American Economic Journal: Economic Policy 2016, 8(2): 322–343
http://dx.doi.org/10.1257/pol.20140283

Assessing Financial Education: Evidence from Boot Camp[†]

By WILLIAM SKIMMYHORN*

This study estimates the effects of Personal Financial Management Course attendance and enrollment assistance using a natural experiment in the US Army. New enlistees' course attendance reduces the probability of having credit account balances, average balances, delinquencies, and adverse legal actions in the first year after the course, but it has no effects on accounts in the second year or credit scores in either year. The course and its enrollment assistance substantially increase retirement savings rates and average monthly contributions, with effects that persist through at least two years. The course has no significant effects on military labor market outcomes. (*JEL* D14, I21, J45)

Financial literacy and education remain popular topics among the media, policymakers, and academics. In the United States, increasing personal responsibility for retirement planning and concerns over savings rates have generated calls for more financial education. Federal government responses have included President George W. Bush's 2008 Financial Literacy Advisory Council, President Obama's 2009 financial literacy campaign, and no less than 16 federal programs among 14 agencies (*Government Accountability Office* (GAO) 2012). Yet there exists little robust evidence that financial education improves individuals' economic decision making.

In this paper, I estimate the causal effects of financial education and enrollment assistance on several financial outcomes using administrative data related to the 2007–2008 rollout of a Personal Financial Management Course (PFMC) in the US Army.[1] The data provide information on credit decisions, retirement savings and labor market outcomes. Staggered implementation of the course across locations

* Office of Economic and Manpower Analysis, Department of Social Sciences, United States Military Academy, 607 Cullum Road, West Point, NY 10996 (e-mail: william.skimmyhorn@usma.edu). I thank Brigitte Madrian, David Laibson, John Smith, David Lyle, James Lee, Luke Gallagher, Susan Carter, Antoinette Schoar, Francis Murphy, Hilary Hoynes, seminar participants at Harvard University, the National Bureau of Economic Research (NBER) Summer Household Finance Session, the Boston Federal Reserve Bank, West Point, the Global Financial Literacy Excellence Center (GFLEC) Financial Literacy Seminar and three anonymous referees for their comments and assistance in developing this research. The opinions expressed herein reflect the personal views of the author and not those of the US Army or the Department of Defense.

[†] Go to http://dx.doi.org/10.1257/pol.20140283 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[1] Bell, Gorin, and Hogarth (2009) evaluated the pilot PFMC at Fort Bliss, Texas and found small beneficial effects. Unfortunately, the use of self-reported data, the low survey response rate and the nonexperimental comparison group limit the reliability of their evaluation, leaving open the question of the PFMC's causal effects.

and time provides exogenous variation in financial education exposure. Using individually matched credit bureau data, I find that the course reduces the probability of having positive debt balances, actual combined account balances, and the probability and number of adverse outcomes (e.g., account delinquencies and legal actions) in the first year after the course. I also find that course attendance and its coupled enrollment assistance have substantial effects on retirement savings contributions through at least two years. The course has no significant effects on adverse employee turnover, current productivity, or retention eligibility and decisions, which are outcomes of interest to employers considering financial education.

To date, the existing research on financial education has struggled to demonstrate a causal relationship between education and behavior. Hastings, Madrian, and Skimmyhorn (2013) provide a recent and detailed review and highlight the challenges for research on this issue. While there is widespread evidence of financial illiteracy (Lusardi 2004, Lusardi and Mitchell 2007), convincing evidence that literacy correlates with financial outcomes (Lusardi and Mitchell 2007) and motivation from advocacy groups and some government agencies to provide more services to individuals, especially the poor, given new evidence on the effects of poverty and time scarcity (Mullainathan and Shafir 2013), the evidence on the effects of education is mixed. Two leading meta-analyses provide mixed conclusions on the causal effects of financial education. Fernandes, Lynch, and Netemeyer (2014) suggest a prior expectation of small effects of financial education on behavior; while Miller et al. (2014) suggests the potential for improving behavior with respect to some outcomes (e.g., savings behavior) but not others (e.g., reducing defaults). In the United States, recent studies that employ experimental or quasi-experimental procedures (e.g., Choi, Laibson, and Madrian 2011; Duflo and Saez 2003; and Cole, Paulson, and Shastry 2013) fail to provide convincing evidence on the causal effects of financial education. Outside the United States, there are positive findings for education in selected contexts including farmers' insurance decisions in India (Gaurav, Cole, and Tobacman 2011) and micro-entrepreneurs' accounting behaviors in the Dominican Republic (Drexler, Fischer, and Schoar 2014). Perhaps unsurprising given the extant findings, Willis (2011); Hastings, Madrian, and Skimmyhorn (2013); and Fernandes, Lynch, and Netemeyer (2014) have all questioned the cost-effectiveness of additional publicly funded financial education.

