ASSESSING FINANCIAL EDUCATION METHODS: PRINCIPLES VS. RULES-OF-THUMB APPROACHES[1]

October 26, 2015

Abstract: Despite thousands of financial education programs and tremendous public and private desire to improve individuals' financial decision-making, little is known about how best to teach financial education. Using an experimental approach, we estimate the effects of two different education methodologies (principles based and rules of thumb) on the knowledge, self-assessed knowledge, financial self-efficacy, motivation to learn, willingness to seek advice, risk preferences, and time preferences of high-performing undergraduate students. We find that both methods increased cognitive measures of knowledge and non-cognitive measures of self-efficacy, motivation to learn, and willingness to take financial risks. We find few differences in the relative effectiveness of each method, though the principles methodology appears to generate larger gains in self-efficacy while the rules of thumb method appears to reduce individuals' willingness to seek advice. This evidence should inform economists on the role of heuristics in economic learning and the policy community on optimal program design for different student populations.

JEL Codes: A22, D14, G11

Corresponding Author:
William L. Skimmyhorn
Office of Economic and Manpower Analysis,
Department of Social Sciences
United States Military Academy
West Point, NY 10996
william.skimmyhorn@usma.edu
(845) 938-4285

Brian Mitchell
Department of Social Sciences
United States Military Academy
West Point, NY 10996
evan.davies@usma.edu

Evan R. Davies
Department of Social Sciences
United States Military Academy
West Point, NY 10996
evan.davies@usma.edu

David Mun
Department of Social Sciences
United States Military Academy
West Point, NY 10996
david.mun@usma.edu

[1] We thank Mike Walker, Brian Forester, Susan Carter, Luke Gallagher and participants at the Cherry Blossom Financial Education Institute and the 5th Annual FDIC Consumer Research Symposium for their assistance in completing this research. We appreciate the support of the Economics and American Politics instructors at USMA. David Walik and Kevin Friedman provided excellent research assistance. All errors are our own. The opinions expressed herein reflect the personal views of the authors and do not reflect the views of the United States Military Academy, the United States Army or the Department of Defense.

I. Introduction

The most recent and widely cited reviews of the financial education literature provide lukewarm support for its general efficacy and leave open the question of optimal teaching methods. While some studies suggest little reason for optimism (Fernandes, Netemeyer and Lynch 2014; Hastings, Madrian and Skimmyhorn 2014), others suggest that education can positively impact select behaviors (Miller et al. 2014). A more recent review (Lusardi and Mitchell 2014) suggests previous non-findings are unsurprising if financial literacy is a human capital investment with groups naturally differing in their optimal decisions. Unfortunately, none of these reviews evaluates the impact of different financial education methods. They are however, united in their calls for more experimental work and better program evaluations.

A few recent studies suggest some promise for financial education. Experimental work in the developing world suggests that financial education can improve the accounting behaviors of micro-entrepreneurs (Drexler, Fischer and Schoar 2014) and rainfall insurance decisions of farmers (Gaurav, Cole and Tobacman 2011). In the U.S., recent research by Lusardi et al. (2014) suggests several delivery mechanisms (e.g., brochures, narratives and videos) that may be effective in improving confidence, self-efficacy and financial literacy. Larger quasi-experimental studies also suggest that education may improve financial decision making for high school students (Brown et al. 2014) and young military recruits (Skimmyhorn 2014). Importantly, while the Lusardi et al. (2014) paper evaluates different delivery mechanisms, the Drexler, Fischer and Schoar (2014) paper demonstrates differential effects by course methodology (principles or rules-of-thumb based training). We build on their methodological evaluation using an experimental approach in a domestic setting of wide interest: undergraduate education.

As an additional motivating factor for our work, economists appear divided on the utility of using heuristics (which we will equate with rules-of-thumb) for individual decisions. To date, there is substantial evidence that individuals use heuristics generally (Tversky and Kahneman 1974) and in financial decisions specifically (see Winter, Schlafmann and Rodepeter 2012 for a review), and we are interested in whether encouraging their use via education might improve financial decision-making. While some suggest that simplifying financial topics may improve behavior (Drexler, Fischer and Schoar 2014), others argue that employing rules-of-thumb in financial decisions might be costly since decisions are increasingly complex and highly individualized (Willis 2011). There may exist a tradeoff between simplifying information in order to increase learning and providing sufficiently complex information to prepare individuals for the variety of financial decisions they make. Since the utility of rules-of-thumb education may vary according to individuals' abilities (Binswanger and Carman 2011), assessing differential treatment effects is an especially important task. And while some (Love 2013 and Cocco, Gomes and Maenhout 2005) have developed improved rules-of-thumb, no research evaluates the effectiveness of teaching their heuristics. Our data afford a more detailed look at the heterogeneous effects of two leading methods.

We estimate the effects of these two methodologies (principles-based and rules-of-thumb) using an experimental approach. While we are interested in actual behavioral outcomes, here we provide initial evidence on the programs using tests of knowledge and self-reports for several behaviors. We find that both methods increased cognitive measures of knowledge and non-cognitive measures of self-efficacy, motivation to learn and willingness to take financial risks relative to a control group. Relative to one another, we find that the principles

methodology generates larger gains in self-efficacy and the rules-of-thumb method reduces individuals' likeliness to seek advice.

Our research makes several unique contributions because it provides evidence on the general effectiveness of financial education and the relative effectiveness of different education methods. We study the effects of financial education in a mandatory course, thereby eliminating pervasive concerns over selection into financial education programs or methodologies. The random assignment of students into courses, of instructors to teaching methods and our ability to randomly assign teaching methods to instructors creates an appealing experimental context with a control and different treatment groups. We also collect data on several important outcomes (i.e., objective knowledge, self-reported knowledge, self-efficacy, motivation to learn, likeliness to seek advice, risk preferences and time preferences) using a pre/post assessment. While ideally we would observe actual financial decisions, we identify and observe a number of outcomes plausibly linked to financial decision-making. This approach enables us to provide insight into the mechanisms under which financial education methods might affect decision-making. These outcomes also provide a more comprehensive look at the potential benefits of financial education than previously considered. Finally, our access to detailed administrative data enables more precise estimation of the causal effects of education and the evaluation of potential heterogeneous treatment effects. Our paper proceeds as follows: In Section II we describe the institutional setting and the data. We describe our identification strategy in Section III and present our results in Section IV. In Section V we provide robustness checks. We discuss our results and conclude in Section VI.

II. Institutional Setting and Data

Background

The United States Military Academy at West Point is a highly selective (mean SAT in this sample is 1317) undergraduate liberal arts institution of moderate size (student population is approximately 4500) with a robust core and Science, Engineering, Technology and Mathematics focused curriculum that results in a Bachelor's of Science degree for all graduates. The core curriculum includes a one semester Principles of Economics course, typically taken during the sophomore year. As part of this semester-long course (40 x 55 minute lessons), students complete lab periods (4 x 2.0 hour) devoted to personal finance.[2] West Point requires completion of these labs to better prepare students for the financial challenges of military service (all graduates become Lieutenants in the Army) and to enable them to advise and assist the soldiers they will lead. During the same year, students must also complete American Politics, a one semester course with similar time requirements. Since the institution randomly assigned students to Principles of Economics or American Politics during the fall semester, we can evaluate the effects of financial education relative to a control group. We discuss the generalizability of our findings given this sample further in Section VI.

Our treatment consists of completion of the Principles of Economics course coupled with four two-hour personal finance lab periods conducted using a principles-based (PB) or rules-of-thumb (ROT) methodology. The design enables us to evaluate the general effectiveness of financial education and the relative effectiveness of the PB and ROT methodologies.

We exploit administrative data from the Academy that captures individual demographics (age, gender, race and prior military service), their SAT score,[3] and individuals' first year

[2] Among the sophomores, some students are enrolled in both courses (N=60). We assign them to the treatment group since they received financial education and we control for their dual attendance with an indicator variable. In addition, a small number of freshman (N=47), juniors (N=25), seniors (N=1) and foreign exchange cadets (N=6) are enrolled in one course and we omit them from the analysis.

[3] We use SAT scores imputed by the school. The scores reflect the maximum of the individuals' actual SAT score or their estimated SAT score using an ACT/SAT concordance table.

academic GPA for members of the Class of 2017 during the fall 2014 semester (N=986).[4] In Table 2 we present summary statistics for these characteristics by control and treatment groups and observe similar characteristics across all groups. We discuss our experimental validity further in Section III.

Financial Education Methodologies

Students in both treatment methods covered the same topics and had the same course attendance requirements. Students completed the labs in a seminar format, with instructors presenting material via lecture, slides, videos and handouts. Instructors assigned students practice exercises that required calculations by-hand, using computers (e.g., Microsoft Excel), or using online calculators and these exercises required students to implement and use unique concepts from each method. Class sizes were small (12-17 students, mean=14.94) and included substantial interaction and active Q&A between the instructors and students. See Table 1 for details on the topics covered in each lab period and the key differences by treatment method. They had common reading assignments prior to each session from the *Guide to Personal Financial Planning for the Armed Forces* (Gayton and Handler 2012) and they completed the same capstone exercise, a personal financial plan consisting of goals and a detailed budget.[5]

We designed the PB method based on traditional personal finance instruction that teaches students general skills such as a financial planning process, evaluating the tradeoffs associated with different types of consumption, budgeting and the time value of money. This method also covers specific topics such as emergency savings, investing, purchasing insurance and the decision to buy/lease a car and buy/rent a home. Similar programs are used in a variety of

[4] N=994 students took the first assessment and N=986 (99.2%) completed the second assessment. We restrict our analysis to individuals with data for both assessments.

[5] The common economics text for both groups was R. Glenn Hubbard and Anthony P. O'Brien, "Economics," 5th Edition.

settings including programs developed by national governments (e.g., the Federal Deposit Insurance Corporations' 'Money Smart' program and the Australian Securities and Investments Commission's program of the same name), international organizations (e.g., UNICEF's Child Social and Financial Education Program) and non-profit organizations (e.g, the National Endowment for Financial Education's 'Smart About Money' and the Financial Industry Regulatory Authority Investor Education Foundation's 'Love Your Money' program).

We designed the ROT method using the latest financial education curriculum of the non-profit organization 'Moneythink.'[6] We adapted their rules-of-thumb and some of the specific examples to the topics we developed for our course. The ROT did not attempt to develop financial planning experts. Instead, this method attempted to simplify the presentation of information, avoided lengthy discussions of the theory or detailed calculations behind decisions and most importantly, provided students with heuristics (17 in total) designed to simplify complex decisions and enable students to make good decisions. See Table 1 for the personal finance session topics and the corresponding rules-of-thumb.

