JOURNAL OF
ECONOMICS and FINANCE EDUCATION

JEFE

AEF
ACADEMY of ECONOMICS and FINANCE

Volume 17 SUMMER 2018 Number 1

Academy of Economics and Finance

A Vanishing Gender Gap in Attitudes Toward Economics: A Pre- and Post-Course Analysis

James Ullmer and Yanju Li[1]

ABSTRACT

This study explores the gender gap in economics. A pre- and post-survey was conducted in three introductory classes in microeconomics to measure any observable difference in attitudes between male and female students toward the discipline. The pre-survey discovered that male students had a statistically significantly more positive attitude than female students in two of the four measures of student attitude, including the overall metric. The post-survey revealed no such difference in attitudes between the genders in any of the measures. These results were consistent in online class, traditional face-to-face class, and face-to-face honors class formats.

Introduction

During the past few decades, there have been several studies examining the underrepresentation of women—as well as racial and ethnic minorities—in the economics profession (Leeds 1992; Kahn 1995; Ginther and Kahn 2003; Ginther and Kahn 2014). Recent evidence in terms of the number of Ph.D. and Bachelor degrees in economics awarded to women indicates that this gender gap will continue into the foreseeable future. For example, in 2014 only 157 of the 500 doctorate degrees awarded to US citizens and permanent residents were awarded to women. Similarly, on the undergraduate level, Integrated Postsecondary Education System (IPEDS) data reveal that only 28 percent of the bachelor degrees in economics were awarded to women, far less than the approximately 43 percent of degrees earned by women in mathematics and statistics (Bayer and Rouse 2016).

This lack of diversity in the profession is mostly attributable to the more negative attitudes held by female students toward economics than their male counterparts. Several explanations have been put forth to explain this gender difference in attitude. One reason cited is women's general lack of interest in the subject matter. This general lack of interest by female undergraduates has been offered by some researchers as the principal cause for the low enrollment of women in economics classes (Feiner and Roberts 1995; Ferber 1995; Bansak and Starr 2010). This is problematic, since that as Calkins and Welki (2006) have found, interest in the subject matter of a discipline is the most significant factor for undergraduate students in their choice of a major. Moreover, Dynan and Rouse (1997) cite this as the chief reason that female students are more than twice as likely as male students to avoid enrolling in a principles of economics course during their freshman year—a time period when many undergraduates choose their major. Bansak and Starr (2010) conjecture that an important part of the gender differential in predispositions comes from the fact that while both male and female students have expectations from studying economics that are way out of line with reality (namely, that it has a lot to do with making money), it is more of a put-off for women than for men. This misperception may contribute to women's negative attitudes concerning the discipline. Bansak and Starr recommend correcting this misimpression at the outset of a principles of economics course to reduce this confusion and get rid of its adverse effects on learning.

Ballard and Johnson (2005) find that women have less confidence in their ability to succeed in economics courses than their male counterparts, and therefore are reluctant to enroll in principles of economics classes. This lack of confidence may also create a self-fulfilling prophecy, and thereby negatively affect their

[1] James Ullmer, Professor of Economics, College of Business, Western Carolina University, Cullowhee, NC, 28723, ullmer@wcu.edu, 828-227-3762; Yanju Li, Research Associate, Student Financial Management Center, Georgia State University, Atlanta, Georgia, yli@gsu.edu, 949-701-6602; The authors would like to acknowledge Stephen Le Beau, Western Carolina University, for his valuable assistance in uploading the Qualtrics Survey into Blackboard

performances in these courses, resulting in lower grades. These lower grades, especially when compared to grades received in their other classes, may discourage female students from majoring in economics. In this regard, Rask and Tiefenthaler (2008), using data from several liberal arts colleges, found that female students are more negatively affected by lower grades in economics—especially when compared to their grades in other coursework—than their male counterparts. Consequently, female students are less willing than male students to continue studies in economics if they receive a lower grade in a principles course of economics than they obtain in their other coursework.

Research has been conducted on the influence of same-gender instructors on women's attitudes toward economics and any subsequent decision to major in the discipline. For instance, Dynan and Rouse (1997) have found that the gender of the instructor has little influence on female students' decisions to major in economics. However, the pedagogical approach employed by the instructor does seem to affect women's success, or lack thereof, in economics classes, and thereby affect their dispositions concerning economics. For example, some research cited the teaching styles of some economics instructors as negatively affecting women's attitudes toward the discipline (Hall and Sadler 1982; Jensen and Owen 2000, 2001).

The negative attitudes of many female students cited above are often acquired and/or reinforced in introductory courses in economics. With that in mind, the primary purposes of this paper are: 1) to explore gender differences in student attitudes toward economics in terms of confidence in learning the subject matter, relevance of the discipline, and general interest in economics; 2) to examine whether these predispositions toward economics, especially as they pertain to the gender gap, might change during a semester in a principles of microeconomics class; 3) to discover whether the class platform—a traditional face-to-face class, an online class, and a face-to-face class with strictly Honors students—might have some influence on the gender gap.

Methodology

In the Fall 2016 semester, three classes of principles of microeconomics were taught at Western Carolina University by the same male professor of economics. Each section covered the same material that is taught in a typical introductory course in microeconomics and used the same text—*Microeconomics*, 20thed, McConnell, Brue, and Flynn. However, each section was different in either platform or audience. There was an online class of thirty-four students comprised primarily of distance learners, many of whom were non-traditional students. There was a face-to-face class consisting of forty-seven traditional undergraduate students. Finally, there was a face-to-face honors course of twenty-eight students, who were all members of the Honors College at Western Carolina University.[2]

A questionnaire was developed using a seven-point Likert-scale to ascertain student attitudes in regard to economics. The questionnaire contained 18 statements—each representing a positive predisposition toward some aspect of economics—to which students were to respond from 1, strongly disagree, to 7, strongly agree. On the basis of these statements, three variables were derived to measure students' attitudes in three distinct areas; *Ability to comprehend economics, Usefulness of economics, and Interest in economics*. A fourth variable, *Overall attitude toward economics*, was also calculated using all of the questions on the survey. The questionnaire also included one open-ended question in which students expressed their general impressions of economics.[3]

A reliability test, Cronbach's alpha, was employed using SPSS 24 to test the internal consistency of the four dependent variables being measured by the questions on the survey. On the pre-test data, the three variables: *Ability to comprehend economics*; *Usefulness of economics*; and *Interest in economics* had a Cronbach's α of .845, .859, and .884, respectively, indicating a relatively high reliability. In the global measure, *Overall attitude toward economics*, Cronbach's α equaled .924, indicative of an extremely strong reliability. For the post-test data, fairly high reliabilities were revealed as well. Cronbach's α equaled .925, .880, and .930 for *Ability to comprehend economics, Usefulness of economics*, and *Interest in economics*,

[2] The Honors College at Western Carolina University is very selective and rated in the top ten nationally in undergraduate research. Presently, there are approximately 1300 students enrolled in the Honors College.

[3] Refer to the appendix for the Questionnaire: The variable Ability to Comprehend Economics is measured by items 1, 2, 3, 4, 10, 14, and 20. The variable Usefulness of Economics is measured by items 7, 8, 11, 12, 15, 17, and 18. The variable General Interest in Economics is measured by items 5, 6, 8, 13, 16, and 18.

respectively. The comprehensive metric, *Overall attitude toward economics*, had a Cronbach's α of .959, demonstrating exceptionally strong reliability.[4]

The questionnaire was administered online using Qualtrics Survey Software. To measure student attitudes toward economics at the beginning of the semester, the questionnaire was made available to students in all three classes during the first two weeks of the fall semester before any student assessment occurred in terms of quizzes or exams. Then, in order to assess student predispositions toward economics at the end of the semester, that same survey was administered during the final two weeks of the fall semester before finals.

Subsequently, inferential statistics in the form of t-tests if the underlying data was normally distributed, or the corresponding non-parametric Mann-Whitney U test if the distribution was non-normal, were performed on both pre-course and post-course empirical data. Next, the statistical results of the pre and post surveys were compared to examine whether differences in male and female attitudes in economics—the gender gap—had changed. Finally, two tests were performed to examine whether student attitudes, especially as they pertain to the gender gap, were influenced by class platform. First, a one-way analysis of variance (ANOVA) for normally distributed data, or the parallel Kruskal-Wallis H test in the case of a non-normal distribution, was performed on both pre- and post-survey results and compared to ascertain whether student attitudes were affected by class format. A second statistical test was performed using multiple regression analysis to further analyze whether the gender gap was affected by the class platform.

Course Gender Differences in Attitude at Outset of Course

A total of ninety-three respondents—forty-five male and forty-eight female students—out of the population of one hundred nine students enrolled in the three classes completed the questionnaire at the beginning of the semester. The mean values of all four variables—*Ability to comprehend economics*, *Usefulness of economics*, *Interest in economics*, and *Overall attitude toward economics*—demonstrate that male students had a generally more positive attitude toward economics than their female counterparts. The mean is the average response to the questionnaire statements as measured on the Likert-scale. The mean values along with their standard deviations are displayed in the first two columns of Table 1.

Table 1: Comparison of Student Attitudes at the Outset of Class and End of Semester

	Outset of Class			End of Semester		
Dependent Variable	Male (N=45)	Female (N=48)	P-Val of Diff	Male (N=46)	Female (N=51)	P-Val of Diff
Ability to comprehend economics	4.78 (.12)	4.35 (.13)	.016*	5.01 (.09)	4.94 (1.21)	.92
Usefulness of economics	5.64 (.089)	5.26 (.13)	.077	5.45 (.74)	5.37 (1.05)	.87
Interest in economics	5.04 (.15)	4.79 (.15)	.248	4.83 (.96)	4.82 (1.37)	.52
Overall attitude toward economics	5.16 (.097)	4.80 (.12)	.020*	5.11 (.76)	5.06 (1.12)	.81

* Statistical significance at .05

Inferential statistical tests were then conducted on the pre-test means to see if these differences in predisposition were statistically significant at a significance level of .05. First, the distributions of the dependent variables were each tested for normality to ascertain whether parametric or non-parametric tests should be performed to test for statistical significance. The Shapiro-Wilk test showed that the variables *Ability to comprehend economics* (p = .374) and *Overall attitude toward economics* (p = .240) were normally distributed. That same analysis found that the variables *Usefulness of economics* (p < .001) and *Interest in economics* (p = .023) were not normally distributed.

[4] A Cronbach's alpha from 0.65 – 0.8 indicates the measure is reliable. A Cronbach's alpha greater than 0.8 is indicative of high reliability.

Therefore, t-tests were conducted on the variables, *Ability to comprehend economics* and *Overall attitude toward economics*. The results revealed that males had a statistically significant more positive predisposition than females on the measure, *Ability to comprehend economics* (p = .016), as well as the metric, *Overall attitude toward economics* (p = .020). A non-parametric test, the Mann-Whitney U test, was then employed on the *Usefulness of economics* and *Interest in economics* dependent variables, because of the violation of normality. The results indicated that there was no statistically significant difference in attitudes between genders on the variables *Usefulness of economics* (p = .077) and *Interest in economics* (p = .248), though males displayed a slightly more favorable attitude on the two measures. The p-values for the difference of means between genders at the outset of the class are displayed in column 3 of Table 1.

Gender Differences in Attitude at the End of Semester

Ninety-seven students out of the population of one hundred nine students completed the questionnaire at the end of the semester. As in the pre course survey, the number of respondents was approximately evenly divided between males and females—forty-six male students and fifty-one female students. A preliminary examination of the descriptive statistics indicates that while males still appear to have a slightly more favorable predisposition toward economics, the gender gap had become smaller by the conclusion of the semester. The mean values of the measures along with their standard deviations are displayed in columns 4 and 5 of Table 1.

Inferential statistics were next applied to the end of the semester survey results to ascertain whether there was a statistically significant difference between males and females on each of the four measures; *Ability to Comprehend economics, Usefulness of economics, Interest in economics*, and *Overall attitude toward economics.* The Shapiro-Wilk test was again used to test for the normality of the underlying distributions of the dependent variables. None of the four dependent variables proved to be normally distributed (p < .001, p < .001, p = .016, and p = .002, respectively).

For that reason, Mann-Whitney U tests were conducted on all four measures of student predispositions toward economics. The striking finding here was that there was no statistically significant difference in attitudes between male and female students in any of the four measures of student attitudes regarding economics. The p-values for *Ability to comprehend economics*, *Usefulness of economics*, *Interest in economics*, and *Overall attitude toward economics* were p = .92, p = .87, p = .52, and p = .81, respectively. It appears that the class experience of female students during the semester had closed the gender gap between male and female attitudes toward economics. Specifically, by the end of the semester female students were just as confident as their male counterparts in their ability to comprehend economics, which was not the case at the beginning of the term. Moreover, the statistically significant gender gap on the overall measure of student attitude evident at the beginning of the semester had also disappeared. The p-values for the difference of means between the genders are displayed in column 6 of Table 1.

Effect of Class Platform on Student Attitudes

Statistical techniques were subsequently employed to examine whether the initial gender gap and its subsequent disappearance were consistent across delivery platforms. Specifically, an ANOVA (if the variable was normally distributed) or the parallel non-parametric Kruskal-Wallis H test (if the variable wasn’t normally distributed) was performed on both pre- and post-survey data to ascertain whether there was any difference in student attitudes attributable to class platform: traditional face-to-face class, online class, and face-to-face class with Honors students.

For the pre-survey results, as noted above, the variable *Ability to comprehend economics* and the metric, *Overall attitude toward economics* were approximately normally distributed. Consequently, a one-way ANOVA was carried out on those variables. The corresponding non-parametric Kruskal-Wallis H test, was then conducted on the variables *Usefulness of economics* and *Interest in economics* because they weren’t normally distributed. The results indicate that there were no statistically significant differences on any of the four dependent variables based on class platform. The p-values on the variables, *Ability to* c*omprehend economics*, *Usefulness of economics, Interest in economics*, and *Overall attitude toward economics* were p = .86, p = .81, p = .66, and p = .85, respectively. Refer to Table 2 for the pre-course survey descriptive and inferential statistics.

For the post-course results the assumption of normality did not hold for any of the dependent variables so the Kruskal-Wallis H test was employed. Based on class platform, there was a statistically significant difference in only one variable, *Ability to comprehend economics*, where students enrolled in the Honors course gained more confidence in their ability to learn economics than students in the other two classes (p=.015). There was no statistical difference based on class format for the other three variables, *Usefulness of economics*, *Interest in economics*, and *Overall attitude toward economics,* where p = .056, p = .37, and p = .496, respectively. Refer to Table 2 for the post course descriptive and inferential statistics.

Table 2: Comparison of Student Attitudes by Platform

	Outset of Class				End of Class			
Dependent Variable	Face-to-Face (N=33)	Honor (N=25)	Online (N=35)	P-Val of Diff	Face-to-Face (N=45)	Honor (N=19)	Online (N=33)	P-Val of Diff
Ability to comprehend economics	4.51 (.91)	4.63 (.78)	4.55 (.91)	.86	4.78 (1.04)	5.55 (.54)	4.90 (1.23)	.015*
Usefulness of economics	5.47 (.60)	5.46 (.99)	5.40 (.75)	.81	5.20 (.96)	5.83 (.61)	5.45 (.93)	.056
Interest in economics	4.75 (1.10)	4.99 (.97)	5.00 (1.00)	.66	4.66 (1.33)	5.19 (.99)	4.85 (1.07)	.370
Overall attitude toward economics	4.92 (.70)	5.03 (.79)	4.98 (.79)	.85	4.89 (1.01)	5.54 (.64)	5.08 (.99)	.496

* Statistical significance at .05

Multiple Regression Analysis

To further examine the effect of class experience on the gender gap across class delivery platforms, a regression analysis was undertaken. The following multiple regression model was specified: Yi' = $b_0 + b_1$ (gender) + b_2 (post) + b_3 (female-post) + b_4 (upper) + b_5 (honors) + b_6 (online). Dummy variables were used in the in the model and were coded accordingly: 1 = female, 0 = male; 1 = post-test =1, 0 = pre-test; 1 =female post, 0 = otherwise; 1 =upper class, 0 = freshman or sophomore; 1 = honors class, 0 = otherwise; 1 = online class, 0 = otherwise. Multiple regressions were then conducted using SPSS 24. Refer to Table 3 for the regression equations.

Table 3: Regression Equations

Dependent Variable	Regression Equations
Ability to comprehend economics	Yi' = 4.527 + (-3.93) x (gender) + (.282) x (post) + (.431) x (female-post) + (.167) x (upper) + (.448) x (honors) + (.034) x online
Usefulness of economics	Yi' = 5.543 + (-.372) x (gender) + (-.156) x (post) + (.282) x (female-post) + (-.076) x (upper) + (.274) x (honors) + (.165) x (online)
Interest in economics	Yi' = 4.902 + (-.250) x (gender) + (-.162) x (post) + (.213) x (female-post) + (-.107) x (upper) + (.345) x (honors) + (.289 x (online)
Overall attitude toward economics	Yi' = 4.995 + (-.343) x (gender) + (-.005) x (post) + (.282) x (female-post) + (.000) x (upper) + (.356) x (honors) + (.156) x (online)

The conditions consistent with a vanishing gender gap are: (1) the coefficient b_1 was negative, indicating that female students had a more negative attitude regarding economics than male students, when controlling for the other predictors; (2) the coefficient b_3 was positive and the absolute value of b_1 was fairly close to b_3, indicating the gender gap had vanished by the conclusion of the semester. The first condition was met in all four regression equations. The coefficients on all of the variables were negative. The Parameter estimates for *Ability to comprehend economics*, *Usefulness of economics*, *Interest in economics*, and *Overall attitude toward economics* were b_1 = -.393, -.372, -.250, and -.343, respectively. The second condition was also met in all four of the regression equations. That is, the coefficient b_3 was positive and the absolute value of b_1 was close to b_3. The parameter estimates for b_1 and b_3 for *Ability to comprehend economics*, *Usefulness of economics*, *Interest in economics*, and *Overall attitude toward economics* were (-.393 to .341, -.372 to .282, -.250 to.213, and -.343 to .282, respectively). Thus, the multiple regression analysis was consistent with the previous statistical analyses. Seemingly, the class experience of female students was such that the gender gap evident at the onset of the class had vanished in all four measures of student attitude by the end of the semester. The following section briefly discusses the pedagogy used in the three classes.

Pedagogical Considerations

The pedagogy employed in all three principles classes is straightforward and involves intuitive understanding of economic concepts—rote learning is minimized—with mathematical analysis then employed to anchor the principles. Finally, current applications of the economic principles are presented. A simple case in point is the concept of equilibrium price in the context of the supply and demand model. The notion that the equilibrium price is the only stable price in a competitive market is illustrated by taking actual prices above and below the market clearing price and showing the concomitant downward pressure on price from a surplus and the upward pressure on price from a shortage. This leads to an intuitive understanding of the concept and an appreciation of the importance of the rationing function of price—supply and demand hopefully becomes more than a cliché. Then, the graphic and algebraic solutions to equilibrium are added. Finally, applications of the model involving current topics are illustrated.

The pedagogical approach employed in all three classes was identical except for one difference in each of the course platforms. The obvious dissimilarity is between the face-to-face classes and the online course in which there is no formal classroom experience. To ensure that online students were receiving the necessary explanations of microeconomic concepts, lectures on all of the principles of microeconomics were prepared and videotaped in an empty classroom, simulating a face-to-face-educational experience as closely as possible. The videos were then posted on blackboard—many students commented on the usefulness of the videos in student evaluations—along with accompanying power-points. The second difference in format involved the honors course where a writing component was included. Honors students were required to comment on five posts—one each week beginning in the middle of the semester. They could write on any post from three designated blog sites: gregmankiwblogspot.com, www.marginalrevolution.com, and econlog.econlib.org. In class evaluations, students commented favorably on the blog assignments.

The reasons for the vanishing gender gap can be gleaned somewhat from the responses of female students to the open-ended question on the survey at the end of the semester. Some typical answers include the following: “I have found throughout the semester it is very important to know when decision making.” “It is becoming more intriguing and interesting to me.” “Since beginning this course, I feel as though economics has become a greater part of my life.” “After taking this class, I am truly interested in economics. The material was taught in such an understandable manner that I feel the concepts are rather easy to grasp.” “I have thoroughly enjoyed this class and taking a deeper dive into economics.” From these representative responses, it appears that the “learning experience” of female students across all three class delivery platforms was the cause of the vanishing gender gap.

Conclusion

The critical finding of this study is that the male-female gender gap need not be taken as a given. This research strongly suggests that learning experience in a principles of microeconomics course can favorably impact the overall attitude toward economics of female undergraduate students, as well as bolster their confidence in the ability to learn economic concepts. Thus, adopting certain teaching practices in introductory economics courses may help alleviate the underrepresentation of women who have majored in the discipline, historically, and thereby close the gender gap in the economics profession. A secondary important finding is that the narrowing of the gender gap occurred independent of class delivery platform. Consequently, these teaching practices can be successfully employed in different classroom or online settings, thus narrowing the gender gap in diverse educational platforms.

Future research is needed to expand the findings of this study. First, a larger sample size would strengthen the inferences deduced from the paper's conclusions. Second, research could be expanded by including other types of institutions of higher learning—for example, women's colleges, small private liberal arts colleges, and large public universities in future studies. Third, additional research could also explore and identify those specific teaching styles and methods that closed the gender gap evidenced in this study.

References

Ballard, Charles, and Marianne Johnson. 2005. "Gender Expectations and Grades in Introductory Economics." *Feminist Economics* 2: 95-122.

Bansak, Cynthia, and Martha Starr. 2010. "Gender Differences in Predispositions toward Economics." *Eastern Economic Journal* 36: 33-57.

Bayer, Amanda, and Cecilia Rouse. 2016. "Diversity in the Economics Profession: A New Attack on an Old Problem." *Journal of Economic Perspectives* 30: 221-242.

Calkins, Lindsay Noble, and Andrew Welki. 2006. "Factors that Influence Choice of a Major: Why Some Students Never Consider Economics." *International Journal of Social Economics* 33: 547-564.

Dynan, Karen E., and Cecilia Elena Rouse. 1997. "The Underrepresentation of Women in Economics: A Study of Undergraduate Economics Students." *Journal of Economic Education* 28: 350-368.

Feiner, Susan, and Bruce Roberts. 1995. "Using Alternative Paradigms to Teach about Race and Gender: A Critical Thinking Approach to Introductory Economics." *America Economic Review* 85: 367-371.

Ferber, Marianne A. 1995. "The Study of Economics: A Feminist Critique." *American Economic Review* 85: 357-361.

Ginther, Donna K., and Shulamit Kahn. 2004. "Women in Economics: Moving Up or Falling Off the Academic Career Ladder." *Journal of Economic Perspectives* 18: 193-214.

Ginther, Donna K., and Shulamit Kahn. 2014. "Women's Careers in Academic Social Science: Progress, Pitfalls, and Plateaus." In *The Economics of Economists,* edited by Alessandro Lanteri and Jack Vromen. Cambridge: Cambridge University Press.

Hall, Roberta M., and Bernice R. Sandler. 1982. *The Classroom Climate: A Chilly One for Women?* Washington D.C. Association of American Colleges.

Jensen, Elizabeth J., and Ann L. Owen. 2000. "Why are Women Such Reluctant Economists? Evidence from Liberal Arts Colleges." *The American Economic Review* 90: 466-470.

Jensen, Elizabeth J., and Ann L. Owen. 2001. "Pedagogy, Gender, and Interest in Economics." *The Journal of Economic Education* 32: 323-343.

Kahn, Shulamit B. 1995. "Women in the Economics Profession." *Journal of Economic Perspectives* 9: 193-206.

Leeds, Michael A. 1992. "Who Benefits from Affirmative Action? The Case of the AEA Summer Minority Program 1986-1990." *Journal of Economic Perspectives* 6: 149-156.

Rask, Kevin, and Jill Tiefenthaler. 2008. "The Role of Grade Sensitivity in Explaining the Gender Imbalance in Undergraduate Economics." *Economics of Education Review* 27: 676-687.

APPENDIX: Questionnaire

Demographic Questions:

Gender

- ○ Male
- ○ Female

Is this class required for your degree?

- ○ Yes
- ○ No

Are you an Honor's College student?

- ○ Yes
- ○ No

Are you or do you plan on pursuing a minor in Economics?

- ○ Yes
- ○ No

What year in school are you?

- ○ Freshman
- ○ Sophomore
- ○ Junior
- ○ Senior

What is your major?

Indicate your opinion on each of the following:

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor Disagree	Somewhat agree	Agree	Strongly Agree
1. My current understanding of economics is good.	○	○	○	○	○	○	○
2. I am able to grasp the essence of current economic issues	○	○	○	○	○	○	○
3. I am able to understand and apply the math in economics.	○	○	○	○	○	○	○
4. Principles of Economics are easy to understand.	○	○	○	○	○	○	○
5. I find economics interesting.	○	○	○	○	○	○	○
6. I enjoy reading about topics in economic.	○	○	○	○	○	○	○
7. Economic models are helpful in understanding economic concepts.	○	○	○	○	○	○	○
8. I would like to learn more economics.	○	○	○	○	○	○	○
9. Economics is an important subject in a student's education.	○	○	○	○	○	○	○
10. I feel at ease with economics.	○	○	○	○	○	○	○
11. Economics is a valuable tool in business decision-making.	○	○	○	○	○	○	○
12. Economics is a valuable tool in personal decision-making.	○	○	○	○	○	○	○
13. Economics can be very exciting.	○	○	○	○	○	○	○
14. I could become really proficient in economics.	○	○	○	○	○	○	○
15. I like using economic concepts to analyze real-life situations.	○	○	○	○	○	○	○
16. I would like to have more opportunity to learn about economics.	○	○	○	○	○	○	○
17. To understand political issues, one needs some knowledge of economic principles.	○	○	○	○	○	○	○
18. I find economic issues stimulating.	○	○	○	○	○	○	○
19. Graphs in microeconomics are important tools for understanding economic principles.	○	○	○	○	○	○	○
20. I understand the economic way of thinking.	○	○	○	○	○	○	○

Briefly describe your overall impression of economics in terms of your interest in the subject matter, its ease or difficulty in understanding, and its overall usefulness and application in personal and business decision-making.

Face-to-Face Versus Online: A Comparison of Student Performance in Introduction to Finance Courses

Vaughn Cox[1]

ABSTRACT

This study compares the performance of students taking Introduction to Finance classes in an online versus a face-to-face setting. The study included 995 students. The analysis compared dropout rates, exam scores, final grades, and case study scores. The dropout rates of students in the online classes was more than twice that of students in the face-to-face classes. Students in the face-to-face classes received higher scores on their exams, case studies, and higher final grades, but the difference was small, less than the difference between the grade of a B and a B-.

Introduction

The question addressed in this study was whether students in an online class perform at the same level as students in a face-to-face class. This is an important question. The trend in higher education is towards more and more online learning with less face-to-face interaction with an instructor and with other students. The question of whether students can have an effective learning experience with the same or similar performance outcomes is a question of concern for all those who facilitate, develop, and teach online courses.

The performance of two groups of students, online and face-to-face, were compared. The class was an Introduction to Finance class that is required of all business majors. Students were either juniors or seniors. The group of students included in the study was made up of 995 students in the Woodbury School of Business at Utah Valley University in Orem, Utah. Of these students, 460 took the class online and 535 took the class face-to-face. The groups were made up of students from 20 different sections (9 online and 11 face-to-face) who took the class between the fall of 2012 and the summer of 2015.

Course Organization

This course is an Introduction to Finance course. It is a required course for all business majors. The majority of students are business management majors, but the class also includes students who are majoring in accounting, marketing, information systems, and finance, as well as a few students from other majors.

The class topics were divided into three major areas with an exam at the end of each. The first topic covered the introduction to the course, understanding financial statements, and financial analysis, which includes ratio analysis, calculating growth rates, and the DuPont Identity. The second topic covered time value of money, bonds and bond valuation, the capital asset pricing model, and common and preferred stock. The third topic covered capital budgeting, cash flow forecasting, and the weighted average cost of capital.

Students were expected to participate in a large number of activities to assist them in learning the course material. For each chapter covered in the text they were assigned to read the chapter, complete a reading review assignment, complete a problem assignment, and take a quiz. They were also assigned to participate in class, complete three case studies, and take three exams.

Significant effort was expended to assure that the online and ground classes covered the same material, had similar levels of rigor, and the same methods of assessment. Both the online and ground classes completed the same reading review assignments, problem assignments, and chapter quizzes. They completed the same case study assignments and took the same exams.

[1] Utah Valley University, 800 W University Parkway, MS 280 Orem, UT 84058-6703; vaughn.cox@uvu.edu.

There primary differences in the classes were in two key areas. The first was attending class, interacting with the instructor in class, and participating in the in-class discussions and activities. The second difference was in student-to-student discussions and relationships. Due to the nature of the online class format, online students had fewer opportunities than in a face-to-face setting to interact with the instructor. Online students also had fewer opportunities to interact with other students, to study with them, and to develop beneficial learning relationships with them.

As a result of these differences, the online course was designed differently than the face-to-face course. A series of video lectures were recorded by the instructor. Using the class notes, the instructor created PowerPoint slide presentations for each chapter. Using Camtasia video recording software, the instructor recorded video lectures. The online students could view these video lectures online. They would see the PowerPoint slides and see and hear the instructor talking. For each chapter there were one or more of these video lectures for the online students to view. In addition, online students were asked to participate in an online discussion board, answering questions and having online conversations with other students. A website, specific to the course, and the textbook were also provided. The website included access to the textbook, practice questions and problems, solutions, and explanations. The website also included videos that explained specific topics and methods for solving and completing problems.