My research contributes to this literature by exploiting plausibly exogenous variation that enables causal estimates of the effects of financial education. It analyzes a variety of new outcomes covering multiple portions of household balance sheets and labor market decisions and it uses rich administrative data that avoids the potential biases of self-reported outcomes and enables estimation of heterogeneous treatment effects. Given the sample's characteristics and diversity, the results provide direct evidence on a population of substantial policy interest: US military service members and important evidence on the effects of financial education for young, moderate-income workers. The paper proceeds as follows: Section I describes the PFMC; Section II summarizes the data; Section III provides the empirical framework; Section IV presents the results; Section V discusses the findings; and Section VI concludes.

Table 1—PFMC Implementation Schedule

Date	Location	Percent of sample
June 2007	Aberdeen Proving Grounds, MD	9.2 percent
	Fort Jackson, SC	4 percent
October 2007	Fort Benning, GA	18.4 percent
November 2007	Fort Eustis, VA	10.2 percent
	Fort Lee, VA	15.2 percent
December 2007	Fort Huachuca, AZ	0.9 percent
March 2008	Fort Sam Houston, TX	7 percent
April 2008	Fort Rucker, AL	1 percent
May 2008	Fort Leonard Wood, MO	12.6 percent
June 2008	Fort Knox, KY	5 percent
	Redstone Arsenal, AL	1.8 percent
August 2008	Fort Gordon, GA	6.7 percent
	Fort Sill, OK	8.1 percent

Note: Percentages reflect the fraction of the administrative data sample (observations = 82,211) from each location.

Source: Author compilation using Department of Defense (DOD) and Army Emergency Relief (AER) data

I. A Unique Natural Experiment

Between June 2007 and August 2008, the US Army's nonprofit relief society, Army Emergency Relief (AER), implemented an eight-hour financial education course. The PFMC was mandatory for new enlistees as part of their job-specific Advanced Individual Training (AIT), which all enlisted soldiers attend following basic training. The purpose of the PFMC was "to assist Service men and women and their immediate families in their efforts to building personal wealth through reducing debt and establishing savings goals" (AER-DOD Memorandum of Understanding (MOU) 2003). AER designed and executed the course with San Diego City College (SDCC). See Table 1 for the PFMC implementation schedule. As described in Section IV, I exploit this unique rollout to provide causal evidence on the effects of financial education.

Treatment includes education, assistance in signing up for savings plans, and advice provided by the instructors during breaks or in response to specific questions. The course was typically completed in two sessions in which civilian instructors, trained by SDCC, gave lectures on the topics in Table 2, following standardized slides and course booklets. The PFMC hours replaced eight hours of leisure time for new soldiers.