We provide two examples to highlight the differences in the treatment arms. First, to teach the time value of money (TVM), the PB course introduced the content, provided visual demonstrations of the effects of compounds interest (i.e., graphs) and had students complete (using by-hand calculators and excel) a variety of exercises to emphasize the effects of changes in the parameters in the basic TVM formula: $FV = PV(1 + i)^n$. In the ROT course, instructors provided similar visual demonstrations of the effects of compound interest and provided students with online tools (e.g., calculators) to determine the present and/or future values of different cash flows. The ROT method briefly showed students the TVM formula but it did not did encourage

[6] For more information on 'Moneythink', see: http://moneythink.org/

its use during practical exercises. As a result, the courses may vary slightly in their content (primarily the same topics but different examples) but especially in the problem solving methods students were required to employ.

Second, when teaching students how to calculate their after-tax income, the PB approach covered all of the intermediate calculations in determining their total tax liability which included calculating adjusted gross income, determining deductions and credits and subtracting FICA Social Security and Medicare taxes from their taxable income. The ROT course, on the other hand, simply highlighted these components and did not take time to show each individual calculation. In both courses, instructors emphasized the importance of understanding the difference between average and marginal tax rates, but in the ROT course, students were taught to determine their after-tax pay by simply referencing their military pay stub. Additionally, students enrolled in the ROT course were only required to list their total after-tax income when completing their budget exercise, while those enrolled in the PB course were required to list each individual component of their tax liability.

Despite these differences in design, there were a number of commonalities between the two methods. As previously mentioned, both shared the same textbook and assigned readings. In addition, the culminating event for the personal finance labs was a personal finance exercise (PFX) completed by both treatment groups and graded on the same scale. While most of the PFX components were common across teaching methods, the ROT method prompted students to use specific rules in their budgeting efforts (e.g., the use of SMART goals [See Table 1] and recommended budget allocations by category). For several concepts, the two methods shared the same examples or practical exercises but they required students to solve these problems using different techniques. In addition, instructor reliance on common teaching examples (e.g.,

teaching compound interest with the rule of 72) exists to some extent. Finally, while the treatment groups differed in the personal finance methods, they completed the same principles of economics lessons, some of which provided instruction on topics that could have influenced students' learning about personal finance (e.g., lessons on inflation, taxes and consumer choice). These factors serve to mitigate the actual differences in the two treatment groups but were accepted as important baseline requirements for the economics program.

Outcomes

Our assessment exploits data on nine different outcomes likely to be related to financial decision making but observable at the completion of our education. These outcomes reflect a number of channels through which financial education could affect financial behavior. See Appendix Table 1 for a complete list of our assessment items by outcome. We collected data for all outcomes from online pre and post assessments (99% completion rate) required as part of the course. Instructors initiated the assessments using common scripts, encouraged students to complete the assessments to the best of their abilities and incentivized student participation using participation grades. All of the items are identical for both assessments with the exception of a final set of items (new knowledge) discussed below.

Our first two outcomes relate to financial knowledge, measured using multiple choice questions, and our motivation is two-fold. First, knowledge development is a primary goal of financial education (Hastings, Madrian and Skimmyhorn 2013; Lusardi and Mitchell 2014) and an appropriate benchmark for financial educators. Second, financial knowledge is positively related to financial decision-making in a number of contexts including avoiding financial mistakes (Martin 2007); risk diversification (Lusardi and Mitchell 2008); hypothetical choices by college students (Chen and Volpe 1998), credit management, saving and investment (Hilgert,

Hogarth and Beverly 2003); avoiding high-cost borrowing (de Bassa Scheresberg 2013); and planning for retirement (de Bassa Scheresberg 2013, Van Rooij, Lusardi and Alessi 2012, Lusardi and Mitchell 2006). We measure topical financial knowledge using custom designed questions that reflect the specific course topics and we report the percentage of 20 items answered correctly. We also measure financial knowledge using the percentage of the standard five items developed by Lusardi and Mitchell (2009) answered correctly.

We use self-assessed knowledge as our third outcome since perceptions of financial knowledge and capabilities may also affect financial decision-making (Algood and Walstad 2013). Our fourth outcome measures financial self-efficacy for handling day to day financial measures since prior research (Robb and Woodyard 2011, FDIC 2007) suggests that financial education might increase individuals' financial confidence.

In addition, since low financial literacy has been linked to a lack of motivation (Mandell and Klein 2007) we measure the effects of the courses on individuals' motivation to learn about financial topics with our fifth outcome. Sixth, since financial advice may serve as a substitute for financial knowledge (Collins 2012) we measure individuals' self-assessed likeliness to seek advice. To simplify our empirical framework and support the use of ordinary least squares (OLS) regressions, we convert the Likert scale answers for outcomes three through six above into indicator variables that reflect "High" levels of the attribute (i.e., =1 for scale answers $\geq$7, =0 otherwise).[7]

Financial education might also affect decision-making by altering individuals' risk preferences. Grable (2000) suggests that individuals with more education and those with higher levels of financial knowledge demonstrate higher levels of risk tolerance. A likely explanation is that financial education alleviates individual concerns by demonstrating risk is common in

[7] In Section V we complete robustness checks using alternate thresholds and find similar results.

financial markets and assumed by many participants, but these effects might also move in the opposite direction increasing individual risk aversion by highlighting the potential for losses. Our seventh outcome measures individuals' self-assessed willingness to take risk and we convert the Likert scale response to an indicator for high willingness to assume risk as above. Eighth, since financial education might affect time preferences (Hastings and Mitchell 2011; Meier and Sprenger 2007), we evaluate a measure of patience using the share of a hypothetical loan allocated to long term savings goals (>10 years).[8]

The final outcome reflects objective knowledge measures of students' financial choices for new financial decisions that were neither covered in the course nor presented in the initial assessment. One potential advantage of a principles-based financial education is its ability to teach students skills that can be used in new contexts. Said differently, a rules-of-thumb approach might make it more difficult for individuals to solve problems when faced with new circumstances (Drexler, Fischer and Schoar 2014). We asked students to make decisions related to paying down debts with different interest rates using a windfall payment, select among tax advantaged savings strategies for children's education and compute the time required to double a planned down payment for a home. We combined these three items into a new problem solving outcome and report the percentage answered correctly. In Table 3 we present summary statistics for all nine outcomes by assessment (pre vs. post) and group (control vs. PB vs. ROT).

III. Empirical Strategy

To identify the causal effects of financial education we exploit the ability to randomly assign students and the ability to assess student outcomes at the beginning and end of the course. The institution randomly assigned students to the American Politics or Economics course prior to

[8] The majority of students at USMA accept a low interest "pre-commissioning loan" midway through their junior year. The loan for the class of 2016 was a 5-Year, $36,000 loan at 0.75%.

the semester. Then the Economics course director, working with the research team, completed a balanced random assignment process prior to the semester wherein each instructor was assigned an equal number of principles and rules-of-thumb sections (See Appendix Table II for the detailed assignments). This effectively randomly assigned students to a treatment method. As a result of these assignments, we can compare the changes in individual outcomes by group assignments. While our assignment processes are plausibly random we utilize the pre/post nature of our assessments in a differences in differences (DD) framework that only requires an assumption of parallel trends across the groups.

As a result, we compare the changes in outcomes for students enrolled in American Politics (control group, N=422) with the changes in outcomes for students enrolled in Economics overall (treatment group, N=574), those Economics students assigned to the PB group (N=291) or those assigned to the ROT group (N=283) using the DD model specified in Equation 1[9]:

$$Y_{ist} = \beta_0 + \beta_1 Post_t + \beta_2 T_i + \beta_3 Post_t \times T_i + X_i^{'}\gamma + \delta_s + \varepsilon_i \quad (1)$$

In this model, Y_{ist} is an outcome of interest (i.e., objective knowledge, self-reported knowledge, financial self-efficacy, motivation to learn, likeliness to seek advice, risk preference, time preference and new knowledge) for individual i in period t in section s. $Post_t$ is an indicator that equals 1 for a student's end of course (post) assessment and equals 0 for the initial (pre) assessment. T_i is an indicator that equals 1 for students assigned to the treatment group (PB or ROT) and equals 0 otherwise. $X_i^{'}$ is a vector of individual characteristics potentially related to student outcomes that includes: age, gender, race, SAT scores, first year GPA and an indicator for prior military service. δ_s is a vector of section fixed effects (the course is offered during one

[9] In Section V we complete robustness checks and find that our estimates are not sensitive to our functional form.

of four standard times). We cluster our standard errors at the instructor level to capture unobserved correlations in the error terms.

We complete four different comparisons. To evaluate the overall efficacy of financial education we compare the combined treatment group ($T_i = 1$) versus the control group ($T_i = 0$). Next, to evaluate the effectiveness of each method independently, we compare the PB treatment group ($T_i = 1$) to the control group and the ROT treatment group ($T_i = 1$) to the control group separately. Finally, to compare the two education methods, we compare the ROT treatment group to the PB treatment group. In each case, β_3 is the DD coefficient of interest and reflects the causal effects of an economics course and 6-8 hour financial education course relative to the control group (in comparisons 1-3) and the causal effects of a ROT course relative to the PB course (in comparison 4). The identifying assumption in this DD framework is parallel trends: we assume that the outcomes for the control group members would trend in the same manner as the outcomes for the members a treatment group. The assumption seems plausible given the random assignment of students to the different courses (American Politics or Economics), the balanced random assignment (within instructor sections) of economics students to the PB and ROT methods and the common sophomore student experiences at West Point. We provide additional evidence below.

For our final outcome (new financial problem solving), we do not have a pre-treatment measure of individual decisions. As a result, we estimate the OLS regression in equation 2:

$$Y_i = \beta_0 + \beta_1 T_i + X_i'\gamma + \delta_s + \varepsilon_i \quad (2)$$

In this model T_i is a binary treatment indicator and the individual characteristics (X_i) and section fixed effects (δ_s) are the same as in equation 1. We cluster the standard errors at the instructor level and complete the same four comparisons described above.

In support of our identification assumptions, we complete two analyses. First we compare individuals assigned to each group by all of our observable characteristics. In Table 2 we provide summary statistics for each our experimental groups and we compare them to the control group and to one another. The results reveal very few differences in the groups across a number of individual characteristics. The ROT group has a higher average SAT score and first year GPA than the Control group and a higher first year GPA than the PB group. We control for these characteristics in our regressions and we also complete a DD specification wherein the baseline differences do not undermine our identification.

Second, we test whether the individuals' observable characteristics are related to their assigned treatment condition using the covariate regressions specified in equation 3:

$$T_i = \alpha_0 + X_i'\sigma + \delta_s + \varepsilon_i \qquad (3)$$

For each of our four cases, we regress an indicator for the assigned treatment condition (Combined, PB_i or ROT_i) on our observable characteristics (X_i) and the section fixed effects δ_s. We then evaluate the partial R-squared for the individual characteristics as suggested by Altonji, Elder & Taber (2005) and test the joint significance of the individual characteristics in predicting treatment assignment. We present the results of these tests in Panel B of Table 2. In all four cases, the individual characteristics only explain a tiny portion of the variation in treatment (0.017, 0.016, 0.024 and 0.020 respectively) and the observable characteristics are jointly unrelated to treatment (p-values equal 0.512, 0.861, 0.254 and 0.255 respectively). Since there appears to be substantial covariate balance by treatment condition (Panel A) and the observable characteristics are jointly unrelated to treatment group assignments (Panel B), we proceed as if the unobservable characteristics are unrelated to treatment condition. This is even stronger evidence than is required for our DD model, which only requires that the groups would trend

similarly between the pre and post assessment. It strongly suggests a valid experimental design. This evidence also supports our identification assumption for our ninth outcome (new problem solving), which we only collect during the post assessment.

