JOURNAL OF
FINANCIAL EDUCATION

Vol. 51 Spring 2025 No. 1

EDUCATIONAL RESEARCH

CREATIVE PEDAGOGY

SPECIAL TOPICS

CASE STUDIES

Financial Education Association

INSTRUCTOR'S CASE SOLUTION

You can obtain a detailed case solution and class discussion guide to any case published in the *Journal of Financial Education* by sending a written request, including your email address, on university stationery to:

Richard J. Fendler, Editor
Journal of Financial Education
Robinson College of Business
35 Broad Street, Suite 1206
Georgia State University Atlanta, GA 30303
jfed@gsu.edu

PERMISSION TO COPY

The *Journal of Financial Education*'s mission includes the dissemination of educational research and teaching aids that will promote/improve the educational process. In keeping with that mission, the *Journal of Financial Education* both encourages and permits copying of any article, in any quantity, for classroom use. The Journal only requests that the *Journal of Financial Education* be given appropriate citation. This blanket authorization to copy for classroom use does not include any articles reprinted from other journals, in which case, permission to reproduce must be obtained from the original journal of publication.

Copyright 2025 by the Journal of Financial Education
Richard Fendler and Milind Shrikhande, Editors
ISBN 00093-3961

The Journal Staff

The Journal of Financial Education is published by the Financial Education Association (FEA). The Executive Director of FEA is Timothy B. Michael (michael@uhcl.edu). For information concerning subscriptions to the journal, please see: http://jfedweb.org/subscrib.html. If you have any questions about the Journal of Financial Education or about the contents of this issue, please contact the journal co-editors (Richard Fendler and Milind Shrikhande) at jfed@gsu.edu.

Can Resilience Mitigate the Effects of Factors Associated with Finance Student Attrition?

Asli Eksi
Salisbury University

David J. Emerson
Salisbury University

Kenneth J. Smith
Salisbury University

This study examines whether resilience can serve as a coping mechanism to reduce the adverse effects of stress and burnout on student departures from the finance major. Using structural equations modeling analyses with survey data collected from 171 finance majors at four different U.S. universities, we examine the relations between role stressors, stress arousal, burnout, and departure intentions. We find that specific role stressors have direct positive associations with stress arousal and academic burnout, and burnout has a significant direct positive association with departure intentions. Moreover, specific role stressors and stress arousal have significant indirect positive associations with departure intentions through their associations with burnout. However, resilience counteracts these associations through its negative associations with stress arousal and burnout, confirming its mitigating effect. Our findings call for incorporating resilience training strategies into the finance curriculum to help students develop the soft skills needed to cope with the fast-paced, competitive work environments they are likely to face in the finance industry.

Key Words: resilience, role stress, academic burnout, departure intentions, finance major

Introduction

The finance profession is known for growth opportunities and lucrative salaries but also for fast-paced and competitive environments with long working hours and high stress levels. Not surprisingly, financial managers, financial analysts, and compliance officers are reported as having some of the most stressful jobs in the country (Williams, 2022). Because of this stress, finance professionals suffer from burnout and frequent turnover. According to a survey by LemonEdge.com, 31% of finance and banking professionals plan to leave their industry due to high pressure, while a further 31% are planning to stay within the industry but leave their current jobs. Moreover, 33% of finance and banking professionals state that their burnout level has increased due to changes in work environment after the COVID-19 pandemic (Pharande, 2022). The issue is especially severe for young professionals. Indeed, after the pandemic, many junior bankers have been unable to cope with the intense work environment created by difficult working conditions and increased uncertainty (Kessler & Hirsch, 2021). Despite significant salary

increases, big bonuses, and other benefits, young finance professionals have quit their jobs at alarming rates (Melin, 2021).

Resilience is a key coping mechanism needed by financial professionals to deal with stress, yet financial education places little or no emphasis on its development. Finance majors acquire multiple competencies as a part of their education, including a vast amount of technical knowledge, as well as effective communication and analytical thinking skills. While these competencies play a vital role in preparing them for their careers, they may lack the ability to withstand stressful working environments. In this paper, we address the importance of resilience as a coping mechanism to deal with stress and burnout at an early stage. Specifically, via structural equations modeling analyses, we evaluate whether resilience can reduce stress and academic burnout among finance students, and in turn decrease voluntary departure intentions from the major.

Occupational stress and burnout pose a serious problem for banking and finance professionals and are associated with high employee turnover (see, e.g., Giorgi et al., 2017 for a review). Abate et al. (2018) document significant relationships between job satisfaction, burnout, and turnover intentions among employees in the U.S. retail banking industry and conclude that unsatisfied employees with burnout are less likely to remain in their jobs. Other researchers (e.g., Belias & Koustelios, 2014; Gidou et al., 2020; Gupta et al., 2018) found similar results using data on banking professionals from different countries, illustrating the international dimension of this phenomenon. Furthermore, recent industry reports highlight an increase in voluntary turnover in financial services, especially among the new generation that is less likely to tolerate stress, and point to the large costs this generates for their employers. In a survey conducted by PricewaterhouseCoopers (2012), around a quarter of CEOs in financial services state that they had to cancel or delay a key strategic initiative over the previous 12 months because they were unable to retain their young talent.

The way young finance professionals handle stress in their careers mimics that in college. The better they can cope with difficult situations in college, the better they are likely to cope in their career. We hypothesize that finance majors with high levels of resilience will experience less stress arousal and academic burnout and will be less likely to leave their major. Our hypotheses stem from a broad array of literature documenting the positive effects of resilience on academic performance (e.g., Hartley, 2010; Ong et al., 2006; Shields, 2001). Moreover, Yeager and Dweck (2012) demonstrated that, rather than being immutable, resilience can be developed in classrooms, leading to academic and social success. We bring this research to the attention of finance educators to help reduce voluntary student departure intentions and prepare a new generation of resilient finance professionals.

Our analyses use psychometric data gathered via self-report surveys from a sample of finance majors at four geographically dispersed universities in the U.S. Utilizing survey data, we analyze the associations among resilience, role stressors, stress arousal, burnout, and intentions to depart the major. Specifically, we test an expanded role stress model with structural equation modeling procedures, with role stressor and resilience measures as antecedents, stress arousal and burnout as mediators, and departure intentions as the outcome variable. We found that resilience serves to reduce intentions to depart the major. Although role stressors are positively associated with stress arousal and academic burnout as well as departure intentions, we found that resilience counteracts these associations through its negative associations with stress arousal and burnout.

Our paper has important implications for finance educators. It draws attention to the high turnover among young finance professionals arising from stress and burnout and suggests that this issue should be addressed at the college level. It calls for incorporating resilience training into the

finance curriculum by showing how resilience can be used to cope with stress and burnout and retain finance majors. Highly resilient students remain focused and committed to completing their degrees. They then carry these habits into their professional lives allowing them to adapt to the fast-based and competitive working environment in the finance industry.

The rest of our paper is organized as follows. The next section provides a brief review of literature on role stress theory as it forms the foundation of our research. We then develop our hypotheses and present our theoretical model. Next, we describe our data and methodology, followed by our results. We conclude with a discussion of our findings for finance educators as well as for the finance industry.

Literature Review

Role stress refers to the pressure that individuals face within the scope of their role in an organization. Kahn et al. (1964) identified three components of role stress: role conflict, role ambiguity, and role overload. Role conflict occurs when a person is faced with simultaneous contradictory expectations, whereas role ambiguity arises from unclear performance expectations. Role overload represents one's perception of excessive work tasks and responsibilities.

Role stressors have been shown to have a negative impact on a variety of desirable job outcomes, and a positive association with turnover intentions. Moreover, research suggests that the effect of role stressors on departure intentions is mediated by two key variables: stress arousal and burnout (Fogarty et al., 2000; Smith et al., 2007). Girdano and Everly (1979) define stress arousal as the fairly predictable arousal of mind-body systems that can damage the system to the point of malfunction or disease. Burnout is a response to prolonged stress characterized by feelings of being emotionally drained and lacking resources.

In an academic context, burnout includes exhaustion, cynicism, and academic inefficacy (Maslach et al., 2016). Exhaustion refers to the feeling of being overburdened and fatigued by one's academic work, while cynicism means taking an indifferent or distant mindset to it. Finally, academic inefficacy is defined as a lack of contentment associated with one's accomplishments and expectations of continued ineffectiveness.

Smith et al. (2007) suggested that both stress arousal and burnout are responses to role stressors and serve as antecedents to various outcomes such as departure intentions. However, while stress arousal is an immediate response, burnout develops through chronic exposure to the same role stressors (LePine & LePine, 2005; Maslach & Schaufeli, 1993). Smith et al. (2007) posited that stress arousal may be directly associated with detrimental role outcomes before burnout manifests itself, but also has a direct positive association with burnout, thereby serving as a mediator between role stressors and burnout.

Previous studies suggest different mitigating factors to overcome the negative effect of role stress and burnout on role outcomes. Of note, Coutu et al. (2002) and Ong et al. (2006) argued that individual traits such as resilience can counteract the negative effect of stressors on job outcomes by increasing the threshold at which stress arousal is triggered, thereby decreasing burnout.

Resilience is defined as one's capacity to withstand adverse conditions and persevere under stress (Connor & Davidson, 2003). Robertson et al. (2015) posited that resilience protects individuals from the negative outcomes of occupational stress and advocate for resilience training in the workplace. Smith et al. (2020a) showed that resilience decreases stress arousal and burnout among auditors and reduces job dissatisfaction and turnover intentions as a consequence. Similarly, Smith et al. (2020b) found that individual resilience levels serve as a protective factor

for accounting students by enhancing psychological health and diminishing academic burnout resulting in reduced departure intentions from the major.

The demanding working conditions for new professionals, arising from uncertainty after COVID-19 and the recent developments in fintech have each increased job stress. These stressors may exacerbate the significant staffing challenges that the finance and banking industry already faces. In this paper, we test whether resilience serves as a coping mechanism to attenuate the adverse effect of stressors among finance majors. Examining these relationships in an academic context is important, as attrition in the finance major further amplifies these challenges by decreasing the supply of entry-level finance professionals.

Hypothesis Development

Figure 1 presents the theoretical model we propose and illustrates the hypothesized associations tested. The model is based on prior psychological and organizational behavior research. Given the scarce literature documenting the proposed relations among college students (e.g., Smith et al., 2020b; Emerson et al., 2022), let alone finance majors, we draw some of our hypotheses from the literature focusing on working professionals.

Prior research documents significant positive associations between role stressors and stress arousal. For example, Smith et al. (2007) found role ambiguity and role overload to have significant positive associations with stress arousal, while Smith and Emerson (2017) found role conflict and role overload to have significant positive associations with stress arousal. Based on this evidence, we propose the following hypotheses:

H1(a). *There is a positive association between role conflict and stress arousal.*
H1(b). *There is a positive association between role ambiguity and stress arousal.*
H1(c): *There is a positive association between role overload and stress arousal.*

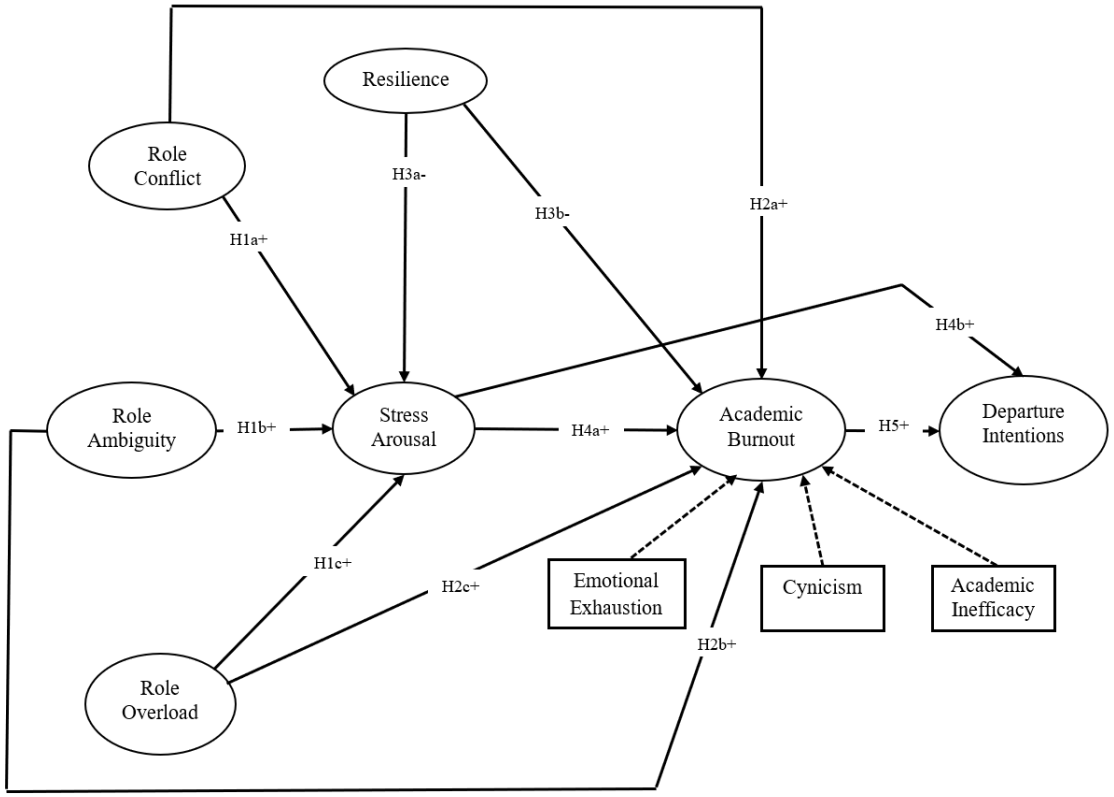
Although stress arousal is expected to mediate the associations between role stressors and burnout, we cannot ignore the potential for high levels of role stress to be directly associated with burnout. In fact, with a sample of auditors working in public accounting, Smith et al. (2020a) found all three role stressors to have significant positive associations with burnout after controlling for the influence of stress arousal. This possibility prompts the following hypotheses:

H2(a). *There is a positive association between role conflict and burnout.*
H2(b). *There is a positive association between role ambiguity and burnout.*
H2(c): *There is a positive association between role overload and burnout.*

Coutu (2002), Connor and Davidson (2003), and Ong et al. (2006) proposed that resilience is a coping mechanism that can mitigate the effect of role stressors by increasing the threshold at which stress arousal is triggered. Consequently, we propose:

H3(a). *There is a negative association between resilience and stress arousal.*

Figure 1
Theoretical Model with Hypothesized Paths



Zunz (1998) found that resilience-specific personality traits have negative associations with individual burnout dimensions and Caverly (2005) shows that resilient employees have lower burnout rates. Moreover, Strumpfer (2003) argued that burnout occurs when the cumulative effects of stress arousal overwhelm an individual's coping mechanisms, suggesting that the effect of resilience on burnout can be mediated by stress arousal. However, Smith and Emerson (2017) found a significant direct association between resilience and burnout even after controlling for the mediating effect of stress arousal. This leads us to posit:

H3(b). *There is a negative association between resilience and burnout*.

The above-mentioned ordering of stress arousal and burnout regarding how stressors affect individuals implies that specific role stressors may not be individually excessive, but their cumulative effect may trigger stress arousal and induce burnout (Fogarty et al., 2000). This suggests a link between stress arousal and burnout, leading to the following prediction:

H4(a). *There is a positive association between stress arousal and burnout.*

Previous role stress studies have found stress arousal has a direct negative association with turnover intentions (Smith et al., 2017). Thus, while burnout is expected to have a significant mediating influence on the associations between stress arousal and departure intentions, the potential for a significant direct association motivates us to explore the possibility that:

H4(b). *There is a positive association between stress arousal and turnover intentions.*

Fogarty et al. (2000) found that burned-out accountants tend to view their organization in adversarial terms and are more likely to depart from the organization. Likewise, prior research documents a positive relationship between burnout and departure intentions among various occupations (Smith et al., 2018). Similar results have been found among both accounting and business students, suggesting that if students are burned-out in their studies, they are more likely to view their program in adversarial terms and change their majors or quit school (Smith et al. 2020b, 2022). This prompts us to predict::

H5. *There is a positive association between burnout and departure intentions.*

Data and Methodology

Sample

Survey data was collected from a sample of 171 finance majors from four universities: one private university in the West (n=30), one public university on the East Coast (n=61), and two public universities in the South (n=52, and n=28). One hundred sixty-nine of the survey instruments had complete data for each of the primary study constructs and were included in the analyses.

All four schools in this study have accredited AACSB business programs. These schools are diverse in terms of size, student composition, geographic location, mission, and funding source. The human subjects committee at each school approved the instrument package that was administered in class at concurrent points during each school's academic term. The instrument package provided assurances of anonymity and an opt-out option. Scale ordering was varied to mitigate the concerns for common method bias.

Males comprised over three quarters of the participants reporting gender (128/167). Of those reporting academic level, 82 percent (140/171) indicated that they were sophomores or juniors. Ninety seven percent (166/171) reported that they were between 18 and 24 years old.

Measures

The following measures are used in this study:

Role Ambiguity: three items from Rizzo, House, and Lirtzman's (1970) 14-item Role Conflict and Role Ambiguity Scale;
Role Conflict: three items from Rizzo et al.'s (1970) scale;
Role Overload: four items from the Beehr, Welsh, and Tabor (1976) scale;
Resilience: the 10-item Connor-Davidson Resilience Scale-10 (CD-RISC 10, Campbell-Sills and Stein 2007);
Stress Arousal: the four-item Stress Arousal Scale-4 (SAS4, Smith et al. 2012);
Academic Burnout: the 16-item Maslach Burnout Inventory: General Survey for Students (MBI-GS(S), Maslach et al. 2016), which includes Emotional Exhaustion, Cynicism, and Academic Inefficacy;

Departure Intentions: three items from the Donnelly and Ivancevich (1975) scale adjusted to an academic context.

These scales have been used extensively in prior research. The items for each construct have demonstrated favorable convergent and discriminant validity as well as reliability in both academic and professional settings. Item wording was modified where appropriate to reflect the academic setting. Except for the SAS4, the measures are five-point Likert scales; the SAS 4 is a four-point Likert scale. The full list of survey items appears in Table 2 below.

In various organizational settings, demographic factors such as age, gender, educational level, etc., have been associated with specific behavioral outcomes (Donnelly et al., 2003; Herda & Martin, 2016). To guard against the possibility that similar factors might confound the tested relationships between role stressors, resilience, stress arousal, burnout, and departure intentions in this academic setting, we include four key demographic factors as controls in our analyses. These control factors appear as Constructs 1-4 in Table 1.

Table 1
Inter-Scale Correlations

Construct	1	2	3	4	5	6	7	8	9	10	11
1. School[1]	-										
2. Gender[2]	-.010	-									
3. Age[3]	*-.335*	.087	-								
4. Academic Level[4]	*.527*	.030	*.594*	-							
5. Resilience	.097	.163	-.038	.017	-						
6. Role Conflict	.016	.010	-.067	.028	-.083	-					
7. Role Ambiguity	.071	-.069	-.036	-.113	.275	.287	-				
8. Role Overload	.009	-.274	.043	.009	-.126	*.472*	-.079	-			
9. Stress Arousal	-.073	-.151	.050	.030	-.273	.025	-.024	.162	-		
10. Burnout	.085	-.185	.012	-.005	*-.143*	*.407*	-.124	*.650*	*.305*	-	
11. Departure Intentions	-.025	-.140	.072	.071	*-.115*	.277	-.206	.287	.168	*.423*	-

Bold correlations significant at $p < .05$; *bold and italicized* at $p < .01$.

[1] Coded '1' if Eastern school ($n = 61$), '2' if Western School ($n = 30$), '3' if Southern School #1 ($n = 52$), and '4' if Southern School #2 ($n = 28$)
[2] Coded '1' if female ($n = 39$), and '2' if male ($n = 128$)
[3] Coded '1' if under 19 ($n = 15$), '2' if 19-21 ($n = 95$), '3' if 22-24 ($n = 56$), and '4' if over 24 ($n = 5$)
[4] Coded '1' if freshman ($n = 31$), '2' if sophomore ($n = 55$), '3' if junior ($n = 85$)

A critical concern was that primary study construct scores might be moderated by school. However, a series of analyses of variance revealed that there were no significant differences in any of the primary study constructs by school. Therefore, we combined the data from all four schools in the subsequent analyses.

Methodology

Ali et al. (2016) noted that common method variance must be examined when data are collected via self-reported questionnaires, particularly when the same person provides data for both the independent and dependent variables (Podsakoff et al., 2003). To address this issue, we conducted Harman's (1976) single-factor test to evaluate whether a single factor accounts for most of the

covariance in the model. We rejected this possibility as several explanatory factors were identified following analysis of the data. In addition, discriminant validity tests of the measurement model provide additional assurance that common method bias does not appear to be a significant issue.

To assess the hypothesized model, we conducted Partial Least Squares Structural Equation Modeling (PLS-SEM) using SmartPLS 4 (Ringle et al., 2024). Following Anderson and Gerbing's (1988) admonition that measurement model assessment must precede the evaluation of structural linkages, we assessed the validity and reliability of the measures and then proceeded to test the structural model. We incorporated a bootstrapping method with 5,000 resamples to assess the significance of the path coefficients and factor loadings (Hair et al., 2013).

Results

Measurement Model Assessment

We used Becker et al.'s (2012) two-stage approach for estimating the measurement and structural models that incorporate higher order formative constructs. In Stage 1, each of the lower-order components of academic burnout, (i.e., emotional exhaustion, cynicism, and academic inefficacy) are associated with the hypothesized antecedents (i.e., resilience, role stressors, and stress arousal) and key outcome (i.e., departure intentions) to assess internal consistency, convergent validity, and discriminant validity among the constructs.

Table 2 presents the results of the Stage 1 convergent reliability and validity assessment of the measurement model. It should be noted that the following scale items were excluded due to low loadings on their respective constructs: Resilience - I try to see the humorous side of things when I am faced with problems; Academic Inefficacy - I feel exhilarated when I accomplish something at the university (reverse-scored item); Cynicism - I just want to get my work done and not be bothered; and Departure Intentions - It is likely that I will actively look for a new major or University next year.

The data supports all three reported internal consistency measures, i.e., Cronbach alpha (CA), rho-A, and Composite Reliability (CR), with values ranging from 0.761 to 0.956, effectively falling within the 0.70 to 0.90 range that Hair et al. (2019) prescribe as satisfactory to good. All of the individual item loadings exceed Chin et al.'s (2008) suggested minimum value of 0.60. The Average Variance Extracted (AVE) for each construct exceed Fornell and Larcker's (1981) prescribed minimum of 0.50, thus establishing adequate convergent validity.

Table 2
Convergent Reliability and Validity Assessment – Stage 1

Construct	Scale Items	Loadings	CA	rho_A	CR	AVE
Resilience	I have been able to adapt when changes occur.	0.699	0.902	0.910	0.919	0.559
	I can deal with whatever comes my way.	0.773				
	Having to cope with stress can make me stronger.	0.654				
	I tend to bounce back after illness, injury, or other hardships.	0.750				
	I believe I can achieve goals, even if there are obstacles.	0.776				
	Under pressure, I stay focused and think clearly.	0.750				
	I am not easily discouraged by failure.	0.788				
	I think of myself as a strong person when dealing with life's challenges and difficulties.	0.809				
	I am able to handle unpleasant or painful feelings like sadness, fear, or anger	0.719				
Role Conflict	I receive homework assignments without the resources to complete them.	0.802	0.811	0.821	0.888	0.725
	I receive incompatible requests from two or more people.	0.873				
	Overall, I often receive conflicting directions.	0.879				
Role Ambiguity (Reverse-scored)	Clear, planned goals/objectives exist for my coursework.	0.864	0.761	0.803	0.859	0.671
	I know how my performance is going to be evaluated.	0.800				
	I know exactly what is expected of me.	0.792				
Role Overload	I am responsible for an almost unmanageable number of concurrent assignments.	0.788	0.866	0.875	0.909	0.713
	I simply have more work to do than can be done in an ordinary day.	0.889				
	I feel that I just don't have time to take an occasional break.	0.826				
	Overall, I have too much work to do.	0.872				
Stress Arousal	Anticipating or remembering upsetting things?	0.786	0.833	0.849	0.887	0.663
	Thinking about things which upset you?	0.864				
	Concerned or worried?	0.791				
	Repeating unpleasant thoughts?	0.814				
Emotional Exhaustion	I feel emotionally drained by my studies.	0.852	0.866	0.875	0.904	0.654
	I feel used up at the end of the day at the university.	0.808				
	I feel tired when I get up in the morning and have to face another day at the university.	0.838				
	Attending classes all day is really a strain for me.	0.687				
	I feel burned out from my studies.	0.846				

Cynicism	I have become less interested in my studies since my enrollment.	0.801	0.829	0.830	0.886	0.660
	I have become less enthusiastic about my studies.	0.843				
	I have become more cynical about whether my university work contributes anything.	0.776				
	I doubt the significance of my studies.	0.843				
Academic Inefficacy (Reverse-scored)	I can effectively solve the problems that arise in my studies.	0.801	0.804	0.818	0.863	0.559
	I feel I am making an effective contribution in my classes.	0.718				
	In my opinion, I am a good student.	0.714				
	I have accomplished many worthwhile things in my studies.	0.721				
	While at the university, I feel confident that I am effective at getting things done	0.780				
Departure Intentions	I often think about quitting.	0.955	0.909	0.911	0.956	0.917
	I often think about dropping out of the University.	0.960				

We also assessed whether the lower order constructs were empirically distinct from one another using Henseler et al.'s (2015) recommended heterotrait-monotrait ratio of correlations. The results, available for review by contacting the corresponding author[1], support the discriminant validity among the constructs.

We next conducted Kock's (2015) multicollinearity diagnostic procedure to further assess whether our model is contaminated by common method bias. Kock (2015) specified that variance inflation factor (VIF) values generated by this procedure of 3.3 or lower indicate that the model is free of common method bias. None of the VIFs in our model exceeded 1.449, thus common method bias did not appear to be a concern with our data.

As specified by Sarstedt et al. (2019), we assessed Stage 2 by running a redundancy analysis of the academic burnout measurement model. In the Stage 2 model, academic burnout is measured with three formatively measured indicators, i.e., the latent variable scores for Emotional Exhaustion, Cynicism, and Academic Inefficacy constructs analyzed in Stage 1. The redundancy analysis related the higher-order academic burnout construct to an alternative global single-item measure of academic burnout. This analysis generated a path coefficient of 0.798, above Sarstedt et al.'s (2019) suggested minimum threshold of 0.700. Next, we found that collinearity did not negatively affect the model given that our assessment generated VIF values no greater than 1.365 for the indicators of the higher-order academic burnout construct. Finally, the weights of all three indictors had a significant effect ($p < .01$) on academic burnout supporting the latter's conceptualization as a reflective-formative second-order factor.

Structural Model Evaluation

We next evaluated the structural model. Following Hair et al. (2019), we evaluated the structural model by examining the structural coefficients (i.e., betas) and their corresponding t-values calculated using a bootstrapping procedure with 5,000 resamples. Figure 2 presents the results of the hypothesis tests and structural model evaluation. Role overload had a significant positive association with stress arousal and academic burnout, supporting H1c and H2c. Role conflict and role ambiguity also had significant positive associations with academic burnout, supporting H2a and H2b. Resilience had a significant negative association with stress arousal and academic burnout, thus supporting H3a and H3b. In turn, H4a and H5 were supported as stress arousal had a significant positive association with academic burnout, which in turn had a significant positive association with departure intentions. However, no significant associations were measured between either role conflict or role ambiguity and stress arousal, or between stress arousal and departure intentions, thus failing to support H1a, H1b, or H4b.

The f2 effect size metric indicates how the removal of a particular predictor construct affects an endogenous construct's R2 value (Hair et al., 2019). According to Cohen (1988), f2 values of 0.02, 0.15, and 0.35 indicate small, medium, and large effect sizes. Figure 2 indicates that there are six small and two medium effect sizes among the eight significant paths.

The illustrated Q2 statistics for each endogenous construct (i.e., stress arousal, burnout, and departure intentions) in Figure 2 all exceed Hair et al.'s (2019) prescribed lower threshold of zero for supporting the predictive accuracy of the model for each construct. The reported R2 values indicate that resilience and role overload explain 7 percent of the variance in stress arousal, the role stressors, resilience, and stress arousal explain 49.3 percent of the variance in academic burnout, and academic burnout explains 19.6 percent of the variance in departure intentions.

[1] Kenneth Smith (kjsmith@salisbury.edu)

As noted above, we found a significant correlation between gender and role overload. Consequently, we tested whether gender moderated role overload's significant associations with stress arousal and burnout. Though not reported in tabular form, neither moderating effect is significant, thus providing reasonable assurance that gender did not confound the tested relationships between role overload and either outcome measure.

As Figure 2 illustrates, neither resilience nor stress arousal have significant direct associations with departure intentions. To gain a deeper understanding of the effects of stress arousal and resilience on academic burnout and departure intentions, we performed a path-analytic decomposition of the direct and indirect effects of resilience and stress arousal on each outcome (Sarstedt et al., 2020). As Table 3 reports, in addition to its direct association with academic burnout, resilience has a significant indirect negative association ($\beta = -.056$; $p < .05$) through its direct negative association with stress arousal. Moreover, resilience has a significant indirect negative association with departure intentions ($\beta = -.102$; $p < .05$) through its direct associations with stress arousal and academic burnout and indirect association with academic burnout. Stress arousal also has a significant indirect association with departure intentions ($\beta = .099$; $p < .05$) via its direct positive association with each academic burnout. Also noteworthy, though not presented in tabular form, role overload has a significant positive indirect effect ($\beta = .234$; $p < .01$) on departure intentions via several pathways illustrated in Figure 2.

To evaluate final structural model's out-of-sample predictive power, we applied Shmueli et al.'s (2019) suggested procedure using the PLSpredict option in SmartPLS. According to Shmueli et al. (2019), the PLS-SEM model has high predictive power if all its indicators have lower predictive errors (i.e., lower RMSE and MAE values, and higher Q2predict values) when compared to the linear benchmark model. Online Appendix B results indicate that the PLS-SEM model has lower RMSE and MAE statistics values, and higher Q2predict values than the benchmark model for both departure intentions indicators, thus indicating high predictive power for the model.

Discussion and Implications

This study examines the potential mitigating role of resilience on the associations between role stressors, stress arousal, academic burnout, and departure intentions of finance majors. In line with theory, our findings indicate that role overload as mediated by stress arousal and academic burnout had a significant positive association with departure intentions. However, higher levels of resilience appear to mitigate this association. Through its negative influence on stress arousal and academic burnout, resilience indirectly attenuates the positive associations between stress arousal and burnout, and burnout and departure intentions. Moreover, our findings support the role of burnout in this model as it fully mediates the influence of its antecedents (i.e., role stressors and stress arousal) on departure intentions.

Figure 2
Final Structural Model
(Only significant paths illustrated for ease of diagramming and interpretability.)
Bold path significant at p < .05; **bold and italicized** at p < .01.
* f-square significant at p < .05; ** at p < .01.
[1] = indicator weight

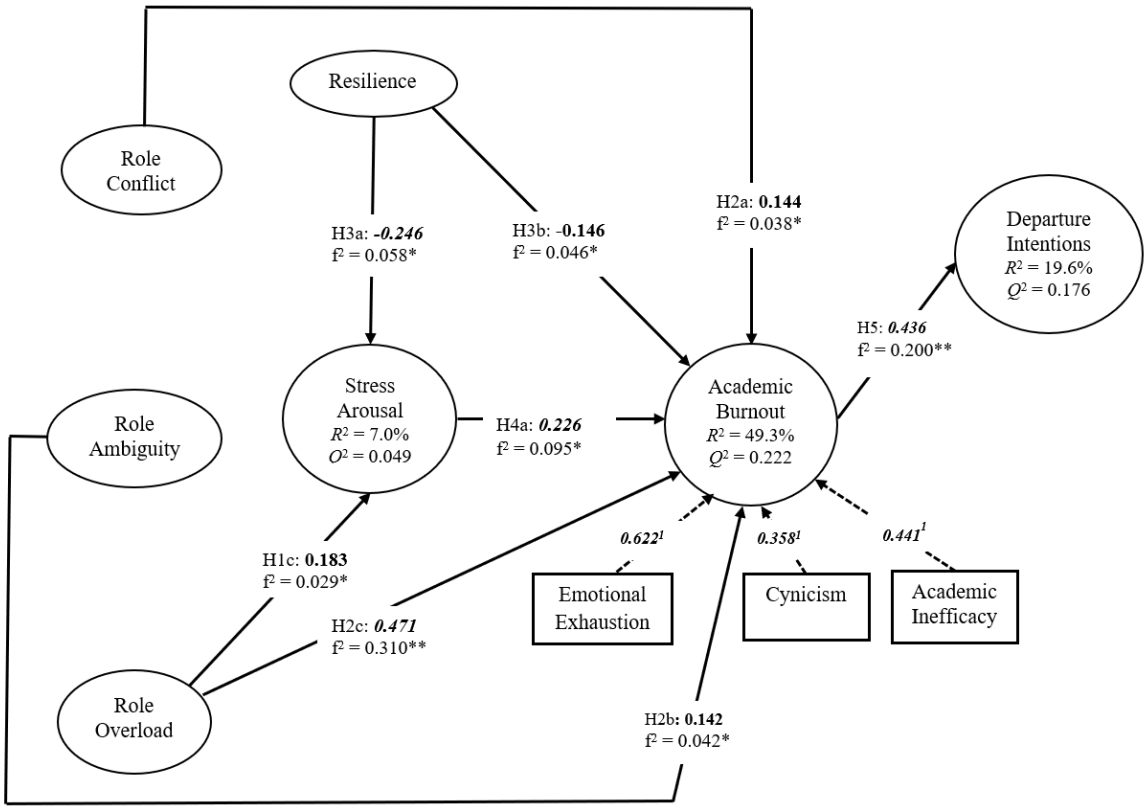


Table 3
Direct and Indirect Effects of Resilience and Stress Arousal on Academic Burnout and Departure Intentions

Predictor Construct	Predicted Construct					
	Academic Burnout			Departure Intentions		
	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect	Total Effect
Resilience	*-.146*	-.056[1]	*-.202*	NA	-.102[2]	-.102
Stress Arousal	*.226*	NA	*.226*	NA	.099[3]	.099

Bold paths significant at $p < .05$; ***bold and italicized*** at $p < .01$.
[1]Product of Resilience → Stress Arousal (-.246) x Stress Arousal → Academic Burnout (.226).
[2]Sum of the following indirect effects: Resilience → Stress Arousal → Departure Intentions (-.014); Resilience → Academic Burnout → Departure Intentions (-0.064); Resilience → Stress Arousal → Academic Burnout → Departure Intentions (-.024)
[3]Product of Stress Arousal → Academic Burnout (.226) x Academic Burnout → Departure Intentions (.436).

Our finding that neither role conflict nor role ambiguity had a significant association with stress arousal may have been due to the significant correlations among role stressors (Schaubroeck et al., 1989). Role conflict and role ambiguity correlate at .287 ($p < .001$), and role conflict and role overload correlate at .472 ($p < .001$). These correlations may have attenuated the associations between both role conflict and role ambiguity and stress arousal. Another explanation for these non-significant paths may lie in the low mean score of 2.424 for role conflict and 2.326 for role ambiguity, compared to role overload's mean score of 2.927. Indeed, there may simply be no corollary among students for conflicting demands and ambiguous expectations reported in the workplace. In contrast, the expanding skillset required of finance majors has undisputedly increased their workload, thus supporting the strength of role overload as a predictor of students' stress, burnout, and departure intentions.

Given that resilience seems to attenuate the adverse impact of stress and burnout on departure intentions from the major, it would appear propitious to incorporate resilience training into the finance curriculum. As Masten (2001) noted, resilience is a dynamic individual characteristic that can be learned and developed. Therefore, providing targeted resilience training can serve as an efficient and cost-effective intervention to alleviate the negative effects of stress among finance students and help them improve their stress coping skills before entering the workforce.

Fortunately, there are several options available for educators who seek to enhance the resilience levels of their students. For example, they can promote an adaptive stress mindset and self-connection as means of fostering resilience and train the students to be more aware of why they chose the finance major and what they expect from it so that students learn how to act in a manner consistent with their motivations and aspirations (Klussman et al., 2021). Moreover, finance educators can utilize different resilience training programs available such as Transforming Lives Through Resilience Education from the University of Texas at Austin, and Penn Resilience Program, and Perma Workshops from the University of Pennsylvania.

Previous literature documents the potential benefits of resilience training programs in academic settings. For example, Steinhard and Dolbier (2008) showed that a group of students who received training had more effective coping strategies compared to a control group. Similarly, Oman et al. (2008) found that students who enrolled in a meditation-based stress management program

recorded significantly lower stress levels. Maddi et al. (2009, 2012) reported higher hardiness levels and GPAs among a large group of undergraduate students who received hardiness training. DeRossier et al. (2013) found that resilience education was associated with enhanced stress coping skills and better adaptation to college life. The evidence from these studies suggests that the key to retaining and successfully educating finance students may rest on efforts to reduce stress and academic burnout through resilience training. Furthermore, our findings indicate that time management should be part of such training to address the students' workload concerns.

Limitations and Conclusion

This study, like others that use cross-sectional designs with self-reported data, has certain limitations. First, the self-report measures might subject the tested associations to the influence of common methods variance. However, each of this study's instruments have demonstrated validity and reliability in prior research. Moreover, our measurement model analyses confirm the validity of our constructs and show that common methods variance fails to explain the interrelationships among them.

Second, the cross-sectional design of this study prompts concern over trajectory bias, i.e., that the measures capture states, not traits. This raises the question as to whether the data would support alternative construct orderings, such an iterative feedback loop between stress arousal and burnout. Even though our proposed theoretical model is well-grounded based on previous research, our cross-sectional data still precludes us from drawing a definitive conclusion about the temporal precedence of one construct over the other.

Despite these potential limitations, this study makes a noteworthy contribution to our understanding of whether and how resilience serves as a coping strategy that counteracts the influence of role stressors on academic burnout and finance majors' intentions to discontinue their studies. Through its direct negative associations with stress arousal and academic burnout, and its indirect negative association with academic burnout, resilience has a significant indirect negative connection with student intentions to leave the major. Being the only effort thus far to simultaneously examine and uncover these associations among finance majors, it represents a critical first step toward mitigating excessive finance major turnover through planning and implementing resilience training programs.

Resilience training programs can especially contribute to the success of students who possess the required intellectual and technical skills for a finance major but are deficient in the soft skills to handle stress that they will inevitably face in the workplace. Moreover, they enhance finance programs' prospects for retaining majors and placing them in prestigious positions upon graduation. In the current era of business school student retention problems, university budget constraints, the projected demographic enrollment cliff, and increasing program accountability, these programs can offer significant benefits to students, faculty, program administrators, and future employers.

References

Abate, J., Schaefer, T., & Pavone, T. (2018). Understanding generational identity, job burnout, job satisfaction, job tenure and turnover intention. *Journal of Organizational Culture, Communications and Conflict, 22*(1), 1-12.

Ali, F., Kim, W. G., & Ryu, K. (2016). The effect of physical environment on passenger delight and satisfaction: Moderating effect of national identity. *Tourism Management, 57*, 213-224.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411.

Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods, 8*(3), 274–289.

Becker, J. M., Klein, K., & Wetzels, M. (2012). Hierarchical latent variable models in PLS-SEM: Guidelines for using reflective-formative type models. *Long Range Planning*, *45*, 359–394

Beehr, T.A., Welsh, J. T., & Tabor, T.D. (1976). Relationship of stress to individually and organizationally valued states: Higher order needs as a moderator. *Journal of Applied Psychology*, *61*(1), 41–47.

Belias, D., & Koustelios, A. (2014). Job satisfaction and job burnout among Greek bank employees. *International Journal of Management*, *5*(1), 33-45.

Boehm, S. A., Dwertmann, D. J., Bruch, H., & Shamir, B. (2015). The missing link?