Whether an eight-hour course is sufficient to meet the program's objectives is unclear. On the one hand, this seems too short given the financial literacy required to succeed in today's economy. On the other hand, time is a commodity in short supply for training programs and a longer course may not be justified if additional time has diminishing returns. Of note, Schreiner, Clancy, and Sheradden (2002) found that an education program for Individual Development Accounts increased savings for low-income households with diminishing effects after eight to ten hours. Drexler, Fischer, and Schoar (2014) found positive effects from an accounting course lasting only 15 hours. Importantly, a course of relatively short duration may have limited

TABLE 2—PFMC CURRICULUM

Lesson	Subject	Topics	Hours
1	Financial ethics	Legal, moral, and ethical aspects of personal financial management	0.75
2	Leave and earnings (pay) statement	Understanding pay statements, military benefits and insurance coverage, educational benefits, payroll deductions, and resolving pay problems	0.25
3	Developing a spending plan	Net worth, debt-to-income ratios, discretionary versus non-discretionary spending	1
4	The essentials of credit	Types of credit, factors affecting credit worthiness, proper credit usage, warning signs of too much debt, credit and debt assistance, consumer protection laws, credit reports	1
5	Consumer awareness	Psychology of advertising, types of deception, identity theft recognition and correction, description of common scams	1
6	Car buying	Personal budget review, contract tips, determining fair price, negotiation tips, effects of car ownership in the military, financing, consumer protection	1.5
7	Meeting your insurance needs	Renters and homeowners, automobile, life, health, insurance frauds and scams, protection tips	0.5
8	Thrift savings plan and investing	Retirement concepts, the thrift savings plan, military retirement programs, compound interest, investments	2
		Total	**8**

Source: Author compilation based on AER and SDCC information on the PFMC

effects on behaviors involving complex combinations of analytic skills, life experience, and self-control.

The PFMC covered both principles (e.g., the time value of money) and some rules of thumb (e.g., obtain a copy of your credit report annually), and focused on the financial decisions young workers are most likely to face (buying a car is included; buying annuities is not). For outcomes related to credit and the labor market, treatment should be thought of as education. For the TSP outcomes, treatment should be thought of as education coupled with assistance, since instructors may have assisted with TSP enrollment at some locations.[2]

II. Army and Credit Bureau Administrative Data

A. *Military Administrative Data*

I use administrative data from the army and a national credit bureau and analyze course topics including credit decisions (e.g., debt levels and delinquencies), retirement savings (e.g., the Thrift Savings Plan), and labor market outcomes (e.g., adverse separations and reenlistment decisions). The military data is a repeated

[2] Information is based on author interviews with AER, SDCC, and PFMC instructors (2011–2012). TSP enrollment assistance varied by location and time (e.g., at some locations enrollment forms were distributed; at others SDCC personnel assisted in form completion and/or submission). Unfortunately, neither AER nor SDCC collected detailed data on the variation in assistance and I cannot separately identify the effects of education and assistance for TSP outcomes.

cross-section and covers all active duty army soldiers entering service from May 2006 to June 2009. I restrict the sample to individuals starting AIT at each location within 12 months of PFMC implementation to minimize time-varying enlistment differences. This generates an administrative data sample of $N = 82{,}211$ individuals for my analyses in the first year after an individual starts AIT. Since individuals progressively leave the military, my samples for years 2–4 are reduced to $N = 70{,}785$, $N = 59{,}765$, and $N = 44{,}946$, respectively. Online Appendix Table 3 shows that attrition is unrelated to treatment.

To avoid contamination between my experimental groups, I omit individuals starting AIT in the month preceding, month of, or month following PFMC implementation. I also omit those individuals whose AIT start date and course length produce overlap with the PFMC implementation date.[3] Since I assume that individuals were treated in the month they began AIT, measurement error arising from outcome observation prior to treatment makes my estimates lower bounds. The army demographic data, measured at AIT start, contain a rich set of characteristics potentially related to financial decision making, including demographic data (age, gender, marital status, number of dependents, and race), human capital data (education, Armed Forces Qualification Test (AFQT) scores and enlistment timing), and economic factors (length of AIT, deployments, and compensation).

B. *Credit Market Outcomes*

Since several of the course topics relate to credit behavior and decisions, I evaluate a number of credit outcomes of interest. Given the cost of the credit bureau data, I match a random subsample of individuals to their credit bureau data from April of each year from 2007 to 2010.[4,5] The credit data contributes to this analysis in several ways. First, the outcomes directly measure the causal effects of education unconfounded by other factors such as direct assistance. Second, the data provide a relatively complete financial profile for several outcomes related to the PFMC's topics, even though it does not capture payday loans, auto title loans, or informal lending. Finally, the data enables more precise estimation of the PFMC effects since it allows me to control for individuals' baseline credit outcomes (e.g., credit scores and balances) for those with credit records for the year prior to their entry into the military.