One additional concern for the DD estimation between the two treatment groups (PB vs. ROT) might be that there exist common shocks or experiences between members of one group that could drive results (e.g., all of the good instructors use the same personal finance education method). To address this concern we balanced instructor assignments across treatment groups to ensure that all economics instructors taught both methods (e.g., instructors with four sections taught two PB sections and two ROT sections). We also cluster our standard errors at the instructor level. We summarize the instructor assignments in Appendix Table 2.

IV. Results

Summary Statistics

In Table 3 we present summary statistics for our outcomes by treatment assignment (control vs. PB vs. ROT) and assessment (pre vs. post). The control group statistics (Panel A) reveal primarily stable mean outcome levels between the pre and post assessments, with small declines among a few variables. The Panel B results suggest large increases in mean outcomes after the course within the PB group for several outcomes (topical knowledge, self-assessed-knowledge, self-efficacy and willingness to take risk), small changes in a few (Big 5 knowledge, motivation to learn, patience) and a small decline in likeliness to seek advice. Similarly, the Panel C results suggest large increases in mean outcomes after the course within the ROT group for several outcomes (topical knowledge, self-assessed knowledge, self-efficacy and willingness to take risk), small changes in a few (Big 5 knowledge, motivation to learn) and a moderate

decline in likeliness to seek advice. For the final outcome, both treatment groups appear to have higher mean levels of performance in solving new problems relative to the control group.

Main effects

In Table 4 we report the OLS estimates for equations 1 and 2 restrict our attention to the main coefficients of interest (i.e., the DD coefficients for Cols. 1-8 and the OLS coefficient for Col. 9). In Panel A we estimate the combined treatment effects (PB and ROT methods) and find large and statistically significant effects for seven of nine outcomes. In reporting our results we refer to the main effects using the regression coefficients using percentage points (pp) and the effects sizes (point estimate divided by control mean) using percentages. For example, an average, the two methods increase topical knowledge (Col. 1) by 9pp, a 15% increase on a control mean of 58.94%, and the effect is statistically significant ($p<0.01$). The effects on Big 5 knowledge are smaller (8%) but significant ($p<0.01$). Using a slightly different measure, individual's self-reports of a high level of financial knowledge (Col. 3), the results reveal even larger increases. The education increases the probability of "high" (i.e., ≥ 7) self-assessments of knowledge by 25pp (114%, $p<0.01$), a very large effect. Taken together these results strongly suggest that together, the methods are effective in increasing individuals' financial knowledge. We cannot determine if the large increases in self-assessed knowledge reflect overconfidence or if students are reporting knowledge on items not included in the other measures.

Turning to our measures or anticipated behavior and preferences, the education increased the probability of "high" (i.e., ≥ 7) assessments of financial self-efficacy (Col. 4) by 15.43pp (29%, $p<0.001$) suggesting that that both methods may impart confidence and enable students to complete more routine financial tasks that leads individuals to rate themselves higher. Encouragingly, the culminating personal financial exercise in the course aimed to do just that.

However, given that the self-efficacy gains (29%) are somewhat larger than the objective knowledge gains (8-15%), these reports could reflect overconfidence or actual increases in basic financial skills. The relationship between financial knowledge and self-efficacy warrants further attention. The course increased the probability of having a "high" level of motivation to learn about personal finance topics on their own (Col. 5, 8.40pp), a moderately sized effect (11%) that is statistically significant ($p<0.01$). The course did not have an economic or statistically significant effect on the probability of being "highly" likely to seek advice on average (Col. 6, $p=0.391$). Here the point estimate suggests that the course could have lowered individuals' likeliness to seek advice. The course increased the probability of having a "high" self-assessed willingness to take financial risk (Col. 7) by 14.54pp, large effects (32%) that are statistically significant ($p<0.01$) and consistent with previous findings that risk tolerance correlates with financial literacy (Cavezzali, Gardenal, and Rigoni 2012) and education more generally (Hallahan, Faff and McKenzie 2004 and Grable and Lytton 1998). On average, the course did not increase individuals' patience (Col. 8, 1.53pp, $p=0.302$). Finally, for new problem solving (Col. 9), the OLS estimates reveal that the financial education increased performance by 15.68pp (37%, $p<0.01$).

In Panels B and C we disaggregate the combined effects above and evaluate the effectiveness of each method compared to the control group. Many of the Panel B estimates suggest very similar results as those above: the PB method increased knowledge (Cols. 1-3), self-efficacy (Col. 4), motivation to learn (Col. 5), willingness to take risk (Col. 7), and the ability to solve new problems (Col. 9). In addition, the PB method increased individuals' patience (Col. 8) by 10% ($p=0.022$). The PB method does not appear to have a statistically or economically significant on individuals' likeliness of seeking advice (Col. 6).

The Panel C estimates suggest that the ROT method is also effective in increasing knowledge, self-efficacy, motivation to learn, willingness to take risk, and the ability to solve new problems. In most cases the point estimates are very similar to the PB estimates. There are two exceptions: the ROT method appears to reduce an individuals' likeliness to seek advice (Col. 6) by about 8%, but the statistical significance is marginal (p=0.102). The ROT method has no effect on our measure of patience (Col. 8, p=0.555).

Finally, in Panel D we provide estimates for the relative effectiveness of each financial education methodology. With three exceptions, the ROT method estimates do not differ significantly from the PB method (omitted group). In terms of differences, the ROT method reduces self-efficacy (Col. 4) by 7.4pp (13%) relative to the PB group (p=0.063). This might suggest that the PB method imparts more technical skills and deeper understanding of personal finance topics that leads individuals to rate themselves higher. However, given that the methods do not differ significantly in their actual knowledge (Cols. 1-2), these reports might reflect overconfidence. The relationship between self-efficacy and knowledge merits further study.

Second, the ROT method reduces the likelihood that an individual will seek financial advice (Col. 6) by 6.8pp (7%, p<0.01). While the mechanism is unknown, one possibility is that the PB's complexity (more equations, use of MS Excel, etc…) leads individuals to conclude that they need assistance in making financial decisions. Another possibility is that the ROT method, by encouraging the use of online resources and providing specific websites, might reduce individual's likeliness of seeking advice inadvertently.

Third, the ROT method reduces an individuals' measure of patience (Col. 8) by 5.1pp (14%, p=0.020). In this case, the PB method's more intensive use of NPV analysis and students' own calculations of the time value of money may have increased their willingness to allocate

money from an anticipated loan towards long-term goals. In this case the PB method may have better demonstrated the time value of money and returns associated with saving early in life, or it might have improved students' understanding of the real income requirements of later life consumption through the use of detailed calculations and MS Excel.

While the majority of our results suggest that both methods are equally effective, there are no outcomes in which the ROT method improves student outcomes relative to the PB method. Conversely, the PB method appears more effective in improving financial self-efficacy, increasing the likelihood of seeking advice, and increasing savings towards long-term goals.

Heterogeneous Treatment

Next we examine whether the methods had differential effects for our nine outcomes within four student groups (females, those with low quantitative ability, low financial knowledge and low initial motivation to learn) and we present the results in Table 5. Given that the focus of this paper is the comparison in teaching methods, we focus our heterogeneous treatment analysis on the comparison between the ROT and PB groups (comparable to Panel D in Table 4).

In Table 5 Panel A we estimate the treatment effects among female students. Persistent interest in the underrepresentation of women in undergraduate economics majors (Goldin 2013) and in the field of finance (GAO 2013) motivates this analysis. In short, our estimates suggest no meaningful differences in the effectiveness of the PB and ROT methods for female students. While most of the point estimates are negative, only a few approach statistical significance (i.e., motivation to learn and likeliness to seek advice). This finding is most likely explained by the self-selection of female students to a school with a robust mandatory STEM curriculum.

The findings and discussion in Drexler, Fischer and Schoar (2014) suggest that ROT methods might be more effective among individuals with low ability, knowledge or motivation.

This could be the case if these methods ease the learning of difficult concepts. In Table 5 Panel B we estimate the treatment effects among relatively lower quantitative ability students (i.e., the lowest quartile of SAT Math scores, score≤ 600). The PB curriculum might be less effective for students with lower quantitative abilities given its emphasis on calculations and analytic approaches. Our results suggest that both methods are similarly effective among this group with one notable exception: the ROT course again reduces an individual's likelihood of seeking advice (Col. 6) by 16.5pp (17%, p<0.01). Given that the mean SAT Math score for this 'low' ability group is still 575 and that they will have completed calculus prior to the economics course, these results are unsurprising.[10] They suggest that even among individuals with moderate quantitative abilities, the PB method appears slightly better than the ROT method.

In Table 5 Panel C we estimate the effects of the methods among individuals with low initial levels of financial knowledge (roughly the lowest quartile, with initial topical knowledge assessment scores ≤ 55%). For these individuals a ROT method might be more accessible and hence more effective. Our results suggest the opposite. All but one of the point estimates are negative, though only one is statistically significant. The results suggest that the ROT method reduces self-assessed knowledge (Col. 3) by 16.8pp (80%), though the result is only marginally statistically significant (p=0.0746). Our explanation is similar to that above; even the 'low' scoring individuals in this sample appear to have substantial financial knowledge.[11] Even so, the PB method again appears to be at least as effective as the ROT method and more so in one area.

Finally, in Panel D we estimate heterogeneous treatment effects for individuals with low initial scores for their motivation to learn (roughly the lowest quintile, with Likert scale scores ≤

[10] The 2012 U.S. SAT Math mean was 514, which is between the 1st and 5th percentiles of our student distribution. http://testprep.about.com/od/SAT_Scores/a/2012_Average_SAT.htm

[11] Using a related outcome (Big 5 knowledge) and the National Financial Capability Study, Hastings, Madrian and Skimmyhorn (2013) report that 14% of respondents with "some college" answer all of the Big 5 questions correctly compared to 17% in our sample. This suggests that our student population is relatively financially knowledgeable.

6 out of 10). One concern in selecting an appropriate teaching methodology is identifying an approach that will prove effective even for those students least interested in the material. Here the two methods appear to be roughly equally effective. None of the point estimates are statistically significant and they appear equally divided between positive and negative. While the zeros are not especially precise, the results do not suggest a penalty for the PB method in this subgroup. We conclude this section by noting that while both teaching methods appear effective for a variety of outcomes, overall the PB method appears to generate more beneficial effects than the ROT method among our subgroups.