Investigating organizational identity strength and transformational leadership climate as mechanisms that connect CEO charisma with firm performance. *The Leadership Quarterly, 26*(2), 156–171.

Campbell-Sills, L., & Stein, M. B. (2007). Psychometric analysis and refinement of the Connor-Davidson resilience scale (CD-RISC): Validation of a 10-item measure of resilience. *Journal of Traumatic Stress*, *20*(6), 1019–1028.

Caverly, N. M. (2005). Mapping out occupational resiliency and coping in a public service work setting. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *66*, 1648.

Chin, W. W., Peterson, R. A., & Brown, S. P. (2008). Structural equation modeling in marketing: Some practical reminders. *Journal of Marketing Theory and Practice*, *16*(4), 287–298

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates, Inc.

Connor, K. M., & Davidson, J. R. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and Anxiety*, *18*(2), 76–82.

Coutu, D.L. (2002). How resilience works. *Harvard Business Review*, *80*, 46-51.

DeRosier, M. E., Frank, E., Schwartz, V., & Leary, K. A. (2013). The potential role of resilience education for preventing mental health problems for college students. *Psychiatric Annals*, *43*(12), 538–544.

Donnelly Jr, J. H., & Ivancevich, J. M. (1975). Role clarity and the salesman: An empirical study reveals that perceived role clarity may be an important factor in maximizing a salesman's job performance. *Journal of Marketing*, *39*(1), 71-74.

Donnelly, D. P., Quirin, J. J., & Bryan, D. (2003). Attitudes toward dysfunctional audit behavior: The effects of locus of control, organizational commitment, and position. *Journal of Applied Business Research, 19*(1), 95–108.

Emerson, D. J., Hair, J. F., & Smith, K. J., (2023). Psychological distress, burnout, and business student turnover: The role of resilience as a coping mechanism. *Research in Higher Education*, *64*(2), 1-32.

Fogarty, T. J., Singh, J., Rhoads, G. K., & Moore, R. K. (2000). Antecedents and consequences of burnout in accounting: Beyond the role stress model. *Behavioral Research in Accounting*, *12*, 31–67.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equations models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50.

Gidou, M., Paltayian, G. and Gavalas, D. (2020). Job burnout and job satisfaction in the banking sector: The perceptions of bank employees in Greece. *Review of Economics and Finance*, *18*, 13-8.

Giorgi, G., Arcangeli, G., Perminiene, M., Lorini, C., Ariza-Montes, A., Fiz-Perez, J., Di Fabio, A. and Mucci, N. (2017). Work-related stress in the banking sector: A review of incidence, correlated factors, and major consequences. *Frontiers in Psychology*, *8*, 2166.

Girdano, D. A. & Everly, G. S. (1979), *Controlling Stress and Tension: A Holistic Approach*, Prentice-Hall.

Gupta, B., Mittal, S., & Mittal, V. (2018). Analyzing the impact of job burnout on job satisfaction: A study on Indian bank employees. *Special Issue for International Youth Symposium*, *10*, 19-30.

Hair, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2013). *A Primer on Partial Least Squares Structural Equation Modelling (PLS-SEM)*. Los Angeles: Sage Publications.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2-14.

Harman, H. H. (1976). *Modern Factor Analysis.* University of Chicago Press.

Hartley, M. T. (2010). Increasing resilience: Strategies for reducing dropout rates for college students with psychiatric disabilities. *American Journal of Psychiatric Rehabilitation*, *13*(4), 295–315.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135.

Herda, D. N., & Martin, K. A. (2016). The effects of auditor experience and professional commitment on acceptance of underreporting time: A moderated mediation analysis. *Current Issues in Auditing, 10*(2), A14–A27.

Kahn, R. L., Wolfe, D. M., Quinn, R. P. & Snoek, J. D. (1964). *Organizational Stress: Studies in Role Conflict and Ambiguity*, John Wiley & Sons, Inc.

Kessler, S., & Hirsch, L. (2021, March 27). Wall street's sleepless nights. Retrieved January 30, 2023, from https://www.nytimes.com/2021/03/27/business/dealbook/banker-burnout.html.

Klussman, K., Lindeman, M. I. H., Nichols, A. L., & Langer, J. (2021). Fostering stress resilience among business students: The role of stress mindset and self-connection. *Psychological Reports, 124*(4), 1462–1480.

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration*, *11*(4), 1–10.

LePine, J., Podsakoff, N. & LePine, M. (2005). A meta-analytic test of the challenge stressor-hindrance stressor framework: An explanation for inconsistent relationships among stressors and performance. *Academy of Management Journal*, *48*(5), 764-775.

Maddi, S. R., Harvey, R. H., Khoshaba, D. M., Fazel, M., & Resurreccion, N. (2009). Hardiness training facilitates performance in college. *The Journal of Positive Psychology*, *4*(6), 566–577.

Maddi, S. R., Harvey, R. H., Khoshaba, D. M., Fazel, M., & Resurreccion, N. (2012). The relationship of hardiness and some other relevant variables to college performance. *Journal of Humanistic Psychology*, *52*(2), 190–205.

Maslach, C. & Schaufeli, W. (1993), *History and Conceptual Specificity of Burnout, Recent Developments in Theory and Research*, Hemisphere.

Maslach, C., Jackson S. E., & Schwab R.L. 2016. *MBI-general Survey (students) MBI-GS (S)*. Mind Garden.

Masten A. S. (2001). Ordinary magic. Resilience processes in development. *The American Psychologist*, *56*(3), 227–238.

Melin, A. (2021, July 16). Wall street: $100,000 pay, bonus, perks aren't enough for young bankers. Retrieved January 30, 2023, from https://www.bloomberg.com/news/articles/2021-07-16/wall-street-100-000-pay-bonus-perks-aren-t-enough-for-young-bankers?leadSource=uverify+wall.

Oman, D., Shapiro, S. L., Thoresen, C. E., Plante, T. G., & Flinders, T. (2008). Meditation lowers stress and supports forgiveness among college students: A randomized controlled trial. *Journal of American College Health*, *56*(5), 569–578.

Ong, A. D., Bergeman, C. S., Bisconti, T. L., & Wallace, K. A. (2006). Psychological resilience, positive emotions, and successful adaptation to stress in later life. *Journal of Personality and Social Psychology*, *91*(4), 730.

Pharande, P. (2022, May 24). Burnout mounts as a third of banking and financial services plan to leave the industry due to high pressure. Retrieved January 30, 2023, from https://www.lemonedge.com/news-and-press-releases/burnout-mounts-as-a-third-of-banking-and-financial-services-plan-to-leave-the-industry-due-to-high-pressure.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903.

PWC. (2012, April 20) Millennials at work reshaping the workplace in financial services. Retrieved January 30, 2023, from https://www.pwc.com/gx/en/financial-services/publications/assets/pwc-millenials-at-work.pdf

Ringle, C. M., Wende, S., & Becker, J.-M. (2024). *SmartPLS 4*. Bönningstedt: SmartPLS. Retrieved from https://www.smartpls.com.

Rizzo, J. R., House, R. J., & Lirtzman, S. I. (1970). Role conflict and role ambiguity in complex organizations. *Administrative Science Quarterly*, *15*, 150–163.

Robertson, I. T., Cooper, C. L., Sarkar, M. and Curran, T. (2015). Resilience training in the workplace from 2003 to 2014: A systematic review. *Journal of Occupational and Organizational Psychology*, 88(3), 533-562.

Sarstedt, M., Hair Jr, J. F., Cheah, J. H., Becker, J. M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian Marketing Journal (AMJ)*, *27*(3), 197-211.

Sarstedt, M., Hair, J. F., Nitzl, C., Ringle, C. M., & Howard, M. C. (2020). Beyond a tandem analysis of SEM and PROCESS: Use PLS-SEM for mediation analyses! *International Journal of Market Research*, *62*(3), 288–299.

Schaubroeck, J., Cotton, J. L., & Jennings, K. R. (1989). Antecedents and consequences of role stress: A covariance structure analysis. *Journal of Organizational Behavior*, *10*(1), 35–58.
Shields, N. (2001). Stress, active coping, and academic performance among persisting and nonpersisting college students. *Journal of Applied Biobehavioral Research*, *6*(2), 65–81.
Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J. H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing, 53*(11), 2322–2347.
Smith, K. J., Davy, J. A., & Everly, G. S. (2007). An assessment of the contribution of stress arousal to the beyond the role stress model. *Advances in Accounting Behavioral Research*, *10*, 127–158.
Smith, K. J., Everly, G. S., & Haight, G. T. (2012). SAS4: Validation of a four-item measure of worry and rumination. *Advances in Accounting Behavioral Research*, *15*, 101-131.
Smith, K. J., & Emerson, D. J. (2017). An analysis of the relation between resilience and reduced audit quality within the role stress paradigm. *Advances in Accounting*, *37*, 1–14.
Smith, K. J., Emerson, D. J., & Boster, C. R. (2018). An examination of reduced audit quality practices within the beyond the role stress model. *Managerial Auditing Journal*, *33*(8/9), 736-759.
Smith, K. J., Emerson, D. J., Boster, C. R., & Everly Jr, G. S. (2020a). Resilience as a coping strategy for reducing auditor turnover intentions. *Accounting Research Journal*, *33*(3), 483-498.
Smith, K. J., Haight, T. D., Emerson, D. J., Mauldin, S., & Wood, B. G. (2020b). Resilience as a coping strategy for reducing departure intentions of accounting students. *Accounting Education*, *29*(1), 77-108.
Steinhardt, M., & Dolbier, C. (2008). Evaluation of a resilience intervention to enhance coping strategies and protective factors and decrease symptomatology. *Journal of American College Health*, *56*(4), 445–453.
Strumpfer, D. J. W. (2003). Resilience and burnout: A stitch that could save nine. *South African Journal of Psychology*, *33*(2), 69-79.
Williams, G. (2022). The 25 most stressful jobs. *US News & World Report*. Retrieved January 30, 2023, from https://money.usnews.com/careers/company-culture/slideshows/the-most-stressful-jobs.
Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*(4), 302-314.
Zunz, S.J. (1998). Resiliency and burnout: Protective factors for human service managers. *Administration in Social Work*, *22*(3), 39-54.

Online Synchronous vs Online Asynchronous modalities: Student Performance in a Core Finance Course

Satish Nargundkar
Georgia State University

Milind Shrikhande
Georgia State University

We compare student performance in two sections of an undergraduate core Finance course, one offered in the online asynchronous modality, and the other in the online synchronous modality. Based on factors identified in the literature such as flexibility, social interaction, online tools available, and initial student motivation, the courses were designed to make best use of each online modality. We find no significant difference in student performance on individual exams between the two online modalities. However, when it came to student performance on team projects, the synchronous online modality was more effective than the asynchronous online modality. We discuss possible reasons for the observed results based on the need for interaction and levels of learning desired.

Key words: online synchronous modality; online asynchronous modality; course design; student performance; undergraduate Finance course

Introduction

In our previously published study (Nargundkar & Shrikhande, 2023) we compared the face-to-face modality with the synchronous online modality for a core course (FI 4000 – *Valuation of Financial Assets*) for Finance majors and the face-to-face modality with the asynchronous online modality for an elective course (FI 4040 – *International Finance*) for undergraduate students.

One limitation of that study was that while the two online modalities were compared with the face-to-face modality, there was no direct comparison between the two online modalities – synchronous and asynchronous. At that time, we did not have the same instructor teaching the same course in both online modalities. Since then, however, the core course which was taught in the synchronous mode in the past, was also taught in both synchronous and asynchronous modalities in the same semester, permitting a direct comparison of the two online modalities with each other, while controlling for the potentially confounding effects of the instructor, course content, and timing.

The findings on student performance, from such a direct comparison between two online modalities in the core course, are presented here as an addendum to the earlier paper to make the study more complete. A similar comparison could be made for the elective course as well, but so far, the instructor has not had the opportunity to teach that course in both online modalities.

Literature Review

An early study on online teaching and learning by Brown & Liedholm (2002) compared student performance at a public research university in a Principles of Microeconomics course for

undergraduate students across three modalities - Live (face-to-face), Hybrid (two-thirds face-to-face and one-third online asynchronous), and Virtual (Online Asynchronous). Their findings were in line with the expectation at the time that the traditional face-to-face modality would result in the best student performance, followed by the hybrid modality, with the fully online modality being the least effective in delivering student performance. Given the state of the technology for online teaching over twenty years ago, this was not a surprising result. Further, they found that female students performed worse than the male students in the face-to-face modality, but there was no significant difference between the genders in the hybrid or online asynchronous modalities. It is possible the female students found the face-to-face setting more inhibiting than the hybrid or online setting.

Zheng and Luo (2023) conducted a meta-analysis of the literature spanning papers from 2002 to 2022, comparing the effectiveness of the various teaching modalities, and found that most studies compared one of the online modalities with the traditional face-to-face modality. Few studies directly compared the online synchronous with the online asynchronous modality – they found only 13 studies that directly compared the two online modalities with each other. Of these, only five studies were in higher education – two in medicine, and one each in computer science, economics, and management. Their analysis showed that overall, students performed slightly better in asynchronous online classes compared to the synchronous classes. The effect size was generally small, and showed no difference across education levels. Also, they found trends that suggested that quantitative courses were more suitable for asynchronous learning.

Some other meta-analyses (Ebner & Gegenfurtner, 2019; Gegenfurtner & Ebner, 2019) found the opposite result, that synchronous courses were slightly more effective than asynchronous ones. Zheng and Luo (2023) point out that this could be due to the difference in the criteria for selection of studies in the sample. Zheng and Luo (2023) only looked at studies that separated the two modalities clearly, while the other meta-analyses included some papers that may have compared a mixture of modalities in some cases.

Le (2022) compared pre-recorded lectures in economics (asynchronous) with live online lectures (synchronous) and found that students with lower ability (lower 50^{th} percentile) based on prior GPA performed worse in the asynchronous modality, but there was no significant difference in the performance of students with higher ability (upper 50^{th} percentile). Further, they found that asynchronous learning in the first few classes of the semester was more harmful to the lower 50^{th} percentile among the students. This result suggests that prior ability of students is important for success in the asynchronous format since this modality relies on the student's ability to learn through self-study.

As an example of a qualitative course, Zhu et. al. (2021) studied the effectiveness of teaching service leadership qualities during the pandemic in both asynchronous and synchronous modalities. They compared pre and post test scores for several key learning objectives of the course, and found that there was significant learning in both modalities for each of the student learning-objectives. However, they did not directly compare the two online modalities to see if one was more effective than the other.

A point of debate in these studies and their conclusions is whether students have a choice over whether to take a particular course online or F2F. That is, the finding of no significant difference depends on students being directly placed in each learning environment. In fact, as Allen & Seaman (2014) suggest, much of the current growth in online course offerings occurs in large universities offering both an online section of a course as well as one in a traditional format. In such settings, students are free to choose which delivery mode they prefer. Some studies examined the self-

selection bias while comparing the face-to-face modality with the online modality (Stanford-Bowers, 2008; Helms, 2014; Johnson & Palmer, 2015). However, none of the studies examined self-selection in comparing asynchronous and synchronous online modalities.

Given the limited number of studies that directly compare the asynchronous and synchronous modalities for learning effectiveness, our study presents such a comparison at a business school in a public research university, to help understand effect on student performance. Our study includes undergraduate Finance majors in a core, quantitative course. Given Zheng and Luo's (2023) meta-analyses-based findings, one might expect students to do better in the asynchronous modality for such a quantitative course.

Method

Two sections of the same 6-credit core course (FI 4000- *Valuation of Financial Assets*) for finance majors were taught by the same instructor in two different online modalities – asynchronous and synchronous. Student performance in the two sections on individual exams as well as team projects was compared. The sample size included 59 students in the synchronous online section and 26 students in the asynchronous online section.

To compare the individual student performance, scores on 3 different midterm exams were averaged, and a two-sample t-test performed to test the hypothesis that the mean scores in the two sections were the same. Likewise, student performance on teams was compared across the two sections, based on their team project scores. To ensure that student ability was not a confounding factor in the results, we compared the average prior GPAs of the students in the two sections.

Course Design

Gilpin's (2020) framework suggests designing the course based on student motivation, degree of interaction desired, and the flexibility offered to students. Holden & Westfall (2008) discuss various technologies and the nature of discourse in different learning environments in their guide to instructional media selection. According to them, online synchronous courses are better suited to environments where symmetric (two-way) interaction between the instructor and students is more important, and there is a need for immediate, real-time clarification of concepts. On the other hand, when such interactions are not critical, and students can benefit from imagery and narration, the asynchronous modality fits well.

The core course in our study is primarily quantitative and generally more challenging (as a foundational, solitary six-credit course in the curriculum) and therefore, we believe, requires significant interaction between the instructor and the students. Immediate feedback to the students from the instructor as well as working together in teams during class to solve problems has traditionally been the approach taken to teach this class. In addition, it is typically assumed that a student's intrinsic motivation is low in a core course because it is mandatory (Nargundkar & Shrikhande, 2012). Therefore, one would expect that students would perform better in this course if offered synchronously rather than asynchronously.

To test the hypothesis that the synchronous modality would be more effective than the asynchronous one, it was necessary to control some variables. Two sections were compared that were taught by the same instructor. Also, to maintain the quality of the course content, care was taken to design the course in both modalities to be as similar in experience for the students as possible. Table 1 highlights some of the key

features of this course, namely, the course structure, assignments, examinations, and the assessment.

Table 1
FI4000 – Online Synchronous vs Online Asynchronous

Aspect	Online Synchronous	Online Asynchronous
Class structure	Synchronous attendance Video recordings provided Breakout exercises; Polls Bi-weekly problems review	No synchronous sessions Videos in classroom setting Bi-weekly problems review Chat sessions on Zoom for Q&A
Assignments	Problem-set submission biweekly Quantitative group-project analysis Project report; Excel analysis	Problem-set submission biweekly Quantitative group-project analysis Project report; Excel analysis
Examinations	Proctored in zoom, with cameras on Help-sheet, financial calculators 150-minute duration exams	Proctored in zoom, with cameras on Help-sheet, financial calculators 150-minute duration exams
Assessment	Self-study and Chat sessions Quality of analysis & accuracy Rubrics for assessing examinations	Self-study and Chat sessions Quality analysis & accuracy Rubrics for assessing examinations

As shown in Table 1, the goal was to make the delivery of both the modalities as similar to each other as possible within the constraints of the modalities. The key difference in the class structure was that by definition, the asynchronous modality did not have any real-time synchronous online sessions, while in the synchronous modality, the students met with the instructor online each week at a set time. During these class meetings online, occasional polling was conducted to gauge student understanding. The asynchronous online students had access to weekly chat sessions on Zoom to ask questions of the instructor, based on their self-study from the materials posted online.

The assignments, examinations, and assessments were identical in both the modalities. For example, students were given a 2.5-hour time slot to take the exam on Zoom with the webcam on during the entire exam, and the instructor would be in the zoom session during the entire exam period to proctor and to answer questions. In other words, the examinations were synchronous, even for the asynchronous section. This helped to minimize the likelihood of cheating on examinations.

Findings from our previous study

The findings from our previous study (Nargundkar and Shrikhande, 2023) are briefly summarized below.

For the finance core course (FI4000), we found that average student performance on individual exams was significantly better in the synchronous online modality (average score of roughly 81%) compared to the face-to-face modality (average score of roughly 72%).

Likewise, for the international finance elective course (FI4040), we found that average student performance on individual exams was significantly better in the asynchronous online modality (again, roughly 81%) compared to the face-to-face modality (roughly 76%).

When it came to team project performance, the findings for the finance core course were the exact opposite, with the face-to-face modality delivering significantly better student performance (roughly 92%) when compared to the synchronous online modality (roughly 88%).

However, for the finance elective course, there was no significant difference between the face-to-face modality (roughly 87%) and the asynchronous online modality (roughly 86%), on the team-based case reports.

We also studied performance by gender, and found no significant differences between male and female student performance in the individual exams for either the core or the elective class.

Extension of the findings: Current paper

The current study, as mentioned, directly compares performance in the online asynchronous modality with the online synchronous modality within the core course.

The individual performance of students was gauged based on their average score on three different midterm exams. These average scores for the two modalities are compared in Table 2.

Table 2
Student Performance on Individual Exams

	Synchronous	Asynchronous	P-value
Mean	80.93	79.09	0.299
Std Dev	8.85	6.75	
Sample Size	59.00	26.00	

As Table 2 shows, there is no significant difference ($p > 0.05$) in average student performance between the synchronous online and asynchronous online modalities. We also separated the data by gender to see if there was a difference in performance between the modalities by gender. Once again, we found no significant difference in performance between the modalities, either for Male students or for Female students.

Similarly, the team performance of the students was gauged based on their team project grades. Table 3 shows a comparison of these project grades by modality.

Table 3
Student Performance on Team Projects

	Synchronous	Asynchronous	P-value
Mean	91.19	87.88	< 0.01
Std Dev	5.11	5.13	
Sample Size	59.00	26.00	

Table 3 shows that student team project performance was significantly better ($p < 0.01$) in the synchronous online modality compared to the asynchronous online modality.

We compared the prior GPAs of the students in each section to check for differences in ability among the students in the two modalities. Table 4 shows their GPAs and the result of the 2-sample t-test to check for significant differences in the average GPA.

Table 4
Student GPAs - Synchronous vs Asynchronous

	Synchronous	Asynchronous	P-value
Mean	3.29	3.25	0.709
Std Dev	0.50	0.55	
Sample Size	59.00	26.00	

The p-value indicates that there is no significant difference in the mean GPAs between the two sections. Thus, the initial ability of the students can be assumed to be equivalent.

Discussion

Our previously published results (Nargundkar and Shrikhande, 2023) showed counterintuitive results, with individual online performance better than in the face-to-face modality, for both synchronous and asynchronous online modalities. We provided several plausible reasons for these findings.

First, the modalities were chosen (synchronous for the core and asynchronous for the elective) based on factors identified in the literature (Gilpin 2020) such as flexibility, social interaction, online tools available, and initial student motivation.

Second, the design of the courses in each modality was carefully thought through to minimize the limitations of each modality. For example, social interaction is weakest in the asynchronous modality, and therefore, care was taken to incorporate opportunities for greater interaction between the students and their peers as well as between the students and the instructor.

Finally, students self-selected into the various modalities, and thus probably chose the modality best suited to their learning preferences. The online modalities also reduce the waste of time and energy in commuting, potentially allowing for more time to focus on studying, which is important considering that online modalities require more self-study.

From the previous study, which showed that individual student performance in both online modalities was better than the face-to-face modality, we found that the average exam scores in each case rose to about 81% (although in two different courses), suggesting that perhaps the two online modalities may not be different from each other in delivering student performance. We did not find any differences in performance across modalities by student gender.

In this study, the direct comparison of the two online modalities bears out that expectation when it comes to individual performance on exams. We found no significant difference between the synchronous and asynchronous modes (Table 2). However, when it comes to team projects, students in the synchronous modality did outperform those in the asynchronous modality. As with the previous study, we checked to make sure that the average GPAs across the modalities were not significantly different from each other, indicating that initial ability of the students was not a factor in the results. Also, as with our previous study, we found no gender-based differences in performance.

Both these results support the idea that when it comes to self-study, courses offered in both these modalities, when properly designed, will result in similar student learning outcomes. However, teamwork requires a greater degree of interaction between the students, and the synchronous mode does have an advantage over the asynchronous in this regard. Also, based on Krathwohl's (2002) updated model of Bloom's taxonomy (Bloom et al., 1956), the learning

objectives that were addressed in the individual exams were primarily a mixture of basic knowledge (remembering), and comprehension (understanding). The team-based projects, however, required higher levels of learning, beginning with analysis and application, and extending to creative thinking. Thus, our results may also indicate that higher levels of learning are better achieved with direct interaction that is afforded by the synchronous modality.

One factor that can influence the results, as discussed in the literature, is student self-selection into the different modalities. Students may choose an asynchronous or a synchronous modality based on their learning preferences or other individual characteristics that influence their performance. A student that values social interaction is perhaps more inclined to choose the synchronous modality and learn better as a result, while a student that is more independent minded, who likes to work alone may prefer the asynchronous modality. However, in any class, there are students that choose the class due to scheduling conveniences rather than a preference for one modality over another. Given that the individual student performance in our study was not significantly different across the two modalities, this can mean one of two things regarding the issue of self-selection. First, perhaps a majority of the students did not base their choice of section on the modality itself, and as mentioned above, chose it for scheduling reasons. Second, assuming students did select the modality based on personal preference for that modality, the results suggest that the advantage gained by students choosing the modality was equal for each modality. In other words, students choosing the synchronous modality may have gained a learning advantage for themselves, but so did the students that chose the asynchronous modality. As stated earlier, the prior GPAs across the modalities were not statistically distinguishable from one another, indicating that one group did not begin with any advantage over another.

However, student team performance was better in the synchronous modality. This could simply be, as discussed before, due to the increased interaction opportunities in the synchronous modality. However, it is possible that self-selection into the synchronous modality by students that prefer social interaction and learn better that way may have further enhanced their performance in team projects.

Combining the results from our previously published study and its extension presented here, we find that for individual performance, the online modality, whether synchronous or asynchronous, proves to be better than the face-to-face modality. Further, there is no significant difference in student performance between the two online modalities, based on the direct comparison shown in this study. For teamwork, in the previously published study, we found that student performance in the face-to-face modality was better than the synchronous online modality in the core course. In this extension of the study, we found that the synchronous online modality was better than the asynchronous modality for team project performance in the core course. Combining the two results yields a sequential progression of performance improvements from asynchronous online to synchronous online to face-to-face modalities. This is consistent with the expectation that student performance in team-based activities should indeed improve as the degree of interaction between students increases.

Implications for instructors

Our results showed that individual student performance was not significantly different across the two online modalities. The individual performances were measured based on examinations that primarily tested at lower levels of learning in Bloom's (1956) taxonomy. However, team performance in group projects tested higher levels of learning, where the synchronous modality

outperformed the asynchronous modality. The implication for instructors (or course coordinators) is that the more challenging and quantitative a course is, and the higher the level of learning desired, the more it might benefit student learning if offered in a synchronous modality. A course that requires students to only comprehend the basics of a subject without needing to engage in much critical thinking may work equally well in either modality.

In addition, it must be noted that student-student and student-instructor interaction of some sort is necessary for effective learning regardless of modality, and effort must be made to ensure that some opportunities for such interactions are provided where possible. For example, in the synchronous modality, breakout rooms for group interaction among students is strongly recommended. In fact, such breakout rooms often provide students with privacy from the instructor to openly discuss things that even a face-to-face class cannot provide. For the instructor-student interactions, the instructor could travel between breakout rooms, or use other techniques like occasional polling to assess student learning and participation.

For an asynchronous session, the instructor should have weekly chat sessions for Q&A that are synchronous. These should be optional, so that the flexibility inherently promised to students in an asynchronous class is maintained. However, instructors should take care to record all such sessions and post them online, so that those that did not attend can still benefit from the interactions that their classmates had with the instructor. Another idea to incorporate the benefits of interaction even in an asynchronous setting is to record the basic lecture videos in the form of an interaction with some students (either real students or actors) to simulate a live classroom session with the question and answer dynamics that are usually lost in a straight lecture video.

Conclusion

Comparing the findings from the early Brown & Liedholm (2002) study with both the previous Nargundkar and Shrikhande (2023) study and the current paper the following conclusions can be reached: a) The finding in the Brown & Liedholm (2002) study that the traditional face-to-face modality would result in the best student performance, followed by the hybrid modality, with the fully online modality being the least effective in delivering student performance, was negated by the finding in earlier Nargundkar and Shrikhande (2023) study that individual online performance was better than in the face-to-face modality, for both synchronous and asynchronous online modalities. Plausible reasons were provided to explain these findings. With the advent of online technologies, the potential for online teaching and learning to become as effective or more effective than face-to-face teaching and learning will only increase over time. Current evidence indicates that at least for individual skill-based learning, online modalities can outperform face-to-face classes.

Finally, our findings in the context of team projects in the earlier (Nargundkar and Shrikhande, 2023) and current study bring into focus the importance of interaction among students for better student performance and accordingly provide strong rationale for our finding that there is a sequential progression of performance improvements from asynchronous online to synchronous online to face-to-face modalities when higher level learning outcomes are desired.

Limitations and Future Research

This study was limited to a core Finance course that was quantitative in nature. The results, therefore, may not generalize to other disciplines. Moreover, our sample size, while adequate for

this study, was limited. For future research, extending such comparisons to other courses within Finance as well as to other business disciplines will help in understanding how the two online modalities affect student learning in different disciplines. A multi-course, multidisciplinary study with a larger sample size will help improve the generalizability of the results. Further, this study compares results in the short term, that is, at the end of the semester. Studies can be conducted to examine the long-term retention of knowledge and skills acquired in each modality.

Student self-selection into one modality or another deserves to be studied in greater depth. For instance, are there student characteristics that influence their choice of modality and their performance?

Further developments in technology, such as the use of AI-based assistants in teaching, may also influence the effectiveness of different teaching modalities.

References

Allen, I. E., & Seaman, J. (2014). *Grade Change: Tracking Online Education in the United States*. Babson Survey Research Group and Quahog Research Group, LLC. http://www.onlinelearningsurvey.com/reports/gradechange.pdf

Bloom, B., Englehart, M., Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. New Longmans, Green.

Brown, Byron W., & Liedholm, Carl E. (2002). Can web courses replace the classroom in principles of microeconomics? *The American Economic Review*, *Papers and Proceedings of the One Hundred Fourteenth Annual Meeting of the American Economic Association*, May, 92(2), 444-448.

Ebner, C., & Gegenfurtner, A. (2019). Learning and satisfaction in webinar, online, and face-to-face instruction: A meta-analysis. *Frontiers in Education*, 4, 92. DOI:10.3389/feduc.2019.00092

Gegenfurtner, A., & Ebner, C. (2019). Webinars in higher education and professional training: A meta-analysis and systematic review of randomized controlled trials. *Educational Research Review*, 28, 100293.

Gilpin, Staci (2020). A framework for fostering emerging online learner persistence: The role of asynchronous and synchronous discussion. *Journal of Teaching and Learning Special Issue*: *Digital Learning in Higher Education,* 14(1), 29–42.

Helms, J. L. (2014). Comparing student performance in online and face-to-face delivery modalities. Online Learning, 18(1). DOI: https://doi.org/10.24059/olj.v18i1.348

Holden J. T., & Westfall. J. (2008). *An instructional media selection guide for distance learning,* Fifth Edition, United States Distance Learning Association (USDLA).

Johnson, D., & Palmer, C. C. (2015). Comparing student assessments and perceptions of online and face-to-face versions of an introductory linguistics course. *Journal of Asynchronous Learning Network*, 19(2).

Krathwohl, D. R. (2002). A revision of bloom's taxonomy: An overview. *Theory Into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Le, K. (2022). Pre-recorded lectures, live online lectures, and student academic achievement. *Sustainability*, 14 (5), 2910.

Nargundkar, S. & Shrikhande, M. (2012). An empirical investigation of student evaluations of instruction—the relative importance of factors. *Decision Sciences Journal of Innovative Education,* 10(1), 117-135.

Nargundkar, S. and M. Shrikhande (2023). Examining student performance in face-to-face, online synchronous and online asynchronous instruction, *Journal of Financial Education*, Spring, 48 (1), 1-35.

Stanford-Bowers, D. E. (2008). Persistence in online classes: A study of perceptions among community college stakeholders. *Journal of Online Learning and Teaching*, 4(1), 37-50.

Zeng, Hang & Luo, Jiuton (2023). Effectiveness of synchronous and asynchronous online learning: A meta-analysis, *Interactive Learning Environments*, DOI: 10.1080/10494820.2023.2197953

Zhu, X., Shek, D. T., & Chan, C. H. (2021). Promoting service leadership qualities and well-being among university students through an online course during covid-19 pandemic. *International Journal of Environmental Research and Public Health*, 18(15), 8162.

Determinants of Financial Literacy and Tax Literacy of Canadian Students

Dorothee Feils
MacEwan University

Michelle Malin
MacEwan University

Eloisa Perez-de Toledo
MacEwan University

Determinants of financial literacy and tax literacy are essential areas of research as low levels of financial literacy and tax literacy can negatively affect the well-being of individuals and the economy as a whole. This study analyzes the determinants of financial literacy and tax literacy of post-secondary students enrolled in a Canadian university. We find that age, gender, year of study, ownership of financial products, confidence in financial knowledge and the access to formal sources of financial information are positively related to financial and tax literacy levels of post-secondary students. In addition, access to informal financial information is positively related to tax literacy levels. Thus, financial and tax education should be provided using a variety of methods (both formal and informal) to post-secondary students to improve financial and tax literacy.

Keywords *financial literacy, tax literacy, post-secondary students, Canada*

Introduction

Literacy is a broad concept that is defined by the Organization for Economic Co-operation and Development (OECD) program for the international assessment of adult competencies as "the ability to identify, understand, interpret, create, communicate and compute, using printed and written materials associated with varying contexts. Literacy involves a continuum of learning in enabling individuals to achieve their goals, to develop their knowledge and potential, and to participate fully in their community and wider society" (Adult Literacy, 2003). Specifically to financial knowledge, The Canadian Task Force on Financial Literacy defines it as "(…) the knowledge, skills and confidence to make responsible financial decisions". More broadly, Lusardi (2015) defines financial literacy as "the ability to process economic information and make informed decisions about financial planning, wealth accumulation, debt, and pensions" (p.260). Similarly, Atkinson and Messy (2012, p.5) define it as "a combination of awareness, knowledge, skill, attitude and behaviour necessary to make sound financial decisions and ultimately achieve individual financial wellbeing." In this context, we define financial literacy as a set of skills developed from the accumulation of financial knowledge acquired formally and/or informally that enables individuals to make effective financial decisions throughout their life. By effective financial decisions we mean decisions that yield optimal outcomes as they are informed by research and a clear understanding of the individual realities and needs.

Financial literacy is considered an essential life skill, with empirical evidence supporting a strong positive association between financial knowledge and financial wealth (Van Rooij et al., 2012). In addition, financial literacy increases the confidence of individuals to invest in the stock market, which may allow them to profit from higher returns on their investments in the long term (Lusardi & Mitchell, 2007; Thomas & Spataro, 2018) and enjoy a more comfortable retirement. The importance of financial literacy is even greater today due to the growth of economic markets and greater accessibility of economic and financial transactions through internet purchasing (Ergun, 2017), the advent of cryptocurrencies, crowdfunding and crowdfinancing (Hua, Huang & Zheng, 2019). At the same time, the number and complexity of financial investment choices are increasing (Lusardi & Mitchell, 2014). Furthermore, Canada and other countries have seen a trend over the last few decades of defined benefit pension plans replaced by defined contribution pension plans, shifting the financial risk for retirement income from the corporation to the individual (Broadbent et al., 2006). Put all together, individuals today are making considerably more financial decisions throughout their lives due to the increased access to new financial products and longer lifespans (Lusardi, 2019). Yet, despite its growing importance, existing research has shown that financial literacy is lacking overall (Lusardi & Mitchell, 2017).

After years of research and the development of myriad of programs aimed at increasing the level of financial literacy of individuals, financial literacy is still low specially among the very young (Lusardi & Mitchell, 2007). Chen and Volpe's (1998) study on college students shows that only 53% answered the financial literacy questions correctly and they conclude that the low level of financial literacy limits students' ability to make informed decisions. In addition, extant financial literacy literature focuses almost exclusively on financial knowledge as a measure of financial literacy. However, there is some evidence indicating that tax literacy is also an important component of financial literacy. Tax literacy is similar to financial literacy but applies specifically to tax knowledge. Tax literacy can be defined as "having the knowledge, skills and confidence to make responsible tax decisions" (Godbout, Genest-Grégoire & Guay, 2017, p. 6). Therefore, tax literacy is central to helping individuals meet the definition of financially literate according to Lusardi and Mitchell (2014), which will contribute to the individual's future success and wellbeing. Similar to financial literacy, studies show that tax literacy of the general population is very low (Godbout et al., 2017; Pham et al., 2020).

The objective of this study is to identify the determinants of financial and tax literacy for a cohort of Canadian post-secondary students. Financial literacy is arguably more important for post-secondary students than other cohorts, as they are generally close to the age where they will be making critical financial decisions, such as purchasing a home, a vehicle and negotiating employment benefits. Tax literacy is also important for post-secondary students as there are a number of credits (such as the tuition credit and student loan interest credit) as well as government benefit programs (such as the Lifelong Learning Plan, Canada Education Savings Grants or Canada Learning Bonds, and Universal Childcare), which specifically benefit this group. We therefore designed this study to investigate what factors impact the financial literacy and tax literacy of students attending an undergraduate university in Canada.

The paper is divided as follows: part 2 provides a review of the extant literature in financial and tax literacy and sets forth our hypotheses. Part 3 describes the data and methodology. Part 4 discusses our results and Part 5 concludes the paper.

Literature Review

Numerous studies into financial literacy indicate that a wide range of demographic factors impacts a young person's understanding of personal finance issues, including income (Peng et al., 2007), age (LaBorde & Mottner, 2016), gender (Chen & Volpe, 2002), and the level of financial literacy of parents (Chen & Volpe, 1998). Specifically, many studies find that females are less likely to answer financial literacy questions correctly than their male counterparts (Chen & Volpe, 1998 & 2002). In addition, several studies have shown that older students perform better on a variety of financial literacy questions related to money management, retirement, debt management, tax, and insurance (Chen & Volpe, 2002; LaBorde & Mottner, 2016). Several socio-demographic characteristics of a young person's family, including income level, investment ownership, and parental figures' education level, are strongly related to improved financial literacy (Cameron et al., 2014; Lusardi et al., 2010). These results suggest that some financial knowledge is transferred from parents to their children (Lusardi et al., 2010), which aligns with the findings from a survey of college students conducted by Chen and Volpe (2002). "When asked where they acquire personal finance knowledge, 74% of women and 68% of men answer that they obtained the knowledge from their parents. The next most frequent answer is that they learned from their own mistakes (70% for women vs. 64% for men)." (Chen & Volpe, 2002, p. 13).

There is evidence showing that life experiences also impact students' financial literacy knowledge. Students are more likely to answer personal finance questions correctly if the concept relates to the student's personal experience, as students performed better on financial literacy questions related to automobile insurance and apartment leases (Chen & Volpe, 1998). Likewise, students performed poorly on questions related to topics with which they have little experience, such as retirement savings and income tax (LaBorde & Mottner, 2016), while students who experienced financial difficulties, struggled to pay for their educational expenses, and those who have loans or debts outstanding, have higher levels of financial literacy (Ergun, 2017). Similarly, students with more work experience performed better than students with fewer years of work experience (Chen & Volpe, 1998 & 2002). In addition, students who have had a bank account for longer tend to perform better on financial literacy questions (Cameron et al., 2014; Peng et al., 2007). These results are consistent with findings from studies of the general public that show that individuals who currently own investments (such as stocks or bonds) displayed greater investment knowledge than those who do not hold any investments (Peng et al., 2007). Several studies show that older students perform better on personal finance questions than younger students (Chen & Volpe, 1998, 2002; LaBorde & Mottner, 2016). These results may be tied to the life experience of older students as older students are more likely to have more years of work experience and living on their own (Chen & Volpe, 1998).

The fact that older students demonstrate significantly higher levels of financial literacy could also be linked to additional years of education. Individuals with higher levels of education are more likely to answer questions on personal finance correctly (Chen & Volpe, 1998; Lusardi et al., 2010). Financial literacy for students increases as their level of education and years of post-secondary education increase (Bhushan & Medury, 2013; Peng et al., 2007). In addition, students with higher levels of intelligence and greater cognitive ability are likely to perform better on financial literacy questions (LaBorde & Mottner, 2016; Lusardi et al., 2010). Thus, education is the most significant factor in improving financial literacy (Albeerdy & Gharleghi, 2015). A study by LaBorde and Mottner (2016) found that a personal finance course provided to college students

could raise overall financial literacy in all areas and decrease the inequality of personal finance knowledge across students by gender, age, and ethnicity.