I focus my analysis on PFMC program goals (i.e., reducing debt) and topics (e.g., develop a spending plan, essentials of credit) by analyzing several credit outcomes in the first two years after AIT. I analyze routine outcomes including credit scores

[3] I also omit treatment group individuals based on this criterion to achieve balance on the AIT length characteristic. I omit individuals for whom the absolute value of their AIT start month cohort (i.e., $[-12, 12]$) minus the length of their AIT course (i.e., $[1, 12]$) is greater than or equal to zero. This results in smaller samples near the month of implementation at each location.

[4] I randomly selected $N = 39{,}485$ records for matching. To test whether treatment is related to the probability of attempted match, I regress an indicator for membership in the subsample on my treatment variable and location and time fixed effects. The coefficient is 0.1043 with a clustered standard error p-value of 0.071 and a cluster wild bootstrap ($N = 1{,}000$ iterations) of $p = 0.198$. This suggests that my credit sample is a random subsample.

[5] Eighty-four percent of the $N = 39{,}485$ Year 1 records matched. 85 percent of the $N = 28{,}496$ Year 2 records matched. See Table 4 (column 1) for evidence that treatment is unrelated to the probability of having a matched credit record.

and having active credit.[6] I also analyze the probability of any balance and the actual balance for five outcomes (credit card balance, automobile balance, finance loan balance, mortgage balance, and the aggregate of all four).[7,8] Finally, I analyze probability of having any accounts in an adverse status and the number of these accounts for three outcomes (adverse legal actions [sum of foreclosures, liens, judgments, and bankruptcies], accounts 60 days past due, and accounts 120 days past due).[9] I winsorize the credit balance and count outcomes at the first and ninety-ninth percentiles. See online Appendix Table 4 and online Appendix Table 6 for non-winsorized results, which are similar to my main results. I measure the credit outcome horizons using annual cross-sections of credit data from the first April after AIT completion (e.g., Year 1 outcomes reflect data from the April that falls between month 1 and 12) and on average, I observe individuals in their first year 6 months after the course and individuals in their second year 18 months after the course.

C. *Retirement Savings Outcomes*

Since the single most significant portion of the curriculum (two of eight hours) is dedicated to retirement savings and the Thrift Savings Plan (TSP), I evaluate TSP decisions (average monthly contributions and the probability of participation in each year) for an individual's first four years in the military. The TSP is a tax-advantaged retirement program available to federal employees, including the military, with participation rules similar to those of a 401(k). Initial enrollment must occur via a paper form; subsequent contributions must occur via payroll deduction; changes can be made online or at a military finance office; and there is a loan option. These features and my use of payroll data minimize the chance of unobserved savings and withdrawals. While military members do not receive matching funds (the army has a separate defined benefit pension), contributions were typically tax-deferred,[10] and individuals could select from several funds, all of which had low expense ratios.[11] In addition, staff members often provided enrollment assistance (e.g., completing and/or submitting enrollment forms) to soldiers.

Whether saving for retirement in a tax-deferred account is optimal for new enlistees is an open question. Financial education might affect decision making by improving numeracy (e.g., computing net present values), increasing literacy (e.g., demonstrating the costs of minimum credit balance payments or the tax advantages

[6] To preserve the credit sample size, I assign zeros for records that are matched but coded as inactive since businesses and the credit bureau have the incentive to report all account balances. See Table 4 for evidence that treatment is unrelated to credit matching or activity. In unpublished results, available upon request, I complete robustness checks for all credit regressions with only matched and active records. The results are nearly identical.

[7] The aggregate balance outcome is the sum of the other four balances listed. Automobile balances include automobile loans and leases. Finance loans are personal loans, credit union loans, or revolving lines of credit from agencies like sales finance companies.

