ramsey's work to the models he challenges and provide analysis of interest to educators.

---

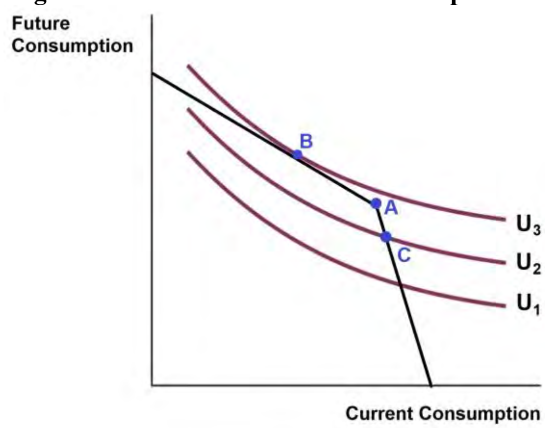
[1] Wood: James Madison University, Harrisonburg VA, 540-568-3243, woodwc@jmu.edu. Niederjohn: Concordia University, Mequon WI, 414-617-3813, scott.niederjohn@cuw.edu. The authors thank Mark Schug and Tawni Hunt Ferrarini for helpful comments.

Ramsey's Fundamental Assumption

The tradition of economics is to assume that people are good at maximizing their own utility, and to treat personal finance as information and techniques that improve the maximization process. Ramsey's approach starts with the opposite assumption: that people are not good at maximizing their own utility. Ramsey's consumers stumble from mistake to mistake, with overconsumption and excessive debt chief among them.

To counter these mistakes, Ramsey recommends a variety of tricks and methods for reducing current consumption and saving more for the future. The techniques include a written budget (Ramsey 2011, pp. 54-56), a monthly cash envelope system (Ramsey 2011, pp. 68-69), and a rigid 15 percent retirement savings goal (Ramsey 2011, pp. 222-223). Ramsey frequently writes and speaks on the assumption that his callers are not optimizing. Figure 1 illustrates.

Figure 1: Current and Future Consumption

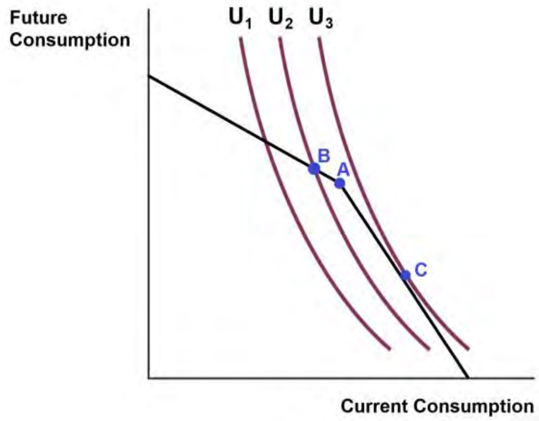


Households have preferences across current and future consumption, usually interpreted as "working life" vs. "retirement." Households start with current and future incomes that place them at point A on the kinked line. The kink comes from households facing different interest rates for borrowing and saving. Moving downward to the right from A, households get a relatively small amount of additional current consumption compared with the future consumption they will give up as they borrow at relatively unfavorable interest rates (say, 18 percent on credit card balances). Moving upward to the left from A, households face the relatively flat slope embodied in the interest rates they receive (for example, as holders of certificates of deposit.) Those who save nothing for the future remain at A and enjoy lifetime utility less than U_3. If rational optimizing households save for retirement, they achieve the highest possible lifetime utility, U_3, by moving to point B—giving up current consumption but getting greater future consumption they value more highly. Differing tastes for current and future consumption are represented by the slopes of the utility curves, with flatter curves reflecting greater patience and a willingness to defer consumption now for greater consumption in the future.

Ramsey's callers, as non-optimizers, often appear not to recognize the tradeoffs. If Ramsey's callers were thinking in terms of Figure 1, they would know that they were at a point like point C. They may not even have enough information to realize they are at a point like C. Instead of saving for the future, they have done the opposite, using credit to expand current consumption at the expense of the future. Callers could increase their utility level by getting out of debt and saving. Thus, Ramsey's techniques for getting them to cut current consumption move them toward a point they would prefer, if only they could get there.

To be sure, it is possible that some individuals do not value retirement or perhaps do not believe that they will be alive to enjoy it. A high degree of current consumption is rational for these individuals, as illustrated in Figure 2. Their steep utility curves reflect a strong desire to consume now rather than in the future. For such consumers, moving to point B by saving more for retirement would result in a lower utility level (U_2 instead of U_3). Those seeking Ramsey's advice, however, seem not to be consuming rationally, but instead to have wandered into a consumption pattern they would prefer to avoid. More generally, Americans' biggest financial regret is consistently reported as not saving enough for retirement or not starting to save for retirement soon enough (Smith 2019).

Figure 2: Current and Future Consumption with Present-Oriented Preferences



Interestingly, Ramsey's rule to save 15 percent of gross income for retirement is not just a floor but also a ceiling, derived from rough experience-based optimization (Ramsey 2011, p. 223):

> *Why not more? You need some of your income left to do the next two steps, college savings and paying off your home early. Why not less? Some people want to invest less or none so they can get a child through school or pay off the home superfast. I don't recommend that because those kids' college degrees won't feed you at retirement.*

However, Ramsey rarely needs to counsel against saving too much for retirement. This is not surprising, given both the audience for his show (unlikely to include many people who are confident in their financial future) and the imperatives of the talk radio format (with its need to get dramatic stories past the call screener and on the air).

Under Ramsey's fundamental assumptions, people are not just non-optimizing, but all too often "stupid." Ramsey (2019b) once told a caller, "When I do something stupid that costs me money, I call it stupid tax." Examples include cosigning a loan for a wasteful relative (Ramsey 2019b) and making a loan to an ex-girlfriend (2019c). The stupid tax involves spending in ways that do not increase utility and, in these examples, make relationships worse.

Consumers paying a stupid tax have no place in traditional optimizing economic theory. We do find them, however, in the increasingly prominent (Costa et al. 2018) behavioral finance and behavioral economics schools (Thaler and Sunstein 2008), doing such things as:

- Exhibiting a "status quo bias" (Samuelson and Zeckhauser 1988) that freezes them in place, preventing them from moving. Someone who says "I will always have a car payment" and therefore rules out working toward a car payoff is exhibiting a status quo bias.
- Making bad decisions on probabilistic matters (Tversky and Kahneman 1974), such as buying overpriced extended warranties. They fail to correctly weigh the low probability and limited payoff of an overpriced extended warranty compared with the up-front and certain large premium.
- Knowing credit card debt is bad, even while assuming that everyone has excessive credit card debt, falling prey to a "herd mentality." They feel secure following what they think the herd is doing, even if they are incorrect about the herd.

Ramsey does not think the caller facing a car repossession needs better optimization instructions, but instead the motivation to stop doing stupid things and then to start saving. Given Ramsey's total rejection of optimizing models, we wondered whether his work is unfamiliar to academic economists. A full-date, full-text search of the American Economic Association's ECONLIT database turned up only five total references to Ramsey. Among those, only two formally cited Ramsey–and even then, the citation was to his website and not any published work. Thus, there are millions of people taking financial advice that is virtually unknown within academic economics. We note this not to disparage the advice, but to emphasize the different knowledge sets of the two groups.

Ramsey's fundamental assumption of non-utility-maximization, influenced by his own bankruptcy, leads to an aversion to debt in all its forms, on both the borrowing side and the lending side. This aversion includes recommending an asset allocation of 100 percent stock and 0 percent bonds for investors, regardless of age, in order to avoid participation in debt. Ramsey frequently tells his audience, "The borrower is a slave to the lender," citing Proverbs 22:7 (Ramsey 2011, p. 94). Ramsey also mocks his former finance professors, optimizers who taught that debt was a tool and "were all broke" (Ramsey 2011, p. 93). Ramsey claims as his most important qualifying credential that "I have done stupid things with zeroes on the end," adding "I have a Ph.D. in D-U-M-B." (Ramsey 2013, p. xvii).

Ramsey celebrates the end of indebtedness with those following his plan using a radio programming feature called the "Debt Free Scream." The individual or household newly out of debt is led through a serious of questions (Ramsey 2019d) about how much debt was paid off and how long it took. Individual motivation tips and stories are also included. When the questioning is done, there is a countdown, "3-2-1, we're debt free!" followed by cheering and studio sound effects. To qualify for a debt free scream, the individual or household must be truly debt free, not even holding a credit card. To Ramsey, the convenience and incentives of a credit card, even one paid off monthly, do not justify the use of a debt instrument.

To fully appreciate Ramsey's aversion to debt, consider his advice on credit scores—not how to raise credit scores, but how to have no credit score at all. Ramsey explains that someone with no active credit accounts and no activity in the past six months will not have a FICO (Fair Isaac) credit score at all (Ramsey 2011, p. 108). This stands in sharp contrast to the near-universal personal finance advice to maintain a good credit score (for example, Billingsley et al. 2017 and Howard 2019). Ramsey acknowledges the inconvenience of not having credit or a credit score, but he considers being debt-free to be a sufficient offsetting reward.

Ramsey's model makes sense for consumers who do not trust their ability to manage debt or who have a moral objection to debt. It does not appear to be optimizing for other consumers. The rational consumer who can borrow at 5 percent and make a 6 percent return on assets can, over time, multiply personal wealth. Ramsey's approach gives up this potential return. This issue comes up frequently when callers balk at Ramsey's advice to sell off stocks to extinguish a home mortgage. Ramsey asks them to consider their situation if they had a paid-off home and no debt. Would they go into mortgage debt to invest in stocks? The usual answer is "of course not." Ramsey's reversal of the question relies on risk aversion and also invokes psychic benefits: "When you pay off your house and burn the mortgage, take off your shoes and walk through the backyard. The grass feels different under your feet" (Ramsey 2019e).

Evaluation of Financial Products

Starting from non-utility-maximization, Ramsey's rule-based approach calls for avoiding certain financial products and embracing others–without further analysis. Because of the difference in approaches, his advice tends to vary in tone and substance from traditional personal finance advice. Here are some financial products that Ramsey dismisses and the often-contrasting perspectives from conventional personal finance textbooks:

- Automobile leases: Ramsey says that leasing a car "is the worst possible way to acquire a vehicle" (Ramsey 2013, p. 32) and repeatedly refers to a lease as a "fleece." Typical personal financial planning textbooks (Billingsley et al. 2017, p. 164; Kapoor et al. 2019, p. 305) take a different approach. Keown (2019, p. 250) writes, "And don't forget about the option to lease," while Ramsey's message is to forget about the option to lease. Billingsley et al. (2017) concede the point that a car lease generally does increase the total cost to buyers compared to buying a car with a loan. However, the decision can be explained by rational factors that affect buyers including rising car prices, the non-deductibility of consumer loan interest, lower monthly payments, driving a more expensive car for the same monthly payments, and minimizing the down payment to preserve or invest cash. Kapoor et al. (2019, p. 305) treat the buy-or-lease decision very much as an open question for an individual consumer to determine. Other easily accessible personal finance material, including free source textbooks on the topic, treat the topic in a similar manner, in contrast to Ramsey's open-and-shut approach.
- Cash value life insurance products, such as whole life and universal life: "I want to be crystal clear here: Cash value life insurance is total garbage" (Ramsey 2011, p. 165). The advice from personal finance texts is to carefully evaluate life insurance options. Conventional personal finance textbooks (Kapoor et al. 2019, p. 434; Billingsley et al. 2017, p. 309; Keown 2019, p. 303) tend to present the advantages and disadvantages of various life insurance products. These texts point out a number of

benefits of whole life policies, including permanent coverage, savings vehicles, and some tax advantages. Of course, such books also point out the higher sales commissions and marketing fees as well as such policies having lower yields than traditional investments. However, they still leave the consumer's decision about cash value policies, open, in contrast to Ramsey's adamant opposition to such products.

- Fixed annuities: "I certainly don't use fixed annuities for anything" (Ramsey 2019f). Once again, Ramsey's advice is in some contrast to the type of advice found in typical personal finance textbooks. Traditionally, fixed annuities can be part of a recommended retirement income strategy (Kapoor et al. 2019, p 645). Billingsley et al. (2017, p. 568) point out that while fixed-rate annuities are conservative, very low risk, and essentially only promise a return of the principal plus a small rate of interest, they don't fluctuate in value as interest rates rise and fall, and so the principal is secure.
- Non-conventional mortgages, such as reverse and adjustable rate mortgages: "Focus only on conventional fixed-rate options, and never—*never*—get a mortgage term longer than 15 years" (Ramsey 2011, p. 300, emphasis in original). Thirty-year mortgages are commonly recommended as the safest and conventional choice; however, personal finance textbooks once again provide more context. Kapoor et al. (2019, p. 311) point out for those interested in debt reduction: "You might pay an additional amount each month (toward the loan principal) so your equity in the home will increase faster. Or you might choose a 15-year mortgage rather than one for 30 years." Billingsley et al. (2017, p. 202) add: "Because the borrower assumes most of the interest rate risk in these mortgages, the *initial rate of interest* on an adjustable-rate mortgage is normally well below—typically by 2 to 3 percentage points—the rate of a standard 30-year fixed-rate loan. Of course, whether the borrower actually ends up paying less interest depends on the behavior of market interest rates during the term of the loan" (emphasis in original).