V. Robustness Checks

We complete a series of robustness checks to further support our findings and we present the results in the Appendix tables. First, given our DD strategy, positive treatment effects might arise from gains among the treatment groups, declines among the control group, or both. Since we observe declines in average outcome levels in the control group for some outcomes (see Table 3), we evaluate whether a lack of attention/motivation among the control group participants in their second evaluation might drive our estimates. While the evaluation contained items related to both Principles of Economics and American Politics, the majority of the items related to economics and personal finance. So control group members may have had less motivated to perform well. We first note that such differential motivation seems unlikely since the evaluation was a mandatory part of each course, though students were only scored on their participation. Second, these concerns do not affect the estimates comparing the PB and ROT methods to one another. Finally, to more formally evaluate this possibility, we included 4 items in the assessment that evaluate each student's effort (e.g., a simple algebra problem, West Point trivia). We created an outcome reflecting the percentage of the 4 items correctly answered and

estimated Equation 1 to determine if there were differential levels of effort/attention between the control and treatment groups from the first to the second assessment. The results in Appendix Table 3 suggest no differential effort levels in any of our comparisons. We conclude that the declines in some outcomes among the control group reflect the best counterfactual for the treatment groups. The simplest explanation may be that economic knowledge may be crowded out among undergraduates, especially in settings with high course loads as here.

Second, given our small number of clusters (N=24 for comparison 1, N=23 for comparisons 2 and 3, and N=13 for comparison 4) we complete the Wild Bootstrap procedure suggested by Cameron, Gelbach and Miller (2008) to ensure that our results are robust to alternate standard error computations. We present the p-values for both standard error methods (and for another clustering choice, instructor by hour) in Appendix Table 4A and note that the results are nearly identical for all of our comparisons.

Third, we estimate our treatment effects using different functional forms. In Appendix Table 4B we present OLS binary treatment estimates (instead of a DD specification) using only the post-assessment outcomes while controlling for initial assessment outcome scores. The results are qualitatively similar, though the PB vs. Control estimates (Panel B) are larger for likeliness to seek advice (Col. 6) and smaller for patience (Col. 8). The PB vs. ROT estimate for patience is smaller and statistically insignificant. In Appendix Table 4C we provide Logit marginal effect estimates for the DD specification for our binary outcomes. These results are also similar to our main DD estimates, though the self-assessed knowledge (Col. 3), self-efficacy (Col. 4), and risk-preference (Col. 7) outcomes are typically smaller and the motivation to learn (Col. 5) and likeliness to seek advice (Col. 6) are slightly larger. Overall, nearly all of our findings are stable to different regression functional forms.

Finally, we complete the main DD specifications (Equation 1) for outcomes three through seven (self-assessed knowledge, self-efficacy, motivation to learn and likeliness to seek advice) using alternate thresholds for "high" (i.e., =6, =7 and =8) outcome levels. We complete this analysis for the combined treatment versus control comparison. The results in Appendix Table 5 demonstrate that our main results are relatively stable to different threshold levels.

VI. Discussion & Conclusion

We estimate the effects of two different financial education methodologies (principles-based and rules-of-thumb-based) on several economic outcomes using a field experiment in a sample of high-performing undergraduate students. We find that both teaching methods increased cognitive measures (i.e., actual and perceived knowledge) and non-cognitive measures (i.e., self-efficacy, motivation to learn and willingness to take risks) of financial literacy.

Interestingly, we find only a few differences in the relative effectiveness of each method. The principles methodology appears to generate larger gains in self-efficacy while the rules-of-thumb method appears to reduce individuals' willingness to seek advice. Since the principles-based method arguably provides a more general toolkit than the rules-of-thumb approach, we expected that it would better prepare students to solve new financial problems. However, both methods proved equally effective in preparing students on this dimension. The most likely explanation for the lack of differential effects for the two methodologies is their similarity in overall content. Despite attempts to make the methods distinct, the common readings, same syllabus topics, nearly identical culminating graded assignment and the likelihood that instructors may have provided comparable instruction provided students with a somewhat

combined methodological experience. Researchers in this area should be aware that their treatment intentions may be more difficult to execute than they anticipate.

We also find a few heterogeneous treatment effects. Notably, we find some evidence that financial education is slightly less effective for female students, though both methods generate important gains relative to the control group. For individuals with low quantitative abilities, initial knowledge scores, or initial motivation levels, the principles method appears to be slightly more effective overall, though the ROT method is effective in many cases.

Relative to existing estimates, ours appear reasonable. Good causal estimates of financial knowledge gains from personal finance instruction are uncommon, but Lusardi et. al. (2014) find knowledge effects from 6-20% and self-efficacy effects of 20%. Our estimates on knowledge (7-16%) are comparable, but our self-efficacy estimates (22-36%) are higher than theirs. Our larger effects seem reasonable given the duration of the course we study (one semester with 8 hours of instruction) versus theirs (about 5 minutes) and the required assignments (e.g., practice problems and a capstone personal financial planning exercise). To our knowledge, we are the first to measure the effects of different teaching methods on risk preferences, time preferences, individuals' likeliness to seek advice and their problem solving skills for new topics.

Perhaps of more direct interest, Drexler, Fischer and Schoar (2014) evaluate these same training methods among micro-entrepreneurs. They find no improvements in self-reported accounting behaviors for the principles method and 8-25% improvements for the rules-of-thumb method. While we do not measure any specific financial behaviors, our estimates for behavioral attributes such as self-efficacy (22-36%) and motivation to learn (11%) appear comparable to theirs. Importantly, relative to their work, we find beneficial effects of both teaching methods and suggestive evidence that the principles-based method is more effective overall. While the

exact reason is unknown, we suspect that while the rules-of-thumb method may be effective for audiences with lower levels of human capital, principles-based methods may be equally or more effective in higher human capital settings. The optimal choice of teaching methods for audiences with human capital levels between our sample and theirs requires additional study.

Some internal validity considerations suggest reason for even more optimism in our sample. Absences among the treatment group will reduce actual differences in education and bias downward our estimates in the first three comparisons. A "John Henry" effects among the control group would have similar effects. For all four comparisons, sharing of course materials and new knowledge between groups, which seems especially likely in this team focused environment than in other undergraduate settings, will result in contamination that also biases our estimates downward. Despite the intervention's design, the commonalities in the teaching methods (see Section II) also produce contamination. The last two concerns may explain why we do not find even more differences between the PB and ROT methods.

However, West Point is a unique institution and so our results should be interpreted carefully. On the one hand West Point is like many other competitive and elite undergraduate settings. For example, using estimated median SAT scores West Point's nearest peer institutions are the University of Wisconsin-Madison and Boston University.[12] National ranking systems tell a similar story with West Point placing between Massachusetts Institute of Technology and the University of Pennsylvania in the Forbes 2015 rankings and between Colgate University and Macalester College in the U.S. News and World Report 2015 Liberal Arts College rankings.[13] Finally, previous research on peer effects in the student body at West Point (Lyle 2009) identifies

[12] We estimate the median SAT score using Integrated Postsecondary Education Data System statistics for the Class of 2017. We estimate the median using the average of the 25th and 75th percentile scores for each institution.
[13] For the Forbes rankings see: http://www.forbes.com/top-colleges/list/. For the U.S. News Rankings see: http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/national-liberal-arts-colleges/data.

Dartmouth and Williams as peer institutions while research on classroom mentor effects at the Air Force Academy (Carrell, Page and West 2010), identifies institutions including Georgia Tech, California Institute of Technology and Virginia Tech as peer institutions. While none of these institutions are the same as West Point, they provide some insight into the educational settings and student populations for which these financial education methodologies might have similar results.

On the other hand, self-selection into West Point (a military institution where students may take instructors' advice especially seriously), the relatively high human capital of the students, and students' professional motivations to learn the material in preparation for their career as a leader in the Army, and their near certain approval of a low-interest $40K loan during their junior year all suggest that our estimates could be upper bounds relative to typical undergraduates.

We omit a detailed discussion of the costs associated with each method but highlight a few important considerations. In our study, cost differences were negligible since we provided all lesson plans and materials to the instructors. In general, the principles method requires the educator to have a more complete understanding of the material, including the mathematical concepts inherent in financial problems (e.g., how to compute a loan payment) and the relevant policy rules for the audience (e.g., IRA eligibility). While these requirements may not exceed those of for normal economics or finance classes, they may not be assured in all educational settings (e.g., Way and Holden 2009 for U.S. high schools). The rules-of-thumb method may be easier to execute once lessons are prepared, but the preparation itself may be more challenging as it requires identification and articulation of an appropriate rule of thumb as well as validation and integration of useful online resources.

In summary, this research provides the first experimental evidence on the overall and relative efficacies of teaching personal finance using a principles-based or rules-of-thumb-based approach in an undergraduate setting. We hope that our research will improve economists' and educators' understanding of the utility of heuristics in economic learning. It should also provide more reliable evidence to policymakers and practitioners on optimal program design among student populations with relatively high levels of human capital.

References

Allgood, Sam and Walstad, William. 2013. "Financial Literacy and Credit Card Behaviors: A Cross-Sectional Analysis by Age," *Numeracy* 6(2): Article 3. doi:10.5038/1936-4660.6.2.3

Altonji, Joseph, Todd Elder and Christopher Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools" *Journal of Political Economy* 113 (1): 151-184.

Binswanger, Johannes & Carman, Katherine Grace. 2011. "The Role of Decision Making Processes in the Correlation between Wealth and Health," Discussion Paper 2011-005, Tilburg University, Center for Economic Research.

Brown, Meta, Wilbert van der Klaauw, Jaya Wen and Basit Zafar. 2014. "Financial Education and the Debt Behavior of the Young." Federal Reserve Bank of New York Staff Report No.634.

Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *The Review of Economics and Statistics* 90(3): 414-427.

Carrell, Scott E., Marianne E. Page, and James E. West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." Quarterly Journal of Economics, 125(3): 1101-1144.

Cavezzali, Elisa and Gardenal, Gloria and Rigoni, Ugo. 2012. "Risk Taking, Diversification Behavior and Financial Literacy of Individual Investors." Università Ca' Foscari Department of Management Venezia Working Paper No. 17/2012. Available at SSRN: http://ssrn.com/abstract=2162783

Chen, Haiyang, and Ronald P Volpe (1998). An analysis of personal financial literacy among college students. *Financial Services Review, 7,* 107-128.

Cocco, Joao, Francisco Gomes and Pascal Maenhout. 2005. "Portfolio Choice Over The Life Cycle." *Review of Financial Studies* 18: 491-533.

Collins, J. Michael, Financial Advice. 2012. "A Substitute for Financial Literacy?" *Financial Services Review* 21 (4): 307-322. doi: 10.2139/ssrn.2046227

de Bassa Scheresberg, Carlo. 2013. "Financial Literacy and Financial Behavior among Young Adults: Evidence and Implications," *Numeracy*: Vol. 6: Iss. 2, Article 5. URL: http://scholarcommons.usf.edu/numeracy/vol6/iss2/art5

Drexler, Alejandro, Greg Fischer and Antoinette Schoar. 2014. "Keeping it Simple: Financial Literacy and Rules-of-thumb." *American Economic Journal: Applied Economics* 6(2): 1-31.