Although education can improve individuals' objective financial knowledge, it also impacts their perception of their financial knowledge. Even providing individuals with some information on diversification and risk, whether in the form of a short video, written narrative, or interactive visual tool, significantly decreases the number of "don't know" responses to financial literacy questions (Lusardi & Mitchell, 2014). Perceived financial knowledge or confidence has a significant impact on financial behavior. Those with higher confidence in their financial literacy are less likely to seek financial advice (Kramer, 2016). Not seeking financial advice can lead to one making poor financial decisions. Overconfidence combined with low levels of objective financial knowledge was positively associated with risky (costly) financial behaviors such as obtaining payday loans or auto title loans (Tokar Asaad, 2015). The results of these studies indicate that financial literacy education needs to focus on improving an individual's knowledge of personal finance topics as well as confidence in their financial knowledge.

In contrast to financial literacy, where there have been numerous studies on the personal finance knowledge of post-secondary students and the general public, there are very few studies into tax literacy, which indicates a need for additional research in this area. Similar to financial literacy, the tax literacy of university students and the general public is low (Chardon et al., 2016; Godbout et al., 2017). Other similarities between financial literacy and tax literacy are that education and income are positively linked to better performance on surveys and questionnaires (Pham et al., 2020). Individuals with more years of education and higher income levels are more likely to answer questions related to personal income tax and the progressive nature of the tax system correctly (Godbout et al., 2017; Pham et al., 2020). Additional years of education are also associated with a greater level of tax knowledge for university students (Chardon et al., 2016; Pham et al., 2020). Unlike financial literacy, where men traditionally demonstrate higher knowledge levels, that pattern does not always hold for tax literacy. For example, in a study of Quebec residents, even though female respondents were more likely to underestimate their tax knowledge, their performance on personal income tax questions was equal to male respondents (Godbout et al., 2017). On the other hand, a study of Australian university students found that male students outperformed female students on income tax questions (Chardon et al., 2016). This result is consistent with findings from a 2019 study of the tax literacy of Canadians, where the average score for men was higher than the average score for women (Pham et al., 2020). Another difference between financial literacy and tax literacy studies results is that work experience, and the added income tax obligations that come with that experience do not impact tax literacy. Individuals who operated an unincorporated business (i.e., self-employed individuals) performed no better or worse on personal income tax questions than other respondents (Godbout et al., 2017). Likewise, employment was not related to income tax knowledge in a study of university students (Chardon et al., 2016).

In the present study, we analyze the determinants of financial and tax literacy in a sample of students enrolled in post-secondary studies at a Canadian university. To achieve this objective, we collected data that allowed us to construct scores to measure the level of financial literacy and the level of tax literacy. Next, we test the relationship between the scores and factors that are known to affect the levels of financial and tax literacy: (1) demographic characteristics, such as age, gender, parent's education, (2) student status, such as program of study and year of study, (3) access to previous financial education, either from formal or informal sources, (4) propensity to save, (5) risk tolerance, (6) money sentiment, and (7) financial self-confidence, the level of confidence

individuals have in their own financial knowledge and how comfortable they feel managing money.

We hypothesize that (H_1) gender, age and parent's education all affect the level of financial and tax literacy as previous studies have shown (Chen & Volpe, 2002), student's program of study (H_2) should impact financial and tax literacy as being enrolled in a Bachelor in Business or Commerce degree as opposed to other careers is hypothesized to have a positive impact on the level of financial literacy. Students that have accessed some type of financial education (H_3), either from formal or informal sources, should have a higher financial and tax literacy level. Propensity to save (H_4) is a construct that measures students decision when faced with the option to save or spend excess cash. Excess cash is defined as money they have earned and accumulated, and that the student does not need to pay for ordinary daily expenses or to pay back debt. Risk tolerance (H_5) is an attribute that measures a person's attitude towards accepting risk. It is also defined as the level of risk that an individual prefers to accept. There is evidence that risk tolerance is a personality trait and as such it does not change much throughout life (Van de Venter et al., 2012), however some studies show that risk tolerance decreases with age and increases with the level of education (Grable & Joo, 2004). Risk tolerance is measured by a self-assessment question that measures the subjective level of financial risk that each student is willing to accept, similar to Hallahan et al. (2004). Money sentiment (H_6) is a construct that measures students' perception of "money" and "debt", we hypothesize that students with negative sentiment towards money and debt will present lower financial and tax literacy scores. Financial self-confidence (H_7) is measured as a score calculated from students' answers to questions related to how confident they feel when managing their finances (debt and investments). It is hypothesized that students that feel more comfortable using a financial calculator and more confident managing money will present higher levels of confidence managing their personal finances and therefore higher levels of financial and tax literacy.

Our hypotheses and variables are summarized in Table 1.

Table 1
Variable description and descriptive statistics

Variable	Measure	Mean	Std. Dev.	Min	Max
Financial Literacy Score	Twelve questions measuring students' basic accounting knowledge (e.g. definition of net worth), risk and return basics (e.g. relate financial product to most appropriate level of interest rate), purchasing power parity effect on real returns, financial decisions (e.g. increase debt vs save for retirement), credit card management, and portfolio diversification.	0.541	0.256	0	1
Lusardi-Mitchell Score	Percentage of correct answers to Lusardi's 3 questions that measure the level of understanding about compound interest rates, inflation, and risk diversification.	0.614	0.339	0	1
Tax Score	Fifteen questions measuring students' knowledge of available tax credits and eligible taxable income.	0.425	0.256	0	1
Gender	1=male, 0=female	0.531	0.499	0	1
Age	≤18, 19-20, 21-24, 25-29, ≥30 years old	2.537	0.974	1	5
Year	Year of studies (1 to 5)	2.150	0.945	1	5
Program	Program of studies: Open Studies=0 Certificate (1 year)=1 Diploma (2 years)=2 Bachelor (4 years)=3	2.73	0.614	0	3

Major	Major not declared=0 major declared=1	0.661	0.474	0	1
Parents' Level of Education	Don't know=0 Less than High School=1 High School=2 Technical training (Trades)=3 Diploma=4 Bachelor's degree=5 Graduate degree (Master's or Doctorate/PhD)=6	3.900	1.777	0	6
Information Source (Informal)	Have used an Internet source to access Personal Finance information (Google, YouTube, etc.) or read a book or consulted a friend/relative =1, otherwise=0	0.711	0.453	0	1
Information Source (Formal)	Have attended a seminar, class or High School course on Personal Finance = 1, otherwise=0	0.404	0.491	0	1
Money Sentiment	Score in the question "Debt is Bad" on a scale from 1 to 7	5.095	1.710	1	7
Confidence Score	Score calculated from student's self-assessment of how confident they are with managing money and debt, using investment products, and how much their family talked about finances while they were growing up.	0.679	0.157	0	1
Financial Product Score	Score calculated if the student owns any of the following financial products: RRSP, TFSA, Mortgage, Stocks, Mutual Funds.	0.256	0.244	0	1
Propensity to Save Score	Score calculated from students' answers to questions related to decision making when faced with the option to save or spend excess cash.	0.671	0.471	0	1
Risk Tolerance	Self-assessment question that measures the subjective level of risk that each individual prefers to accept.	0.426	0.186	0	1

Data and Methodology

We examine these issues using a questionnaire that included the three basic questions proposed by Lusardi and Mitchell (2014) and Mitchell and Lusardi (2015) as well as other questions designed to measure the levels of financial and tax literacy of undergraduate students at a Canadian university. All respondents completed the questionnaire on the first day of class of their introductory accounting or finance course in January of 2020. In total, 596 students completed the survey. Table 1 shows descriptive statistics for our sample.

We generated two different scores to measure the financial literacy of the students: first, the Lusardi-Mitchell Score, using the three fundamental financial literacy questions proposed by Lusardi and Mitchell (2014) and second, a comprehensive Financial Literacy Score that is composed by twelve questions that measure students' basic knowledge of accounting (e.g. definition of net worth), of risk and return (e.g. relate a financial product to the most appropriate level of interest rate), of the effect of inflation on real returns, of sound financial decision making (e.g. increase personal debt vs save for retirement), of credit card management, and of portfolio diversification.

In addition, we generated a Tax Score to measure the tax literacy of the students in the sample. The Tax Score is composed by fifteen questions that measure two dimensions of tax literacy: (1) student's knowledge of available tax credits to their demographics and (2) student's knowledge of all income that would be considered taxable by the current tax legislation normally earned by their demographics.

Other questions in the questionnaire were related to demographics of each participant as well as the level of previous exposure to financial education from formal and informal sources. We also collected data on the factors that are known to affect financial and tax literacy in young adults, such as risk tolerance, propensity to save, money sentiment and whether students hold an investment account or not. A summary of descriptive statistics of our sample demographics is reported in Table 2. Fifty three percent of surveyed students were male and 47% female, 87% are under 24 years of age, 72% is enrolled in either the first or second year of a Bachelor or Diploma program and 50% of their parents hold a graduate or undergraduate degree.

Table 2
Sample demographics

	Frequency	**Percentage**
Gender		
Female	279	46.81%
Male	316	53.02%
Age group		
⩽ 18	58	9.73%
19-20	275	46.14%
21-24	183	30.70%
25-29	45	7.55%
30+	35	5.87%
Year of studies		
First	146	24.58%
Second	283	47.64%
Third	106	17.85%
Forth	48	8.08%
≥ Fifth	11	1.85%
Program		
Open Studies	17	2.85%
Certificate	3	0.50%
Diploma	102	17.11%
Degree	474	79.53%
Parent's level of education		
≥ High School	15	2.52%
High School Diploma	60	10.07%
Trades, Diploma, CEGEP, etc.	161	27.01%
Undergraduate studies	209	35.07%
Graduate Studies	92	15.43%
Don’t know	59	9.90%

Table 3 compiles the results obtained in each of the scores (Financial Literacy Score, Lusardi-Mitchell Score and the Tax Score) for the entire sample and aggregated using different demographic groupings. Overall, the percentage of correct answers for the Lusardi-Mitchell Score is higher than the other scores, although the financial literacy score shows similar results. The Tax Score has a lower percentage of correct answers as it requires specialized knowledge whose acquisition is normally triggered by an event that has tax implications for the individual taxpayer, such as joining the workforce or opening a registered savings account (i.e. a retirement savings plan). For this reason, the tax scores increase substantially with age, with students in the age group 30+ years old scoring considerably higher than their younger peers. Consistent with previous research, our results show gender and age as important determinant factors of the level of financial literacy, with older and male students showing a higher level of financial literacy, especially for the Lusardi-Mitchell Score. The Tax Score shows a smaller difference between male and female students, although still statistically significant. Tax knowledge is influenced by age, formal financial education and by the ownership of financial investments, especially registered savings plans, as those investments have tax implications.

Table 3
Correct responses to each score
(Financial Literacy Score, Lusardi-Mitchell Score and Tax Literacy Score)

		% of correct answers		
		Financial Literacy Score	**Lusardi-Mitchell Score**	**Tax Literacy Score**
All sample		0.541	0.614	0.425
Gender*	Female	0.460	0.4970	0.403
	Male	0.614	0.718	0.445
Age group				
	⩽ 18	0.411	0.489	0.320
	19-20	0.525	0.619	0.392
	21-24	0.559	0.608	0.439
	25-29	0.620	0.667	0.573
	30+	0.693	0.733	0.596
Year of studies				
	Year 1	0.425	0.505	0.337
	Year 2	0.559	0.635	0.422
	Year 3	0.612	0.660	0.508
	Year 4	0.629	0.688	0.531
	Year 5	0.545	0.697	0.424
Level of risk tolerance				
	Low	0.514	0.587	0.402
	High	0.581	0.652	0.458

Has accessed formal sources of financial education before				
	No	0.523	0.580	0.402
	Yes	0.569	0.663	0.459
Has accessed informal sources of financial education before				
	No	0.516	0.583	0.383
	Yes	0.552	0.626	0.442
Hold investments (TFSA/RRSP)				
	No	0.476	0.537	0.355
	Yes	0.580	0.662	0.466

(*) 0.2% of our sample did not report gender.

Table 4 reports descriptive statistics for other factors identified as determinants of the degree of financial and tax literacy of post-secondary students. The level of confidence when dealing with money was measured by a series of self-assessment questions related to how confident students feel using a calculator, managing money, investments and debt, as well as how much their family talked about finances when they were growing up. The Confidence Score was built based on these questions and is reported in the first column of Table 4. The next columns show the scores for each question by different demographic groupings. In general, the level of confidence with using a financial calculator was consistent among different gender, age and program of studies. The lower level of confidence was reported when managing investments, which is expected as only a small percentage of university students hold investments (26% of our sample). Family Talk measures how much family members talked about finances when the student was growing up. Our results show a consistent level across all demographics with the lower levels reported for mature students (30+ years of age) and students enrolled in a certificate program (1 year duration).

The last three columns report the scores for the Propensity to Save Score which is calculated from students' answers to questions related to decision making when faced with the option to save or spend excess cash. The 25 to 29 years old demographics presents the higher propensity to save. Money sentiment is a score that measures student's perception about money and debt (how much do you agree with the statement "*debt is bad*"). The higher the score the more negative the perception about money and the sentiment towards taking on debt. The results show that female students score higher in this construct which aligns with the results of higher propensity to save score.

The last column reports the results for Risk Tolerance. Financial risk tolerance is defined as an individual's attitude towards risk, and it has been found to be a determinant factor in asset selection when investors decide which assets to include in their portfolios (Droms, 1987). It is an important factor to consider when analyzing financial literacy, as risk tolerance influence financial behavior. However, there is no consensus in the literature on the determinants of individuals' level of tolerance to risk, although some studies indicate that factors such as age, gender and education are related to risk tolerance levels. More specifically, risk tolerance reduces with age, which is expected as younger investors have more time to recover from a market downturn than older people. Similarly, female investors seem to have a lower tolerance to risk than males (Fisher & Yao, 2017). Finally, higher levels of education are related to a higher capacity to assess and evaluate risks, therefore individuals with a higher level of education tend to be more risk tolerant and assume more risks. The results obtained from our sample confirm the evidence that risk tolerance increases with education, as it is higher among students in the last year of studies (Chaulk

et al., 2003; Grable & Joo, 2004). However, contrary to the expected, our sample shows that female students are more tolerant to risk than male students.

TABLE 4
Determinants of Financial Literacy and Tax Literacy

		I feel confident:							
	Confidence Index	***Family talk***	***using a calculator***	***managing money***	***managing credit card debt***	***managing investments***	**Propensity to Save**	**Money Sentiment**	**Risk Tolerance**
Entire sample	0.679	*0.570*	*0.855*	*0.730*	*0.734*	*0.507*	0.671	0.722	0.426
Gender									
Male	0.682	*0.540*	*0.852*	*0.734*	*0.741*	*0.544*	0.651	0.684	0.391
Female	0.678	*0.606*	*0.861*	*0.728*	*0.729*	*0.466*	0.694	0.768	0.466
Age group									
≤ 18	0.652	*0.574*	*0.862*	*0.717*	*0.638*	*0.468*	0.638	0.788	0.429
19-20	0.672	*0.560*	*0.843*	*0.728*	*0.727*	*0.501*	0.635	0.717	0.430
21-24	0.691	*0.591*	*0.859*	*0.729*	*0.759*	*0.517*	0.683	0.707	0.413
25-29	0.703	*0.619*	*0.892*	*0.733*	*0.743*	*0.527*	0.822	0.721	0.430
30+	0.692	*0.465*	*0.869*	*0.771*	*0.812*	*0.543*	0.743	0.727	0.448
Year of studies									
year 1	0.648	*0.569*	*0.818*	*0.707*	*0.679*	*0.468*	0.641	0.759	0.435
year 2	0.686	*0.562*	*0.869*	*0.737*	*0.746*	*0.518*	0.668	0.723	0.408
year 3	0.701	*0.586*	*0.867*	*0.756*	*0.765*	*0.532*	0.708	0.702	0.453
year 4	0.680	*0.571*	*0.863*	*0.702*	*0.753*	*0.512*	0.729	0.631	0.424
year 5	0.673	*0.610*	*0.831*	*0.714*	*0.779*	*0.429*	0.545	0.766	0.525
Program of studies									
Certificate	0.600	*0.381*	*0.905*	*0.619*	*0.714*	*0.381*	-	0.762	0.481
Diploma	0.676	*0.599*	*0.838*	*0.727*	*0.718*	*0.496*	0.676	0.759	0.411
Bachelor	0.682	*0.566*	*0.860*	*0.732*	*0.740*	*0.514*	0.681	0.710	0.428
Open Studies	0.614	*0.500*	*0.796*	*0.714*	*0.663*	*0.398*	0.429	0.816	0.421
Parent's education									
≤ High School	0.697	*0.619*	*0.867*	*0.752*	*0.743*	*0.505*	0.800	0.810	0.356
High School	0.670	*0.574*	*0.845*	*0.707*	*0.726*	*0.500*	0.667	0.690	0.426
Technical	0.665	*0.519*	*0.859*	*0.734*	*0.749*	*0.465*	0.772	0.664	0.480
Diploma	0.683	*0.575*	*0.843*	*0.756*	*0.753*	*0.486*	0.720	0.730	0.407
Bachelor	0.706	*0.618*	*0.867*	*0.738*	*0.759*	*0.546*	0.644	0.753	0.420
Graduate	0.690	*0.561*	*0.887*	*0.753*	*0.730*	*0.520*	0.663	0.722	0.425
Don't know	0.588	*0.458*	*0.782*	*0.646*	*0.615*	*0.441*	0.542	0.685	0.420

Analysis

In this section we analyze the results we obtain using OLS regressions to analyze and test our hypotheses using the survey data. The results of the regression analyses are reported in Tables 5, 6 and 7. First, we run a regression on the entire sample and next we analyze the sample by gender.

Table 5 reports the results of the regressions on the Tax Score, the Financial Literacy Score and the Lusardi-Mitchell Score (Lusardi & Mitchell, 2014) for the entire sample. We find that gender, age and parental educational levels are significantly positively related to the tax literacy score and the financial literacy score, consistent with (H_1). As predicted in H_2, the year of study, program of study and major declared are positively related to the financial literacy score, but only year of study and major (or specialization) are significantly related to the Tax Literacy Score. The use of formal information sources is positively related to the Tax Literacy Score and the Financial Literacy Score, consistent with (H_3), while the use of informal information sources is significantly related to the Tax Literacy Score. Contrary to expectations (H_6), we find no relationship between risk tolerance and financial or tax literacy. As expected, the confidence score and the financial product score are positively related to the tax literacy score and the financial literacy score (H_5 and H_7).

Table 5
Regression Analysis of the Financial Literacy, Tax Literacy and Lusardi-Mitchell Scores for the entire sample.

Table 5 Panel a. Regression Analysis of the Tax Literacy Score for the entire sample.				
	(1)	(2)	(3)	(4)
VARIABLES	Tax Score			
Gender	0.045**	0.033	0.040**	0.028
	(0.020)	(0.020)	(0.020)	(0.020)
Age	0.053***	0.036***	0.056***	0.040***
	(0.011)	(0.011)	(0.011)	(0.011)
Year	0.029**	0.027**	0.027**	0.025**
	(0.012)	(0.012)	(0.012)	(0.012)
Program	0.011	-0.007	0.004	-0.008
	(0.016)	(0.016)	(0.017)	(0.016)
Major Declared	0.052**	0.043*	0.045**	0.037*
	(0.023)	(0.022)	(0.023)	(0.023)
Money Sentiment		-0.010		-0.011*
		(0.006)		(0.006)
Financial Product Score		0.176***		0.166***
		(0.043)		(0.044)
Propensity to Save Score		0.047**		0.046**
		(0.021)		(0.021)
Risk Tolerance Score		0.025		0.021
		(0.021)		(0.021)

Confidence Score		0.224***		0.205***
		(0.072)		(0.072)
Parents' Level of Education			0.013**	0.009*
			(0.006)	(0.006)
Information Source (informal)			0.054**	0.043*
			(0.022)	(0.022)
Information Source (formal)			0.043**	0.024
			(0.020)	(0.020)
Constant	0.140**	0.060	0.055	0.013
	(0.055)	(0.076)	(0.061)	(0.079)
Observations	593	589	586	585
R-squared	0.104	0.172	0.127	0.184

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5 Panel b.
Regression Analysis of the Financial Literacy Score for the entire sample.

	(5)	(6)	(7)	(8)
VARIABLES	Financial Literacy Score			
Gender	0.150***	0.131***	0.149***	0.130***
	(0.019)	(0.019)	(0.020)	(0.019)
Age	0.045***	0.027**	0.046***	0.029***
	(0.011)	(0.011)	(0.011)	(0.011)
Year	0.029**	0.023**	0.028**	0.022**
	(0.011)	(0.011)	(0.011)	(0.011)
Program	0.053***	0.033**	0.046***	0.032**
	(0.016)	(0.015)	(0.016)	(0.015)
Major Declared	0.057***	0.048**	0.053**	0.044**
	(0.022)	(0.021)	(0.022)	(0.021)
Money Sentiment		-0.024***		-0.024***
		(0.006)		(0.006)
Financial Product Score		0.195***		0.191***
		(0.041)		(0.041)
Propensity to Save Score		0.026		0.027
		(0.020)		(0.020)
Risk Tolerance Score		0.007		0.005
		(0.019)		(0.019)
Confidence Score		0.250***		0.244***
		(0.067)		(0.068)
Parents' Level of Education			0.008	0.005
			(0.005)	(0.005)

Information Source (informal)			0.020	0.010
			(0.021)	(0.021)
Information Source (formal)			0.033*	0.014
			(0.020)	(0.019)
Constant	0.103**	0.112	0.066	0.089
	(0.052)	(0.071)	(0.059)	(0.074)
Observations	593	589	586	585
R-squared	0.190	0.274	0.196	0.277

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5 Panel c.
Regression Analysis of the Lusardi-Mitchell Score for the entire sample.

	(9)	(10)	(11)	(12)
VARIABLES	Lusardi-Mitchell Score			
Gender	0.216***	0.201***	0.213***	0.198***
	(0.026)	(0.027)	(0.027)	(0.027)
Age	0.028*	0.012	0.028*	0.013
	(0.015)	(0.015)	(0.015)	(0.015)
Year	0.030*	0.022	0.027*	0.022
	(0.015)	(0.015)	(0.015)	(0.015)
Program	0.049**	0.030	0.041*	0.030
	(0.021)	(0.022)	(0.022)	(0.022)
Major Declared	0.049*	0.048	0.046	0.043
	(0.030)	(0.029)	(0.030)	(0.030)
Money Sentiment		-0.021***		-0.021***
		(0.008)		(0.008)
Financial Product Score		0.217***		0.204***
		(0.057)		(0.057)
Propensity to Save Score		-0.011		-0.009
		(0.028)		(0.028)
Risk Tolerance Score		-0.007		-0.013
		(0.027)		(0.027)
Confidence Score		0.156*		0.137
		(0.095)		(0.096)
Parents' Level of Education			0.008	0.006
			(0.007)	(0.007)
Information Source (informal)			0.027	0.023
			(0.029)	(0.029)
Information Source (formal)			0.071***	0.056**
			(0.027)	(0.026)

Constant	0.198***	0.270***	0.150*	0.231**
	(0.071)	(0.100)	(0.079)	(0.103)
Observations	593	589	586	585
R-squared	0.144	0.182	0.155	0.190

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In order to explore the importance of gender for tax literacy and financial literacy more in-depth, we repeated the regression analysis by gender. The results of these regressions are reported in Tables 6 and 7. Most of the results are consistent with the results for the entire sample. The main differences are that the parental level of education has no significant relationship to tax and financial literacy for females but has a significant positive relationship for male students. On the other hand, formal sources of information are significantly positively related to tax and financial literacy for females, but we find no significant relationship for male students. Finally, the program of study is negatively related to the tax score for females and positively related for male students.

Table 6
Regression Analysis of the Financial Literacy, Tax Literacy and Lusardi-Mitchell Scores by Gender. Results reported are for female students.

Table 6 Panel a. Regression Analysis of the Tax Literacy Score by Gender. Results reported are for female students.				
	(1)	(2)	(3)	(4)
VARIABLES	Tax Score			
Age	0.058***	0.040**	0.055***	0.039**
	(0.015)	(0.016)	(0.015)	(0.016)
Year	0.030*	0.028*	0.029*	0.027*
	(0.015)	(0.015)	(0.016)	(0.015)
Program	-0.041**	-0.051**	-0.034	-0.044**
	(0.021)	(0.021)	(0.021)	(0.021)
Major Declared	0.050	0.038	0.042	0.032
	(0.032)	(0.032)	(0.032)	(0.033)
Money Sentiment		-0.009		-0.010
		(0.009)		(0.009)
Financial Product Score		0.175***		0.157***
		(0.059)		(0.059)
Propensity to Save Score		0.046		0.036
		(0.030)		(0.030)
Risk Tolerance Score		-0.000		0.001
		(0.030)		(0.030)
Confidence Score		0.137		0.121
		(0.102)		(0.103)
Parents' Level of Education			-0.008	-0.007
			(0.008)	(0.008)

Information Source (informal)			0.043	0.037
			(0.029)	(0.029)
Information Source (formal)			0.087***	0.073***
			(0.028)	(0.028)
Constant	0.267***	0.232**	0.229***	0.219**
	(0.071)	(0.106)	(0.079)	(0.108)
Observations	279	279	277	277
R-squared	0.138	0.181	0.173	0.206

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6 Panel b
Regression Analysis of the Financial Literacy Score by Gender. Results reported are for female students.

	(5)	(6)	(7)	(8)
VARIABLES	Financial Literacy Score			
Age	0.050***	0.029*	0.047***	0.026*
	(0.015)	(0.016)	(0.015)	(0.016)
Year	0.021	0.017	0.020	0.017
	(0.016)	(0.015)	(0.016)	(0.016)
Program	0.014	0.001	0.018	0.005
	(0.021)	(0.021)	(0.022)	(0.021)
Major Declared	0.061*	0.046	0.059*	0.044
	(0.033)	(0.032)	(0.034)	(0.033)
Money Sentiment		-0.018**		-0.018**
		(0.009)		(0.009)
Financial Product Score		0.211***		0.207***
		(0.059)		(0.060)
Propensity to Save Score		0.031		0.028
		(0.030)		(0.030)
Risk Tolerance Score		0.010		0.012
		(0.030)		(0.030)
Confidence Score		0.253**		0.261**
		(0.101)		(0.103)
Parents' Level of Education			-0.007	-0.007
			(0.009)	(0.008)
Information Source (informal)			0.028	0.020
			(0.030)	(0.029)
Information Source (formal)			0.049*	0.028
			(0.029)	(0.028)

Constant	0.209***	0.164	0.196**	0.166
	(0.073)	(0.106)	(0.082)	(0.109)
Observations	279	279	277	277
R-squared	0.096	0.177	0.108	0.185

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6 Panel c
Regression Analysis of the Lusardi-Mitchell Score by Gender. Results reported are for female students.

	(9)	(10)	(11)	(12)
VARIABLES	Lusardi-Mitchell Score			
Age	0.031	0.007	0.028	0.007
	(0.021)	(0.023)	(0.022)	(0.023)
Year	0.007	0.003	0.004	0.000
	(0.022)	(0.022)	(0.023)	(0.023)
Program	0.018	0.005	0.021	0.008
	(0.030)	(0.030)	(0.030)	(0.031)
Major Declared	0.078*	0.066	0.071	0.060
	(0.047)	(0.047)	(0.047)	(0.048)
Money Sentiment		-0.013		-0.012
		(0.013)		(0.013)
Financial Product Score		0.247***		0.234***
		(0.086)		(0.086)
Propensity to Save Score		0.026		0.028
		(0.043)		(0.044)
Risk Tolerance Score		0.008		0.006
		(0.043)		(0.044)
Confidence Score		0.200		0.184
		(0.147)		(0.150)
Parents' Level of Education			0.000	0.000
			(0.012)	(0.012)
Information Source (informal)			0.026	0.017
			(0.042)	(0.042)
Information Source (formal)			0.099**	0.079*
			(0.041)	(0.041)
Constant	0.302***	0.265*	0.253**	0.235
	(0.103)	(0.154)	(0.115)	(0.158)
Observations	279	279	277	277
R-squared	0.031	0.075	0.052	0.090

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7
Regression Analysis of the Financial Literacy, Tax Literacy and Lusardi-Mitchell Scores by Gender. Results reported are for male students.

	(1)	(2)	(3)	(4)
VARIABLES	Tax Score			
Age	0.047***	0.051***	0.031*	0.037**
	(0.017)	(0.017)	(0.017)	(0.017)
Year	0.027	0.030*	0.025	0.026
	(0.018)	(0.018)	(0.017)	(0.017)
Program	0.072***	0.055**	0.049*	0.043*
	(0.025)	(0.026)	(0.026)	(0.026)
Major Declared	0.052	0.037	0.045	0.034
	(0.032)	(0.032)	(0.031)	(0.031)
Money Sentiment			-0.010	-0.010
			(0.008)	(0.008)
Financial Product Score			0.179***	0.162**
			(0.064)	(0.064)
Propensity to Save Score			0.041	0.035
			(0.030)	(0.030)
Risk Tolerance Score			0.041	0.040
			(0.028)	(0.028)
Confidence Score			0.279***	0.262**
			(0.103)	(0.103)
Parents' Level of Education		0.028***		0.020***
		(0.008)		(0.008)
Information Source (informal)		0.065*		0.051
		(0.034)		(0.033)
Information Source (formal)		0.015		-0.011
		(0.029)		(0.029)
Constant	0.033	-0.091	-0.092	-0.175
	(0.083)	(0.093)	(0.108)	(0.113)
Observations	314	309	310	308
R-squared	0.102	0.142	0.185	0.210

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7 Panel b
Regression Analysis of the Financial Literacy Score by Gender. Results reported are for male students.

	(5)	(6)	(7)	(8)
VARIABLES	Financial Literacy Score			
Age	0.038**	0.038**	0.024	0.026*
	(0.015)	(0.015)	(0.015)	(0.015)
Year	0.036**	0.041**	0.027*	0.030*
	(0.016)	(0.016)	(0.015)	(0.016)
Program	0.098***	0.084***	0.075***	0.072***
	(0.023)	(0.024)	(0.023)	(0.023)
Major Declared	0.053*	0.040	0.047*	0.038
	(0.029)	(0.029)	(0.028)	(0.028)
Money Sentiment			-0.027***	-0.027***
			(0.007)	(0.007)
Financial Product Score			0.188***	0.176***
			(0.057)	(0.057)
Propensity to Save Score			0.019	0.016
			(0.027)	(0.027)
Risk Tolerance Score			0.004	0.001
			(0.025)	(0.025)
Confidence Score			0.245***	0.234**
			(0.092)	(0.093)
Parents' Level of Education		0.020***		0.015**
		(0.007)		(0.007)
Information Source (informal)		0.010		-0.005
		(0.031)		(0.029)
Information Source (formal)		0.029		0.011
		(0.027)		(0.026)
Constant	0.131*	0.072	0.154	0.108
	(0.075)	(0.085)	(0.097)	(0.102)
Observations	314	309	310	308
R-squared	0.144	0.155	0.239	0.251

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7 Panel c.
Regression Analysis of the Lusardi-Mitchell Score by Gender. Results reported are for male students.

	(9)	(10)	(11)	(12)
VARIABLES	Lusardi-Mitchell Score			
Age	0.021	0.021	0.009	0.010
	(0.020)	(0.020)	(0.020)	(0.021)
Year	0.051**	0.052**	0.039*	0.043**
	(0.021)	(0.021)	(0.021)	(0.021)
Program	0.083***	0.065**	0.062**	0.059*
	(0.030)	(0.032)	(0.031)	(0.032)
Major Declared	0.025	0.019	0.029	0.020
	(0.038)	(0.039)	(0.038)	(0.038)
Money Sentiment			-0.024**	-0.025**
			(0.010)	(0.010)
Financial Product Score			0.207***	0.189**
			(0.078)	(0.079)
Propensity to Save Score			-0.045	-0.047
			(0.037)	(0.037)
Risk Tolerance Score			-0.021	-0.028
			(0.035)	(0.035)
Confidence Score			0.146	0.128
			(0.126)	(0.127)
Parents' Level of Education		0.015		0.012
		(0.009)		(0.010)
Information Source (informal)		0.029		0.021
		(0.041)		(0.040)
Information Source (formal)		0.055		0.044
		(0.035)		(0.035)
Constant	0.303***	0.250**	0.417***	0.359**
	(0.098)	(0.112)	(0.132)	(0.140)
Observations	314	309	310	308
R-squared	0.072	0.076	0.114	0.124

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Conclusion

Financial literacy and tax literacy are essential areas of research as low levels of financial literacy and tax literacy can negatively affect the well-being of individuals and the economy as a whole. Although financial literacy has been studied extensively and numerous programs have been implemented to improve financial literacy, the financial literacy of the general population and post-secondary students remains very low. On the other hand, tax literacy is a relatively new construct, and very few studies have explored the level of tax literacy in the population. Even though very little research has been done on the tax literacy of the general population and the tax literacy of students, the few studies conducted all indicate that tax literacy levels are deficient (Chardon et

al., 2016; Pham et al., 2020). This lack of tax literacy has negative consequences for the Canadian economy. Lack of knowledge on what sources of income are subject to tax can lead to increased financial hardship, resulting in reduced consumer spending and economic retraction. In addition, a poor understanding of ones' country's tax system can lead to the misconception that the system is not fair. It can result in increased tax avoidance and evasion instances, which costs the government significant amounts in terms of time and resources dedicated to tax compliance (Alm et al., 2010). Lastly, low tax literacy can lead to the under-utilization of government benefits programs, such as old age security and pension plan benefits, as individuals may not understand how to access government funding by filing accurate income tax returns (Godbout et al., 2017).

This study aims to determine the financial literacy and tax literacy of post-secondary students and to determine factors that impact this group's financial literacy and tax literacy. Overall results of this study indicate that post-secondary students' financial literacy and tax literacy levels are very low. Slightly over half of the financial literacy questions in our survey were answered correctly by students. These results are worrisome as several Canadian provinces (notably Ontario and Alberta) have introduced financial literacy topics into the curriculum of high school students (Arthur, 2012). It could be that, throughout their post-secondary education, students have forgotten many of the essential financial literacy concepts covered in high school. A study conducted by Mandell and Klein (2009) showed that students did not retain personal finance knowledge five years after a high school level course. Students may not have understood or held personal finance concepts from high school, because they did not need this information at the time it was taught. The most appropriate timing for financial literacy education is when students recognize an immediate need for personal finance concepts, which could also improve the preservation of this information in students' memory (Mandell & Klein, 2009). This theory is supported by the results from a study by Peng et al. (2007), who administered a financial literacy survey on alumni of a large midwestern university. Their study found that students who took a personal finance course in college performed better on financial literacy survey questions than those who took a personal finance course in high school and those who took a personal finance course in both college and high school.

Consistent with previous research, this study found that the demographic characteristics of age, gender and parent’s education level are all positively related to financial literacy and tax literacy. Males performed better on financial literacy and tax literacy questions than females, but the difference in scores between males and females is smaller for the tax literacy questions. Other studies have found that personality traits such as risk tolerance and confidence also impact the financial literacy and tax literacy of students (Tokar Asaad, 2015, Kramer, 2016). Although this study found that confidence was positively related to financial and tax literacy, we found no relationship between risk tolerance and the student’s financial and tax literacy levels. Experience with financial products positively impacts financial, and tax literacy as students who own more financial products tend to perform better on both the tax and financial literacy questions. Also consistent with previous research, education is found to impact financial and tax literacy (Chen & Volpe, 1998; Lusardi et al., 2010; Pham et al., 2020). Students who have completed more years of post-secondary education or have accessed formal sources of financial and tax information performed better on the financial literacy and tax literacy questions. This study found that the year of study, major and use of formal sources of information are positively related to financial and tax literacy.

The results of this study indicate that providing additional education, whether in the form of formal education such as a course or a seminar is positively related to students' financial literacy and tax literacy. Future research should be conducted into which form of formal education has the

most significant impact on post-secondary students' financial and tax literacy and how formal and informal education interact. Another area for future research is the impact of personal finance education of post-secondary students on their financial and tax literacy. The results of this study indicate that the overall level of financial literacy and tax literacy for post-secondary students is low. The development of education programs and further research are required to determine how best to improve post-secondary students' financial literacy and tax literacy.

References

Adult Literacy 2003. Retrieved from https://www.oecd.org/education/innovation-education/adultliteracy.htm

Albeerdy, M. I., & Gharleghi, B. (2015). Determinants of the financial literacy among college students in Malaysia. *International Journal of Business Administration, 6*(3), 15-24. https://doi.org/10.5430/ijba.v6n3p15

Alm, J., Cherry, T., Jones, M., & McKee, M. (2010). Taxpayer information assistance services and tax compliance behavior. *Journal of Economic Psychology, 31*(4), 577-586. https://doi.org/10.1016/j.joep.2010.03.018

Arthur, C. (2012). Financial literacy education *Financial Literacy Education* (pp. 1-12): Brill. https://doi.org/10.1007/978-94-6091-918-3_1

Atkinson, A., & Messy, F.-A. (2012). Measuring financial literacy: Results of the OECD/International Network on Financial Education (INFE) pilot study. https://www.oecd-ilibrary.org/finance-and-investment/measuring-financial-literacy_5k9csfs90fr4-en

Bhushan, P., & Medury, Y. (2013). Financial literacy and its determinants. *International Journal of Engineering, Business and Enterprise Applications, 4*(2), 155-160.

Broadbent, J., Palumbo, M., & Woodman, E. (2006). The shift from defined benefit to defined contribution pension plans–implications for asset allocation and risk management. *Reserve Bank of Australia, Board of Governors of the Federal Reserve System and Bank of Canada, 1*, 54. https://www.researchgate.net/publication/242072140_The_Shift_from_Defined_Benefit_to_Defined_Contribution_Pension_Plans_-_Implications_for_Asset_Allocation_and_Risk_Management

Cameron, M. P., Calderwood, R., Cox, A., Lim, S., & Yamaoka, M. (2014). Factors associated with financial literacy among high school students in New Zealand. *International Review of Economics Education, 16*, 12-21.

Chardon, T., Lee, C., De Zwaan, L., & Liu, Y. (2016). University students and tax literacy: Opportunities and lessons for tax teaching. *Journal of the Australasian Tax Teachers Association, 11*(1), 85-102. https://static1.squarespace.com/static/5e88791f0c333872e0b33272/t/5ee61733095ea02582e2450d/1592137546120/JATTA-2016-All-Articles.pdf

Chaulk, B., Johnson, P. J., & Bulcroft, R. (2003). Effects of marriage and children on financial risk tolerance: A synthesis of family development and prospect theory. *Journal of Family and Economic Issues, 24*, 257-279.

Chen, H., & Volpe, R. P. (1998). An analysis of personal financial literacy among college students. *Financial Services Review, 7*(2), 107-128. https://doi.org/10.1023/a:1025495221519

Chen, H., & Volpe, R. P. (2002). Gender differences in personal financial literacy among college students. *Financial Services Review, 11*(3), 289-307. https://doi.org/10.1016/s1057-0810(99)80006-7

Droms, W. G. (1987). Investment asset allocation for PFP clients. *Journal of Accountancy, 163*(4), 114.

Ergün, K. (2017). Financial behaviour and financial literacy among university students. *Research in Economics and Business: Central and Eastern Europe, 9*(2), 77-94.

Fisher, P. J., & Yao, R. (2017). Gender differences in financial risk tolerance. *Journal of Economic Psychology, 61*, 191-202. https://doi.org/10.1016/j.joep.2017.03.006

Godbout, L., Genest-Grégoire, A., & Guay, J.-H. (2017). The knowledge deficit about taxes: who it affects and what to do about it. *CD Howe Institute Commentary, 484*, 1-15. https://www.cdhowe.org/sites/default/files/2021-12/Commentary%20484.pdf

Grable, J. E., & Joo, S.-H. (2004). Environmental and biophysical factors associated with financial risk tolerance. *Journal of Financial Counseling and Planning, 15*(1), 73-82.

Hallahan, T. A., Faff, R. W., & McKenzie, M. D. (2004). An empirical investigation of personal financial risk tolerance. *Financial Services Review, 13*(1), 57-78.