To understand the financial products that are recommended by Ramsey, consider his optimal long-term investment portfolio, which is unconventional in that it is overbalanced toward growth. Specifically, Ramsey (2011, p. 209) recommends 25 percent allocations for each of four categories of stock mutual funds: (1) growth, (2) aggressive growth, (3) growth and income and (4) international. Consistent with his aversion to debt, he recommends zero percent in bonds. Notice that Ramsey's advice fails if the past high performance of such portfolios does not continue into the future. Downplaying "Past performance does not guarantee future results" and similar formulations (Newall and Parker 2019), Ramsey places a high weight on mutual funds' track records: "Always look at the track record of mutual funds before you buy one" (Ramsey 2011, p. 211).

Because consumers and investors in Ramsey's world are not rational, they need an investment advisor "with the heart of a teacher" (Ramsey 2011, p. 199) to patiently show them how their money should be managed and allocated. Ramsey has a network of local providers ("SmartVestor Pros") who subscribe to his teachings and pay him a fee. These providers in turn work on commission for clients. Ramsey evaluates providers using customer feedback but, reflecting his dismissive view of credentials, does not impose credentialing requirements for participating investment advisors. Ramsey's revenue from this part of his business model is substantial, with one back-of-the envelope calculation estimating $900,000 per month (Kelly 2017). With this and multiple additional sources of income pushing the total beyond $1 million per month (Harrison 2015), it is not surprising that Ramsey's net wealth is estimated at $55 million (The Richest 2019).

Although investment expenses associated with his approach are greater than for strategies such as buying and holding no-load mutual funds, Ramsey (2016) explains that advisors are necessary to maintain investor confidence (and to keep individuals from getting out of stocks when the market is down). The rational optimizer of economic theory does not need an advisor "with the heart of a teacher," referred by someone with a conflict of interest in recommending commission-based advising. However, someone who has repeatedly paid the "stupid tax" may do better paying high commissions, if only by avoiding stupid investment moves.

Ramsey's investment advice is at sharp odds with conventional personal finance:

- Index funds, strongly recommended in many quarters, do not play an important part in Ramsey's recommended strategy because they are not prominently featured by commission-based investment advisers. In a recent (February 2019) Initiatives of Global Markets (IGM) Survey by the University of Chicago's Booth School of Business, the following statement was evaluated: "In general, absent any inside information, an equity investor can expect to do better by holding a well-diversified, low-fee, passive index fund than by holding a few stocks." On this item, 57% responded "Strongly Agree," 36% responded "Agree," and the remaining 7% did not respond. It is clear that mainstream economists stand in sharp opposition to Ramsey's teachings on this topic.

- Adjusting an investor's asset allocation for age and risk tolerance, a commonly recommended risk reduction strategy (Kapoor et al. 2019, p 470; Billingsley et al. 2017, p. 435; Keown, 2019 p. 381), is not endorsed by Ramsey because he opposes bonds on moral grounds. A survey of dozens of financial literacy textbooks and websites was unable to find a single source in agreement with this philosophy. Every conventional financial planning resource suggests that bonds are a key component of a balanced investment portfolio.
- While traditional personal finance texts point out the advantages and disadvantages of commission-based advisors (Billingsley et al. 2017, p. 28), following the Ramsey approach means engaging a "SmartVestor Pro," typically compensated by commission. Personal finance texts tend to emphasize choosing Certified Financial Planners (CFP) or Chartered Financial Consultants (ChFC) while acknowledging CPAs, attorneys, investment managers, and other professionals may provide sound financial advice (Billingsley et al. 2017, p. 28). That said, such books and mainstream textbooks also focus on how an advisor is compensated (Kapoor et al. 2019, p 117). In most cases, the advice is to assure that they benefit when you benefit (growing your investment portfolio or providing an unbiased financial plan) rather than through commissions from buying investments or through trades.

The Baby Steps

How should people order their financial lives? The rational optimizers of economic theory continuously solve their complex life-cycle problem. Ramsey's non-optimizing consumers, in contrast, must be guided by a rigid sequence of actions. Ramsey calls his recommended sequence of actions "Baby Steps." A clear statement of the Baby Steps is found in Ramsey (2011, pp. 7-8):

- Baby Step 1: Put $1,000 in a beginner emergency fund.
- Baby Step 2: Pay off debt using the debt snowball (paying smallest principal amounts first).
- Baby Step 3: Put three to six months of expenses into savings as a full emergency fund.
- Baby Step 4: Invest 15% of household income into Roth IRAs and pretax retirement plans.
- Baby Step 5: Begin college-funding for your kids.
- Baby Step 6: Pay off your home early.
- Baby Step 7: Build wealth and give.

Ramsey rarely departs from the Baby Steps in advice to radio callers. A frequent early response to a caller's situation is "Where are you in the Baby Steps?" If, for example, a caller asks about transferring a credit card balance to a lower interest rate, Ramsey will insist that the $1,000 beginner emergency fund be complete before any debt management activity begins. Ramsey clearly believes in the rules-based approach derived from his experience in counseling individuals (Ramsey 2011, p. 6). Notice that Ramsey's approach agrees with conventional personal finance in some aspects, such as in recommending an emergency fund and systematic retirement saving. The striking contrast lies in Ramsey's aversion to debt and his insistence on adherence to rules. In the Baby Steps, only a $1,000 emergency fund is saved before an all-out assault on debt begins. Further, recall that exactly 15 percent of income is to be saved for retirement–not more and not less.

Baby Step 2's "debt snowball" provides a useful contrast of approaches. Mathematically, consumers will pay the least interest and get out of debt soonest if they tackle high-interest-rate debt first. This is the traditional advice of personal finance texts (Billingsley et al. 2017, p. 250). Ramsey's approach calls on consumers to forget the interest rate and focus on the smallest balance outstanding. Thus, a consumer following the debt snowball would pay off a $50 balance on an old store credit card before applying funds to the $2000 balance of a high-interest-rate card. The psychological lift of totally extinguishing one account, in Ramsey's view, far outweighs any disadvantages–and this advantage builds as additional small credit accounts are eliminated. Marketing research suggests (Gal and McShane 2012) that in practice, Ramsey's approach may work better despite being non-optimizing in a sense. Because home mortgages will typically be the largest debts consumers have as well as the lowest interest rate, in this case the "debt snowball" approach is likely consistent with traditional personal finance advice to pay off the lowest interest loans last.

Baby Step 7 shows a final contrast with traditional personal finance, which is studiedly neutral on the moral content of individual decisions. Ramsey's approach is specifically Biblical, relying on such texts as Psalm 24:1 ("The earth is the Lord's, and everything in it, the world, and all who live in it") and Luke 11:42 (the Christian discipline of giving away a tenth of income). It is not clear whether Ramsey's approach is consistent with optimizing models of maximizing lifetime wealth. Ramsey (2011, p. 308) sometimes seems to argue that giving wealth away will, through behavioral change, increase wealth in the long run. Sacrificial giving (Ramsey 2019g), on the other hand, would imply lower wealth but higher lifetime utility. This higher utility would come from a "warm glow" effect, inherent in the act of giving (Null 2011). Some of the utility would even come from knowing that wealth was being held by ethical people: "If you are a good person, it is your spiritual duty to possess riches for the good of mankind" (Ramsey 2013, p. 198).

Implications for Educators

Students often have been exposed to Dave Ramsey's principles before they enter our classrooms. It is important to be able to put these principles into context. The instructor who fully buys into Ramsey's approach might be tempted to discard conventional texts and just teach the Baby Steps.

However, doing that would be a disservice to students. Ramsey's approach is focused on the individual following his advice more than on learning the analytical principles of personal finance. The apprendix illustrates the issues, using the Council for Economic Education's National Standards for Financial Literacy (2013). The left-hand column contains the six major standards and the right-hand column summarizes Ramsey's treatment of that area. In some areas, such as budgeting, Ramsey's approach fundamentally agrees with the standards. In other areas, such as investment, the differences are large. In each case the national standards are detached and analytical, helping students understand how people behave, while Ramsey's approach is to command people. In other words, the standards focus on a social science approach to how people behave, but Ramsey tells people what to do. From this fundamental difference, at least three issues arise.

First, the Baby Steps, as fully implemented with associated advice from Ramsey, only work as long as the underlying rules continue to be accurate. Since the underlying rules are based on current price structures, future changes in financial services could dramatically degrade their quality. In the future, Ramsey's approach could involve giving up good deals if competition pushed pricing anomalies into line–for example, competition that made car leases or whole life insurance more favorable. The Ramsey approach avoids the informational problem (that gathering information on pricing anomalies is costly) by simply counseling adherence to the rules. Thus, Ramsey counts on a high degree of pricing inertia in financial products.

Second, Ramsey's aversion to debt is sufficiently out of the mainstream that a personal finance class would not be complete if it were the only viewpoint of debt taught. For many consumers, the temptation to overuse debt is so strong that making the case for extreme debt aversion is a useful exercise. But for many others, including highly successful builders of net wealth, the time and transaction cost of totally avoiding credit is excessive.

Third, Ramsey's evidence of success has important selection and causation problems. Ramsey (2011, p. 54) insists that a written budget goes along with success—but is that effect caused by the budget, or do people who successfully budget also have other characteristics that are associated with success? Investors who have financial advisors have larger portfolios (Ramsey 2016), but what causes what? Only studies that carefully control for endogeneity (Marsden et al. 2011) can provide useful knowledge on the value of seeking financial advice. Finally, consider how Ramsey (2019h) conducted a survey of millionaires among his audience to share their strategies. Although this approach has a long tradition and is useful for gathering success stories (Stanley 1996), self-selection by self-identified millionaires leaves it powerless to determine causality.

It is also fair to ask what motivation lies behind Ramsey's enterprises. A natural first hypothesis from economics would be that he maximizes his own net wealth–and rather successfully, at that (The Richest 2019). Ramsey himself says that his company defines success "by the number of lives changed" (Ramsey 2019i). These two goals are not necessarily conflicting.

Conclusion

It is hard to argue with Dave Ramsey's commercial success, even as reservations about some of his financial teachings apply. His Baby Steps make a great deal of sense when applied to those with significant debt who do not know where to start or what to do next. At the same time, his preferences for which financial products to emphasize and to avoid are far from the mainstream. Whether educators agree or disagree with his approach, they need to know what he is saying.

References

Billingsley, Randy S., Lawrence J. Gitman, and Michael D. Joehnk. 2017. *Personal Financial Planning* (14th ed.). Boston, MA: Cengage Learning.

Booth School of Business IGM Forum. 2019. *http://www.igmchicago.org/.*

Costa, Daniel F., Francisval de Melo Carvalho, and Bruno César de Melo Moreira. 2018. "Behavioral Economics and Behavioral Finance: A Bibliometric Analysis of the Scientific Fields." *Journal of Economic Surveys* 33: 3-24.

Council for Economic Education. 2013. "National Standards for Financial Literacy." *https://www.councilforeconed.org/resource/national-standards-for-financial-literacy/*

Faulkner, Ashley E. 2017. "Financial Literacy Education in the United States: Exploring Popular Personal Finance Literature." *Journal of Librarianship and Information Science* 49: 287-298.

Gal, David, and Blakeley B. McShane. 2012. "Can Small Victories Help Win the War? Evidence from Consumer Debt Management." *Journal of Marketing Research* 49: 487-501.

Harrison, Scott. 2015. "Dave Ramsey's Company Reveals How Much Money It Brings In." *https://www.bizjournals.com/nashville/blog/money/2015/09/dave-ramseys-company-reveals-how-much-money-it.html.*

Howard, Clark. 2019. "Credit Score Guide." *https://clark.com/story/credit-score-guide-what-to-know-how-to-improve/.*

Kapoor, Jack, Les Dlabay, and Robert J. Hughes. 2019. *Personal Finance* (13th ed.). New York: McGraw-Hill Education.

Kelly, Bruce. 2017. "Advisers Profit Through Association with Dave Ramsey–but Not by as Much as in the Past." *https://www.investmentnews.com/article/20170407/FREE/170409952/advisers-profit-through-association-with-dave-ramsey-x2014-but-not.*

Keown, Arthur J. 2019. *Personal Finance: Turning Money into Wealth* (8th ed.). New York: Pearson North America.

Marsden, Mitchell, Cathleen D. Zick, and Robert N. Mayer. 2011. "The Value of Seeking Financial Advice." *Journal of Family and Economic Issues* 32: 625-643.