Course Grading and Assessment

Class Participation and Attendance

The students in the face-to-face class received points for attendance and class participation. In the online class, students received points for participating in a weekly discussion board question. Online students were asked to make a post answering a specific question about the topic of the week and to comment on two of the posts made by other students in the class.

Reading Assignments

Students were assigned to read each chapter before the chapter is discussed in class. After the chapter reading, they were assigned to complete a reading review assignment. The assignment varied. It could have been completing a short concept review quiz or writing a short reflection on the reading. In the online sections, students were assigned to read the chapter and complete the reading review assignment before completing any of the other chapter assignments. This assignment was identical for both the face-to-face and online students.

Chapter Problem Assignments

As part of the study of each chapter, students were assigned one or sometimes two problem assignments. The assignment was made up of problems similar to those found at the end of the chapter. This assignment was identical for both the online and face-to-face students.

Chapter Quizzes

At the end of the study of each chapter, students were assigned to take a multiple-choice, ten-question quiz. The quiz was typically made up of conceptual and application questions from the chapter material. These quizzes were identical for both the online and face-to-face students.

Exams

The students took three exams. The first exam covered the introduction to the class, financial statements, and financial statement analysis. The second exam covered time value of money, bond principles and valuation, the Capital Asset Pricing model, and stock principles and valuation. The third exam covered capital budgeting, cash flow forecasting, and the cost of capital. The exams were the same for both the online and face-to-face classes. The exams were objective multiple choice exams that could be taken either online or as a paper exam. There was no cumulative final exam in the class.

Case Studies

Students were assigned to complete three case studies during the course. They were expected to be able to complete the financial calculations and analysis and to write a report that explains the analysis performed, results, and conclusions of the analysis. Students were taught that to be successful in finance you must first be able to perform the analysis and then you must be able to clearly explain what you have done to someone not experienced in finance. The case studies were graded by the course instructor using a standardized grading rubric.

The first case study was a detailed financial analysis assignment. Students were asked to complete a financial analysis of a company comparing it to a competitor or the industry. The second case study was a complex time value of money problem that included completing a retirement plan or a college savings plan for a family including a sensitivity analysis. The third case study was a complex capital budgeting problem. Students were asked to consider different investment opportunities and to create the cash flow forecasts and perform the analysis to determine which project will provide the best financial results to the company.

Literature Review

The use of online classes in post-secondary education has been increasing at rates much faster than enrollment. In 2012, nearly 7.4 million or 32.0% of all students participated in one or more distance education classes (NCES 2015). For students participating in a 4-year baccalaureate program the number is even higher at 53.0%. The number of students taking online classes increased at an annual rate of 8% from the 2003-2004 academic year to 2011-2012 (NCES 2016). Over this same period, the number of students participating exclusively in online programs has increased from 4.9% to 6.5%, nearly 1.5 million students (NCES 2016).

Retention and Dropout Rates

As the number of students taking online courses has increased, one of the most significant concerns is the low retention rates in many courses (Heyman 2010). Herbert (2006) concluded that the dropout rate of online courses is often 10% to 20%. Smith (2010) found that in some online classes as many as 40% of students do not finish the course. Levy (2007) found that students at a lower level of college were at a higher risk of dropping out, primarily because they were less prepared to deal with the academic rigors of an online class. One of the reasons for this high dropout rate may be the positive effect of face-to-face interaction on a student’s success, which is not part of most online classes (Shea and Bidjerano, 2010). On the opposite spectrum, Bawa (2016) found that students who do not dropout and succeed in online courses, have a sense of control and feeling of competency after completing the class.

Student Success in Online Classes

Many studies have identified that the student success of online classes is similar to that of traditional face-to-face classes. Lou et al. (2006) concluded that distance education students seem to learn equally as well as students in traditional classes. Bernard et al. (2004) determined that classroom instruction and distance education are comparable. Tallent-Runnels et al. (2006) found that both methods of delivery, online and face-to-face, are adequate. Larson and Sung (2009), using student grades and final exam scores, found no significant difference regarding student satisfaction, learning effectiveness, and faculty satisfaction. Kearns et al. (2004) found that students in a traditional class were more satisfied with the course than students in a web-based class. Howsen and Lile (2008) examined the test scores of students in macroeconomics courses and found that online students had lower test scores than traditional face-to-face students. Summers et al. (2005) surveyed students in statistics courses and found that the students in the online classes were less satisfied than the students in traditional classes. Iverson et al. (2005) found that online learners had a higher level of enjoyment and utility and a significantly stronger intent to transfer their learning. Other studies have identified that student success is a function of student motivation or class design. Klein et al. (2006)

determined that the characteristics of the learner, specifically student motivation, is significantly related to the course outcome. Lim and Morris (2009) concluded that both student motivation and the quality of the online instructor were significant variables influencing course outcomes. Sankaran and Bui (2001) determined that the learning strategies used in the class and the level of student motivation affected learning outcomes.

The Use of Technology in Online Classes

The use of computer technology in a course tends to increase the interest level of students (Shrivastava 1999). Tamim et al. (2011) conducted a meta-analysis of 1055 primary studies that examined the use of technology in the classroom and concluded that the use of technology enhances learning. Computer learning is effective because it provides cognitive efficiencies; a variety of tasks and tools that facilitate cognitive learning can be effectively incorporated in a computer based online class (Cobb 1997). Technology enhances learning and learning is best supported when the student is engaged in active, meaningful exercises via technological tools that improve cognitive support (Schmid et al. 2014).

Class Design and Learning Tools in Online Classes

Class design and the inclusion of appropriate learning tools are a significant factor in student success in online classes. Moulton et al. (2006) identified that mixed practice, which is the inclusion of a variety of topics and problems in an assignment, helps a student to learn and remember more effectively than massed practice, which is the inclusion of many similar problems in an assignment. This process of interleaving topics in an assignment increases the retention level of students (Cepeda et al. 2006). Roediger and Karpicke (2006) found that spaced remembering, the act of learning information and then recalling that information at some time in the future increases the learning level of students. Teaching students and then later quizzing them on the information prior to the test resulted in increased test scores (Roediger et al. 2011). The greater the frequency of quizzes or other forms of spaced remembering, the better the test scores (McDaniel et al. 2011). The learning tool of generation is another activity that can be used to increase learner retention (Richland et al. 2009). Generation is the idea of creating learning opportunities where students are required to answer a question or solve a problem. An example of generation is a fill in the blank or short essay question. Generation has been shown to increase student learning and retention (Jacoby 1978). Reflection, another learning tool, is the act of restating a learned concept in one's own words. This can be done verbally or in writing. A student that uses reflection has an increased level of learning and retention (Gingerich et al. 2014). Courses that are designed to include these types of learning tools will result in a higher level of student success.

Research Question and Hypotheses

The research question addressed in this study was whether students who take a face-to-face Introduction to Finance class learn and perform at the same levels as students who take the class in an online environment. This question was addressed in two ways: first, do the students complete the class or dropout, and second, for those students who complete the class, does the performance of students in face-to-face classes differ from the performance of students in online classes.

It should be noted that the two groups of students were not randomly selected, rather they were self-selected. Each student selected an online or face-to-face class based on their personal needs and preferences. This raises the issue of selection bias. Any difference in performance results may not be due to the differences in the teaching and class delivery of the online or face-to-face classes, but rather due to differences in the samples as a result of the self-selection. Even though both groups were self-selected, all the students had similar characteristics. First, they were all college students, matriculated into the business school, and required to take the class. Second, they were all experienced, juniors or seniors seeking a business degree that had all completed the same required prerequisite classes.

There are two hypotheses in this study:

$H1_0$: There is no difference in the dropout rate of face-to-face and online students who took the Introduction to Finance class.

$H1_a$ There is a difference in dropout rate of face-to-face and online students who took the Introduction to Finance class.

$H2_0$: There is no difference in the performance of face-to-face and online students who took the Introduction to Finance class.

$H2_a$ There is a difference in performance of face-to-face and online students who took the Introduction to Finance class.

Methodology

Two areas of student performance were measured in this study. The first was course completion. Did the student stick with the course and complete it or did the student drop out before the course was completed? The second area of student performance was how well did the students perform in a variety of assessments completed during the course. In other words, how did the grades of the face-to-face students compare with the grades of the online students?

Student grades were compared in three areas. First, in the area of exams, students took three exams taken during the course. Second, in the area of case studies, students completed three case studies during the course. The third and final area of comparison was the final grade of the students in the course.

Dropout Rates

A dropout was defined as any student that officially dropped the course, any student that gave up and quit during the course, and any student who missed most of the assignments and did not take the final exam. The method of determining the completion or dropout rates was simply to determine which students dropped out and then to calculate the percentage dropout rate and the percentage completion rate.

Student Performance

There were nine different areas where student performance was compared. Four were in the area of exams. A comparison was made for each of the three exams and for the combination of all exam scores. Four comparisons were also made in the area of case studies. A comparison was made for each of the three case studies and for the combination of all case study scores. The ninth and final comparison was for the final grade or final percentage score for the course.

Only the scores of completed exams and case studies were included in the data. If a student did not complete a case study or take an exam, that student's score, which was zero, was not included in the data. The intent of the study was to compare the completed work of the face-to-face students with the completed work of the online students.

For each of the nine areas of comparison the descriptive statistics, including the mean, median, and standard deviation, were calculated for the online students, the face-to-face students, and the full sample of students. A two-sample t-test was also performed comparing the results of the online students to the face-to-face students.

Analysis and Results

Completion and Dropout Rate

The completion and dropout rates are presented in Table 1. Of the 460 online students that took the class, 401 or 87.2% completed the class and 59 or 12.8% did not complete the class. Of the 59, 24 withdrew from the class and 35 quit participating in the class. Of the 535 face-to-face students that took the class 503 or 94.0% completed the class and 32 or 6.0% dropped out. Of the 32 that dropped out seven withdrew from the class and 25 quit participating in the class. The dropout rate for the online students was more than twice the dropout rate of the face-to-face students.

Table 1: Completion and Dropout Rates

	Online Students		Face-to-Face Students	
Total Students	460	100.0%	535	100.0%
Formal Withdrawal	24	5.2%	7	1.3%
Quit-No Withdrawal	35	7.6%	25	4.7%
Total Withdrawn	59	12.8%	32	6.0%
Completed	401	87.2%	503	94.0%

Student Performance – Exams

The descriptive statistics for the three exams are presented in Table 2. Each exam was worth 100 points. The mean score for all students was 80.4. For each exam, the mean score was higher among the face-to-face students than the online students, but the difference was small. On average, the face-to-face scores were 1.53 points higher than the online scores. The range varied from a low of 1.1 to a high of 2.1 points.

Table 2: Descriptive Statistics: Exams

	Online Exams				Face-to-Face Exams				
	Exam 1	Exam 2	Exam 2	All Online	Exam 1	Exam 2	Exam 3	All Face-to-Face	All Exams
Number of Students	400	400	399	1,199	503	503	503	1,509	2,708
Mean Score	80.8	78.9	78.9	79.5	82.2	81.0	80.0	81.1	80.4
Median Score	82.0	80.8	80.5	81.0	83.8	82.5	82.5	82.5	82.0
Standard Deviation	12.2	14.6	14.6	13.7	10.9	11.7	13.7	12.2	13.0

A two-sample t-test was performed comparing the results of online students with the results of the face-to-face students for each exam and for the combined exams. The t-test identified that there is a statistically significant difference between the students' scores on each exam and on all exams combined. The results are presented in Table 3.

Table 3: T-Test: Exams

2 Sample T-test: Online & Face-to-Face	Estimate of Difference	Confidence Interval 95%	T - Value	P - Value	Degrees of Freedom
Exam 1	-1.336	-3.86, 0.188	-1.720	0.086	807
Exam 2	-2.092	-3.86, -0.33	-2.330	0.020	755
Exam 3	-1.139	-3.01, 0.73	-1.200	0.231	829
All Exams	-1.522	-2.52, -0.54	-3.000	0.003	2,403

The results of the t-test are that there was a statistically significant difference between the scores of the online students and the face-to-faces students. Since there is a difference between the results of the two groups, a Cohen's d effect size test was performed. Cohen's d is an effect size used to indicate the standardized difference between the two means (Cohen 1988). The results of the Cohen's d test are presented in Table 4. An effect size of 0.2 or less is considered small. An effect size between 0.2 and 0.7 is considered medium, and an effect size of 0.8 or more is considered large (Becker 2000).

Table 4: Cohen's d Effect Size: Exams

	Exam 1	Exam 2	Exam 3	Combined Exams
Cohen's d Effect Size	0.1215	0.1587	0.0780	0.0190

For each exam and for the exams combined the Cohen's d effect size is small. From a grading perspective, it is less than the difference between say, a B and a B-.

Student Performance – Case Studies

The descriptive statistics for the three case studies are presented in Table 5. Each case study was worth 30 points. The mean score for all students was 25.8. For each case study, the mean score was higher among the face-to-face students than the online students. On average, the face-to-face scores were 1.17 points higher than the online scores. The range varied from a low of 0.4 to a high of 2.5 points.

Table 5: Descriptive Statistics: Case Studies

	Online Students				Face-to-Face Students				All
	Case 1	Case 2	Case 3	Combined	Case 1	Case 2	Case 3	Combined	Combined
Num. Students	381	382	360	1,123	472	479	477	1,428	2,551
Mean Score	25.7	25.0	24.8	25.2	26.3	25.4	27.3	26.3	25.8
Median Score	26.0	26.0	25.0	26.0	26.0	26.0	26.0	27.0	26.0
Std, Deviation	2.9	3.0	2.9	2.9	3.2	3.5	2.7	3.2	3.1

A two-sample t-test was performed comparing the results of online students with the results of the face-to-face students for each case study and for the combined case study scores. The t-test identified that there was a statistically significant difference between the students' scores on each case study and on all case studies combined. The results are presented in Table 6.

Table 6: T-Test: Case Studies

2 Sample T-Test: Online & Face-to-Face	Estimate for Difference	Confidence Interval–95%	T - Value	P - Value	Degrees of Freedom
Case 1	-0.664	-1.07, -0.26	-3.22	-0.001	839
Case 2	-2.455	-0.83, 0.04	-1.80	0.072	756
Case 3	-2.445	-2.84, -2.07	-12.630	0.000	756
All Cases Combined	-1.165	-1.40, -0.93	-9.570	0.000	2,500

The results of the t-test were that there was a statistically significant difference between the scores of the online students and the face-to-faces students. Since there is a difference between the results of the two groups, a Cohen's d effect size test was performed. Cohen's d is an effect size used to indicate the standardized difference between the two means. The results of the Cohen's d test are presented in Table 7.

Table 7: Cohen's d Effect Size: Case Studies

	Case 1	Case 2	Case 3	All Cases
Cohen's d Effect Size	0.2227	0.12367	0.883	0.381

The Cohen's d effect size was different for each case. For Case 1 it was medium in size, for case 2 it was small, and for case 3 it was large. For the combined cases, it was medium in size. From a grading perspective, the difference in case 1 would be the equivalent of the difference between a B and a B-. For case 2 the difference is less than the difference between say a B and a B-. For case 3 the difference is similar between the difference between a B and a C grade.

Student Performance – Final Grades

The descriptive statistics for the final grades are presented in Table 8. The grades are presented as percentages. The mean score for all students was 83.1. The mean grade for the face-to-face students was 83.8 compared to 82.3 for the online students, a difference of 1.5 percentage points.

Table 8: Descriptive Statistics: Final Grades

	Online Students	Face to Face Students	All Students
Number of Students	401.0	503.0	904.0
Mean Score	82.3	83.8	83.1
Median Score	82.9	84.9	83.8
Standard Deviation	9.9	9.4	9.7

A two-sample t-test was performed comparing the results of online students with the results of the face-to-face students. The t-test identified that there is a statistically significant difference between the students' final grades. The results are presented in Table 9.

Table 9: T-Test: Final Grades

2 Sample T-Test: Online & Face to face:	Estimate for Difference	Conf. Interval 95%	T - Value	P – Value	Degrees of Freedom
Final Grades	-1.474	-2.747, -0.202	-2.270	0.023	839

The results of the t-test were that there was a statistically significant difference between the final grades of the online students and the face-to-faces students. Since there is a difference between the results of the two groups, a Cohen's d effect size test was performed. The results of the Cohen's d test are presented in Table 10. For the final grades, the Cohen's d effect size was small. From a grading perspective it is less than the difference between, say a B and a B-.

Table 10: Cohen's d Effect Size: Final Grades

	Final Grade
Cohen's d Effect Size	0.15529

Implications, Recommendations, & Conclusions

The conclusions of the study are both positive and negative. The positive aspect is that the performance results for online learning, while not the same as with face-to-face learning, were similar. Ideally, the performance results would be the same. The success of online students would be equal to, or perhaps better, than traditional face to face learning. But, the results, identifying that the performance of online and face-to-face results were similar, supports the conclusion that online learning is effective. Efforts to improve online instruction should continue with the goal of making online instruction as effective or more effective as face-to-face instruction.

The negative aspect is the poor retention rate of online classes. How can online classes be modified so that there is better student retention? Solving this problem, is, perhaps, the most significant challenge facing online learning today.

Regarding the first hypothesis, that there is no difference between the dropout rates in the online classes compared to the face-to-face classes, we must reject the null hypothesis. Regarding the second hypothesis, that there is no difference in student performance in the areas of scores and grades between the two groups of students, must also be rejected. It was found that in the three areas of comparison, exams, case studies, and final grades, the face-to-face students performed better than the online students did. While the face-to-face students performed better overall, the difference was small. The average final grade percentage was 82.3% for the online students compared to 83.8% for the face-to-face students. This difference of 1.5 points is not large.

References

Bawa, Papia. 2016. "Retention in Online Courses: Exploring Issues and Solutions – A Literature Review." *Sage Open* 6: 1-11.

Becker, Lee. 2000. "Effect size (ES)." Retrieved from: https://www.uv.es/~friasnav/EffectSizeBecker.pdf

Bernard, Robert M., Philip C. Abrami, Yiping Lou, Eugene Borokhovski, Anne Wade, and Lori Wozney. 2004. "How Does Distance Education Compare With Classroom Instruction? A Meta-Analysis of the Empirical Literature." *Review of Educational Research* 74: 379-439.

Cepeda, Nicholas J., Howard Pashler, Edward Vul, John T. Wixted, and Doug Rohrer. 2006. "Distributed Practice in Verbal Recall Tasks: A Review and Quantitative Synthesis." *Psychological Bulletin* 132: 354-380.

Cobb, Tom. 1997. "Cognitive Efficiency: Toward a Revised Theory of Media." *Educational Technology Research and Development* 45: 21-35.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Mahwah NJ: Lawrence Erlbaum Associates.

Gingerich, Karla J., Julie M. Bugg, Sue R. Doe, Christopher A. Rowland, Tracy L. Richards, Sara Anne Tompkins, and Mark A. McDaniel. 2014. "Active Processing via Write-to-Learn Assignments: Learning and Retention Benefits in Introductory Psychology." *Teaching of Psychology* 41: 303-308.

Herbert, Michael. 2006. "Staying the Course: A Study in Online Student Satisfaction and Retention." *Online Journal of Distance Learning Administration* 9.

Heyman, Errin. 2010. *Overcoming Student Retention Issues in Higher Education Online Programs*: *A Delphi Study*. Retrieved from ProQuest Dissertations and Theses database.

Howsen, Roy M., and Stephen Lile. (2008). "A Comparison of course Delivery Methods: An Exercise in Experimental Economics." *Journal of Economics and Finance Education* 7: 21-28.

Iverson, Kathleen M., Debora L Colky, and Vincent L. Cyboran. 2005. "E-learning Takes the Lead: An Empirical Investigation of Learner Differences in Online and Classroom Delivery." *Performance Improvement Quarterly* 18: 5-18.

Jacoby, Larry. 1978. "On Interpreting the Effects of Repetition: Solving a Problem Versus Remembering a Solution." *Journal of Verbal Learning and Verbal Behavior* 17: 649-667.

Kearns, Lauren E., Jennifer R. Shoaf, and Marguerite B. Summey. 2004. "Performance and Satisfaction of Second-Degree BSN Students in Web-Based and Traditional Course Delivery Environments." *Journal of Nursing Education* 43: 280-284.

Klein, Howard J., Raymond A. Noe, and Chongwei Wang. 2006. "Motivation to Learn and Course Outcomes: The Impact of Delivery Mode, Learning Goal Orientation, and Perceived Barriers and Enablers." *Personnel Psychology* 59: 665-702.

Larson, David, and Chung-Hsien Sung. 2009. "Comparing Student Performance: Online Versus Blended Versus Face-to-Face." *Journal of Asynchronous Learning Networks* 13: 31-42.

Levy, Yair. 2007. "Comparing Dropouts and Persistence in e-learning Courses." *Computers & Education* 48: 185-204.

Lim, Doo Hun, and Michael Lane Morris. 2009. "Learner and Instructional Factors Influencing Outcome Within a Blended Learning Environment." *Educational Technology & Society* 12: 282-293.

Lou, Yiping, Robert Bernard, and Philip Abrami. 2006. "Media and Pedagogy in Undergraduate Distance Education: A Theory-Based Meta-Analysis of Empirical Literature." *Educational Technology Research and Development* 54: 141-176.

McDaniel, Mark A., Pooja K. Agarwal, Barbie J. Huelser, Kathleen B. McDermott, and Henry. L. Roediger III. 2011. "Test Enhanced Learning in a Middle School Science Classroom: The Effects of Quiz Frequency and Placement." *Journal of Educational Psychology* 103: 399-414.

Moulton, Carol-Anne, Adam Dubrowski, Helen MacRae, Brent Graham, Ethan Grober, and Richard Reznick. 2006. "Teaching Surgical Skills: What Kind of Practice Makes Perfect?" *Annals of Surgery* 244: 400-409.

NCES. 2015. U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Spring 2013 and Spring 2014, Enrollment component. (This table was prepared December 2014.)

NCES. 2016. U.S. Department of Education, National Center for Education Statistics. *Digest of Education Statistics, 2014* (Table 311.15).

Richland, Lindsey E., Nate Kornell, and Liche Sean Kao. 2009. "The Pretesting Effect. Do Unsuccessful Retrieval Effects Enhance Learning?" *Journal of Experimental Psychology: Applied* 15: 243-257.

Roediger III, Henry L., and Jeffrey D. Karpicke. 2006. "Test Enhanced Learning: Taking Memory Tests Improves Long-Term Retention." *Psychological Science* 17: 249-255.

Roediger III, Henry L., Pooja K. Agarwal, Mark A. McDaniel, and Kathleen McDermott. 2011. "Test Enhanced Learning in the Classroom: Long-Term Improvements from Quizzing." *Journal of Experimental Psychology: Applied* 17: 382-395.

Sankaran, Siva R., and Tung Bui. 2001. "Impact of Learning Strategies and Motivation on Performance: A Study in Web-Based Instruction." *Journal of Instructional Psychology* 28: 191-198.

Schmid, Richard F., Robert M. Bernard, Eugene Borokhovski, Rana M Tamim, Philip C. Abrami, Michael A. Surkes, C. Anne Wade, and Jonathan Woods. 2014. "The Effects of Technology Use in Postsecondary Education: A Meta-Analysis of Classroom Applications." *Computers & Education* 72: 271-291.

Shea, Peter, and Temi Bidjerano. 2010. "Learning Presence: Towards a Theory of Self-Efficacy, Self-Regulation, and the Development of a Communities of Inquiry in Online and Blended Learning Environments." *Computers & Education* 55: 1721-1731.

Shrivastava, Paul. 1999. "Management Classes as Online Learning Communities." *Journal of Management Education* 23: 691-702.

Smith, Belinda G. 2010. *E-Learning Technologies: A Comparative Study of Adult Learners Enrolled on Blended and Online Campuses Engaging in a Viral Classroom.* Retrieved from ProQuest Dissertations and Theses database.

Summers, Jessica J., Alexander Waigandt, and Tiffany A. Whittaker. 2005. "A Comparison of Student Achievement and Satisfaction in an Online Versus a Traditional Face-to-Face Statistics Class. *Innovative Higher Education* 29: 233-250.

Tallent-Runnels, Mary K., Julie A. Thomas, William Y. Lan, Sandi Cooper, Terence C. Ahern, and Shauna. M. Shaw. 2006. "Teaching Courses Online: A Review of the Research." *Review of Educational Research* 76: 93-135.

Tamim, Rana M., Robert M. Bernard, Eugene Borokhovski, Philip C. Abrami, and Richard F. Schmid. 2011. "What Forty Years of Research Says About the Impact of Technology on Learning: A Second-Order Meta-Analysis and Validation Study." *Review of Educational Research* 81: 4-28.

On the Value of Teaching Edgeworth Boxes in Introductory Economics Courses

***M. Garrett Roth*[1]**

ABSTRACT

Edgeworth boxes, which illustrate the movement toward equilibrium via trade, are almost universally relegated to intermediate microeconomics courses. I argue for their inclusion in introductory courses as a natural bridge between production / consumption possibility frontiers and supply / demand curves. When presented in an intuitive, graph-based fashion, the Edgeworth box model provides a supplementary illustration (to the Ricardian model) of trade as a Pareto improvement and emphasizes production for the purpose of utility via consumption rather than for its own sake. A course which only includes production as such is, therefore, incomplete.

Introduction

"Two people are stranded on a deserted island …" This child-like prompt motivates a sophisticated economic model: the Edgeworth box (Edgeworth 1881). Despite its potential complexity, however, a simple, qualitative and, most of all, intuition-based exposition of this idea provides a necessary addendum within the general topic of the gains from trade and, more broadly, emphasizes the social value of market exchange. Accordingly, this paper lays out a brief pedagogical argument for the inclusion of Edgeworth boxes, particularly in conjunction with production possibility frontiers (PPFs), in introductory (micro)economics courses.

The inclusion of Edgeworth boxes has several advantages upon which I will subsequently expand. First, it provides a much needed bridge between an illustration of the "gains from trade" via two-person (or country), two-good consumption possibility frontiers and ordinary, *n*-person markets modeled via supply and demand curves. Secondly, Edgeworth boxes, in modeling efficiency in allocation, explain how a society arrives at a particular point on the production possibility frontier (PPF), thereby clarifying the valid yet open-ended assertion that all points on the PPF and consumption possibility frontier (CPF) are "equally good" from a social perspective. Students are, thus, transported from a Stalin-esque world where an efficient production point is specified arbitrarily to a world where individuals produce, trade, and consume as part of a Hayekian "spontaneous order." Thirdly, Edgeworth boxes introduce the microeconomic cornerstone of indifference curves without the more sophisticated ideas of substitution and income effects. Thus, the introduction of an indifference curve within the Edgeworth box is a useful "warm-up" to the pairing of indifference curves with budget lines in intermediate microeconomics courses; the foundational concept of Pareto optimality is, likewise, introduced.

Fourthly, Edgeworth boxes provide a second illustration of the necessary "gains from trade," as shown by the ever-shrinking core, which omits much of the (potentially valid) "winners and losers" aspect of country-to-country trade typical of PPF-CPF analysis; English wine-makers and Portuguese wool-makers, to use Ricardo's classic example, may be worse off if their products are now imported rather than domestically produced (Ricardo 1817). Thus, Edgeworth boxes bolster economic arguments for free trade by making universal, mutual gain from trade by individuals unambiguously superior to self-sufficiency. Lastly, the interrelation of Edgeworth boxes and PPFs provide a springboard for an exposition of equity concerns via utility possibility curves. Deriving, rather than hypothesizing, a grand utility possibility curve, provides a basis for the introduction of Social Choice theory, which itself serves as a convenient motivation

[1] Assistant Professor of Economics, Dahlkemper School of Business, Gannon University, 109 University Square, Erie, PA 16501, mgarrettroth@gmail.com. The author wishes to thank participants at the 2017 Robert Morris University Teaching Economics Conference, Prof. William P. McAndrew, and two anonymous referees for helpful comments on earlier versions of this paper.

for a discussion of voting rules ("how do we decide which social welfare function is correct?") and public choice theory ("why should we expect our hypothetical dictator to be benevolent?").

The remainder of the paper proceeds as follows: The next section describes the current sequence of topics in introductory (micro)economics courses. I also address the argument that Edgeworth box analysis, properly delivered, is too mathematically complex for an introductory course. The subsequent section presents more thorough arguments for the inclusion of Edgeworth boxes within an introductory curriculum *as such*. For the sake of expositional completeness, I then briefly summarize the long-established relationship between Edgeworth boxes, PPF's, and utility possibility curves. The final section briefly concludes.

Exchange in Introductory Economics Texts

In nearly all introductory economics textbooks, production and exchange are largely separate topics, at least as a substantive, quantitative matter. Trade-offs are first presented solely within the context of production possibilities, opportunity costs, and comparative advantage and subsequently abstracted into a buyer-seller relationship within the ordinary, modern context of a money-for-goods exchange. Where some bridge exists between the two subjects (Baumol and Blinder 2011; Gwartney et. al. 2014; Hall and Lieberman 2012), it is terse and qualitative.[2] The systematic, *quantitative* treatment of where an economy *will* emerge on a production or consumption possibility frontier is, thus, relegated to a perfunctory afterthought. Similarly, the gap between a goods-for-goods exchange typical of the country-level "gains from trade" problem simply morphs into *n*-person supply and demand curves intersecting at equilibrium. The topic of two-person exchange via endowments, as modeled by an Edgeworth box diagram is, consequently, sorely neglected. This omission is tantamount to explaining the first, but not second or third section of a novel; students know that economic agents produce according to comparative advantage but not how such (efficient) production relates to barter and consumption.