Hua, X., Huang, Y., & Zheng, Y. (2019). Current practices, new insights, and emerging trends of financial technologies. *Industrial Management & Data Systems, 119*(7), 1401-1410. https://doi.org/10.1108/IMDS-08-2019-0431

Kramer, M. M. (2016). Financial literacy, confidence and financial advice seeking. *Journal of Economic Behavior & Organization, 131(Part A)*, 198-217. https://doi.org/10.1016/j.jebo.2016.08.016

LaBorde, P. M., & Mottner, S. (2016). Can a College-Level Personal Finance Course Close the Knowledge Gaps? *Journal of Financial Education, 42*(1-2), 35-55. https://www.jstor.org/stable/90000835

Lusardi, A. (2015). Financial literacy: Do people know the ABCs of finance? *Public Understanding of Science, 24*(3), 260-271. https://doi.org/10.1177/0963662514564516

Lusardi, A. (2019). Financial literacy and the need for financial education: evidence and implications. *Swiss Journal of Economics and Statistics, 155*(1), 1-8. https://doi.org/10.1186/s41937-019-0027-5

Lusardi, A., & Mitchell, O. S. (2007). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics, 54*(1), 205-224. https://doi.org/10.1016/j.jmoneco.2006.12.001

Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *American Economic Journal: Journal of Economic Literature, 52*(1), 5-44. https://doi.org/10.1257/jel.52.1.5

Lusardi, A., & Mitchell, O. S. (2017). How ordinary consumers make complex economic decisions: Financial literacy and retirement readiness. *Quarterly Journal of Finance, 7*(03), 1750008. https://doi.org/10.1142/s2010139217500082

Lusardi, A., Mitchell, O. S., & Curto, V. (2010). Financial literacy among the young. *Journal of Consumer Affairs, 44*(2), 358-380. https://doi.org/10.1111/j.1745-6606.2010.01173.x

Mandell, L., & Klein, L. S. (2009). The impact of financial literacy education on subsequent financial behavior. *Journal of Financial Counseling and Planning, 20*(1), 15-24.

Mitchell, O. S., & Lusardi, A. (2015). Financial literacy and economic outcomes: Evidence and policy implications. *The Journal of Retirement, 3*(1), 107. https://doi.org/10.3905/jor.2015.3.1.107

Peng, T.-C. M., Bartholomae, S., Fox, J. J., & Cravener, G. (2007). The impact of personal finance education delivered in high school and college courses. *Journal of Family and Economic Issues, 28*, 265-284. https://doi.org/10.1007/s10834-007-9058-7

Pham, A., Genest-Grégoire, A., Godbout, L., & Guay, J.-H. (2020). Tax literacy: a Canadian Perspective. *Canadian Tax Journal, 68*(4), 987-1008. https://doi.org/10.32721/ctj.2020.68.4.pham

Thomas, A., & Spataro, L. (2018). Financial literacy, human capital and stock market participation in Europe. *Journal of Family and Economic Issues, 39*, 532-550. https://doi.org/10.1007/s10834-018-9576-5

Tokar Asaad, C. (2015). Financial literacy and financial behavior: Assessing knowledge and confidence. *Financial Services Review, 24*(2), 101-117. https://doi.org/10.61190/fsr.v24i2.3236

Van de Venter, G., Michayluk, D., & Davey, G. (2012). A longitudinal study of financial risk tolerance. *Journal of Economic Psychology, 33*(4), 794-800. https://doi.org/10.1016/j.joep.2012.03.001

Van Rooij, M. C., Lusardi, A., & Alessie, R. J. (2012). Financial literacy, retirement planning and household wealth. *The Economic Journal, 122*(560), 449-478. https://doi.org/10.1111/j.1468-0297.2012.02501.x

Acknowledgments

We are grateful for the financial support from AIMCO Foundation for Financial Education, CPA Education Foundation of Alberta, and MacEwan University – School of Business.

Enhancing Exam Assessment by Weighting Exam Questions Using Asset Pricing Principles

David L. Stowe
Ohio University

John D. Stowe
Ohio University

For most exams, question weights are specified in advance, often with equal weights for all questions. Unfortunately, the values that questions have for assessing student mastery of the underlying content are not equal (or proportional to pre-specified weights). Exam design has analogs to investment portfolio design. For decades, we have explored portfolio performance where risks are not proportional to portfolio weights and asset returns depend on systematic risks rather than total risk. We recommend adjusting question weights based on their systematic contributions to exam dispersion. Betas can be estimated for exam items much like betas are estimated for common stocks with the CAPM. When adjusting question weights to match their betas, we demonstrate that an exam has better psychometric properties, is more efficient, and provides better measurement of individual student performance. Examination time and resources, like everything else, are scarce commodities, and we advocate a method of weighting exam questions to enhance the information produced by an exam.

Keywords: Exam reliability, exam discrimination, asset pricing models, exam efficiency, CAPM

Introduction

When returning exams to a class, we have never had a student raise their hand and ask: "Professor, What is the beta for question 1? How much credit should I receive if I get question 1 correct? You use the CAPM for everything else. Why don't you use it to score our exams?"

Great questions. In capital markets, we assign a risk premium based on the systematic risk of an asset. On an exam, the same logic applies—assign credit based on a question's systematic dispersion. Our discipline has made immense contributions to capital markets, and it is time to apply asset-pricing principles to one of our mundane professorial tasks, assessing student performance on examinations. It turns out that this is easy to do.

Most exams use prespecified question weights. For large multiple-choice exams, equal question weighting is the norm. In contrast, investment portfolios do not use equal weights for each investment. Just like investments, the value of each question should vary and depend on its properties. Using asset-pricing principles to assign question weights can substantially improve the psychometric properties of an exam.

Assume an equal-weighted multiple-choice exam has been administered and machine-scored and a statistical/psychometric analysis of the exam is in hand. At this point, we can calculate the betas for each question (the item standard deviation times the item-total correlation divided by the

exam standard deviation). Then we recalculate exam scores and exam statistics based on these new beta weights. (Equivalently, you can estimate betas for each question by regressing the scores on the question against the total scores on the exam.) Because these new weights are informative, the student scores are adjusted and the overall exam performance, as indicated by Cronbach's alpha (Cronbach, 1961, Furr, 2018, and Price, 2017), a common measure of exam reliability, increases substantially.

If you currently use machine scoring for objective exams, you simply paste a matrix of student scores into Excel and do the calculations described here. You could also design an Excel template or write a program in SAS or Python. The psychometric performance of the exam is greatly enhanced (exam alpha increases) and the discrimination of a given exam is increased to a level that normally would require a much longer exam to achieve.

We demonstrate the procedure with a short hypothetical exam (small number of students and questions) to highlight its properties. Then we present the analysis for a larger actual exam and discuss some of the implementation issues an instructor may have. The increases in discrimination across students and in the reliability of the overall exam come by reassessing the exam information that is already at hand.

We present two alternatives for scoring exams—item weighting and option weighting. With item weighting, a beta (score) is estimated for the correct answer to each question and each incorrect answer is scored as a zero. For option weighting, betas are estimated for all answers to each question, right and wrong. For example, if A is the correct answer, it might have a beta of 1.10. Incorrect answers B, C, and D might have betas of -0.10, -0.30, and -0.70. Incorrect answers are not equally bad and do not have equal penalties for choosing them. Both of these alternatives are explored in this paper. Item weighting increases the reliability and discrimination possible with equal-weighted exams. Option weighting goes even further, increasing the reliability and discrimination possible with item weighting.

Background

Financial models express the risk premium for a risky asset as a function of its systematic risk: risk premium = $\boldsymbol{r_i - r_f = \frac{\rho_{i,m}\sigma_i\sigma_m}{\sigma_m^2}\left(r_m - r_f\right) = \rho_{i,m}\sigma_i \frac{\left(r_m - r_f\right)}{\sigma_m}}$. For an exam, the value or weight of each question is based on its measure of systematic risk—either its beta ($\frac{\rho_{i,m}\sigma_i\sigma_m}{\sigma_m^2}$) or its $\boldsymbol{\rho_{i,m}\sigma_i}$. That basic intuition from asset pricing, rewarding systematic risks only, is applied to valuing or weighting the questions on an exam.

On an exam with equally weighted questions, the student gets the same reward, one point, for each correct answer. Instead of setting the value of such a question equal to 1, we recommend setting the value of a question proportional to its systematic dispersion, which we can define as its "item beta" (the slope coefficient from regressing an item score against the total exam score, which is the item standard deviation times the item-total correlation divided by the exam standard deviation). Some faculty may prefer to use a related concept, the "item reliability" (which is the product of the question standard deviation and the question item-total correlation). We focus on the item beta to shorten the paper and because of its intuitive simplicity and similarity to betas on financial assets. We argue that the reward for getting each question correct should be based on its item beta. We apply that intuition to show the effects on the overall psychometric properties of an exam and on relative student performance.

Of course, student assessment is only one component of the educational process. Faculty are responsible for curriculum design, its delivery, test design, and assignment of grades. Our focus in this paper, enhancing exam information about student performance, has value when these other responsibilities are also well done. Several researchers have studied the writing of effective multiple-choice items (see Haladyna (2004), Haladyna, Downing, & Rodriguez (2002), and Kehoe (1995)). Many universities and professional societies (law, medicine, and others) have workshops and guidelines for question writers. Ardjmand, Stowe, and Stowe (2020) apply psychometric and finance principles to a related problem—exam construction (selecting questions) based on investment portfolio principles. This paper takes the exam as given and re-weights the questions to enhance the information from that exam, its reliability and discrimination.

While the psychometric properties of an exam are useful, they should not be the sole focus of exam analysis. For example, Cronbach's alpha is considered a measure of exam reliability, capturing the ability of the exam to measure the degree of mastery of the underlying material across students. A low alpha may signal that the exam is of poor quality and does not provide a sound basis for assigning grades. On the other hand, specialists warn of trying to create an alpha that is too high. On a typical exam, many questions that fail to discriminate across students may be overweighted, and other questions with important content that do effectively discriminate may be underweighted. The methods we describe in this paper to reweight items based on their systematic dispersions can substantially improve exam reliability.

When varying question weights are used, the most common way to assign weights is expert opinion. The weights may be based on the time needed to answer a question, the difficulty or the importance of a question, or the cognitive levels required. In some cases, instructors or experts may be working at cross purposes, where one is trying to facilitate high scores or grades by using a lot of easy questions while another may be trying to gauge student mastery of difficult concepts. We recommend using asset pricing concepts as a guide and basing the possible credit for questions on their systematic dispersions.

Selecting question weights resembles the classic mean variance optimization problem defined long ago by Markowitz (1952). In practice, mean-variance portfolio optimization can be challenging because the solution, based on historical estimates, can result in a sparse number of active weights. This puzzle has been explored extensively by Michaud (1989), Fisher and Statman (1997), and many others. Our approach, which revises question weights around their original weights based on their betas (as in Sharpe's (1964) CAPM), does not go to mathematical extremes. In our approach, the insight the instructor uses to construct the exam is the starting point, and those weights are revised upward or downward to improve the exam's reliability.

A Basic Example

In this section, we use a simple example to demonstrate the method of weighting exam questions based on their item betas and show that the overall exam properties (Cronbach's alpha and other overall exam metrics) are substantially enhanced along with the measurement of relative student performance. The relative performance of individual students can change where some will score (and rank) higher than their equal-weighted scores. This can occur for two reasons. The adjusted scores have less idiosyncratic volatility or noise than the equal-weighted scores, reducing the randomness in student ranking. Additionally, students who correctly answer more of the highly valued questions will have higher scores. Correspondingly, of course, other students will have lower scores.

In Table 1, we present the raw scores for six students taking an exam with four equally weighted questions. The cells in the shaded area of the table show a score of 1 if a student got that question correct and a score of 0 if incorrect. The right-hand side of the table gives the total score for each student, ranging from 4 down to 1. No student had a total score of zero and no student left any questions blank. The average student score is 2.33. Student total scores are expressed as a decimal fraction in the rightmost column (1.00 = 100%, and the mean decimal score is 0.5833). The bottom rows show the *p*-value for each item (the proportion correct). The scores follow a binomial distribution and the mean (p-value), variance, standard deviation, and item-total correlation for each item are given.

Table 1
Analysis of Raw Scores (Equal-weighted)

Student	Item Scores (0,1)				Total Score (sum)	Total Score %
	Q1	Q2	Q3	Q4		
1	1	1	1	1	4	1.00
2	1	1	1	0	3	0.75
3	1	1	0	1	3	0.75
4	1	0	1	0	2	0.50
5	1	0	0	0	1	0.25
6	0	0	1	0	1	0.25
Sum	5	3	4	2	14	
Mean (*p*-value)	0.8333	0.5000	0.6667	0.3333	2.3333	0.5833
Variance	0.1389	0.2500	0.2222	0.2222	1.2222	0.0764
Standard deviation	0.3727	0.5000	0.4714	0.4714	1.1055	0.2764
Item-total correlation	0.5394	0.9045	0.2132	0.7462		
Item reliability*	0.2010	0.4523	0.1005	0.3518		
Beta**	0.7273	1.6364	0.3636	1.2727	4	
Number of items		4				
Mean total score (out of 4)		2.3333				
Variance total score		1.2222				
Standard deviation total score		1.1055				
Coefficient of variation		0.4738				
Sum (item variances)		0.8333				
Cronbach's alpha		0.4242				
Standard error of measurement		0.8389				

*Item reliability = Item standard deviation × Item-total correlation
**Item beta = Item std dev × Item-total correlation / Std dev total score (%)

When the item-total correlations are multiplied by the item standard deviations, this gives the item reliability for each item, as shown in the table. The item beta, discussed more extensively below, can be calculated in two ways. From the data in a typical machine-graded exam analysis, the item beta is estimated as the item standard deviation times the item-total correlation divided by the standard deviation of the exam total score (expressed as a decimal fraction). With the raw data on exam scores, the item beta also is estimated by regressing the students' scores on an item

against the exam total scores (expressed as a decimal fraction). Both procedures provide identical betas for each item. The item reliabilities and item betas are both shown in the table.

Summary statistics for the exam are also given in Table 1. The table shows the mean total score, variance, and standard deviation. The coefficient of variation (0.474) is the standard deviation of total scores divided by its mean, which shows the relative amount of score dispersion. Cronbach's alpha, the coefficient of reliability, is the proportion of the total variance that is systematic, or not idiosyncratic. The formulas for Cronbach's alpha are discussed in the Appendix. The reliability index of the exam is 0.424. The standard error of measurement, showing the amount of the standard deviation of the total score that is unsystematic, is 0.839. The standard error of measurement = sqrt(variance total score)×sqrt(1 – Cronbach's alpha).

At this point, we abandon the assumption of equal-weighted questions and use the item beta for each item as its weight. Instead of giving each answer a weight of 1 (out of 4), we replace the correct score for each item (for each student) by its beta.

The item beta is the same beta used in the capital asset pricing model: using $\boldsymbol{\sigma_x^2}$ and $\boldsymbol{\sigma_x}$ for the variance and standard deviation of the exam total scores (in decimal form), $\boldsymbol{\sigma_i}$ for the item standard deviation, and $\boldsymbol{\rho_{ix}}$ for the item-total correlation, the item beta is $\boldsymbol{\beta_i} = \frac{\boldsymbol{\rho_{ix}\sigma_i\sigma_x}}{\boldsymbol{\sigma_x^2}} = \frac{\boldsymbol{\rho_{ix}\sigma_i}}{\boldsymbol{\sigma_x}}$. The item beta is simply the item reliability ($\boldsymbol{\rho_{ix}\sigma_i}$) times a constant, which for the example in Table 1 is $(\boldsymbol{1/\sigma_x}) = \boldsymbol{1/0.2764 = 3.618}$. Each item beta is this constant (3.618) times the item reliability. For the example in Table 1, the highest item beta is for item 2 (1.6364) and the lowest is for item 3 (0.3636). As in the CAPM where a stock beta is estimated by regressing its returns against market returns, the item betas can also be estimated by regressing the student scores on each item on the student total scores on the exam, producing the same beta estimates.

In Table 2, the student receives the item's beta for each correct answer and 0 for each incorrect answer. The total score for each student is the sum of the item betas for the questions correct, shown in the column to the right in Table 2. The correlations between the adjusted scores for each item and the adjusted total scores will also change. Using beta-weighting instead of equal weighting provides more information about student and exam performance. The coefficient of variation increases from 0.474 in Table 1 to 0.684 in Table 2—the dispersion of total scores relative to the mean total score increases. Importantly, Cronbach's alpha is substantially increased to 0.596 compared to a lower 0.424 for equal weighting. Changing from equal-weighting the exam items to beta-weighting increases the exam reliability. While the exam total score standard deviation, coefficient of variation, and Cronbach's alpha are increasing, the standard error of measurement in this example also increased, but relatively less than these.

When changing from equal weighting to beta-weighting, the average score on the exam dropped from 2.333 (58.3%) in Table 1 to 2.091 (52.3%). This occurred because the easier questions had a lower beta and harder questions had a higher beta. A perfect score under both weighting systems was a score of 4 (100%). An instructor can always curve the scores on the beta-weighted exam to have the same average as the equal-weighted exam . In this case, an instructor could construct an average score of 2.333 (58.3%) in Table 2 by multiplying the item reliabilities by 4.038 instead of by 3.618. The standard deviation of total scores would also increase proportionally, while the coefficient of variation, and Cronbach's alpha would stay the same. With this curve, a perfect score increases from 4 (100%) to 4.464 (111.6%). Curving the scores by multiplying by a constant does not affect the reliability of the exam.

Table 2
Analysis of Beta-weighted Scores

Student	Item Scores (0, Beta)				Total Score (sum)	Total Score %
	Q1	Q2	Q3	Q4		
1	0.7273	1.6364	0.3636	1.2727	4.0000	1.0000
2	0.7273	1.6364	0.3636	0.0000	2.7273	0.6818
3	0.7273	1.6364	0.0000	1.2727	3.6364	0.9091
4	0.7273	0.0000	0.3636	0.0000	1.0909	0.2727
5	0.7273	0.0000	0.0000	0.0000	0.7273	0.1818
6	0.0000	0.0000	0.3636	0.0000	0.3636	0.0909
Sum	3.6364	4.9091	1.4545	2.5455	12.5455	3.1364
Mean	0.6061	0.8182	0.2424	0.4242	2.0909	0.5227
Variance	0.0735	0.6694	0.0294	0.3600	2.0468	0.0764
Standard deviation	0.2710	0.8182	0.1714	0.6000	1.4307	0.2764

Number of items	4				
Mean total score	2.0909				
Variance total score	2.0468				
Standard deviation total score	1.4307				
Coefficient of variation	0.6842				
Sum (item variances)	1.1322				
Cronbach's alpha	0.5958				
Standard error of measurement	0.9096				

The student total scores in Table 2 are based on item betas, while the equal-weighted scores in Table 1 include more idiosyncratic dispersion. Reducing this extra noise can change student scores and values of the test items. Each item originally had a weight of 1, while the reweighted scores have beta weights ranging from a high of 1.6364 for item 2 down to a low of 0.3636 for item 3. Student 1 made a 100% under in both weighting systems. Getting the valuable questions correct gave student 3 a substantially higher score compared to their equal-weighted score. Students 2 and 3 had an equal-weighted score of 3 (in Table 1), but student 3 has a much greater beta-weighted score than student 2 in Table 2. Similarly, in Table 2 students 5 and 6 had the same equal-weighted score, but student 5 has a higher beta-weighted score than student 6 by getting a more valuable question correct. Student 4 scored much lower under beta-weighting (27.3% versus 50%) because the questions she got correct had low values (low betas). When the items are reweighted based on their systematic variation, the assessment of each student's degree of understanding can change substantially.

In summary, the empirical benefits sought are:

- Increases (decreases) in the relative weights of the more (less) informative questions.
- An increase in Cronbach's alpha, the systematic variation of exam answers.
- Increase in the relative dispersion (coefficient of variation) of total exam scores.
- Increases in the item-total correlations of the more informative items.
- Instructor can curve the scores (rescaling the weights by a constant and changing the distribution of total scores), which does not change the psychometric properties above.

Results For a Live Exam

The analysis of an exam with 25 multiple-choice questions administered to 57 students in an introductory investments class further illustrates the application of our beta-weighting method.

Table 3
Analysis of Raw Scores (Equal Weights)

Question	Item Mean (p)	Item Variance	Item Std Dev	Item-total Correlation	Item Reliability	Item Beta
1	0.877	0.108	0.328	0.329	0.108	0.812
2	0.404	0.241	0.491	0.574	0.281	2.120
3	0.386	0.237	0.487	0.337	0.164	1.235
4	0.825	0.145	0.380	0.357	0.136	1.024
5	0.789	0.166	0.408	0.353	0.144	1.085
6	0.912	0.080	0.283	0.128	0.036	0.273
7	0.579	0.244	0.494	0.102	0.050	0.379
8	0.807	0.156	0.395	0.194	0.077	0.577
9	0.754	0.185	0.430	0.268	0.115	0.868
10	0.561	0.246	0.496	0.472	0.234	1.763
11	0.596	0.241	0.491	0.127	0.062	0.468
12	0.649	0.228	0.477	0.337	0.161	1.212
13	0.667	0.222	0.471	0.321	0.152	1.141
14	0.386	0.237	0.487	0.293	0.143	1.075
15	0.772	0.176	0.420	0.353	0.148	1.116
16	0.333	0.222	0.471	0.116	0.055	0.411
17	0.772	0.176	0.420	0.290	0.122	0.917
18	0.754	0.185	0.430	0.366	0.158	1.187
19	0.860	0.121	0.347	0.383	0.133	1.002
20	0.579	0.244	0.494	0.487	0.241	1.812
21	0.860	0.121	0.347	0.459	0.159	1.201
22	0.860	0.121	0.347	0.292	0.101	0.763
23	0.368	0.233	0.482	0.458	0.221	1.664
24	0.719	0.202	0.449	0.192	0.086	0.651
25	0.702	0.209	0.457	0.071	0.032	0.244
Sum	16.772	11.018			3.319	25
Mean	0.671					1

Table 3 shows summary data for the 25 questions, which are equally weighted, worth 1 point for a correct answer and 0 for an incorrect answer. For question 1, 0.877 of the students got the question correct, which gives an item variance of 0.108 and item standard deviation of 0.328. Item 1 had a correlation of 0.329 with the total exam score, so the item reliability (item standard deviation × item-total correlation) is 0.108. The item beta of 0.812 can be estimated by dividing the item reliability by the exam total score standard deviation (which was 0.1328). Item betas can also be calculated by regressing item scores against total exam scores. Using beta weights, the student should receive 0.812 points for a correct answer and 0 for an incorrect answer. The same statistics are shown for the other 24 items in the table. The item reliabilities across items range from 0.032 to 0.281 and the item betas range from 0.244 to 2.120. The item betas sum to 25 (the number of questions) and average 1.0. If the item betas are estimated with 25 simple regressions,

the slope coefficients are the item betas and the intercepts (not shown) sum to zero and average zero. The sum of the item means is 16.772, which is 67.09% out of the 25 possible. The item beta weights for this exam are a constant 1 / 0.1328 = 7.532 times the item reliabilities (where 0.1328 is the standard deviation of exam total scores).

For the two weighting systems, equal weights versus beta weights, the question weights for each item are shown in Table 4, along with the mean student score for each item for the two weighting systems. The weights for equal weighting had a mean of 1 and, obviously, a standard deviation of zero, and the weights for beta weighting had a mean of 1 and a standard deviation of 0.477. If they were normally distributed, the beta weights would differ from equal weights by at least 0.477 almost 32% of the time. The mean points scored for each item also differ substantially for the two systems, as the table shows.

Table 4
Weights for the Two Weighting Systems

Question	Equal Weighted		Beta Weighted	
	Weights	Mean score	Weights	Mean Score
1	1	0.877	0.812	0.712
2	1	0.404	2.120	0.855
3	1	0.386	1.235	0.477
4	1	0.825	1.024	0.844
5	1	0.789	1.085	0.857
6	1	0.912	0.273	0.249
7	1	0.579	0.379	0.220
8	1	0.807	0.577	0.466
9	1	0.754	0.868	0.655
10	1	0.561	1.763	0.990
11	1	0.596	0.468	0.279
12	1	0.649	1.212	0.787
13	1	0.667	1.141	0.761
14	1	0.386	1.075	0.415
15	1	0.772	1.116	0.861
16	1	0.333	0.411	0.137
17	1	0.772	0.917	0.708
18	1	0.754	1.187	0.895
19	1	0.860	1.002	0.861
20	1	0.579	1.812	1.049
21	1	0.860	1.201	1.033
22	1	0.860	0.763	0.656
23	1	0.368	1.664	0.613
24	1	0.719	0.651	0.468
25	1	0.702	0.244	0.171
Total	25	16.772	25	16.019
Mean	1	0.671	1	0.641
Std Dev	0		0.477	
Maximum	1		2.120	
Minimum	1		0.244	

Table 5 summarizes the overall exam properties for these weighting systems. The purpose of re-valuing or reweighting the questions is to improve the exam properties over the equal-weighted exam. Cronbach's alpha increased (from 0.5931 to 0.6693), which means that there is more systematic variation relative to noise in the total scores (and grades) after reweighting.

Table 5
Exam Properties for Two Weighting Systems

	Equal Weighted	Beta Weighted
Number of items	25	25
Mean total score	16.772	16.019
Variance	11.018	17.859
Standard deviation	3.319	4.226
Coefficient of variation	0.198	0.264
Sum (item variances)	4.744	6.383
Cronbach's alpha	0.5931	0.6693
Std error of measurement	2.117	2.430

Table 6
Example Student Report (Total Score Based on Betas for Correct Answers)

Name	Arnold Jones
Title	Fin 300 First Exam
Date	April 1, 2024

Question	Correct	Incorrect	Value
1	B		0.812
2	C		2.120
3	B		1.235
4	D	A	0
5	C		1.085
6	C		0.273
7	D		0.379
8	C		0.577
9	B	D	0
10	B		1.763
11	B	D	0
12	B	D	0
13	D	C	0
14	C		1.075
15	C		1.116
16	B	A	0
17	C		0.917
18	A		1.187
19	D		1.002
20	A		1.812
21	B		1.201
22	C	D	0
23	D		1.664
24	B		0.651
25	C		0.244
Total			19.113

19.113	Total value of correct answers
25	Number of questions on test
76.45%	Your percentage score

Question values for a correct answer typically range from 0.5 to 1.5, depending on how much value each question contributes to relative exam performance. Question values center around 1.0.

Table 6 presents an example of a report that each student might receive for their exam. Without attempting to explain all details, the instructor might simply comment that the more challenging and interesting questions, determined statistically, have a value above 1, while other questions with a low relation to exam performance tend to have a value below 1. All questions count, but some are simply more valuable than others. When a student misses a question, their incorrect answer is indicated in the report, and they receive 0 points for that question.

Implementation Issues

The purpose of weighting the questions based on their systematic dispersions (betas) is to enhance the information about overall exam performance, performance of individual questions, and relative performance of individual students compared to the same exam with equal-weighted questions. There are several benefits and costs associated with beta-weighting questions to produce more reliable exams that are discussed here.

Test Length and Reliability

By using beta-weighting instead of equal-weighting for the items on the live test above, the reliability of the test (Cronbach's alpha) increased from 0.5931 to 0.6693. Is this an important increase? A traditional way to increase exam reliability is to increase the length of an exam. The Spearman-Brown prophesy formula (Spearman, 1910; Brown, 1910) addresses the relation between test length and Cronbach's alpha:

$$\boldsymbol{\alpha^{new} = \frac{m\alpha^{old}}{1 + (m-1)\alpha^{old}}}$$

where *m* is the test length of the new test divided by that of the old test that would be required to generate the Cronbach's alphas of the old and new tests. For our alpha increase from 0.5931 to 0.6693, *m* would be 1.389, requiring an increase in test length of 38.9%. For our test of 25 items, to reach this increased alpha, with equal-weighting, would require a new test of 34.7 items—roughly 10 additional items of the same quality as the existing items. Of course, this would also require that students have the time and ability to take a test that is 39% longer.

The exam characteristics will determine the benefit from the increased alpha from using systematic weighting. Four other examples of this are in Table 7.

Table 7
Increase in Effective Test Size from Beta-weighting

Test	Number of Items	Cronbach's alpha		Value of m	Effective increase in items	Effective % increase in length
		Equal-weighted	Beta-weighted			
1	25	0.5585	0.5960	1.166	4.2	16.6%
2	20	0.7154	0.7351	1.104	2.1	10.4%
3	27	0.6687	0.8331	2.472	39.8	147.2%
4	20	0.7027	0.8485	2.370	27.4	137.0%

The first two tests in Table 7 have a modest increase in alpha, and an increase in effective test length of only about 16% and 10%. The last two tests had larger increases in alpha which resulted in dramatic increases in effective test length of 147% and 137%. The live example with beta-

weighting had an effective length 39% greater than the equal-weighted exam. If an exam has homogeneous systematic dispersions (the betas were all identical and the item standard deviations and item-total correlations were also all identical), there would be no gain from re-weighting of the equally weighted items. The resulting m = 1 and there is no effective increase in test length. As a practical matter, test items will not be homogeneous. There will be differences in *p*-values, item standard deviations, and item-total correlations creating a dispersion of item betas and there will be opportunities to enhance exam reliability through reweighting.

Reweighting the questions, reducing the weights for low value questions and increasing the weights for higher value questions, achieves an increase in reliability at low cost. On the other hand, increasing exam length (and sticking with equal weighting) would incur the costs associated with designing, administering, and grading longer exams. The amount of time available for testing may need to be increased. For high-stakes exams such as the CFA exam or the CPA exam, the curriculum is extensive and testing time is an extremely scarce commodity. Making the best use of limited testing time is a high priority. For any rigorous college class, making better use of the available testing time is desirable.

Duhachek and Iacobucci (2004) describe Cronbach's alpha as an increasing function of the number of items and the mean correlation among items and not related to sample size and covariance heterogeneity. They also show that the alpha standard error is decreasing with sample size, number of items, and the mean correlation among items and an increasing function of covariance heterogeneity. Increasing the effective test length would increase alpha and increase its significance. Fortunately, the cost of beta-weighting instead of equal weighting may be modest intellectually and administratively.

Using "Corrected" Item-total Correlations

Most exam analysis software packages present the item-total correlation between an item's performance and the exam total scores that include the item's score. Of course, it is possible to calculate the item-total correlation between an item's performance and total exam scores with the item removed, which can be termed the "corrected" item-total correlation. The corrected item-total correlations will be lower as will be "corrected" item betas. This will result in lower exam scores, where a perfect exam will have the sum of the betas equal less than the number of questions (and a perfect score is less than 100%). You can multiply the betas by a constant adjustment factor to correct this bias. Just as with the beta-weighting system (uncorrected), Cronbach's alpha using the "corrected" scores is also substantially greater than the equal-weighted Cronbach's alpha. Either weighting system lowers (increases) the weights for weaker (stronger) questions and enhances the overall exam properties. We present an analysis with the uncorrected correlations because most canned test packages do not use corrected correlations and because the interpretation of the betas is simpler.

Computer Adaptive Testing

Computer adaptive testing (CAT) alters an exam in real time (while the student is taking the exam) to achieve a more accurate estimate of a student's mastery of a curriculum. With a limited amount of test time, CAT attempts to achieve the reliability of a longer test. Of course, CAT is useful only for large-scale tests that are computerized. Our procedure achieves increased reliability not by changing the test, but by reweighting the items on the test based on their systematic

dispersions. This more accurate ranking of students is possible for even small test administrations and modest increased resources, much unlike CAT.

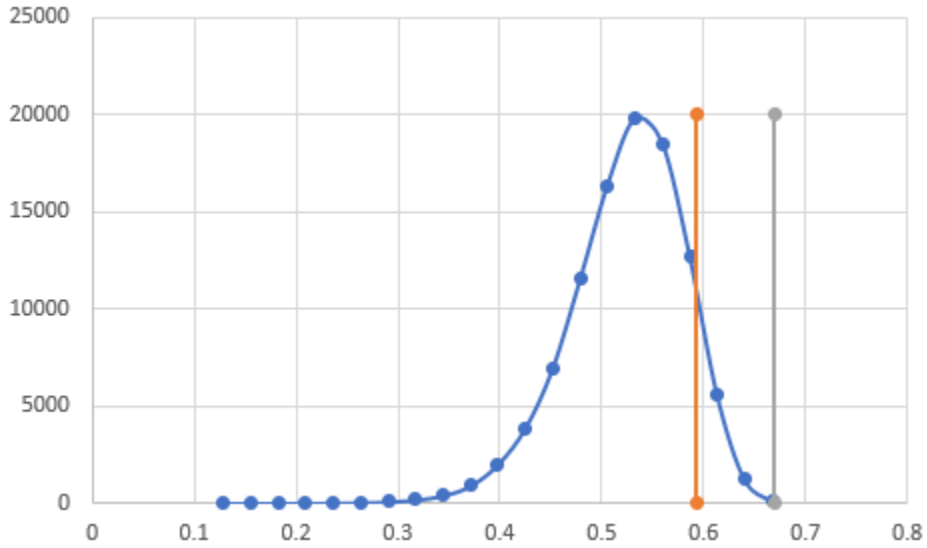
A Simulation Experiment

For perspective on question weights chosen based on their empirical systematic dispersions (betas), we conducted a simulation experiment. We randomly chose a set of 25 question weights for the 25 questions drawn from a uniform distribution and rescaled as needed so that the weights sum to 25. We applied these weights to the data from the 57-student class and calculated the exam properties including Cronbach's alpha (and coefficient of variation and standard error of the estimate). Iterating this process 100,000 times, the distribution of the simulated Cronbach alphas is shown in Table 8 and Figure 1.

Table 8
Simulated Cronbach alphas

Range (bottom of bin)	Frequency
0.1285	1
0.1555	0
0.1824	2
0.2094	4
0.2363	11
0.2633	32
0.2903	65
0.3172	173
0.3442	390
0.3711	895
0.3981	1992
0.4250	3850
0.4520	6930
0.4789	11513
0.5059	16351
0.5328	19772
0.5598	18502
0.5867	12675
0.6137	5542
0.6406	1220
0.6676	80

Figure 1
Frequency Distribution of Simulated Cronbach alphas



Compared to the Cronbach's alpha of 0.6693 using beta weighting for the 25 questions for question weights, none of the 100,000 simulated outcomes reached this alpha level. Several were close. Eighty outcomes (0.08%) were within 0.0017 and 0.0287 of the ideal. Most were far below, much like the equal-weighted outcome is far below. The assumed dispersion for the simulated weights was actually greater than that of the beta weights, which should have made it easier for a randomly chosen set of weights to dominate the beta-weighted rule. The beta weights performed well compared to a very large number of randomly chosen exam weights. Equal weights and a large number of randomly chosen weights were consistently inferior.

Equal Original Weights Are Not Required

Although the examples in this paper were for equally weighted questions that were graded right or wrong (binary), these are not necessary assumptions. The approach of weighting questions based on their systematic dispersions can work perfectly well an exam with questions of differing original weights and for questions that have partial credit. For example, one question could be binary and worth one point, and another question could be worth up to eight points, including any value between 0 and 8. For machine graded exams (which are usually equally weighted multiple-choice exams), you can paste the matrix of student question scores into a template and submit. For exams that are not machine graded (and can have a variety of question weights and scoring), the instructor would have the extra burden of putting the question scores into the matrix.

The beta-weighting system reflects the structure that the instructor used when designing the test. The item betas depend on the item standard deviations and item-total correlations (and total exam standard deviation). Those items with large values for these two statistics will have large betas, and those with low values with have small betas. For example, if the instructor had 25 equally weighted questions, the procedure would probably not create heavy new weights on a handful of questions and zero weights on the rest. The estimated beta weights will vary above and below 1 and sum to 25. The revised weights will increase the exam reliability (Cronbach's alpha) without abandoning the pedagogical goals of the instructor's original exam.

If the instructor assigned varying weights for the questions instead of equal weights, the variance-covariance matrix would reflect those weights. The final weights and makeup of the exam reflect both the instructor's priorities as well as the statistical behavior of the group taking the exam. Exam content (and instructor exam design) and statistical behavior jointly determine the beta weights for an exam that is more reliable and discriminates across students more efficiently than an equal-weighted system or a system of pre-specified weightings.

Our approach of re-weighting the questions on a test based on their systematic variation can be applied to various types of tests, such as equally weighted right-wrong multiple-choice tests, essay tests with unequal weights that include partial credit, or tests with a mixture of types of questions.

Beta-Weighting Each Item Option

A possible next step beyond weighting questions according to their covariances (betas) with total exam scores is to is to weight the incorrect options for each item based on their covariances (betas). In other words, if A is the correct answer, instead of assuming that B, C, and D have equal values (such as 0), we estimate betas for B, C, D which can vary considerably.

Abandoning the assumption that questions are equally weighted and weighting them based on their systematic dispersions (relative to the total exam score) allows substantial increases in the information available from an exam. At this point, we take the next step and also abandon the assumption that all incorrect answers should be weighted equally.

Test item option weighting has a history, as described in Claudy (1978). There are several scoring methods. Number right scoring assigns 1 point for a correct answer and 0 for incorrect answers or omits. Correction for guessing scoring assigns 1 point for a correct answer and -1/(N-1) points for incorrect answers, where N is the number of options; omits receive 0 points. Guttman weights scoring assigns a z-score to every choice (correct or incorrect) where the z-score is based on the mean score on all other test options besides the present one. Biserial weights scoring is based on the item-total correlations between an option (correct or incorrect) and the score on all other options on the test. Proportional weights scoring weights each option by the proportion of students in an upper scoring group (such as the top 25%) selecting each option.

Assigning variable penalties for each incorrect answer is a logical extension of our approach, which assigned a variable bonus for each correct answer and the same penalty for incorrect options that were selected. To fairly consider item option weighting, the weighting system we use (based on betas against total exam scores) is applied to all options, including correct and incorrect choices.

Incorrect answers are not created equal. Commonly, the correct answer is worth 1 point and there is no penalty for incorrect choices. For an item with four choices, with A the correct choice and B, C, and D as incorrect choices, the student would receive 1 point for A under equal weighting or beta points for A under beta-weighting, and 0 points for B, C, or D. When taking multiple-choice tests, a well-known test-taking strategy, when the student does not know the answer, is to try to eliminate one or two of the distractors and then select one of the remaining better answers. Exam writers are trained to create distractors (wrong answers) that could be logical mistakes that a student could make when they don't have the right answer. The mistakes that exam writers anticipate may reveal differing levels of knowledge or lack thereof. Clearly, exam takers and exam writers do not consider all incorrect answers to be equal.

To show the properties of estimating betas for all answers, correct and incorrect, we continue the example 25-question test with 4 choices per item. Table 9 presents the betas for each of the 25

items, with 25 beta estimates for the correct answers and 75 beta estimates for the three incorrect answers to each item.

Table 9
Using Betas for All Answers, Correct and Incorrect

Item	Estimated betas for each option					Mean score for item
	Highest			Lowest		
1	0.812	0	-0.022	-0.791		0.642
2	2.120	-0.242	-0.432	-1.445		0.167
3	1.235	-0.265	-0.309	-0.660		0.209
4	1.024	-0.101	-0.450	-0.472		0.776
5	1.085	-0.212	-0.242	-0.631		0.773
6	0.273	0.128	0.058	-0.460		0.237
7	0.379	0.006	-0.072	-0.313		0.175
8	0.577	0.058	0.018	-0.653		0.388
9	0.868	-0.163	-0.172	-0.534		0.569
10	1.763	-0.295	-0.414	-1.055		0.699
11	0.468	0.228	-0.255	-0.441		0.265
12	1.212	0	-0.546	-0.666		0.574
13	1.141	-0.022	-0.300	-0.819		0.534
14	1.075	0	-0.191	-0.884		0.042
15	1.116	-0.141	-0.150	-0.825		0.709
16	0.411	0.189	-0.166	-0.434		0.017
17	0.917	0.098	-0.340	-0.675		0.593
18	1.187	0	-0.110	-1.076		0.648
19	1.002	-0.132	-0.420	-0.450		0.816
20	1.812	-0.242	-0.626	-0.944		0.757
21	1.201	0	0	-1.201		0.864
22	0.763	-0.101	-0.212	-0.450		0.618
23	1.664	0.051	-0.242	-1.472		0.037
24	0.651	0	0	-0.651		0.285
25	0.244	0.049	0.009	-0.302		0.093
Sum	25	-1.110	-5.588	-18.303		11.486
Mean	1	-0.044	-0.224	-0.732		0.459
Std dev	0.477	0.138	0.184	0.316		0.274

In the table, the four betas for each item are presented in descending order of size. In this case, the highest betas for the 25 items were, indeed, for the correct answer for each item, although the highest beta for an item might not be for the correct answer if the question is poorly constructed or keyed incorrectly. The other three betas for each item are for the three incorrect choices.