Newall, Philip W.S., and Katie N. Parker. 2019 "Improved Mutual Fund Investment Choice Architecture." *Journal of Behavioral Finance* 20: 96-106.

Null, Clair. 2011. "Warm Glow, Information, and Inefficient Charitable Giving." *Journal of Public Economics* 95: 455-465.

Ramsey, Dave. 2011. *Dave Ramsey's Complete Guide to Money: The Handbook of Financial Peace University*. Brentwood, TN: Lampo Press.

Ramsey, Dave. 2013. *The Total Money Makeover: A Proven Plan for Financial Fitness*. Nashville, TN: Thomas Nelson.

Ramsey, Dave. 2016. *Stress and Anxiety Surrounding Retirement: A Ramsey Solutions Research Study*. *https://www.daveramsey.com/research/stress-anxiety.*

Ramsey, Dave. 2019a. "How to Avoid Bankruptcy." *https://www.daveramsey.com/blog/the-basics-of-bankruptcy*.

Ramsey, Dave. 2019b. "Big Stupid Tax – Ask Dave." *https://www.daveramsey.com/askdave/debt/big-stupid-tax*.

Ramsey, Dave. 2019c. "A Grand in Stupid Tax–Ask Dave." *https://www.daveramsey.com/askdave/relationships/a-grand-in-stupid-tax.*

Ramsey, Dave. 2019d. "Schedule Your Debt-Free Scream." *https://www.daveramsey.com/show/debt-free-scream.*

Ramsey, Dave. 2019e. "The Peace of a Paid Off Mortgage." *https://www.daveramsey.com/askdave/home-buying-and-selling/the-peace-of-a-paid-off-mortgage*

Ramsey, Dave. 2019f. "Annuities for Long-Term Retirement?" *https://www.daveramsey.com/askdave/investing/annuities-for-long-term-retirement.*

Ramsey, Dave. 2019g. "Guidance for Giving from God." *https://www.daveramsey.com/askdave/stewardship/guidance-for-giving-from-god.*

Ramsey, Dave. 2019h. "How to Go from Zero to Millionaire in 25 years." *https://www.daveramsey.com/blog/how-to-go-from-zero-to-millionaire.*

Ramsey, Dave. 2019i. "How It Started." *https://www.daveramsey.com/careers/about-dave.*

Ramsey, Dave, and Rachel Cruze. 2014. *Smart Money Smart Kids*. Brentwood, TN: Lampo Press.

Salmon, Felix, and Susie Poppick. 2013. "Save Like Dave Ramsey... Just Don't Invest Like Him" *Money* 42: 72-78.

Samuelson, William, and Richard Zeckhauser. 1988. "Status Quo Bias in Decision Making." *Journal of Risk and Uncertainty* 1: 7-59.

The Richest. 2019. "Dave Ramsey Net Worth." *https://www.therichest.com/celebnetworth/celeb/tv-personality/dave-ramsey-net-worth/.*

Smith, Kelly Ann. 2019. "Americans Say Their Biggest Financial Regret Is Not Saving for Retirement Sooner." *https://www.bankrate.com/banking/savings/financial-security-may-2019.*

Stanley, Thomas J. 1996. *The Millionaire Next Door: The Surprising Secrets of America's Wealthy*. New York: Taylor Trade Publishing.

Thaler, Richard H., and Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185: 1124-1131.

APPENDIX

National Standards vs. Dave Ramsey's Approach

CEE Financial Literacy Standards (2013)	Dave Ramsey's Approach (2011)
I. Earning Income. Income for most people is determined by the market value of their labor, paid as wages and salaries. People can increase their income and job opportunities by choosing to acquire more education, work experience, and job skills. The decision to undertake an activity that increases income or job opportunities is affected by the expected benefits and costs of such an activity. Income also is obtained from other sources such as interest, rents, capital gains, dividends, and profits.	"Work in your strengths" because you are unlikely to change fundamentally and you will learn and grow the most in areas where you are already strong (2011, ch. 11). Take part-time and side jobs as necessary to achieve short-term goals like getting out of debt (2011, 266-267). Work hard but "you are not your job," so play hard when you are not working (2011, 269-270).
II. Buying Goods and Services. People cannot buy or make all the goods and services they want; as a result, people choose to buy some goods and services and not buy others. People can improve their economic well-being by making informed spending decisions, which entails collecting information, planning, and budgeting.	Use a written budget to control your spending, assigning every dollar of expected income a destination in advance, using only cash with a budget envelope system if necessary (2011, 68-69).
III. Saving. Saving is the part of income that people choose to set aside for future uses. People save for different reasons during the course of their lives. People make different choices about how they save and how much they save. Time, interest rates, and inflation affect the value of savings.	Save a $1000 "baby emergency fund" until you are out of debt, then an emergency fund of three to six months' income, then 15 percent of gross income for retirement plus additional amounts for children's future education (2011, 7-8).
IV. Using Credit. Credit allows people to purchase goods and services that they can use today and pay for those goods and services in the future with interest. People choose among	Do not use credit. The only acceptable use of credit is to purchase a house after you are out of debt, limiting the mortgage payment to 25 percent of take-home income on a 15-year fixed-rate

different credit options that have different costs. Lenders approve or deny applications for loans based on an evaluation of the borrower's past credit history and expected ability to pay in the future. Higher-risk borrowers are charged higher interest rates; lower-risk borrowers are charged lower interest rates.	mortgage with 10 percent down (2011, 30) Do not seek a high credit score; instead seek to have no credit score at all (2011, 108)
V. Financial Investing. Financial investment is the purchase of financial assets to increase income or wealth in the future. Investors must choose among investments that have different risks and expected rates of return. Investments with higher expected rates of return tend to have greater risk. Diversification of investment among a number of choices can lower investment risk.	Allocate all of your invested funds to mutual fund (growth, aggressive growth, growth and income and international) and zero percent to bonds (2011, 207). Purchase these investments through a local professional "with the heart of a teacher" (2011, 199)
VI. Protecting and Insuring. People make choices to protect themselves from the financial risk of lost income, assets, health, or identity. They can choose to accept risk, reduce risk, or transfer the risk to others. Insurance allows people to transfer risk by paying a fee now to avoid the possibility of a larger loss later. The price of insurance is influenced by an individual's behavior.	Do not go uninsured, but instead immediately purchase homeowner's (or renter's), auto, health, disability, long-term care (for those over 60), identity theft and life insurance (2011, 149), preferring large deductibles (2011, 150); own ten times your income in term life insurance—and never cash value life insurance of any form (2011, 164-167).

A Student Fund in a One-Hour Lab

C. Alan Blaylock[1]

ABSTRACT

First-hand experiences considering relevant research are used to examine the benefits and difficulties in implementing a student-managed investment fund through a repeatable one-hour lab format at a small regional state university. Implementing a student-managed fund in a lab format with fewer hours than a regular course reduces demands on faculty teaching load as well as on available hours in a student's academic program. The multi-semester approach allows a stepwise progression of learning over multiple semesters, but its major difficulty is managing the mix of student skill levels.

Introduction

Since its inception, the Ross Whipple Student Managed Investment Fund at Henderson State University (HSU) was incorporated into the finance program in various ways. For a few years prior to spring 2015, it was implemented in a three-hour senior-level course that was offered in both the fall and spring semesters. Beginning spring semester 2015, fund management began to be conducted in a split-level, one-hour lab that could be taken for credit multiple times. This was a result of working within the constraints of the characteristics of a small student population, available program hours, the university's liberal arts mission, and faculty teaching loads. This paper is a case study examining the following characteristics in implementing a student managed investment fund (SMIF) in a finance program at a small public regional university: one credit hour, participants of differing experience levels, lab format, multiple semester exposure, and prerequisite requirements. HSU's experience is similar to that of Austin College, described in Bruce and Greene (2014), which dealt with the same constraints.

Sections that follow lay out the reasoning for using a repetitive one-hour lab open to various skill levels. The first section describes common issues when establishing a SMIF. The second section describes the circumstances at HSU, circumstances common among small regional universities. The third section explains the structure HSU has adopted. The fourth section discusses experiences using that structure, followed by a section describing program adjustments moving forward. The last section summarizes.

SMIFs

The benefits of student managed investment funds (SMIFs) are well accepted (Charlton et al. 2015; North and Stevens 2012; Gradisher et al. 2016; Boughton and Jackson 2018). A component of SMIFs that leads to such benefits is active or experiential learning. SMIFs exhibit many of NSSE's Engagement Indicators, most notably, Higher-Order Learning, Reflective Learning, and Collaborative Learning, all of which are positively related to first-year retention and six-year graduation rates (NSSE 2016). SMIFs are also a tool to engage alumni, investment professionals, and the community (Gradisher et al. 2016; Boughton and Jackson 2018).

Research has viewed SMIFs through a variety of dimensions such as funding sources, academic structure, decision processes, continuity, costs, allowable investments, and participant selection (Lawrence 1994; Johnson et al. 1996; Lawrence 2008; Bruce and Greene 2014; Gradisher et al. 2016; Boughton and Jackson 2018).

These characteristics, particularly participant selection, time, and structure, in many instances are shaped by the unique situations and characteristics of the fund's host university, and, in turn, are major factors in how the fund is implemented in the curriculum. For instance, characteristics of regional schools

[1] Professor of Finance, Henderson State University, Arkadelphia, AR. Email: blayloc@hsu.edu.

such as limited funding, small student populations drawn from a geographically limited student base, and a limited number of faculty may necessitate offering a student fund open to all who are able to enroll in a single required course in the finance curriculum (Macy 2010). Thus, student selectivity is not an option. Such was the case with the Austin College case described in Bruce and Greene (2014). These constraints, together with students' work and family obligations, also limit the activities of the fund (e.g., trading only once per year).

A student fund can be offered as an extra-curricular activity or as a for-credit course. If it is offered as a for-credit course, the number of hours dedicated to fund management needs to be determined. Such were concerns considered by Austin College (Bruce and Greene 2014). If the student fund is managed as an extra-curricular endeavor, students must weigh the immediate and pressing importance of dedicating time to the extra-curricular activity at the expense of time dedicated to coursework applied toward earning a degree. Students would tend to place more importance on coursework, resulting in lessened benefits of the student fund. Bruce and Greene describe grades as a "form of currency with which to compensate students who participate in the management of the portfolio. It might seem controversial to think of course grades as currency. However, the point is that a grade provides an incentive mechanism that allows the interest of the student to be aligned with that of the fund's client" (Bruce and Greene 2014, p. 54). Bruce and Greene go on to advocate that the compensation of knowledge – the benefit of the practical experience – "mitigates much of the incentive alignment problem" (Bruce and Greene 2014, p. 55). However, it is reasonable to speculate that this may be true for students in a highly selective and competitive SMIF program, but not so true for students from a small population who participate in managing the student fund simply as a program requirement. In addition, a formal course may also incentivize faculty advisors to devote more time and effort to the SMIF program.

Given the non-selective and non-competitive nature of HSU's SMIF program, managing the fund through not only a for-credit but also a required course was deemed a necessity. Although accountability may be elicited by managing real money, the fund still belongs to someone besides the student. Receiving a grade in a course required for graduation would seem to incentivize students to a higher level of accountability and effort. The formal course also allows more structure and guidance than in a student club.

Understandably, given the benefits of active learning, a preference for both students and instructors may be to increase the number of program hours allocated to fund management - perhaps a three-hour course every semester. There are several problems with such an approach. First, diminishing marginal returns are sure to set in. While a student managed fund enhances finance and investments education, it cannot serve as a complete investment education. There is just so much more that needs to be learned than what can be learned in managing a student fund. The more time dedicated to the fund is less time to cover other needed topics. As hours allocated to a SMIF increase, the marginal benefit of other activities would begin to outweigh the marginal benefit of fund management.

Second, the benefit of a required for-credit course has the cost of consuming available hours in a student's degree program and increasing faculty workloads. There are a limited number of credit hours that can be allocated to fund management in a finance program. Even if a student enrolls in the fund as an elective, scheduling, graduation timing, and cost constraints would still limit the number of hours a student can allocate to the fund. Every hour allocated toward a SMIF is one less hour in another course. This applies to both students and teachers. This constraint is particularly acute for liberal arts institutions (Bruce and Greene 2014).

If fund management is limited to, say, three hours in a finance program, a question can be raised as to which is better, one three-hour course or three one-hour courses over multiple semesters? HSU tried the latter for 9 semesters. Reducing the credit hours per semester reduced faculty loads and requiring students to repeat the course for credit over multiple semesters not only maintained the amount of student credit hours, but also provided additional benefits for the students and the program. This spreading of hours over multiple semesters is discussed in more depth in another section. First, an overview of HSU's situation is presented.