Federal Deposit Insurance Corporation. 2007. “A Longitudinal Evaluation of the Intermediate-term Impact of the ‘Money Smart’ Financial Education Curriculum upon Consumers’ Behavior and Confidence.” April. https://www.fdic.gov/news/news/financial/2007/fil07046.html

Fernandes, Daniel, John G. Lynch, Jr. and Richard G. Netemeyer. 2014. “Financial Literacy, Financial Education and Downstream Financial Behaviors,” forthcoming in *Management Science.*

Gaurav, Sarthak, Shawn Cole and Jeremy Tobacman. 2011 “Marketing Complex Financial Products in Emerging Markets: Evidence from Rainfall Insurance in India.” *Journal of Marketing Research* 41: S150-S162.

Gayton, S. Jamie and Scott P. Handler. 2012. *Guide to Personal Financial Planning for the Armed Forces*. Mechanicsburg, PA: Stackpole.

Goldin Claudia. 2013. “Notes on Women and the Undergraduate Economics Major.” *CSWEP Newsletter*. (Summer):4-6, 15.

Grable, John E. and Ruth H. Lytton. 1998. Investor risk tolerance: testing the efficiency of demographics as differentiating and classifying factors. *Financial Counseling and Planning 9,* 61–74.

Grable, John E. 2000. “Financial Risk Tolerance and Additional Factors That Affect Risk Taking in Everyday Money Matters.” *Journal of Business and Psychology* 14(4): 625-630.

Hallahan, Terrence A., Robert W. Faff, and Michael D. McKenzie. 2004. "An Empirical Investigation of Personal Financial Risk Tolerance," Financial Service Review (13:1), pp. 57-78.

Hastings Justine S., Brigitte C. Madrian and William L. Skimmyhorn. 2012. “Financial Literacy, Financial Education and Economic Outcomes.” *Annual Review of Economics* 5: Submitted. doi: 10.1146/annurev-economics-082312-125807.

Hastings Justine S. and Olivia S. Mitchell. 2011. “How Financial Literacy and Impatience Shape Retirement Wealth and Investment Behaviors.” NBER Work. Pap. 16740

Hilgert, Marianne A., Jeanne M. Hogarth and Sondra G. Beverly. 2003. “Household Financial Management: The Connection Between Knowledge and Behavior.” *Fed. Reserve Bull* 89(7):309–22

Love, David A. 2013. “Optimal Rules-of-thumb for Consumption and Portfolio Choice.” *The Economic Journal* 123: 932–961. doi: 10.1111/ecoj.12002

Lusardi, Annamaria and Olivia S. Mitchell. 2014. "The Economic Importance of Financial Literacy: Theory and Evidence." *Journal of Economic Literature* 52(1): 5-44.

Lusardi, Annamaria and Olivia S. Mitchell. 2009. “How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness.” *NBER Working Paper 15350*.

Lusardi, Annamaria, and Olivia S. Mitchell. 2008. "Planning and Financial Literacy: How Do Women Fare?" *American Economic Review*, 98(2): 413-17.

Lusardi, Annamaria and Olivia S. Mitchell. 2006. *Financial literacy and planning: Implications for retirement well-being.* Working Paper No. 1, Pension Research Council, Wharton School, University of Pennsylvania.

Lusardi, Annamaria, Anya Savikhin Samek, Arie Kapteyn, Lewis Glinert, Angela Hung and Aileen Heinberg. 2014. “Visual Tools and Narratives: New Ways to Improve Financial Literacy.” NBER Working Paper No.20229.

Lyle, David S. 2009. “The Effects of Peer Group Heterogeneity on the Production of Human Capital at West Point.” American Economic Journal: Applied Economics, 1(4): 69-84.

Mandell, Lewis and Linda S. Klein. 2007. "Motivation and Financial Literacy." *Financial Services Review* 16: 105-116.

Martin, Matthew, A Literature Review on the Effectiveness of Financial Education (June 15, 2007). FRB Richmond Working Paper No. 07-03. Available at SSRN: http://ssrn.com/abstract=2186650

Meier, Stephan and Charles Sprenger. 2008. "Selection into Financial Literacy Programs: Evidence from a Field Study." Federal Reserve Bank of Boston Public Policy Discussion Paper No. 07-5.

Miller, Margaret, Julia Reichelstein, Christian Salas and Bilal Zia. 2014. “Can You Help Someone Become Financially Capable? A Meta-analysis of the Literature." World Bank Policy Research Working Paper No.6745.

Skimmyhorn, William. 2015. “Assessing Financial Education: Promising Evidence From Boot Camp.” Forthcoming, *American Economic Journal: Economic Policy*.

Tversky, Amos and Daniel Kahneman. 1974. “Judgment Under Uncertainty: Heuristics and Biases.” *Science* 185: 1124-1131.

United States Government Accountability Office. 2013. “Diversity Management: Trends and Practices in the Financial Services Industry Agencies after the Recent Financial Crisis,” Report to the Ranking Member, Committee on Financial Services, U.S. House of Representatives. April.

Van Rooij, Maarten., Lusardi, Annamaria, ALESSI, Rob. 2012. Financial literacy, retirement planning and household wealth, *Economic Journal*, Vol. 122 (May), pp.449-478.

Way, Wendy and Karen Holden. 2009. Teachers' Background and Capacity to Teach Personal Finance: Results of a National Study. *Journal of Financial Counseling and Planning* 20(2): 64-78.

Winter, Joachim Kathrin Schlafmann and Ralf Rodepeter. 2012. "Rules-of-thumb in Life-cycle Saving Decisions." *The Economic Journal* 122: 479–501. doi: 10.1111/j.1468-0297.2012.02502.x

Willis, Lauren E. 2011. "The Financial Education Fallacy." *American Economic Review* 101: 429-434.

Table 1: Personal Finance Curriculum & Methods

Lesson	Subject	Principles Based (PB) Methodology	Rules of Thumb (ROT) Methodology	Applicable Rule of Thumb	Hours
1	Personal Finance for Service Members	Goal Setting: Group Brainstorming Exercise	Goal Setting: Group Brainstorming Exercise to develop SMART goals	1.1 Develop SMART Goals and Track Your Progress	2
		Net Worth: Powerpoint Presentation of Equation	Net Worth Exercise: Online Calculator	1.2 Net Worth Equals Assets Minus Liabilities	
		Taxes: PPT and Board Work to Calculate Marginal vs Average Tax Rates	Taxes: PPT showing where to find taxes on pay stub	None	
		Budgeting: Group Brainstorming Exercise	Budgeting: 20/50/30 Brainstorming Worksheet	1.3 Pay Yourself First 1.4 Do Not Spend More than you Make 1.5 Create a Budget using the 20/50/30 Rule	
2	Personal Finance Basics/Major Financial Decisions	TVM: Excel Based Exercise with Explanation of Equations; Example Board Problems Using Equations	TVM: Online Calculator Exercise with Explanation of Equations; Example Board Problems Using Online Calculators	2.1 A Dollar Today is Worth More Than a Dollar Tomorrow	2
		Credit Card Balance Example Using Excel Based Equation	Credit Card Balance Example Using Online Calculator	2.2 Always Pay Your Bills on Time 2.3 Always Pay Off Your Credit Card Balance	
		Pay Day Loan Example Using Excel Based Equation	Pay Day Loan Example Using Online Calculator	None	
		New versus Used Car: Excel Based Exercise with NPV Equations	New versus Used Car: Online Calculator Based Exercise without Providing NPV Equations	2.4 An Automobile Costs $10,000 a year	
		Rent vs Buy Exercise: Online Calculator	Rent vs Buy Exercise: Online Calculator	2.5 Do Not Purchase a Home for More Than 2.5 Times Your Annual Income	
3	Investing for your Future	Diversification Exercise: Excel Based Efficient Portfolio Presentation and Online Calculator Based Exercise	Diversification Exercise: Online Calculator Based Exercise	3.1 Do Not Put all Your Eggs in One Basket 3.2 Invest in Low-Fee Index Funds	2
		DCA Exercise: Excel Based Equations	DCA Exercise: Online Calculator	3.3 Keep it Simple by Investing Monthly and Being Disciplined	
		Emergency Fund Exercise: Excel Based Equations	Emergency Fund Exercise: Online Calculator	3.4 Build an Emergency Fund 3.5 Inflation will Erode your Purchasing Power	
4	Retirement & Insurance	Traditional vs Roth IRA Exercise: Excel Based	Traditional vs Roth IRA Exercise: Online Calculator	4.1 Minimize Taxes by Investing Within a Tax-Sheltered Account	2
		Insurance Needs Exercise: Excel Based	Insurance Needs Exercise: Online Calculator	4.2 Always Carry the Appropriate Insurance	
				Total	8

Table 2. Summary Statistics

	Full Sample	Control Group	Treatment Group: Combined		Principles (PB)		Rules of Thumb (ROT)		
	Mean (SD) (1)	Mean (SD) (2)	Mean (SD) (3)	Diff. from Control (4)	Mean (SD) (5)	Diff. from Control (6)	Mean (SD) (7)	Diff. from Control (8)	Diff. from PB (9)
Panel A. Individual Characteristics									
Age	18.60	18.57	18.63	0.06	18.66	0.09	18.60	0.04	-0.05
	(1.11)	(1.05)	(1.16)	[0.37]	(1.15)	[0.29]	(1.17)	[0.67]	[0.588]
Female	0.15	0.14	0.16	0.02	0.15	0.01	0.18	0.04	0.03
	(0.36)	(0.35)	(0.37)	[0.31]	(0.35)	[0.79]	(0.38)	[0.16]	[0.291]
Black	0.08	0.10	0.08	-0.02	0.08	-0.02	0.08	-0.02	-0.001
	(0.28)	(0.29)	(0.26)	[0.27]	(0.27)	[0.36]	(0.26)	[0.33]	[0.950]
Hispanic	0.12	0.11	0.12	0.01	0.15	0.04	0.10	-0.01	-0.05
	(0.32)	(0.31)	(0.33)	[0.50]	(0.35)	[0.14]	(0.30)	[0.63]	[0.071]
Other Race	0.11	0.12	0.10	-0.02	0.08	-0.03	0.12	0.001	0.03
	(0.31)	(0.32)	(0.30)	[0.40]	(0.28)	[0.14]	(0.32)	[0.98]	[0.173]
White	0.69	0.68	0.70	0.02	0.69	0.02	0.71	0.03	0.02
	(0.46)	(0.47)	(0.46)	[0.44]	(0.46)	[0.67]	(0.45)	[0.38]	[0.672]
SAT Score	1317	1304	1326	22.32	1318	14.12	1335	30.74	16.62
	(124.97)	(122.08)	(126.33)	[0.01]	(120.89)	[0.13]	(131.37)	[0.002]	[0.117]
Prior Enlisted	0.21	0.22	0.20	-0.01	0.20	-0.01	0.20	-0.02	-0.005
	(0.41)	(0.41)	(0.40)	[0.62]	(0.40)	[0.74]	(0.40)	[0.62]	[0.886]
First year GPA	2.98	2.92	3.03	0.11	2.97	0.05	3.09	0.17	0.12
	(0.59)	(0.58)	(0.59)	[0.003]	(0.53)	[0.21]	(0.63)	[0.0003]	[0.014]
Observations	991	422	569		289		280		
Classes (Sections)	73	35	38		19		19		
Instructors	24	11	13		12		13		
Panel B. Covariate Regression Results									
Partial R2 for Indiv. Char.				0.017		0.016		0.024	0.020
p-value for F-Test of Joint Sig. of Indiv. Char.				0.512		0.861		0.254	0.255
Observations				986		706		698	568

Note: DOD Data. The table presents summary statistics from administrative and baseline assessmeent data. Standard deviations of each variable appear in parentheses and p-values for the differences in means appear in brackets. We describe the treatment groups in Section 2. In Panel B, the partial R-squared and p-values at the bottom of columns 4, 6, 8, and 9 report the results from Equation 3. In all cases, the observable characteristics are unrelated to the assigned treatment conditions.