The betas for the 25 correct answers are identical to those estimated previously and presented in Table 4. The average beta for the correct answers is 1 and their sum is 25 (the number of items). For each item in Table 9, the betas for the three incorrect answers sum to minus the beta for the correct answer, and the four betas together sum to zero. For example, if the correct beta is 1.30, the three incorrect betas could be values such as -0.70, -0.40, and -0.20.

The betas for each item (correct and incorrect) sum to zero, and the sum of all of the betas for the exam (100 in total) is equal to zero. The betas for the 25 correct answers sum to 25 and have an average of 1, while the betas for the 75 incorrect answers sum to -25 and average -0.333.

Previously, we argued that if the correct answers have widely varying covariances (and betas) with exam total scores, the correct answers should be weighted accordingly. Similarly, the incorrect answers to the questions also have widely varying covariances (and betas) with exam total scores, and ignoring this variation destroys a rich trove of information about student performance. Hence, instead of arbitrarily saying that each incorrect answer is worth 0 (or perhaps –1/3 using a correction for guessing), each incorrect answer should be valued based on its covariance or beta. Many incorrect answers may have a small beta, and occasionally a modest positive beta, while other incorrect answers can have larger, negative betas. The student's exam score is the sum of the betas for all of the questions that the student got correct, plus the sum of the betas for their incorrect choices (which are usually negative).

In Table 9, the mean and standard deviation of the betas for the 75 incorrect questions is -0.333 and 0.368, with maximum and minimum values of 0.228 and -1.472. Eight betas were 0, which occurs whenever that choice was never selected. Eleven of the incorrect betas were positive, and the majority (56) were negative. Clearly, assuming that all incorrect answers should be scored the same ignores the information in their wide-ranging covariances with exam total scores.

The rightmost column of Table 9 gives the mean score for each item, which is the average of the four betas, weighted by the proportion of students choosing each answer. The contribution of each item to the overall exam mean score depends on the beta values for the item and the distribution of the options selected by students. The overall mean score is 11.486 value-weighted questions, or 45.9% of the 25 possible.

The effects of beta weighting all item choices on overall exam statistics are summarized in Table 10. Compared to beta-weighting correct answers only, introducing betas for incorrect answers results in a decrease in the mean total score. The total score variance and standard deviation, coefficient of variation of the total score, and the sum of item/option variances all increase. However, because the covariances are increasing relatively more than the variances, Cronbach's alpha increases substantially. Weighting the items by their betas and introducing the weighting of incorrect answers based on their betas allows student exam scores to capture additional information about exam performance that was suppressed when using equal-weighted items and when assuming that all incorrect answers should be equally valued.

Table 10
Overall Statistics for Four Weighting Systems

	Weighting System			
	Equal weighted		Beta weighted	
	Correct only	Correct and penalty for incorrect	Correct only	Betas for correct and incorrect
Score for correct answers	1	1	Varying betas	Varying betas
Score for incorrect answers	0	-0.333	0	Varying betas
Number of items/options	25	25	25	100
Mean total score	16.772	14.029	16.019	11.486
Variance total score	11.018	19.589	17.859	47.139
Std dev total score	3.319	4.426	4.226	6.866
Coeff variation total score	0.198	0.315	0.264	0.598
Sum (item variances)	4.744	8.434	6.383	9.322
Cronbach's alpha	0.593	0.593	0.669	0.810
Std error of measurement	2.117	2.823	2.430	2.990

When the exam is scored with four different weighting schemes, the discrimination changes. Compared to equal weighting of correct answers only (1 for correct and 0 for incorrect), equal weighting of correct answers and a fixed penalty for incorrect answers (1 for correct, -0.333 for incorrect) reduces mean score and increases the dispersion of the total score. However, Cronbach's alpha is unchanged (at 0.593). Beta weighting the correct answers with no penalty for incorrect answers substantially improves Cronbach's alpha (to 0.669). Item weighting all answers (with betas for correct answers and betas for incorrect answers which are frequently negative) provides even greater discrimination across students, increasing Cronbach's alpha to 0.810.

Table 11 presents a student report where each student receives the beta values for their correct or incorrect responses to each of the 25 items. This student received various positive beta values for the 18 correct items and negative beta values for 6 of the 7 items missed. Item 11 was interesting because this student received 0.228 points for incorrect answer D. Correct answer B would have been worth 0.468 points (and incorrect answers A and C had negative betas). Incorrect answer D had a modest positive correlation to overall exam performance and hence had a positive beta and positive contribution to the student's score. This student received 19.113 points for the 18 correct answers and a net of -1.598 points for the 7 incorrect answers, for a total score of 17.515 (70.06% of the 25 points possible).

Table 11
Example of Student Report (Total Score Based on Betas for Correct and Incorrect Answers)

Name	Arnold Jones
Title	Fin 300 First Exam
Date	April 1, 2024

Question	Correct	Incorrect	Value
1	B		0.812
2	C		2.120
3	B		1.235
4	D	A	-0.450
5	C		1.085
6	C		0.273
7	D		0.379
8	C		0.577
9	B	D	-0.163
10	B		1.763
11	B	D	0.228
12	B	D	-0.546
13	D	C	-0.022
14	C		1.075
15	C		1.116
16	B	A	-0.434
17	C		0.917
18	A		1.187
19	D		1.002
20	A		1.812
21	B		1.201
22	C	D	-0.212
23	D		1.664
24	B		0.651
25	C		0.244
Total			17.515

17.515	Total point value
25	Number of questions on test
70.06%	Your percentage score

Question values for correct answers typically range from 0.5 to 1.5, depending on how much value each question contributes to relative exam performance. Correct question values center around 1.0.
Incorrect question values typically range from -0.1 to -0.5 and center around -0.333.

In the usual equal-weighted exam with 4 choices and no penalty for guessing, the expected student score with random guessing is 25%. In the beta-weighted exam with no penalties for guessing, the expected score with random guessing is also 25%. For an equal-weighted exam (with 4 options) and an adjustment for guessing, the score is right minus (1/3) wrong and the expected score is 0%. For beta-weighting all options (which has negative weights for most incorrect answers), the expected score for random guessing is also 0%. (If an instructor wanted to restore the gift from guessing without changing Cronbach's alpha, the student scores could be 25% plus 75% of the exam score based on correct and incorrect betas. With an average total score of 67.12%, the curved average score would be 25% + 0.75(67.12%) = 75.34%.)

Conclusion

Poor quality questions that are over-weighted relative to their values waste resources and introduce useless noise in exam scores. Likewise, high quality questions that are under-weighted cause useful information about student performance to be wasted. We introduce a simple method of re-weighting questions based on their systematic dispersions. This process discards some useless idiosyncratic variability resulting in greater discrimination and higher reliability coefficients. Much like investment portfolios where portfolio weights can be chosen to improve portfolio performance, item weights can be selected to improve exam performance measurement. And like asset pricing models such as the CAPM, where asset prices reflect systematic risk, item weights should reflect item systematic dispersions.

The process for an equally weighted multiple-choice exam would be to score and tabulate a set of student exams as if the questions were equal weighted. Then estimate item betas ($\frac{\rho_{ix}\sigma_i\sigma_x}{\sigma_x^2} = \frac{\rho_{ix}\sigma_i}{\sigma_x}$) and rescore all exams using item betas as question weights. Then distribute student reports (showing student item scores and total scores) and an exam analysis for instructors with summary statistics (such as mean, median, standard deviation and variance, Cronbach alpha, coefficient of variation, efficiency gain (equivalent length of weighted exam compared to equal weighted exam), and standard error of the estimate and the list of student scores).

The weighting system can be extended to use item and option weighting (where betas are estimated for both correct and incorrect answer choices). Both item-weighting and option-weighting provide additional information about student performance that is lost in the common equal-weighting of exam answers.

The revised exam weights will improve overall psychometric outcomes and provide greater discrimination across students. Of course, improved results will not occur when the test instrument is poor (e.g., all students scored 100%) or if the students are uniformly clueless. Generally, however, with substantive exams and with heterogeneous student populations, the beta-weighted exams can be much more efficient and informative than their naïve, equal-weighted counterparts.

Enhancing the analysis of exam results obviously does not change faculty obligations to design and deliver a high-quality curriculum and to produce exams that are based on learning objectives and fairly assess student mastery of that curriculum. Improving exam analysis by weighting exam questions based on their item betas (or scoring exam questions based on the betas of both correct and incorrect answer choices) can prove to a useful part of the educational enterprise by enhancing the reliability and discrimination considerably beyond what is achieved with prespecified question weights. Just as asset-pricing principles have contributed to investment practice and the functioning of capital markets, similar principles can be applied to enhance the useful information about exam performance.

References

Ardjmand, E., Stowe, D., & Stowe, J. (2020). Using Portfolio Theory to Design Better Exams. *Journal of Financial Education*, 46(2), 271–297.

Brown, W. (1910). Some Experimental Results in the Correlation of Mental Abilities. *British Journal of Psychology*, 3, 296–322.

Claudy, J. (1978). Biserial Weights: A New Approach to Test Item Option Weighting. *Applied Psychological Measurement*, 2(1), 25–30.

Cronbach, L. (1961). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 16(3), 297–334.

Duhachek, A., & Iacobucci, D. (2004). Alpha's Standard Error (ASE): An Accurate and Precise Interval Estimate. *Journal of Applied Psychology*, 89(5), 792–808.

Fisher, K., & Statman, M. (1997). The Mean-Variance-Optimization Puzzle: Security Portfolios and Food Portfolios. *Financial Analysts Journal*, 53(4), 41–50.

Furr, R. (2018). *Psychometrics: An Introduction* (3rd ed.). Sage Publications, Inc.

Kuder, G., & Richardson, M. (1937). The Theory and Estimation of Test Reliability. *Psychometrika*, 2(3), 151–160.

Haladyna, T. (2004). *Developing and Validating Multiple-Choice Test Items* (3rd ed.). Routledge.

Haladyna, T., Downing, S., & Rodriguez, M. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Applied Measurement in Education*, 15(3), 309–333.

Kehoe, J. (1995). Writing Multiple-Choice Test Items. *Practical Assessment, Research, and Education*, 4(9), 1–3.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, 7(1), 77–91.

Michaud, R. (1989). The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal*, 45(1), 31–42.

Price, L. (2017). *Psychometric Methods: Theory into Practice*. The Guilford Press.

Sharpe, W. (1964) Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, 19(3), 425–442.

Spearman, C. (1910). Correlation Calculated from Faulty Data. *British Journal of Psychology*, 3(3), 271–295.

Appendix
Background on Cronbach's alpha

Commercially available test analysis software provides statistics about the overall exam and about each separate test item. For the overall exam, summary statistics for the total exam scores include the mean and median scores and the variance and standard deviation. Other data may include the high and low scores, selected percentiles (25th and 75th), and skewness.

For us, the most important statistic is Cronbach's alpha, which shows how much of the total variance is systematic (as opposed to idiosyncratic). Cronbach's alpha is based on the variance-covariance matrix of exam scores. The variance of total scores on the exam ($\boldsymbol{s_x^2}$) can be broken into the sum of the covariances between items and the sum of the item variances:

$$\boldsymbol{s_x^2 = \sum\sum c_{ij} + \sum s_i^2}$$

where $\boldsymbol{c_{ij}}$ and $\boldsymbol{s_i^2}$ are the covariance and variance terms in the covariance matrix. In our work, we are using formulas for the population variance rather than the sample variance (dividing by n instead of n-1), although you will see either approach in use.

Cronbach's alpha is estimated as:

$$\boldsymbol{\alpha = \left(\frac{k}{k-1}\right)\left(\frac{\sum\sum c_{ij}}{s_x^2}\right) = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum s_i^2}{s_x^2}\right)}$$

where *k* is the number of questions on the exam. The factor $\left(\frac{k}{k-1}\right)$ sets the upper bound on $\boldsymbol{\alpha}$ as 1. The greater the item covariances relative to the item variances (and the greater the correlations among items), the greater is Cronbach's alpha and the systematic relative to idiosyncratic variance for an exam.

The Kuder-Richardson 20 (KR-20) calculates the exam alpha for binary (0,1) responses (Kuder & Richardson, 1937). For this case, Cronbach's alpha formula and the KR-20 formula are equivalent, and both can be used. For binary responses, the variance for an item is $\boldsymbol{s_i^2 = pq}$ where *p* is the probability of a correct answer and $q = 1-p$. The KR-20 formula is:

$$\boldsymbol{\alpha = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum pq}{s_x^2}\right)}$$

Since the KR-20 is a special case of Cronbach's alpha, we focus on the more general statistic and our comments are applicable to the KR-20.

Based on Cronbach's alpha, the standard error of measurement = sqrt(variance total score)×sqrt(1 – Cronbach's alpha):

$$\boldsymbol{SEM = \sqrt{s_x^2} \times \sqrt{1-\alpha}}$$

The SEM is inversely related to $\boldsymbol{\alpha}$, indicating how much of the exam standard deviation is idiosyncratic as opposed to systematic.

Statistics about each test item often include the p-value (% correct), item standard deviation, breakdown of choices (if there are four choices, how many students chose each one), and an item-total correlation. The item-total correlation may be computed with the item score included in the total score or with the item score removed from the total score. The former is more common.

Finally, an item discrimination statistic may be provided, which is the mean score on the item for the top x% of the students minus the mean score of the bottom x% of students.

Online Students' Perception of a Finance Project using Wharton Research Data Services (WRDS)

Oscar Solis
West Texas A&M University

Eric McKee
West Texas A&M University

The purpose of this study is to investigate undergraduate business students' perceptions of a hands-on research project utilizing the Wharton Research Data Services (WRDS) database. We developed and administered an online survey. Both quantitative and qualitative data were collected from students enrolled in three online introductory finance courses. The findings suggest that utilizing a WRDS database in a hands-on research project has positive implications for undergraduate students, finance faculty, and business schools. The paper also highlights the benefits and challenges of the WRDS platform and project for students and finance faculty teaching introductory finance courses.

Keywords: Finance, financial analysis, hands-on research project, online students, WRDS platform

Introduction

An integral component of the business school curriculum is finance courses, and most accredited business schools require students to complete at least one finance course (AACSB Business Standards, 2020; Campbell et al., 2006; Jones et al., 2022; Root et al., 2007). Introductory finance courses can be challenging for instructors, especially courses with a mixture of finance and non-finance majors. The challenges are further exacerbated when introductory courses are taught online (Jones et al., 2022) as opposed to face-to-face (F2F).

In the fall 2021 semester, the West Texas A&M University's Paul and Virginia Engler College of Business (PVECOB) gained access to the *Wharton Research Data Services* (WRDS) database. WRDS is widely used by researchers in finance, economics, and many other areas, and it also offers materials suitable for the classroom. The PVECOB School of Accounting, Economics, and Finance (AEF) has implemented a WRDS learning activity in numerous graduate and undergraduate courses. The AEF program's goal is to enhance learning opportunities in FIN 3320 *Business Finance* for finance and non-finance majors. An additional goal is to provide students with hands-on research experience with emerging technologies, which aligns with the standards and teaching practices promoted by the AACSB.

This study aims to explore undergraduate students' perceptions of the WRDS learning project in the introductory finance course FIN 3320 *Business Finance*. Additionally, we will identify the benefits and challenges that online students perceive while completing a hands-on research project using the WRDS platform. Lastly, we address students' perception of the WRDS project in relation to the course objectives.

Literature Review

Finance programs across the United States are offering more courses online and are enhancing the hands-on learning projects that are woven into their curricula (Batu et al., 2018; Sebastianelli & Tamimi, 2011). In a study by Sebastianelli and Tamimi (2011), students perceived that the most valuable features in quantitative online business courses are the encouragement of professor-student interaction and the use of customized videos. Additionally, Moreale & Zaynutdinova (2018) highlighted that students who engage in hands-on learning opportunities are able to more effectively transition from an academic to an industry environment.

The driving force in online learning is technology (Yu et al., 2021), which is consistently changing. Innovative and forward-thinking instructors are aware of the impact of technology on learning and they understand the way students learn is changing. Fortunately, online learning tools such as WRDS can positively impact student performance (Batu et al., 2018; Jones et al., 2022; Biktimirov & Klassen, 2008; Hadsell, 2009). Furthermore, Payne and Tanner (2011) found that finance courses that implement hands-on learning assignments using technology (e.g., WRDS) to solve finance problems are able to ensure that students graduate with the skills necessary to succeed in the workplace.

One of the most notable and recognized educational frameworks is Bloom's Taxonomy, a hierarchical classification of objectives and skills. The taxonomy has been widely applied to primary, secondary, and higher education practices (Krathwohl et al., 1964; Krathwohl, 2002). Practitioners in business and academic disciplines have routinely promoted higher levels of learning for students and the promotion of developing students' business acumen (Alshurafat et al., 2023; Betts, 2008; Hallows & White, 2016; Handy & Basile, 2005; Reeves, 1990; Shulman, 2002; Soffe, Marquardt & Hale, 2011). For our study, the following Bloom's Taxonomy categories include how the FIN 3320 *Business Finance* project promotes a high level of learning: (1) Knowledge: the use of WRDS platform; (2) Comprehension: the understanding of company financial statements and ratios; (3) Application: the use of WRDS data to calculate financial ratios; (4) Synthesis: the use of liquidity, profitability, debt, inventory, and operating ratios to write a financial analysis report; and (5) Evaluation: the determination of the company's performance over three years and whether its performance is on par, better than, or worse than the industry average.

Hallows and White (2016) stated that finance projects promote higher levels of student learning and help students increase their critical thinking skills. In similar fashion, the WRDS project was designed to advance the PVECOB's learning objectives, specifically the objectives of *Communication, Critical Thinking*, and *Business Knowledge*. The PVECOB defines these learning objectives as demonstrating effective writing for business (Communication), analyzing complex problems and recommending feasible solutions (Critical Thinking), and demonstrating a basic understanding of the common body of knowledge in various business disciplines (Business Knowledge), which for this project included finance as well as accounting. The project was graded based on clarity of writing, analysis of financial ratios, and calculation of common financial ratios.

Purpose and Research Questions

The purpose of this study is twofold. First, we provided one of the first exploratory surveys related to students' perceptions of a WRDS project in an online introductory finance course. Second, given that most business schools require introductory finance courses and follow an unwritten rule to maintain academic rigor in these courses, we examined students' perceptions of

pedagogical approaches in teaching the WRDS project in an online course (FIN 3320). The findings will help business schools and finance faculty design and implement hands-on learning projects using WRDS in online courses, in addition to helping faculty employ WRDS for academic research. As WRDS is already available at many universities for research purposes, implementing a student learning project using WRDS would not require any additional resources for most faculty. Additionally, if schools are unable to justify the expense of using WRDS purely as a research tool (even though WRDS is heavily used by researchers in disciplines such as finance and accounting), the application of WRDS as both a teaching and research tool may help provide the rationale for gaining access.

This study explores the following research questions:

1. What perceptions do students have of using the WRDS platform and the WRDS research project in introductory business finance courses?
2. How do students in introductory business finance courses describe the benefits and challenges of using the WRDS platform and the WRDS project?
3. Did students believe that the WRDS project covered the *Communication, Critical Thinking,* and *Core Business Knowledge* course objectives?

Description of the Project

We both taught FIN 3320 *Business Finance* during the 2021-2022 academic year, had the same learning objectives, and administered a similar hands-on project using the WRDS platform. Among the three online sections, four to five companies were selected from various industries for each online course. We assigned one publicly traded company to students; therefore, about eight to ten students received the same company. Students were instructed to individually write a three-to-five-page report, which included the company background, calculations, analysis of ten common financial ratios for the company, and an industry comparison. The WRDS project was worth 15% of the overall course grade. See Appendix A for an example of the project used.

First, students set up a WRDS account, and once the PVECOB designated administrator approves the account, students have access to WRDS. Second, students have access to and are encouraged to review WRDS tutorial videos that we provided. The videos were available on the Blackboard course site for each online course and instructed students on navigating the WRDS websites, extracting financial statements from the database, and accessing industry-average information. Third, students used the company's ticker symbol to access the financial statements and download the data using Microsoft Excel. The online courses' financial statements and analysis sections focused on several liquidity, leverage, efficiency, profitability, and market value ratios. We selected ten financial ratios for students to calculate. Students were asked to calculate the ten financial ratios based on financial information found in WRDS.

Lastly, along with the calculations, students analyzed their company's performance by comparing the calculated financial ratios over the past three years and with the industry average, also found through the WRDS platform. Students had between four to seven weeks to complete the project. We graded their respective online courses based on calculations, financial analysis, grammar, and formatting and provided grades and thorough feedback to students within two weeks.

Methods

The *Student Perceptions of WRDS* online survey served two purposes. First, we wanted to examine the perceptions of online students using WRDS. Second, we wanted to explore the benefits and challenges of the hands-on project using WRDS. WRDS is a common research platform, so if we could also apply it for instructional content, it would be a more efficient use of the department's limited resources. One of us taught one section with 39 students, and the other taught two sections with 41 students and 42 students enrolled in FIN 3320 *Business Finance*. Appendix B provides the survey questions given to the students.

WRDS Perception Survey

We developed a 24-question online survey which included ten demographic and academic information questions, eleven experiences with WRDS questions, and three open-ended questions about WRDS. The questions were derived from the identified themes and categories from our experience teaching online finance courses and the previous implementation of a WRDS project in the fall 2021 semester. In compliance with the research protocol established by the university and to protect the rights of the participants in this study, Institutional Review Board (IRB) approval was secured prior to the implementation of the *Student Perception of WRDS* online survey.

During the spring 2022 semester, 122 undergraduate students 18 years of age or older enrolled in introductory finance courses at the regional university. Participants in the study were recruited from three introductory finance courses, totaling a final sample of 119 students. In April 2022, we sent an e-mail to students enrolled in the courses requesting that they complete the online survey hosted by Qualtrics. The e-mail informed students of the purpose of the study, consent information, and a link to the survey. The online survey was open for three weeks.

Both qualitative and quantitative data were collected by means of an online survey created and administered by the researchers. Because of the descriptive nature of this research, the qualitative data will provide the thick, rich description essential to understanding the student's perception of WRDS and the project. The quantitative data will provide matters of measurement and effectiveness of the WRDS research project and the utilization of students' skills.

Seventy-four undergraduate students anonymously participated in this survey, yielding a response rate of 62%. Of the 74 participants, a total of one questionnaire was removed due to incompletion, thus, making the final sample for this study 73 participants (61.34%). This response rate is consistent with Sheehan and McMillan's (1999) report that online survey response rates range from 6 to 75%.

To ensure reliability, a Cronbach Alpha coefficient analysis was performed for the study's WRDS scale. For a sufficient reliability level, alpha values should be above 0.70 (Kline, 2011; Lavrakas, 2008) and should not exceed 0.94 (Taber, 2018). The scale reliability coefficient was 0.9246. When grouping questions into the two main categories, it was 0.9390 for the questions related to learning objectives of the WRDS project (See Table 4) and 0.8262 for questions related to the WRDS experience (See Table 2).

Results and Discussion of Findings

The *Student Perceptions of WRDS* online survey captured the following demographic information: age, gender, majors, GPA, and classification at the beginning of the spring 2022 semester. Seventy-three undergraduate students completed the online survey.

We purposefully crafted the *Student Perceptions of WRDS* online survey to target student perceptions and feedback concerning WRDS platform and the WRDS project in online introductory finance courses. The findings are presented accordingly.

Twenty-four (32.9%) of the 73 participants were male and 49 (65.8%) were female, and one participant did not disclose gender, which was consistent with the enrollment in the online classes. Only eight participants were finance majors and 65 (89.0%) were non-finance majors. The undergraduate students were mostly juniors (46.6%) followed by seniors (39.7%) and the smallest group were sophomores (12.3%). One participant declined to report class standing.

Table 1
Demographics of Students (N=73)

	Number	Percentage of Total
Male	24	32.9%
Female	49	65.8%
Finance Majors	8	11%
AEF Majors	17	23.3%
Non-AEF Majors	56	76.7%
High GPA	46	63.0%
Low GPA	26	37.0%
Age 22 or below	34	46.6%
Age 23 or above	39	53.4%

WRDS Platform and Research Project

In Table 2, the mean for each WRDS project statement was computed for the participants' responses. The scale ranged from strongly disagree (1) to strongly agree (5). A brief report of the mean and standard deviation of each learning assessment is shown in Table 2. The highest mean for the WRDS project statement was, "I was challenged by the WRDS assignment" (4.3288). The lowest mean students reported was, "I learned more from completing the WRDS assignment than just listening to a traditional lecture" (4.1644). Overall, the respondents were positive on WRDS project.

Table 2
Students' Report on WRDS Assignment (N=73)

WRDS Assignment	1	2	3	4	5	Mean	SD
I was challenged by the WRDS assignment	1	2	9	21	40	4.3288	0.8984
I prefer courses that utilize hands-on learning activities such as the WRDS assignment	0	1	16	26	30	4.1644	0.8167
I learned more from completing the WRDS assignment than just listening to a traditional lecture	1	6	15	23	28	3.9726	1.0270
I gained experience in collecting, analyzing, and interpreting numerical information through the WRDS assignment	1	0	12	28	32	4.2329	0.8253
The WRDS assignment should be used in future FIN 3320 classes	2	0	14	25	32	4.1644	0.9282

Note: The scale ranged from Strongly Disagree (1) to Strongly Agree (5)

As this is an introductory class with both finance and non-finance majors, any differences in perception between majors are noteworthy. Table 3 presents the T-stat comparisons of the answers of Finance majors compared to all other majors and Accounting, Economics, and Finance (AEF) majors compared to all other majors. T-stats appear in parentheses below the coefficient estimates. *, **, and *** indicate significance at the 10%, 5%, and 1% respectively.

The only statistically significant difference between majors was the agreement in preferring hands-on projects like the WRDS project, preferring hands-on projects to a traditional lecture, and that the project helped the student gain experience working with data. Finance majors were only slightly in agreement with the three statements while the non-finance majors had stronger agreement. A similar, albeit weaker, pattern was observed comparing AEF and non-AEF majors. A non-introductory course or one with a heavy weight towards Accounting, Economics, and especially Finance majors may benefit from a more in-depth project.

Table 3
WRDS Assignment Perception Separated by Majors (N=73)

	Finance	Non-Finance	Difference	AEF	Non-AEF	Difference
Preferred Hands-on Assignment	3.5	4.246	.746 (0.013)**	3.706	4.304	.598 (.007)***
Preferred to Traditional Lecture	3	4.092	1.092 (0.003)***	3.647	4.071	.424 (.137)
Gained Experience with Data	3.5	4.323	.823 (0.007)***	3.882	4.339	.457 (.045)**

Note: The scale ranged from Strongly Disagree (1) to Strongly Agree (5). *, **, and *** indicate significance at the 10%, 5%, and 1% respectively. All other questions showed no statistical difference between majors.

Perceptions of the WRDS Research Project

We coded the responses to open-ended questions from the Student Perceptions of WRDS online survey. Responses were independently coded by both of us through several iterations of reading the responses and maintaining a list of unique words or phrases students used to describe their experiences in each of these three introductory finance courses.

The first open-ended question was, "What challenges did you face using WRDS?" The responses relating to the identified theme of the challenges were captured in the following categories: (1) navigating throughout the WRDS website; (2) calculation, interpretation, and analysis of ratios; (3) signing up for WRDS; and (4) understanding finance terminology.

Over 30% of respondents stated that the WRDS did not pose a challenge for them, although many of these respondents mentioned that it would have been a challenge if not for the materials provided by their professor to assist in navigating the website. This is a testament to the organization of the courses, instructional videos, and Zoom Live sessions we provided. Despite the effort put into providing resources for the students, 38% of the respondents reported that navigating the website was the biggest challenge. Implementing a project using WRDS, especially in an online course with limited capacity for guidance in the classroom, will likely require the instructor to provide significant resources to the students. Selected examples of students' comments supporting this theme are as follows:

<u>Challenges</u>

- The biggest challenge I faced was the new experience. I had never used a platform like this or done an assignment this deeply thought out in a "math" course.
- I have never used or analyzed any of this type of business financial information before. Distinguishing the different terms that, to me, sounded very much alike was challenging.
- The site can be confusing if you have never used it before but my professor did a good job at explaining everything.
- Navigating the system.
- I had challenges interpreting the data and being able to understand the differences in the ratios.
- The WRDS assignment was challenging, but it forced me to use my critical thinking skills, trying to compare increases and decreases in a company to solve each equation.
- WRDS was tedious, but it does teach you how to solve problems manually and helps you take every step you need to solve or explain equations.

The second open-ended question was "What benefits do you feel you received from using WRDS?" The responses relating to the identified theme of the challenges were captured in the following categories: (1) better understanding of ratios and finance; (2) gained hands-on experience with real-world data; (3) improved communication, research, and analytical skills; and (4) experience using a financial database. Selected examples of student's comments supporting this theme are as follows:

Benefits

- I gained a more thorough understanding of the financial ratios. It was also helpful to see the ratios in comparison to the industry as a whole and different businesses within that industry. The visualization that vastly different ratios in one section means very little by itself, however, when put together with other financial ratios, it gave a clearer picture of the financial soundness of a company solidified the material taught in the class in my mind.
- WRDS helped me visualize the financial ratios and how they impact the financial reports.
- WRDS was a neat way to learn and walk through the assignment. From an online student perspective, it was better than listening to a video or reading a textbook.
- I feel using WRDS was very beneficial. It helped take book knowledge and turn it into real world knowledge. I feel that using something like WRDS helps you understand the material on a deeper level.
- I learned how to analyze financial ratio better than I would have by just a lecture.
- The benefits I received from using WRDS would be critical thinking which has helped me understand the importance of researching yearly analytics to have a better understanding of a company's progress.
- I learned so many things about finance while using WRDS. I learned multiple new formulas and skills, while also learning how to apply these things in the real world.

Meeting Course Objectives

Finally, we asked a series of questions to determine what learning objectives the WRDS project covered. For the present study, the mean for each learning assessment was computed for the participants' responses. The scale ranged from strongly disagree (1) to strongly agree (5). A brief report of the mean, standard deviation, and median of each learning assessment is shown in Table 4. Students reported the two highest means for the WRDS project utilized the analytical skills (4.3562) and critical thinking skills (4.3151). The lowest mean students reported was communication skills (3.9452).

With an in-depth project and non-AEF majors who tend to be weaker at quantitative skills, any differences in learning outcomes between strong and weak students is also important to consider.

Table 4
Students' Report on Learning Assessment (N=73)

Learning Assessment	1	2	3	4	5	Mean	SD
I gained a more in-depth understanding of financial ratio analysis through the WRDS assignment	1	1	7	31	33	4.2877	0.8076
The WRDS assignment utilized my problem-solving skills	1	0	8	34	30	4.2603	0.7643
The WRDS assignment utilized my analytical skills	1	0	6	31	35	4.3562	0.7522
The WRDS assignment utilized my communication skills	1	2	20	27	23	3.9452	0.9112
The WRDS assignment utilized my critical thinking skills	1	0	7	32	33	4.3151	0.7615
The WRDS assignment helped me develop specific skills, competencies, and points of view needed by professionals in business	1	1	11	31	29	4.1781	0.8390

Note: The scale ranged from Strongly Disagree (1) to Strongly Agree (5)

Table 5 presents the T-stat comparisons of the answers of high GPA students compared to all other students. Students that reported a GPA of 3.0 or better were considered to have a high GPA. T-stats appear in parentheses below the coefficient estimates. *, **, and *** indicate significance at the 10%, 5%, and 1% respectively. Strong students were more positive about WRDS than weak students, particularly in questions about improving their business and analytical skills.

Table 5
WRDS Assignment Perception Separated by GPA (N=73)

	High GPA	Low GPA	Difference
Understood Financial Ratios	4.435	4.037	.398 (0.041)**
Tested Analytical Skills	4.478	4.148	.330 (0.070)*
Developed Business Skills	4.326	3.926	.400 (0.048)**
Gained Experience with Data	4.391	3.963	.428 (0.031)**
Should Use Assignment	4.323	3.889	.437 (0.051)*

Note: The scale ranged from Strongly Disagree (1) to Strongly Agree (5). *, **, and *** indicate significance at the 10%, 5%, and 1% respectively. All other questions showed no statistical difference between students.

Overall, students reported that they enjoyed having a hands-on research project and preferred it to a traditional lecture. Students found the WRDS research project challenging and said it helped improve their problem-solving and analytical skills. They also reported that the project helped them with *Communication, Critical Thinking,* and *Core Business Knowledge*: the three learning objectives of the project.

Discussion

We identified four benefits of this scholarship. First, an examination of the effectiveness of the WRDS platform and hands-on projects for introductory finance classes is limited. Therefore, the study contributes to the gap in the finance literature. Second, the information in this study will assist finance faculty in developing and strengthening current and future WRDS research activities and other hands-on projects. Third, we evaluate the effectiveness of the WRDS research project and improve any shortcomings identified by the study. Lastly, the survey results help finance instructors and researchers recognize the benefits and challenges of using WRDS in introductory finance courses.

The findings from this study highlight the benefits and challenges of incorporating WRDS platform and the WRDS project in introductory finance courses. The course and instructional videos included as part of the WRDS project play an important part in improving student performance, as illustrated by the findings of Sebastianelli and Tamimi (2011). For instructors, the effectiveness of the hands-on project using the WRDS platform will depend upon course organization, instructional videos, and availability via office hours, phone, or Zoom live sessions (come-and-go sessions for students on designated times and days) to address complex finance terminology and WRDS project's challenges.

The benefits for instructors include: (a) providing students with a hands-on experience; (b) utilization of a financial database; (c) providing a more interactive and engaging online course with real-world application; and (d) strengthening students' transferable skills, (e.g., analytical, communication, critical thinking skills, etc.) for upper-level courses and future careers. From experience and the research evidence, the benefits of incorporating the project using the WRDS platform outweigh the challenges faced by both students and instructors.

Implementation and Modifications

Instructors have numerous ways to modify the project to better engage students or address other concerns. While we chose ten common financial ratios for analysis, other financial ratios can be added or removed to best fit what the instructor wishes to emphasis. Likewise, ratios involving societal impact like the gender or diversity of the board of directors can also be added. Students were assigned to one of four companies, but students can be allowed to pick their own companies, or the number of companies chosen could be increased or decreased. This is at the instructors' discretion. It should also be noted that before instructors allow students to select their own company, instructors should both consider the workload of downloading financial statements from WRDS and administering calculations and analysis for numerous companies as well as if the companies chosen provide the necessary data to calculate all the assigned ratios.

Navigation of the WRDS website to access the financial data required guidance from instructors. Providing video tutorials and Zoom sessions to assist students appeared to aid in website navigation. Instructors can view how often students watch or access that information. Unfortunately, the videos that we used were shared across the department and, therefore, included data from students enrolled in different classes. If the project is given to an on-campus class, some classroom time to demonstrate using the website would be needed to provide a similar level of assistance.

Cheating is an ever-present danger no matter the class format. Ensuring that students are assigned to unique companies would make sharing work more difficult; however, the burden on

the instructor also increases. As this is partially a writing assignment, plagiarism and AI generation is also a potential cheating problem. Besides assigning unique companies, instructors may have access to plagiarism detection software. Turnitin, one such software available to the authors, also offers an AI detection option.

Limitations

There are limitations of this study that should be acknowledged. First, the student participants came from one public university. It is unknown how these perceptions might compare to students at other institutions. Second, the information was self-reported and focused on three online finance courses. Self-reported data may not accurately reflect if the project tested the course objectives. Despite the limitations, the findings reported here corroborate what we believe to be the implementation of projects using the WRDS database, which is a beneficial addition to introductory online financial economics courses.

Conclusions and Future Research

The purpose of this paper is to explore student perceptions and illustrate an instructional tool using WRDS database in introductory online finances courses. Students reported that they were generally positive towards WRDS and the hands-on project, and the WRDS project tested them on learning outcomes. We highlighted instructional strategies and provided evidence supporting the incorporation of the WRDS platform in introductory finance classes with many non-majors. The WRDS project described in this paper supports the development and strengthening of students' skillsets for upper-level courses and career development.

The implications for future research are numerous, and the authors are currently in the process of conducting a second research study. We are interested in expanding the study to compare F2F and online finance courses and providing pretest and posttest assessments. Additionally, future opportunities entail researching different disciplines and graduate students and exploring gender differences and other demographics.

Ethical Approval

IRB approval has been obtained from the authors' institution, confirming ethical standards.

References

AACSB Business Standards. (2020). Retrieved from https://www.aacsb.edu/accreditation/standards/business.

Alshurafat, H., Alaqrabawi, M., & Al Shbail, M. O. (2023). Developing learning objectives for forensic accounting using bloom's taxonomy. *Accounting Education*, 1-17.

Batu, M., Bower, N., Lun, E., & Sadanand, A. (2018). Testing the effectiveness of online assignments in theory of finance. *Journal of Education for Business, 93*(3), 119–127. doi:10.1080/08832323.2018.1425660

Betts, S. C. (2008). Teaching and assessing basic concepts to advanced applications: Using Bloom's taxonomy to inform graduate course design. *Academy of Educational Leadership Journal*, *12*(3), 99.

Biktimirov, E.N., & Klassen, K.J. (2008). Relationship between use of online support materials and student performance in an introductory finance course. *Journal of Education for Business*, *83*(3), 153–158. doi:10.3200/JOEB.83.3.153-158

Campbell, N. D., Heriot, K.C., & Finney, R. Z. (2006). In defense of silos: An argument against the integrative undergraduate business curriculum. *Journal of Management Education*, *30*(2), 316–332. doi:10.1177/1052562905277231

Hadsell, L. (2009). The effect of quiz timing on exam performance. *Journal of Education for Business*, *84*(3), 135–140. doi:10.3200/JOEB.84.3.135-141

Hallows, K., & White, S. (2016). The introductory financial management course—do students develop critical thinking skills?. *Journal of Financial Education*, *42*(1-2), 1-34.

Handy, S. A., & Basile, A. (2005). Improving accounting education using Bloom's taxonomy of educational objectives. *Journal of Applied Research for Business Instruction*, *3*(3), 1-6.

Jones, J. S., Ragan, K. P, Witte, H. D., & Zhang, Y. J. (2021). The impact of online learning tools on student success in college of business core finance classes. *The Journal of Education for Business*, *97(*4), 277–284.

Kline, R. B. (2011). *Methodology in the social sciences. Principles and practice of structural equation modelling* (3rd ed.). Guilford Press.

Krishnan, V. S., Bathala, C. T., Bhattacharya, T. K., & Ritchey, R. (1999). Teaching the introductory finance course: What can we learn from student perceptions and expectations? *Financial Practice and Education*, *9*, 70–82.

Lavrakas, P. J. (2008). *Encyclopedia of survey research methods.* Sage Publications, Inc., Thousand Oaks.

Moreale, J., & Zaynutdinova, G. R. (2018). A Bloomberg terminal application in an intermediate finance course. *Journal of Financial Education*, *44*(2), 262-283.

Payne, J. D., & Tanner, G. (2011). Experiential learning and finance: A hands-on-approach to financial modeling. *Journal of Financial Education*, *37*(3/4), 82-100.

Reeves, M. F. (1990). An application of Bloom's taxonomy to the teaching of business ethics. *Journal of Business Ethics*, *9*, 609-616.