The Conditions at the University

HSU is a regional university with many of the attributes described by Macy (2010). Total enrollment has been hovering around 3,000 students, with undergraduate business and finance students numbering 400 and 40, respectively. Although small, the School of Business is AACSB accredited. The "finance major" is actually an emphasis area under the Business Administration major for the BBA. About ten students in the finance emphasis area graduate per year. Many HSU students are first-generation college students, and although very bright and accomplished students go through the finance program, enough students who may not be well-prepared for college level instruction warrants consideration when developing the finance curriculum.

The constraints of available program hours and faculty resources are acute at HSU. HSU is the state's only public liberal arts university and a member of the Council on Public Liberal Arts Colleges (COPLAC). Thirty-two hours of the 120-hour baccalaureate degree requirements remain after accounting for the liberal arts and business core courses. The finance emphasis area and free electives take up the remaining thirty-two hours. Emphasis areas for the Business Administration major usually amount to eighteen hours, leaving fourteen hours for free electives. While this may be enough hours to construct an adequate finance program, having only two finance faculty members limits the number of finance courses that may be offered. Of course, finance faculty teach extra sections of the basic finance course for the non-finance students taking the business core and for a popular and ever-growing MBA program. In short, a small student population, less prepared students, limited available student hours for the finance program, and limited faculty teaching hours are key considerations when implementing a SMIF in the finance program at HSU.

While these constraints exist, and although growth in the student body and faculty is desired, HSU embraces its liberal arts heritage and student characteristics. This is not to say HSU advocates a lack of college preparedness, but rather, the philosophy of HSU, and especially the finance program, is to take even less-prepared students and bring them to their full potential to be competitive with any student from a flagship state university.

Initial Lab Implementation

The Ross Whipple Student Managed Investment Fund was established in 2001 at HSU through a donation of $250,000 made by Ross Whipple. Monies are held by the university's foundation but are under the exclusive control of the School of Business to provide students practical experience in investment management. On December 31, 2018, the fund's approximate value was $523,000. It has no restrictions on investable assets, and no funding requirements. In fall 2017, authorization was acquired from the donor and university administration to withdraw monies to pay for resources directly related to managing the fund. The appropriation amount is to be no greater than 0.25% of the total value of the fund at the end of each calendar quarter, which is a nominal annual rate of 1% of total assets.

Fund management is conducted through a split-level lab framework. FIN 3241 Investments Lab is the lower-level lab, and FIN 4221 Portfolio Management Lab is the upper-level lab. Both labs are one credit hour and meet for two hours per week at the same time. Both are offered in the fall and spring semesters. FIN 3241 Investments Lab can be taken a maximum of four times for credit; FIN 4221 Portfolio Management Lab can be taken a maximum of three times for credit. Although a student could conceivably earn seven credit hours (seven semesters) participating in the two labs, only one credit for FIN 3241 Investments Lab and two credits for FIN 4221 Portfolio Management Lab is required in the finance program. It is hoped that a student would graduate before utilizing the seventh semester, but the option is there to accommodate student schedules. For instance, some could take four semesters of FIN 3241 Investments Lab and only two semesters of FIN 4221 Portfolio Management Lab, while others may only take one or two semesters of FIN 3241 Investments Lab but three semesters of FIN 4221 Portfolio Management Lab. The approach of using a repeatable one credit-hour course may be somewhat similar to the fund at University of Missouri, St. Louis as described in Lawrence (2008) and Austin College as described by Bruce and Greene (2014).

One-Hour Lab

Given the constraints on faculty workloads and the limited number of available hours in an academic program, the multi-semester approach is only viable if the course is one credit hour. However, a downside to limiting the fund to a one-hour course is the amount of time available for fund management. In addition, the one-hour course elicits less incentive for students to put forth an effort than a three-hour course. These problems may be alleviated, at least partially, by assigning some fund management activities to the two upper-level investment courses, FIN 4103 Investment Management and FIN 4213 Security Analysis, where the upper-level lab, FIN 4221 Portfolio Management Lab, is a co-requisite. This is the current practice used in the Ross Whipple Student Managed Investment Fund. Just a small weight in these courses would offer a greater incentive since the weight affects the grade for a three-hour course.

There is also the decision to conduct fund management through a regular class or classify the course as a lab. In a regular course, every credit hour would provide just fifty minutes of class time per week but entail outside-of-class study requirements. For a lab, class time can be extended, to say two to three hours, but outside-of-class study requirements are eliminated. For a course of average difficulty, a general rule of thumb is for students to allocate two hours of study outside of class per week for every one credit hour (Utah State University 2018). The standard Carnegie Unit for higher education equates to one class hour and two hours of homework (Silva et al. 2015). This would allow a total of about three hours (two hours and fifty minutes) per week available for fund management.

The concept of a lab, on the other hand, is that the class period is longer but with no expectation of work outside of the lab period. Therefore, a tradeoff exists: a lab format would only allow two to three hours per week for fund management, but it extends class time to work together with the faculty advisor and students; a one-hour course would allow outside work but limit the amount of time in class with others. HSU opted for a lab that meets two hours per week. A possible alternative includes having a lab format for beginning students to allow for more instructor class contact hours per week, and a one-hour course for upper-level students to allow for additional outside work.

Even a one-hour course or lab adds an additional hour to the faculty advisor's teaching load, but the one hour is minimal and deemed acceptable. However, the reduction in hours does not translate into a reduction of actual time and effort put into advising students managing the fund. The one-hour lab consumes just as much time as a regular three-hour course or more. Lawrence (2008) states that most professors who advise SMIFs as a formal class equate the workload to a 4.5-credit hour course, and some schools have difficulty replacing faculty advisors because of the increased workload. Because of the added complexity of managing students at various levels, advising the Ross Whipple Student Managed Investment Fund has at times consumed over half of the advisor's teaching efforts.

In short, a course offering more credit hours would be preferable, but considering available program hours and faculty teaching load constraints, the reduced hour, multi-semester approach was deemed more valuable.

Prerequisites

Requiring students to repeat the one credit hour lab over multiple semesters offers a mixed bag of obstacles and benefits. A major obstacle is the need for adequate prerequisites prior to taking the first SMIF lab. A one-hour lab provides little time for instruction, and requiring the SMIF lab be repeated three times means the first lab occurs in the student's junior year at the latest. If multiple semesters are encouraged, the starting semester could be even earlier. This also means the lab includes students of different experience levels. Students in their third (or later) semester managing the fund are in the same class as those in their first. In short, the combination of little time for instruction, starting early, and differing student experience levels necessitates a need for adequate prerequisites before taking the first semester of the course sequence. Austin College recognized the importance of employing prerequisites earlier so juniors could participate in its fund (Bruce and Greene 2014).

Usual courses that would suffice as a prerequisite for a SMIF usually would be junior or senior level courses that require their own prerequisites. This would push the starting point of the multiple semester sequence of SMIF courses too late in a student's program. A prerequisite was needed that could be taken in the freshman or sophomore year to provide adequate time to fit more semesters of the SMIF course in a four-year program. A simple solution could have possibly been to create a new course. However, creating an extra course to exist solely as a prerequisite to the student fund would have been deemed to use too many resources for such a limited benefit. It would be preferable to have a course that would already have been generally accepted in university curriculums or otherwise provided academic or professional benefits beyond simply serving as a prerequisite for the fund.

The answer came purely by coincidence when CFA® Institute unveiled its new Investment Foundations Program at the same time the one-hour SMIF course was being considered. The Investment Foundations Program, formerly the Claritas Program, is designed for those working in the investment industry who do not make investment decisions, such as those in human resources, accounting, or the legal department. The program culminates in an exam and subsequent Investment Foundations Certificate. The free curriculum for the program covers a broad overview of the global investment industry and serves as an excellent introduction to investments. It is also suitable to use in a freshman or sophomore course. As part of CFA Institute's Investment Foundations Academic Program, the School of Business created FIN 2233 Beginning Investments, a sophomore level course that covers the program's curriculum and prepares students to sit for the exam. FIN 2233 has no prerequisites. This became the perfect solution for a prerequisite to FIN 3241 Investments Lab, the first lab in the sequence of SMIF courses. It was a new course, but it not only served as a prerequisite to the labs, it also provided instruction leading to a certificate from a globally recognized professional organization.

Ideally, students would take FIN 2233 Beginning Investments in the spring semester of their freshman year and then participate in the lab every semester thereafter for a total of six semesters, depending on the student's graduation timetable and desired electives. Although the finance program requires one hour of the lower-level lab and two hours of the upper-level lab, students are encouraged to enroll in the lab every semester.

An alternate prerequisite to FIN 3241 Investments Lab is FIN 3043 Business Finance, the junior level principles course required of all business majors. Having this course as a prerequisite allows any business student who has taken it an opportunity to participate in the fund. Requiring prerequisites to the lab provides some foundational and essential knowledge so students can more quickly take on the tasks of fund management. This is especially important when those in the lower-level lab, FIN 3241 Investments Lab, engage with students with more fund management and coursework experience, such as those in the upper-level lab, FIN 4221 Portfolio Management Lab, or those who have multiple semesters of FIN 3241 Investments Lab.

FIN 4221 Portfolio Management Lab is a corequisite to two required upper-level investment courses, FIN 4103 Investment Management and FIN 4213 Security Analysis. FIN 3241 Investments Lab and FIN 2233 Beginning Investments are two of the prerequisites for both FIN 4103 Investment Management and FIN 4213 Security Analysis. In addition to each serving as a corequisite, each of these two upper-level courses also serve as an alternate prerequisite for the upper-level lab. So, a student could conceivably take just one of these courses with the lab as a corequisite and then continue to enroll in the lab for two more semesters without having to take the second upper-level investment course. This may occur when a student changes majors from finance but still wants to participate in the lab. Also, a student could take one of the upper-level investment courses as an elective or part of another program such as part of one of our certificate programs.

Multiple Semesters

The problematic early start to the sequence of SMIF courses also offers major advantages to the concept of requiring multiple semesters. Starting early means students are engaged in their major of choice earlier in their academic career. They get to see if finance is the right choice for them. For some, finance is not their right choice, and they change majors. Making such a change is better if done earlier than later. The low-level prerequisite and the minimal burden on a student's schedule could also be inviting to any non-finance or non-business students interested in investments and could also be used as a recruiting tool to the finance program. The early start coupled with a multiple semester sequence offers a pathway of continual engagement through active learning. If the active learning environment through a SMIF is so great, should not the SMIF experience be limited to those who have supposedly already learned the most? Austin College's SMIF as described in Bruce and Greene (2014) spans four semesters with the first two courses counting 0.25 credit hours and the last two semesters counting 0.75 credit hours.

The multi-semester approach is also a way to continually refresh and build on previously learned material. Deliberate practice, not rote repetition, not only increases performance in recalling new information but also the ability to take on more complex problems (Brabeck et al. 2018). Taking the initial prerequisite followed by multiple labs can be viewed as a type of "spaced practice." Spaced practice is having an initial study and subsequent practice spaced out over time and has been shown to result in durable long-term learning, not only in rote recall, but also in the application of knowledge (Kang 2016). Such spacing effects have been shown to last up to two years (Kerfoot 2009).

The multi-semester approach is intended to provide a stair-step progression of active learning so that each additional semester builds on the previous. This allows a slow growth process and is especially important for slower students. The repetitive nature of the sequence reinforces previously learned knowledge (students are prone to forget), adjusts and corrects any incorrect understanding of previously learned material, and garners familiarity with investment topics that leads to more confidence. The multi-semester approach is very similar to that prescribed in North and Stevens (2012), where students rotate and progress through various assignments in a SMIF program over a two-year period.

The early start and frequent exposure to the student fund also may provide better preparation for advanced student experiences. Students may be better prepared for advanced investment courses. The slow growth process can substitute for needed prerequisites that otherwise could not be required due to limited resources. Additionally, students may be able to secure earlier and better internships.

Experiences from Initial Implementation

The advisor creates the Investor Policy Statement (IPS), which is reviewed annually. As described in Lawrence (2008), investment performance for a student fund is secondary to its educational mission. This is true of the Ross Whipple Student Managed Investment Fund as well. The donor has been clear that student education is the goal. Bruce and Greene (2014) posit that focusing on the goal of investment performance can achieve the educational goal. However, care should be taken in how that can be interpreted. The performance objective may hinder the educational objective. Some investment strategies involving day trading, options, futures, short selling, etc. may have a goal to generate positive returns, but they may be ignored because of fears of not being able to accomplish that goal. With the Ross Whipple Student Managed Investment Fund, its risk tolerance is very high and would warrant an asset allocation in all equities as well as other risky assets and strategies. However, doing so would limit the opportunity for students to realize investing is not just beating the S&P 500, and that measuring performance may entail a benchmark of mixed indices. For this reason, the Ross Whipple Student Managed Investment Fund has an allocation to fixed-income with a corresponding benchmark using Barclay's U.S. Aggregate Bond Index. The current benchmark is 90% S&P 500 and 10% fixed-income. The current asset allocation also includes a 10% allocation to alternatives.