I expect that the customary objection to including an exposition of Edgeworth boxes in an introductory course is that such analysis is either too conceptually challenging or that it necessarily requires the use of calculus. Consider, however, the almost universal inclusion of budget lines and indifference curves, though often relegated to "Topics for Further Study" as in Mankiw (2014, Ch. 21), in introductory texts. Their widespread presence indicates that the concept of indifference curves is generally considered appropriate, if somewhat challenging, for introductory courses. Given the absence of such complex notions as "substitution effects" and "income effects," which are customary to any basic exposition of consumer choice, Edgeworth boxes are, in fact, simpler conceptually than the consumer choice analysis that is usually included, sometimes very early (e.g. Hall and Lieberman 2012), in introductory texts. In short, Edgeworth boxes introduce indifference curves without the need to also consider budget lines or any detailed explanation of the dual effects caused by relative price changes (despite the central role of relative prices in the movement toward equilibrium in an Edgeworth box).

Empirical evidence gathered from my own microeconomics classes reinforces the accessibility of the material, when explained as qualitatively as possible, at the introductory level.[3] In the first exam of the spring 2018 semester, overall multiple choice percentages and the frequency of correct answers on the Edgeworth box multiple choice question were less than one percentage point apart. In the final exam of the previous semester, the corresponding scores fell within two percentage points (with the Edgeworth box percentage higher than the overall percentage). Thus, student comprehension of the Edgeworth box material was not at all dissimilar from the more orthodox elements of the first section of the course, both when first introduced and at end of semester.

The notion that a thorough exposition of Edgeworth boxes *requires* calculus is also a misnomer. While it is often difficult to convey *any* marginal concept without differentiation, the story of "two castaways who trade what they've found" is a simple narrative which is easy to also simultaneously explain graphically.

[2] This is solely a criticism of major introductory textbooks and is not meant to imply that individual professors do not or could not (independently) find a suitable theoretical means by which to unify the production possibility frontier framework with a supply and demand framework exclusive of Edgeworth boxes.

[3] IRB approval was obtained to report exam statistics (Gannon University IRB #18-02-10).

The contract curve and core, as such, are not crucial to basic comprehension of trade for mutual benefit.[4] That any post-trade allocation must lie within the core is, to the contrary, self-evident when assuming that both parties will only trade on order to be strictly better off. One can, therefore, replace a term such as "marginal rate of substitution" with "relative value" which, although identical in meaning, is nevertheless more intuitive and less technical. The equilibrium point where traders' indifference curves are tangent can be colloquialized into "just touching at a point," etc. Thus, in this particular case, the calculus component can easily be excluded to no great detriment as with comparable, universal introductory course topics like elasticity and marginal revenue.

The Value of Trade, Production for Utility, and Social Choice

The previous section asserted that (i) some bridge between comparative advantage in (international) trade and supply-demand analysis is important and (ii) Edgeworth box diagrams are within the bounds of graphical abstraction of economic behavior that is typical of an introductory economics course. In this section, I discuss the advantages in both the argument for free trade and spontaneous order that Edgeworth boxes, in conjunction with a PPF, can convey to beginners in economic theory.

As addendum to the previous section, consider the goal of economists in espousing the value of markets: the net utility gain to individuals inherent in free trade *rather than the output of goods and services as such*. The production possibility frontier, though a paradigmatic example of the phenomenon of scarcity, does not illustrate the utility value in exchange but in efficient production under resource constraints.[5] Likewise, examples of specialization and comparative advantage are focused on obtaining more tangible goods and services rather than more utility as such.[6] Successive trading of initial endowments at the individual level, however, clearly illustrates the utility gains that must be made (in expectation) by each individual trader. Thus, an Edgeworth box, moreso than dual consumption possibility frontiers, illustrates the necessary gains in *utility* via voluntary exchange. Consequently, the popular myth that economists are obsessed with production for its own sake is replaced by output which is (efficiently) produced and traded strictly for the sake of bestowing satisfaction (via increased utility) to traders.

Also consider the indeterminacy of both consumption and production possibility frontiers. The goal in these constructions is to produce and trade efficiently from the *sole* perspective of output maximization. Accordingly, the link between production, exchange, and consumption for utility is only implied; the particular production / consumption point chosen is peripheral to the analysis. However, nesting an Edgeworth box within a production or consumption possibility frontier completes the narrative; expanded total endowments ultimately lead to increased post-trade utility by all individuals.[7] Though no single, optimal point exists on a particular PPF, production for maximum utility via consumption makes apparent the instinctual idea that a mix of (complementary) goods is typically preferable to consumers over extrema toward either the x or y axis.[8]

Lastly, consider the theoretical leap from strict efficiency considerations to questions of equity, which are necessarily political in nature. The mainstream of economics, by and large, operated on a "benevolent dictator" model of equity considerations via Social Choice theory until the Public Choice revolution of the 1960's as catalyzed by Buchanan and Tullock (1962). The fundamentals of Social Choice as a quantitative exercise follow directly from the pairing of Edgeworth boxes and PPF's. The hypothetical "benevolent dictator" should prefer a Pareto optimal point of utility distribution to a Pareto inferior point, *ceteris*

[4] The contract curve is, however, important in the construction of utility possibility curves, which can (and probably should) obviously be introduced at a later point in an introductory course.

[5] Efficiency in production is, of course, derived from iso-quant and iso-cost curves which are certainly beyond the scope of any introductory economics course. We (as teachers) are, therefore, left to decry as intuitively silly the idea of putting the best gun-smiths in charge of dairy farming and vice-versa.

[6] Naturally, this distinction is, under the premise of non-satiation, on *philosophical* emphasis rather than any actual mechanical discrepancy.

[7] This relationship is given in detail in a separate section of the paper.

[8] The criterion for joint efficiency in production and trade is not strictly necessary but, if provided colloquially rather than mathematically, could nevertheless help to illustrate why some "local" equilibria on a particular contract curve are not "globally" Pareto optimal.

paribus.[9] Establishing the necessary (Pareto) efficiency of markets in reaching the grand utility possibility curve (UPC) provides a natural bridge to questions of fairness. Having challenged students with the interrelation between efficiency in production and allocation, students should both: (i) understand the mechanical complexities of arriving at a point on the grand UPC (which, nonetheless, requires *no central direction* of any sort) and (ii) be skeptical of the ability of any individual (or group of individuals), benevolent or otherwise, to reallocate resources such that society will arrive at some other, more equitable point on the grand UPC.[10] Social Choice theory, properly motivated, thus becomes a departure point to topics in political economy such as voting rules and Public Choice, which are often included in introductory texts (e.g. Mankiw 2014, Ch. 22.2).

The Interrelationship between Edgeworth Boxes and Production Possibility Frontiers

Throughout this paper, I have taken the relationship between Edgeworth boxes and production possibility frontiers as common knowledge. However, this interplay may not be completely apparent or perhaps fuzzy, dependent upon one's particular microeconomic background. While intermediate and advanced microeconomic textbooks (e.g. Schotter 2008) often make this plain, I shall, nonetheless, briefly explain why the leap from PPF to Edgeworth box is both graphically eloquent and intuitively obvious for the sake of completeness in exposition. I emphasize that, for the purpose of brevity, my overview will be *dense* and *terse*. This is *not* the phrasing I would employ to a non-academic audience and *certainly* not within the context of an introductory class. A series of brief lesson modules for an introductory course exposition of Edgeworth boxes (both in and of themselves as well as their relationship to PPF's) is provided in Appendix A.

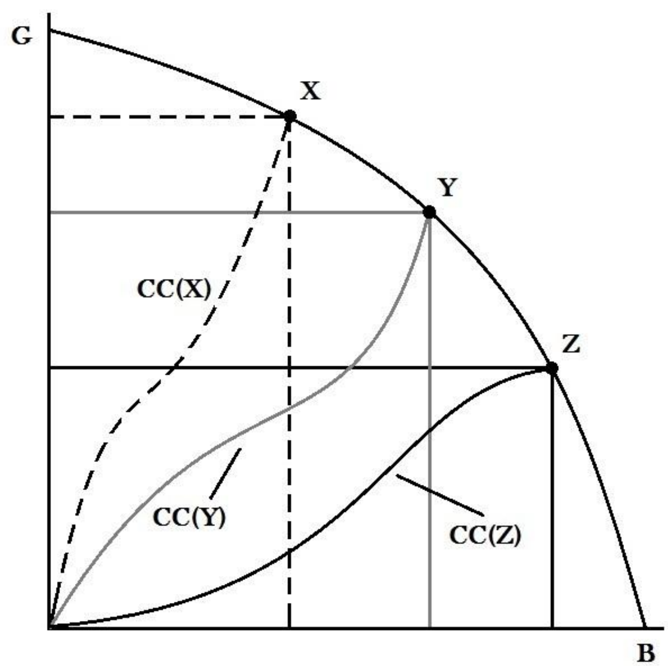


Figure 1a – Pareto Optimality in Goods Space

[9] See Roemer (1996) for a summary of formal Social Choice theory.

[10] Mankiw (2014, pp. 147-48) emphasizes such skepticism toward central planning in the allocation of production or distribution.

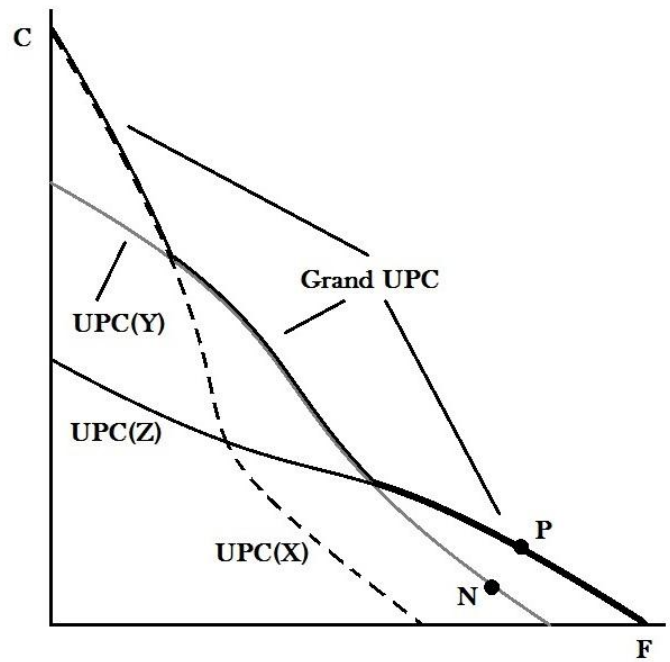


Figure 1b – Pareto Optimality in Utility Space

As illustrated in Figure 1a, the PPF is the potential output of an economy (two-person or otherwise) given input constraints. Thus, any particular point on the PPF is efficient in production. Having arrived at a particular point of production on the PPF, we have, as previously noted, no particular intuitive path to utility via consumption. However, any point of production on the PPF forms the total output constraints that characterize the endowment economy of an Edgeworth box scenario. Given that we have produced *b* units of butter and *g* units of guns, *b* and *g* are the boundaries of a good-for-good exchange scenario. Thus, any particular point on the PPF generates a unique Edgeworth box for a given set of resource constraints. As illustrated in Figure 1a, production point *X* generates the dashed Edgeworth box, point *Y* generates the solid gray box, and point *Z* generates the solid black box nested within the PPF.

The goods produced will subsequently be exchanged until the marginal rate of substitution (i.e. relative value) is equal across both persons. The PPF dictates efficiency strictly in production. The Edgeworth box, in contrast, depicts allocations of that particular (efficient) level of production; any point on the contract curve (CC), which spans each box horizontally from southwest to northeast corner, is efficient in allocation. The particular outcome of trading will be contingent on the initial endowments of the traders. Furthermore, each point of efficient allocation in a given Edgeworth box exists in guns-butter (output) space as in Figure 1a, but also generates specific levels of utility for Robinson Crusoe and Friday via their post-trade consumption bundles. We can, therefore, postulate a utility possibility curve (UPC) that translates each efficient allocation from the guns-butter goods space in Figure 1a to the Crusoe-Friday utility space depicted in Figure 1b. In this instance, the points on CC(*X*) correspond to the UPC(*X*), where the extreme southwestern point on the CC is one axis intercept in Figure 1b and the extreme northeastern point the other intercept. At each of these intercepts, one person has all of the output and, therefore, all of the utility. UPC's *Y* and *Z* are generated by the same translation of CC's *Y* and *Z*.

However, consider a point such as *N*, which lies on UPC(*Y*). Point *N* is clearly Pareto inferior to a point such as *P*, given that both *C* and *F* enjoy a greater level of utility at *P*. Thus, points on a UPC(*Y*) are only optimal *given* the production point *Y*. When considering output and exchange in tandem, there is room for Pareto improvement. Consequently, the departure point for questions of equity is based on simultaneous (Pareto) efficiency in production *and* exchange. Graphically, this is the outer envelope of the (infinite) collection of UPC's, called the grand utility possibility frontier. Points on this curve are characterized by a slope of the production possibility frontier (the marginal rate of transformation) which is equal to the slope of Crusoe and Friday's indifference curves at equilibrium within the Edgeworth box (the marginal rate of substitution).[11]

Conclusion

In previous sections, I have presented a brief case for the inclusion of Edgeworth boxes in introductory economics courses, particularly in conjunction with production possibility frontiers, which are currently both a staple economic concept and graphical "square one" in introductory classes. Given that the concept of preference relations is both intuitive and foundational to microeconomics, early introduction of the idea is both appropriate in terms of rigor and pedagogically advantageous for students who may later take intermediate and advanced classes. Moreover, the subsequent introduction to Edgeworth boxes (i) reinforces the win-win nature of free trade (in expectation) at the individual level, (ii) explicitly emphasizes efficient production for the sake of consumption and utility rather than production for its own sake and, conveniently but less importantly, (iii) serves as a valuable precursor to Social Choice Theory and political economy. To jump directly from "comparative advantage" and "terms of trade" to prices, markets, supply and demand is to tell only half of the economic story.

References

Baumol, William J., and Alan S. Blinder. 2011. *Microeconomics: Principles and Policy. 12th Edition.* Boston: Cengage Learning.

Buchanan, James M., and Gordon Tullock. 1962. *The Calculus of Consent*. Ann Arbor, MI: University of Michigan Press.

Edgeworth, Francis Y. 1881. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences.* London: C. K. Paul and Company.

Gwartney, James D., Richard L. Stroup, Russel S. Sobel, and David A. Macpherson. 2014. *Microeconomics: Private and Public Choice. 15th Edition*. Boston: Cengage Learning.

Hall, Robert E., and Marc Lieberman. 2012. *Economics: Principles and Applications. 6th Edition*. Boston: Cengage Learning.

Mankiw, N. Gregory. 2014. *Principles of Microeconomics. 7th Edition*. Mason, OH: South-Western College.

Ricardo, David. 1817. *On the Principles of Political Economy and Taxation*. London: John Murray.

Roemer, John E. 1996. *Theories of Distributive Justice*. Cambridge, MA: Harvard University Press.

Schotter, Andrew R. 2008. *Microeconomics: A Modern Approach*. Mason, OH: South-Western College.

APPENDIX A
Lesson Procedures for Teaching Edgeworth Boxes in an Introductory Course

Suggested Semester Sequencing

In my experience, introductory students are typically uncomfortable with graphs, particularly as a tool for the illustration of economic logic, at the outset of the semester. Consequently, I find it advantageous to

[11] In an introductory class, the realization that "global" Pareto optimality is matter of spontaneous coordination in both production and exchange is, I believe, much more important than the graphical details which support the idea. Likewise, the Edgeworth box analysis in general provides a useful illustration of Pareto improvement.

introduce graphing as simply as possible via the thorough analysis of the production possibility frontier and the inputs-to-outputs relationship that it depicts. The relatively straightforward depiction of a single line in the graphical space builds to comparative advantage, the gains from trade, and consumption possibility frontiers. A growing familiarity with the way in which graphs are used to model economic interaction thereby provides a much needed stepping stone to understanding the more sophisticated Edgeworth box diagram. It also flows chronologically in illustrating production, trade, and consumption for utility in the appropriate sequence. Following Mankiw's *Principles of Economics* ordering, this places the Edgeworth box material between Chapter 3 (Interdependence and the Gains from Trade) and Chapter 4 (The Market Forces of Supply and Demand).

Optional Introductory Activity for Smaller Classes (5 minutes)

Required Materials**:** One Fun Dip stick and powder packet for each student in the class, cut at the seam between the stick and packet.

Procedure**:** Ask the class (rhetorically) whether or not production is necessary in order for people to gain from trade. Define an endowment economy as a scenario in which potential traders are simply given a bundle of goods to exchange as they see fit. Further assert that you are going to create such an economy for the class. Give one half of the class *two* Fun Dip sticks each and the other half *two* Fun Dip packets each. Tell them that they are free to engage in trade with their classmates for the next five minutes along whatever terms of trade are mutually agreeable. The obvious exchange ratio of one stick for one powder packet will provide intuition about why traders in the more complex Edgeworth box scenario behave as they do. Accordingly, I suggest an occasional break from the "Crusoe and Friday trading beer and pizza" exposition to draw the appropriate parallels to their own motivations in trading endowments with their classmates.

Teaching Module 1 - Edgeworth Box Set-Up (45-60 minutes)

Lesson Procedures:

1. Begin with a Robison Crusoe scenario (two people in an island economy) to motivate a limited quantity of two goods (e.g. beer and pizza) available for trading. This allows for a graphical space as shown in Figure A1, wherein a two-dimensional graphical box illustrates any possible allocation of the good between the two traders. Make sure students understand what any given point, such as ω in Figure A1, indicates about the prevailing allocation of goods between Robinson Crusoe (RC) and Friday (F).

2. Motivate indifference curves via preference relations using hypothetical ordinal rankings of two-good bundles. Use the nature of complementary goods (beer and pizza) to plot several points on a single indifference curve. Start at extremes (e.g. lots of beer, very little pizza) and work inward, as illustrated in Figure A2, to utilize students' intuition about their own psychological trade-offs between bundles. You may want to ask students to draw indifference curves for close substitutes and/or compliments to test their understanding of how the shape of indifference curves is deduced.

3. Return to the preference relation material to illustrate a full set of indifference curves on a stand-alone graph. It is important for students to understand that we are merely plucking particular illustrative curves from an infinite series of parallel lines (resembling a quarter of an onion when halved width-wise) and that such curves exist independent of any economic interaction by traders. Point out that whatever can be graphically deduced of Crusoe's situation at the standard origin point can also be applied to Friday when rotated 180 degrees, since the northeast corner is the corresponding origin point for Friday. Thus, indifference curves are roughly symmetric diagonally (although indifference curves are obviously idiosyncratic and, thus, not symmetric *per se*). End with a fully labeled Edgeworth box containing two or three indifference curves for each trader. Make sure students understand which trader's indifference curves correspond with a particular origin point (where a trader has zero of both goods).

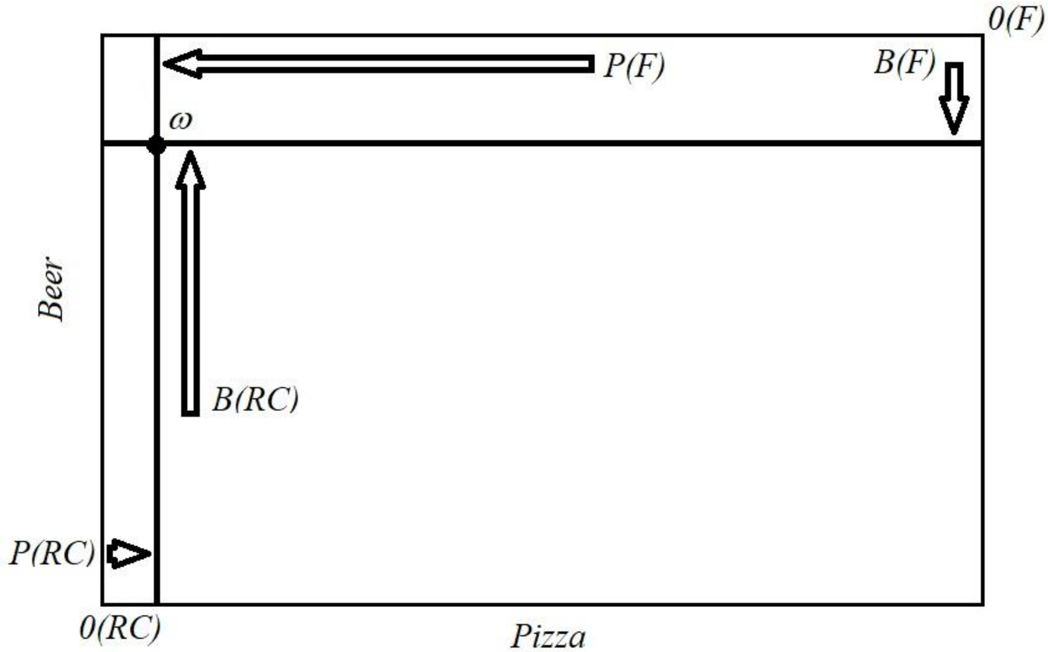


Figure A1 – Initial Allocation in an Edgeworth Box

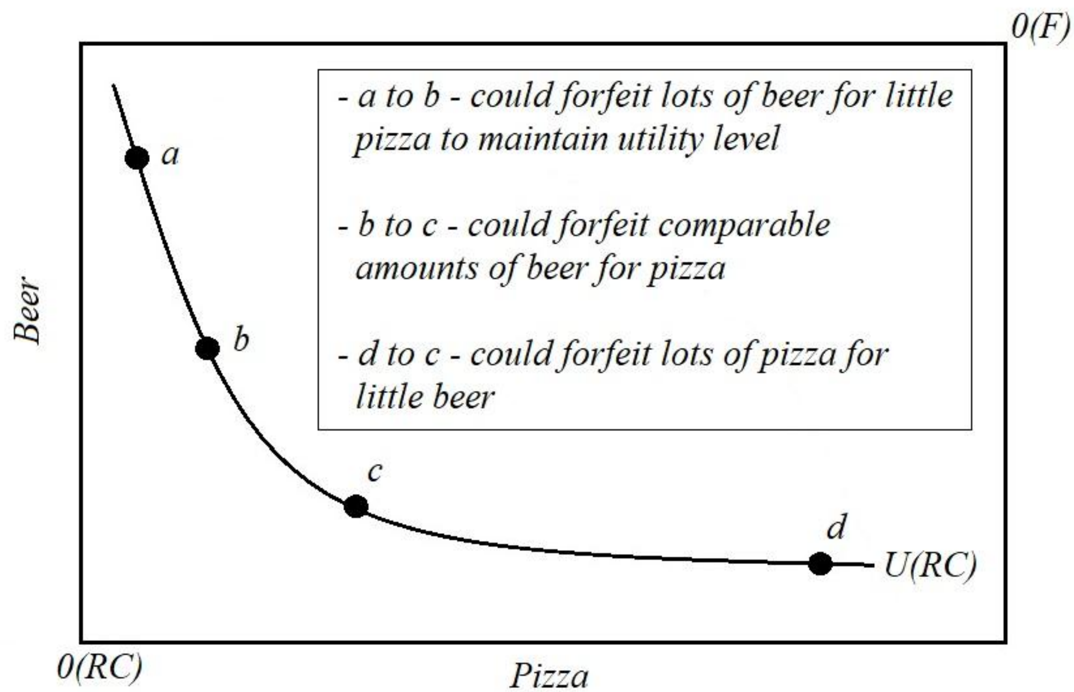


Figure A2 – Motivation for a Typical Indifference Curve

Teaching Module 2 – Exchange in an Edgeworth Box (45-60 minutes)

Lesson Procedures:

1. Stipulate an arbitrary endowment point well toward either the northwest or southeast corner of the Edgeworth box. It is, of course, most convenient to choose a point where two indifference curves drawn in the previous teaching module intersect, such as ω in Figure A3. Ask students to identify how goods are apportioned between traders at this particular point and emphasize the intuitive gains from trade when the asymmetry of relative values between the two goods is large (e.g. Crusoe has lots of pizza and is very thirsty, Friday has lots of beer and is very hungry).

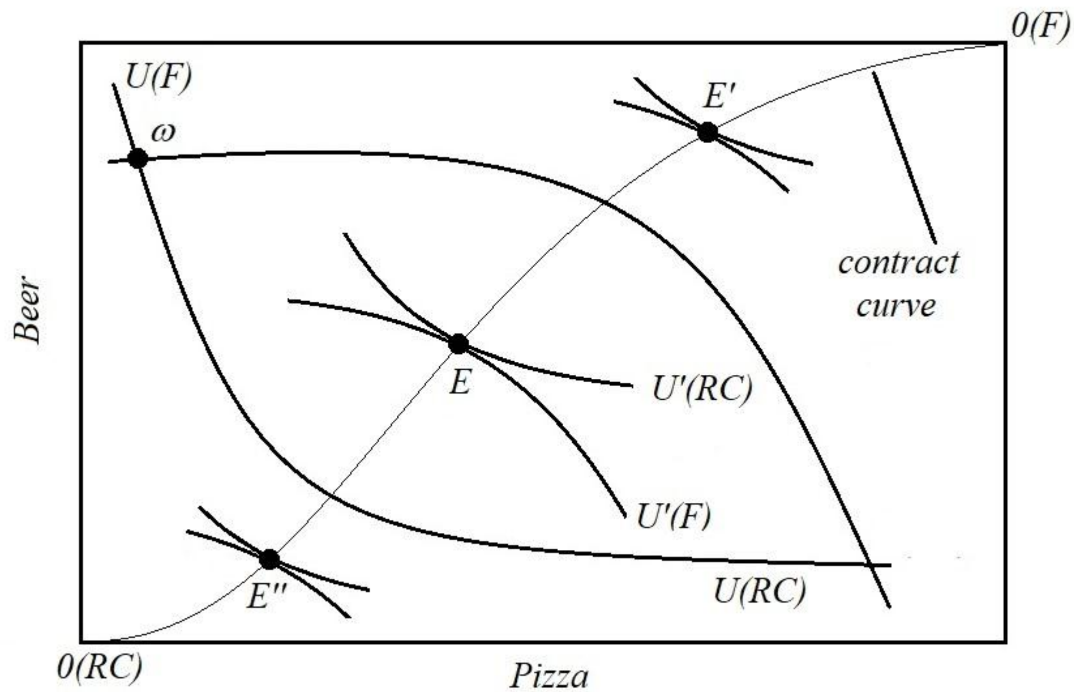


Figure A3 – Trade to Equilibrium and Alternative Equilibria

2. Introduce the concept of Pareto improvement (as well as inferiority, efficiency, etc.) by showing which bundles are superior and inferior to the endowment points for each trader. Ruling out bundles that will make either trader worse off (by coloring the inferior bundles in different shades for each person) leaves the elliptical core of potential post-trade allocations.
3. Stipulate an arbitrary trade. If applicable, point out the correspondence between differing relative values here and differing opportunity costs in a PPF-CPF trade scenario. Emphasize that the terms of trade, as in a PPF-CPF scenario, are based on relative negotiating skill and are, in no way, dictated by the parameters of the situation. However, because trade must make both traders better off, any trade must lie between relative values as opportunity costs of consumption comparable to those of production.
4. With one or two trades, illustrate how successive trading diminishes the core and brings traders' relative values closer together. Ask students why or when trading would stop. Emphasize how the logic of the situation (equal relative values across traders) relates to the graph (a core with an area of zero). Define Pareto efficiency in relation to the equilibrium point, labeled *E* in Figure A3.
5. Briefly illustrate how alternative endowment points generate alternative (Pareto efficient) equilibria, such as *E'* and *E"* in Figure A3. Ask students whether the southwest and northeast corners of the box are Pareto efficient. Using these corners and the equilibria previously illustrated, draw and explain a contract curve, as shown in Figure A3. End with a restatement of the principle that trade, even without

production, is necessarily mutually beneficial (in expectation) and that economists are ultimately concerned with individual "utility" (and not with goods in and of themselves).

Teaching Module 3 – Relating Edgeworth Boxes and PPF's (15-25 minutes)

Lesson Procedures:

1. I suggest breaking down market activity into: (i) production, (ii) exchange and (iii) consumption. Explain to students that PPF's describe step (i), while Edgeworth boxes describe steps (ii) and (iii). Without expectations of a correct response, ask students to consider how these two diagrams might be related.
2. Draw a PPF and choose a particular point of efficient production. Draw lines to the x and y – axis from the arbitrary production point and ask what you have effectively drawn (answer: an Edgeworth box).
3. Choose several points on the PPF and illustrate how each forms an Edgeworth box with a corresponding contract curve, as previously depicted (in the body of the paper) by Figure 1a. Point out to students that in a production economy with specialization of inputs, you are likely to begin trading very close to either the northwest or southeast corner of the box (where Crusoe has only produced one good and Friday only the other), which creates great possibilities for both traders to gain via market exchange. In closing, remind students of the purpose of the diagram(s): to illustrate the foundational principle that voluntary exchange *must* be jointly beneficial (in expectation) and also that markets effectively and efficiently coordinate complex economic interactions without central planning or direction.
4. **Optional:** If you think students are capable of comprehending more sophisticated concepts, you can introduce the distinction between Pareto efficiency in exchange and Pareto efficiency in both production and exchange. Show that a point on a contract curve might have a very different slope than the point on the PPF that formed the Edgeworth box (that contains that particular contract curve). Consequently, the relative values of traders (at equilibrium) are not reflected in the opportunity costs of the production of goods. Thus, to have "global" (as opposed to "local") Pareto efficiency, the trade-off in production must be equal to the trade-off in consumption.