Root, T., Rozycki, J., Senteza, J., & Suh, I. (2007). The undergraduate finance curriculum in the new millennium: A comprehensive survey. *Journal of Financial Education*, *33*, 1–27.

Sebastianelli, R., & Tamimi, N. (2011). Business statistics and management science online: Teaching strategies and assessment of student learning. *Journal of Education for Business*, *86*(6), 317–325. doi:10.1080/08832323.2010.525545

Sheehan, K. B., & McMillan, S. J. (1999). Response variation in e-mail surveys: An exploration. *Journal of Advertising Research, 39*(4), 45-54.
Shulman, L. S. (2002). Making differences: A table of learning. *Change: The Magazine of Higher Learning*, *34*(6), 36-44.
Soffe, S. M., Marquardt, M. J., & Hale, E. (2011). Action learning and critical thinking: A synthesis of two models. *Action Learning: Research and Practice*, *8*(3), 211-230.
Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*, 1273-1296.
Yu, J., Li, H., & Haley, J. D. (2021). The development and implementation of asynchronous online finance courses. *Journal of Financial Education*, *47*(2), 1–18. https://www.jstor.org/stable/48738990

Appendix A
WRDS Financial Ratio Assignment

This assignment is designed to utilize your financial analysis and critical thinking skills. For this assignment, you will calculate financial ratios for the three years 2020, 2021, and 2022 (inclusive) for a designated company and compare the results with industry averages found in *WRDS Financial Ratios Visualization*. Furthermore, you will provide a brief background of the company and address how the company is being managed based on the financial ratios.

Step 1: Your last name will determine which company you will analyze. Locate your designated company below in Table A1.

Table A1
Company Assignments

Assigned (by Last Name)	Company	Symbol (Ticker)	GICS Sector (Industry)	Headquarters Location
A-Ca	T-Mobile	TMUS	Telecommunications	Bellevue, WA
Ch-D	Costco	COST	Consumer Staple	Issaquah, WA
G-K	ConocoPhillips	COP	Energy	Houston, TX
L-Q	Auto Zone	AZO	Consumer Discretionary	Memphis, TN
R-Z	Tesla	TSLA	Consumer Discretionary	Austin, TX

Step 2. Log into WRDS and download balance sheets and income statements for the above three years. Compute the financial ratios for your company. For this assignment, use the designated formulas listed in Table A3. You are welcome to use EXCEL or a calculator.

For additional help finding the financial statements, see the instructions below and review the step-by-step video provided in WTClass.

Step 3. Go to *WRDS Financial Ratios Visualization* for (GICS Sector) industry information and compare your results with the industry averages you find. For this assignment, please use the industry averages located in WRDS. For additional help finding the Industry averages, see the instructions below and review the step-by-step video provided in WTClass.

Step 4. Put your analysis in the format below or create your own so long as it is easy to read in Table A2.

Table A2
Example Table for the Assignment

Financial Ratios and Industry Comparative for Solee Energy						
Ratios	**2020**	**2021**	**2022**	**2022 Numerator**	**2022 Denominator**	**Industry**
Current Ratio	1.27	1.02	1.24	$30,266,000	$23,872,000	1.32
Quick Ratio	…	…	…	…	…	…
O. Profit Margin	…	…	…	…	…	…
…	…	…	…	…	…	…

As you know, we have covered several financial ratios this semester. For this assignment, you will only need to complete a table with the following **ten** ratios **(use the formulas listed in Table A3)** and Industry averages:

Table A3
Ten Ratios and Formulas Used for WRDS Assignment

Ratio	Formula
Current Ratio	$\frac{\text{Current Assets}}{\text{Current Liabilities}}$
Quick Ratio	$\frac{(\text{Current Assets} - \text{Inventory})}{\text{Current Liabilities}}$
Operating Profit Margin	$\frac{\text{EBIT}}{\text{Sales}}$
Net Profit Margin	$\frac{\text{Net Income}}{\text{Sales}}$
(Operating) Return on Assets	$\frac{\text{EBIT}}{\text{Total Assets}}$
Return on Equity	$\frac{\text{Net Income}}{\text{Total Equity}}$
Inventory Turnover	$\frac{\text{Cost of Goods Sold}}{\text{Inventory}}$
Total Asset Turnover	$\frac{\text{Sales}}{\text{Total Assets}}$
Debt/Equity Ratio	$\frac{\text{Total Liabilities}}{\text{Total Equity}}$
Debt Ratio	$\frac{\text{Total Liabilities}}{\text{Total Assets}}$

Step 5. Go to *50/50 Women on Boards (*https://5050wob.com/). Scroll down to *Research* and then *Directory*. Search for your company name and complete the chart. If your company is not listed on the website, please go to your designated company's website and search for the board of directors. You will then find the number of women on the board and divide the number by the total number of board members to calculate the percentage of women on the board.

Table A4
Example Table for the Assignment

	Total number of Board members	Number of Women on Board	% of Women on Board	Does the Board receive a rating of Gender-Balanced (GB)?	U.S. Average
Solee Energy	13	6	46%	Yes	27.3%

*27.3% is the number everyone will use for comparison because it is the number for the Russell 3000 index as of the start of this semester

Step 6. After you have computed the ratios for the years 2020, 2021, and 2022, then answer the following questions:

Question 1. Provide a brief background of the company.
Question 2. What do the ratios tell you about the company and how it is being managed from year to year? For this question, have one paragraph for each ratio. Be sure to address 50/50 Women on Boards information.
Question 3. Compare the company ratios to the Industry. How is the company being managed relative to the average firm? You want to address whether they are doing better or worse relative to the Industry and **why**.

To access a balance sheet and the income statement for a company using WRDS, please follow the ensuing steps.

1. Go to **WRDS**.
2. Login to your **WRDS** account. ...
3. Select "Classroom" tab.
4. Scroll down to the "Balance Sheet" or "Income Statement" and select one or the other.
5. Select the "Link to Platform" tab.
6. Select the date range.
7. Enter the Ticker Symbol for the company.
8. Submit Query.
9. After submitting the query, a new screen will appear with the requested information.

Please follow the ensuing steps to access **WRDS** Financial Ratios Visualization (for Industry).

1. Go to **WRDS**.
2. Login to your **WRDS** account. ...
3. Select "Classroom" tap.
4. Select "**ACCOUNTING**"
5. Scroll down to the "**Financial Ratios Visualization**" and select.
6. Then select the "Link to Platform" tab.
7. After you select, a new screen will appear with a table of the Industry Financial Ratios.

Format for Assignment:

- The report should be written in your own words. (No AI for this assignment.)
- Double-spaced with 1-inch margins on all sides
- Times New Roman, Cambria, or Calibri 12-point font
- No more than 4 pages written (not including exhibits)
- Be sure to proofread and utilize the Writing Center. Use Grammarly.
- Punctuation, sentence form, spelling, grammar, neatness, and organization will be considered in the grading process

Grading Sections:

- Calculations: 70 points
- Content/Questions: 60 points
- Language/Formatting 20 points
- **Total Points** **150 points**

Appendix B
Student Perception of WRDS Research in Business Finance Courses

1. **Age**:
2. **Gender**: (Male; Female; Other)
3. **Ethnicity** (African American/Black; Asian; Caucasian; Hispanic/Latino; Other)
4. **Classification**: (Freshman; Sophomore; Junior; Senior)
5. **College of Business** (Major)
6. **College of Business** (Minor)
7. **Overall GPA**: (3.5 – 4.0; 3.0 > 3.5; 2.5 > 3.0; 2.0 > 2.5)
8. Which semester were you enrolled in **FIN 3320** Business Finance?
9. How did you take **FIN 3320** Business Finance? (Online; Face-to-Face)
10. Prior to enrolling into **FIN 3320**, have you ever used **WRDS** before?: (Yes; No)

Perception of WRDS I

Based on your experience with the WRDS assignment, please indicate to what extent you agree or disagree with each statement.

Responses (Questions 11-17): (Strongly agree; Agree; Neutral; Disagree; Strongly disagree)

11. I gained a more in depth understanding of financial ratio analysis through the WRDS assignment
12. The WRDS assignment utilized my problem solving skills
13. The WRDS assignment utilized my analytical skills
14. The WRDs assignment utilized my communication skills
15. The WRDS assignment utilized my critical thinking skills
16. The WRDS assignment helped me develop specific skills, competencies, and points of view needed by professionals in business
17. I was challenged **by the WRDS assignment**

Perception of WRDS II

Based on your experience with WRDS, please indicate to what extent you agree or disagree with each statement.

Responses (Questions 18-21): (Strongly agree; Agree; Neutral; Disagree; Strongly disagree)

18. I prefer courses that utilize active learning activities such as the WRDS assignment
19. I learned more from completing the WRDS assignment than just listening to a traditional lecture
20. I gained experience in collecting, analyzing, and interpreting numerical information through the WRDS assignment
21. The WRDS assignment should be used again in the course in the future FIN 3320 classes

Please describe what you have learned using the WRDS platform. Feel free to provide examples.

22. What challenges did you face using WRDS? (Open-ended Question)
23. What benefits do you feel you received from using WRDS? (Open-ended Question)
24. What additional comments do you have pertaining to the WRDS assignment? (Open-ended Question)

The art of the deal[2]: Navigating health insurance through double oral auctions, the iterated prisoner's dilemma, and risk management[3]

Roger Lee Mendoza
California State University-Los Angeles

Insurance is among the major transactions covered by the financial sector. Whether privately purchased or publicly sourced, insurance is the largest healthcare payer in any industrialized country. In the United States, over 82 percent of healthcare expenditures, including drug costs, is financed by insurance. We devised a simulation game — partly adapted from the double oral auction and iterated prisoner's dilemma games — as an instructional tool for healthcare finance and insurance courses. It teaches and engages students about the special characteristics of financial risk in healthcare, how insurance mitigates such risk, and why insurance markets work. We find its game-theoretic elements (players, rules, actions, strategies, outcomes, and payoffs) helpful in bridging hands-on learning and real-life experiences in healthcare coverage and encouraging wider student engagement in otherwise highly technical courses like insurance. Individually and collectively, undergraduate and graduate students can acquire thinking and practical skills in risk and insurance management, including knowledge-in-action, processes, problem-solving, decision-making, and communication. Immersive experiences and gameplay mechanics further provide student players in this three-part game with a unique playground for customization and creation. Finally, we share debriefing and feedback, and suggest that insurance courses and segments consider incorporating game-theoretic simulation as an instructional modality.

Keywords*: Uncertainty; health insurance; insurance markets; healthcare finance; game-theoretic simulation; insurance education.*

Introduction

Pioneered by mathematicians Emile Borel and John Von Neumann almost a century ago, game theory evolved to discover whether an optimal or "best" strategy for any game exists. Any game-theoretic model contains six basic elements: 1) competitive players (decision-makers); 2) preset rules governing players' behavior; 3) actions that players take within a given set of rules; 4) strategies which represent the complete plan of actions that players take in all possible scenarios; 5) game outcomes (each of which results from choices made by players in consideration of others

[2]With apologies to Trump DJ with Schwartz T (1987). *Trump: The art of the deal*. New York, NY: Random House.

[3]Revised version of paper presented at the 25th annual conference of the Academy of Business Education (ABE), Panel session on teaching innovations, San Antonio, Texas, September 19-21, 2024. The author acknowledges with thanks the helpful comments and suggestions of the four anonymous peer reviewers and the journal editor, Dr. Richard Fendler, and the invaluable research and editorial assistance of Lucy DuMez. As with any work of this nature, the usual caveat applies.

at any given point in the game); and 6) payoffs accrued by players as a result of each possible outcome. Along with a solution concept, these basic elements help game theorists deduce a set of equilibrium strategies for each player. That is, any game with a finite number of players, each of whom has an equally finite number of strategic options to choose from, would have at least one Nash equilibrium, where no player may profit by unilaterally deviating from a chosen strategy after considering the other player's decision (Smith, 2003). Various games and strategy matrices, based on player cooperation with one another, or defection (i.e., competition and conflict), have been developed over the years. Game theory has found enormous applications in many fields of study, such as economics, politics, logic, biology, marketing, military strategy, and computer science, and in many sectors, including healthcare (Gokhale & Traulsen, 2010).

Based on two game-theoretic models, namely, the double oral auction (DOA) and (iterated or repeated) prisoner's dilemma, we devised and piloted a three-part simulation game as a teaching tool in an introductory health insurance course and the health insurance segment of a healthcare finance course. Insurance is a core or foundational course in healthcare management programs at the undergraduate and graduate levels. It is offered either as a stand-alone course or a segment of a healthcare finance course (Mendoza, 2023). Whether privately purchased or publicly sourced, insurance — one of the major transactions covered by the financial sector — is the largest healthcare payer in every industrialized country in the world. In the United States, over 82 percent of all healthcare expenditures, including drug costs, is financed by insurance (Mendoza, 2022).

Through our game-theoretic simulation undergraduate and graduate students learn about three important aspects of insurance, in general, and health insurance, in particular: 1) insurance markets (non-group versus group); 2) insurance pricing (including risk exposure and the relationship between healthcare spending and out-of-pocket price); and 3) the central problems of insurance, namely, adverse selection and moral hazard. How and when student players manage to "close" or "hammer" a deal (reflective of the title of this paper) is of paramount interest to us in teaching and learning insurance. In choosing gaming pedagogy, the insurance lesson seeks not just to teach, but actively engage, students about the special, but often misunderstood, characteristics of financial risk in healthcare, and how insurance mitigates such risks at a given cost. In this context game-theoretic simulation allows us to assess students acquisition of hands-on and thinking skills, including knowledge-in-action, processes, problem-solving, decision-making, and communication. Immersive experiences and gameplay mechanics further provide student players with a unique playground for customization and creation.

Related Literature

Simulation games have been used in healthcare management courses for sometime now (Asrafian, Darzi & Athanasiou, 2011; Shurkin, 2013). Typically adapted for classroom use, these games aim to foster competitive problem-solving and decision-making, and conclude with a debriefing that covers healthcare management concepts, principles, methods, and problems, and their practical implications. The advantages of gaming in healthcare finance and insurance include: 1) heightened student involvement; 2) opportunities for real-life problem-solving, including in unusual situations; 3) memory recall; 4) process illustration; and 5) recognition and resolution of decisional constraints (Zelmer & Zelmer, 1980; Macleod & Smith, 1984; Celly, 2007; Asrafian, Darzi & Athanasiou, 2011).

Among the earliest of health insurance simulation games was a board game devised by Macleod and Smith (1984) to simulate real-life experiences of paying healthcare-related bills given

many other financial considerations. Student players — each of whom is assigned a type of health insurance policy — roll dice and move their pieces around the periphery of the board, landing on spaces that may be purchased or on spaces that require the player to draw a card from a pile in the center of the board. Besides improving final examination scores, the authors observed that the game afforded greater sensitivity about the advantages of healthcare coverage as well as the financial burden it imposes or imposes unequally (e.g., for the elderly). Weinberg et al. (2013) appear to extend these objectives in devising another game consisting of four rounds. It focuses on risk pooling to explore students' "understanding of risk and their own level of risk aversion, and apply the learned concepts to the policy debates over healthcare insurance coverage" (p. 128). Round One introduces the health risks in medical care and hospitalization, leading most participants to join risk pools of health insurance organizations in Round Two to limit potential liability. In Round Three, insurance pools attempt to expel high-risk individuals, including those with pre-existing conditions (now prohibited under the Affordable Care Act of 2010), while generally charging catastrophic coverage higher membership premiums in Round Four. Mellor's two-part experimental game (2005), on the other hand, deals with the central problems of asymmetric information. Part 1 simulates a market in which the low-risk and high-risk can purchase (the same) healthcare coverage, with instances of government regulation that constrains sellers from using buyer type to underwrite premiums. The outcomes suggest why asymmetric information produces adverse selection. In Part 2, insurers offer moderate and generous coverage resulting in adverse selection which leads to inefficient buyer sorting across plans under government-mandated community rating, limits on premium increases, and reduced purchases of the more generous plan/s. Debriefing found this game useful in illuminating the causes and consequences of adverse selection for consumers and insurers, and potential solutions to employer and government-sponsored coverage.

The foregoing simulations offer creative ways of teaching and learning basic concepts of healthcare risk, risk pooling, and insurance trade-offs and helping build student competencies for analyzing continuing policy debates over healthcare coverage. These games were particularly insightful to us in developing our own experimental game. However, taking a cue from Celly (2007), because understanding risk and the role of insurance in risk management can be difficult to learn and challenging to teach, we seek to fill in the gaps in the previously reviewed literature by introducing students to health insurance markets, and how third-party intervention affects insurance supply and demand, especially considering that risk differentiation is an equally important concept in acquiring critical learning skills (Asrafian, Darzi & Athanasiou, 2011).

For these reasons, in adapting a DOA with multiple buyers and sellers, we added some key elements of the prisoner's dilemma game that turn the various auction rounds into an indefinitely iterated game. Buyers and sellers in a DOA can simultaneously call out offers to buy (bid) or sell (ask), and, subject to their respective pay-off calculations, can mutually agree on an offer at any point in the trading process. Each buyer's reservation value (or maximum allowable bid) and seller's production (or opportunity) cost for the auctioned commodity is their private information, and preferences are quasi-linear (Assadi et al., 2017). An experimental game version was previously created by Gillette (1996) and then used by Gazda et al. (2012) to test for moral hazard and distinguish between insurance purchasing, first without and subsequently with third-party intervention. In Gillette's DOA (1996), buyers are informed after a few free trading rounds that a "benevolent dictator" will pay 80 percent of the negotiated insurance price, allowing students to understand the trade-offs between market efficiency and equity arising from universal healthcare coverage. Gazda et al. (2012) essentially obtained the same outcomes as Gillette, but statistically

tested them based on a null and alternative hypothesis. They conclude that market prices arising from third-party intervention and accelerated purchasing raise average insurance price and reduce market efficiency, consequently producing moral hazard.

Gillette (1996) and Gazda et al. (2012) help us simulate the effects of third-party intervention by government. We expand on their DOA model through a three-phase simulation game that allows for a broader and more in-depth understanding of the potential implications of non-group and group insurance on healthcare accessibility, quality, and cost, with and without third-party intervention by government. In addition, we incorporate an indefinitely repeated game into our DOA that fairly resembles the iterated prisoner's dilemma. In this case, buyers and sellers alike put forward offers without knowing how many auction rounds there will be and how long each one might last, including when the auctioneer (or "pit boss") might discretionarily close an auction round, and whether another round would be called. Recognition (of player propensities, strategies, and risks), player reputation, and reciprocity — the key elements of the iterated prisoner's dilemma game (Axelrod, 1984) — equally underpin our simulation.

In contributing to the literature on experiential learning in health insurance, we situate our game-theoretic simulation within the broader scholarship of teaching in finance education as we analyze game outcomes and by way of conclusion.

Methods

Designed as a three-part experimental game, ours was piloted in a three-hour undergraduate class in healthcare finance (i.e., for its insurance segment) and three-hour graduate/master's-level class in healthcare risk and insurance in the Fall Term of 2023. Similar to the game simulation literature we reviewed, both undergraduate- and graduate-level games were held in-person for ease of instruction, assessment, and feedback. Briefing was conducted in class using slides and emailed written instructions the week before the DOA game was played. Debriefing based on the objectives and outcomes of each of the three game segments took place at the end of the entire game, although we share them in this paper at the end of each of the following three sections.

The three game parts (insurance trading and coverage value, insurance markets, and market intervention) were played over a period of approximately 75 minutes prior to debriefing, with a 10-12 minute break and quick briefings by the instructor in between. The games were closely supervised by the instructor as auctioneer and a student recorder of offers. Each buyer and seller was assigned a reservation value and production cost, respectively, in all three parts for which each participant was allowed to trade until they managed to "seal a deal" for one single unspecified commodity (i.e., health insurance). In Part 1, undergraduate and graduate students are separately and randomly assigned as uninsured buyers and sellers. They trade for the unspecified commodity based on their reservation values and production costs, which help differentiate healthcare financial risks with and without insurance. In Part 2, students are randomly assigned as single buyers or as one group of buyers and single sellers of the unknown commodity. This arrangement is intended to represent non-group and group health plans and simulate their respective markets, bargaining power, and pricing. No distinction is made in terms of market size (i.e., small or large group insurance). In Part 3, students are randomly assigned again as single buyers and sellers of the unnamed (or by this time already known) commodity, but then third-party intervention is introduced and changes the incentives and disincentives behind player offers. This is meant to simulate the role and consequences of government intervention in the insurance marketplace and social insurance for certain (vulnerable) populations, particularly the elderly (Medicare) and the

low-income (Medicaid). Because we set reservation values higher than production costs in all three parts of the game to induce trade, the competitive equilibrium (CE) derives from the number of buyers willing to purchase at a given price and the number of sellers willing to sell at that price. Any price below production cost is not an equilibrium price since it will create excess demand. Conversely, any price above reservation price is not an equilibrium price either since it will create excess supply. In this regard, the supply is based on the number of sellers with a production cost at or below the price and the demand would be the number of buyers with a reservation value greater than or equal to the price.

Expected outcomes from each of the three game parts could resemble a prisoner's dilemma, but in a way that shifts from the classic one-shot to an indefinitely iterated game (Axelrod, 1984). In making an offer, buyers and sellers in our DOA always have an incentive to strategize in a way that creates a less than optimal outcome for the others. In this sense, players receive the greatest payoff if they "betray" the other rather than "cooperate" as they maximize their benefit and minimize their cost. However, because the sheer length and number of DOA rounds in our game are not definitely known to players, the classic one-shot prisoner's dilemma can evolve into an indefinitely iterated game, enabling the same players to learn more about each other over time based on recognition, reputation, and reciprocity. It then becomes likelier for buyers and sellers to formulate game strategies that eventually reward cooperation. Strategic thinking between players at this point need not necessarily lead to suboptimal outcomes (Axelrod, 1984).

Outcomes from the three game parts are summarized in univariate tables for which measures of dispersion were used. They seek to fulfill our three-fold objective outlined in the introductory section of this paper. Because our game is intended for classroom use only to introduce students to the subject of (health) insurance, it is exempt from IRB approval of research projects that undertake "systematic investigations" involving human subjects to contribute to generalizable knowledge and develop or test a theory applicable to a wider population.

Limitations

While we identify clear learning objectives in each of the three parts of our game-simulation, we cannot fully replicate and capture the whole range of incentives, valuation processes, pricing dynamics, and motivations and attitudinal changes at play in health insurance markets and transactions. After all, this is a classroom game we piloted in two classes with student players experientially learning about insurance for the first time, rather than insurers, individual purchasers, firms, and governments whose cost-calculus are far more sophisticated and better informed based on actions and strategies available to them and their expected outcomes. While our simulation should not be confused with an experimental study aiming for generalizable results, the buyer and seller stipulations and subsidies we added to our game help players understand how problem-solving and decision-making in health insurance might evolve and change among players over time.

Insurance Trading and Coverage Value

In Part 1 of our game, students in the undergraduate finance and graduate insurance classes were randomly assigned as buyers or sellers (an uneven number makes no difference). One student in each class was designated as recorder (of offers and winners).

Each buyer was handed an index card, labeled "Buyer," and containing a not-to-be-disclosed reservation value (e.g. $100) to purchase a single but unknown product. That means they could bid for any amount no higher than that value (e.g., $1-100 if reservation value = $100). Every seller also received an index card labeled, "Seller," that contained their confidential production cost (e.g. $70) for the same single but unrevealed product that they will sell in a DOA for any amount no less than that cost (≥ $70 if production cost = $70). Like in any DOA, buyers were expected to bargain for the lowest possible transaction price, while sellers negotiated for the highest possible price, to maximize their respective gains or surpluses based on the principle, "don't pay more than your value, and don't sell for less than your cost." Hence, each player's payoff as indicated in their respective index cards:

Buyer's pay-off = reservation value - transaction price
Seller's pay-off = transaction price - production cost

After reminding players of the rules, and giving them enough time for questions and comments and to strategize, trading in $1 increments began. A buyer could trade by openly stating their bid. A seller could accept, counter-offer (ask), or remain silent, which a buyer could likewise do. Any bid is written on the board and remains there until accepted, canceled, or replaced by a higher bid or lower ask. Buyers and sellers may trade for as many times as they wish until the instructor (as auctioneer) closes the auction round. They drop out of the market if their offer to buy or sell their single product is accepted in which case auctioneer and recorder will check their index cards, record the transaction, and announce the winning offer (after chanting "going, going, gone!"). The auctioneer could also end trading anytime with a "soft close." In that case, buyers and sellers alike are asked if they will accept the last offer on the table. And only the auctioneer can choose whether or not to call for another auction (round) if time permits.

If no offer is accepted after a reasonable wait permitted by the auctioneer, the round is also canceled and a new round may commence at the auctioneer's discretion. However, for this succeeding round, a new bid should be above the highest one from the canceled round (e.g., if no one wanted to sell $80 in the previous round, the new offer must at least be $81). Similarly, any new offer to sell or ask should be lower than the lowest one received in the canceled round (e.g., if no one wanted to pay $90, a new offer to sell must not be higher than $89). Ostensibly, the goal at this point is to keep things moving toward a transaction with the new bids and asks kept in $1 increments.

There is a catch to the failure of any buyer to secure a mutually acceptable trade in Part 1. A buyer stipulation distinguishes our DOA game from others: Buyer's reservation price is automatically reduced by 50 percent at the end of Part I if no seller accepted any of their bids for that single unrevealed product (i.e., health insurance). This is communicated to all students prior to game start when they are briefed about the rules at which time they can also ask questions. The halving represents the subsidy that a buyer shopping for insurance does not receive or gain because they remain uninsured and therefore pay for healthcare at full market price. Halving also changes the buyer's value of the service. Studies show that insurer-negotiated rates can range up to 72 percent lower than a healthcare provider's full price, depending on the service (Amin et al., 2024). Considering that the reservation price of each buyer is unique but not far different from other buyers (≤ 10 percent), any buyer faces a significant loss in failing to secure a deal at the end of Part 1.

Game outcomes are reported in Table 1. Trading in both undergraduate and graduate classes was highly spirited in all rounds. In the undergraduate class, sellers in the first two rounds managed to negotiate relatively higher transaction or trading prices despite resistance exhibited by buyers

through bids and counter-bids. Eventually, bids markedly declined in rounds 3 and 4. Undergraduate sellers seem to realize that buyers were countering sellers' asks that were above $84, gradually resulting in player cooperation and mutually acceptable offers of $81 in round 3 and $83 in the terminal round. In the graduate class, trading prices went down and up slightly in all five rounds. Regardless of the variation, several graduate student buyers likewise appeared to be signaling to sellers but were more cooperative in negotiating prices that were not too far off from each other and ranging from $83 to $86.

Table 1
Uninsured buyer and seller trading

Game components	Undergraduate Class*	Graduate Class*
Rounds	4	5
# of buyers**	9	4
# of sellers	9	4
Buyer's reservation value ($ range)	90-100	90-100
Seller's production cost ($ range)	70-78	70-78
Transaction price ($)		
Round 1	85	84
Round 2	86	83
Round 3	81	86
Round 4	83	84
Round 5	n/a	85
Average (μ)	83.75	84.4
Variance (σ^2)	3.687	1.040
Standard deviation (σ)	1.920	1.019

*One student designated as recorder did not play.
**Buyer's stipulation: reservation value reduced by 50 percent at game end if unable to win bid in any round.

In terms of variability, transaction prices at the undergraduate and graduate DOAs fell below 2 standard deviations, suggesting that offers and counters were close in value to the mean, and hence, homogenous enough, especially in the graduate health insurance class. Mean undergraduate and graduate prices that closed a deal were also very close at $83.75 and $84.40, respectively. A relatively low standard deviation and closing price homogeneity suggest that trading in Part 1 easily converged to the CE solution. All offers were below reservation values and above production costs. This lends support to the nearly universal finding from three decades of experimentation that price and quantity quickly converge to CE values where supply equals demand. As some economists aptly put it: "The rapid flow of information combined with the ability of traders to

instantly undercut an outstanding bid or ask makes the [DOA] perhaps the closest embodiment of the economist's notion of a perfect frictionless market" (Rust, Palmer & Miller, 1989, p. 2).

During the debriefing, held as a structured Q&A at the end of the entire game, the results appearing in Table 1 were shared in class. Students were first asked if they had any idea what they were bidding for. Several students in both classes said they thought they were bidding for some type of health insurance, though others simply said insurance, as our game was played in the undergraduate finance class during the first of three weeks devoted to the subject of insurance and risk management, while the graduate class was specifically on health insurance. Students in both classes were then asked why they thought buyers (but not sellers) would lose a substantial amount (50 percent) of their reservation value if they either chose not to bid or failed to secure a winning bid. After all, more than 65 percent of buyers from both classes received the buyer's reduction. No one was able to supply the correct answer from the undergraduate finance class. One did from the graduate insurance class, but they explained it in less coherent terms. The instructor then proceeded to the discounting effect of health insurance. He did so in terms of the significant healthcare cost-reductions, pointing out that subsidization on average is up to 72 percent of a provider's full price (Amin et al., 2024), and the subsidy increases in more complex health issues and riskier and/or longer medical treatments and services. Adverse selection was then introduced in terms of how health status or conditions and needed treatments can induce higher-risk individuals to buy — or buy more generous — insurance by way of taking advantage of this subsidy.

Thereafter, debriefing served to analyze the four or five trading rounds in Table 1 based on the theory of healthcare risk and risk management through insurance, health plan shopping on the part of the enrollee and marketing by the insurer (underscoring a DOA model's relevance), open enrollment windows, and the (now unenforceable) individual mandate of the Affordable Care Act. The CE relative to insurance market pricing was next covered. Finally, students were asked if they knew what the DOA's transaction price in a mutually binding insurance contract would be equivalent to. Two undergraduate students and four graduate students correctly indicated that they believed they were negotiating for an optimal premium for healthcare coverage. Other replies included various cost-sharing plan provisions, healthcare costs, medical loss ratio, etc. Risk premium and medical loss ratios were then discussed at this point.

Insurance Markets

Part 2 is the longest. It consists of one game played by sub-group A, and another by sub-group B, in both the undergraduate finance and graduate insurance classes. Students were randomly assigned to form these two sub-groups per class and designate buyers and sellers under each sub-group.

Sub-group A in each class consisted of single sellers and buyers who may place bids on their own for a commodity which students in Part 1 had already correctly guessed to be health insurance. In contrast, buyers from sub-group B of each class can place bids only as one group and based on what they had agreed upon (e.g., by consensus or voting) for their minimum and maximum offers prior to the auction. Only their designated group leader can speak and bid on their behalf. Hence, sellers from sub-group B can only accept or ask based on the bid put forward by sub-group B's chosen leader in each class. For Part 2, sub-group A played first, followed by sub-group B, each for a series of three rounds owing to time constraints. One student was assigned again as recorder.

Similar to Part 1, buyers and sellers in sub-groups A and B of the two classes were allotted reservation prices and production costs in index cards. Reservation prices and production costs

differed from each other by no more than 10 percent. Auction rules, actions, strategies, and expected payoffs were the same as Part 1, except that any buyer's reservation value was not reduced for failing to win a bid in any round. However, a seller's stipulation was added to sub-group A's sellers. The stipulation requires sellers to increase their offer to sell (at or above production cost) by at least $10 (e.g., if production cost of one seller is $75, they can only accept or ask for at least $85 from the buyer). This stipulation is also revealed to buyers in sub-group A during the instructor's briefing prior to game start. Unsurprisingly, sellers in sub-group A ended up having considerably less flexibility to play given the (narrow) range they had to consider in asking or accepting. Without telling any of the players prior to debriefing, the $10 was intended to resemble the surcharge for higher risk and administration cost to insurers of individual (or non-group) health plans. Non-group plans are typically purchased by the self-employed, unemployed, and job-transitioning workers, as well as workers whose employers do not offer comprehensive or any plan options and might instead offer them a sum of money to purchase marketplace insurance on their own (Gillen, 2018). The seller's stipulation serves to exemplify and distinguish from production cost these risk factors and administrative burden that increase non-group premiums. Absent a seller's stipulation, it is difficult to impress on students, particularly in game form, the principle or rationale behind costlier non-group insurance premiums.

In contrast, sub-group B buyers and sellers in either class appear to replicate trades for (small or large) group plans sponsored by employers and associations of employers (e.g., association health plans). The buyer group and single sellers in sub-group B traded freely without any reduction to buyer reservation value for failure to win a bid and without the $10 minimum seller stipulation. In this regard, we hasten to point out that health plans in group insurance markets are much less expensive (and consequently, more affordable) than those sold in the non-group market as risk is more spread across a larger number of plan enrollees, and more people, through their employers and associations, are incentivized to buy into these plans (Morrisey, 2020).

Table 2 presents the results from our DOA Part 2. In both classes, we observed that buyers, especially in the undergraduate sub-group A, were more challenged in finding a mutually acceptable offer. Many sellers in the undergraduate sub-group A could not give in to any offer or ask below the range of $80 to $85, considering that they were assigned anywhere from $70 to $75 in production costs to begin with (i.e., before adding in the $10 minimum seller's stipulation). A few sellers from both classes also sought to increase gain beyond the $80 to $85 range. DOA rounds in sub-group B of either class, by comparison, resembled the iterated prisoner's dilemma rounds of Part 1 of the game, even if buyers had to bid as a single group. CE prices were much higher on average in sub-group A than B for both the undergraduate finance and graduate insurance classes.

Round 1 in the undergraduate sub-group A had to be canceled after three minutes, as single buyers and sellers failed to strike a deal between $80 and $85 at minimum. It also took these players longer to close one in rounds 2 and 3 compared to sub-group B. Nonetheless, there seemed to be greater excitement among the undergraduate players from sub-group A than sub-group B. Our student recorder observed that the single buyers were constantly looking into sellers' eyes as if to nudge them to give in to transaction prices less than $85 or so. On several occasions during the three rounds, sub-group A players were held up by repetitive buyer offers that the instructor and recorder flagged, in contrast to sub-group B. The resulting trades in round 2 ($85) and round 3 ($83) of sub-group A of the undergraduates were a lot higher than their counterparts in sub-group B (buyer group and single sellers), and therefore, closer to the spread of buyer reservation values than seller production costs.

Table 2
Non-group versus group insurance buyer and seller trading

Game components	Undergraduate Class*		Graduate Class*	
	Sub-group A	Sub-group B	Sub-group A	Sub-group B
Rounds	3	3	3	3
# of buyers	5	4	4	4***
# of sellers	5**	4	4**	4***
Buyer's reservation value ($ range)	86-90	90-95	86-90	90-95
Seller's production cost ($ range)	70-75	70-75	70-75	70-75
Transaction price ($)				
Round 1	0	83	85	82
Round 2	85	80	83	79
Round 3	83	79	84	78
Average (μ)	84	80.67	84	79.67
Variance (σ^2)	1.000	2.888	0.666	2.888
Standard deviation (σ)	1.000	1.699	0.816	1.699

*One student designated as recorder did not play.
**Seller's stipulation: any buyer's bid or seller's ask must be ≥ $10 + production cost.
***Sub-group A buyers and sellers were all randomly (re)assigned to sub-group B due to small graduate class size.

On the other hand, trading prices for undergraduate sub-group B (ranging from $79 to $83) were closer to the dispersion of seller (stipulation-free) production costs. The four buyers that made up one buyer group in the undergraduate sub-group B managed to negotiate lower prices in all three rounds, as sellers had more wiggle room in the absence of a seller's stipulation (to cover the risk and administrative surcharge of non-group premiums). Table 2 thus shows that, at CE, mean transaction price for the three rounds was $84 for sub-group A and $80.67 for B which had a wider dispersion of transaction prices. Like its variance, standard deviation was consequently much lower for sub-group A (1.0) than B (1.70), although both were close in value to the mean.

A similar trading pattern is noticeable in the graduate insurance class. However, we saw a higher level of cooperation and, consequently, reciprocity among the graduate players of sub-group A. No auction cancellation took place. This could mainly be a function of class size: only eight participating and more closely-knit graduate students compared to 18 undergraduate students. Small player number combined with inter-player recognition and trust seem helpful in mitigating bidding deadlocks. Nonetheless, unable to go lower than the same production cost range which was the seller's stipulation, graduate student sellers in sub-group A concluded trades with single buyers averaging $84 which was the same for the undergraduate sub-group A, albeit for all three

(not two) rounds. Sub-group B of the graduate insurance class traded at fairly similar prices (μ = \$79.67) as their undergraduate finance counterparts (μ = \$80.67). Standard deviation was equal for the two sub-groups B (1.70). With reservation values and production costs purposely held constant across similar sub-groups for experimental purposes, Part 2 outcomes in the graduate class came close to the undergraduate class.

Debriefing for Part 2 was also held at the end of the entire game but is discussed here for a more organized analysis relative to game outcomes. During the debriefing for Part 2 several students in the undergraduate and graduate classes said they believed they were once again bidding for health insurance. Sub-group A buyers in each class were then asked: Did you find it more or less challenging to offer in any or all rounds of Part 2? Without disagreement, buyers, especially among undergraduates, found bidding more "cumbersome" or "complicated" in light of the seller's stipulation, as sellers seemed "less willing to give in" or "cooperate" this time and would "only go as far as a few dollars less" than what buyers bid for.

In explaining the purpose of the seller's stipulation, sub-group A sellers were asked: What did you think the \$10 added to production cost was intended to cover/pay for as a seller's stipulation? Did you find it a barrier to trade? Sellers could not tell in the graduate insurance class. One buyer in the undergraduate finance class, to which a second student concurred, referred to it as the risk differential of a higher premium, which priced more buyers out of the insurance market. At that point, the instructor introduced the concept of individual or non-group risk and premium, indicated that the seller's stipulation was meant to resemble in game form, albeit imperfectly, the higher actuarial risk of small pools in non-group markets, and illustrated how (actuarial or claims) experience relates to insurance pricing but in a way that is typically more burdensome to buyers in the non-group market. Barriers to trading arising from the seller's stipulation was further discussed in terms of the smaller number of insurers and buyers in the non-group market and the protections imposed by the Affordable Care Act on non-group plans, including the prohibition against using pre-existing conditions and gender for enrollment eligibility and underwriting, age and smoker premium caps, maximum wait time, guaranteed issue and renewal, etc., all of which serve to increase risk and cost on the part of the insurer (Gillen, 2018; Morrisey, 2020).

The next part of the debriefing for Part 2 focused mainly on sub-group B in the two classes. They were asked: Did you find the auctions in Part 2 any different from Part 1? If so, in what way/s? Responses from undergraduate and graduate students did not considerably vary. Buyers generally felt that each Part 2 round was slower, or at times held up, since they had to decide as one single group (of 4 students in either class) and occasionally needed to be consulted by their chosen group leader while the auction was going on. Because graduate student players in sub-group A were randomly (re)assigned to sub-group B owing to their small class size, some had the advantage of hindsight. While they conceded that such instances of group strategizing may have slowed down the rounds a bit, it also gave them more time to "try to read each other's minds" and reach a "compromise" with transaction prices. Market prices for sub-group B of the graduate insurance class declined to the CE in each round, from \$82 to \$78. These were below reservation price and above production cost ranges, as can be gleaned from Table 2.

Thereafter, buyers and sellers in the two classes were asked: Why did you think buyers were constituted as one single group in sub-group B? No one answered correctly. But some students suggested that the "group approach" was intended to differentiate individual and corporate buyers, bids, and purchasing power and/or explore group decision-making in sub-group B, especially after these students had learned about the non-group market during the debriefing. At that point, group insurance was introduced by the instructor. Much discussion and feedback addressed how a larger

pool diversifies risk (loss) and, consequently, lowers the cost of risk factors and plan administration, as average closing prices in Table 2 imply ($79.67 to $80.67 for the two sub-groups B in contrast to $84 for the two sub-groups A). The law of large numbers, being foundational to group plan design and pricing, was next covered in terms of the entry of additional risks (i.e., more group enrollees) to an insured pool which reduces variation in the average loss per insured around the expected value (Smith & Kane, 1994). The choices employers have to make in selecting a suitable employee group health plan, with assistance from brokers/agents and insurers (Morrisey, 2020), was subsequently covered in the context of certain player dynamics that the undergraduate and graduate buyer groups (from sub-group B) experienced hands-on.