Some raise the issue of the implied short time horizons due to the students' short tenure of only one or a few semesters (Block and French 1991; Lawrence 2008). This really should not be a concern since a portfolio's time horizon is not defined by manager tenure. The problem of constant turnover (Johnson et al. 1996) and changing management style (Mallet et al. 2010) should not be an issue if an adequate IPS is in place. In addition, unlike sporadic and unexpected turnover of a real investment firm, student turnover can be planned and managed (Bruce and Greene 2014). In describing their own fund at the University of Richmond, Charlton et al. (2015) report that having existing student managers aid in the selection of new student managers creates a more long-term view for both the existing and new managers. For the Ross Whipple Student Managed Investment Fund, a mix of different experience levels may also develop such an interest, since entering students are exposed to and work alongside more experienced students.

The repeatable multi-semester approach with differing skill levels has its challenges. Obviously, the most difficulty arises from managing students of different experience and skill levels. Optimally, students in each level would work on a task appropriate for their level; students then progress from one level to the next. A problem has been students reverting to the mean, where less-skilled students aspire to do better while more skilled students use the less-skilled students as a benchmark and excuse to not apply themselves more. The faculty advisor has found himself advising the students as one large group instead of by experience level as should have been done.

A related issue is allocating the proper amount of time to complete tasks. Some tasks that were expected to take 15 minutes took almost the entire two-hour lab for some students. In a regular course, with expectations of students working outside of class, this would not be a problem since each student could take as much time as needed. Part of the solution to these two problems has been to enforce stricter grading. Currently, a certain amount of time is allocated to a task, and if students do not complete it in the allocated time, they must complete it outside of the lab.

When fund management began to be implemented in the new lab format, an attempt was made to align lab activities in the upper-level lab, FIN 4221 Portfolio Management Lab, with topic coverage in its two corequisites, FIN 4103 Investment Management and FIN 4213 Security Analysis. However, the coursework began to limit what could be done in the lab and the lab activities began to limit what could be done in the course. Adding to the problem is having second-semester seniors in the investment courses mixed with first-semester seniors; half of the seniors planned to graduate at the end of the semester while the other half had another semester after the current one. Currently, very little alignment exists between the lab and the corequisite courses.

Students with differing skill levels in the same lab required a mechanism to tailor assignments and instruction to the disparate levels. This is accomplished in the university's learning management system (LMS), currently Canvas. Students can be assigned refresher readings and videos and group and individual assignments with written and video instructions. Introductory teaching was needed to acclimate new students to the investment philosophy of the fund. The first semester this was tried, the refresher readings and introductory learning modules received positive feedback, but they left too little time for actually managing the fund and interacting with the more experienced students. Currently, the learning modules are employed but play a minor role.

As mentioned previously, a key benefit to the multi-semester approach is the progression of learning that builds and reinforces previous learning. Each semester in the sequence has tasks unique to the semester order while also allowing some tasks to be shared between semesters. For SMIF programs that span multiple semesters, student assignments vary by semester. Bruce and Greene (2014) offer a variety of roles and responsibilities. Student activities in the Ross Whipple Student Managed Investment Fund are just an example of what could be done. Students taking FIN 3241 Investments Lab for the first time (known herein as first-semester students) may screen for new stocks and go through a rudimentary analysis and make a short presentation on each stock. The whole lab votes, and selected stocks go through another round of more in-depth analysis. Those in the first semester of FIN 4221 Portfolio Management Lab (herein known as second-semester students) may assist the first semester students in their analysis and contribute additional research of their own. The first-semester students would also complete additional tasks such as performance measurements, learning modules, and rebalancing assessments for practice. In addition to assisting the first-semester students, the second-semester students may do more in-depth analysis of companies already in the portfolio and perform an economics analysis and create a sector rotation strategy (using ETFs). They may also work with those in FIN 4221 Portfolio Management Lab for the second time (herein known as third-semester students) in performance measurement and rebalancing assessments. Students of all levels participate in voting on investment decisions, except perhaps for trades simply to rebalance the portfolio.

The primary task of the third-semester students is updating and reevaluating portfolio and individual metrics and preparing a semester written and oral report. They deliver the oral presentation to the donor and university administration at the end of every semester. Many have a problem with public speaking and have difficulty grasping all that needs to be reported. Currently, to assist in their preparation, presenters for the semester present during the very first lab period to introduce the fund to everyone. The Dean and Associate Dean are invited to the presentation to provide a sense of importance of the presentation to the students. It is a precursor to the final presentation. The presentation is counted as part of FIN 4103 Investment Management or FIN 4213 Security Analysis, the lab's corequisites. The first day of class for FIN 4103 Investment Management or FIN 4213 Security Analysis usually occurs before the first lab period. This allows the presentation to be assigned in either FIN 4103 Investment Management or FIN 4213 Security Analysis and completed in the first meeting of the lab. Second-semester students, also taking the same corequisite, are assigned a rebalancing assessment and present their findings during the first lab period.

Note that some students may have taken FIN 3241 Investments Lab multiple times and would not really be considered "first-semester" students. Students who repeat FIN 3241 Investments Lab are assigned tasks of both first-semester and second-semester students based on their skill level. They may also be allowed some flexibility in choosing their assignments. This would add some incentive to taking FIN 3241 Investments Lab more than once. Those taking FIN 4221 Portfolio Management Lab for a third time are allowed to expand into advanced activities such as covered call strategies. All these activities just described are not set in stone and are simply illustrations of current practices.

Planned Adjustments

Few students have taken advantage of taking the labs more than the minimum number of three semesters. Since the lab's inception in the fall semester of 2015 to the spring semester of 2019, 68 students have passed or were currently enrolled in either FIN 3241 Investments Lab or FIN 4221 Portfolio Management Lab. Of course, most of these students enrolled for more than one semester, resulting in a total enrollment of 124 students over eight semesters. Only three students enrolled beyond the required minimum of three semesters. Each was enrolled in the student's fourth semester in spring 2019.

Beginning fall 2019, the finance program is structured into two tracks, a financial planning track and a financial analysis track. Both have a central theme of investments, which would allow the Ross Whipple Student Managed Investment Fund to have an integral role in both tracks. However, the more structured approach to the sequence of investment courses required adjusting the way the student fund was incorporated into the program. Such incorporation sought to address the difficulty in teaching multiple skill levels and the lack of student participation in more than three semesters. The limitations of faculty course load and student program hours remain.

First and foremost, the fund is no longer taught as a one-hour lab. In the new program, students are required to take the following sequence of courses: FIN 2233 Beginning Investments is the usual starting prerequisite in the freshman or sophomore year. Fundamentals of Investments (two credit hours) is taken in the spring of the junior year, and Investment Planning (three credit hours) is taken in the fall of the senior year. Applications in Portfolio Management (one credit hour - formerly FIN 4221 Portfolio Management Lab) is taken in the spring of the senior year. The student fund is implemented in the last three courses. Time allocated to the fund in each course is equivalent to one credit hour. Thus, three credit hours of fund management is maintained in the program. FIN 3241 Investments Lab is retained but becomes a one-hour elective course offered in the spring for sophomores who took FIN 2233 Beginning Investments in their freshman year.

The expectation is that students will become more incentivized to put forth more effort toward fund management in the courses, since their performance in the first two courses impacts two and three hours of credit instead of just one. In the fall semester, fund management is limited to students enrolled in Investment Planning. In the spring semester, those students who have completed Investment Planning will take Applications in Portfolio Management and meet with students in Fundamentals of Investing and FIN 3241 Investments Lab. The issue of varied skill levels was not eliminated but was ameliorated. The hope is that those who receive dedicated attention in the fall in Investment Planning will be able to work more independently in the spring, freeing up instructor time to concentrate on those in Fundamentals of Investments and FIN 3241 Investments Lab. Both sets of students in these lower-level courses should be at a comparable skill level.

Summary

We offer fund management through a one credit hour lab with minimal prerequisites to accommodate the environment of a small regional university. This allows early participation through multiple semesters within the constraints of student time and faculty resources. The multi-semester approach provides a stair step progression of active learning so that each additional semester builds on the previous and allows a slow growth process, which is especially important for slower students. Through a one-hour lab format, faculty teachings loads are not incumbered as much as a full course, and student participation over multiple semesters is not hindered due to limited number of hours in student coursework. The multi-semester approach is not without its problems, the most notable being the managing of different student skill levels. Insights gained from using this format helped develop a new approach for the future.

References

Block, Stanley B., and Dan W. French. 1991. "The Student-Managed Investment Fund: A Special Opportunity in Learning." *Financial Practice and Education* 1: 35-40.

Boughton, Charles M., and Katherine L. Jackson. 2018. "Variation in the Structure and Administration of Student-Managed Investment Funds." Paper presented at Academy of Economics and Finance, Houston, TX.

Brabeck, Mary, Jill Jeffrey, and Sara Fry. 2018. "Practice for Knowledge Acquisition (Not Drill and Kill)." *https://www.apa.org/education/k12/practice-acquisition.aspx*.

Bruce, Brian, and Jason Greene. 2014. *Trading and Money Management in a Student-Managed Portfolio*. Oxford: Academic Press.

Charlton Jr., William T., John Earl II, and Jerry J. Stevens. 2015. "Expanding Management in Student Managed Investment Funds." *Journal of Financial Education* 41: 1-23.

Gradisher, Suzanne, Douglas R. Kahl, John M. Clinebell, and Jerry L. Stevens. 2016. "Fiduciary and Legal Considerations for Student-Managed Investment Funds." *Journal of Education for Business* 91: 83-89.

Johnson, David W., Joe F. Alexander, and Garth H. Allen. 1996. "Student-Managed Investment Funds: A Comparison of Alternative Decision-Making Environments." *Financial Practice and Education* 6: 97-101.

Kang, Sean H.K. 2016. "Spaced Repetition Promotes Efficient and Effective Learning: Policy Implications for Instruction." *Instructional Strategies* 3: 12-19.

Kerfoot, B. Price. 2009. "Learning Benefits of Online Spaced Education Persist for Two Years." *Journal of Urology* 181: 2671-2673.

Lawrence, Edward C. 1994. "Financial Innovation: The Case of Student Investment Funds at United States Universities." *Financial Practice and Education* 4: 47-53.

Lawrence, Edward C. 2008. "Student Managed Investment Funds: An International Perspective." *Journal of Applied Finance* 18: 67-83.

Macy, Anne. 2010. "Student-Managed Investment Funds: A Framework for Regional Schools." *Academy of Educational Leadership Journal* 14: 47-62.

Mallet, James E., Lawrence J. Belcher, and G. Michael Boyd. 2010. "Experiment No More: The Long-Term Effectiveness of a Student-Managed Investments Program." *Journal of Financial Education* 36: 1-15.

North, David, and Jerry L. Stevens. 2012. "The Two-Year Student Managed Investment Fund: Coordination of Applied Investment Skills across Disciplines and Courses." Working paper, University of Richmond.

NSSE. 2016. "Annual Results 2016."
http://nsse.indiana.edu/NSSE_2016_Results/pdf/NSSE_2016_Annual_Results.pdf#page=14

Silva, Elena, Taylor White, and Thomas Toch. 2015. "The Carnegie Unit: A Century-Old Standard in a Changing Education Landscape." Carnegie Foundation Report.
https://www.carnegiefoundation.org/wpcontent/uploads/2015/01/Carnegie_Unit_Report.pdf.

Utah State University. 2018. "Estimate Study Hours."
https://www.usu.edu/asc/assistance/pdf/estimate_study_hours.pdf.

Operating a Student-Managed Investment Fund in a Small Class Setting

Qian Li[1]

ABSTRACT

The student managed investment fund (SMIF) at Midwestern State University was founded in 2010 and offers students an experiential learning opportunity. This article discusses its operational structure, investment policies, and strategies. It also discusses efforts to motivate students to run the SMIF as a productive team.

Introduction

"Successful experiential learning requires just the right amount of structure to manage the complexities of open-ended, real world situations while helping students navigate their requirements and ambiguities."
--The University of Texas at Austin Faculty Innovation Center[2]

Student managed investment funds (SMIFs) are an increasingly popular experiential learning tool among universities and students (Lawrence 2008). The benefits of students managing real money over traditional lecture-based investment courses or simulations can be tremendous: The SMIF allows students to apply basic investment theories and concepts in real markets and understand the complexity of these markets. Moreover, it provides opportunities for students to improve their skills in many areas. Clinebell and Murphy's (2016) survey results show that SMIFs aid in the long-term development of students and in achieving learning goals such as discipline knowledge, communication skills, leadership skills, and communication skills.

A SMIFs operation and success may depend on its university environment and student body. For example, Kahl (1997) discussed the challenges and opportunities faced by a metropolitan university, University of Akron. Mallett et al. (2010) summarize experiences from a pioneer SMIF at Stetson University. This article introduces Midwestern State University's (MSU Texas) SMIF and discusses the unique challenges it has faced. It also discusses MSU's experience in motivating the students to run the SMIF as a productive team.