APPENDIX B
Sample Exam Questions for Edgeworth Box Material

1. Which two goods would likely have the flattest (least bowed) indifference curves in an Edgeworth box diagram?
   a. Cheese and crackers
   b. Coffee and tea
   c. Both (a) and (b) will likely have the same shape for a typical person

2. Trade in an endowment economy requires …
   a. Different relative values of goods
   b. Different opportunity costs of production
   c. Equal division of the gains from trade

3. Efficiency in production is desirable because …
   a. Maximizing output is an end in and of itself
   b. It ultimately leads to greater levels of utility for traders
   c. It rules out trades which are not mutually beneficial

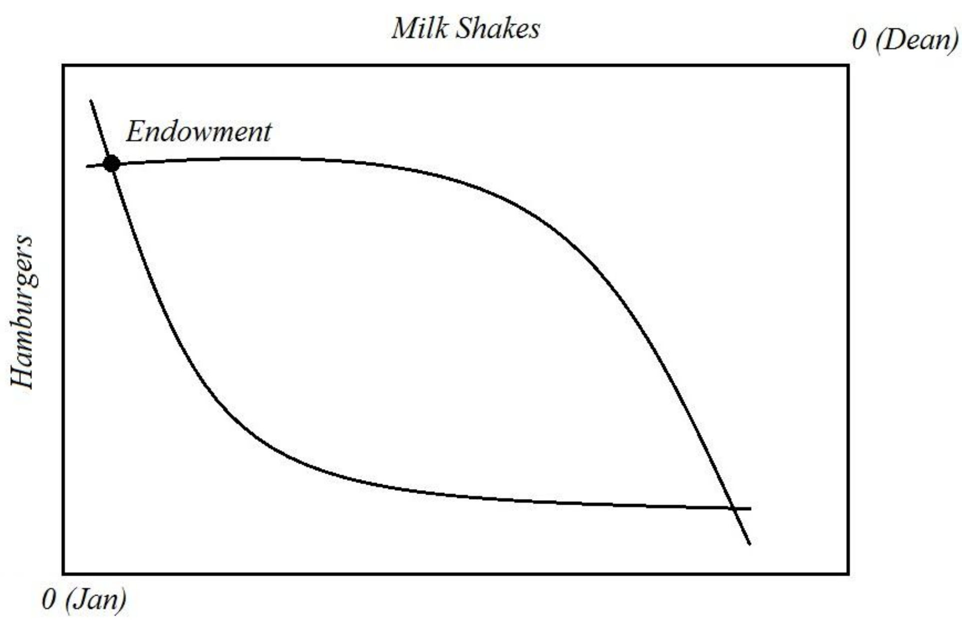
4. The endowment point in an Edgeworth box can **never** be …
   a. Pareto optimal
   b. At the northeast or southwest corner of the box
   c. Neither (a) nor (b)

5. A **single**, mutually beneficial trade in an Edgeworth box diagram will **necessarily** …
   a. Shrink the area of possible future trading
   b. Create a point of Pareto optimality
   c. Evenly distribute the benefits of trade

6. Do equilibrium points in an Edgeworth box diagram have to be Pareto optimal?
   a. Yes, otherwise further trading would occur
   b. No, because the endowment point, where trade begins, will never be optimal.
   c. Sometimes, but only when both traders are equally good at negotiation.

7. How do Edgeworth boxes relate to PPF's?
   a. Each potential equilibrium in an Edgeworth box corresponds to a separate PPF.
   b. Each point on the PPF creates an Edgeworth box.
   c. They don't relate because PPFs show production and Edgeworth boxes show trade.

8. Graphing Problem



On the graph above:
(a) Clearly label Jan's indifference curve *U(J,W)* and Dean's indifference curve *U(D,W)* at the endowment point.
(b) At the endowment point, whose value of milk shakes to hamburgers is relatively high? (**Jan** / Dean / they are equal) Whose value is relatively low? (Jan / **Dean** / they are equal)
(c) Draw two more indifference curves depicting a potential equilibrium and label them U(D,E) and U(J,E). Label the equilibrium point E.
(d) At the equilibrium point, whose value of milk shakes to hamburgers is relatively high? (Jan / Dean / **they are equal**) Whose value is relatively low? (Jan / Dean / **they are equal**)

Interest Rate Swap Valuation Since the Financial Crisis: Theory and Practice

Ira G. Kawaller and Donald J. Smith[1]

ABSTRACT

The financial crisis of 2007-09 revealed the importance of counterparty credit risk in the valuation of non-collateralized interest rate swaps. In theory, these valuations rest on assumed default probabilities and recovery rates. These assumptions, however, should be reflected in the risk-adjusted discount rates of the counterparties. Thus, in practice, swap valuations can be generated by discounting prospective swap settlements using risk-adjusted discount rates, cash flow by cash flow. This article demonstrates this method, discerning risk-adjusted discount rates from data that are readily available on the Bloomberg information system. Critically, if the inputs for the two methodologies are mutually consistent, theory and practice should yield identical valuations.

Introduction

The financial crisis of 2007-09 marked a fundamental change in methodologies used for valuing LIBOR-based interest rate swaps. Prior to the crisis, both dealers and end-users used a *single curve* methodology to determine the fair value for these derivatives. The exercise required projecting future cash flows based on a LIBOR forward curve and then discounting these projected cash flows using discount rates from that same LIBOR-based yield curve. With the bankruptcy of Lehman brothers, however, it became widely appreciated that this single curve methodology failed to address the issue of counterparty credit risk appropriately for these contracts. Moreover, accounting rules specifically initiated the requirement that derivatives be carried on the balance sheet at fair values that reflect their credit quality; but this requirement was generally not satisfied by the single curve methodology.

Since then, practitioners have largely coalesced around two different methods to value interest rate swaps. Both methods start with a LIBOR forward curve, but after that common starting step, the two methods take different approaches with respect to addressing credit risk. The more intuitive approach uses the standard discounted cash flow (DCF) analysis, applying discount rates that reflect the risk premiums applicable to the owing party, cash flow by cash flow. Thus, this “risk-adjusted DCF” approach uses up to three yield curves: the LIBOR yield curve (to generate the expected cash flows) and the yield curves pertaining to each of the parties to the contract (to derive discount factors for each of the parties with an obligation to pay).

The alternative methodology employs a two-step process for valuing interest rate swaps. We call this approach the *ex-post adjustment* (EPA) method. Under this approach, we first value the derivative without considering credit quality. That is, we discount projected cash flows using discount factors derived from overnight indexed swaps (OIS) curve[2] to generate a starting swap valuation that assumes a probability of default that approaches zero. These discount rates reflect spot and forward overnight fed funds rates, and they are typically taken to be proxies for risk-free interest rates. We then modify this initial valuation by making further adjustments that relate to credit considerations.

When applying the EPA method to a generalized swap where cash flows move in both directions (i.e., some inflows and some outflows), a credit valuation adjustment (CVA) would be made relating to cash flows expected to be received and a distinct debit valuation adjustment (DVA) would be made for cash flows

[1] Ira G. Kawaller is the founder of Derivatives Litigation Services, igkawaller@gmail.com. Donald J. Smith is Associate Professor of Finance, Questrom School of Business, Boston University, donsmith@bu.edu.

[2] OIS discounting has become standard practice to get a value for a swap assuming no default. See Hull and White (2013) and Smith (2013).

expected to be paid. Thus, the CVA for one party would be a DVA for the other, and vice versa. Given the reciprocal nature of CVAs and DVAs, it's not uncommon for many market participants to ignore the distinction and simply refer to both adjustments as CVA adjustments. In any case, as both adjustments serve to shrink the values of their associated risk-free cash flows, whether valuing the swap from the perspective of the asset position or the liability position, the same *absolute value* would result for both parties. A positive value would reflect an asset position, while a negative value would reflect a liability.

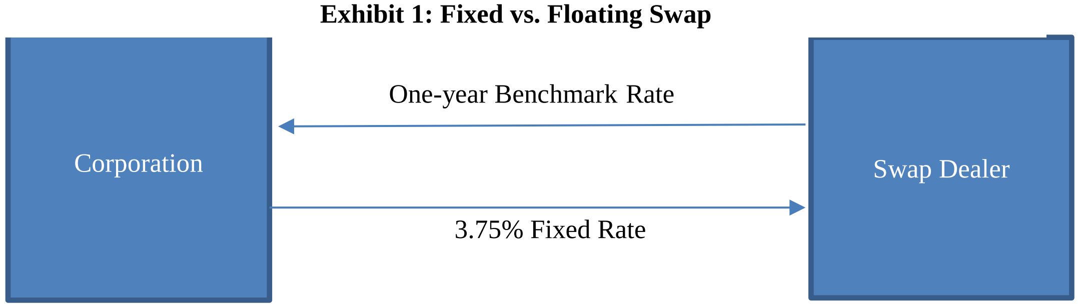
It should be clear that if we start with a valuation generated by the risk-adjusted DCF method, we could solve for respective default probabilities and recovery assumptions such that the DCF valuation could be replicated using the EPA method. Alternatively, starting with an EPA valuation, for any given CVA and DVA, we can solve for a unique credit spread for the owing party, such that if we had a consistent data set, the two valuations, again, would be identical.[3]

The first section of this paper presents the current state of valuation theory where we pose a numerical example to demonstrate the way forward interest rates, discount factors, and swap market values may be determined. Then, in the second section, we introduce practical considerations to demonstrate the way in which analysts may employ readily accessible data to generate their desired valuations.

Interest Rate Swap Valuation in Theory

Our numerical example assumes a five-year swap with a fixed rate of 3.75%, where resets and settlements based on an assumed one-year accrual period and benchmark rate. We assume the counterparties are a corporate end-user (fixed payer) and a swap dealer (fixed receiver). The annual accrual period is simply a pedagogic expedient to simplify our math by eliminating the need for showing fractional time periods in our equations. The swap is pictured in Exhibit 1 in the familiar box-and-arrow format.

Exhibit 1: Fixed vs. Floating Swap



Settlements in this swap are determined on a net basis in arrears. That means that if the one-year benchmark rate were 2.00% on any reset date (i.e., at the start of the accrual period), the corporation would pay the dealer 1.75% of the swap's notional principal at the end of the year: (3.75% - 2.00%) = 1.75%. If the one-year rate were 5.00%, the cash flow would move in the opposite direction. That is, the dealer would pay the corporation 1.25% of the notional in arrears: (5.00% - 3.75%) = 1.25%.

A key step in this swap valuation exercise is to project the forward rates for the floating reference rate. These forward rates would be observable if there were an actively traded futures market for one-year benchmark rates. Otherwise, forward rates would likely be inferred from observed data on one-year through five-year annual payment benchmark bonds. The procedure to get these rates is called *bootstrapping*. Exhibit 2 displays the assumed coupon rates and prices on these benchmark bonds.

[3] In recent years, some money-center banks have begun to include other risks and costs to their derivatives business in addition to the CVA and DVA. The term "XVA" has emerged to summarize these valuation adjustments. These include the FVA (funding valuation adjustment), KVA (capital valuation adjustment), LVA (liquidity valuation adjustment), and MVA (initial margin valuation adjustment). See Gregory (2015), Green (2016), Ruiz (2015), Lu (2016), and Smith (2017) for detailed discussion of these adjustments, as well as a few more.

Exhibit 2: Market Data on Benchmark Bonds (Par = 100)

Projected Settlements	Coupon Rate	Price
1	0.00%	99.75
2	0.25%	99.25
3	1.50%	100.125
4	1.75%	98.25
5	2.75%	100.25

We first derive benchmark discount factors that are consistent with the principle of no arbitrage. These discount factors, denoted DF_1 through DF_5, are obtained iteratively by applying the principle that the observed bond price of a zero-coupon bond equals its future value times its associated present value factor. Thus, we solve for the one-year discount factor, DF_1, as follows:

$99.75 = 100 * DF_1$; $\quad DF_1 = 0.997500$

The hallmark of a bootstrapping procedure is that the output from one step is used as an input in the next step. That is, in the next period (reflecting the two-year bond), we discount the coupon paid at the end of year 1 using DF_1, and then at the second year, the second (and final) coupon payment plus par is discounted by the second discount factor, DF_2. [Note: all calculations are done on a spreadsheet and rounded results are reported in the text].

$99.25 = (0.25 * 0.997500) + (0.25 + 100) * DF_2$; $\quad DF_2 = 0.987537$

Continuing iteratively for each bond, we determine the remaining discount factors as follows:

$100.125 = (1.50 * 0.997500) + (1.50 * 0.987537) + (101.50 * DF_3)$; $\quad DF_3 = 0.957118$

$98.25 = (1.75 * 0.997500) + (1.75 * 0.987537) + (1.75 * 0.957118) + (101.75 * DF_4)$;
$DF_4 = 0.915000$

$100.25 = (2.75 * 0.997500) + (2.75 * 0.987537) + (2.75 * 0.957118) +$
$(2.75 * 0.915000) + (102.75 * DF_5)$; $\quad DF_5 = 0.872436$

Given discount factors for each date, we are now able to derive associated one-year forward rates, denoted as $R_{1,2}$, $R_{2,3}$, $R_{3,4}$, and $R_{4,5}$. For example, $R_{4,5}$ is the one-year rate between dates 4 and 5 that is consistent with no arbitrage opportunities in the benchmark bond market. The spot one-year rate on the yield curve is simply the yield to maturity on the one-year zero-coupon bond, determined as follows:

$P_1*(1+R_{0,1}) = 100$ implies $R_{0,1} = (100/P_1) – 1$ or $R_{0,1}= (100/99.75) – 1 = 0.2506\%$

where P_1 is the spot price for the one-year bond, and $R_{0,1}$ is its yield to maturity. The one-year forward rates that follow that spot rate are calculated from the following progression of equations:

$DF_2*(1+R_{0,1})*(1+R_{1,2})=1$

$DF_3*(1+R_{0,1})*(1+R_{1,2})*(1+R_{2,3})=1$

$DF_4*(1+R_{0,1})*(1+R_{1,2})*(1+R_{2,3})*(1+R_{3,4})=1$

$DF_5*(1+R_{0,1})*(1+R_{1,2})*(1+R_{2,3})*(1+R_{3,4})*(1+R_{4,5})=1$

Recall, however, that $(1+R_{0,1}) = 1/DF_1$; and thus:

$R_{1,2} = DF_1/DF_2 - 1 = 0.997500/0.987537 - 1 = 0.010088$ $\quad R_{1,2} = 1.0088\%$

$R_{2,3} = DF_2/DF_3 - 1 = 0.987537/0.957118 - 1 = 0.031783$ $\quad R_{2,3} = 3.1783\%$

$R_{3,4} = DF_3/DF_4 - 1 = 0.957118/0.915000 - 1 = 0.046030$ $\quad R_{3,4} = 4.6030\%$

$R_{4,5} = DF_4/DF_5 - 1 = 0.915000/0.872436 - 1 = 0.048787$ $\quad R_{4,5} = 4.8787\%$

Exhibit 3 displays the projected cash flows for the swap assuming a notional principal of 100. Values shown in columns 2 through 4 reflect amounts from the perspective of the corporate entity – i.e., the fixed-rate payer. Positive values reflect cash inflows to the corporate; negative values reflect cash outflows. The projected variable payments reflect the one-year spot and subsequent forward rates derived above, times the notional principal (= 100). The contractual fixed payments from the corporation to the dealer reflect the 3.75% fixed rate, also times the notional principal. Finally, the fourth column shows the projected *net* settlements (i.e., the sum columns 2 and 3).

Exhibit 3: Cash Flows for the Swap (Notional Principal = 100)

Year	Variable Settlements	Contractual Fixed Settlements	Projected Net Settlements
1	0.2506	-3.7500	-3.4994
2	1.0088	-3.7500	-2.7412
3	3.1783	-3.7500	-0.5717
4	4.6030	-3.7500	0.8530
5	4.8787	-3.7500	1.1287

Because the corporation and the swap dealer are not default-free entities, we need to discount the projected net settlement payments using *risk-adjusted discount factors*. These discount factors are found by performing another bootstrapping procedure that incorporates assumptions about the credit risk parameters for the two counterparties. The corporation is assumed to have a 1.50% probability of default for each year (known as the *hazard rate*) – *conditional* on not defaulting in a prior period. If a default were to occur, however, some portion of the exposure would likely be recovered. We assume a recovery rate of 40% for the corporate entity. The swap dealer, on the other hand, might reasonably be expected to have a lower annual probability of default and a lower recovery rate. We arbitrarily assume values of 0.50% and 10%, respectively, for these inputs. In practice, the credit risk parameters can be inferred from the prices of credit default swaps or received directly from a credit rating agency or consultancy.[4]

For illustrative purposes, we solve for the risk-adjusted price of a five-year zero-coupon corporate bond in Exhibit 4. While we describe this process for only one maturity, this same procedure needs to be repeated to solve for the risk-adjusted prices for zero-coupon bonds with maturities associated with all the settlements throughout the horizon of the swap under consideration – for both the corporate entity and the swap dealer.

[4] Hull (2014) provides an example of valuing a credit default swap (CDS) given assumptions about the conditional probability of default and the recovery rate. Given the observed price for the CDS, the example could be revised to back out the probability of default consistent with the assumption for recovery. For an example of a consultancy that provides probabilities of default, see Kamakura Corporation at kamakuraco.com. See Jarrow (2012) for caveats about simple use of credit default swaps to get the default probabilities.

Exhibit 4: CVA Computation for a Five-Year Zero-Coupon Corporate Bond

Year	Spot and Forward Rates	Exposure	LGD	POD	DF	CVA
1	0.2506%	87.4623	52.4774	1.5000%	0.997500	0.7852
2	1.0088%	88.3446	53.0068	1.4775%	0.987537	0.7734
3	3.1783%	91.1525	54.6915	1.4553%	0.957118	0.7618
4	4.6030%	95.3482	57.2089	1.4335%	0.915000	0.7504
5	4.8787%	100.0000	60.0000	1.4120%	0.872436	0.7391
				7.2783%		**3.8099**

To calculate the CVA adjustment, we use the previously derived benchmark spot and forward rates to determine the default loss exposure for each year. We start with the most distant settlement date in column 3 and work up the table, year by year. If the corporation were to default at the end of the fifth year, the investor could lose the entire par value (100). However, if the default were to occur at the end of the fourth year, the exposure would be 95.3482, which is 100 discounted back for one year using the benchmark forward rate of 4.8787% ($R_{4,5}$). That is, 95.3482 = 100/1.048787. For default at the end of year 3, the exposure is 91.1525 = 100/(1.048787 * 1.046030), etc.

Given the year-by-year exposures, our loss given default (LGD) value simply reflects the expected loss that will ultimately be realized with a default. In our example, we assume a recovery rate of 40%, such that the LGD values in column 4 are the exposures for default loss for each year, times (1 – 0.40). This amount, (1 – 0.40), is known as the *loss severity*.

Next, we calculate probabilities of default for year n (POD_n) in column 5, as follows:

$$POD_n = \left(1 - \sum_{i=0}^{n-1} POD_i\right) * 0.0150\%$$

This equation is used, iteratively, year by year. Again, by assumption, the first year's POD is 1.50% such that the probability of survival is 98.50%. In general, the survival probability for each year is the parenthetical term in the above equation. The POD for the second year, conditional on no prior default, is 1.50% * (100 - 1.50)% = 1.50% * 98.50% = 1.4775%. Analogous calculations of the above equation are required to calculate the PODs for all remaining accrual periods. Given these respective PODs, the *cumulative probability of default* over the lifetime of the bond is their sum – in this example, 7.2783%. Finally, we reproduce the previously calculated benchmark discount factors in column 6 of Exhibit 4, enabling the calculation of the annual contributions to CVA by multiplying the LGD times the POD times the DF in each year. These products are shown in the far-right column of Exhibit 4. The CVA for the five-year zero-coupon corporate bond is the sum of this column, 3.8099.

This same methodology is required to calculate cumulative CVAs for the complete set of zero-coupon corporate bonds having maturities ranging from the first accrual period through the last, which, in this case, means one year to five years. These respective CVAs are shown in Exhibit 5. We subtract the corporate CVAs from the corresponding benchmark zero-coupon bond prices, which are just the benchmark discount factors multiplied by 100. From the zero-coupon corporate bond prices we obtain the discount factors that apply to the fixed-rate payer on the interest rate swap. For example, the five-year benchmark zero-coupon bond price is 87.2436. That less the CVA of 3.8099 gives a five-year zero-coupon bond price of 83.4337 (= 87.2436 – 3.8099) and a discount factor of 0.834337**.** Analogous content is shown in Exhibit 6 relating to the calculations of the discount factors relating to the credit risk of the swap dealer, the fixed-rate receiver.[5]

[5] Note that the back-up calculations shown in the text pertain exclusively to the determination of the corporate discount factor for the discrete five-year horizon. Analogous processing (not reported) is required for all settlement date horizons.

Exhibit 5: CVA Computation for the Corporate Entity (Five years; Par = 100)

Year	Benchmark DF	Benchmark Bond Price	Corporate CVA	Corporate Bond Price	Corporate DF
1	0.997500	99.7500	0.8978	98.8522	0.9885
2	0.987537	98.7537	1.7642	96.9895	0.9699
3	0.957118	95.7118	2.5456	93.1662	0.9317
4	0.915000	91.5000	3.2206	88.2794	0.8828
5	0.872436	87.2436	3.8099	83.4337	0.8343

Exhibit 6: CVA Computation for the Dealer (Five years; Par = 100)

Year	Benchmark DF	Benchmark Bond Price	Dealer CVA	Dealer Bond Price	Dealer DF
1	0.997500	99.7500	0.4489	99.3011	0.9930
2	0.987537	98.7537	0.8866	97.8671	0.9787
3	0.957118	95.7118	1.2857	94.4261	0.9443
4	0.915000	91.5000	1.6347	89.8653	0.8987
5	0.872436	87.2436	1.9434	85.3002	0.8530

Given these discount factors, the most direct method to incorporate credit considerations into the valuation of an interest rate swap is to identify the counterparty projected to make the net settlement payment and then use the discount factor corresponding to that date and entity. We show this in Exhibit 7, which displays the projected net settlements from the perspective of the corporate fixed-payer (Exhibit 3), the discount factors that apply to the corporation (Exhibit 5), and the discount factors for the swap dealer (Exhibit 6). The first three net settlements are negative—they are liabilities of the corporation and are discounted using the corporate discount factors. The last two are positive—they are assets to the corporation and are discounted using the swap dealer's discount factors. The net result is a swap value of -4.9212 to the corporation and + 4.9212 to the swap dealer.

$$(-3.4994 * 0.988522) + (-2.7412 * 0.969895) + (-0.5717 * 0.931662)$$

$$+ (0.8530 * 0.898653) + (1.1287 * 0.853002) = -4.9212$$

Exhibit 7: Risk Adjusted Swap Valuation (Corporate Perspective)

Year	Projected Net Settlements	Corporate DF	Dealer DF	Swap PV
1	-3.4994	0.988522	0.993011	-3.4592
2	-2.7412	0.969895	0.978671	-2.6587
3	-0.5717	0.931662	0.944261	-0.5326
4	0.8530	0.882794	0.898653	0.7666
5	1.1287	0.844337	0.853002	0.9628
				-4.9212

Interest Rate Swap Valuation in Practice

To value a swap using the credit-adjusted discounted cash flow method, we simply apply the standard theoretical valuation model using discount factors that reflect the credit quality of the owing counterparty, cash flow by cash flow. Our discussion uses the Swap Manager function on Bloomberg to value an interest rate swap on 6/30/16. The swap in question has an end date of 12/31/16, a fixed coupon of 1.4875%, and a variable leg tied to one-month LIBOR. Settlements are netted monthly, reflecting the actual/360-day count convention. These features are captured in Exhibit 8 and Exhibit 9, reflecting the valuations first from the perspective of the *fixed payer* and then from the perspective of the *fixed receiver*. Note that the same underlying curve selections and conventions apply to both screens, such that both show NPVs of the same absolute value. The positive value reflects the swap being an asset for the fixed rate payer; the negative value reflects the swap being a liability for the fixed rate receiver.

The next three Bloomberg screen shots show the cash flow calculations from the perspective of the fixed-rate payer. Exhibit 10 shows the fixed cash flows; Exhibit 11 shows the variable cash flows; and Exhibit 12 puts them together to show the net cash flows, along with their discounted values.[6] The variable and fixed settlements shown in these exhibits are found by multiplying the notional amount, times the respective interest rates, times time. In our example, we assume the actual/360 day-count method for calculating time, but some swaps may apply alternative day-count conventions.[7] Present value amounts are simply the product of payment amounts times their associated discount factors.

[6] In each of these exhibits, when the last day of the month happens to be a weekend day, which occurs in July and December), the settlement date and accrual period end reverts by convention to the last business day of that month.

[7] Note that some confusion may arise in terms of terminology, as ISDA swap definitions identify the *reset date* as the first date in which the new variable rate will be applied, while Bloomberg calls the reset date the date at which the new variable rate is determined.

Exhibit 8: Main Screen - Fixed Rate Payer

Exhibit 9: Main Screen - Fixed Rate Receiver

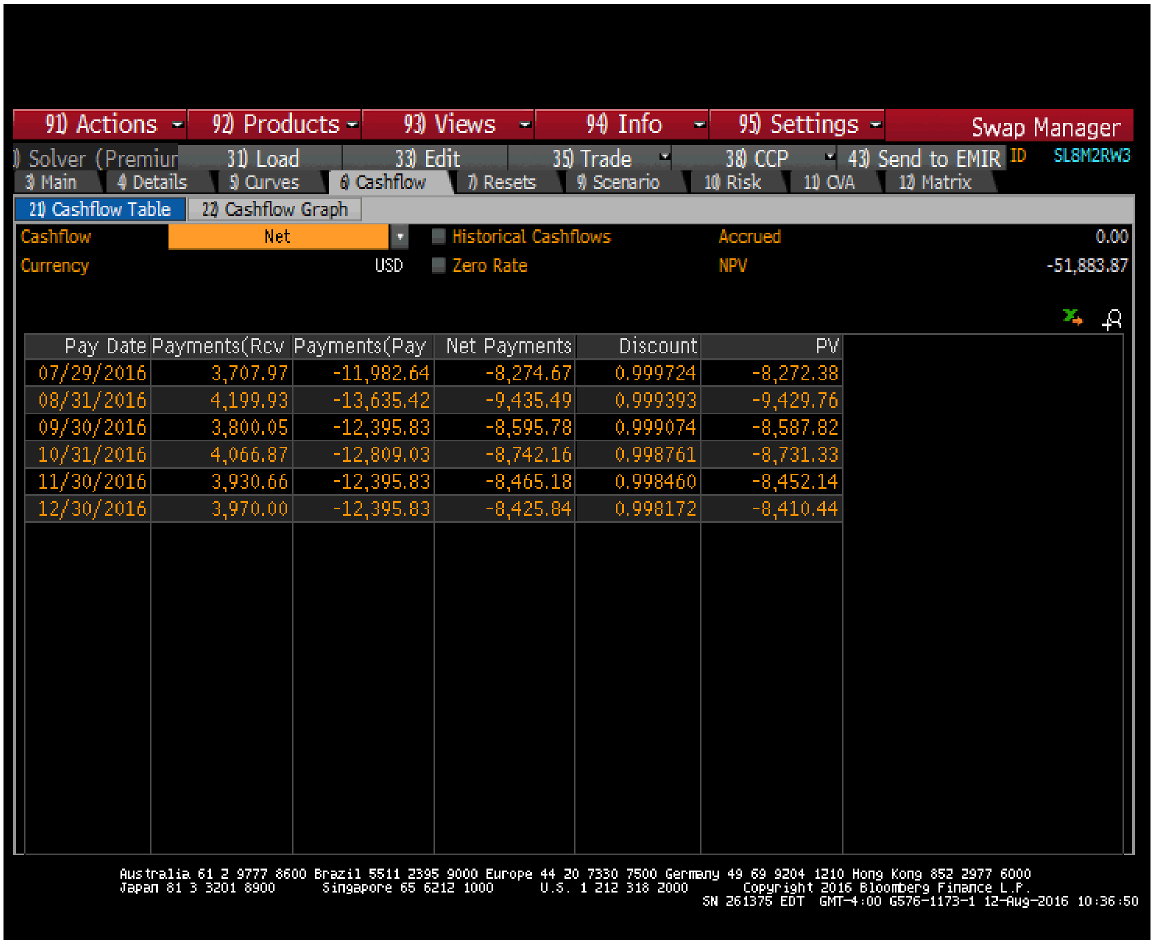
Exhibit 10: Fixed Cash Flows

Pay Date	Accrual Start	Accrual End	Days	Notional	Principal	Payment	Discount	PV
07/29/2016	06/30/2016	07/29/2016	29	-10,000,000.00	0.00	-11,982.64	0.999724	-11,979.33
08/31/2016	07/29/2016	08/31/2016	33	-10,000,000.00	0.00	-13,635.42	0.999393	-13,627.14
09/30/2016	08/31/2016	09/30/2016	30	-10,000,000.00	0.00	-12,395.83	0.999074	-12,384.35
10/31/2016	09/30/2016	10/31/2016	31	-10,000,000.00	0.00	-12,809.03	0.998761	-12,793.15
11/30/2016	10/31/2016	11/30/2016	30	-10,000,000.00	0.00	-12,395.83	0.998460	-12,376.75
12/30/2016	11/30/2016	12/30/2016	30	-10,000,000.00	0.00	-12,395.83	0.998172	-12,373.18

Exhibit 11: Variable Cash Flows

Pay Date	Accrual Start	Accrual End	Days	Notional	Principal	Reset Date	Reset Rate	Payment
07/29/2016	06/30/2016	07/29/2016	29	10,000,000.00	0.00	06/28/2016	0.46030	3,707.97
08/31/2016	07/29/2016	08/31/2016	33	10,000,000.00	0.00	07/27/2016	0.45817	4,199.93
09/30/2016	08/31/2016	09/30/2016	30	10,000,000.00	0.00	08/26/2016	0.45601	3,800.05
10/31/2016	09/30/2016	10/31/2016	31	10,000,000.00	0.00	09/28/2016	0.47228	4,066.87
11/30/2016	10/31/2016	11/30/2016	30	10,000,000.00	0.00	10/27/2016	0.47168	3,930.66
12/30/2016	11/30/2016	12/30/2016	30	10,000,000.00	0.00	11/28/2016	0.47640	3,970.00

Exhibit 12: Net Cash Flows



Pay Date	Payments(Rcv	Payments(Pay	Net Payments	Discount	PV
07/29/2016	3,707.97	-11,982.64	-8,274.67	0.999724	-8,272.38
08/31/2016	4,199.93	-13,635.42	-9,435.49	0.999393	-9,429.76
09/30/2016	3,800.05	-12,395.83	-8,595.78	0.999074	-8,587.82
10/31/2016	4,066.87	-12,809.03	-8,742.16	0.998761	-8,731.33
11/30/2016	3,930.66	-12,395.83	-8,465.18	0.998460	-8,452.14
12/30/2016	3,970.00	-12,395.83	-8,425.84	0.998172	-8,410.44

With this foundation, we now must perform the following steps:

1. Infer the risk-free OIS interest rates that underlie the discount factors.
2. Identify the yield curve of any owning party of the swap. (Only one yield curve would apply if all the cash flow obligations are in one direction. Two yield curves would be required if both parties expect to have a net payment obligation.)
3. Determine the spread between the risk-free rate and the rate pertaining to the owing party (i.e., the risk premium) at the maturity corresponding to the time remaining until the swap's end date.
4. To find the risk-adjusted discount rate for each cash flow, add this risk premium of the owing party to that cash flow's risk-free discount rate.
5. Given the set of risk-adjusted discount rates, calculate the discount factors for each of the component net settlements.
6. Find the risk-adjusted PV by multiplying each anticipated net settlement by its respective risk-adjusted discount factor, and sum the elements.