Market Intervention

In Part 3, undergraduate and graduate students were again randomly assigned as single buyers and sellers in their respective classes for indefinitely repeated rounds of trading for the same single commodity, which by now the instructor had confirmed as health insurance. All buyers and sellers were issued new index cards containing their respective reservation values and production costs just like in Part 1 for ease of outcomes comparison. Game rules, available player actions and strategies, target outcomes, and payoffs in Part 3 were essentially the same as Part 1. The difference lies in a buyer's stipulation that came in the form of a uniform buyer subsidy (%) in Part 3. Buyers and sellers alike were informed before playing that "the government" had opted to intervene in the insurance marketplace by granting an 80 percent (a game variation can be a flat dollar) subsidy of the negotiated or closing price to the winning buyer at each iterated round. The experimental game devised by Gillette (1994) in our literature review conceals this third-party subsidy from sellers, thereby altering player bidding incentives.

In Part 3, a buyer who, for example, is assigned an $80 reservation value and hammers a deal with a seller for a $70 transaction price will not simply gain the $10 buyer payoff in Part 1, but as much as $66 or over six times more after the buyer's obligation is reduced to $14:

Buyer payoff = reservation value - transaction price + (transaction price x government subsidy)

$$= \$80 - \$70 + (\$70 \times 0.80)$$
$$= \$10 + \$56$$
$$= \$66$$

Sellers, on the other hand, have to ask based on their assigned production cost just like in Part 1. Trading in Part 3 goes on until a buyer's bid or a seller's ask is accepted, or the round is canceled, or another round is called by the auctioneer. Three rounds in Part 3 were allowed by the auctioneer for the undergraduate finance and graduate insurance classes prior to debriefing. Unlike the other games in our reviewed literature, payoffs from all three parts of our DOA were tallied at the end of Part 3 to determine which students will qualify for credit (either in-kind like candies or bonus test points).

Part 3 outcomes are summarized in Table 3. Deal-closing prices per round climbed to the mid-$90s in the two classes, save for round 1 of the graduate class, which one might consider as the starting point of their learning curve and which still stood high at $89. Such closing prices are in stark contrast to the high $70s and mid $80s of closing prices in every round of this game's Parts 1 and 2. Moreover, mean closing prices listed in Table 3 were unusually high (graduate = $92; undergraduate = $95) compared to Parts 1 and 2 (ranging from $79.67 to $84.40). Undergraduate and graduate sellers managed to successfully drive up most bids per round, knowing that the 80 percent government subsidy would leave buyers with much higher payoffs. By this time, too,

bargaining, cooperation, and reciprocity seem to have evolved between buyers and sellers in many instances, as variations of the same game had been played and repeated without a known terminus or final round. Collectively, the results of Part 3 support Gazda et al. who concluded that "in aggregate, third party interventions seemed to result in a rise of the market price" (2012, p. 52).

Table 3
Government-subsidized buyer and seller trading

Game components	Undergraduate Class*	Graduate Class*
Rounds	3	3
# of buyers**	9	4
# of sellers	9	4
Buyer's reservation value ($ range)	90-100	90-100
Seller's production cost ($ range)	70-78	70-78
Transaction price ($)		
Round 1	93	89
Round 2	95	93
Round 3	96	94
Average (μ)	94.67	92
Variance (σ^2)	1.555	4.666
Standard deviation (σ)	1.247	2.160

*One student designated as recorder did not play.
**Buyer's stipulation: buyer's government subsidy = 80 percent of transaction price per round.

When the buyer's government subsidy is taken into account, what might initially appear surprising becomes comprehensible. The 12-minute break prior to play afforded students on both sides of the game enough time to craft bidding and asking strategies. Buyers came to understand (some confirming with the instructor) that a higher bid that came closer to, or even at the maximum of, their reservation values would still be largely replaced by the uniform third-party subsidy for the deal-closing (premium) price. The higher number and amount of buyer bids offered in Part 3 in our two classes signify intensified risk-seeking than those we had previously witnessed in Parts 1 and 2 of this DOA game. At the same time, we observed that sellers, particularly in the undergraduate finance class, were more persistent in asking higher prices. Undergraduate trading prices that resulted from these push-and-pull exchanges ranged from $93 to $96 and corresponded to buyer payoffs in the range of $71.40 and $80.80. The slightly wider spread of our graduate insurance class's trading prices from $89 to $94 meant potential buyer payoffs that were not far off at $72.20 to $81.20. By marked comparison, buyer payoffs in the two classes did not exceed $19 in Part 1 and $17 in Part 2 of our game.

Part 3 trading prices in the undergraduate finance class were above CE, with several bids breaching buyer reservation values; those for the graduate insurance class had one exception that stood just outside of buyer reservation values (round 1 = $89). Standard deviation in both classes was higher than 1.0, but it was considerably more pronounced in the graduate class at 2.16 as a result of the much wider distribution of the data sets compared to Parts 1 and 2 of this DOA game. The same may be said about the variance in Table 3.

Part 3 proved to be the most worked-out but also the most animated DOA. Most buyers in the finance and insurance classes confirmed that they bid based on their calculated payoffs after reducing their offers by the government subsidy when asked during the debriefing why their bids were much higher and up to the mid-$90s. The few buyers that either did not methodically consider the subsidy in bidding or miscalculated it were quickly overtaken by other buyers (and sellers) in both classes. Several sellers, on the other hand, stated that they were intent on asking for much higher prices in Part 3 because they knew that buyers would also profit by much more, although they did not know the buyers' reservation values. Many were surprised to find out during the debriefing that the buyer payoffs ranged as high as $70+ to $80+. The instructor then explained that the higher payoffs directly induced by the buyer-only subsidy breed the asymmetric informational problem of hidden action or "moral hazard" in (health) insurance. That is because buyers (or the insured) have the incentive to increase their exposure to risk when they do not bear the full cost, and therefore, cannot "internalize" the consequences. This causes "rational agents [to become] less rational in terms of average market price, after intervention of a third party on the market … which raises the average market prices presenting a manifestation of moral hazard" (Gazda et al., 2012, p. 53). We also discussed during the debriefing how people who adversely select into insurance could possibly increase moral hazard (i.e., moral hazard on demand).

Potential sources of adverse selection and moral hazard under the Affordable Care Act were then identified, including the 10 essential health benefits, prohibition against benefit caps and pre-existing conditions for enrollment eligibility, etc. (Mendoza, 2017). One student from the graduate class asked whether physicians "could also be a moral hazard?" The instructor indicated that physicians, clinics, labs, hospitals, and many other providers profit from provider moral hazard "by carrying out more services than insured patients really need" or inducing insured patients "to want or 'need' more tests and procedures, [such as h]igh-margin testing and elective surgery" (Light, 2021).

Finally, the instructor clarified that third-party intervention does not necessarily have to be exclusively by government. Any fully funded or self-funded insurance coverage and marketplace health plans that accept premium tax credits are, technically, forms of third-party intervention. We devoted the last portion of our debriefing to discussing how entitlement programs, notably Medicare, Medicaid and the Children's Health Insurance Program (CHIP), have helped increase healthcare costs in the U.S. to the point where about 18 percent of national income now goes to healthcare expenditures alone. The U.S. federal government, since the enactment of the Affordable Care Act that expanded Medicare and Medicaid, has also turned into the single largest healthcare payer amid a ballooning budget deficit and debt burden and countless instances of Medicare and Medicaid fraud and abuse (Morrisey, 2020; Light, 2021). For these reasons, entitlement payments are expected to remain a legislative reform priority for years to come (Morrisey, 2020).

Gaming Pedagogy: Future Directions for Assurance of Learning

Our findings suggest that gaming pedagogy is a valuable method that combines game mechanics with traditional instruction to create engaging learning experiences at the undergraduate and graduate levels and help finance and insurance students apply concepts in the real world. However, assessment of learning outcomes from our game-simulation may be limited and largely anecdotal during this pilot phase of this gaming project. To make learning assessment more systematic and broaden assurance of learning, an instructor or researcher using or adapting our game might consider the following suggestions:

1. Conduct pre- and post-knowledge tests among the student players. These should at least cover basic risk and insurance concepts and principles, insurance market players, and some decision-making approaches and strategies available to these players. These tests can track and assess changes and improvements in comprehension and analytical skills as well as motivational and attitudinal changes among undergraduate and graduate students, and how they might vary.
2. Test statistically for differences in undergraduate and graduate game outcomes in DOA Parts 1, 2, and 3. Test results might point to the need to vary, simplify, enhance, and/or clarify game rules and incentives, the composition of and dynamics arising from player groups, the breadth and depth of debriefings and feedback, and use of appropriate metrics to evaluate learning gains and challenges.
3. Gather, organize, and evaluate data from player bids and asks in a way that allows the instructor and/or researcher to assess how indefinite iterations or repetitions of game processes and outcomes may have changed and improved player skills and competencies, attitudes, and actual performance.
4. Besides determining and reporting (e.g., announcing in class) who qualified for credit (candies, bonus point, etc.), it might help to track which and how many of the players qualified at each part of the DOA game to assess whether their motivations and attitudes also adjusted or changed (e.g., to gain more or maximize credit by "strategizing" differently, seeking to close deals faster, "warding off" potential deals that their fellow players could gain, etc.).
5. Bridge theoretical concepts with practical applications by drawing from and relating them to students' personal knowledge, preferences, and experiences concerning risks and risk-taking, including their psychosocial and cultural implications. Integrating the social experiences of game-based learning and bringing them into the classroom will help increase assurance of learning in healthcare finance and insurance courses.
6. Consider adapting this DOA game to different class sizes and time and other resource constraints. This can be done after assessing player performance, motivational and attitudinal changes, and game outcomes relative to the games' (varying) time requirements/allocation, role assignments, and players' comparative advantages and challenges. For instance, one could assess whether the reassignment of graduate students from one group (or sub-group) to another during the same part of the game, owing to the generally smaller size of graduate classes, affects game outcomes in comparison to their undergraduate counterpart/s.

Conclusion

In a three-part simulation game that combines a DOA with the iterated elements of the prisoner's dilemma, we sought to introduce students to the basic concepts, principles, methods, strategies, and problems of healthcare risk management through the vehicle of insurance. We find a DOA, with bidding rounds that could repeat indefinitely, relevant and practical in engaging students hands-on. Student learning in a course whose content is technical, if not dry, can be irrigated with creative gaming pedagogy which encourages active, continuous, and simultaneous participation, such as a game of multiple buyers and sellers offering at the same time. This game format further allows for signaling, reputational recognition and trust, and the reciprocal benefits of cooperation to gradually evolve, especially as the game repeats with some variations on a theme. Bargaining and compromise — as we saw in all three parts of our game — constitute the "art of the deal" to borrow from the title of one widely read book (Trump with Schwartz, 1987).

Because finance as a field of study refers to the practice of managing money, it encompasses how individuals, firms, and governments acquire and spend funds. Risk management is critical in identifying, assessing, and mitigating potential financial risks associated with these funds. By building in buyer and seller stipulations that create and differentiate between incentives (and disincentives) in an indefinitely iterated DOA, game-theoretic simulations, such as the one we have devised for (health) insurance purposes, also allows for comparative financial analysis: Healthcare risks with and without insurance coverage, group versus non-group markets, and the effects of third-party subsidy intervention on pricing and asymmetric information, particularly adverse selection and moral hazard. Game iterations demonstrate and encourage us to assess how players respond and adjust to the opportunities and challenges before them as they craft their own "art of the deal" to maximize healthcare coverage value and minimize financial risk.

In light of the foregoing, we suggest finance and insurance courses consider incorporating game-theoretic simulation as an instructional modality. Game theory can be adapted for classroom use and carefully designed to captivate and inspire student learning, and from what we observed in several instances of our DOA game, perhaps even induce some belly laughs (for instance, as students became deadlocked over offers and counter-offers). We find that such approach leads to a creative blend of formal and fun to make magic in the classroom. And that by itself is a challenge and a reward.

References

Amin K, Cox C, Ortaliza J, Wager E (2024). Health care costs and affordability. In Altman, D, ed. Health Policy 101, np. Available at: https://www.kff.org/health-policy-101-health-care-costs-and-affordability/.

Asrafian H, Darzi A, Athanasiou T (2011). Evidence synthesis: Evolving methodologies to optimise patient care and enhance policy decisions. In Athanasiou T, Darzi A, eds. Evidence Synthesis in Healthcare, 1-47. London: Springer-Verlag.

Assadi S, Khanna S, Li Y, Vohra R (2017). Fast convergence in the double oral auction. ACM Transactions on Economics and Computation, 5(4): 1-18.

Axelrod, R. (1984). The evolution of cooperation. New York: Basic Books, Inc.

Celly K (2007). Fostering critical thinking in business courses: Pedagogical innovations and strategies. The Business Review, 8(1): 148-153.

Gazda V, Kubák M, Maliková Z, Gróf M (2012). Double oral auctions and tendencies toward moral hazard. International Journal of Humanities and Social Science, 2(17): 49-54.
Gillen E (2018). Why grow the nongroup insurance market: The freelance economy demands it. RTI International, Deb]cember 11. Available at: https://www.rti.org/insights/why-grow-nongroup-insurance-market-freelance-economy-demands-it
Gillette D (1994). Double oral auctions and health Care. Paper presented at the Economics Science Association Annual Meeting (November 1994)
Gokhale CS, Traulsen A (2010). Evolutionary games in the multiverse. Proceedings of the National Academy of Sciences, 107(12): 5500–5504.
Light DW (2021). The three moral hazards of health insurance: The troubling history and misuse of the term. Princeton, NJ: Institute for Advance Study. Available at: https://www.ias.edu/ideas/three-moral-hazards-health-insurance
MacLeod NM, Smith SR (1984). Coverage: a simulation game to teach health care financing. The Journal of Family Practice, 18(4): 587-589.
Mellor J (2005). Illustrating adverse selection in health insurance markets with a classroom game. Southern Economic Journal, 72(2): 502-515.
Mendoza RL (2017). Information asymmetries and risk management in healthcare markets: The U.S. Affordable Care Act in retrospect. Journal of Economic Issues, 51(2): 520-540.
Mendoza RL (2022). Benefits and costs of financing accounts receivable portfolios in the healthcare industry. Business Forum, 29(1): 39–51.
Mendoza RL (2023). Does health economics inform healthcare management in graduate business programs? Journal of International Business Education, 18(1): 209–228.
Morrisey, MA (2020). Health Insurance, 3rd ed. Chicago, IL: Health Administration Press.
Rust J, Palmer R, Miller J (1989). A double auction market for computerized traders. SFI working paper: 89-001, Santa Fe Institute, San Fe, New Mexico. Available at: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=ab5d32fe20083e98c976ecfb532593be7a21f144.
Shurkin, J. (2013). Game theory tackles rising health care costs. Inside Science, April 12. Available at: https://www.insidescience.org/news/game-theory-tackles-rising-health-care-costs.
Smith MS (2003). Game Theory. In Burgess G & Burgess H, eds. Beyond intractability, np. Boulder: Conflict Information Consortium, University of Colorado. Available at: http://www.beyondintractability.org/essay/prisoners-dilemma.
Smith ML, Kane SA (1994). The law of large numbers and the strength of insurance. In Gustavson SG & Harrington SE, eds. Insurance, risk management, and public policy, pp. 1-27. New York: Springer Science+Business Media, LLC.
Trump DJ with Schwartz T (1987). Trump: The art of the deal. New York, NY: Random House.
Weinberg S, Mcllwain TF, Stegall S, Willis K (2013). The health insurance risk game: A simulation for meaningful learning. The Journal of Health Administration Education, 30(2). 127-136
Zelmer AE, Zelmer ACL (1980). Health care. In Horn RD & Cleaves A, eds. The Guide to Simulation/Games for Education and Training. Beverly Hills, CA: Sage Publications.

From Stage to Stocks: Understanding IPO Underpricing using Taylor Swift's the Eras Tour

Matthew Faulkner
San José State University

When students first learn about Initial Public Offerings (IPOs), the idea of pricing a soon-to-be publicly traded company seems opaque and difficult. While exploring theories of IPOs, this paper presents a simple way to illustrate the concept of IPO underpricing using a pop icon as an analogous example, Taylor Swift ticket prices for the Eras Tour. By comparing the initial sale prices of concert tickets to their secondary market prices, we effectively demonstrate the concept of IPO underpricing. This analogy clarifies a theory of IPO pricing and engages students by relating it to a familiar pop-culture context. Following psychology and pedagogy literature, this approach enhances overarching comprehension and retention of the IPO process, making a complex financial topic more accessible, relatable, and memorable.

JEL Classification: A20, A22, A23, G32

Keywords: IPO, IPO Underpricing, Taylor Swift, The Eras Tour, Pedagogy

Introduction

Understanding Initial Public Offering (IPO) underpricing theory, and IPOs in general, can be a difficult topic for many students. Students (and investors for that matter) are often intrigued with the amount of capital raised in an IPO, the excitement of new publicly traded companies they can invest in, the "going public" process, and the returns IPOs experience – specifically, the rather large first-day gains observed in IPO pricing data.

When teaching topics such as first-day gains and theories around IPO underpricing as an aspect of the IPO landscape, students are often perplexed. Yet, confusion comes in many forms during the IPO process (e.g. selecting an underwriter, the book-building process, and/or getting to the final offer price). In this paper, we discuss how Taylor Swift's *Eras Tour* can provide a simple, effective, and memorable way, to teach about one major IPO concept, IPO underpricing.

Teaching IPOs can be particularly challenging given that students have commonly not been involved with an IPO, have not discussed the depths of IPOs in introductory finance courses, nor do IPOs feel relatable to anything they have typically done in their lives thus far; which is where this paper strives to close a gap. In addition, most students do understand the basic IPO idea – a company transitions from private ownership to public by selling ownership (shares) to raise capital. This serves as a way to exit or realize gains for founders and early investors, as well as, increase capital that the company can use to expand and grow.

Interestingly, in the secondary market for the *Eras Tour* concert tickets, a notable phenomenon has been observed: tickets, once purchased at the initial sale price (Price A), often command significantly higher resale prices almost immediately (Price B). This phenomenon is not necessarily unique to Taylor Swift's concert tickets to other concerts or events, but has become increasingly visible in mainstream media. The immediate markup in secondary ticket prices can

be likened to the phenomenon of IPO underpricing in financial markets. IPO underpricing occurs when the initial offering price of a stock is initially sold below its closing price on the first day of trading. This leaves a lot of "money-on-the-table", the change between the offer price and first closing market price multiplied by the number of shares outstanding for the firm, or by number of tickets for Taylor Swift. It also results in immediate gains for the initial IPO investors. Similarly, initial purchasers of tickets often experienced substantial gains if they chose to sell their tickets shortly after purchase.

Building upon pedagogical and psychological research, we attach Taylor Swift concert ticket sales to the field of finance, using ticket pricing and resales from a contemporary titan in the entertainment industry as a tool to help memory and understanding. Pedagogical research shows that it is necessary for learners to attach a new piece of information to an old one, or it just won't stick (Willis, 2006 p.44). Moreover, personal connections between the information to be learned and students' lives build greater memory storage and more efficient memory retrieval pathways (Tyng et al., 2017), and regardless of being a fan, Taylor Swift commonly adds a personal connection for most students given her mainstream popularity (Coscarelli et al., 2024). Additionally, moments that are experienced, seen, or heard that hold some personal meaning are more likely to build long-term recollections available for later retrieval (Eich, 1995); we hope this pedagogical teaching tool accomplishes this for a majority of students. Ultimately, it is important that students build an understanding of the IPO process and theories that exist, whether in corporate finance, capital markets, or investments courses.

A Teaching Exercise for IPOs and Underpricing

Preface

This project specifically addresses an IPO underpricing teaching tool. A typical course would certainly cover the IPO process prior to addressing the theory of underpricing. This project does not create a teaching tool for the IPO process, but recommends that instructors address why a company goes public prior to teaching concepts such as IPO pricing and returns. These may include life-cycle theories and market-timing theories (Ritter and Welch, 2002). Additionally, one should explore the process of going public, from selecting an underwriter to understanding the book-building process with road shows. Instructors may also explore the number of shares offered in an IPO and the types of cash offer IPO underwriting styles: firm commitment, best efforts, and Dutch auction. Table 1 provides a sample teaching topic outline for IPOs that can be adapted to the instructor's course.

As a course progresses from the topic of a company being sold via an IPO during the primary issuance to its availability on an exchange for the secondary market transactions, the pricing theories begin to take shape. As discussed earlier, IPO underpricing occurs when the initial offer price is sold below its closing price at the end of the first day of trading. IPO researchers have continually tried to understand why IPO underpricing exists by developing theories, briefly discussed in closing steps section.

When teaching IPO underpricing, instructors typically first address *what* is IPO underpricing, then to move onto *why* IPO underpricing exists, if an instructor even continues to this second question at all. This project sits at the front end of the discussion of *what* is IPO underpricing. In the course setting, the analogy of having an alternate real-world event to the *what* is underpricing question allows the students to then move onto theories of *why* underpricing exists with a clear

picture of IPO underpricing as a concept. The underpricing analogy of the Taylor Swift the *Eras Tour* tickets does little to help understand the *why* IPO underpricing exists, as entertainment ticket pricing and IPO stock pricing have unique dynamics. Therefore, we leave it to its straightforward purpose: to help students understand *what* is IPO underpricing, building a clear foundation to continue into *why* IPO underpricing exists in future class discussions.

Table 1
How to use the Eras Tour in IPO Teaching Sample

This table provides a sample outline of teaching the IPO process and where to incorporate the ideas presented in this paper. Specifically, this project provides an analogy to employ for the Step 4 IPO Pricing: What is IPO Underpricing? Other steps are generic recommendations.

Step	Topic	Description
1	IPO Process	Why a company goes public? ● For example, life-cycle or market-timing theories.
2	IPO Process	Selecting an underwriter, syndicates, the book-building process, road shows. Getting to the Offer Price
3	IPO Pricing and Taylor Swift	What is IPO underpricing? ● Taylor Swift's the Eras Tour analogy. ● Showcase Figure 1, Table 2, Table 3 and Table 4
4	IPO Pricing	Why underpricing exists. For example: ● The winner's curse (Rock, 1986), ● Prospect theory and non-pecuniary benefits (Loughran and Ritter 2002), ● The changing risk composition, the realignment of incentives, and the changing issuer objective function hypotheses (Loughran and Ritter, 2004)
5	IPO Performance	Long-run performance of IPOs. For example: ● IPO performance in the long-run (Ritter 1991; Loughran and Ritter 2004)

Furthermore, this tool not only provides a pedagogical analogy to understand IPO underpricing as a concept, but it additionally allows students to better understand the entire IPO process through engagement, discussions, and recollection of learning. When students take information or a concept and connect it to their existing knowledge and experiences it achieves deep learning that lasts (Barkley and Major, 2020). Anecdotally, the use of this tool in a classroom has indeed increased the understanding of the entire IPO process. Even if the Taylor Swift analogy has been built for only one direct purpose, its benefits were far-reaching.

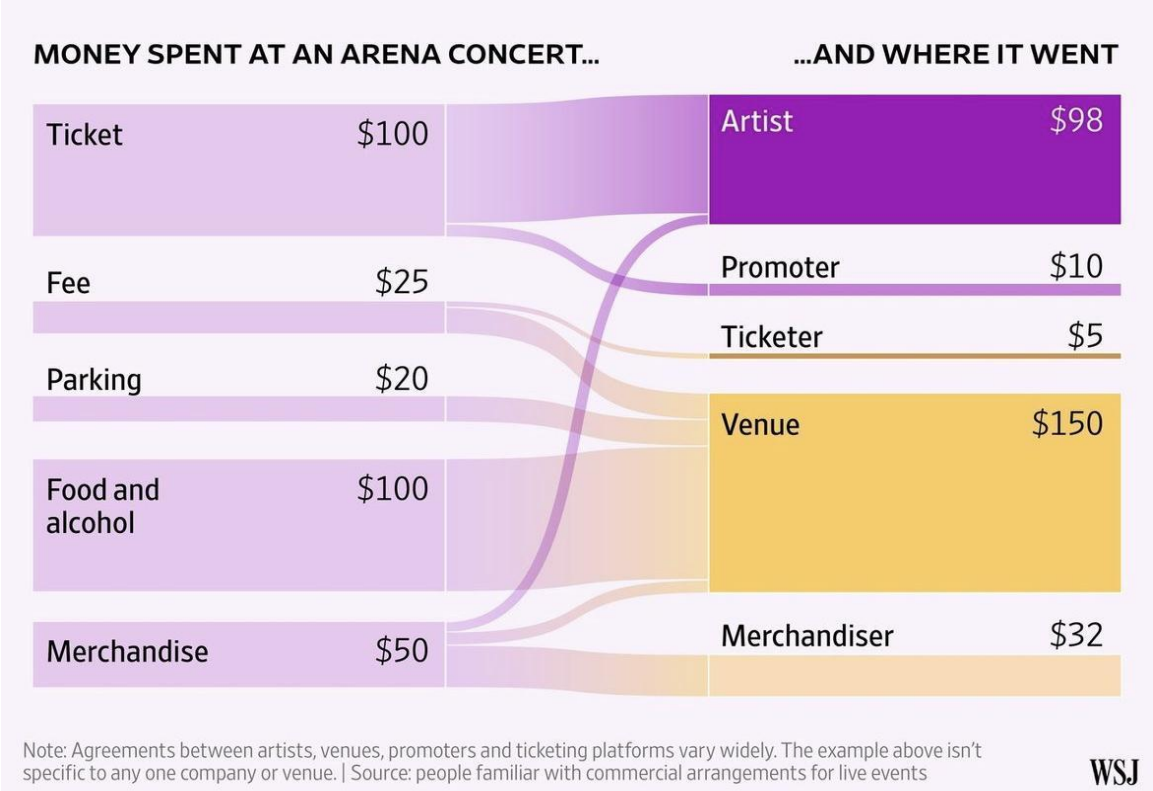
Using the Teaching Tool

To use this teaching methodology, we recommend the following procedure. First, showcase Taylor Swift's the Eras Tour original average ticket prices relative to resale ticket prices, we assume these average resale prices occurred the day they were bought (i.e. typically on ticket release day). As of July 2023, the average face value standard ticket price for the Eras Tour was $253, ranging from $49 to $499, a sizeable amount (Fox, 2023). However, the resale average ticket prices reached astonishing heights with an average of $1,088, ranging from $800 to $11,000 (Gendron, 2023; Vega, 2023). The students may well be wide-eyed and bewildered at that difference, although many may not be all that surprised if they were keeping up with daily media reporting. Continuing, correlate that these tickets were bought at the offer price in the primary market, the analogous "IPO". They were then [immediately] sold in the secondary market for a large markup. The obvious thought, one anecdotally invoked by many students, may be "wow,

Taylor Swift, could/should have sold the original tickets for more!" Since much of the ticket money flows to the artist, we likened this to initial flows of capital to a company; hence, Taylor Swift's the Eras Tour [IPO] underpricing. Figure 1 displays a typical breakdown of concert revenue streams and to whom the money flows.

Figure 1
Concert Money Flows

This figure shows a standard structure of money spent on concert tickets and at concerts, as well as, to whom the money flows, courtesy of the Wall Street Journal (Steele and Dapena, 2024). Generally, most of the ticket revenue flows to the artist.



Next, introduce students to historical IPO data to showcase IPO underpricing in financial markets. Table 2, courtesy of Jay Ritter (2024b, p. 2), showcases the largest underpricing events in U.S. stock market history. Interestingly, the largest IPO underpricing in U.S. market history is not shown in Table 2 since it was an American Depository Receipt (ADR) IPO; that of Alibaba. ADRs are outside the scope of this project; however, we'd be remised not to define it in a pedagogical article. An instructor could spend a moment to introduce what is an ADR depending on the course if they choose to use the Alibaba example. An ADR is a negotiable certificate issued by a U.S. bank representing a specified number of shares (or a fraction of a share) in a foreign company's stock. ADRs trade on U.S. stock markets like regular domestic shares and are a popular way for American investors to invest in foreign companies without dealing with the complexities of foreign stock markets.

Table 2
Money Left on the Table

This table, courtesy of research by Jay Ritter (2024b), displays the money left on the table for IPOs, listed by size of money left on the table. The column Dollar amount left on the table captures the difference between the columns Offer Price and First closing market price multiplied by column Number of shares offered. Columns for Company, IPO Date, and Ticker Symbol are also provided. Find the entire IPO underpricing table beginning on page 2: https://site.warrington.ufl.edu/ritter/files/money-left-on-the-table.pdf

Dollar amount left on the table	Company	IPO date	Offer Price	First closing market price	Number of shares offered	Ticker symbol
$5,075,000,000	Visa	080319	$44.00	$56.50	406,000,000	V
$3,937,028,063	Airbnb	201210	$68.00	$144.71	51,323,531	ABNB
$3,750,040,000	Snowflake	200916	$120	$253.93	28,000,000	SNOW
$3,477,690,000	Rivian Automotive	211110	$78.00	$100.73	153,000,000	RIVN
$2,887,830,000	DoorDash	201209	$102	$189.51	33,000,000	DASH
$1,852,500,000	Coupang	210311	$35.00	$49.25	130,000,000	CPNG
$1,586,300,000	United Parcel Service*	991110	$50.00	$68.125	87,520,000	UPS
$1,540,730,469	Corvis	000728	$36.00	$84.71875	31,625,000	CORV
$1,496,000,000	Snap	170302	$17.00	$24.48	200,000,000	SNAP
$1,365,500,000	Bumble	210211	$43.00	$70.31	50,000,000	BMBL
$1,323,000,000	Twitter	131107	$26.00	$44.90	70,000,000	TWTR
$1,312,437,500	Palm	000302	$38.00	$95.0625	23,000,000	PALM
$1,281,747,555	Royalty Pharma PLC	200616	$28.00	$44.50	77,681,670	RPRX
$1,189,598,400	Affirm Holdings	210113	$49.00	$97.24	24,600,000	AFRM
$1,096,956,250	Sycamore Networks	991022	$38.00	$184.75	7,475,000	SCMR
$1,072,687,500	Akamai Technologies	991029	$26.00	$145.1875	9,000,000	AKAM
$959,100,000	Goldman Sachs*	990504	$53.00	$70.375	55,200,000	GS
$939,435,000	Abcellera Biologics	201211	$20.00	$58.90	24,150,000	ABCL
$920,700,000	VA Linux	991209	$30.00	$239.25	4,400,000	LNUX
$846,809,000	Kenvue	230504	$22.00	$26.90	172,812,560	KVUE
$835,200,000	Free Markets	991210	$48.00	$280.00	3,600,000	FMKT
$816,000,000	Avanex Corporation	000204	$36.00	$172.00	6,000,000	AVNX
$811,468,774	PagSeguro Digital Ltd	180124	$21.50	$29.20	105,385,555	PAGS
$801,281,304	Qualtrics	210128	$30.00	$45.50	51,695,568	XM
$782,595,000	C3.ai	201209	$42.00	$92.49	15,500,000	AI
$772,650,000	Agilent Technologies*	991118	$30.00	$42.750	60,600,000	A
$754,800,000	Elanco Animal Health	180920	$24.00	$36.00	62,900,000	ELAN
$728,242,000	Webmethods	000211	$35.00	$212.62	4,100,000	WEBM
$726,750,000	AT&T Wireless*	000427	$29.50	$31.875	306,000,000	AWE
$726,000,000	VMware	070813	$29.00	$51.00	33,000,000	VMW
$722,120,000	Yandex NV	110524	$25.00	$38.84	52,174,088	YNDX
$719,531,250	McData Corp.	000809	$28.00	$85.5625	12,500,000	MCDT
$686,375,000	Array Technologies	201015	$22.00	$36.45	47,500,000	ARRY
$656,250,000	Foundry Networks	990928	$25.00	$156.25	5,000,000	FDRY
$629,100,000	Doximity	210624	$26.00	$53.00	23,300,000	DOCS
$616,401,360	KPMG Consulting Inc	010208	$18.00	$23.48	112,482,000	KCIN
$605,769,220	GoodRX	200923	$33.00	$50.50	34,615,384	GDRX
$604,035,000	Chewy	190614	$22.00	$34.99	46,500,000	CHWY
$600,000,000	Genentech	990720	$97.00	$127.00	20,000,000	DNA
$580,000,000	Intersil Holding	000225	$25.00	$54.00	20,000,000	ISIL
$578,500,000	ZoomInfo Technologies	200605	$21.00	$34.00	44,500,000	ZI
$569,250,000	Storage Networks	000630	$27.00	$90.25	9,000,000	STOR
$568,820,000	General Motors	101118	$33.00	$34.19	478,000,000	GM
$553,181,250	Finisar	991112	$19.00	$86.875	8,150,000	FNSR
$552,000,000	Intrepid Potash	080421	$32.00	$50.40	30,000,000	IPI
$550,375,000	Vroom	200609	$22.00	$47.90	21,250,000	VRM
$547,200,000	Petco Health and Wellness	210114	$18.00	$29.40	48,000,000	WOOF
$542,608,690	Zoom Video Communic.	190418	$36.00	$62.00	20,869,565	ZM
$541,875,000	Charter Communications*	991109	$19.00	$22.75	144,500,000	CHTR
$541,328,968	Blackstone Group LP	070621	$31.00	$35.06	133,333,000	BX

In 2014, Alibaba Group Holding (BABA) issued ADRs in the U.S. and, one could argue, left over $8 billion on the table. Alibaba's IPO occurred in September 2014 at the stock price offering of $68 per share. At the end of the first day of trading, Alibaba's share price closed at $93.89, a 38.1 percent increase. Arriving at the $8 billion undervaluation requires some basic arithmetic and the instructor may ask students how this number was reached. Walking through it step by step, on Alibaba's IPO date, the company sold 320.1 million shares. With an underpricing of $25.89 on 320 million shares, Alibaba left approximately 320 million x $25.89 = $8.3 billion on the table. Again, emphasize that money left on the table is the difference between the offer price and the first-day closing stock price multiplied by the number of shares issued in the IPO. Note to students that the implied notion of money left on the table is that if the company's stock is selling for $X price at the end of the first day of trading, then the company, in theory, could've sold it for $X for the offer price of the IPO. Other notable IPOs that students would recognize include VISA, Airbnb, Snap, Bumble, Twitter, or Chewy each leaving $5 billion, $4 billion, $1.5 billion, $1.4 billion, and $600 million, on the table, respectively.

To offer a comparison, Table 3 provides an analogy of the IPO underpricing of VISA to Taylor Swift's the *Eras Tour* tickets. We cannot provide exact underpricing numbers for the *Eras Tour* as we do not have exact data after countless attempts to work with Ticketmaster and comparable firms. Secondly, a distinction in ticket "underpricing" to IPO underpricing is that tickets may, or may not, occur in a single-day transaction, (i.e. bought Monday morning and sold Monday afternoon); therefore, underpricing would only be analogous to IPOs for the purchase prices and resale prices on the first sale date. If, for example, the average *Eras Tour* ticket price has an offer price of $253 and a closing same-day resale price of $1,088 and a show sold 70,000 tickets (the average for the tour), this would lead to $835 x 70,000 = $58,450,000 left on the table per show. Additionally, Table 4 provides U.S. IPO underpricing data per year from 1980-2023 (Ritter, 2024a, p. 2). We note IPO underpricing, whether small or large, occurs in every year except 2008 and 2023 showcasing the prevalence of the subject.

Table 3
Visa Underpricing analogy to Taylor Swift Underpricing

This table is for illustrative purposes. The left segment of the table shows the VISA IPO. First, the offer price of the IPO was $44. The first day closing market price was $56.50. The firm left $5 billion on the table. The right segment of the table shows the Taylor Swift analogy. The average ticket price was $253. The average resale price was $3,801. We have tried multiple attempts to gather ticket pricing data from Ticketmaster and other platforms to no avail. Therefore, we rely on relevant news sources. However, the exact ticket resale values are not necessary, the data we have serves well for illustrative and pedagogical purposes.

VISA IPO		**The Eras Tour**	
Offer Price	$44	**Ticket Face Value**	$253
First Day Closing Price	$56.50	**Resale Price**	$1,088
Shares	406 million	**Tickets**	72,000/show
Underpricing	$12.50	**Underpricing**	$835
Money Left on the Table	$5,075,000,000	**Money Left on the Table**	$60,120,000/show 152 shows $9,138,240,000

Table 4
IPO Underpricing by Year

This table, courtesy of research by Jay Ritter (2024a), displays IPOs by year from 1980 to 2023. Column *Number of IPOs* shows the count of IPOs in the respective year. The subsequent two columns capture mean first-day returns, followed by the median first-day return. These three columns capture IPO underpricing percentages. *Aggregate Proceeds* exclude overallotment options- the option for underwriter to sell extra shares in high demand issues. The amount of money left on the table is defined as the closing market price on the first day of trading minus the offer price, multiplied by the shares offered. Page 2 of https://site.warrington.ufl.edu/ritter/files/IPOs-Underpricing.pdf

Year	Number of IPOs	Mean First-day Return		Median First-day Return	Aggregate Amount Left on the Table	Aggregate Proceeds
		Equal-weighted	Proceeds-weighted			
1980	71	14.3%	20.0%	6.9%	$0.18 billion	$0.91 billion
1981	192	5.9%	5.7%	0.3%	$0.13 billion	$2.31 billion
1982	77	11.0%	13.3%	3.7%	$0.13 billion	$1.00 billion
1983	451	9.9%	9.4%	2.5%	$0.84 billion	$8.89 billion
1984	171	3.7%	2.5%	0.0%	$0.05 billion	$2.02 billion
1985	186	6.4%	5.6%	2.8%	$0.23 billion	$4.09 billion
1986	393	6.1%	5.1%	1.1%	$0.68 billion	$13.40 billion
1987	285	5.6%	5.7%	1.3%	$0.66 billion	$11.68 billion
1988	105	5.5%	3.4%	2.3%	$0.13 billion	$3.88 billion
1989	116	8.0%	4.7%	4.0%	$0.27 billion	$5.81 billion
1990	110	10.8%	8.1%	5.6%	$0.34 billion	$4.27 billion
1991	286	11.9%	9.7%	7.5%	$1.50 billion	$15.39 billion
1992	412	10.3%	8.0%	4.4%	$1.82 billion	$22.69 billion
1993	510	12.7%	11.2%	6.3%	$3.52 billion	$31.44 billion
1994	402	9.6%	8.3%	4.2%	$1.43 billion	$17.18 billion
1995	462	21.4%	17.5%	13.2%	$4.90 billion	$27.95 billion
1996	677	17.2%	16.1%	10.0%	$6.76 billion	$42.05 billion
1997	474	14.0%	14.4%	9.4%	$4.56 billion	$31.76 billion
1998	283	21.9%	15.6%	8.9%	$5.25 billion	$33.66 billion
1999	476	71.2%	57.4%	37.5%	$37.11 billion	$64.67 billion
2000	380	56.3%	45.8%	27.9%	$29.68 billion	$64.80 billion
2001	80	14.0%	8.4%	10.2%	$2.97 billion	$35.29 billion
2002	66	9.1%	5.1%	8.2%	$1.13 billion	$22.03 billion
2003	63	11.7%	10.4%	8.7%	$1.00 billion	$9.54 billion
2004	173	12.3%	12.4%	7.1%	$3.86 billion	$31.19 billion
2005	159	10.3%	9.3%	5.8%	$2.64 billion	$28.23 billion
2006	157	12.1%	13.0%	5.6%	$3.95 billion	$30.48 billion
2007	159	14.0%	13.9%	6.8%	$4.95 billion	$35.66 billion
2008	21	5.7%	24.7%	-1.7%	$5.63 billion	$22.76 billion
2009	41	9.8%	11.1%	5.7%	$1.46 billion	$13.17 billion
2010	91	9.4%	6.2%	3.1%	$1.84 billion	$29.82 billion
2011	81	13.9%	13.0%	8.5%	$3.51 billion	$26.97 billion
2012	93	17.7%	8.9%	11.1%	$2.75 billion	$31.11 billion
2013	158	20.9%	19.0%	13.0%	$7.89 billion	$41.56 billion
2014	206	15.5%	12.8%	5.8%	$5.40 billion	$42.20 billion
2015	118	19.2%	18.9%	10.3%	$4.16 billion	$22.00 billion
2016	75	14.5%	14.2%	5.0%	$1.77 billion	$12.52 billion
2017	106	12.9%	16.0%	9.0%	$3.68 billion	$22.98 billion
2018	134	18.6%	19.1%	11.6%	$6.39 billion	$33.47 billion
2019	113	23.5%	17.6%	17.9%	$6.95 billion	$39.28 billion
2020	165	41.6%	47.9%	26.2%	$29.66 billion	$61.86 billion
2021	311	32.1%	24.0%	17.0%	$28.65 billion	$119.36 billion
2022	38	48.9%	14.2%	9.3%	$0.99 billion	$6.98 billion
2023	54	11.9%	16.1%	-0.5%	$1.92 billion	$11.92 billion
1980-2023	**9,181**	**18.9%**	**20.5%**	**7.0%**	**$233.32 billion**	**$1,140 billion**

Closing Steps

Once students have now grasped the idea of what IPO underpricing is using Taylor Swift's the Eras Tour, the original goal of this pedagogical exercise, the discussion would move forward into the financial theories as to why IPO underpricing exists. While outside the scope of this project, we recommend a few points that may help in building and facilitating future class discussion(s). Depending on the level of desired depth around the topic of IPOs, the next progressive step is to discuss underpricing theory from the lenses of oversubscribed issuances (i.e. demand), risk (e.g. litigation risk avoidance), and/or the general difficulty in accurate valuation. Additionally, more formal finance theories may be used at the instructor's discretion. To note a few: the winner's curse (Rock, 1986), prospect theory and non-pecuniary benefits (Loughran and Ritter 2002), the changing risk composition, the realignment of incentives, and the changing issuer objective function hypotheses (Loughran and Ritter, 2004). Lastly, the instructor may choose to examine IPO performance in the long-run (Ritter 1991; Loughran and Ritter 2004).