The article is organized as follows: The next section provides a brief introduction of the SMIF in MSU's Dillard College of Business Administration (DCOBA). The third section describes the fund's investment guidelines and strategies. The fourth section discusses the practices used in the SMIF class to motivate students and improve teamwork. The fifth concludes the article.

Overview of the Dillard College SMIF

MSU Texas is a small public university in Texas with an enrollment of about 5,600 students. Its DCOBA has an enrollment of slightly under 1,000 undergraduate students, of whom about 100-120 major in finance, and 50-60 MBA students. Hence, DCOBA's SMIF operates in a smaller scale environment than many other SMIFs.

DCOBA's SMIF was created in 2010 with a generous donation of $400,000 from the Dillard family, a long-term supporter of the college. The Dillard fund also covers expenses for a Bloomberg Professional Services annual subscription and the equipment in the computer lab where the Bloomberg terminals are installed. Since the SMIF's inception, there have been no additional donations into or withdrawals from the fund.

---

[1] Dillard College of Business Administration, Midwestern State University, 3410 Taft Blvd, Wichita Falls, TX 76308, qian.li@msutexas.edu.

[2] https://facultyinnovate.utexas.edu/experiential-learning

As of the end of August 2018, the fund is worth approximately $680,000. The fund is owned by MSU's University Foundation and managed by the students enrolled in the SMIF class. The SMIF students are responsible for producing a performance report to the stakeholders at the end of each semester. Currently, the return and income from the SMIF's investments stay in the fund, and the participating students cannot receive compensation or charge fees for managing the portfolio. While the current size of the portfolio is not large enough to support regular use of the income on scholarships or student activities such as field trips, it may become a good resource for such uses in the future as the portfolio grows over time.

The SMIF was initially set up to be an investment club for students. However, the faculty soon realized that the club attracted little interest from students and that student participation was not reliable in a club setting. Therefore, the faculty decided to change the fund to a course-based format. The course format has made student attendance and participation more structured, and the grades/credits serve both as a reward to the participating students and as a tool to hold students more accountable.

MSU's SMIF course is cross-listed for both graduate (MBA) and undergraduate students. Every semester approximately 6-12 students enroll in this class, with 60%-80% of them being undergraduate students. The SMIF literature does not provide much information about average enrollment in SMIF programs across schools, as the literature focuses more on the size of the portfolio than on the size of the class that manages the portfolio. However, based on (informal) discussions with faculty members from other programs, MSU's SMIF class has a smaller number of student participants than found at other universities. In addition, while students have the option to take the course for two semesters, earning up to 6 credit hours, most students only take the SMIF class for one semester. This makes the student/management turnover rate even higher than in schools that require students to take the SMIF course for two consecutive semesters.

The SMIF course has the following learning goals for students:

1. Critical thinking, problem-solving, and decision-making skills: The SMIF experience should improve students' critical thinking, problem-solving, and decision-making skills, as it requires the students to synthesize information, analyze economic and financial market conditions, evaluate firms' financial conditions and value, and make investment recommendations and decisions.

2. Knowledge integration: To understand the value of a company, conduct reasonable valuation, and make sound investment decisions, students need to integrate their knowledge in various areas such as economics, management, business strategy, and accounting as well as portfolio management.

3. Business communication skills: The SMIF course is a reading-, writing- and presentation-intensive course. In this class, students write several reports and make several presentations based on their research of specific sectors/industries and individual companies/stocks. Students are also expected to deliver a formal report and presentation about the fund performance to the college at the end of the semester. This is one of the best opportunities offered in the business school for students to practice and improve their writing and oral communication skills.

4. Technology skills: The SMIF students are expected to use Excel intensively in financial analysis and valuation models. They also practice using Bloomberg Terminals to collect and interpret market and financial information. Students are also required to obtain the Bloomberg Market Concepts (BMC) certificate, which can demonstrate their proficiency in the Bloomberg system and their market and industry knowledge to potential employers. In addition, students are also encouraged to use data analytics tools such as R or Python to retrieve and summarize market data.

5. Team collaboration and leadership skills: The SMIF has a diverse group of students from various backgrounds and provides an opportunity for students to work with others to achieve group goals. In addition, students are expected to step up to the leadership roles and ensure smooth and effective operation of the fund.

As explained in the next section, these learning goals are implemented throughout the course by the investment process and the assignments and tasks assigned to the SMIF students.

The Investment Process

MSU's SMIF students need clear guidelines on which stocks they may or may not select. Currently, students are only allowed to invest in U.S. domestic stocks traded in the U.S. market. In addition, according to the SMIF's current Investment Policy Statement, students should avoid companies with total market capitalization lower than $2 billion, and stocks with prices lower than $5 per share. Other guidelines include that no single stock should weight more than 5% of the portfolio value, and that the portfolio's beta should not exceed 1.5. Bonds, derivatives, and alternative investments are not allowed in the portfolio, mainly because MSU currently does not offer classes in these areas in its curriculum. In addition, students are discouraged from investing in index ETFs or mutual funds because, as noted by Cooley and Hubbard (2012), passively indexing the market would fail to develop students' analytical and decision marking skills and portfolio management experience. As of the end of May 2018, approximately 18% of the SMIF's portfolio is in cash, while the rest is invested in 33 common stocks.

Each semester every student is required to conduct at least one independent stock research and analysis and to make an investment recommendation on the stock he or she picks. Students are expected to make investment recommendations and decisions based on the valuation of stocks and focus on fundamental analysis. Students are instructed to follow a top-down approach and start by analyzing the macroeconomic environment and identifying which sectors' industries are likely to perform better under the current and future economy. The next step is an industry analysis in which students examine the competitive environment in the industry, industry profitability, and growth prospects. After that they try to identify a company that is the strongest player in the industry and perform the company analysis.

After completing a thorough stock analysis, the student analyst is required to make a presentation as the "stock pitch" to the class. The stock pitch also comes with a target price and stop-loss price recommendation. The stop-loss order is to protect our investment in case the stock price falls substantially after the purchase, especially during the winter and summer breaks. Each analyst's presentation takes about 20-30 minutes, which allows the student to introduce the stock with detailed information and analysis. There is also a 5-10-minute Q&A session after each presentation, during which other SMIF members can ask the presenter questions regarding how reasonable the information, the analysis, and the assumptions are. After the Q&A session, the group then votes on the stock recommendation. If the majority of the group votes to buy, the instructor places the buy order. This process is a good combination of individual effort and team collaboration. The presenter knows that they need to prepare for the analyst report, the presentation, and questions from the audience. The audience practices its critical thinking by processing the information it learns during the presentation and asking questions.

Once a stock is approved and added to the SMIF's portfolio, a student manager is assigned to follow up with the stock and is responsible for informing the class of any news and material information regarding the company and the stock, including market developments, earnings release updates, revised outlooks, and mergers and acquisitions. The designated student also makes recommendations to the class regarding whether to continue holding an existing stock, to sell an existing stock, or to set new parameters such as target price and/or stop-loss price on an existing stock. If a majority of the class votes to sell the stock or to change stop-loss prices, the instructor places the order accordingly. Depending on the class size and the number of stocks held in the portfolio, each student has 2-4 stocks to cover each semester. This practice helps students understand what drives the price movement in stocks and how the stock market reacts to different information. We also try to assign students stocks in the same or similar industry so that they can become our "industry expert" and gain substantial industry-specific knowledge.

In addition to watching individual stocks in our portfolio, we also keep an eye on the entire portfolio. Throughout the semester, students track the returns, risks, and sector/industry diversification of the portfolio, and make group decisions to rebalance the portfolio when necessary.

Motivating Students and Improving Teamwork in the SMIF

One of the unique features of MSU's SMIF course compared to other lecture-based courses is that it is supposed to be student-led, and the success of the SMIF greatly depends on students' participation and collaboration. In the early years of MSU's SMIF, faculty observed that one or two students' delay and inaction negatively affected the SMIF group's performance and demotivated other students. Though there is no perfect solution, several approaches have been tried to motivate students and improve teamwork.

Assigning Students Specific Operational Tasks

Students are given clear guidelines at the beginning of each semester about what they need to do for the fund. To promote accountability, students are assigned various management roles and positions based on their interests and personalities. This is similar to the practices of the University of San Francisco's SMIF, as discussed by Chincarini and Le (2018), but MSU SMIF's positions are more clearly defined and more concentrated on the routine operation and tasks of the fund because of the smaller number of students in its program. These positions generally include:

- Class Presidents whose tasks include but are not limited to the following:
  - Presiding over class meetings.
  - Taking class votes and reporting the results to the instructor.
  - Coordinating with other SMIF members to ensure each project and assignment, including final presentations, can be finished on time.
- Fund Economists whose tasks include but are not limited to the following:
  - Following the economy and financial markets and providing weekly class updates.
  - Explaining to the class how economic and financial market movements would affect the SMIF.
- Senior Analysts whose tasks include but are not limited to the following:
  - Working with the economists to discuss the sectors of interest and make informed recommendations on new investment opportunities that could enhance or diversify the portfolio.
  - Ensuring the understanding and use of appropriate valuation models by all fund members.
- Portfolio Evaluation Analysts whose tasks include but are not limited to the following:
  - Evaluating the SMIF's holdings and performance at the beginning of the semester, at mid-semester, and at the end of the semester.
  - Reporting to the class about the allocation, diversification, and performance of the portfolio and explaining the reason(s) for changes in allocation and performance.
  - Leading the discussions of portfolio performance in the final presentation and report.
- Public Relation/Communication whose tasks generally include the following:
  - Making arrangements with guest speakers.
  - Planning the final presentation event, including sending invitations and reminders.
  - Attending the college's recruiting events and talking to parents and students about their experience in the SMIF course.

Since this system was implemented in 2015, students have become more active in the fund management, more likely to finish the tasks on time, and feel more involved and more accomplished. This system also saves the instructor a lot of time from having to do what the students are supposed to do, which is important considering the instructor's heavy teaching load and research obligations.

Teamwork Peer Evaluation

A peer review process has been used since Fall semester 2017. At the beginning of the semester, a peer evaluation form is distributed to the class. The form makes students aware of some basic elements of good teamwork such as attending group meetings regularly, providing constructive ideas and opinions to group discussions, and completing team assignments on time. A sample peer evaluation form is provided in the Appendix.

At the end of the semester, the instructor collects completed evaluation forms from group members and uses them as inputs for determining students' teamwork grade component. This peer review system has helped hold each student accountable and reduce friction and complaints among students. Overall, the teamwork peer evaluation system, together with a set of clearly defined management roles, helps mitigate the free riding problem frequently seen in student group projects and assignments.

Conclusion

SMIFs are great experiential learning tools for business school students. They have helped students grow into competent and confident job candidates in the very competitive finance job market. It is not surprising that these programs have become an important recruiting tool for business colleges. However, without a reasonable system to motivate students and hold them accountable, SMIF programs can suffer from free riding and irresponsible investment decisions. This article discusses approaches used in MSU's SMIF classes to help students achieve numerous learning goals and contributes to the discussion on best practices in experiential learning and in organizing and operating a student-led learning activity such as a SMIF.

References

Chincarini, Ludwig, and Tuan Anh Le. 2018. "Managing a Student-Managed Fund: The University of San Francisco Experience." *The Journal of Trading* 13: 84-111.

Clinebell, John, and Jody Murphy. 2016. "Student Managed Investment Funds: The Perspectives of Alumni." *Journal of Financial Education* 42: 304-312.

Cooley, Philip L., and Carl M. Hubbard. 2012. "FAQs about Student Managed Funds." *Advances in Financial Education* 10: 72-84.

Kahl, Douglas R. 1997. "The Challenges and Opportunities of Student-Managed Investment Funds at Metropolitan Universities." *Financial Service Review* 6: 197-200.

Lawrence, Edward C. 2008. "Student Managed Investment Funds: An International Perspective." *Journal of Applied Finance* 18: 67-83.

Mallett, James E., Lawrence J. Belcher, and G. Michael Boyd. 2010. "Experiment No More: The Long-Term Effectiveness of a Student-Managed Investments Program." *Journal of Financial Education* 36: 1-15.

APPENDIX: SAMPLE PEER TEAMWORK EVALUATION FORM

Write the name of each of your team members in a separate column. For each person, indicate the extent to which you agree with the statement on the left, using a scale of 1-4 (1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree). Total the scores in each column.

Evaluation Criteria	Student #1	Student #2	Student #3	Student #4	Student #5
Attends group meetings (online and offline) regularly and arrives on time.					
Contributes meaningfully to group discussions.					
Completes team assignments on time.					
Prepares work in a quality manner.					
Demonstrates a cooperative and supportive attitude.					
Contributes significantly to the success of the project.					
TOTAL					

Some Potential Issues with a Student Managed Investment Program at Small Liberal Arts Colleges: The Berry College Experience

Ken Johnston, John Hatem, Arman Kosedag, and Quentin McTeer[1]

ABSTRACT

This paper focuses on potential issues with student managed investment funds (SMIFs) with emphasis on small liberal arts colleges. A potential solution to alleviate some of these issues is proposed (a policy portfolio) and implemented. The initial results are presented and the subsequent influence of these results on club behavior is discussed.