We discuss each step, in turn.

Given that the yield curves that we will shortly be referencing reflect interest rates that employ the semi-annual compounding convention, we derive the underlying discount rate from the following formula:

$$DF = \frac{1}{\left(1 + \frac{Rate}{2}\right)^{\left(days/\left(\frac{365}{2}\right)\right)}} \tag{1}$$

Rate = The implied risk-free discount rate, annual yield, semiannual compounding
DF = The risk-free discount factor
days = The total number of days between the valuation date and the settlement date.

In equation (1), *Rate*/2 is the rate per semiannual period, a time frame that has 182.5 (= 365/2) days.[8] The denominator is the future value per unit invested for the fraction of the semiannual period. Therefore, the exponent to (1 + *Rate*/2) is days/182.5.

Equation (2) rearranges (1) to isolate *Rate*:

$$Rate = \left(\left(\frac{1}{DF}\right)^{\left(\left(\frac{365}{2}\right)/days\right)} - 1\right) * 2 \tag{2}$$

In the general case where we cannot expect to find actively traded credit derivatives that would provide risk premiums for the required maturities, we're forced to rely on aggregated data to extract reasonable estimates for the required risk premia. Various data providers publish such information. For instance, Bloomberg provides curves for assorted credit ratings in the following sectors:

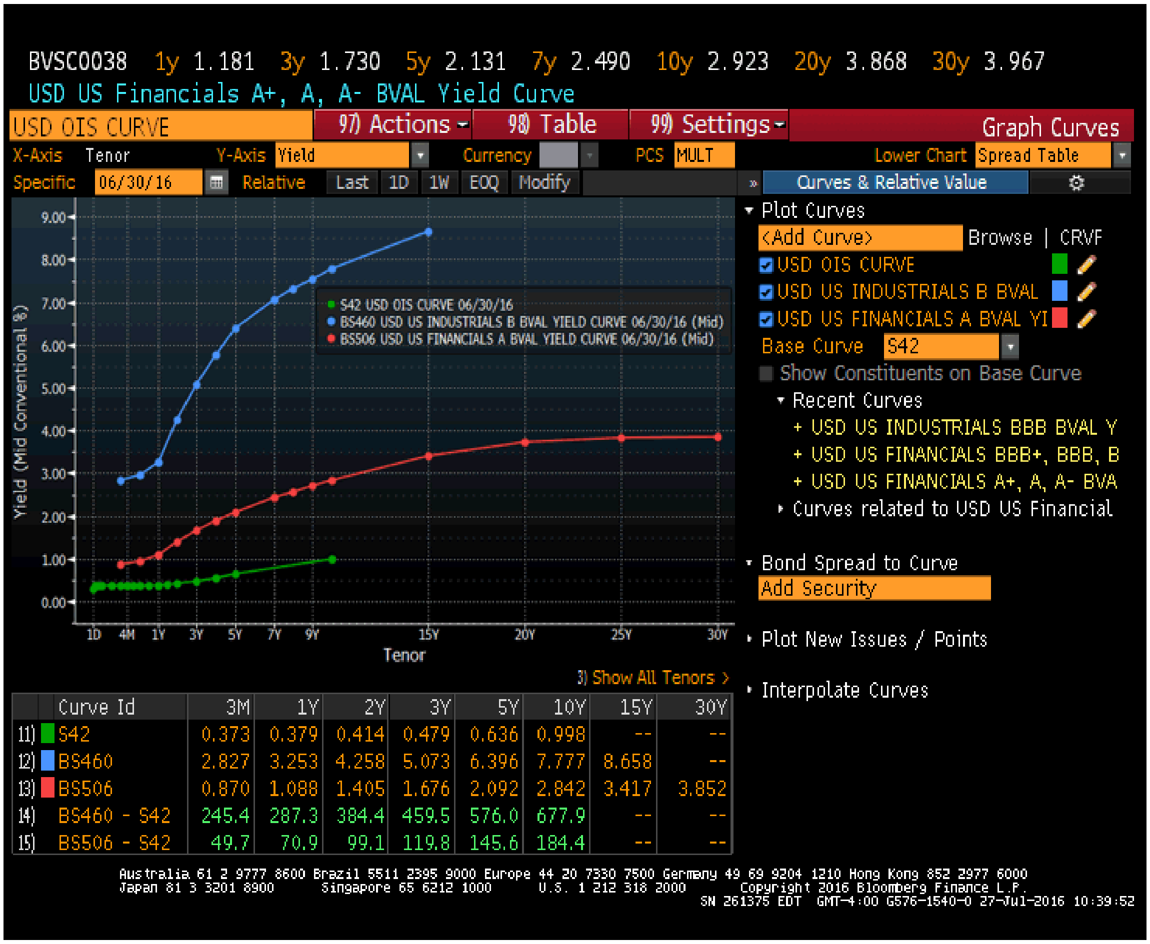
- Communications
- Consumer Discretionary
- Consumer Stables
- Energy
- Financials
- Government and Agency
- Health Care
- Industrials
- Materials
- Technology
- Utilities

Bloomberg yield curves are constructed with rates that reflect closing prices each day for senior unsecured fixed-rate bond financing in each of the respective sectors. For the purposes of our example, the parties to the trade are assumed to be an A-rated financial firm (e.g., the dealer) and a B-rated industrial (e.g., an end-

[8] The assumption of a common accrual period length for all periods (i.e., 182.5 days) may not be universally adopted. That is, some analysts might choose to measure the length of each accrual period, precisely.

user). Exhibit 13 reproduces the screen shot from Bloomberg, reflecting these curves as of the desired value date (6/30/16).

Exhibit 13: Sector Yield Curves



In our example, the time remaining until the swap's end date is six months. Thus, we interpolate results for the three-month horizon and the one-year horizon. For the A-rated financial firm, the three-month yield is shown in the bottom line of Exhibit 13 to be 49.7 basis points and the one-year yield is 70.9 basis points. The interpolated six-month yield is thus 56.8 basis points, calculated using simple weights of 2/3 and 1/3. That is, (2/3 * 49.7) + (1/3 * 70.9) = 56.8. For the B-rated industrial, the three-month and one-year yields are 245.4 and 287.3 basis points, respectively, in the line above giving an interpolated six-month rate of 259.4 basis points, (2/3 * 245.4) + (1/3 * 287.3) = 259.4.

As the swap's value will differ depending on which party bears the obligation to pay the fixed rate, we generate two sets of adjusted discount rates – one associated with the situation where the A-rated financial firm has the obligation to pay, and the other where the B-rated industrial has that obligation. Given each of the respective risk-adjusted discount rates, we find their associated risk-adjusted discount factors using equation (3), which is a restatement of equation (1). The credit-risk-adjusted discount factor (*AdjDF*) corresponds to the risk-adjusted rate (*AdjRate*), which is the sum of the implied OIS rate for each settlement date and the interpolated credit risk premium.

$$AdjDF = \frac{1}{\left(1 + \frac{AdjRate}{2}\right)^{\left(days/\left(\frac{365}{2}\right)\right)}} \tag{3}$$

Results of applying this equation are shown for both counterparties and all settlement dates in Exhibit 14.

Exhibit 14: Risk Adjusted Discount Factors

Valuation Date:	6/30/16						
Accrual End & Settlement	**Days to Settlement**	**OIS Discount Factor**	**Implied OIS Rate**	**Spread for Financial A**	**Financial A Adjusted Discount Factor**	**Spread for Industrial B**	**Industrial B Adjusted Discount Factor**
7/29/16	29	0.9997	0.3477%	0.568%	0.9995	2.594%	0.9980
8/31/16	62	0.9994	0.3578%	0.568%	0.9990	2.594%	0.9956
9/30/16	92	0.9991	0.3679%	0.568%	0.9986	2.594%	0.9935
10/31/16	123	0.9988	0.3682%	0.568%	0.9981	2.594%	0.9914
11/30/16	153	0.9985	0.3680%	0.568%	0.9976	2.594%	0.9893
12/30/16	183	0.9982	0.3653%	0.568%	0.9972	2.594%	0.9872

For example, the 92-day OIS discount factor associated with the third settlement is 0.9991 (from Bloomberg Exhibits 10 and 12). That value corresponds to a rate of 0.3679%, using equation (2):

$$Rate = \left(\left(\frac{1}{0.9991}\right)^{\left(\left(\frac{365}{2}\right)/92\right)} - 1\right) * 2 = 0.003679$$

The interpolated yield for the A-rated financial firm is 0.568% (56.8 basis points), as calculated above. Using equation (3) gives an adjusted discount rate of 0.9986:

$$AdjDF = \frac{1}{\left(1 + \frac{0.00568}{2}\right)^{\left(92/\left(\frac{365}{2}\right)\right)}} = 0.9986$$

The remaining implied OIS rates and the adjusted discount factors for Financial A and Industrial B in Exhibit are calculated in the same manner.

Given these adjusted discount rates, we can now value the swap under two alternative assumptions in Exhibit 15: (1) that the A-rated financial firm bears the obligation to pay, and (2) that the B-rated industrial bears the burden to pay.

Under the prevailing market conditions of this example, the fixed-rate *payer* (irrespective of credit quality) would be carrying the swap as a liability, while the fixed-rate *receiver* would be in the asset position. Thus, if the fixed-rate payer is the A-rated financial entity, the swap's absolute value will be $51,784.33 for both the asset and the liability; and if the fixed payer is the B-rated industrial, the common asset/liability value would be $51,466.99.

Exhibit 15: Risk Adjusted Valuations

Fixed Rate:	1.4875%	(paid by A-rated Financial firm)					
Valuation Date:	6/30/16	Notional:	10,000,000				
Accrual End & Settlement	**Accrual Days**	**Projected 1-month LIBOR Reset Rate**	**Variable Settlement**	**Fixed Settlement**	**Net Settlement**	**Financial A Adjusted Discount**	**PV**
7/29/16	29	0.460%	3,708	(11,983)	(8,275)	0.9995	(8,268)
8/31/16	62	0.458%	4,200	(13,635)	(9,436)	0.9990	(9,419)
9/30/16	92	0.456%	3,800	(12,396)	(8,596)	0.9986	(8,574)
10/31/16	123	0.472%	4,067	(12,809)	(8,742)	0.9981	(8,712)
11/30/16	153	0.472%	3,931	(12,396)	(8,465)	0.9976	(8,429)
12/30/16	183	0.476%	3,970	(12,396)	(8,426)	0.9972	(8,383)
							(51,784)

Fixed Rate:	1.4875%	(paid by B-rated Industrial firm)					
Valuation Date:	6/30/16	Notional:	10,000,000				
Accrual End & Settlement	**Accrual Days**	**Projected 1-month LIBOR Reset Rate**	**Variable Settlement**	**Fixed Settlement**	**Net Settlement**	**Industrial B Adjusted Discount**	**PV**
7/29/16	29	0.460%	3,708	(11,983)	(8,275)	0.9980	(8,254)
8/31/16	62	0.458%	4,200	(13,635)	(9,436)	0.9956	(9,385)
9/30/16	92	0.456%	3,800	(12,396)	(8,596)	0.9935	(8,528)
10/31/16	123	0.472%	4,067	(12,809)	(8,742)	0.9914	(8,650)
11/30/16	153	0.472%	3,931	(12,396)	(8,465)	0.9893	(8,355)
12/30/16	183	0.476%	3,970	(12,396)	(8,426)	0.9872	(8,295)
							(51,467)

It bears repeating that while this example reflects consistency of the paying party throughout the remaining life of the swap, in the general case, some net cash flows might be in one direction and others in the opposite direction. In such a case, the applicable adjusted discount rates would reflect the factor associated with the paying party, cash flow by cash flow.

In determining the appropriate risk premia, the methodology described herein relies on yield curves that relate to unsecured senior debt. Moreover, the objective data we use are an aggregation of portfolios of bond that exhibit at least some degree of heterogeneity. Individual users may perceive their circumstances to be better or worse than those reflected in the available yield curves, and in those cases, some subjective adjustments to the inputs may reasonably be justified.

Conclusions

Since the financial crisis, market participants have been required to address counterparty credit risk explicitly when valuing interest rate swaps for presentation in financial reports. We discuss two methods to accomplish this: the ex-post adjustment (EPA) method and the risk-adjusted discounted cash flow (DCF) method. With the EPA method, the projected settlement payments are initially discounted using OIS rates, which are proxies for "risk-free" rates; and then a further adjustment for credit risk is made to this starting valuation. The DCF method, on the other hand, uses the same projected payments under the swap, but the risk-adjusted valuation is found using discount rates that reflect the credit quality of the owing counterparty, settlement-by-settlement.

The challenge for both methods is determining the relevant inputs. The EPA calculation requires estimates for the probabilities of default and recovery rates for the parties of the swap, and the DCF method requires a determination of appropriate risk-adjusted discount rates. If the underlying data sets for these two methods happen to be mutually consistent, these two methodologies would yield identical results, but such consistency is most unlikely.

References

Green, Andrew. 2016. *XVA: Credit, Funding and Capital Valuation Adjustments*. West Sussex, UK: Wiley.

Gregory, Jon. 2015. *The XVA Challenge: Counterparty Credit Risk, Funding, Collateral, and Capital*, 3rd ed. West Sussex, UK: Wiley.

Hull, John. 2014. *Fundamentals of Futures and Options Markets*, 8th ed. Boston: Pearson.

Hull, John, and Alan White. 2013. "LIBOR vs. OIS: The Derivatives Discounting Dilemma." *Journal of Investment Management* 11: 14-27.

Jarrow, Robert A. 2012. "Problems with Using CDS to Infer Default Probabilities." *Journal of Fixed Income* 21: 6-12.

Lu, Dongsheng. 2016. *The XVA of Financial Derivatives: CVA, DVA and FVA Explained*. Hampshire, UK: Palgrave Macmillan.

Ruiz, Ignacio. 2015. *XVA Desks – A New Era for Risk Management: Understanding, Building and Managing Counterparty, Funding and Capital Risk*. Hampshire, UK: Palgrave Macmillan.

Smith, Donald J. 2013. "Valuing Interest Rate Swaps Using Overnight Indexed Swap (OIS) Discounting." *Journal of Derivatives* 20: 49-59.

Smith, Donald J. 2017. *Valuation in a World of CVA, DVA, and FVA: A Tutorial on Debt Securities and Interest Rate Derivatives.* Singapore: World Scientific.

Clarifying the Meaning of Noninteger N Values in Annuity Calculations

John S. Walker[1] and Jonathan K. Kramer[2]

ABSTRACT

Time value of money problems are integral to finance. We focus attention on the issues raised when a noninteger solution results from solving for the variable *N* (number of periods) in a time value of money annuity problem. We show that such a numerical solution, while mathematically correct, can deviate from the narrative of the problem. This deviation from the narrative could cause misunderstandings between instructors and their students and practitioners and their clients. We demonstrate how instructors and practitioners can explain the deviation from the narrative in such cases, and how cash flows can be adjusted to avoid missing goals.

Introduction

Virtually every undergraduate and graduate finance course covers time value of money (TVM) concepts. Consistent with that coverage, almost every college textbook on financial management and corporate finance—including Master of Business Administration (MBA) textbooks that contain more advanced content—dedicates a chapter to TVM. For example, Chapter 2 in Brealey, Myers, and Allen's (2017) *Principles of Corporate Finance* is titled "How to Calculate Present Value." In that chapter, the authors cover present value, future value, perpetuity, and annuity problems. They specifically define an annuity as "an asset that pays a fixed sum each year for a specified number of years" (p. 28). Brigham and Houston (2017, p. 155) provide a more general definition of an annuity, saying, "When the payments are equal and are made at fixed intervals, the series is an annuity." Generally, finance textbooks define annuities as a *series* of *equal payments* that occur at *fixed intervals*, such as months or years.[3] However, many "annuities" (such as noninteger *N* annuities) do not fit this description, and this is not made clear in finance textbooks.

Our review of the TVM treatment of lump sum and annuity problems in 31 finance textbooks finds great precision in the calculation of the missing variable, whether that is interest rate (*I/Y*), present value (*PV*), payment (*PMT*), or future value (*FV*). Likewise, this precision is also found when authors find *N*—the number of compounding periods for future value problems or discounting periods for present value problems—with noninteger *N* values a common occurrence in introductory books.[4] For example, Brigham and Houston (p. 155) find $N = 15.7473$ as the number of compounding periods needed to grow \$500,000 to \$1,000,000 at 4.5% interest. When it comes to the calculation of *N* for an annuity problem—the number of years (or periods) needed to reach a future value goal—the explanation of a noninteger *N* value is either nonexistent or insufficient in finance textbooks, which we assert is a significant omission. For lump sum problems, noninteger *N* values represent the number of discounting or compounding periods when finding the present value equivalent of a future value or vice versa. When dealing with an annuity problem, the *N* value not only

[1] Professor of Finance, Department of Business Administration, Kutztown University of Pennsylvania, email: walker@kutztown.edu.

[2] Professor of Finance, Department of Business Administration, Kutztown University of Pennsylvania, email: jkramer@kutztown.edu.

[3] Most annuities consist of a *finite* number of payments. If the series of payments is *infinite*, then the financial instrument is termed a "perpetuity."

[4] A noninteger is a real number that has both a whole number and decimal component. For example, 7.5 is composed of a whole number component, 7, and a decimal component, 0.5.

conveys information about the number of discounting or compounding periods; it also provides information about the *number* and *magnitude* of the payments. Because annuities with noninteger *N* values are inconsistent with the narrative of equal payments that occur at fixed intervals, finance textbooks should be more thorough on this topic than is generally found. Books lacking this clarity will leave students with incomplete information about annuity cash flows and their timing for noninteger *N* scenarios.

From a pedagogical standpoint, finance professors want students to have a solid understanding of TVM concepts. This is evident by the fact that TVM is covered in virtually every finance textbook, and professors teach or review TVM concepts in nearly all finance courses. Furthermore, an understanding of TVM concepts is the foundation that students need in order to progress to more advanced topics in finance. The lack of explanation regarding noninteger *N* values for annuity problems is a significant gap in finance textbooks that does not appear to be addressed in more advanced books. It makes sense for instructors to clarify the meaning behind noninteger *N* values during introductory finance classes to reinforce and contrast the implications for lump sum versus annuity problems. Lump sum problems have just one payment regardless of whether *N* is an integer or not. In contrast, annuity problems shift from equal payments for integer *N* scenarios to a series of equal plus one unequal payment for noninteger *N* scenarios. This change in the structure of cash flows deserves thorough treatment in finance textbooks at all levels.

From a practical standpoint, financial planners who are working with clients on retirement planning will often estimate the “nest egg” needed to retire comfortably. Based on that estimate and an assumed rate of return on savings, the planner can outline how much to invest each year to reach the client’s goal. For example, suppose the planner tells his client that the goal can be reached in 18.2 years. What does that mean exactly in terms of the magnitude and timing of cash flows? If the client makes 18 annual payments and then waits an additional 0.2 of a year for those payments to grow at an assumed rate of return, will he reach his goal? Or, does the client need to make a partial, prorated nineteenth payment equal to 20% of the full payment to reach his goal? Perhaps surprisingly, the answer is “no” to both of these questions.

The scenario of a noninteger *N* annuity problem can also arise in a legal environment. For example, Kellison (2009, p. 89) makes the point that noninteger *N* values arise in court cases dealing with economic damages in personal injury and wrongful death lawsuits. An individual’s life expectancy is often a noninteger value (for example, 15.7 years), which needs to be transformed into what is termed “an annuity-certain.” Thus, this topic is relevant from a pedagogical, financial planning, and legal perspective.

In this paper, we build a case showing the inadequate treatment of the meaning behind noninteger *N* values in annuity problems, and then clarify its meaning. Given how thorough finance textbooks are generally with respect to nearly all other TVM calculations, we recommend that authors writing finance textbooks not gloss over this topic but, rather, treat it more thoroughly, consistent with the treatment of other TVM calculations. Interestingly, we find a more adequate treatment in textbooks used for teaching actuarial science; however, most finance students do not take courses in actuarial science.[5] If the meaning of noninteger *N* values is more thoroughly explained in financial management, corporate finance, and personal finance textbooks, this will benefit finance students in general and certainly those students aspiring to careers in financial planning, who often need to explain to clients how annuities work.

This paper is organized as follows: The next section provides a review of the treatment provided by finance textbooks on TVM annuity problems that have noninteger *N* values and contrasts that treatment to what we found in actuarial science textbooks. The third section discusses how this gap in the textbooks was discovered and goes into greater depth as to why this is an issue from a pedagogical standpoint. The fourth section discusses the mathematics behind the noninteger *N* annuity problem. The fifth and final section provides conclusions and recommendations.

Review of Textbook Treatment

Appendix A contains the list of finance textbooks examined as part of our research. The books are both undergraduate and MBA level, each one having a chapter or section within a chapter dedicated to TVM concepts. The books are grouped into three categories: (1) books that calculate a noninteger *N* value for an annuity problem with some discussion about the partial payment; (2) books that calculate a noninteger *N* value but do not include a discussion; and (3) books that neither calculate a noninteger *N* value nor discuss

[5] We thank Ian Haberle, a Kutztown University of Pennsylvania student studying actuarial science, and Victoria Geyfman, a professor of finance at Bloomsburg University of Pennsylvania, for pointing us to the actuarial science textbooks.

it. For the sample of 25 financial management and corporate finance books we investigated, four (16%) calculate a noninteger *N* value and discuss the partial payment, eight (32%) calculate a noninteger *N* value but do not discuss it, and 13 (52%) neither calculate a noninteger *N* value nor discuss it. In addition, we examined six personal finance textbooks and found that two of them calculate a noninteger *N* value, but none of them discuss partial payments.[6]

In contrast to what we found with finance textbooks, the noninteger *N* value is explained thoroughly in actuarial science textbooks that we examined, as well as those written with the actuarial scientist in mind. For example, *The Theory of Interest* (Kellison, 2009, pp. 89–93), a book targeting those preparing for the actuarial exam, contains a section titled "Unknown Time" that discusses annuity problems involving unknown time that "[do] not produce exact integral answers for *N*." In the example given in the Kellison book, the *N* value found is 14.2067. The interpretation is that there are 14 end-of-period equal annuity payments, with one additional, partial payment needed. Kellison discusses the "inconvenience and confusion of making a [partial] payment at a date which is not an integral number of periods from the dates all other payments are made." He provides a numerical example showing the partial payment needed at three different points on the timeline: at 14 years, at 14.2067 years, and at 15 years. He also provides terminology that is not used in the finance textbooks that we investigated. If an earlier partial payment is made at 14 years, it is termed a "balloon payment" because it would be in addition to the regular end-of-period payment. If the partial payment is made at 14.2067 years, it is termed an "exact payment." Finally, if the partial payment is made at 15 years, it is termed a "drop payment." In a textbook written for students studying actuarial science, Chan and Tse (2013, pp. 62–63) present a similar approach to solving noninteger *N* annuity problems.

While Kellison and others writing for the field of actuarial science provide the formula needed to calculate the exact payment, two other authors who specifically target students preparing for the actuarial science exam, Cherry and Gorvett (2008), show a less mathematical approach and reverse-engineer the numbers by finding what additional payment is needed to equate the annuity stream to the present value. (They look at present values of annuities while we examine future values of annuities.) In our analysis, we show the derivation of the factor that can be used to calculate the exact amount of the partial payment, based on the future value of an annuity formula.

Identification of the Gap in Finance Textbooks

Our interest in this topic surfaced when we were preparing pedagogical materials for an online course on financial management. When we got to the creation of an annuity example, specifically finding the number of years, *N*, needed to reach a future value goal, the variables, which were arbitrarily selected, were *I/Y* = 7.00%, *PV* = $0.00, *PMT* = $1,000.00 (which is entered into the calculator with a negative sign), and *FV* = $10,000.00. The problem statement given to students could read, "Find the number of years it would take to build a $10,000.00 nest egg if the client invests $1,000.00 at the end of each year at 7.00%." Typically, a financial calculator, such as the Texas Instruments (TI) BA II Plus, is used when teaching finance. The keypad for this particular calculator (shown in Appendix B) includes five keys in the third row that correspond to the five variables used for TVM problems: *N*, *I/Y*, *PV*, *PMT*, and *FV*. The typical setup for a TVM problem is to provide values for four of the five variables, leaving one unknown. Once the above variables have been entered by students, they then press the CPT key and then the *N* key to calculate an answer of 7.8427 (rounded to the ten-thousandths place if the calculator has been set for four decimal places).

When TVM topics are introduced in a finance course, the first step is to discuss the timeline. Professors ask students to show the cash flows (CFs) and the timing. For the above ordinary annuity problem with CFs occurring at the end of each period, there would be a constant amount of $1,000.00 shown at the end of years 1 through 7. Yet, if those are the only CFs, by the time an investor reaches 7.8427 years, the future value is only $9,161.79. This would be $838.21 or 8.4% below the goal of $10,000.00, *which is a substantial shortfall* (see Table 1).

[6] We checked the TVM chapters and sections, as well as examples and illustrations found on those pages of each textbook. If we missed any relevant material on noninteger *N* annuity problems, we are responsible for that error and apologize to the authors.

Table 1: Shortfall for Noninteger *N* Annuity without Partial Payment

N	PMT	FV (as of N=7.8427)
1	$1,000.00	$1,588.79
2	$1,000.00	$1,484.85
3	$1,000.00	$1,387.71
4	$1,000.00	$1,296.92
5	$1,000.00	$1,212.08
6	$1,000.00	$1,132.78
7	$1,000.00	$1,058.67
	Goal	$10,000.00
FV of All Payments		$9,161.79
	Shortfall	$838.21

Closing this shortfall with an additional payment leads to a twofold question: how much is the additional payment and when does it occur? In contrast to what we find with actuarial science textbooks, our survey of finance textbooks reveals that most authors ignore or skirt these issues. Some authors design their problems to give an integer value for *N*. Yet, if we view *N* as a continuous random variable, the probability of obtaining an exact integer value for a TVM annuity problem is *zero*. Thus, textbooks that show integer solutions to unknown *N* annuity problems are presenting unrealistic examples that sidestep the ambiguity and significance of noninteger *N* values.

The authors that we found who do the best job of addressing the noninteger *N* annuity problem in textbooks targeting finance students are Smart, Megginson, and Gitman (2007, pp. 104–105). Their illustration involves different numbers than ours and is for a loan repayment (thus, a present value annuity problem) rather than building a nest egg. They solve for the number of payments and find $N = 14.2$ based on a payment amount of $3,000. They proceed to say that a payment of "about $600 (0.20 × $3,000) [is needed] at the end of 14.2 years in order to fully repay the $20,000 loan at 12 percent." Thus, they are suggesting an amount that is proportional to the decimal component of *N*, although by saying "about" they are alluding to the fact that a prorated number based on time is not exact.

To be clear, there are two relevant issues our research highlights; one is more practical and the other is more pedagogical in nature. The significant practical issue is that without another payment, a substantial shortfall can exist with a noninteger *N* annuity problem. This could lead to underfunding of goals and an upset client. The second issue, which is, arguably, more important from a pedagogical than a practical perspective, is to understand how to calculate the *exact* payment. For example for the problem we outlined above (see Table 1), if the prorated approach is used, the *FV* = $10,004.52. The difference between the goal of $10,000.00 and the projected value of $10,004.52 is small, showing that a prorated extra payment almost closes the entire shortfall. Similarly, if a prorated payment is used in the Smart et al. example at 14.2 years, the *PV* = $20,005.41. Again, this is a small difference—just 0.03% greater than the target *PV* of $20,000, but it lacks the numerical precision we typically see in all other TVM topics discussed in finance textbooks.