Other Teaching Considerations

Broker, Dealers, Primary, and Secondary markets

This teaching exercise allows for exploring broader financial market concepts through extended analogies. For example, instructors may engage students in thinking about how the Eras Tour tickets, or any entertainment or sports tickets, fluctuate in price over time, much like stock prices. Students can explore the distinction between primary and secondary markets by comparing the initial ticket sale (via Ticketmaster or artist presales) to an IPO, and subsequent resales on platforms like StubHub to secondary market trading on the NYSE or NASDAQ.

Additionally, instructors may draw parallels between financial market structures and the ticket ecosystem. For example, Ticketmaster, loosely speaking, behaves like a dealer or underwriter by holding and distributing initial ticket inventory, while secondary platforms resemble brokers, facilitating transactions between buyers and sellers and charging a fee for the service. These comparisons, while beyond the core scope of IPO underpricing, can help students deepen their understanding of market intermediaries.

The core analogy in this paper assumes tickets are purchased and resold the same day, aligning with IPO underpricing. However, instructors should clarify that in practice, tickets can also be held for long periods or exchanged multiple times before the event, just like stocks. These deviations present further opportunities for discussion and teaching about holding periods, liquidity, and market structures. Each of these are simply additional considerations that an instructor can take or leave as they see fit.

Ethics and Market Regulation

While the parallels between concert ticket pricing and IPO underpricing offer helpful classroom understanding, there are also ethical and regulatory distinctions worth noting. Therefore, in addition to the IPO landscape listed in steps 1 through 5 of Table 1, professors can consider how ticket scalping may lead to thoughtful classroom discussions about business ethics or market regulations.

In financial markets, IPO underpricing is often framed as a strategic decision balancing issuer incentive, investor demand, and long-term relationships with underwriters. While it may result in "money left on the table," it is legal and subject to regulatory disclosure rules.

In contrast, the secondary market for concert tickets, particularly the practice of ticket scalping, raises more contentious ethical issues. Scalpers may use bots and bulk purchasing strategies to acquire tickets at face value, only to resell them at inflated prices, effectively extracting surplus from fans without adding value. Unlike IPO investors, scalpers are typically not capital providers or long-term stakeholders; they profit by exploiting inefficiencies in the original allocation process. This contrast raises broader discussions for the classroom on business ethics and market regulations.

Conclusion

Using pop-culture as a pedagogical tool can be valuable, and with due diligence by the course instructor, other events could be used if they find they would be more relatable to the class, such as the Super Bowl. This illustration was utilized in a senior level undergraduate course on Advanced Corporate Finance (i.e. Financial Management: Theory and Practice) over the two semesters from Fall 2023 and Spring 2024. Anecdotally, it has proved to be immensely useful for its designated purpose in assisting with the understanding and relatability of what IPO underpricing is and why it matters conceptually and pragmatically in the field of finance- Table 5 designs an assessment approach for future implementation. Additionally, it has had larger effects on the understanding and engagement of the entire IPO process. Since the Eras Tour was new, exciting, record-breaking, and widely discussed in major media outlets and across social circles, every student had at least heard about the 152+ show tour, selling an average of 70,000 tickets per show and becoming the highest grossing concert tour of all time (McCluskey, 2024). Many students were also familiar and aware of the broader economic impact which estimates suggest exceeded $10 billion (U.S. Travel Association, 2023).

The result of including this analogy in sessions on IPOs means that every student was able to hinge onto the topic and not only understand it, but also remember it for future discussions. Anecdotally, students have referenced the personal value in communicating their learning to others - having used this analogy in their personal conversations with friends and family to explain IPO underpricing. The Eras Tour may be an outlier in our lifetime, having broken many records for ticket sales and economic impact; its utility as a pedagogical tool is arguably universal. Overall, the takeaway is that contemporary events and pop culture analogies can be powerful pedagogical tools that enhance student engagement and learning in ways that ensure both knowledge retention and educational effectiveness.

Table 5
Design of Effectiveness Assessment

This table provides a design of a future assessment to analyze the effectiveness of using Taylor Swift's the *Eras Tour*, or other event to improve students' understanding of IPO underpricing.

Objective: To evaluate the effectiveness of the Taylor Swift IPO underpricing analogy in improving student understanding, retention, and engagement with IPO concepts.

Step	Topic	Description
1	Recommended Participants	Undergraduate finance students across multiple course sections (e.g., Advanced Corporate Finance or Investments) taught by different instructors, ideally across multiple institutions for generalizability.
2	Experimental Design	Control Group: Receives standard IPO underpricing instruction (textbook/theory-based). Treatment Group: Receives the same instruction but supplemented with the Taylor Swift the *Eras Tour* examples.
3	Data Collection	Pre-test: Administered before the IPO unit to assess baseline knowledge of IPOs and underpricing. Post-test: Administered after the IPO unit to measure gains in conceptual understanding. Retention Test: Administered 4–6 weeks later to measure long-term retention. Surveys: Gather student feedback on engagement, clarity, and perceived relevance via Likert-scale and open-ended questions. Focus Groups: Conducted with a subset of students to gather qualitative feedback on the teaching method's impact.
4	Assessment Metrics	Quantitative: Improvement in test scores (pre- to post-), differences between treatment and control groups, retention score variance. Qualitative: Student-reported engagement, relatability, and perceived clarity of IPO underpricing after using the analogy.
5	Analysis	Statistical paired t-tests to assess learning gains. Thematic analysis of open-ended responses to identify recurring qualitative patterns in student reflections.

References

Barkley, E. F., & Major, C. H. (2020). Student engagement techniques: A handbook for college faculty. Jossey-Bass, a Wiley Brand.

Coscarelli, J., Cox, C., & Bierman, F. (2024, May 17). How big is Taylor Swift? The New York Times. https://www.nytimes.com/interactive/2024/05/17/arts/music/taylor-swift-sales-tours-grammys.html

Eich, E. (1995). Searching for mood dependent memory. Psychological Science, 6(2), 67–75.

Fox, E. (2023) $11,000 to see Taylor Swift? How concert tickets got so expensive. https://www.latimes.com/entertainment-arts/story/2023-07-27/explaining-the- complicated-world-of-ticketing

Gendron, W. (2023) The average price of a resale ticket for the Eras Tour was $3,801. That's up 2,321% from Taylor Swift's last tour. https://www.businessinsider.com/resale-price- taylor-swift-eras-tour-reputation-ticket-pitchfork-report-2023-8

Loughran, T., & Ritter, J. R. (2002). Why don't issuers get upset about leaving money on the table in IPOs? *The Review of Financial Studies*, *15*(2), 413-444.

Loughran, T., & Ritter, J. (2004). Why has IPO underpricing changed over time? *Financial Management*, *33(3)*, 5-37.

McCluskey, M. (2024) As Taylor Swift's Eras Tour Hits One Year, Let's Take a Look at Its Staggering Numbers. https://time.com/6900208/taylor-swift-eras-tour-anniversary-stats/

Ritter, J. R., & Welch, I. (2002). A review of IPO activity, pricing, and allocations. *The Journal of Finance*, *57*(4), 1795-1828.

Ritter, J. R. (1991). The long-run performance of initial public offerings. *The Journal of Finance*, *46(1*), 3-27.

Ritter, J. R. (2024a). Initial public offerings: Underpricing. University of Florida. https://site.warrington.ufl.edu/ritter/files/IPOs-Underpricing.pdf

Ritter, J. R. (2024b). Money left on the table in IPOs by firm. University of Florida. https://site.warrington.ufl.edu/ritter/files/money-left-on-the-table.pdf

Rock, K. (1986). Why new issues are underpriced. *Journal of Financial Economics*, *15*(1-2), 187-212.

Steele, A., Dapena, K. (2024). Who is Making Money on Your Concert Tickets? It's Not Just Taylor Swift. *Wall Street Journal, https://www.instagram.com/p/C7ke5a8tHko/?igsh=MWQ1ZGUxMzBkMA*

Tyng, C. M., Amin, H. U., Saad, M. N., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in psychology*, *8*, 1-22.

U.S. Travel Association (2023) The Taylor Swift Impact – 5 Months and $5+ Billion. https://www.ustravel.org/news/taylor-swift-impact-5-months-and-5-billion

Vega, N. (2023) Taylor Swift is 2023′s most in-demand artist—Eras Tour tickets sell for an average of $1,088. https://www.cnbc.com/2023/12/06/2023-stubhub-top-artists-taylor-swift-beyonce-adele.html

Willis, J. *(2006).* Research-based strategies to IGNITE student learning: Insights from a neurologist and classroom teacher. *Association for Supervision and Curriculum Development,* p. 44.

Load Shedding in South Africa: What should Sozo do?

Karyl B. Leggio
Loyola University Maryland

As available power supply falls well short of demand, South Africans experience lengthy, daily power outages known as load shedding. The lack of power is disruptive to businesses and slows progress towards accomplishing organizational goals. Sozo Foundation, a non-profit organization seeking to educate and empower youth through training to achieve employment, frequently finds its operations disrupted when power fails. Sozo has several options: adjust to power outages and wait for the South African government to upgrade the power grid; obtain funding and purchase batteries or a generator; or purchase solar panels to produce solar energy sufficient to meet the needs of Sozo's operations and sell excess power into the network. What should Sozo do?

Introduction

Vrygrond, an impoverished community in Cape Town, South Africa, is less than one square mile in size. It has a high level of poverty, rampant drug and alcohol abuse, extensive gang activity, and prostitution, leading to many street children. Unemployment levels in Vrygrond exceed 77% (versus 35% country wide), and high school graduation is not common[4]. Because many jobs in South Africa require a high school diploma, the cycle of poverty and abuse continues.

Enter the Sozo Foundation. Founded in 2010, this non-profit has as its mission to improve the lives of those living in Vrygrond and empower the youth of Vrygrond to "live with dignity, purpose, and hope by creating pathways to employment."[5] Sozo helps students earn their GED's and provides students with training for a trade such as coding, web development, baking, barista skills, and hair and makeup training to enable students to work for organizations or open their own businesses. Since 2011, over 891 youth's lives have been impacted, with an 81% GED pass rate in 2021.[6]

Making matters worse, because of an insufficient power supply in South Africa, the country practices load shedding, the process of regularly turning off the power in a region for a period of time. This practice makes it difficult for Sozo to serve its clientele since it needs power for the computers and equipment necessary to teach skills to the students. To create a long-term solution, Sozo is considering developing its own power supply through the use of batteries, generators and/or solar panels; it is trying to decide which alternative power supplies make the most sense.

Sozo Foundation[7]

In 2010, native South African spouses Anton and Elana Cuyler returned to Cape Town, South Africa (after living in England for years) to volunteer as youth workers. They created a non-profit,

[4] https://www.reuters.com/world/africa/south-africas-unemployment-rate-rises-329-first-quarter-2024-05-14/
[5] https://thesozofoundation.org.za/ . This compares to a national average graduation rate of 55% . https://businesstech.co.za/news/government/743573/south-africas-real-matric-pass-rate-is-only-55/
[6] https://thesozofoundation.org.za/educentre/
[7] See the Sozo Foundation homepage for much of this content https://thesozofoundation.org.za/

Sozo, and formed a youth group in Vrygrond township, an overcrowded and under-resourced four block square community home to 45,000 residents[8]. Given the community has no police force, high school or medical facilities, Vrygrond residents suffer from criminal activity, gangs, drug abuse and prostitution, and an under-educated, poverty-stricken, unemployed population.

Sozo is a Greek word meaning to keep safe, to rescue from danger or destruction. The Sozo Foundation has four primary areas of focus: education, skills development, social enterprise, and entrepreneurship.[9] The education activities represent 40% of the Foundation's work and resources; the skills development work is 35% of Sozo's efforts; social enterprise represents 15% of the work and resources, with entrepreneurship comprising the final 10%.

The education focus is helping Vrygrond youth earn a GED to help to secure a job to end the cycle of poverty. In the area of skills development, community members can learn new trades or improve upon existing knowledge. The goal is to provide tools to generate an income and to help community members rise out of poverty. Social enterprise places youth in apprenticeships and on-the-job training in industry[10]. And entrepreneurship helps young people start and run their own businesses.

Figure one is a graphic representation of the Sozo model. As Anton Cuyler, CEO and Founder notes,

> *"The youth development model has evolved in the more than twelve years of community-based youth work and continues to innovate and evolve. In contrast, our impact goal has remained unchanged and simple; develop youth employability, leading to youth employment. Just as any healthy natural ecosystem has an inflow and outflow that enables life to exist and thrive, we also apply an inflow-outflow philosophy. We don't "keep kids busy and off the streets" nor subscribe to a "settling mentality". We offer a variety of inflow opportunities to youth who find themselves at various stages of life and in different states of mind and heart. Once a young person becomes part of the ecosystem, there are many opportunities for growth, development and movement. These interactions are where the ecosystem begins to take care of itself. We realise youth developmental states are uniquely different, and youth will exit the ecosystem at different stages. It remains core to our vision that every youth enters a pathway that leads to the biosphere of further education, training and, ultimately, youth employment. Our most essential nutrient within the ecosystem is our cultivated value of UNCONDITIONAL LOVE."*

A. Education

South Africa's math and science scores are rated the worst of the 148 countries surveyed.[11] A poor educational system is not the only problem: poverty, lack of resources and unsafe spaces

[8] Note: "youth" is defined as individuals between the ages of 18-30.

[9] https://www.biblestudytools.com/lexicons/greek/nas/sozo.html

[10] The differentiation between skills development and social enterprise is the following: during skills development, students are being trained in areas such as baking, construction, and hairdressing, among other skills. For social enterprise, the students actually go out and work in small businesses owned by Sozo that support the community. There is a roastery business that will bring a coffee truck to your event; there is a construction company, Kingdom Builders, that will perform carpentry and construction work. The students are able to gain real world experience with the skills that they learned.

[11] World Economic Forum, 2014. Global Competitiveness Report for 2014–2015

have led to over 60% of South Africans not completing high school, leading to a cycle of high unemployment, disillusionment about their future, gang culture and drug and alcohol abuse.[12]

The Vrygrond area is mired in poverty and crime, with educational attainment well below South Africa's abysmal levels. Sozo provides students with the resources to complete a self-paced, fully online GED program. Completion of the program provides students with access to employment where a high school diploma is a minimum requirement. It also provides students with the opportunity to enroll in college or technical school. Sozo also provides a GED bridging school program that prepares students for the GED program by closing the numeracy and literacy gaps over the course of a year to allow these students to be up to the GED standard when enrolling in the GED program. This helps greatly with the success rate for GED students.

In addition to the GED, Sozo offers life skills courses, after-school tutoring and mentoring, and personal development courses. All students of Sozo's programs have direct access to psycho-social-intervention support which is a crucial component of Sozo's model, given the severely disadvantaged backgrounds of the students. This dedicated team of professional social workers also develops internal resilience and mental health self-regulation skills to support the students to deal with, and process, trauma. Sozo believes students cannot learn when they are hungry, so nutritional meals are provided. And students cannot aspire to that which they do not know, so students are exposed to career opportunities through career-inspiring initiatives.

B. Skills Development

Youth unemployment is among the highest in the world, with a rate of 68% as of 2019, growing to 77% by 2024.[13] To prepare the youth of Vrygrond for employment, Sozo Foundation works on the following:

- Life Skills

  Life Skills, also referred to as 'employability skills', are those that are completely transferable between industries and occupations. Lack of life skills are often the reason employees lose jobs. Sozo teaches the students skills of goal setting, financial management, sexual and emotional health, timekeeping and communication. The sessions also address themes of resilience, teamwork, and leadership.

- Vocational Skills Schools

  The six months of vocational skills courses are vital to entry level access to the job market for students. The majority of students have been out of the educational system for many years. Sozo provides them with an employable skill through in-depth training, as well as peer to peer learning. Skills range from barista and baking skills to hairdressing and beautician skills, construction and coding (which allows students to earn certification through Amazon Web Services.)

- Job Shadowing

  Sozo provides real marketplace experience for students. This is often the first time students have been in a real job environment. Sozo built a network of partnerships for job shadowing

[12] Department of Basic Education, Republic of South Africa, 2015
[13] Expanded youth unemployment rate. Statistics South Africa Labour Force Survey 2019

opportunities with local businesses and companies. Often these businesses offer students jobs upon graduation.

- Job Readiness

  Students are assisted with developing an up-to-date resume with which to apply for jobs. Students engage in job searching, interview preparation and practice. Sozo works with a variety of community partners and businesses to help implement this phase.

C. Social Enterprise

Of the 6.7 million unemployed youth in South Africa as of 2019, 75% have no work experience; in Vrygrond, over 90% of the youth lack work experience.[14] Sozo has three social enterprises to provide students with work experience.

- The Wild Goose Artisan Bakery and Artisan Baking School is a social enterprise seeking to train and support unemployed young people. Proceeds from the bakery are used to support the training of future bakers.
- The Dancing Goat Roastery provides mobile coffee events with trained baristas and a mobile coffee cart to serve espresso-based and other hot drinks at various venues for events.
- The Kingdom Builders is a social enterprise maintenance and construction crew that seeks to train and employ young people through its apprenticeship program.

D. Genesis Incubation Hub

In 2023 Sozo officially launched a new incubation hub for 16 youth-owned micro-businesses to create local employment through one-on-one mentorship and coaching. Sozo provides hub space for internet access, business training and workshops to increase the likelihood of success and growth of these businesses.

Rolling Power Blackouts

Eskom, the monopoly power supplier in South Africa, supplies over 95% of all power and was once considered state of the art. Now the power system, almost entirely comprised of aging coal-fired generators, is subject to increased demand for power by South Africans, aging, less efficient power plants, poor management, corruption, and sabotage. Rolling power blackouts, known as load shedding in South Africa, are a fact of life for South Africans.[15] There is not enough available power to meet demand, and there is no plan in place for Eskom to rectify the situation. As a result, Eskom has begun turning off the power for customers for parts of each day. This process is known as load shedding.

Load shedding distributes demand for electrical power across multiple power sources and is used to relieve stress on the power supply when demand for electricity is greater than supply. To prevent the power grid from overloading, load shedding rotates power outages by reducing consumption until capacity is available. In South Africa's case, the need for load shedding arises due to insufficient generation capacity at power plants.

[14] Harambee 2019 "Youth is defined as 18-30.
[15] Cohen, 2023, https://www.bloomberg.com/news/articles/2023-12-13/why-south-africa-is-mired-in-an-electricity-crisis-quicktake

During load shedding, customers without alternative sources of power (batteries, generators, solar) have no power.[16]

Customers can experience rolling power blackouts for 10-12 hours per day during peak power demand periods. Of course, these outages are not felt equally. With over 60% of the country living below the poverty line, those who cannot afford generators fall further behind without access to the internet and technology when the power is out.[17]

Figure two shows South Africa's dependence upon coal.[18] Some predict it may be 10-15 years before Eskom can build adequate new generation facilities to meet the needs of South Africans.

Because of the crisis state in South Africa, in 2021 President Cyrl Ramaphosa announced the Electricity Regulation Act that will allow independent power producers to produce up to 100 MW of generated power.[19] With the assistance of funds from Western nations, the monopoly of Eskom may be ending as South Africa takes advantage of its wind and solar resources to generate power. Load shedding, in February 2023, has reached stage six, meaning Eskom needs to remove demand for up to 6,000 megawatts from the national grid, leaving South Africans without alternative power sources in the dark for up to ten hours per day.[20]

The Electricity Regulation Act is in place to allow those who can supply power via renewable resources to do so. These alternative power sources reduce consumers' dependence upon Eskom's system to provide power. It also leads to the use of renewable power sources, lessening demand for fossil fuels and improving the environmental impact given some power generation will be coming from solar and wind instead of coal.

The transition to renewable energy sources is not without its costs: many worry about the loss of jobs in coal mines from the move away from coal-fired power plants. However, others warn the entire power grid may collapse leaving the country in darkness. Sozo must determine whether it has the resources to fund alternative sources of power for its operations or if it needs to seek support from additional donors to fund alternative power sources.

Alternative Power Sources

Cape Town, South Africa has extreme weather: it is very hot in the summer months and extremely cold in winter months, with temperatures similar to cities in the Southeast United States such as Raleigh, North Carolina. Global warming is leading to even more extreme temperature swings.

The current way to maintain a comfortable environment in Sozo buildings is not terribly efficient: there are individual office air conditioners for summer and individual heaters in the winter. The monthly power bills for Sozo amount to 50,000 RND (or approximately $2,500 per month). Eskom has been increasing the cost of power at least twice each year, and the size of the increases is not capped by regulators: in July, 2023 alone, power costs increased by 18%[21].

[16] https://www.techtarget.com/searchdatacenter/definition/load-shedding

[17] https://kleinmanenergy.upenn.edu/news-insights/shedding-the-load-power-shortages-widen-divides-in-south-africa/

[18] https://www.bbc.com/news/world-africa-65671718

[19] https://mg.co.za/news/2021-06-10-ramaphosa-unveils-policy-amendments-to-help-struggling-eskom/

[20] https://www.reuters.com/world/africa/south-africas-eskom-ramps-up-power-cuts-stage-6-2023-02-20/

[21] These increases are completely unpredictable. Some months saw 4-5% increases in power prices; other quarters saw 20-30% price increases.

The Sozo Foundation works out of four different buildings located throughout Vrygrond. To power the buildings, the foundation has numerous choices (see Exhibit One for a summary of the pros and cons of each power source):

- it can continue to use Eskom power exclusively, meaning that Sozo's operations will be subject to power outages during load shedding;
- It can install lithium batteries; the batteries provide backup power during load shedding and need to be recharged after each use. Recharging can occur through the use of solar power;
- It can install a diesel generator. The generator can be used during load shedding to maintain consistent power;
- It can install solar panels. The panels can generate power to be used to pull Sozo off of the Eskom grid to reduce the monthly power bill for Sozo; and
- It can purchase inverters to take the excess power created from the solar panels and sell that power into the network, thus creating a revenue source for Sozo.
- Sozo can also combine several of these power sources and create a new power source for its operations, and may be able to go off of the Eskom power grid altogether. For example, in order to no longer be dependent upon power from Eskom, Sozo would need generators, solar panels and inverters. This would enable Sozo to generate revenue in addition to saving costs.

The power needs in each building are as follows (See Exhibit Two):

Education center: The Education Center has solar panels and inverters that were installed last year. These devices are not, however, sufficient to generate adequate power to run the center; they supply supplemental power during load shedding. However, as load shedding has extended to 8-12 hours per day, solar and inverters are not sufficient to supply power to the Education building given the power needs of the computers and other equipment. For full power coverage during load shedding, the Education Center needs either batteries or a diesel generator.

If Sozo chooses to use batteries instead of generators, Sozo could purchase 25 kw batteries (lithium) at a cost of 37,000 RND each; the Education Center would need battery power from five batteries per hour, and with double load shedding, the Foundation would need to quadruple the number of batteries at a total cost 750,000 RND per load shed. Batteries do not generate power to sell into the power grid since it takes a good deal of time to recharge the batteries at night. As a result, even though there are solar panels on this building, the power needs exceed the supply created by the solar panels; there will be no power generated to sell into the network if the battery option is selected.

However, the alternative power source, a diesel generator, would be able to generate adequate power to allow power to be generated by the solar panels, stored in the inverters, and sold into the power network/grid. The generators require the purchase of a 25 kva generator and peripherals to maintain consistent power during load shedding. This building currently has solar panels that were installed last year, and has the correct inverters to store power and sell it into the grid. The generator cost is 170,000 RND[22].

Skills center: The Skills Center load is 25 kw and would therefore need five x five kw batteries per hour of load shedding at a cost of 30,000 RND per battery. Sozo needs to prepare for 2.5 hours

[22] Presume one South African Rand is equal to $.05 US.

of load shedding in this building; therefore it would require 13 batteries at 5 kw each for a total cost of 390,000 RND.

If purchasing a generator, the skills center needs a 25 kva generator at a cost of 170,000 RND. The skills center building does not currently have solar panels. The Center would need a full solar system to save on power costs to lower the electricity bills and to generate power to resell into the grid. Sozo needs an additional 400,000 RND to purchase solar panels to generate power and sell into the network.

Entrepreneurship center: The load for the entrepreneurship center is 15 kw. This building would need 3 x 5 kw batteries per hour at a cost of 30,000 RND per hour. To prepare for 2.5 hours with no power during load shedding, the entrepreneurship center would need to spend 225,000 RND for battery power.

The entrepreneurship center currently has no alternative power sources. The cost for the generator for this building is 102,000 RND. If Sozo wants to also be able to sell power into the power grid, the entrepreneurship center additionally needs a 5 kw inverter, 15 kw backup batteries, and solar system with generation capacity. The solar system for this smaller building costs 300,000 RND. This Center uses less power and can therefore more easily manage with batteries as backups instead of generators.

Social Enterprise: The social enterprise center is similar in size to the entrepreneurship center and requires 15 kw of power. For battery power, this requires an investment of 225,000 RND. For generators, this building needs a 25 kv generator (125,000 RND) because of the power tools that are running and using a lot of power; the building also needs a solar system for 250,000 RND if Sozo wants to sell excess power into the grid.

Since the school shuts down for holidays, summer vacation, etc. this accounts for approximately nine weeks of shut down time per year; all power generated in buildings with solar systems during these shutdowns could sell that power into the grid. Also the Center closes at 5:00 pm each evening, and the sun can shine until 9:00 pm in Cape Town, Alton believes there will be excess power that can be generated and sold each evening.

The assumptions in developing the costs associated with these alternative power solutions are as follows:

- Load shedding will exist for ten more years as Eskom works to overcome its challenges to provide adequate power for the citizens of South Africa
- The useful economic life of batteries is ten years; generators also have a ten year life while solar panels have a 20 year economic life
- Although power costs increased by 18% in July, 2023, to be conservative, we presume power will increase in cost by 5% annually
- Revenue will increase by 3% per year
- Operating costs will increase by 3% annually
- If Sozo is able to generate its own power through the use of solar panels, it will save 33% of operating costs annually
- The market risk premium in South Africa in 2023 is 8.7%[23]
- The cost of debt for short-term government bonds in South Africa is 8.57%[24]
- If Sozo is able to purchase solar panels, it is estimated it will have 1,704 hours of power to sell into the network in 2023; this volume of sales is anticipated to increase by 3% annually

[23] https://www.statista.com/statistics/664880/average-market-risk-premium-south-africa/
[24] https://www.ceicdata.com/en/indicator/south-africa/short-term-government-bond-yield

The Cost of No Backup Power Supplies

Alton quantified the cost of being without power during load shedding. As an example, he discussed the coding school.

> *"There are licenses for 12 students, and to participate in the coding certificate program, it costs Sozo 360,000 RND annually. These aspiring web developers don't have an alternative activity when the power is out. When Sozo is load shedding and cannot go online, students don't come into the Center on those days since they cannot access the lessons. The lessons go on even if the students are not able to access the content. If the students cannot pass the exam, they cannot get certified. Currently, I estimate it takes students 33% more time to complete the certification since content cannot be delivered on time."*

Overall, the costs for all programs are about one-third higher due to delays and waiting time caused by load shedding. Fixed costs, salaries, etc. all stay the same regardless of load shedding, but output is not increasing. Across all programs, approximately one-third of productivity time is lost due to power outages. Anton estimates the cost is 4 million RND per year.

More concerning to the colleagues at Sozo is the worry about students becoming disenchanted due to the delays in their education due to load shedding. For a student lost to the system, the cost of incarceration in South Africa is 300,000 RND per person per year. While this is not a cost to Sozo, it certainly is a cost to the community and to South Africa.

Alton believes the investment in alternative power will pay back within the first 18 months; do you agree?[25]

[25] Presume excess power can be sold back at half of its retail price; presume Sozo will be able to generate enough excess power to sell the equivalent of 10% of its current monthly power bill back to Eskom.

Figure 1
The Sozo mission model

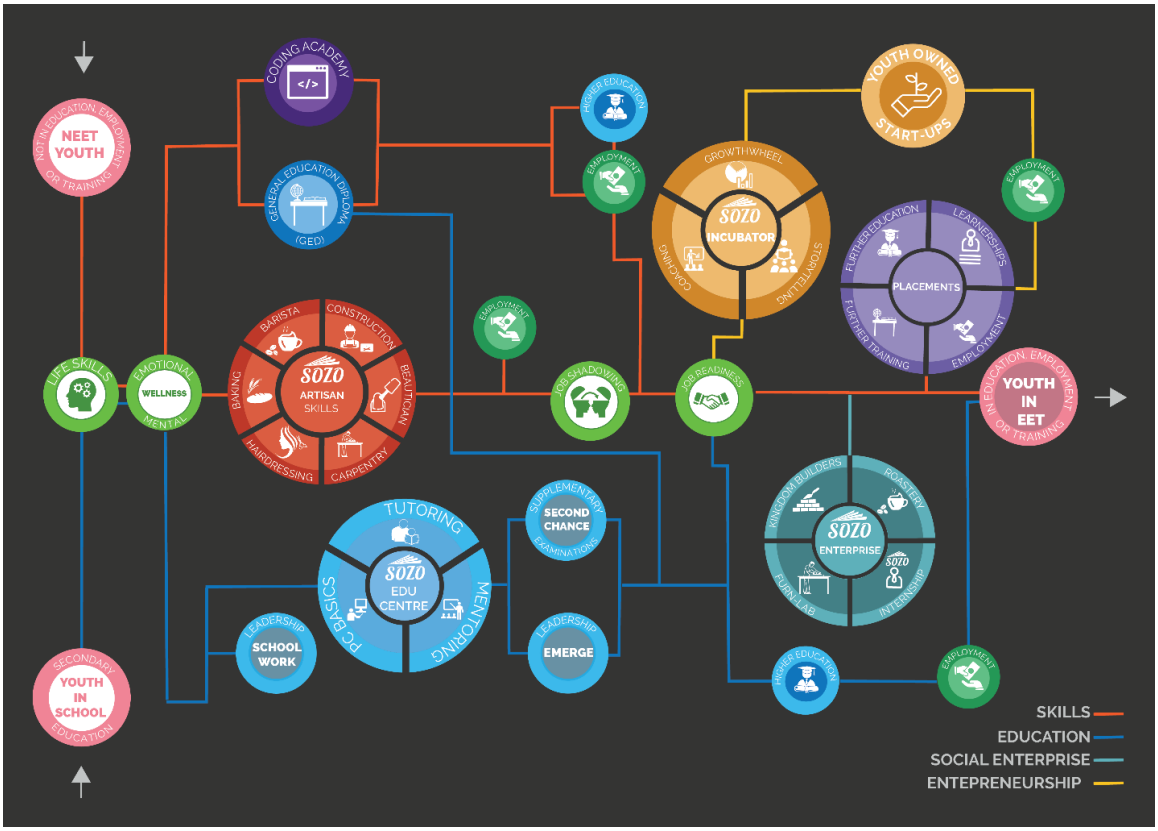


Figure 2
Sources of Power in South Africa

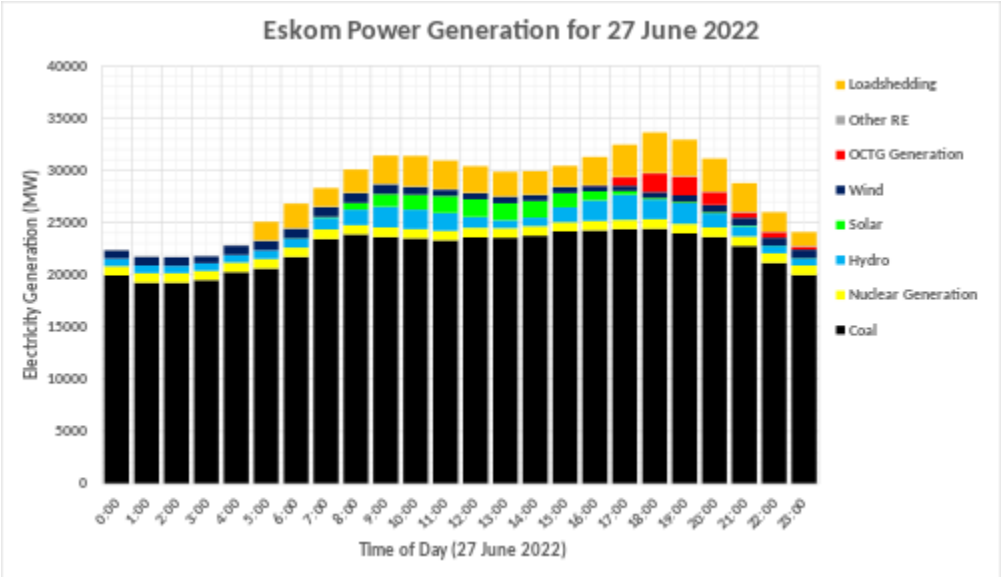
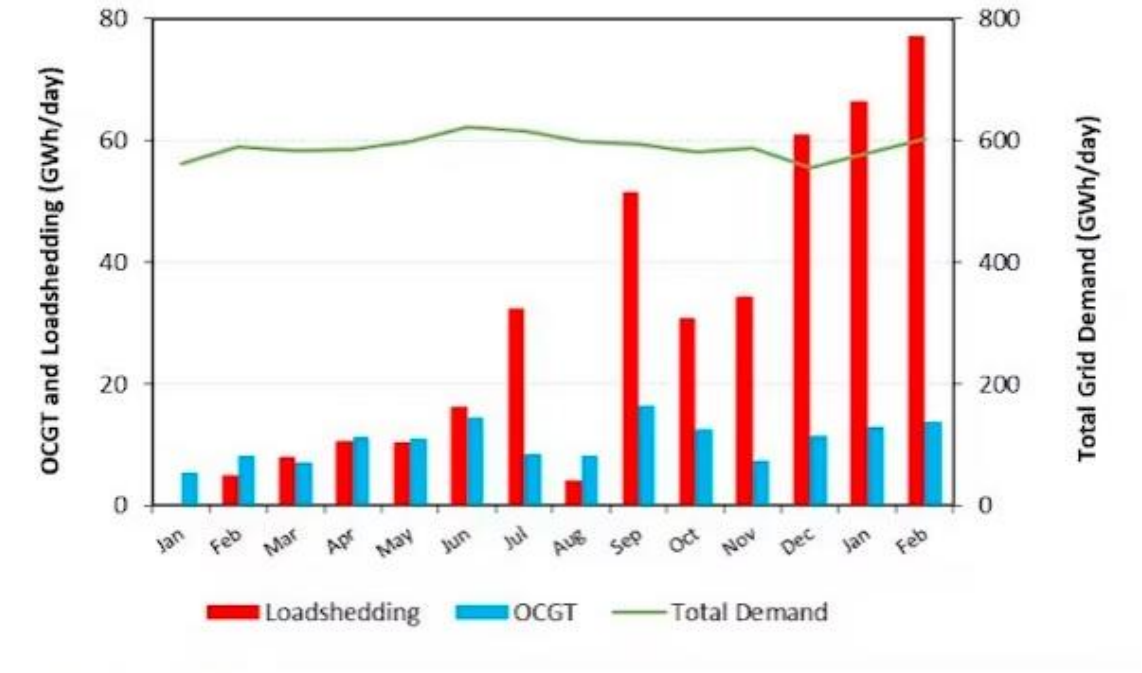


Figure 3
Load Shedding in South Africa[26]



David Richard Walwyn

Exhibit 1
Alternative Power Sources

Do Nothing

Pros:

- No initial outlay

Cons:

- Subject to increases in Eksom's power prices
- Power outages
- Students fall behind in lessons
- Sozo Center closes when no power and students become disenchanted
- Students exposed to street activity (crime, prostitution) when the Center is closed

Lithium Batteries

Pros:

- Backup power during load shedding

Cons:

- Batteries have a limited number of hours of run time
- Batteries need recharging
- Batteries still require purchasing power from Eskom at the going rate

[26] https://www.esi-africa.com/industry-sectors/generation/why-south-africas-new-electricity-minister-should-listen-to-what-the-former-utility-ceo-said/ OCGT is generated power.

Diesel Generators

Pros:

- Consistent power during load shedding (no need to recharge)

Cons:

- Noisy and smelly diesel fuel
- Generators are only for load shedding times; Sozo still needs to purchase power from Eskom
- Generators are subject to theft

Solar Panels:

Pros:

- Efficient, renewable energy source
- Many hours of sunshine daily in Cape Town, allowing the generation of power
- Less dependent upon purchasing power from Eskom

Cons:

- More expensive initial outlay

Inverters

Pros:

- Inverters give Sozo an additional revenue source since excess power can be sold into the network
- Less dependent upon power from Eskom

Cons:

- Cost

Combination:

Pros:

- Allows Sozo to go off of the Eskom network altogether thus eliminating Sozo's exposure to Eskom price increases
- Additional revenue source

Cons:

- Cost

Exhibit 2
Sozo Buildings

Education Center

- Already has solar panels and inverters
    - If generator is purchased, adequate solar power is generated to sell into the grid
    - If batteries are purchased, no solar power to sell into the grid
- Batteries will supply power during load shedding; require charging at night
- Generators provide power during load shedding and support solar power production

Skills Center

- Batteries will supply power during load shedding
- Generators will provide power during load shedding and support solar
- The center does not currently have solar panels and will require the purchase of panels if this is the decision

Entrepreneurship Center

- Batteries can support the center during load shedding
- Generators are an alternative to batteries
- There are currently no solar panels nor inverters so they would need to be purchased if this option is selected
- Given the small size of this center, the solar panels will support all of its power needs and have excess power to sell into the network

Social Enterprise Center

- Batteries are an option for load shedding power supply
- Generators will provide power in lieu of batteries
- Solar panels and inverters will need to be purchased

Figure 4
Sozo Foundation Income Statement, 2022

The Sozo Foundation Trust
Formerly The Nceda Trust
(Registration number: IT 1524/2002)
Annual Financial Statements for the year ended 28 February 2022

Statement of Comprehensive Income

Figures in Rand	Note(s)	2022	2021
Revenue	9	11,693,896	8,963,117
Other income	10	643,695	451,220
Operating expenses		(11,226,615)	(8,329,512)
Operating surplus		**1,110,976**	**1,084,825**
Investment revenue	11	130,391	108,024
Capital expenditure	12	(916,198)	(374,295)
Surplus for the year		**325,169**	**818,554**