Introduction

Student managed investment funds (SMIFs) have grown significantly in number over the last 30 years (Lawrence 1990, 1994, 2008). Several older papers have described how SMIFs were created and how they are run (for example, Block and French 1991; Johnson et al. 1996; and Kahl 1997). Goff and Cox (2016) provide a detailed description of the process of starting and operating a SMIF and provides an example of how experiential learning can enhance the educational process while providing other benefits to students and their college.

Previous papers tend to focus on the benefits of having a SMIF. This paper focuses on potential issues with SMIFs at small liberal arts colleges. A potential solution to alleviate some of these issues is proposed and implemented. The initial results are presented and the subsequent influence of these results on club behavior is examined.

Background

The Berry Investment Group (B.I.G.), a student organization housed in Berry's Campbell School of Business, was created in 1997 after an individual donated $100,000 in initial funding for students to invest and manage the portfolio as a learning experience. Eligible assets for investment include stocks, ETFs, bonds, mutual funds, and cash.[2] In accordance with the B.I.G. constitution, trading of futures or options and short selling of any individual security are excluded. Additionally, no other alternative investment strategies such as private equity or direct loans are allowed.

A portion of each year's gains is harvested and used to enhance the vibrancy of Berry College. These gains are calculated by taking the original $100,000 investment for the group and adjusting that amount for inflation. The difference between the current market value of the portfolio and the inflation adjusted $100,000 is the amount that can potentially be donated to the school. One such donation enhances club members' educational experience each year by funding a trip to the G.A.M.E. conference in New York City. The B.I.G. officers, in consultation with the faculty advisor, determine the attendees. The main factors are the level of participation by each member in club activities and overall expense. Some other examples of recent financial contributions to Berry College include funding for air conditioning in the chapel, outdoor basketball courts, and a ticker tape in the Campbell School of Business lobby.

---

[1] Johnston (corresponding author): Campbell School of Business, Berry College, Mt. Berry, GA 30149, kjohnston@berry.edu. Kosedag: Campbell School of Business, Berry College, Mt. Berry, GA 30149, akosedag@berry.edu. McTeer: Campbell School of Business, Berry College, Mt. Berry, GA 30149, quentin.mcteer@vikings.berry.edu. Hatem: Parker College of Business, Georgia Southern University, Statesboro, GA 30460, jhatem@georgiasouthern.edu.

[2] Although allowed by the B.I.G. constitution, students are discouraged from investing in ETFs and mutual funds.

Issue 1: Club Makeup

Berry College is a small liberal arts college north of Atlanta which graduates a couple dozen finance majors in a typical year; a quarter or so of those majors are members of the club that manages the fund. Since Berry College does not frequently offer courses with fewer than eight students, organizing the SMIF as a club was more appropriate than offering a finance course incorporating students' active management of real funds.[3] To increase the number of students participating in the club, B.I.G. is open to all Berry students, regardless of major, and has no formal application process to screen students. Being a club open to all majors presents a number of issues/opportunities not faced by 'finance student only' clubs. This approach to club membership is a unique feature of Berry College that promotes financial literacy among students across the college.

B.I.G.'s student composition brings at least two limitations. First, meetings have to be much less quantitative in nature and more instructional than the typical SMIF meeting. Officers of the club (finance majors only) have to conduct numerous seminars during the weekly meetings to teach the basics of the investment process (along with the language of business). Second, the faculty supervisor does not attend the weekly meetings, as it tends to intimidate the non-finance majors and hurts retention. In its place there is a weekly meeting with the faculty supervisor and the officers where quantitative issues are addressed.

A benefit of the club organization, as opposed to a course, is that it allows participants to gain portfolio management and asset allocation skills over multiple years. This is important because it allows participants to see the effects of changing economic conditions, political climates, and interest rates on the portfolio. A semester or year-long course would not necessarily be long enough to observe such changes.

Although having majors from outside the business school has provided additional challenges, it has been a strength for the club. The non-finance majors have pitched stocks in their areas of study, where they have some informational advantage over the typical finance student. Examples include a pre-med student pushing for a company that has a drug in an FDA phase three clinical trial and a computer science major advocating for a new high tech firm. When pitching a stock, non-finance students are paired with finance majors to help them analyze the company and follow the investment decision process.

Issue 2: Investment Decision Process Given Club Makeup

In the first couple of meetings each fall, stock pitch sheets are introduced. Care is taken not to overwhelm non-finance majors with too much technical finance information too quickly. Table 1 shows a condensed version of the first sheet which contains security and portfolio considerations. Table 2 shows a condensed version of the second sheet which deals with qualitative observations. Clearly, the focus is on fundamental analysis while understanding the implications of market efficiency.[4]

Discounted cash flow techniques (dividend discount models, free cash flow to firm, free cash flow to equity, etc.), and price multiple approaches where a stock is worth some multiple of its future earnings (such as price-to-earnings ratios or enterprise value-based ratios) are popular subjects in finance text books in stock valuation. Developing accurate future estimates for the inputs of these models in the real world is difficult and somewhat reduces their relevance[5]. Thus, although the concept of intrinsic value exists, searching for it is impractical. While the analysis does not try to come up with the true (intrinsic) price, it does examine the financial ratios and measures along with the thought process used in fundamental analysis. This technique is similar to Warren Buffet's approach, when 'buying a stock is buying a business.' The second page of the stock pitch sheets (Table 2), qualitative observations, addresses this as it examines why the club wants to be in this business in the future.

The current president of B.I.G. summarizes the investment decision making process as follows: "Students in the club decide on specific investments through a top-down approach. We first analyze our current asset allocation and measure which industries we have no exposure to and ask if we need to get into them. After

[3] Lawrence (2008) reports that most student managed funds (71%) are tied to an existing class.

[4] In the past, some technical analysis techniques have been employed to try and time market trades. This was not successful, and the practice has been abandoned. However, the finance officers teach technical analysis to any interested members.

[5] Neumann (2017) focuses practical challenges (specifically on the estimation of long-run growth rate) of dividend discount model within the context of student managed investment funds and argues that neither average annual growth rate nor the sustainable growth equation (ROE*b, return on equity times retention ratio) provide a reliable estimate.

we pinpoint a certain equity-type that we think we need access to, we assign members to those individual equity-types and give them the responsibility to identify the best stock investment in that particular equity-type."

Table 1: Stock Pitch Sheet

SECURITY CONSIDERATIONS	
SALES GROWTH, PROFIT MARGINS	
EARNINGS GROWTH	
DIVIDEND GROWTH	
PRICE RELATIVE TO HIGH-LOW	
P/E RELATIVE TO INDUSTRY	
INSTITUTIONAL OWNERSHIP	
ROE/ROA/EBITDA CURRENT RATIO	
DEBT RELATIVE TO COMPETITORS	
DEBT RELATIVE TO INDUSTRY	
BETA (RELATIVE TO MARKET)	
MARKET CAP	
INSIDER ACTIVITY	
R&D SPENDING	
PORTFOLIO CONSIDERATIONS	
INDUSTRY GROUPING	
CURRENT ASSET ALLOCATIONS	
PERCENTAGE HOLDING RECOMMENDATION	
ADDITIONAL COMMENTS:	

Neely and Cooley (2004), after surveying instructors of SMIFs included in the Association of Student Managed Investment Programs, found that virtually all of the respondents indicated that their fund engages in stock picking, but the split was 50/50 between top-down and bottom-up analysis. In a more recent survey, Lawrence (2008) found that bottom-up was the most employed method (37%), then top-down (27%), followed by buy and hold (11%).

Table 2: Qualitative Stock Pitch Sheet

QUALITATIVE OBSERVATIONS
What do we know that Wall Street doesn't?
Where is the expected value coming from?
What is the competitive advantage?
What does the industry look like in the future?
What does the economy look like in the future?
What is your expected time frame?
ADDITIONAL COMMENTS:

Neely and Cooley (2004) found that 90 percent of the SMIFs that responded to their survey had an investment policy. Mallett and Lerro (2002) believe that the process of stock selection in a SMIF should be directly related to the guidelines expressed in the constitution or mission statement of the organization. While the B.I.G. constitution states that the portfolio is to be managed in this top down fashion, in reality, over the years, it has typically started with security selection, and students tend to concentrate their efforts in areas they have interest in regardless of the portfolio implications. The faculty advisor has to redirect the officers of the club to examine if this potential investment is a good one, given economic and industry (sector) expectations and portfolio implications (discussed in more detail in Issue 3).

One of the most frequent issues B.I.G.'s faculty advisor faces is determining the amount of time an investment is held. The faculty advisor frequently asks the officers to explain why each security is currently held. Many times, the original reason the security was bought no longer applies. If that is the case, the faculty advisor has the officers reanalyze the company to determine if it should still be held. If not, it needs to be sold and the money invested in a new security, in other companies currently in the portfolio, or held as cash while other opportunities are researched.

Table 3 shows the performance of B.I.G.'s portfolio for the last seven years. Over this period, B.I.G.'s portfolio beat the S&P 500 by 1.03% on average.[6] However, results vary significantly; B.I.G.'s portfolio

[6] Jones and Swaleheen (2014) look at the performance of a student managed fund at the University of Florida. They demonstrate that a student managed fund can do as well as the S&P 500 from a risk/return perspective. However, Lei and Li (2015) demonstrate

outperformed the S&P 500 for three years but underperformed it for four years. Most students tend to be members of the club less than two years. If they happen to beat the S&P 500 in those years, the students tend to believe that they have the ability to consistently beat the market. This issue, along with others discussed in the paper, will be addressed in the future by the creation and use of a policy portfolio.

Table 3: B.I.G.'s Performance

Time Period	B.I.G.'s Return	S&P 500 Return	B.I.G. - S&P Return
8/1/2011-7/31/2012	1.39%	9.21%	-7.82%
8/1/2012-7/31/2013	33.42%	25.36%	8.06%
8/1/2013-7/31/2014	12.93%	15.37%	-2.44%
8/1/2014-7/31/2015	6.02%	11.47%	-5.45%
8/1/2015-7/31/2016	3.85%	5.77%	-1.92%
8/1/2016-7/31/2017	17.83%	16.06%	1.77%
8/1/2017-7/31/2018	30.92%	15.88%	15.04%
Average	15.19%	14.16%	1.03%

Issue 3: Size Affects the Extent of Portfolio Analysis

The size of the club affects analysis: large programs tend to have sector leaders and multiple analysts for each sector (i.e., discretionary, energy, healthcare, industrials, media and telecom, real estate and miscellaneous, staples, technology). Thus, they have advocates for each sector. Given that B.I.G. does not have leaders and/or analysts allocated to each sector, asset allocation often takes a backseat to security selection even though it is discussed in the constitution and at meetings. Students tend to concentrate on security selection and not on asset allocation.

Table 4 shows B.I.G.'s asset allocation compared to the S&P 500 (as of December 31, 2017). B.I.G.'s sector weightings are substantially different from the S&P 500's weights. As of December 31, 2017, about 69% of the portfolio is in just two sectors (Technology and Consumer Discretionary).

Table 4: B.I.G.'s Asset Allocation as of December 31, 2017

Sector	S&P 500 Weight	B.I.G. Weight	Difference	Over/Under?
Financials	14.79%	8.42%	-6.37%	Underweighted
Healthcare	13.78%	13.18%	-0.60%	Underweighted
Consumer Discretionary	12.21%	37.14%	24.93%	Overweighted
Industrials	10.27%	0.00%	-10.27%	Underweighted
Consumer Staples	8.20%	4.65%	-3.55%	Underweighted
Energy	6.08%	0.00%	-6.08%	Underweighted
Utilities	2.94%	0.00%	-2.94%	Underweighted
Real estate	2.89%	0.00%	-2.89%	Underweighted
Materials	2.99%	4.60%	1.61%	Overweighted
Telecom	2.06%	0.00%	-2.06%	Underweighted
Technology	23.78%	32.01%	8.23%	Overweighted

From discussions with other faculty advisors at the R.I.S.E and G.A.M.E conferences, most SMIFs choose a stock market index as a benchmark, typically the S&P 500. Given the allocation results from Table 4,

that the out-performance of the portfolio is due to using the S& P 500 index without its dividend component. When the S&P 500 total return is used as the benchmark, the out-performance of the equity portfolio disappears.

clearly this is not an appropriate yard stick for B.I.G. In any given year, the sector holdings differ substantially from the S&P 500 index.

Issue 4: Investment Monitoring and Control Procedures

Trades can only occur after a vote at the weekly meetings. Given the volatile nature of the stock market, sell stop orders are placed on all positions (1.25 standard deviations on the downside). This limits intermeeting downside exposure. It also helps with sentimental attachment to securities (loss aversion). The faculty advisor leaves it up to the officers of the club to determine what data to use to come up with the distribution to determine the standard deviation. An option for the group to "reset" the bound should be left open only on the upside. If the security explodes during a given week, then a new downside stop could/should be placed on the security in order to raise the sell stop price.