Future research could survey practitioners in the financial planning profession to determine whether this lack of exactness is important in a practical setting. In larger transactions, the perhaps seemingly immaterial difference could be viewed as more substantive. In terms of textbook treatment, it is interesting that actuarial scientists include the rigor in their books to ensure that students understand the balloon, exact, and drop payment solutions to the unknown *N* annuity problems, while authors of finance textbooks have elected to essentially ignore these concepts. Perhaps authors of finance textbooks feel that a full explanation of the math behind the partial payment is pedantic and unnecessary. Yet, an understanding of the timing options and derivation of these partial payment amounts reinforces the understanding of annuity problems specifically, and TVM problems in general, and certainly the need for a partial payment of *some* magnitude should be made clear by the authors of finance textbooks.

The Mathematics Behind Noninteger *N* Annuity Problems

As shown by Smart, Megginson, and Gitman (2007) and discussed above, the intuitive way to approach this problem is to estimate the missing payment using a prorated formula. However, their approach is not exact—it is an approximation. Another practical approach—which *is* exact—is to find the future value of the annual payments, compounding them for *N* number of years (7.8427 years in our example), and then find the additional lump sum payment needed at the end of the annuity period to make up for the shortfall. In our example depicted in Table 1 above, the future value of the $1,000 annuity payments is $9,161.79, meaning that an additional payment is required to cover the shortfall and to reach the goal of $10,000.00. This accounting approach reveals that the exact amount needed to close the shortfall is $838.21 (shown in Table 1 above). This exact shortfall is $4.52 less than the prorated estimate of $842.73 ($1,000 × 0.84273).

One approach to understanding why the prorated estimate is not exact is to examine the factor used to find the future value of an annuity. In Appendix C we show the derivation of the future value of an annuity with *N* periods (FVA_N). The derivation shows that FVA_N is precisely equal to the payment (*PMT*) times a factor ($\frac{(1+I)^N-1}{I}$). For ordinary annuity calculations, this is the factor that provides the exact future value of an annuity that spans *N* periods. The significance of this derivation for this research is that this factor can serve as a *proportionality* factor that enables us to calculate the *exact* partial payment needed to achieve the future goal for an annuity problem. For example, if we use I = 7.00% and N = 0.8427 from the example above (found by removing the integer component from the noninteger *N* value initially found for the problem), then the proportionality factor is 0.8382, and the exact partial payment is $1,000 × 0.8382 = $838.21.[7] This additional payment exactly matches the shortfall amount shown above in Table 1.

The actuarial science textbooks show three options for the timing of the additional payment: (1) at the end of the last whole year before *N* on the time line (found by rounding *N* down), (2) at exactly time *N* on the time line, or (3) at the end of the next whole year after *N* on the time line (found by rounding *N* up to the closest integer). For our ongoing example shown in Table 1, the additional payment can occur (1) at the end of year 7 (a balloon payment that includes the periodic fixed payment plus the partial payment), (2) in between years 7 and 8 at 7.8427 years (an exact partial payment), or (3) at the end of year 8 (a drop payment). To find the balloon payment, the exact payment (as calculated above) must be discounted by the fractional period in the noninteger *N*, or 0.8427 years for our example, giving a value of $791.75. Likewise, the drop payment is found by compounding the exact payment by 1 minus the fractional period, or 1 − 0.8427 = 0.1573 for our example, giving a value of $847.17. If a balloon payment is made, the discounted value of the exact payment ($791.75) must be added to the annuity payment. For our example, the balloon payment is $1,000 + $791.75= $1,791.75. For all three options, which are graphically depicted in Figure 1—the balloon payment, the exact payment, and the drop payment—the future value as of N = 7.8427 years is exactly $10,000, which is the goal. Practitioners would likely choose to implement the balloon payment (option 1) or the exact partial payment (option 2). Although both of these options are outside of the narrative of an annuity (equal payments at fixed intervals), we think that they would be the easiest and most intuitive options to explain to a client.

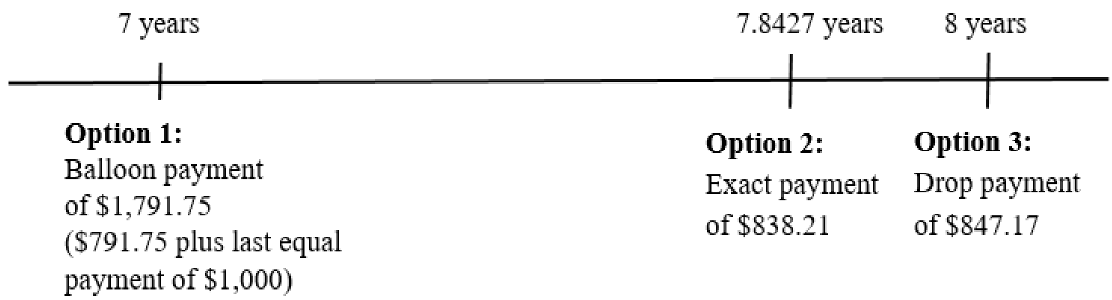
Quantification of the Error and Location of Its Maximum

Once we verified that the FVA_N factor not only determines the future value of an annuity but also the partial payment for a noninteger *N* annuity problem, we then wanted to investigate exactly how much error is introduced by estimating the partial payment using the prorated approach. Also, we wanted to see where, along the interval, the error is maximized. While our analysis focuses on future value of ordinary annuity problems, the same factor can also be used to determine the exact partial payment for present value of ordinary annuity problems. Likewise, our analysis can be extended to annuity due problems, but the proportionality factor needs to be multiplied by the factor (1+*I*).[8]

[7] The computation of the proportionality factor of 0.8382 is shown in the next section.

[8] For an ordinary annuity with a noninteger *N* solution, the balloon payment occurs at the end of the year prior to *N*, while the drop payment occurs one year later. For example with N = 7.8427, the balloon payment occurs at N = 7, the exact payment occurs at N = 7.8427, and the drop payment occurs at N = 8. We did not find an example of the timing for the balloon, exact, and drop payments for noninteger *N annuity due* problems in the actuarial science literature that we checked. Given that cash flows for an annuity due

Figure 1: Three Options for Final Payment: (1) Payment Added to Last Regular Payment, (2) An Exact Payment at *N* = 7.8427, or (3) A Final Payment at the Next Fixed Interval After the Last Fixed Payment



In virtually all cases, solving for *N* in an annuity problem will lead to a noninteger value, meaning there is a whole number component and a decimal component to *N*. For a noninteger, the decimal part of the number falls in the interval (0,1), where the parentheses denote an *open* interval. A closed interval is not used because the decimal part of a noninteger is greater than 0 and less than 1. This research addresses the noninteger *N* scenario.

In our inspection of well-known finance textbooks from the major textbook publishers for higher education, the most specific recommendation found for calculating the partial payment to make for noninteger *N* annuity problems is to prorate the payment based on the fractional time involved. If we define x as the decimal part or fractional time of the noninteger *N* value that falls in the interval (0,1), then the prorated factor (y) used to find the partial payment is simply given as $y = x$, and the partial payment estimate is $PMT \times y$. For our example, the prorated factor $y = 0.8427$, derived from the decimal part of $N = 7.8427$, and the partial payment estimate is $1,000 × 0.8427 = $842.73. However, if the *exact* proportionality factor for determining the partial payment is used for y instead, then we have $y = \frac{(1+I)^x-1}{I}$. Now for our example, rather than $y = 0.8427$, we instead have $y = \frac{(1.07)^{0.8427}-1}{0.07} = 0.8382$, and the partial payment is $1,000 × 0.8382 = $838.21 (based on $I = 7.00\%$).

We define the absolute error (*AE*) or difference between the approximate factor ($y = x$) and the exact factor ($y = \frac{(1+I)^x-1}{I}$) as:

$$AE = x - \frac{(1+I)^x-1}{I}. \tag{1}$$

The absolute error function does not require absolute value notation because the prorated factor always exceeds the exact factor; thus, the absolute error function is always positive. The partial derivative with respect to x of Equation (1) is:

$$\frac{\partial AE}{\partial x} = 1 - \frac{1}{I}\ln(1+I)\,(1+I)^x \tag{2}$$

In order to determine whether the absolute error function reaches a maximum or minimum over the interval (0,1), we find the second derivative:

$$\frac{\partial^2 AE}{\partial x^2} = -\frac{1}{I}[\ln(1+I)]^2(1+I)^x \tag{3}$$

Given that I and x are positive, which ensures that all terms in Equation (3) are positive, the overall sign of the right-hand side of Equation (3) is negative because of the negative coefficient. The negative sign on the second derivative ensures us that the absolute error reaches a maximum rather than a minimum for the x interval (0,1). Furthermore, at the maximum, $\frac{\partial AE}{\partial x} = 0$ or:

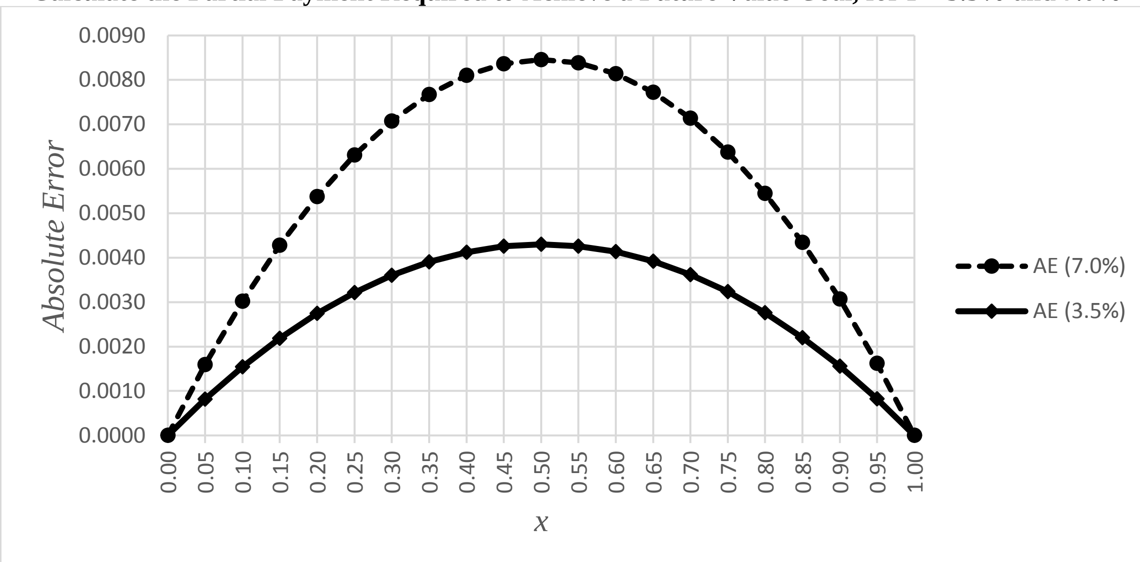
$$1 - \frac{1}{I}\ln(1+I)(1+I)^x = 0 \tag{4}$$

Figure 2 graphically shows and confirms that the function reaches a maximum rather than a minimum for the x interval (0,1), and visually it looks like this occurs around the midpoint of the interval. Note that the absolute error is not only dependent on x, it is also dependent on the interest rate, I, as evident from the contrast between the absolute error for $I = 3.5\%$ versus $I = 7.0\%$. As I approaches zero, the apex of the

occur one year earlier, one possibility would be for the balloon and drop payments to occur one year earlier than shown for an ordinary annuity.

function approaches zero. If I = 0%, then the timing of cash flows is immaterial and the future value of the annuity is simply equal to the sum of the nominal cash flows, independent of timing. Conversely, as I increases, the function's maximum increases, meaning that the error between the exact and approximate payments increases.

Figure 2: The Absolute Error (*AE*) between the Prorated Factor and the Exact Factor Needed to Calculate the Partial Payment Required to Achieve a Future Value Goal, for *I* = 3.5% and 7.0%



At this point of the analysis we skip several steps, but we rearrange Equation (4) and solve for x in order to find the value for x where the absolute error function is maximized for any particular value of *I:*

$$x = \frac{\ln\left[\frac{I}{\ln(1+I)}\right]}{\ln(1+I)} \tag{5}$$

For example, if I = 7.00%, as used earlier in our illustration, we find that x = 0.5028. What this shows is that the difference between the prorated factor and the exact factor reaches a maximum around the midpoint of the (0,1) interval (namely, close to x = 0.5), *but not exactly at the midpoint.* As the interest rate increases, the value of x where the function reaches a maximum monotonically increases. In turn, once we find the x value where the function is maximized, we can use Equation (1) to calculate the maximum absolute error. In the case of a 7.00% interest rate, the error is 0.008457. In all cases of a finite interest rate, the prorated estimate will overestimate the exact payment needed to meet the annuity goal. For our 7.00% interest rate example, we find the absolute error is less than 1% across the entire (0,1) interval.

Another way to measure the difference between the prorated and the exact partial payment is to find the *relative* error. The relative error function (*RE*) is:

$$RE = \frac{x}{\frac{(1+I)^x - 1}{I}} - 1 = \frac{xI}{(1+I)^x - 1} - 1 \tag{6}$$

Because the prorated factor is always an overestimation of the exact factor, the relative error will always be positive given the specification used for Equation (6), with the prorated factor in the numerator and the exact factor in the denominator. The additional information provided by this calculation is worthwhile because it quantifies the error in the estimated partial payment for an annuity problem using the prorated approach in relative terms. For the ongoing example we have used in this article, the absolute error between the prorated factor and exact factor is 0.45%, while the relative error is 0.54%. This is not much of a difference. However, recall that earlier we showed that the *absolute* error reaches a maximum around the center of the (0,1) interval. In the case of the *relative* error, it monotonically increases as x approaches zero, and the difference between the absolute error and relative error becomes more significant.

Conclusions

What exactly is the meaning of *N* in a noninteger annuity problem? Does it represent the number of periods or the number of payments? Does it tell us the magnitude of cash flows? Our paper has demonstrated that the narrative for noninteger *N* annuity problems is ambiguous on several levels. If 100 finance professors were surveyed and asked how they would explain the timing and magnitude of cash flows for a noninteger *N* annuity problem, we expect there would be a variety of explanations, many of them incorrect. Brigham and Houston (2017, p. 156) state, "Annuities must have *constant payments* at *fixed intervals* for a *specified number of periods"* (emphasis provided by the authors). They go on to say, "If these conditions don't hold, then the payments do not constitute an annuity." If their statements are taken literally, then an annuity problem with a noninteger *N* solution is an oxymoron, because the payments are *not* constant and the payments are *not* necessarily at fixed intervals. Only the balloon and drop payment solutions maintain the fixed interval pattern; but they fail the constant payment test.

Our research has uncovered a gap—namely, an inadequate explanation of the meaning of noninteger *N* values for annuity problems—in finance textbooks, including books on financial management, corporate finance, and personal finance. The explanation that comes closest to satisfactory is found in Smart, Megginson, and Gitman (2007). They recognize that an additional payment is needed, and they estimate this payment based on a prorated factor, with the timing of that payment occurring at the end of the annuity period. Our analysis shows that using a prorated factor is a good approximation; some will likely feel that using a prorated factor is sufficient. We show how the future value of an annuity factor—which is well known in finance—can be used as a proportionality factor to find the exact partial payment for a noninteger *N* annuity problem. We have not seen this explanation in the finance textbooks that we reviewed. In contrast, we find that the actuarial scientists are exact in the determination of partial payments for annuities, leaving us wondering why authors of finance books—especially MBA textbooks on corporate finance, which should be more thorough and rigorous—have elected to gloss over the mathematics behind the noninteger *N* annuity problem. The clarification that a partial payment is needed and how to calculate it precisely could be handled efficiently by inserting a footnote in the annuity section of a textbook.

As for professors teaching personal financial planning courses, they need to give special attention to ensuring that the numbers and narrative are consistent. Eventually these students will be dealing directly with clients and their credibility is at stake. Financial goals are quantified and strategies are designed to achieve those goals. It could be unsettling to a client engaging in a planning issue that involves an annuity product if the magnitude and timing of cash flows needed to obtain a future value goal are not explained correctly. Often, consumers will opt for annuity products in order to reduce uncertainty in the outcome. Thus, it will benefit future financial planners to fully understand and be able to explain to their clients the mathematics inherent to annuities, including noninteger *N* scenarios.

References

Brealey, Richard A., Stewart C. Myers, and Franklin Allen. 2017. *Principles of Corporate Finance* (12e). New York, NY: McGraw–Hill Education.

Brigham, Eugene F., and Joel F. Houston. 2017. *Fundamentals of Financial Management* (Concise, 9e). Boston, MA: Cengage Learning.

Chan, Wai-Sum, and Yiu K. Tse. 2013. *Financial Mathematics for Actuaries*. Singapore: McGraw-Hill Education.

Cherry, Harold, and Rick Gorvett. 2008. *ASM Study Manual for Exam FM/Exam 2: Financial Mathematics & Financial Economics (*9e). Westbury, NY: Actuarial Study Materials.

Kellison, Stephen G. 2009. *The Theory of Interest* (3e). New York, NY: McGraw–Hill Irwin.

Smart, Scott B., William L. Megginson, and Lawrence J. Gitman. 2007. *Corporate Finance* (2e). Mason, OH: Thomson South-Western.

APPENDIX A

Below is a list of financial management, corporate finance, and personal finance books that we inspected, along with our condensed research notes. Each book inspected is categorized in one of three ways. If the author or authors calculate a noninteger *N* value for an annuity problem and discuss the need for a partial payment, the book is labeled "C/D." If the author or authors calculate a noninteger *N* value for an annuity problem but do not discuss a partial payment, the book is labeled "C/DD." Finally, if the author or authors neither calculate a noninteger *N* value for an annuity problem nor discuss it, then the book is labeled "DC/DD."

Financial Management and Corporate Finance Textbooks

1. Benninga, S. (2006). *Principles of Finance with Excel*. New York, NY: Oxford University Press, Inc. In a section titled "How Long Will It Take to Pay Off a Loan?" (pp. 112–114) he uses the NPER Excel function and finds the number of periods for an annuity problem to be 5.4. However, there is no discussion about a partial payment. (C/DD)

2. Berk, J., & DeMarzo, P. (2017). *Corporate Finance* (3e). Boston, MA: Pearson Education, Inc. In Chapter 4 Appendix, they have a section titled "Solving for the Number of Periods" (pp. 141–142) where they solve for *N* for an annuity problem. The problem is designed to produce an integer *N* value. (DC/DD)

3. Berk, J., DeMarzo, P., & Harford, J. (2012). *Fundamentals of Corporate Finance* (2e). Boston, MA: Pearson Education, Inc. There is a section on annuities in Chapter 4, but they neither calculate nor discuss noninteger *N* value solutions. (DC/DD)

4. Block, S. B., Hirt, G. A., & Danielson, B. R. (2017). *Foundations of Financial Management* (16e). New York, NY: McGraw-Hill Education. In a section titled "Finding the Number of Annuity Payments" (p. 275), they say, "Excel's NPER function calculates the interest rate that equates the annuity payment stream with the present value or future value of the annuity. The NPER function assumes that each payment is at the end of the period as shown in the previous timeline." The narrative is either an error or confusing because the NPER function in Excel finds the number of periods *not* the interest rate. The example they show finds $N = 4.00$. These authors have carefully designed the example to give an integer value for *N*, so they have sidestepped the noninteger *N* scenario. (DC/DD)

5. Booth, L., Cleary, W. S., & Peterson-Drake, P. (2014). *Corporate Finance: Financial Management in a Global Environment*. Hoboken, NJ: John Wiley & Sons. In a section titled "Determining the Number of Periods or Interest Rates" (pp. 165-166), they say, "using a financial calculator or a spreadsheet, we calculate the number of payments as 39.7217…because these are semiannual payments, it will take forty payments (with the last payment not quite $4 million to repay the loan)." (C/D)

6. Brealey, R. A., Myers, S. C., & Allen, F. (2017). *Principles of Corporate Finance* (12e). New York, NY: McGraw-Hill Education. There is no mention of solving for *N* in a rather limited discussion on annuities. (DC/DD)

7. Brigham, E. F., & Daves, P. R. (2016). *Intermediate Financial Management* (12e). Boston, MA: Cengage Learning. They discuss the present value and future value of annuities in Chapter 28, but they do not calculate or discuss *N* values. (DC/DD)

8. Brigham, E. F., & Ehrhardt, M. C. (2011). *Financial Management* (13e). Mason, OH: South-Western Cengage Learning. In a section titled "Finding the Number of Periods, *N*" (p. 145), they find a noninteger *N* values for an ordinary annuity, but they do not discuss a partial payment. Their treatment is a condensed version of what is found in Brigham & Houston. (C/DD)

9. Brigham, E. F., & Houston, J. F. (2017). *Fundamentals of Financial Management* (Concise, 9e). Boston, MA: Cengage Learning. In a section titled "Finding the Number of Periods, *N*" (p. 163), they find

noninteger *N* values for an ordinary annuity and an annuity due, but they do not discuss partial payments. (C/DD)

10. Brooks, R. M. (2016). *Financial Management: Core Concepts* (3e). Boston, MA: Pearson Education, Inc. He discusses annuities and has a section titled "Waiting Times and Interest Rates for Annuities" (Section 4.7 in the e-book). In Example 4.3, "So you want to be a millionaire?," the question is: How long will it take to build $1 million in savings by investing $5,000 per year at 6%? He finds *N* = 44.0192 years, with no mention of a partial payment. (C/DD)

11. Cornett, M. M., Adair, T., Nofsinger, J. R. (2012). *Finance: Applications & Theory* (2e). New York, NY: McGraw-Hill Irwin. In a section titled "Compute the Time Period" (p. 164), they say, "To find the number of periods, you can solve equation 5–9 for *N*—the number of payments—but the equation becomes quite complicated." They provide a footnote that gives the equation for *N*. They go on to say "Many people just use a financial calculator or spreadsheet." (DC/DD)

12. Damodaran, A. (1997). *Corporate Finance Theory and Practice*. Hoboken, NJ: John Wiley & Sons. He does not calculate or discuss noninteger *N* values for annuity problems. (DC/DD)

13. Emery, D. R., & Finnerty, J. D. (1991). *Principles of Finance with Corporate Applications* (1e). St. Paul, MN: West Publishing Company. This is an old MBA textbook. It discusses annuities thoroughly, but nothing pertaining to solving for *N*. However, it is notable that they have a section titled "Annuities with Partial Time Periods" (p. 59) that begins by saying, "annuity formulas are not normally used with fractional time periods, except in the case of a contract that requires partial payments for partial time periods." They then give two examples using fractional time periods for *integer N* annuity scenarios, with no examples of noninteger *N* values. (DC/DD)

14. Gitman, L. J. (2006). *Principles of Managerial Finance* (11e). Boston, MA: Pearson Addison-Wesley. In a section titled "Finding an Unknown Number of Periods" (pp. 197–199), he finds a noninteger *N* value for an annuity of 8.15. He does not discuss the need for a partial payment. (C/DD)

15. Gitman, L. J., & Zutter, C. J. (2012). *Principles of Managerial Finance* (6e). Boston, MA: Pearson. In a section titled "Finding an Unknown Number of Periods" (pp. 183–184), they calculate a noninteger *N* value of 8.15 years for the present value of an annuity and go on to say, "[The borrower] will still have a small outstanding balance." (C/D)

16. Graham, J. R., Smart, S. B., & Megginson, W. L. (2010). *Corporate Finance* (3e). Mason, OH: South-Western Cengage Learning. In a section titled "Number of Compounding Periods" (pp. 92–93), the same example found in Smart, Megginson, & Gitman is given. (C/D)

17. Lasher, W. R. (2008). *Practical Financial Management* (5e). Mason, OH: Thomson South-Western. He does not calculate or discuss noninteger *N* values for annuity problems. (DC/DD)

18. Moyer, R. C., McGuigan, J. R., & Rao, R. P. (2005). *Contemporary Financial Management Fundamentals*. Mason, OH: Thomson South-Western. In a section titled "Solving for the Discount Rate, Annuity, or Length of Annuity" (pp. 154–155), they calculate a noninteger *N* value for an annuity, but do not discuss the need for a partial payment. (C/DD)

19. Parrino, R., Kidwell, D. S., & Bates, T. W. (2012). *Fundamentals of Corporate Finance* (2e). Hoboken, NJ: John Wiley & Sons, Inc. There is no mention of solving for *N*, despite an extensive discussion on annuities in Chapter 6. (DC/DD)

20. Ross, S. A., Westerfield, R. W., & Jaffe, J. (2010). *Corporate Finance* (9e). New York, NY: McGraw-Hill Irwin. There is no mention of solving for *N* in a rather limited discussion on annuities. However, there is an annuity problem at the end of the chapter that produces a noninteger *N* solution. (DC/DD)

21. Ross, S. A., Westerfield, R. W., & Jordan, B. D. (2013). *Fundamentals of Corporate Finance* (10e). New York, NY: McGraw-Hill Irwin. In a section titled "Finding the Number of Payments" (p. 159), they say essentially the same thing they say in their *Essentials of Corporate Finance* book. (C/DD)

22. Ross, S. A., Westerfield, R. W., & Jordan, B. D. (2014). *Essentials of Corporate Finance* (8e). New York, NY: McGraw-Hill Irwin. In a section titled "Finding the Number of Payments" (p. 136), they find N = 93.11. They go on to say, "Also, some financial calculators won't report a fractional value for N; they automatically (without telling you) round up to the next whole period (not to the nearest value)." (C/DD)

23. Smart, S. B., Megginson, W. L., & Gitman, L. J. (2007). *Corporate Finance* (2e). Mason, OH: Thomson South-Western. In a section titled "Number of Compounding Periods" (pp. 104–106), they find N = 14.20 and say a payment of "about \$600 (0.20 × \$3,000) at the end of 14.20 years [is required] in order to fully repay the \$20,000 loan at 12 percent." This is an interesting example because the precise amount is \$577.48; their number is 3.9% higher than the exact number. Unlike what is discussed in the actuarial science books, there is no mention that the partial payment could be made at the end of year 14 or year 15. (C/D)

24. Titman, S., Keown, A. J., & Martin, J. D. (2014). *Financial Management: Principles and Applications* (12e). Boston, MA: Pearson Education, Inc. In a section on annuities (pp. 158–171), they discuss solving for N and they designed an example to give them an N = 8.0 (an integer), so they skirt the issue. (DC/DD)

25. Welch, I. (2009). *Corporate Finance: An Introduction*. Boston, MA: Pearson Education, Inc. He has a brief section on annuities in Chapter 3, but does not solve for N values. (DC/DD)

Personal Finance Textbooks

1. Dalton, M. A., Dalton, J. F., Gillice, J. M., & Langdon, T. P. (2016). *Fundamentals of Financial Planning* (5e). Metairie, LA: Money Education. Chapter 7, titled "Time Value of Money," gives extensive coverage to calculating present and future values, but does not discuss the calculation or interpretation of N values for annuities. (DC/DD)

2. Garman, E. T., & Forgue, R. (2012). *Personal Finance* (11e). Mason, OH: South-Western Cengage Learning. In problem 3(h) on p. 31, the authors ask the reader to solve for the amount of time that a fund will last if a person makes an annual withdrawal in the form of an annuity. The answer is a noninteger N, but there is no discussion of how to interpret it. (C/DD)

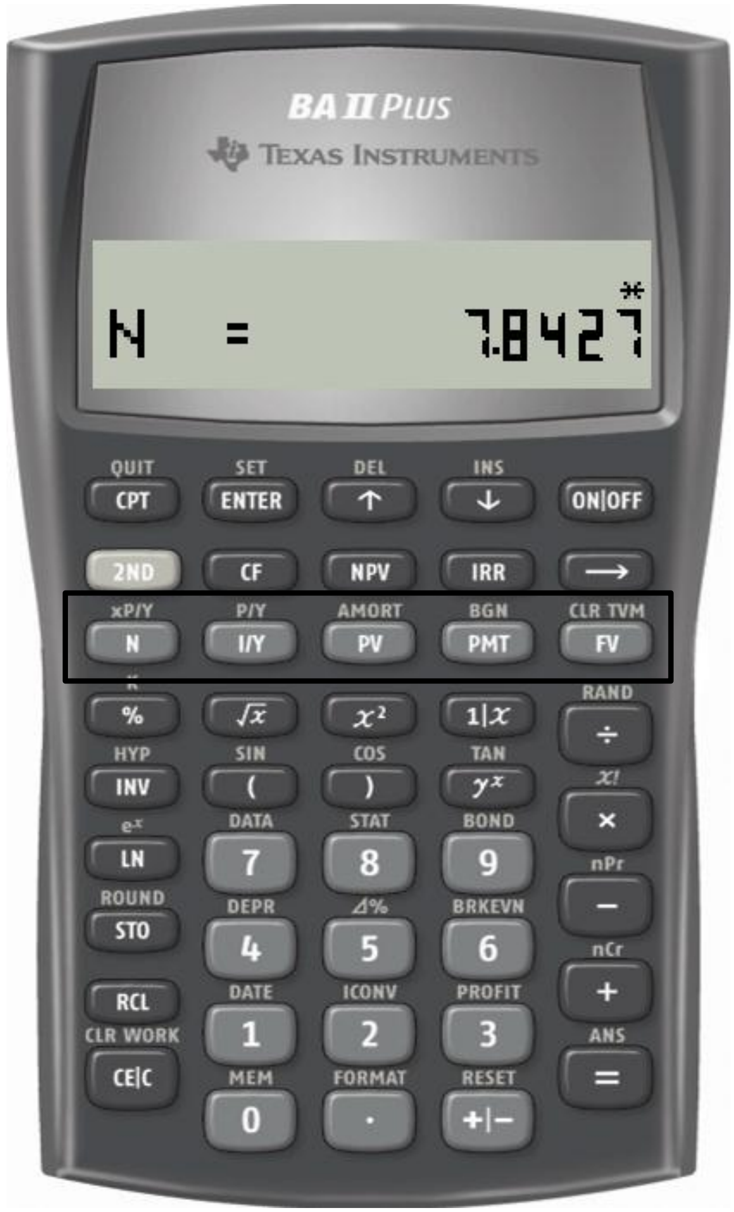
3. Gitman, L. J., Joehnk, M. D., & Billingsley, R. S. (2014). *Personal Financial Planning* (13e). Mason, OH: South-Western Cengage Learning. Chapter 2 contains a section titled "The Time Value of Money: Putting a Dollar Value on Financial Goals" that gives extensive coverage to calculating present and future values, but does not discuss the calculation or interpretation of N values for annuities. (DC/DD)

4. Kapoor, J. R., Dlabay, L. R., & Hughes, R. J. (2012). *Personal Finance* (10e). New York, NY: McGraw-Hill Irwin. The Appendix of Chapter 1 gives thorough coverage of how to calculate present and future values of lump-sums and annuities. It also covers how to calculate loan payments. However, the Appendix does not discuss the calculation or interpretation of N values. (DC/DD)

5. Keown, A. (2016). *Personal Finance: Turning Money into Wealth* (7e). Boston, MA: Pearson Education, Inc. In problem 24 on p. 87, the author asks the reader to solve for the amount of time to repay a debt with the payment being an annuity. The advice for noninteger N solutions is to "round up." (C/DD)

6. Madura, J. (2014). *Personal Finance* (5e). Boston, MA: Pearson Education, Inc. Chapter 3 titled "Applying Time Value Concepts," gives extensive coverage to calculating present and future values, but does not discuss the calculation or interpretation of N values for annuities. (DC/DD)

APPENDIX B



APPENDIX C

The formula for the future value of an ordinary annuity (FVA_N) is given as a geometric series, based on the annuity payment (PMT), the interest rate (I), and the number of periods (N):

$$FVA_N = PMT(1+I)^{N-1} + PMT(1+I)^{N-2} + \ldots + PMT(1+I)^{N-N} \tag{C.1}$$

Note that we have used upper case for variables I and N, consistent with the keys on row three of the Texas Instruments BA II Plus financial calculator (see Appendix B). Equation (C.1) is a geometric series, with each term on the right-hand side equal to the previous term multiplied by a common factor: $\frac{1}{1+I}$. This common factor is the one-period discount factor. A second equation is created by multiplying Equation (C.1) by this common factor:

$$FVA_N\left(\frac{1}{1+I}\right) = PMT(1+I)^{N-1}\left(\frac{1}{1+I}\right) + PMT(1+I)^{N-2}\left(\frac{1}{1+I}\right) + \ldots + PMT(1+I)^{N-N}\left(\frac{1}{1+I}\right),$$

and then collecting terms:

$$FVA_N\left(\frac{1}{1+I}\right) = PMT(1+I)^{N-2} + PMT(1+I)^{N-3} + \ldots + PMT(1+I)^{-1}. \tag{C.2}$$

By subtracting Equation (C.2) from Equation (C.1), we obtain:

$$FVA_N - FVA_N\left(\frac{1}{1+I}\right) = PMT(1+I)^{N-1} - PMT(1+I)^{-1},$$

which is rearranged to obtain an expression for FVA_N:

$$FVA_N = PMT\frac{(1+I)^N - 1}{I}. \tag{C.3}$$

Thus, the future value of an ordinary annuity is the product of two factors: PMT and $\frac{(1+I)^N - 1}{I}$.