Table 3. Financial Outcome Summary Statistics

Outcome	Description	Panel A. Control		Panel B. PB Treatment		Panel C. ROT Treatment	
		Pre	Post	Pre	Post	Pre	Post
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
1	Topical Knowledge, %	59.93	57.95	64.24	71.35	64.34	71.34
		(14.84)	(17.14)	(12.06)	(12.63)	(11.63)	(13.11)
2	Big 5 Knowledge, %	73.06	70.00	74.40	77.01	73.93	77.00
		(21.20)	(22.47)	(16.65)	(15.78)	(16.90)	(17.17)
3	Pr(Self-Assessed Knowledge≥7), %	22.97	20.33	19.72	41.73	20.71	41.43
		(42.11)	(40.30)	(39.10)	(49.44)	(40.60)	(49.35)
4	Pr (Self-Efficacy≥7), %	54.07	50.96	51.76	64.08	53.57	62.14
		(49.89)	(50.05)	(50.09)	(47.46)	(49.96)	(48.59)
5	Pr (Motivation to Learn≥7), %	78.71	71.53	82.57	83.80	81.07	82.50
		(40.99)	(45.18)	(36.70)	(35.70)	(39.24)	(38.06)
6	Pr(Likeliness to Seek Advice≥7), %	84.69	82.78	93.31	88.56	94.29	86.07
		(36.05)	(37.80)	(26.61)	(28.71)	(23.25)	(34.69)
7	Pr(Willingness to Take Risk≥7), %	44.74	44.98	38.20	52.99	37.50	51.43
		(49.78)	(49.81)	(48.83)	(49.88)	(48.50)	(50.07)
8	Loan Allocation to Long Term Savings, %	38.97	40.49	37.20	40.25	39.64	40.07
		(24.77)	(23.50)	(22.09)	(20.79)	(24.12)	(20.76)
9	New Problem Solving, %	-	42.34	-	59.98	-	61.19
		-	(32.09)	-	(30.22)	-	(30.82)

Note: DOD data. N=986. Outcomes described in Section 2. The final outcome (new knowledge) was only collected during the post assessment.

Table 4. OLS Estimates of Main Program Effects

Variable	(1) Topical Knowledge	(2) Big 5 Knowledge	(3) Self-Assessed Knowledge	(4) Self-Efficacy	(5) Motivation to Learn	(6) Likeliness to Seek Advice	(7) Self-Assessed Risk Pref	(8) Patience		(9) New Problem Solving
					Outcomes					
					Panel A. Combined Treatment vs. Control					
Control Mean	0.5894	0.7153	0.2165	0.5251	0.7512	0.8373	0.4486	0.3973		0.4234
PostxT	0.0907***	0.0566***	0.2463***	0.1543***	0.0840***	-0.0283	0.1454***	0.0153	**T**	0.1568***
	(0.0106)	(0.0154)	(0.0291)	(0.0291)	(0.0287)	(0.0326)	(0.0352)	(0.0145)		(0.0213)
R2	0.3019	0.1298	0.0711	0.0431	0.0500	0.0233	0.0633	0.0217		0.2208
Obs	1972	1972	1972	1972	1972	1972	1972	1972		986
					Panel B. PB Method vs. Control					
Control Mean	0.5894	0.7153	0.2165	0.5251	0.7512	0.8373	0.4486	0.3973		0.4234
PostxPB	0.0917***	0.0521***	0.2589***	0.1908***	0.0821**	0.0052	0.1538***	0.0409**	**PB**	0.1474***
	(0.0122)	(0.0183)	(0.0318)	(0.0327)	(0.0312)	(0.0320)	(0.0349)	(0.0166)		(0.0226)
R2	0.2840	0.1435	0.0680	0.0522	0.0616	0.0225	0.0715	0.0218		0.2247
Obs	1412	1412	1412	1412	1412	1412	1412	1412		706
					Panel C. ROT Method vs. Control					
Control Mean	0.5894	0.7153	0.2165	0.5251	0.7512	0.8373	0.4486	0.3973		0.4234
PostxROT	0.0897***	0.0613***	0.2334***	0.1168***	0.0860***	-0.0630	0.1368**	-0.0109	**ROT**	0.1625***
	(0.0117)	(0.0161)	(0.0322)	(0.0363)	(0.0293)	(0.0369)	(0.0551)	(0.0182)		(0.0344)
R2	0.2963	0.1290	0.0514	0.0387	0.0441	0.0280	0.0589	0.0212		0.2177
Obs	1396	1396	1396	1396	1396	1396	1396	1396		698
					Panel D. ROT Method vs. PB Method					
PB Mean	0.6775	0.7594	0.3038	0.5799	0.8455	0.9167	0.4670	0.3762		0.3022
PostxROT	-0.0020	0.0091	-0.0254	-0.0740*	0.0038	-0.0682***	-0.016	-0.0518**	**ROT**	0.0087
	(0.0110)	(0.0154)	(0.0267)	(0.0378)	(0.0188)	(0.0216)	(0.0599)	(0.0198)		(0.0392)
R2	0.2777	0.1137	0.0955	0.0480	0.0412	0.0209	0.0734	0.0319		0.1750
Obs	1136	1136	1136	1136	1136	1136	1136	1136		568

Note: DoD Data. Columns 1-8 report the Difference-in-Differences estimates for Equation 2 for each outcome listed. Column 9 reports OLS estimates of Equation 2 for the new knowledge outcome that was only included on the final assessment. All regressions include section fixed effects. Heteroskedasticity robust standard errors, clustered at the instructor level, are depicted in parentheses. ***, **, and * reflect p<0.01, 0.05, and 0.10 respectively.

Table 5. OLS Estimates of Heterogeneous Program Effects (PB vs. ROT)

	Outcomes									
Variable	(1) Topical Knowledge	(2) Big 5 Knowledge	(3) Self-Assessed Knowledge	(4) Self-Efficacy	(5) Motivation to Learn	(6) Likeliness to Seek Advice	(7) Self-Assessed Risk Pref	(8) Patience		(9) New Problem Solving
Panel A: Females										
Control Mean	0.6345	0.6976	0.0952	0.5238	0.6310	0.9643	0.1905	0.3640		0.4444
PostxROT	-0.018	-0.038	0.0323	-0.022	-0.199	-0.103	-0.002	-0.023	**ROT**	0.0658
	(0.0215)	(0.0406)	(0.0616)	(0.1342)	(0.1303)	(0.0691)	(0.0777)	(0.0490)		(0.0456)
R2	0.3221	0.2557	0.1190	0.1437	0.0825	0.0509	0.1736	0.0625		0.2602
Obs	184	184	184	184	184	184	184	184		92
Panel B: Low Math SAT Scores (Score≤600)										
Control Mean	0.6330	0.6964	0.3036	0.5179	0.8393	0.9464	0.5804	0.3634		0.4345
PostxROT	-0.016	-0.020	-0.084	-0.142	0.0013	-0.165***	0.0429	-0.033	**ROT**	0.0388
	(0.0187)	(0.0341)	(0.1062)	(0.1594)	(0.0690)	(0.0473)	(0.1277)	(0.0462)		(0.0541)
R2	0.2348	0.0912	0.1330	0.1163	0.1147	0.1118	0.1038	0.0870		0.1831
Obs	220	220	220	220	220	220	220	220		110
Panel C: Low Initial Knowledge Scores (Score≤0.55)										
Control Mean	0.5167	0.6750	0.2083	0.5139	0.7222	0.9861	0.4444	0.3977		0.5000
PostxROT	-0.019	0.0551	-0.168*	-0.117	-0.055	-0.100	-0.094	-0.042	**ROT**	-0.054
	(0.0426)	(0.0843)	(0.0857)	(0.1771)	(0.1489)	(0.0664)	(0.1114)	(0.0692)		(0.0000)
R2	0.3949	0.1935	0.1004	0.1252	0.1793	0.1695	0.1928	0.1502		0.2664
Obs	150	150	150	150	150	150	150	150		75
Panel D: Low Initial Motivation Scores (Score≤6)										
Control Mean	0.6391	0.7457	0.1196	0.4130	0.2717	0.9348	0.3804	0.3647		0.5072
PostxROT	0.0246	0.0223	-0.082	-0.092	0.0036	-0.094	-0.057	-0.048	**ROT**	0.0191
	(0.0228)	(0.0364)	(0.0945)	(0.1168)	(0.0609)	(0.0656)	(0.1107)	(0.0463)		(0.0633)
R2	0.3681	0.1980	0.1322	0.0619	0.4167	0.0584	0.1750	0.0662		0.1970
Obs	198	198	198	198	198	198	198	198		99

Note: DoD Data. Cols. 1-8 report the DD estimates for the outcome in each column for the group in each Panel. Col. 9 reports OLS estimates of the new knowledge outcome for the group identified in each Panel. All regressions include section fixed effects. Heteroskedasticity robust standard errors, clustered at the instructor level, are depicted in parentheses. ***, **, and * reflect p<0.01, 0.05, and 0.10.