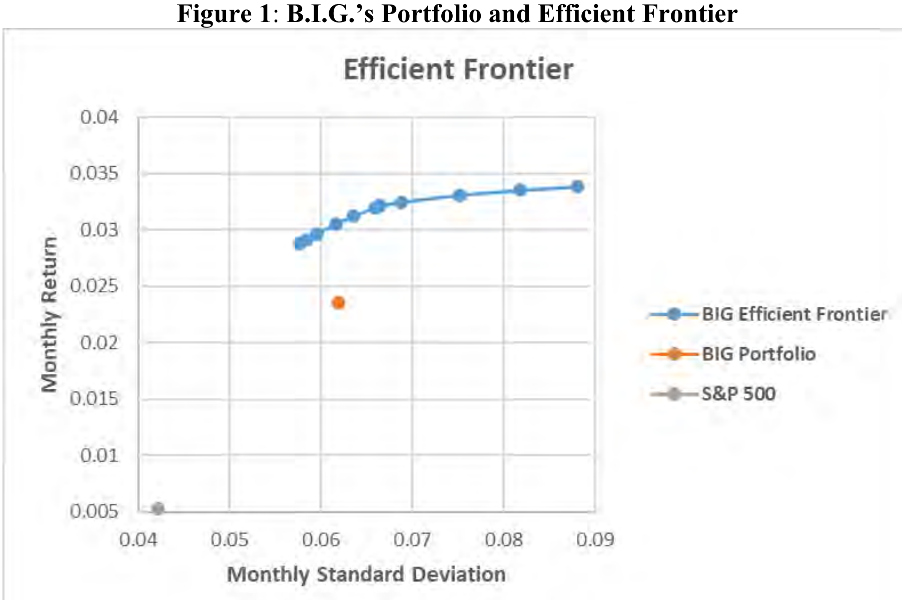
During summers, the same procedures are used. Given the size of Berry College's finance program, typically no classes are offered in the summer. The club does not meet in the summer and therefore cannot make any decisions in the summer, as voting is required to make trades. A wider range is applied, giving the security a little more time to recover (1.85 standard deviations sell stops on the downside). This eliminates the possibility of a black swan event decimating the portfolio in the summer.

Issue 5: Are SMIFs Taking on Enough Risk?

SMIFs have an infinite time horizon since they are owned by perpetual institutions. This allows them to focus on strategy and stock picking. The portfolio does not need to be adjusted for risk due to the age of the owner. However, appropriate allocation shifts should be made during economic downturns and expansions. Given the portfolio's time horizon, the asset allocation can be more risky in comparison to those with a short(er) time horizon. Since the club follows conventional wisdom that the higher the perceived risk, the higher the expected return, the portfolio can afford to take on appropriately higher risk. The group must recognize, however, that with the added levels of risk, volatility would likely increase as well.

In the spring 2018 portfolio management class (which all officers of the club take), students were asked to create an historical efficient frontier of the current stocks in B.I.G.'s portfolio using ten years of monthly data. The constraints and data were adjusted to see how the efficient frontier changes. Results typically looked like Figure 1, where B.I.G.'s portfolio is significantly below the efficient frontier and near the most risk averse section of it. When asked to examine what drives this result, the students came to the conclusions that the S&P 500 is not an appropriate benchmark and that they are not taking on enough risk given the infinite timespan of the college. The faculty advisor had two of the students present their results to the club.

Figure 1: **B.I.G.'s Portfolio and Efficient Frontier**



Using a Policy Portfolio for Realistic Performance Expectations

To obtain an appropriate benchmark, the faculty advisor proposed that the B.I.G. officers develop a policy portfolio/customized benchmark. This involves using ETFs for the various sectors the officers want to invest in and historical data to create the efficient frontier. Substantial learning opportunities occur through this process of developing the efficient frontier of sector ETFs to determine the policy portfolio. Students get an understanding of how choices affect results: length of historical time period tradeoffs (whole business cycle vs. partial data), weekly vs. monthly data, and how results can be driven by constraints (minimum and maximum percentage allowed for each sector). The faculty advisor would allow the students to make the decisions, but lead the discussion on the tradeoffs. This exercise would also teach the members how to deal with incomplete data because some data will be incomplete since the ETFs are relatively new. This introduces the finance students to synthetic data creation.

Through examining the sector ETFs' efficient frontier, the finance students determine the policy portfolio (or a specific benchmark). Thus, the policy portfolio has the normal weightings and asset classes that would be held under normal economic conditions while satisfying the long-term investment objectives and constraints. The use of a policy portfolio now brings the discussion of asset allocation to the forefront. Voting would be required when significant adjustments are made to sector weights away from benchmark. Given the unlimited timespan, the chosen allocation for the benchmark needs to be somewhere near the top third of the efficient frontier. This forces the club, if it wants to beat the custom index, to take on more risk.

Having a policy portfolio/customized index will allow the students to decompose any excess return into security selection and asset allocation effects. When the investment portfolio return differs from the policy portfolio return, the club now has the ability to separate security selection effects (beating the market) from asset allocation effects (having a different asset allocation than the policy portfolio).

Asset Allocation Effect

The asset allocation effect results from adjusting the portfolio sector weights away from the designated benchmark weights; it is sometimes called tactical asset allocation or market timing effects. It is calculated using the following formula for each sector and then summing over all asset classes/or sectors:

Asset allocation effect = (actual portfolio weight – policy portfolio weight) x (asset class return in – total return of policy portfolio)

Security Selection Effect

The security selection effect is the additional return earned over the policy benchmark while asset allocation remains constant, so it gives the club's ability to beat the market. It is calculated using the following formula for each sector and then summing over all asset classes/or sectors:

Security selection effect = actual portfolio weight x (asset class return in actual portfolio– asset class return in policy portfolio)

The implicit assumption is that the club's portfolio is weighted the same as the policy portfolio in each sector, thus any difference here is due to security selection. By breaking the difference in the portfolios' and benchmarks' return into these two components, club members can get a much more realistic idea on how difficult it is to actually beat the market.

Implementing the Policy Portfolio: Results and Learning Outcomes

In spring 2018, after B.I.G.'s officers had created efficient frontiers with B.I.G.'s stock holdings in the portfolio management class, they created a policy portfolio. They used monthly index ETF data from December 2008 – December 2018 and examined the resulting efficient frontiers while changing the constraints related to maximum and minimum allocations. Table 5 shows the policy portfolio. The officers chose to invest in seven sectors, with the highest weight being in healthcare (35%) and the lowest weights being in energy, financial services, consumer cyclical, and real estate (5%).[7]

[7] Comparing Tables 4 and 5, the number of sectors was decreased from 11 to 7. This was driven by the efficient frontier results (similar efficient frontiers with fewer sectors to examine).

Table 5: B.I.G.'s Policy Portfolio

Sectors	ETF	Weight
Energy	VDE	5.00%
Financial Services	VFH	5.00%
Technology	QQQ	25.00%
Consumer Cyclical	XLP	5.00%
Healthcare	XLV	35.00%
Consumer Defensive	VDC	20.00%
Real Estate	RWO	5.00%
		100.00%

B.I.G. approved the policy portfolio near the end of the spring semester in 2018. Table 6 shows the results for the S&P 500, B.I.G.'s policy portfolio, and B.I.G.'s portfolio. Periods are created whenever a trade occurs. In every period, the policy portfolio's return is higher than the S&P 500's; this reaffirmed that the S&P 500 was an inappropriate benchmark. B.I.G.'s portfolio outperformed the policy portfolio benchmark in two subperiods and underperformed in the other two. There was no consistency in the cause of under or over performance; sometimes it was due to security selection, and other times it was due to asset allocation.

Table 6: Return Performance–July 31, 2018-December 31, 2018

	S&P 500 Return	Policy Portfolio Return	Big Portfolio Return	Over/Under performance	Asset Allocation Effect	Security Selection Effect
July 31 - Sept 21 2018	4.03%	4.20%	7.32%	3.12%	-0.30%	3.42%
Sept 21 - Oct 29 2018	-9.84%	-6.98%	-11.87%	-4.88%	1.55%	-6.44%
Oct 29 -Nov 28 2018	3.88%	4.24%	4.11%	-0.13%	-0.70%	0.57%
Nov 28 - Dec 31 2018	-8.64%	-8.20%	-7.37%	0.83%	0.74%	0.09%

In discussions with B.I.G's the officers, it is clear that the policy portfolio and decomposition of returns influenced how the club makes decisions. Asset allocation now plays a more important role in the process. In addition, by tracking results going forward, there will be an historical record of B.I.G. performance and what drives it which should help students have a more realistic view of the difficulty in beating the market (finding undervalued securities).

Currently the policy portfolio has 85% invested in three sectors; is this optimal? It's driven by historical sector returns and correlations; are they chasing the past? There is more learning to come.

Conclusion

Berry College emphasizes "learning by doing," as has been evident from its comprehensive student-work program since the college's inception in 1902. Providing opportunities for practical applications and hands-on experience still remains a distinguishing feature of the college. Affording Berry students (finance majors in particular, as well as any other majors across the college) an opportunity to learn about stock valuation and portfolio management by using real money did not start until 1997 (the 1980s mark the proliferation of SMIFs as reported by Lawrence 2008). As such, it could be considered a relatively young program. One important feature of the program is that it operates as an investment club essentially being run autonomously by students, and it is open to all students regardless of major. Not limiting membership to finance and/or business majors only helps to promote financial literacy across the college. In addition to learning practical aspects of investments and empowering their critical thinking, club members develop ethical decision-making skills by allotting part of the gains for the betterment of the college and surrounding community. The use of a policy portfolio is suggested as a solution to the lack of focus on asset allocation and limited risk exposure while also developing more realistic performance expectations.

References

Block, Stanley B., and Dan W. French. 1991. "The Student-Managed Investment Fund: A Special Opportunity in Learning." *Financial Practice and Education* 1: 35-40.

Goff, Delbert C., and Don R. Cox. 2016. "A Faculty-Driven Approach to Starting and Operating a Student-Managed Investment Fund." *Journal of Economics and Finance Education* 15: 14-22.

Jones, Travis L., and Mushfiq Swaleheen. 2014. "Student Managed Investment Fund Performance: A Look at Equity Portfolio Data." *Journal of Economics and Finance Education* 13: 58-63.

Johnson, David W., Joe F. Alexander, and Garth H. Allen. 1996. "Student-Managed Investment Funds: A Comparison of Alternative Decision-Making Environments." *Financial Practice and Education* 6: 97-101.

Kahl, Douglas R. 1997. "The Challenges and Opportunities of Student-Managed Investment Funds at Metropolitan Universities." *Financial Services Review* 6: 197-200.

Lawrence, Edward C. 1990. "Learning Portfolio Management by Experience: University Student Investment Funds." *Financial Review* 25: 165-173.

Lawrence, Edward C. 1994. "Financial Innovation: The Case of Student Investment Funds at United States Universities." *Financial Practice and Education* 4: 47-53.

Lawrence, Edward C. 2008. "Student managed Investment Funds: An International Perspective." *Journal of Applied Finance* 18: 67-83.

Lei, Adam Y.C., and Huihua Li. 2015. "Portfolio Performance Evaluation Benchmark: A Note." *Journal of Economics and Finance Education* 14: 1-7.

Mallett, James, and Anthony Lerro. 2002. "Establishing and Running a Student Managed Investment Program." Unpublished paper. *http://www.slideshare.net/khanyasmin/establishing-and-running-a-Student-Managed-investment-program.*

Neely, Walter, and Philip L. Cooley. 2004. "A Survey of Student Managed Funds." *Advances in Financial Education* 2: 1-8.

Neumann, John. 2017. "A Framework for Cashflow Valuation Models within Student-Managed Investment Funds." *Journal of Financial Education* 43: 120-140.

Editor's Note

E. Frank Stephenson

The authors of the Johnston et al. paper appearing in this issue are affiliated with Berry College and Georgia Southern University. Hence, they are colleagues of mine or finance co-editor Bill Yang. To avoid any appearance of conflict of interest, Professor Yang and I both recused ourselves from the paper's review process. Professor Edward Graham of UNC-Wilmington, a member of the journal's editorial board, supervised the paper's review.

By coincidence, the Li and Blaylock student-managed investment fund papers published in this issue were submitted for review shortly after the Johnston et al. paper. Since Professor Graham was already supervising the review of the Johnston et al. paper, he kindly agreed to oversee the Li and Blaylock papers as well. We greatly appreciate Professor Graham's service on the editorial board and his assistance with these three papers appearing in this issue.

We are also extremely grateful to Johnathan Norman of Berry College for his diligent copy editing assistance. This issue and several others have benefitted from his attention to detail.