A similar derivation finds the present value of an ordinary annuity is the product of two factors: PMT and $\frac{1-\frac{1}{(1+I)^N}}{I}$. We do not show it here, but factors for present and future values of an annuity due can also be found. Moreover, Equation (C.3) can be rearranged to find N:

$$N = \frac{\log\left(1+\frac{FVA \times I}{PMT}\right)}{\log(1+I)}. \tag{C.4}$$

Does Math Confidence Matter? How Student Perceptions Create Barriers to Success in Economics Classes

Abdullah Al-Bahrani, Whitney Buser, and Darshak Patel[1]

ABSTRACT

One of the most common obstacles in the economics classroom is facing students' disinclination to perform tasks requiring basic quantitative skills. Economics, relative to other disciplines, is particularly bridled by this challenge since mastery of economics requires sufficient mathematical proficiency to elicit anxiety and resistance in many students but is not widely regarded as math intensive enough to generate a selection effect of highly quantitative students. This paper attempts to measure undergraduate economics students' perceptions of their level of "mathiness" or mathematical abilities and anxieties and then identifies the impact of those perceptions on the students' performance in economics courses.

Introduction

At the graduate and professional levels, the field of economics is becoming increasingly quantitative. At the undergraduate level, introductory courses remain a staple in general education core curricula and enroll students of all mathematical levels. Despite its introductory nature, the basic mathematical principles of economics are still present in these courses. Thus, economics professors face a unique challenge: they are unable to avoid mathematical tools altogether as perhaps in humanities courses, nor are their students a self-selected quantitatively eager group as perhaps in STEM courses.

This dichotomy between the quantitative nature of the field and student attitudes often leads to complaints, unpreparedness, and anxiety on the part of the students. We postulate that lack of mathematical confidence is an important contributor to the occasionally negative classroom climate that arises in undergraduate economics and, perhaps, an explanatory variable for the gender gap in economics. Hence, we seek to determine if increased math confidence contributes to higher levels of economics classroom performance.

Previous work finds that perception of ability, even when not correlated with true ability, impacts confidence and success in academic settings (Everingham et al. 2013). It is important for professors who frequently encounter students with math anxiety to understand the role of mathematical confidence on performance. Understanding this role has pedagogical implications for all students, but might be especially helpful in shedding light on females' reluctance to participate in economics.

Economics education research thoroughly establishes a link between mathematical ability and performance in economics courses (Ballard and Johnson 2004; Elzinga and Melaugh 2009; Arnold and Straten 2012; Ullmer 2012) as well as gender asymmetries in economic inclination (Calkins and Welki 2006; Jensen and Owen 2001; Ashworth and Evans 1999). However, a gap in the literature remains. No

[1] Al-Bahrani: Assistant Professor of Economics, Haile/U.S Bank College of Business, Northern Kentucky University, Highland Heights, KY 40199, albahrania1@nku.edu; Buser: Assistant Professor of Business and Public Policy, Business and Public Policy Department, Young Harris College, Young Harris, GA 30582, wtdouglasbuser@yhc.edu; Patel (corresponding author): Senior Lecturer of Economics, Department of Economics, University of Kentucky, 225P Gatton College of Business and Economics., Lexington, KY 40506, darshak.patel@uky.edu. The authors wish to thank Kim Holder for her assistance with the survey and math questions. The authors would also like to thank the discussants from Conference on Teaching and Research in Economic Education, Southern Economic Association, Kentucky Economic Association and American Economic Association meetings. All errors, mistakes etc. in the article are exclusively our own.

studies yet have separated mathematical ability and mathematical confidence and examined the role of confidence.

This research expands on the existing literature to include confidence as a determinant of performance in the undergraduate economics classroom. We distinguish between math confidence and mathematical ability. Results indicate that while math ability matters for both men and women, math confidence (holding math ability constant) plays a large role in determining success in economics classes. Moreover, for women, their perceived confidence is a greater predictor of success than their actual math knowledge.

Literature Review

A clear link has been established between math ability and performance in undergraduate economics. Most studies show a significant positive relationship between math ACT/SAT scores and the Test of Understanding in College Economics (Becker 1997; Siegfried and Walstad 1998). Math ACT and SAT scores are found to be an essential determinant of performance when measuring math aptitude from a multidimensional perspective and are very important indicators of both introductory and intermediate performance (Ballard and Johnson 2004; Butler et al. 1994; Siegfried and Walstad 1998; Arnold and Straten 2012; Elzinga and Melaugh 2009; Ullmer 2012). These data seem to be robust to whether they are self-reported or administratively collected (Haley et al. 2010).

In economics, the quantitative SAT score gender differential can explain approximately 16% of the gender gap (Turner and Bowen 1999). The remaining gap is left to be explained. A clear distinction in preferences and anxiety over grades has been shown to affect some of this differential. Females are less likely than males to continue into a second economics course (Horvath et al. 1992). Females may be more responsive to poor grades and less likely to continue with the discipline if their perceived performance is weak (Rask and Tiefenthaler 2008; Horvath et al. 1992; Jensen and Owen 2001). Chizmar (2000) concludes that when controlling for grade differences the persistence gap between genders disappears.

Literature has hence shown that math ability has a significant impact in students’ selection and performance in economics. Furthermore, the expectation of failing an economic class is more likely to push away female students. Our study therefore looks at the impact of math confidence on undergraduate economics classes. Level of confidence may be relevant to the current gender gap in economics and may also be affected by math placement and timing.

Swope and Schmitt (2006) find that better quantitative skills result in a higher economics final grade and Bosshardt and Manage (2011) find that math aptitude exceeds math training in importance. Schuhmann et al. (2005) find that the fundamental skills needed are the ability to solve systems of equations, compute a percentage, and interpret increases and decreases on a graph. However, the timing of when math is introduced matters. Sabot and Wakeman-Linn (1991) and Anderson et al. (1994) both find that taking calculus in high school is a significant determinant of success in college level economics. Lagerlöf and Seltzer (2009) find that remedial math programs in college do not improve outcomes in introductory economics classes for students with low math aptitude.

Confidence in the classroom can often eclipse objective ability. Engagement and attitude towards learning can be affected by anxiety (Everingham et al. 2013). Lyons and Beilock (2012) find that the math anxious tend to have the same response to anticipating math as to anticipating pain. Chipman et al. (1992) find that math anxiety measures are a much stronger determinant of future career decisions than objective test scores measuring ability. Allgood et al. (2015) finds that, at the college level, course expectations regarding math requirements affect achievement.

There are asymmetric gender and income effects of confidence on performance. Gunderson et al. (2012) confirms that girls tend to have higher anxiety than boys when it comes to mathematics. Jackson and Leffingwell (1999) finds that K-12 teacher behavior is a prime determinant of math anxiety on both genders; low confidence is more easily disseminated to students in the early years (Geist 2015). Beilock et al. (2010) find evidence that the impact of an anxious teacher is worse on female students than on male students, while Mahigir et al. (2012) find that socioeconomic class, not gender, is the biggest driver of math anxiety differences.

While confidence and ability are distinct, recognizing one’s objective level of high ability may mitigate anxiety. Benedict and Hoag (2002) find proficiency measured by ACT Math score to have a negative effect on student anxiety. This may indicate that pedagogical efforts to increase mathematical skills and awareness of objective proficiency might increase confidence and thus performance in the classroom.

Therefore, our study hopes to evaluate the impact of both math ability and confidence at an undergraduate economic classroom by further teasing out results by gender.

Study Design

The goal of this study is to discover the impact of math confidence on performance in economics courses. The data represents individuals in principles of economics classes taught at two regional universities: Northern Kentucky University and University of Kentucky, and a liberal arts college: Young Harris College. The data represents principles of economics classes only.

Data for our study come from two sources: a 10-question math ability test and a perceptions survey. Data collection began during the second week of classes after the add/drop date. A 10-question math ability test was administered to test students' math knowledge (see Appendix Table 1 for the math questions). The questions were selected from SAT and ACT test banks and reflect concepts that are frequently covered in economics (Schuhmann et al. 2005).

In addition to the math test, students completed a survey that provided information on demographics, previous economics courses, and perceptions questions regarding their level of math confidence. Table 1 provides the summary statistics for selected areas of interest. To avoid framing biases, students were given the perceptions surveys before they were aware they would be taking a math quiz.

To assess confidence, math students were asked "How confident are you in your math abilities?" and asked to rate their answers on a Likert scale where 1= not at all confident and 5= very confident.[2] Although this measure is highly subjective, and we would not expect perceptions to be ordinal from participant to participant, these questions accurately gauge exactly what we want to measure: self-perception. We expect some students to be overly confident and some under confident relative to their abilities. For this reason, we measure objective ability separately and group students according to both measures in the analysis.[3]

For the purposes of this paper, confidence is simply a measure of self-perception. It is not ordinal and cannot be compared to another students' rating of their own confidence. It is simply a means of gathering information on student beliefs in regard to their ability to perform mathematical calculations. Motivation for this measurement came from the author's own experiences with students expressing anxiety about test questions involving "math" or statements of "I can't do math." Since this is a clear trend among students, this paper attempts to examine this trend among students. When students express these sentiments, they are certainly not expressing measurable or precise statements regarding their efficacy, their definition of confidence or anxiety, or their definition of "math." Therefore, the confidence measurement in this paper is purely an attempt to gather information on self-perceptions only.

Because confidence can cover a wide range of definitions, several survey questions were included to try and capture an understanding of self-perceptions of confidence. All students were given a pre-test of ability and ability is controlled for throughout the analysis below. This is an attempt to tease about self-beliefs and perceptions versus confidence based on ability. To tease out the effect of math specific confidence instead of general academic confidence, students were also asked to rate their verbal confidence. These two variables did not have a strong correlation, indicating that math confidence is distinct from other types of academic confidence.

In attempt to distinguish what students mean when they cite nervousness about "math," participants further investigated when participants first experience a lack of confidence. Although there were some respondents stating that they have had low confidence in math as long as they could remember, and some respondents stating that they have never felt a lack of confidence, most participants stated that their confidence began to wane in the middle school years. Finally, because confidence self-perceptions are so subjective, participants were also asked to rate belief in ability and belief in own ability relative to peers. After review of all confidence measures collected, the authors selected only the variable "math confidence" to be included in the main analysis of the paper.

Performance is measured by students' final course grade. Our teaching and assessment of principles of economics was similar across the three institutions. Each instructor gave three mid-semester exams and a

[2] Other variation of questions asked in other studies to measure "I feel confident in my abilities to solve mathematics problems."

[3] To ensure we were picking up a distinct measure of math confidence, we asked students several variations of the perceptions questions regarding their overall academic confidence, math confidence, verbal confidence, and assessment of these abilities relative to others.

comprehensive final. Exams at all institutions were based on multiple choice questions derived from verified test bank questions. Exams comprised the majority of the grade (80-90%) with quizzes, homework assignments, and participation making up the remainder (10-20%). Classroom instruction was based on lecture and active learning methods.

Descriptive Statistics

Of the 1193 students asked to complete the survey, 684 did so, for an overall participation rate of 59% (we exclude the dropped students when determining our participation rate). The participation rate by institution was 67%, 53%, and 82% for NKU, UK, and YHC respectively. There were 397 non-respondents, 38 students who chose to opt out of the study, 10 students who started the survey but never completed it, 37 students who completed the survey but were missing key variables, and 27 students who dropped the course.[4]

The main differences across institutions are reflected in the class sizes. Average class sizes are 400, 50, and 30 for UK, NKU, and YHC respectively. The higher participation rate at YHC is likely due to the closer student-faculty contact due to the small class sizes and residential liberal arts environment present at this institution. Moreover, the larger the institution and class size, the smaller the participation rate for our study. The strictness of IRB standards also likely affected participation rate. YHC students were able to receive their survey directly from their instructor and use a small amount of class time to complete it. To fulfill IRB requirements at NKU and UK, the authors' colleague had to summarize the research idea to students, provide a hard copy of the consent to sign and then email the students their surveys.

To examine the possibility of participation bias, we compared our individual population class averages to our samples class averages. For UK, the population average was 80.3% and the sample average was 82.7%. For NKU the measures were 76.4% and 82.73% respectively, and for YHC 81.2% and 80.1% respectively. Given the administration method, non-participation at YHC was almost universally due to absence for a school sanctioned event the day the survey was administered. At NKU and UK, students' non-participation was due to students being absent; the reasons for their absences were not documented since attendance was incentivized by quizzes but was not mandatory. In the case of UK, several students chose to complete the survey online but did not sign the consent forms provided in class. Overall, our sample is quite representative of our population.

Table 1 contains student-level descriptive statistics. The final sample contains 684 students who completed the survey, of which 50.1% were female. Students are classified as 35% freshmen, 40% sophomores, 20% juniors and 5% seniors. The average student age is 20.8 years. The sample is 77% white/non-Hispanic, 8% white/Hispanic, 7% Black, 4% Asian, and 2% other race. Approximately 15% transferred from other institutions, 66% are from within the respective institution's state, 7% are international students and 59% live on campus.

We asked students to identify what bracket of college GPA they fall under: 0% have a college GPA between 0-0.99, 1% have a college GPA between 1-1.99, 9% have a college GPA between 2-2.49, 21% have a college GPA between 2.5-2.99, 31% have a college GPA between 3-3.49 and 38% have a college GPA between 3.5-4. All but 2% of the students have some sort of math course either in high school or college and approximately 45% of the sample work while they attend college. On average, students attempt 5-6 courses per semester.

Students were asked to respond to the question "How confident are you with your overall mathematic abilities?" by using a Likert scale rating where 1=not at all confident and 5= very confident. Students on average were fairly confident in their math ability (3.434). The average final course grade is 82.52 and the average math quiz grade is 51.98. Math confidence and math ability (as measured by the quiz grade) were not correlated.

For further analysis, we replicate Allgood and Walstad (2016) for more a detailed understanding. Having two key math variables, we split each measure in terms of high and low categories. The math ability (quiz) measure is split using the mean of the composite score. Any student with a score greater than the mean of 51.98 is placed in the "High Ability" (53%) category and the rest in the "Low Ability" category.

[4] The following provides the breakdown by institution: Total Enrolled: NKU - 215, UK – 848, YHC – 130. Completion (Lower Level Classes): NKU – 140, UK – 437, YHC – 107.

Table 1: Descriptive Statistics

	All		NKU		UK		YHC	
Variable	**Mean**	Std. Dev	**Mean**	Std. Dev	**Mean**	Std. Dev	**Mean**	Std. Dev
Final Course Grade	**82.520**	9.848	**83.745**	9.518	**82.730**	9.319	**80.093**	11.861
Math Quiz Grade	**51.982**	21.245	**53.571**	22.507	**53.918**	20.833	**41.700**	18.149
Math Confidence	**3.434**	1.056	**3.688**	0.995	**3.387**	1.098	**3.299**	0.903
Age	**20.814**	2.260	**21.938**	2.836	**20.382**	2.098	**21.122**	1.419
HS GPA	**3.615**	1.194	**3.554**	0.577	**3.654**	1.359	**3.491**	0.414
Semester Classes	**5.136**	0.753	**5.119**	0.723	**5.123**	0.770	**5.206**	0.723
Female	**0.501**	0.500	**0.464**	0.501	**0.517**	0.500	**0.486**	0.502
Male	**0.497**	0.500	**0.529**	0.501	**0.483**	0.500	**0.514**	0.502
White/Non Hispanic	**0.774**	0.419	**0.793**	0.407	**0.765**	0.425	**0.785**	0.413
White/Hispanic	**0.086**	0.281	**0.037**	0.190	**0.090**	0.287	**0.131**	0.339
Black	**0.077**	0.267	**0.104**	0.306	**0.071**	0.258	**0.065**	0.248
Asian	**0.044**	0.205	**0.052**	0.223	**0.052**	0.223	**0.000**	0.000
Other Race	**0.020**	0.139	**0.015**	0.121	**0.021**	0.145	**0.019**	0.136
Transfer	**0.150**	0.358	**0.261**	0.441	**0.131**	0.338	**0.084**	0.279
Instate	**0.656**	0.475	**0.706**	0.457	**0.615**	0.487	**0.757**	0.431
International Student	**0.069**	0.254	**0.080**	0.272	**0.048**	0.215	**0.142**	0.350
On Campus	**0.587**	0.493	**0.284**	0.452	**0.597**	0.491	**0.925**	0.264
Private School	**0.215**	0.411	**0.224**	0.418	**0.233**	0.423	**0.131**	0.339
First Economc Class	**0.707**	0.455	**0.470**	0.501	**0.844**	0.364	**0.467**	0.501
College GPA (0-0.99)								
College GPA (between 1-1.99)	**0.014**	0.116	0	0	**0.019**	0.137	**0.009**	0.097
College GPA (between 2-2.49)	**0.091**	0.288	**0.067**	0.251	**0.102**	0.304	**0.075**	0.264
College GPA (between 2.5-2.99)	**0.206**	0.405	**0.216**	0.413	**0.193**	0.395	**0.243**	0.431
College GPA (between 3-3.49)	**0.307**	0.462	**0.276**	0.449	**0.331**	0.471	**0.252**	0.436
College GPA (between 3.5-4)	**0.383**	0.486	**0.440**	0.498	**0.355**	0.479	**0.421**	0.496
Job	**0.452**	0.498	**0.754**	0.432	**0.373**	0.484	**0.383**	0.488
Math	**0.984**	0.126	**0.986**	0.119	**0.982**	0.134	**0.991**	0.097
Freshman	**0.351**	0.477	**0.022**	0.146	**0.530**	0.500	**0.047**	0.212
Sophmore	**0.401**	0.490	**0.522**	0.501	**0.297**	0.458	**0.664**	0.475
Junior	**0.202**	0.402	**0.312**	0.465	**0.154**	0.362	**0.252**	0.436
Senior	**0.047**	0.212	**0.145**	0.353	**0.018**	0.135	**0.037**	0.191
Spring	**0.613**	0.488	**0.500**	0.502	**0.703**	0.458	**0.393**	0.491
Total # of Students	684		140		437		107	
Percent of students	100%		20%		64%		16%	

Similarly, the confidence self-ratings were split as "High Confidence" (48%) for those who responded with a "4" or "5" and "Low Confidence" (52%) for those who responded with a "1", "2", or "3". We then created for math perception groups: a) High Ability and High Confidence (32%) (b) High Ability and Low Confidence (21%) (c) Low Ability and High Confidence (17%) (d) Low Ability and Low Confidence (30%). Table 2 provides a more detailed descriptive of these variables by institution and gender.

We find several similarities and differences among students at the three institutions.[5] There are similarities in the average ages between the three institutions (range from 20.3 to 21.93). We find that there are more men in the NKU (52.9%) and YHC (51.4%) sample (compared UK (48.3%)). The majority of the students at all 3 institutions are white (more than 75%). While black students account for the next largest race proportion at NKU, Hispanics represent the next largest group at UK and YHC.

All three institutions have a large in-state student base. The YHC sample (92.5%) has a large number of students who live on campus followed by UK (59.7%). Only 28% of NKU sample live on campus. There is a big variation between the samples in relation to working: almost 75% of the NKU sample work compared to 37% at UK and 38% at YHC.

We find a noticeable difference in class standings between UK and the other institutions. At UK, the sample is heavily weighted towards freshmen and sophomores. At the other two campuses, the sample is heavily weighted towards sophomores and juniors. Other interesting observations include massive differences between the institutions on whether the current class is the students' first economic class. More

[5] We tested the differences between the means of the characteristics with each institution and with the overall average as well. There were some statistically significant (5% level) differences for some covariates. These t-stats are available upon request. By surveying institutions with different student bodies, we hope to gain insight from a diversity of attitudes and perceptions from different students and different institutions.

students have had prior economics at UK than YHC and NKU. Students at NKU were more confident about their math ability and perceived math aptitude when compared to UK and YHC. Students from NKU and UK received higher final grades and math quiz scores compared to YHC.

Table 2: Actual Ability and Perceived Confidence

	All		NKU		UK		YHC	
Variable	**Mean**	Std. Dev	**Mean**	Std. Dev	**Mean**	Std. Dev	**Mean**	Std. Dev
High Math Ability	**0.53**	0.50	**0.54**	0.50	**0.57**	0.50	**0.39**	0.49
High Confidence	**0.48**	0.50	**0.61**	0.49	**0.46**	0.50	**0.40**	0.49
High Ability & High Confidence	**0.32**	0.47	**0.36**	0.48	**0.33**	0.47	**0.20**	0.40
High Ability & Low Confidence	**0.21**	0.41	**0.17**	0.38	**0.23**	0.42	**0.20**	0.40
Low Ability & High Confidence	**0.17**	0.37	**0.24**	0.43	**0.13**	0.34	**0.21**	0.41
Low Ability & Low Confidence	**0.30**	0.46	**0.22**	0.42	**0.30**	0.46	**0.40**	0.49

Methodology

Our main evaluation criterion is the final course grade (Final Grade in percentages) calculated without any curve (un-curved). Thus, the dependent variable is measured on a 0 to 100 scale. The baseline ordinary least squares (OLS) regression specification is:

$$Final\ Grade_{ij} = \beta_0 + \beta_1 MA_1 + \beta_2 MC_2 + \beta_3 X_3 + \cdots . + \beta_n X_n + \varepsilon$$

where X is a set of individual specific control variables and ε_j is the stochastic error term. This is a standard education production function. Our key variables of interests are: MA and $MC,$ our new measures of math ability and math confidence. The measure of math confidence is collected through the survey conducted at the beginning of the semester and math ability through the math quiz administered during the second week of classes as explained in our data section.[6]

For our choice of control variables, we relied primarily on the educational outcomes literature. The set of control variables includes demographic characteristics (age, race, gender, academic standing), dummy variables for whether or not the individual attended a private high school, whether the student is taking economics for the first time, whether the student is an international student, whether the students transferred to the institution, the self-reported number of hours worked per week, number of courses taken by the students during the semester, the total number of high-school and college math courses the student has taken, and self-reported cumulative college GPA.[7]

[6] For robustness checks we reran the model replacing math confidence with perceived math aptitude. To determine this key variable student were asked not about their confidence levels but about how well their abilities rank in relation to their peers. Results were robust to wording of the perception question and can be found in Tables 3 and 4 of the Appendix.

[7] For robustness checks the general specification of the regression was rerun with a different dependent variable. Final exam score replaced final course average. Initial thinking was that the cumulative final exam would be a better measure of performance as it only measures knowledge and does not include any non-test grade factors. However, after close examination of final exam scores, they did not correlate with previous exam scores. Possibly, this is due to students focusing on other exams during exam week, sleeplessness, discouragement, lack of motivation for students who are already earning good grades in the course, and other factors. Results can be provided upon requests.

Results

This section tests the impact of math confidence controlling for mathematical ability in undergraduate economics. The general regression model attempts to explain the final grade in economics classes as a function of math ability, math confidence, and educational outcome variables. After considering duplicates, incomplete math quizzes and survey responses, and students who registered for upper level classes, the total number of observations is reduced to 684.

Table 3 displays the results for the initial regression specification. Iteration 1 adds math ability only to the control vector of education outcomes variables. Iteration 2 adds math confidence only. Iteration 3 includes both math ability and math confidence. Iterations 4 and 5 are a replication of iteration 3 for males and females respectively.

In every specification, math ability is significant and positive. This finding is consistent with previous literature: stronger math skills increase economic performance. Holding all else constant, scoring one additional correct question on the SAT based math quiz is associated with a 1.2 percentage point increase in the overall course grade on average. Math confidence is positively associated with performance for the whole sample, even when controlling for given levels of ability. Splitting the sample by gender shows that confidence, when controlling for given levels of abilities, is a significant factor in success.[8]

As expected, students' cumulative GPA is also positively correlated with student performance within class for every specification. Specification (1), (2), (3) and (4) found some race effect where Black students performed overall worse than White non-Hispanic students. Juniors performed better than freshman by at least 2-5 percentage points as evident by specification (2) and (5) (this variable was insignificant in specification 1, 3 & 5). Lastly, students at YHC scored lower than students at UK for all 5 specifications. More interesting in our results is that the number of math courses taken is statistically insignificant in all specifications. Perhaps this is because math ability is controlled for in the regressions. Another explanation is that the number of math classes is not necessarily a good measure of their math ability. A better measure to indicate their ability or the quality of their consumption of math would be the highest level of math class taken.

To gain a deeper understanding of the impact of math ability and confidence on success in economics courses, we follow the direction of Allgood and Walstad (2016) and split students in four categories: (a) High Ability and High Confidence (b) High Ability and Low Confidence (c) Low Ability and High Confidence (d) Low Ability and Low Confidence. This provides a much cleaner understanding as it provides an estimate on performance based on the direction of movement from high to low math ability and math confidence. Using Low Ability and Low Confidence as a comparison group, we test to see how other groups perform. Table 4 provides the analysis for the full sample and then broken down for male and female, respectively.

Intuitively, individuals with high mathematical ability and high mathematical confidence perform the best out of the comparison groups. This result is consistent with previous results. Those with high ability and low confidence perform better than those with low ability and low confidence, yet the marginal effect is not as great as the group with both high ability and high confidence, suggesting that confidence still matters even among individuals with high levels of achievements. Among those with low ability and high confidence, the marginal effect was only significant for women, suggesting that confidence is a key factor in success for women with weak math skills.

[8] In the above specifications math confidence entered the regressions as a categorical variable (1 through 5 with 5 being the highest level of confidence.) Because there is reason to believe that the difference between a rating of 1 and 2 and 2 and 3 is not demonstratively the same, the analysis is repeated changing math confidence into a dichotomous variable. Respondents answering they were 4-confident or 5-very confident were given placed into the "high confidence" category. Respondents were answered 3-neutral, 2- not confident, or 1-not at all confident were placed in the "low confidence" category. A similar math variable was created to differentiate between "High Ability" and "Low Ability." Students whose score on the math quiz was greater than the mean were placed in the "High Ability Category" and those who scored below the mean were placed in the "Low Ability Category." Regressions were re-estimated replacing the measure of math confidence and math ability with the variables created as explained above. All results hold as before. However, in this scenario, confidence does matter for male students and the ability co-efficient are much closer for both male and female students. Results of these regressions can be found in Table 2 of the appendix.