Appendix Table 1. Financial Education Outcome Instrument

Outcome		Items (Correct Answers)
1	Topical Knowledge	1 You have assets and liabilities with the following values: Home: $150,000, Investment Accounts: $50,000, Bank Accounts: $3,000, Credit Card Debt: $500 Based on the information above, what is your total net worth?
		2 Why diversify your investments? (Because buying Intel and Microsoft exposes you to the same sector)
		3 If you have a child, a job, a home, and do not own a car, which of the following insurance policies should you most likely not purchase? (Renter's Insurance)
		4 Which of following best describes a financial goal? (Saving $30,000 for a down payment on a home in 7 years.)
		5 What is the difference between a Mutual Fund and an Exchange Traded Fund (ETF)? (A mutual fund is priced at the end of the trading day, and an ETF can be traded during the trading day)
		6 A ROTH IRA allows you to contribute income_____ (Post-tax, paying federal income taxes in the current year.)
		7 If Hannah has an average tax rate of 15% and a marginal tax rate of 25%, what rate will her next dollar of income be taxed at? (25%)
		8 What is the primary advantage of starting to save for retirement early? (You take advantage of compounding
		9 As you approach retirement, your investments should become _____? (Less Risky)
		10 If you invested for retirement in an IRA instead of a traditional account, you would have _____, given the same rate of return for both accounts. (More after-tax savings than if you invested in a taxable account)
		11 A budget is important for all of the following reasons (Both A[Spend less than you earn] and C [Track your expenses over time])
		12 What are the two most important determinants of your credit score? (Your credit usage and payment history)
		13 A fund with a front load means that _____. (Brokers get their commission up front)
		14 What is Dollar Cost Averaging? (Buying a fixed dollar amount of an investment regardless of the share price)
		15 If you have a marginal tax rate of 25%, what is your tax savings in the current year if you invest $1000 in a traditional IRA? ($250)
		16 When deciding between renting versus buying a home/condo/etc. which factor matters least in your financial analysis of the decision? (Prevailing interest rates for auto loans)
		17 Index funds are _____. (A specific type of mutual fund or ETF that matches a market index)
		18 _____ life insurance provides a stated benefit for a fixed period of time and fixed premium payment. (Term)
		19 Why is it important to understand your risk tolerance and time horizon when saving for short term, medium term, and long term goals? (Different savings and investment assets do not have the same interest rates)
		20 What financial asset can you purchase within your IRA account? (All of the Above [Stocks, Bonds, Mutual Funds, Exchange Traded Funds])
2	Big 5 Knowledge	1 Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow? (More than $102)
		2 Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, would you be able to buy more than today, exactly the same as today, or less than today with the money in this account? (Less than today)
		3 Do you think that the following statement is true or false: Buying a single company stock usually provides a safer return than a stock mutual fund? (False)
		4 Do you think that the following statement is true or false: A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest over the life of the loan will be less? (True)
		5 If interest rates rise, what will typically happen to bond prices? (They will fall)
3	Self-Assessed Knowledge	How would you assess your overall financial knowledge? (1=Very Low & 10=Very High)
4	Self-Efficacy	I am good at dealing with day-to-day financial matters, such as checking accounts, credit and debit cards, and tracking expenses. (1=Strongly Disagree & 10=Strongly Agree)
5	Motivation to Learn	I am motivated to learn about personal finance topics on my own. (1=Strongly Disagree & 10=Strongly Agree)
6	Likeliness to Seek Advice	When facing an important financial decision, how likely are you to seek assistance or advice? (1=Not At All Likely & 10=Very Likely)
7	Willingness to Take Risk	When thinking of financial investments, how willing are you to take risks? (1=Not At All Willing & 10=Very
8	Time Preference	Assume that you have just received $40,000 for your Cow Loan and that you have no other debts. Write the amount that you would allocate to each option in the space provided: Present consumption (0-6 months after receipt of loan), such as spring break, car, gifts, and clothing Short-term savings (6 - 18 months after receipt of loan), such as class ring, uniforms, and furnishings; Medium-term savings (within 10 yrs of graduation), such as wedding, real estate, and graduate school tuition; Long-term savings (more than 10 yrs in the future), such as
9	New Knowledge	1 One of your Soldiers asks your advice regarding what he should do with $1,000 he recently inherited. The Soldier has $2,500 in credit card debt with an APR of 18%, a $5,000 car loan with an APR of 6%, $500 in pay-day loan debt with an APR of 260%, and a $10,000 loan from his credit union at 7%. How much money from his inheritance do you recommend he allocate to each type of debt? ($500 to Payday Loan, $500 to Credit Card)
		2 Education savings. Imagine that you are a newly promoted Captain with two children and you are trying to decide how best to save for their college expenses. Assume that you have already expended your GI Bill benefit. Which of the following plans would afford you the most money available for your children's college expenses in 15-20 years? (A tax-advantaged savings account with an estimated real return of 5% wherein your savings contributions each year are made with after tax dollars, the contributions grow each year without being taxed, and you pay no taxes upon withdrawal (much like a ROTH IRA, but for college expenses).)
		3 Time value of money. You are interested in purchasing a home when you retire and you currently have $15,000 saved for a down payment. How long will it take you to double your down payment to $30,000 assuming a 6% real rate of return? (12 years)

Appendix Table 2. Teaching Methods by Time and Instructor					
		Course Time			
Instructor	Building	PM1	PM2	AM1	AM2
1	A	-	-	ROT	PB
2	B	ROT	ROT	PB	PB
3	B	PB	PB	ROT	ROT
4	B	PB	ROT	ROT	PB
5	B	ROT	PB	ROT	PB
6	B	PB	ROT	ROT	PB
7	B	PB	PB	ROT	ROT
8	B	ROT	ROT	PB	PB
9	B	PB	ROT	-	-
10	B	ROT	PB	-	-
11	A	-	-	-	PB
12	A	-	-	PB	ROT
13	B	-	-	-	ROT
PB Total		5	4	3	7
ROT Total		4	5	6	4
Total		9	9	9	11
Note: We describe the PB and ROT methods in Section 2.					

Appendix Table 3. OLS Estimates of Program Effects on Student Motivation/Attention

	Outcomes				
	(1)	(2)	(3)	(4)	(5)
	Overall	Academy Mascots	Algebra Problem	Word Puzzle	West Point Trivia
Variable	Attention, %	(Q7)	(Q15)	(Q35)	(Q47)
	Panel A. Combined Treatment vs. Control				
Control Mean	0.9031	0.9139	0.9342	0.8170	0.9474
PostxPB	0.0118***	0.0183	0.0124	0.0187	-0.0020
	(0.0000)	(0.0114)	(0.0186)	(0.0137)	(0.0240)
R2	0.0528	0.0211	0.0299	0.0258	0.0186
Obs	1972	1972	1972	1972	1972
	Panel B. PB Method vs. Control				
Control Mean	0.9031	0.9139	0.9342	0.8170	0.9474
PostxPB	0.0098	0.0167	0.0214	0.0011	-0.0000
	(0.0145)	(0.0204)	(0.0170)	(0.0286)	(0.0246)
R2	0.0576	0.0199	0.0282	0.0375	0.0241
Obs	1396	1396	1396	1396	1396
	Panel C. ROT Method vs. Control				
Control Mean	0.9031	0.9139	0.9342	0.8170	0.9474
PostxROT	0.0138	0.0199	0.0037	0.0358	-0.0039
	(0.0145)	(0.0217)	(0.0134)	(0.0317)	(0.0238)
R2	0.0586	0.0253	0.0427	0.0261	0.0181
Obs	1412	1412	1412	1412	1412
	Panel D. ROT Method vs. PB Method				
Control Mean	0.9366	0.9566	0.9740	0.8628	0.9531
PostxROT	-0.0040	-0.0032	0.0177	-0.0346	0.0038
	(0.0180)	(0.0200)	(0.0138)	(0.0378)	(0.0301)
R2	0.0307	0.0246	0.0148	0.0171	0.0240
Obs	1136	1136	1136	1136	1136

Note: DoD Data. Columns 1-5 report the DD estimates of Equation 1 for the outcome listed. Col. 1 is the total percentage correct and Cols. 2-4 are binary outcomes for the correct answer to each question. Heteroskedasticity robust standard errors, clustered at the instructor level, are depicted in parentheses. ***, **, and * reflect p<0.01, 0.05, and 0.10 respectively.

Appendix Table 4A. Alternate Standard Error Estimates for Main Program Effects

	Outcomes								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable	Topical Knowledge	Big 5 Knowledge	Self-Assessed Knowledge	Self-Efficacy	Motivation to Learn	Likeliness to Seek Advice	Self-Assessed Risk Pref	Patience	New Problem Solving
				Panel A. Combined Treatment vs. Control					
Control Mean	0.5894	0.7153	0.2165	0.5251	0.7512	0.8373	0.4486	0.3973	0.4234
PostxT	0.0907***	0.0566***	0.2463***	0.1543***	0.0840***	-0.0283	0.1454***	0.0153	**T** 0.1568***
	(0.0106)	(0.0154)	(0.0291)	(0.0291)	(0.0287)	(0.0326)	(0.0352)	(0.0145)	(0.0213)
Instr. Cluster SE p-value	*0.0000*	*0.0013*	*0.0000*	*0.0000*	*0.0078*	*0.3933*	*0.0004*	*0.3018*	*0.0000*
InstrxHour Cluster SE p-value	*0.0000*	*0.0002*	*0.0000*	*0.0000*	*0.0061*	*0.2577*	*0.0003*	*0.3374*	*0.0000*
Wild Bootstrap SE p-value	*0.0000*	*0.0020*	*0.0000*	*0.0000*	*0.0100*	*0.4060*	*0.0020*	*0.3320*	*0.0000*
R2	0.3019	0.1298	0.0711	0.0431	0.0500	0.0233	0.0633	0.0217	0.2208
Obs	1972	1972	1972	1972	1972	1972	1972	1972	986
				Panel B. PB Method vs. Control					
Control Mean	0.5894	0.7153	0.2165	0.5251	0.7512	0.8373	0.4486	0.3973	0.4234
PostxPB	0.0917***	0.0521***	0.2589***	0.1908***	0.0821**	0.0052	0.1538***	0.0409**	**PB** 0.1474***
	(0.0122)	(0.0183)	(0.0318)	(0.0327)	(0.0312)	(0.0320)	(0.0349)	(0.0166)	(0.0226)
Instr. Cluster SE p-value	*0.0000*	*0.0093*	*0.0000*	*0.0000*	*0.0150*	*0.8711*	*0.0002*	*0.0219*	*0.0000*
InstrxHour Cluster SE p-value	*0.0000*	*0.0038*	*0.0000*	*0.0000*	*0.0302*	*0.8371*	*0.0003*	*0.0274*	*0.0000*
Wild Bootstrap SE p-value	*0.0000*	*0.0100*	*0.0000*	*0.0000*	*0.0140*	*0.9040*	*0.0000*	*0.0280*	*0.0000*
R2	0.2840	0.1435	0.0680	0.0522	0.0616	0.0225	0.0715	0.0218	0.2247
Obs	1412	1412	1412	1412	1412	1412	1412	1412	706
				Panel C. ROT Method vs. Control					
Control Mean	0.5894	0.7153	0.2165	0.5251	0.7512	0.8373	0.4486	0.3973	0.4234
PostxROT	0.0897***	0.0613***	0.2334***	0.1168***	0.0860***	-0.063	0.1368**	-0.010	**ROT** 0.1625***
	(0.0117)	(0.0161)	(0.0322)	(0.0363)	(0.0293)	(0.0369)	(0.0551)	(0.0182)	(0.0344)
Instr. Cluster SE p-value	*0.0000*	*0.0010*	*0.0000*	*0.0040*	*0.0076*	*0.1015*	*0.0211*	*0.5552*	*0.0000*
InstrxHour Cluster SE p-value	*0.0000*	*0.0010*	*0.0000*	*0.0032*	*0.0147*	*0.0351*	*0.0107*	*0.5469*	*0.0000*
Wild Bootstrap SE p-value	*0.0000*	*0.0020*	*0.0000*	*0.0060*	*0.0060*	*0.1140*	*0.0260*	*0.5380*	*0.0000*
R2	0.2963	0.1290	0.0514	0.0387	0.0441	0.0280	0.0589	0.0212	0.2177
Obs	1396	1396	1396	1396	1396	1396	1396	1396	698
				Panel D. ROT Method vs. PB Method					
PB Mean	0.6775	0.7594	0.3038	0.5799	0.8455	0.9167	0.4670	0.3762	0.0000
PostxROT	-0.002	0.0091	-0.025	-0.074*	0.0038	-0.068***	-0.016	-0.051**	**ROT** 0.0087
	(0.0110)	(0.0154)	(0.0267)	(0.0378)	(0.0188)	(0.0216)	(0.0599)	(0.0198)	(0.0392)
Instr. Cluster SE p-value	*0.8539*	*0.5575*	*0.3493*	*0.0629*	*0.8392*	*0.0045*	*0.7795*	*0.0158*	*0.8257*
InstrxHour Cluster SE p-value	*0.8329*	*0.6396*	*0.5889*	*0.0869*	*0.9221*	*0.0088*	*0.7545*	*0.0097*	*0.7815*
Wild Bootstrap SE p-value	*0.8500*	*0.5940*	*0.3380*	*0.0780*	*0.8600*	*0.0080*	*0.8100*	*0.0200*	*0.7960*
R2	0.2777	0.1137	0.0955	0.0480	0.0412	0.0209	0.0734	0.0319	0.1750
Obs	1136	1136	1136	1136	1136	1136	1136	1136	568