Table 3: Results Using Composite Scores and Self-Rating Confidence Responses

VARIABLES	(1) Final Grade	(2) Final Grade	(3) Final Grade	(4) Final Grade (M)	(5) Final Grade (F)
Math Quiz	0.12***		0.11***	0.08***	0.13***
	(0.02)		(0.02)	(0.02)	(0.03)
Math Confidence		1.64***	1.13***	1.12***	0.89***
		(0.32)	(0.33)	(0.43)	(0.52)
Male	1.93***	2.66***	1.51***		
	(0.68)	(0.68)	(0.68)		
Age	-0.04	-0.08	0.03	0.05	0.04
	(0.18)	(0.19)	(0.18)	(0.26)	(0.27)
White (Hispanic)	-0.94	-0.68	-0.80	-1.48	-0.80
	(1.16)	(1.16)	(1.16)	(1.35)	(1.98)
Black	-2.92***	-4.23***	-3.15***	-5.50***	-1.90
	(1.26)	(1.25)	(1.25)	(1.80)	(1.79)
Asian	-2.28	-1.85	-2.03	-4.51***	-0.67
	(1.73)	(1.75)	(1.72)	(2.57)	(2.39)
Other Race	1.09	0.25	1.20	1.64	-0.06
	(2.31)	(2.28)	(2.29)	(2.48)	(4.53)
Sophomore	0.32	0.44	0.63	1.30	-0.35
	(0.88)	(0.89)	(0.85)	(1.13)	(1.34)
Junior	1.34	1.93***	1.62	4.56***	-1.28
	(1.13)	(1.11)	(1.11)	(1.45)	(1.79)
Senior	1.71	1.00	1.28	1.54	1.76
	(1.92)	(1.91)	(1.91)	(2.22)	(3.61)
Private School	-0.49	-0.16	-0.33	-1.04	0.08
	(0.80)	(0.81)	(0.80)	(0.97)	(1.30)
Job	-1.09	-1.20***	-0.90	-1.10	-0.95
	(0.69)	(0.69)	(0.68)	(0.90)	(1.04)
First Time Econ	-2.16***	-2.75***	-2.04***	-2.42***	-2.50***
	(0.78)	(0.78)	(0.77)	(1.00)	(1.19)
Number of Courses	0.47	0.67	0.36	0.58	0.18
	(0.45)	(0.45)	(0.44)	(0.60)	(0.68)
International Student	-0.22	-0.39	-0.46	-0.94	0.24
	(1.55)	(1.51)	(1.54)	(2.10)	(2.34)
Transfer	-1.54	-2.14***	-1.46	-2.56***	-0.72
	(1.03)	(1.02)	(1.02)	(1.21)	(1.84)
# of Math Classes	-2.02	-3.26	-3.32	-3.29	-4.88
	(3.17)	(3.25)	(3.16)	(4.36)	(4.72)
Cumulative GPA	3.87***	4.05***	3.77***	3.26***	4.17***
	(0.33)	(0.33)	(0.33)	(0.39)	(0.55)
NKU	-0.09	-0.65	-0.47	-3.09***	1.43
	(0.97)	(0.99)	(0.97)	(1.24)	(1.53)
YHC	-3.37***	-4.72***	-3.27***	-4.32***	-3.10***
	(1.08)	(1.06)	(1.06)	(1.39)	(1.68)
Spring	-1.00	-1.19***		-1.31	-0.76
	(0.71)	(0.71)		(0.90)	(1.09)
Constant	55.67***	55.81***	53.02***	57.95***	52.97***
	(5.71)	(5.79)	(5.68)	(8.17)	(8.35)
Observations	616	642	615	301	313
R-squared	0.37	0.34	0.38	0.46	0.36
Prob > F-Stat	0.00	0.00	0.00	0.00	0.00

Standard errors in parentheses, * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 4: Results Using Categorical Groups: High to Low Ability and Confidence

VARIABLES	(1) Final Grade	(2) Final Grade (M)	(3) Final Grade (F)
High Ability and High Confidence	6.19***	5.23***	6.27***
	(0.83)	(1.10)	(1.30)
High Ability and Low Confidence	4.29***	3.63***	4.49***
	(0.91)	(1.24)	(1.38)
Low Ability and High Confidence	2.64***	1.14	3.26***
	(0.98)	(1.34)	(1.46)
Male	2.19***		
	(0.66)		
Age	-0.04	0.03	-0.04
	(0.18)	(0.26)	(0.27)
White (Hispanic)	-0.55	-1.31	-0.34
	(1.14)	(1.32)	(1.97)
Black	-3.63***	-5.36***	-2.69
	(1.23)	(1.82)	(1.73)
Asian	-2.01	-4.84***	-0.15
	(1.71)	(2.52)	(2.40)
Other Race	-0.19	1.49	-1.88
	(2.23)	(2.51)	(4.15)
Sophomore	0.42	1.57	-0.45
	(0.87)	(1.13)	(1.34)
Junior	1.49	4.63***	-1.17
	(1.09)	(1.39)	(1.76)
Senior	1.40	2.30	0.70
	(1.87)	(2.18)	(3.46)
Private School	-0.15	-0.98	0.24
	(0.79)	(0.97)	(1.28)
Job	-1.27***	-1.20	-1.46
	(0.68)	(0.88)	(1.04)
First Time Econ	-2.36***	-2.65***	-2.48***
	(0.77)	(1.00)	(1.19)
Number of Courses	0.63	0.89	0.43
	(0.44)	(0.59)	(0.67)
International Student	-0.46	-0.35	0.16
	(1.49)	(2.00)	(2.28)
Transfer	-1.56	-2.75***	-0.65
	(1.01)	(1.20)	(1.81)
# of Math Classes	-2.55	-2.48	-4.69
	(3.19)	(4.38)	(4.74)
Cumulative GPA	4.06***	3.53***	4.51***
	(0.32)	(0.38)	(0.54)
NKU	-0.11	-2.66***	1.65
	(0.97)	(1.24)	(1.52)
YHC	-3.65***	-4.10***	-3.71***
	(1.05)	(1.36)	(1.64)
Spring	-0.57	-0.75	-0.33
	(0.70)	(0.91)	(1.08)
Constant	56.97***	59.34***	57.51***
	(5.66)	(8.05)	(8.25)
Observations	643	315	327
R-squared	0.37	0.44	0.34
Prob > F-Stat	0.00	0.00	0.00

Standard errors in parentheses, * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Our results consistently show that mathematical ability matters for both men and women. For women, however, confidence also plays a role. Women with higher confidence do better than women with lower confidence, holding all else constant. Results indicate that confidence in students has a positive impact on their class performance. From a pedagogical perspective, spending time combating math disinclination and building confidence is important. Such pedagogy can help attract more female students to the economics discipline. It is important for future research to discover how to build mathematical confidence in students. This research should note the differences in math anxieties between the genders.

Limitations

All sample collection followed institution specific Institutional Review Board standards. Due to differences in school culture and IRB requirements, response rates varied across institutions and generated variation in overall response rate. One concern with this is the issue of participation bias. Although this does not seem to be the case for UK and YHC, the difference in the sample and population means for NKU seems to indicate that high achieving students were more likely to participate than low performing students. Although a negative for study, we believe this helps support our findings. Our study shows that even for motivated students, ability and confidence matters. We believe these factors could have a much bigger impact for the low motivated students with ability and confidence concerns.

Other concerns that may influence the overall significance of the research are issues that are typical in survey datasets. More importantly, the issue of omitted variables. We feel the variables we collect from our students are sufficient for a standard education production function. With the data as is, it is highly doubtful that any of the econometric techniques would be feasible in solving the issue. A common method to use is an Instrumental Variable (IV) approach or even a matching technique. With our data and sample, these are highly unlikely. A robustness check using high school GPA rather than college GPA to control for ability and for any endogeneity issues is also estimated and results hold. We use the typical education production function (Emerson and Taylor 2004) and insert math confidence as a determinant of outcome. Our approach falls in line with other research examining the determinants of success in the economics classroom.

Conclusion

The determinants of what helps students succeed in economics courses and the economics major has received a lot of attention. We examine the impact of both objective math ability and perceived math confidence on student learning outcomes. Due to the mathematical nature of economics, it is presumed that math proficiency is positively correlated with success at the introductory level courses and in the major. We find that students' current quantitative aptitude does impact student-learning outcomes positively. Furthermore, student confidence in their math abilities is a large predictor of their success in economics relative to their actual math ability. While we find positive associations between confidence and performance for the whole sample, further investigation reveals that this effect is being driven by females only.

The implication of these findings affects the way economics is taught. To help increase student learning of economics, instructors need to devote some effort to encouraging, motivating, and building mathematical confidence, especially for female students. This gives way to further research. What is the best way to increase math confidence? Do these approaches need to differ by gender? Would this increase in mathematical confidence have further positive effects on the educational process? If math differences are present between genders and math ability and math confidence contribute to success in economics, our findings may also have implications in explaining the gender gap in economics and eventually seeking to close it as well.

Furthermore, the educational psychology literature distinguishes between many self-belief concepts but two of them are more relevant to this paper--confidence and self-efficacy. The two are interrelated, but self-efficacy is a domain-specific construct, which refers to beliefs about one's ability to perform specific tasks. Confidence is a more general personality trait that refers to one's belief about one's self-worth and likelihood of success. In recent years, distinction between confidence and self-efficacy has increasingly been made in some business/economics literature. To measure efficacy, one could survey students on 3-5 questions asking about a student's belief in performing specific tasks in algebra or geometry that would be useful in an economics course. Self-efficacy matters in the math literature and should also matter in Economics. The trick is to be able to separate the impacts of each.

It is a common assumption among economic educators that students are anxious about the math and graphing component in economics courses. Our study was motivated by this thought and looked to possibly explain a reason for the lower numbers of female economics majors. This study is new to the field of economics. Our hope is that this will kick-start more in-depth discussion on students' math confidence or efficacy on students' success in economics courses.

References

Allgood, Sam, and William B. Walstad. 2016. "The Effects of Perceived and Actual Financial Literacy on Financial Behaviors." *Economic Inquiry* 54: 675-697.

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

Anderson, Gordon, Dwayne Benjamin, and Melvyn A. Fuss. 1994. "The Determinants of Success in University Introductory Economics Courses." *The Journal of Economic Education* 25: 99-119.

Arnold, Ivo J. M., and Jerry T. Straten. 2012. "Motivation and Math Skills as Determinants of First-Year Performance in Economics." *The Journal of Economic Education* 43: 33-47.

Ashworth, John, and Lynne Evans. 1999. "Lack of Knowledge Deters Women from Studying Economics." *Educational Research* 41: 209-227.

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

Becker, William E. 1997. "Teaching Economics to Undergraduates." *Journal of Economic Literature* 35: 1347-1373.

Beilock, Sian L., Elizableth A. Gunderson, Gerardo Ramirez, and Susan C. Levine. 2010. "Female Teachers' Math Anxiety Affects girls' Math Achievement." *Proceedings of the National Academy of Sciences* 107: 1060–1063.

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

Bosshardt, William, and Neela Manage. 2011. "Does Calculus Help in Principles of Economics Courses? Estimates Using Matching Estimators." *The American Economist* 56: 29-37.

Butler, J. S., T. Aldrich Finegan, and John J. Siegfried. 1994. "Does More Calculus Improve Student Learning in Intermediate Micro and Macro Economic Theory?" *American Economic Review* 84: 206-210.

Calkins, Lindsay N., and Andrew Welki. 2006. "Factors that Influence Choice of Major: Why Some Students Never Consider Economics." *International Journal of Social Economics* 33: 547–564.

Chipman, Susan F., David H. Krantz, and Rae Silver. 1992. "Mathematics Anxiety and Science Careers among Able College Women." *Psychological Science* 3: 292-296.

Chizmar, John F. 2000. "A Discrete-Time Hazard Analysis of the Role of Gender in Persistence in the Economics Major." *The Journal of Economic Education* 31: 107-118.

Elzinga, Kenneth G., and Daniel O. Melaugh. 2009. "35,000 Principles of Economics Students: Some Lessons Learned." *Southern Economic Journal* 76: 32-46.

Emerson, Tisha L. N., and Beck A. Taylor. 2004. "Comparing Student Achievement across Experimental and Lecture-Oriented Sections of a Principles of Microeconomics Course." *Southern Economic Journal* 70: 672-693.

Everingham, Yvette, Emily Gyuris, and Justin Sexton. 2013. "Using Student Feedback to Improve Student Attitudes and Mathematical Confidence in a First Year Interdisciplinary Quantitative Course: From the Ashes of Disaster!" *International Journal of Mathematical Education in Science and Technology* 44: 877-892.

Geist, Eugene. 2015. "Math Anxiety and the 'Math Gap': How Attitudes Toward Mathematics Disadvantages Students as Early as Preschool." *Education* 135: 328-336.

Gunderson, Elizabeth A., Gerardo Ramirez, Susan C. Levine, and Sian L. Beilock. 2012. "The Role of Parents and Teachers in the Development of Gender-Related Math Attitudes." *Sex Roles* 66: 153-166.

Haley, M. Ryan, Marianne F. Johnson, and M. Kevin McGee. 2010. "A Framework for Reconsidering the Lake Wobegon Effect." *The Journal of Economic Education* 41: 95-109.

Horvath, Jane, Barbara Q. Beaudin, and Sheila P. Wright. 1992. "Persisting in the Introductory Economics Course: An Exploration of Gender Differences." *The Journal of Economic Education* 23: 101-108.

Jackson, Carol D., and R. Jon Leffingwell. 1999. "The Role of Instructors in Creating Math Anxiety in Students from Kindergarten through College." *The Mathematics Teacher* 92: 583-586.

Lagerlöf, Johan N. M., and Andrew J. Seltzer. 2009. "The Effects of Remedial Mathematics on the Learning of Economics: Evidence from a Natural Experiment." *The Journal of Economic Education* 40: 115-137.

Jensen, Elizabeth J., and Ann L. Owen. 2001. "Pedagogy, Gender, and Interest in Economics." *The Journal of Economic Education* 32: 323-343.

Lyons, Ian M., and Sian L. Beilock. 2012. "When Math Hurts: Math Anxiety Predicts Pain Network Activation in Anticipation of Doing Math." *PLoS One* 7.

Mahigir, Faroogh, Ali Khanehkeshi, and Ayatollah Karimi. 2012. "Psychological Treatment for Pain among Cancer Patients by Rational-Emotive Behavior Therapy—Efficacy in both India and Iran." *Asian Pacific Journal of Cancer Prevention* 13: 4561–4565.

Rask, Kevin, and Jill Tiefenthaler. 2008. "The Role of Grade Sensitivity in Explaining the Gender Imbalance in Undergraduate Economics." *Economics of Education Review* 27: 676-687.

Sabot, Richard, and John Wakeman-Linn. 1991. "Grade Inflation and Course Choice." *Journal of Economic Perspectives* 5: 159-170.

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

Siegfried, John J., and William B. Walstad. 1998. "Research on Teaching College Economics." In *The Principles of Economics Course: A Handbook for Instructors*, edited by William B. Walstad and Phillip Saunders. New York: McGraw Hill.

Swope, Kurtis J., and Pamela M. Schmitt. 2006. "The Performance of Economics Graduates over the Entire Curriculum: The Determinants of Success." *The Journal of Economic Education* 37: 387-394.

Turner, Sarah E., and William G. Bowen. 1999. "Choice of Major: The Changing (Unchanging) Gender Gap." *Industrial & Labor Relations Review* 52: 289-313.

Ullmer, James. 2012. "Student Characteristics, Peer Effects and Success in Introductory Economics." *Journal of Economics and Economic Education Research* 13: 79.

APPENDIX

Appendix Table 1: Math Quiz

Quiz 1

Student Instructions: you have 15 minutes to take the following quiz. No calculators or other electronic devices are allowed. Please indicate your answer for each question on the blank provided beside each question number. Make sure to include your student ID on your scantron.

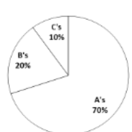
___ 1. A special lottery is to be held to select the student who will live in the only deluxe room in a dormitory. There are 100 seniors, 150 juniors, and 200 sophomores who applied. Each senior's name is placed in the lottery 3 times, each junior's name 2 times, and each sophomore's name 1 time. What is the probability that a senior's name will be chosen?

A. 1/8
B. 2/9
C. 2/7
D. 3/8
E. 1/2

___ 2. A car averages 27 miles per gallon. If gas costs $4.04 per gallon, which of the following is closest to how much the gas would cost for this car to travel 2,727 typical miles?

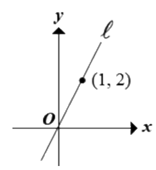
A. $ 44.44
B. $109.08
C. $118.80
D. $408.04
E. $444.40

___ 3. The distribution of Jamal's high school grades by percentage of course credits is given in the circle graph below. What is Jamal's grade point average if each A is worth 4 points; each B is worth 3 points; and each C is worth 2 points?



A. 3.0
B. 3.4
C. 3.6
D. 3.7
E. Cannot be determined from the given information

___ 4. In the x y-coordinate plane below, line contains the points (0, 0) and (1, 2). If line (not shown) contains the point (0, 0) and is perpendicular to , what is an equation of ?



A. $= -\frac{1}{2}$

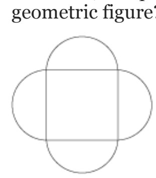
B. $= -\frac{1}{2} \quad + 1$

C. $= -$

D. $= - \quad + 2$

E. $= - 2$

___ 5. The geometric figure shown below consists of a square and 4 semicircles. The diameters of the semicircles are the sides of the square, and each diameter is 10 centimeters long. Which of the following is the closest approximation of the total area, in square centimeters, of this geometric figure?



A. 100
B. 160
C. 260
D. 400
E. 730

___ 6. A DVD player with a list price of $100 is marked down 30%. If John gets an employee discount of 20% off the sale price, how much does John pay for the DVD player?

A. $86.00
B. $77.60
C. $56.00
D. $50.00
E. $44.00

7. Based on the system of equations below, what is the value of the product *xy*?

$$4x - y = 3y + 7$$
$$x + 8y = 4$$

A. – 3/2
B. 1/4
C. 1/2
D. 11/9

8. What value of x satisfies both of the equations below?

$$|4x - 7| = 5$$
$$|3 - 8x| = 1$$

Please record your answer: *x* = _____
A. ¼
B. ½
C. 3
D. 5

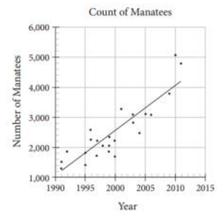
9. A survey was conducted among a randomly chosen sample of U.S. citizens about U.S. voter participation in the November 2012 presidential election. The table below displays a summary of the survey results. According to the table, for which age group did the greatest percentage of people report that they had voted?

Reported Voting by Age (in thousands)

	VOTED	DID NOT VOTE	NO RESPONSE	TOTAL
18- to 34-year-olds	30,329	23,211	9,468	63,008
35- to 54-year-olds	47,085	17,721	9,476	74,282
55- to 74-year-olds	43,075	10,092	6,831	59,998
People 75 years old and over	12,459	3,508	1,827	17,794
Total	132,948	54,532	27,602	215,082

A. 18- to 34-year-olds
B. 35- to 54-year-olds
C. 55- to 74-year-olds
D. People 75 years old and over

10. The scatterplot below shows counts of Florida manatees, a type of sea mammal, from 1991 to 2011. Based on the line of best fit to the data shown, which of the following values is closest to the average yearly increase in the number of manatees?



A. 0.75
B. 75
C. 150
D. 750

Appendix Table 2: Results Using Two Categorical Groups: High vs. Low Ability & High vs. Low Confidence

VARIABLES	(1) Final Grade	(2) Final Grade	(3) Final Grade	(4) Final Grade (M)	(5) Final Grade (F)
High Math Ability	4.50***		3.95***	3.86***	3.85***
	(0.67)		(0.68)	(0.88)	(1.05)
High Confidence		3.17***	2.24***	1.42***	2.56***
		(0.67)	(0.67)	(0.84)	(1.07)
Male	2.41***	2.90***	2.18***		
	(0.67)	(0.67)	(0.66)		
Age	-0.09	-0.11	-0.05	0.03	-0.04
	(0.18)	(0.19)	(0.18)	(0.26)	(0.27)
White (Hispanic)	-0.62	-0.73	-0.54	-1.29	-0.20
	(1.15)	(1.17)	(1.14)	(1.32)	(1.96)
Black	-3.49***	-4.17***	-3.57***	-5.40***	-2.63
	(1.23)	(1.25)	(1.22)	(1.81)	(1.73)
Asian	-2.30	-1.81	-1.96	-4.91***	-0.15
	(1.72)	(1.75)	(1.71)	(2.51)	(2.39)
Other Race	-0.44	0.34	-0.16	1.50	-1.62
	(2.25)	(2.29)	(2.23)	(2.50)	(4.13)
Sophomore	0.28	0.56	0.38	1.60	-0.50
	(0.87)	(0.89)	(0.87)	(1.12)	(1.34)
Junior	1.32	2.18***	1.47	4.63***	-1.17
	(1.10)	(1.11)	(1.09)	(1.39)	(1.76)
Senior	1.69	1.48	1.37	2.34	0.77
	(1.88)	(1.91)	(1.86)	(2.17)	(3.46)
Private School	-0.27	-0.21	-0.15	-0.95	0.32
	(0.79)	(0.81)	(0.79)	(0.96)	(1.27)
Job	-1.29***	-1.30***	-1.25***	-1.20	-1.43
	(0.68)	(0.69)	(0.68)	(0.88)	(1.03)
First Time Econ	-2.36***	-2.69***	-2.37***	-2.64***	-2.52***
	(0.77)	(0.79)	(0.77)	(0.99)	(1.18)
Number of Courses	0.71	0.69	0.62	0.90	0.41
	(0.44)	(0.45)	(0.44)	(0.59)	(0.67)
International Student	-0.02	-0.52	-0.38	-0.43	0.16
	(1.49)	(1.52)	(1.48)	(1.98)	(2.28)
Transfer	-1.59	-2.36***	-1.56	-2.75***	-0.70
	(1.02)	(1.02)	(1.01)	(1.20)	(1.81)
# of Math Classes	-1.64	-3.19	-2.56	-2.59	-4.90
	(3.20)	(3.26)	(3.18)	(4.36)	(4.73)
Cumulative GPA	4.12***	4.14***	4.05***	3.54***	4.49***
	(0.32)	(0.33)	(0.32)	(0.38)	(0.54)
NKU	0.31	-0.53	-0.07	-2.68***	1.73
	(0.97)	(0.99)	(0.97)	(1.24)	(1.51)
YHC	-3.65***	-4.68***	-3.62***	-4.11***	-3.63***
	(1.06)	(1.06)	(1.05)	(1.36)	(1.63)
Spring	-0.50	-1.08	-0.58	-0.74	-0.33
	(0.71)	(0.71)	(0.70)	(0.91)	(1.08)
Constant	56.99***	59.75***	57.30***	59.14***	58.12***
	(5.67)	(5.75)	(5.63)	(8.00)	(8.20)
Observations	643	643	643	315	327
R-squared	0.36	0.33	0.37	0.44	0.34

Standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Appendix Table 3: Results Using Composite Scores and Self-Rating Aptitude Responses

VARIABLES	(1) Final Grade	(2) Final Grade	(3) Final Grade	(4) Final Grade (M)	(5) Final Grade (F)
Math Quiz	0.12***		0.11***	0.08***	0.13***
	(0.02)		(0.02)	(0.02)	(0.03)
Perceived Math Aptitude		1.81***	1.23***	1.06***	1.09***
		(0.36)	(0.37)	(0.48)	(0.57)
Male	1.93***	2.70***	1.55***		
	(0.68)	(0.67)	(0.68)		
Age	-0.04	-0.11	0.01	0.01	0.05
	(0.18)	(0.19)	(0.18)	(0.26)	(0.27)
White (Hispanic)	-0.94	-0.86	-0.92	-1.54	-0.97
	(1.16)	(1.16)	(1.15)	(1.35)	(1.97)
Black	-2.92***	-4.28***	-3.16***	-5.36***	-2.00
	(1.26)	(1.25)	(1.25)	(1.81)	(1.79)
Asian	-2.28	-1.73	-1.95	-4.32***	-0.67
	(1.73)	(1.75)	(1.72)	(2.58)	(2.38)
Other Race	1.09	0.30	1.23	1.80	-0.08
	(2.31)	(2.28)	(2.29)	(2.48)	(4.52)
Sophomore	0.32	0.51	0.66	1.52	-0.40
	(0.88)	(0.89)	(0.85)	(1.13)	(1.33)
Junior	1.34	2.14***	1.71	4.80***	-1.19
	(1.13)	(1.11)	(1.11)	(1.44)	(1.78)
Senior	1.71	1.34	1.50	1.95	1.80
	(1.92)	(1.91)	(1.90)	(2.22)	(3.60)
Private School	-0.49	-0.09	-0.28	-0.94	0.00
	(0.80)	(0.81)	(0.80)	(0.98)	(1.28)
Job	-1.09	-1.12	-0.87	-1.10	-0.92
	(0.69)	(0.69)	(0.68)	(0.91)	(1.04)
First Time Econ	-2.16***	-2.75***	-2.08***	-2.41***	-2.58***
	(0.78)	(0.78)	(0.77)	(1.01)	(1.19)
Number of Courses	0.47	0.67	0.37	0.61	0.15
	(0.45)	(0.45)	(0.44)	(0.60)	(0.68)
International Student	-0.22	0.05	-0.13	-0.62	0.52
	(1.55)	(1.51)	(1.53)	(2.11)	(2.32)
Transfer	-1.54	-2.04***	-1.36	-2.68***	-0.53
	(1.03)	(1.03)	(1.02)	(1.21)	(1.83)
# of Math Classes	-2.02	-3.18	-3.18	-3.10	-4.94
	(3.17)	(3.25)	(3.16)	(4.37)	(4.70)
Cumulative GPA	3.87***	3.97***	3.71***	3.25***	4.09***
	(0.33)	(0.33)	(0.33)	(0.39)	(0.55)
NKU	-0.09	-0.43	-0.35	-3.11***	1.62
	(0.97)	(0.98)	(0.97)	(1.25)	(1.50)
YHC	-3.37***	-4.72***	-3.25***	-4.39***	-3.07***
	(1.08)	(1.06)	(1.06)	(1.39)	(1.67)
Spring	-1.00	-1.02		-1.14	-0.75
	(0.71)	(0.71)		(0.91)	(1.08)
Constant	55.67***	55.75***	52.87***	58.14***	52.48***
	(5.71)	(5.79)	(5.68)	(8.22)	(8.34)
Observations	616	643	616	301	314
R-squared	0.37	0.34	0.38	0.45	0.36

Standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Appendix Table 4: Results Using Categorical Groups: High to Low Actual and Perceived Ability

	(1)	(2)	(3)
VARIABLES	Final Grade	Final Grade (M)	Final Grade (F)
High Ability and Perceived High Aptitude	4.99***	3.77***	5.50***
	(0.78)	(1.02)	(1.24)
High Ability and Perceived Low Aptitude	2.91***	3.03***	2.37
	(1.33)	(1.77)	(2.03)
Low Ability and Perceived High Aptitude	1.41	-0.08	2.40***
	(0.93)	(1.25)	(1.39)
Male	2.48***		
	(0.67)		
Age	-0.11	-0.06	-0.05
	(0.18)	(0.26)	(0.27)
White (Hispanic)	-0.93	-1.73	-0.56
	(1.15)	(1.34)	(1.98)
Black	-3.83***	-5.50***	-3.14***
	(1.24)	(1.86)	(1.75)
Asian	-1.90	-4.89***	-0.15
	(1.74)	(2.60)	(2.42)
Other Race	0.13	1.20	-0.35
	(2.26)	(2.55)	(4.17)
Sophomore	0.37	1.53	-0.52
	(0.88)	(1.14)	(1.34)
Junior	1.62	4.72***	-0.94
	(1.11)	(1.42)	(1.76)
Senior	1.54	2.18	1.83
	(1.89)	(2.21)	(3.49)
Private School	-0.16	-0.97	0.15
	(0.80)	(0.99)	(1.29)
Job	-1.00	-0.99	-1.27
	(0.69)	(0.91)	(1.05)
First Time Econ	-2.53***	-2.66***	-2.81***
	(0.78)	(1.02)	(1.20)
Number of Courses	0.73***	0.93	0.45
	(0.44)	(0.60)	(0.68)
International Student	0.22	0.20	0.46
	(1.50)	(2.01)	(2.31)
Transfer	-1.63	-2.71***	-1.16
	(1.02)	(1.23)	(1.81)
# of Math Classes	-2.27	-1.79	-5.11
	(3.24)	(4.43)	(4.90)
Cumulative GPA	3.95***	3.47***	4.41***
	(0.33)	(0.39)	(0.55)
NKU	0.06	-2.48***	1.90
	(0.98)	(1.27)	(1.51)
YHC	-3.82***	-4.29***	-3.81***
	(1.07)	(1.41)	(1.65)
Spring	-0.16	-0.25	-0.08
	(0.78)	(1.05)	(1.17)
Constant	58.41***	61.35***	58.83***
	(5.78)	(8.15)	(8.51)
Observations	643	315	327
R-squared	0.35	0.43	0.33

Standard errors in parentheses, * p<0.10, ** p<0.05, *** p<0.01