Note: DoD Data. Columns 1-8 report the Difference-in-Differences estimates for Equation 2 for each outcome listed. Column 9 reports OLS estimates of Equation 2 for the new knowledge outcome that was only included on the final assessment. All regressions include section fixed effects. Heteroskedasticity robust standard errors, clustered at the instructor level, are depicted in parentheses. ***, **, and * reflect p<0.01, 0.05, and 0.10 respectively. We provide the p-values for the clustered standard errors, for standard errors clustered at the instructorxhour level, and for the Wild Bootstrap procedure in italics for each outcome for comparison purposes.

Appendix Table 4B. Alternate Estimation Models (Binary Treatment Controlling for Initial Scores) for Main Program Effects

	Outcomes							
Variable	(1) Topical Knowledge	(2) Big 5 Knowledge	(3) Self-Assessed Knowledge	(4) Self-Efficacy	(5) Motivation to Learn	(6) Likeliness to Seek Advice	(7) Self-Assessed Risk Pref	(8) Patience
				Panel A. Combined Treatment vs. Control				
Control Mean	0.5795	0.7000	0.2033	0.5096	0.7153	0.8278	0.4498	0.4049
T	0.1076***	0.0672***	0.2247***	0.1315***	0.1086***	0.0137	0.0981***	0.0016
	(0.0142)	(0.0172)	(0.0334)	(0.0286)	(0.0310)	(0.0229)	(0.0313)	(0.0147)
R2	0.4885	0.2417	0.1996	0.1702	0.2012	0.1270	0.2389	0.1327
Obs	986	986	986	986	986	986	986	986
				Panel B. PB Method vs. Control				
Control Mean	0.5795	0.7000	0.2033	0.5096	0.7153	0.8278	0.4498	0.4049
PB	0.1088***	0.0710***	0.2155***	0.1583***	0.1165***	0.0419*	0.1056***	0.0129
	(0.0142)	(0.0174)	(0.0360)	(0.0315)	(0.0286)	(0.0237)	(0.0311)	(0.0147)
R2	0.4781	0.2768	0.2069	0.1889	0.2137	0.1483	0.2574	0.1478
Obs	706	706	706	706	706	706	706	706
				Panel C. ROT Method vs. Control				
Control Mean	0.5795	0.7000	0.2033	0.5096	0.7153	0.8278	0.4498	0.4049
ROT	0.1023***	0.0612***	0.2340***	0.1021***	0.0964**	-0.0147	0.0855*	-0.0075
	(0.0149)	(0.0180)	(0.0355)	(0.0341)	(0.0355)	(0.0243)	(0.0492)	(0.0187)
R2	0.4926	0.2710	0.2123	0.1655	0.2088	0.1184	0.2396	0.1357
Obs	698	698	698	698	698	698	698	698
				Panel D. ROT Method vs. PB Method				
PB Mean	0.7135	0.7701	0.4201	0.6597	0.8507	0.9097	0.5451	0.4043
ROT	-0.0080	-0.0014	-0.0231	-0.0586**	-0.0193	-0.0601***	-0.0210	-0.0179
	(0.0058)	(0.0116)	(0.0237)	(0.0247)	(0.0181)	(0.0160)	(0.0419)	(0.0151)
R2	0.3820	0.1363	0.1626	0.1567	0.1703	0.1420	0.2229	0.1284
Obs	568	568	568	568	568	568	568	568

Note: DoD Data. Columns 1-8 report OLS estimates for the main program effects for each outcome listed using a binary treatment indicator in lieu of a Difference-in-Differences specification. The sample is restricted to one observation (post assessment) per student. We omit results for Column 9 since the main specification is an OLS model for the new knowledge outcome that was only included on the final assessment. All regressions control for the initial assessment scores for each outcome and include section fixed effects. Heteroskedasticity robust standard errors, clustered at the instructor level, are depicted in parentheses. ***, **, and * reflect p<0.01, 0.05, and 0.10 respectively.

Appendix Table 4C. Alternate Estimation Models (Logit) for Select Main Program Effects

	Outcomes								
Variable	(1) Topical Knowledge	(2) Big 5 Knowledge	(3) Self-Assessed Knowledge	(4) Self-Efficacy	(5) Motivation to Learn	(6) Likeliness to Seek Advice	(7) Self-Assessed Risk Pref	(8) Patience	(9) New Problem Solving
				Panel A. Combined Treatment vs. Control					
Control Mean			0.2033	0.5096	0.7153	0.8278	0.4498		
T	-	-	0.0828***	0.1196***	0.1266***	0.0477*	0.0692*	-	-
	-	-	(0.0292)	(0.0328)	(0.0346)	(0.0247)	(0.0376)	-	-
Pseudo-R2	-	-	0.0319	0.0244	0.0442	0.0241	0.0388	-	-
Obs	-	-	1972	1972	1972	1972	1972	-	-
				Panel B. PB Method vs. Control					
Control Mean			0.2033	0.5096	0.7153	0.8278	0.4498		
PB	-	-	0.0731**	0.0469	0.0985***	0.0742***	-0.0002	-	-
	-	-	(0.0311)	(0.0325)	(0.0235)	(0.0235)	(0.0304)	-	-
Pseudo-R2	-	-	0.0361	0.0305	0.0545	0.0291	0.0477	-	-
Obs	-	-	1412	1412	1412	1412	1412	-	-
				Panel C. ROT Method vs. Control					
Control Mean			0.2033	0.5096	0.7153	0.8278	0.4498		
ROT	-	-	0.0930***	0.0370	0.0728***	0.0536**	-0.0004	-	-
	-	-	(0.0314)	(0.0329)	(0.0274)	(0.0274)	(0.0344)	-	-
Pseudo-R2	-	-	0.0254	0.0261	0.0371	0.0277	0.0392	-	-
Obs	-	-	1396	1396	1396	1396	1396	-	-
				Panel D. ROT Method vs. PB Method					
PB Mean			0.4201	0.6597	0.8507	0.9097	0.5451		
ROT	-	-	-0.0020	-0.0114	-0.0216	-0.0157	-0.0136	-	-
	-	-	(0.0209)	(0.0198)	(0.0189)	(0.0189)	(0.0212)	-	-
Pseudo-R2	-	-	0.0334	0.0231	0.0410	0.0184	0.0392	-	-
Obs	-	-	1136	1136	1136	1136	1136	-	-

Note: DoD Data. Columns 3-7 report Logit marginal effect estimates for the main program effects for each outcome listed in lieu of an OLS specification. We omit results for non-binary outcomes (i.e., Columns 1, 2, 8 and 9). All regressions include section fixed effects. Heteroskedasticity robust standard errors, clustered at the instructor level, are depicted in parentheses. ***, **, and * reflect p<0.01, 0.05, and 0.10 respectively.

Appendix Table 5. OLS Estimates of Program Effects Using Alternate Thresholds for "High"

		Outcomes				
		(1)	(2)	(3)	(4)	(5)
Variable		Self-Assessed Knowledge	Self-Efficacy	Motivation to Learn	Likeliness to Seek Advice	Self-Assessed Risk Pref
		Panel A: "High" Outcome >= 6				
	Control Mean	0.4450	0.7201	0.8517	0.9139	0.6531
PostxT		0.3657***	0.1396***	0.0752**	0.0168	0.1657***
		(0.0378)	(0.0323)	(0.0318)	(0.0229)	(0.0332)
R2		0.1080	0.0502	0.0348	0.0191	0.0589
Obs		1972	1972	1972	1972	1972
		Panel B: "High" Outcome >= 7 (Main Specification)				
	Control Mean	0.2165	0.5251	0.7512	0.8373	0.4486
PostxT		0.2463***	0.1543***	0.0840***	-0.028	0.1454***
		(0.0291)	(0.0291)	(0.0287)	(0.0326)	(0.0352)
R2		0.0710	0.0423	0.0494	0.0212	0.0631
Obs		1972	1972	1972	1972	1972
		Panel C: "High" Outcome >= 8				
	Control Mean	0.0634	0.3218	0.5443	0.7356	0.2057
PostxT		0.0707***	0.1468***	0.0741**	-0.051	0.0878***
		(0.0196)	(0.0352)	(0.0307)	(0.0373)	(0.0268)
R2		0.0302	0.0289	0.0345	0.0141	0.0409
Obs		1972	1972	1972	1972	1972

Note: DoD Data. Columns 1-5 report the Difference-in-Differences estimates for the outcome listed in each column using a "High" level indicator set at the number identified in each panel. All regressions include the coaviariates specified in Equation 1 and section fixed effects. Heteroskedasticity robust standard errors, clustered at the instructor level, are depicted in parentheses. ***, **, and * reflect p<0.01, 0.05, and 0.10 respectively.