[8] I omit mortgage balances from my analysis given their low prevalence (1 percent to 2 percent) in my sample. This low prevalence is unsurprising given the young, primarily single, mobile, and low-income nature of the sample.

[9] While 30 days past due is often used by credit bureaus, I use the 60-day horizon based on my data availability.

[10] On October 1, 2012 the TSP established a Roth (post-income tax, tax-free) option for all members. Time fixed effects account for this change.

[11] The default fund is a government securities fund. Other funds include fixed income securities, common stock, small cap stock, international stock, and lifecycle funds. Since 2006, the average net expense ratio has not exceeded 0.031 percent.

of the TSP), or lowering enrollment costs (psychological or time). The final two outcomes seem especially likely in light of the PFMC's bundled education and assistance. So while economists might argue over the optimality of the course and its welfare effects, I forego a formal welfare analysis. Instead I note that the army decided to conduct the course and I evaluate it against its stated goals of increasing savings and reducing debt.

I observe monthly TSP contributions and measure the TSP outcome horizons relative to an individual's AIT start month (e.g., Year 1 outcomes reflect an individual's average monthly TSP contributions from the month after they start AIT through 12 months, and the participation indicator reflects any contributions made during this same period). I winsorize the annual monthly average TSP contributions at the first and ninety-ninth percentiles. Mean control group participation is 12 percent, 15 percent, 16 percent, and 17 percent in Years 1 through 4, respectively. While I observe all TSP contributions, my view of an individual's retirement portfolio is incomplete as I cannot observe IRAs or other household 401(k) accounts. But the TSP is an important part of active duty army members' retirement plans, and the incomplete picture is less concerning for this young population with limited labor market experience.

D. *Military Labor Market Outcomes*

Since financial stress may undermine job performance (Carrell and Zinman 2014, AER-DOD MOU 2003), I use army data to evaluate four labor market outcomes potentially related to financial decision making.[12] To evaluate job performance, I observe whether an individual is adversely separated from the military. To evaluate current productivity, I observe whether an individual is rapidly promoted to a supervisory position (sergeant) during their first term and whether they are offered the opportunity to extend their service by re-enlisting. These outcomes may reflect an individual's ability to focus more on job performance when they have a better financial situation. Finally, to evaluate firm attachment, I observe whether individuals opt for another term in the army given the opportunity to re-enlist. I measure these labor market outcomes during an individual's first enlistment term. While employer-employee relations in the military differ from those in other public and private sector jobs, the US military is a volunteer force and these outcomes might provide insight into whether employer-funded financial education offers a return on investment in the form of lower turnover or higher productivity. Since these outcomes use restricted samples from the full administrative sample for those with terms less than or equal to four years, I evaluate whether treatment is correlated with presence in these samples. The results in online Appendix Table 2, panel B, show that treatment is unrelated to presence in my labor market samples; thus, my estimates might reasonably be applied to all new enlistees.

[12] The army administrative data, which is only available through August 2013, censors my visibility of the final treatment group (AIT started August 2009) to those with initial enlistment terms less than or equal to four years. As a result, I complete these analyses for those with initial terms less than or equal to four years. Online Appendix Table 2 shows that presence in this sample is uncorrelated with treatment.

III. Using the Staggered Rollout to Estimate the PFMC's Effects

While the PFMC implementation month at each base is known, the exact course start dates are not. In addition, individual level data on course attendance is unavailable. As a result, I impute my treatment variable (PFMC attendance) and define an individual's treatment status using their estimated AIT start date relative to the PFMC start date at their AIT location. See online Appendix Table 1 for a description of the imputation procedure, which uses administrative data on individual entry dates, basic training durations, and future assignments (including individuals' AIT locations). The procedure may lead to a downward bias on my estimates for at least two reasons. First, if control group members experience training or travel problems that delay their AIT arrival, then they may actually attend the PFMC once the course commences at that location. Second, I assume that all individuals receive treatment but some may in fact miss the training (e.g., due to medical appointments or other ad hoc training requirements).

Individuals who started AIT after the PFMC began at their location are assigned a value of one for treatment; those who started before are assigned a value of zero. As a result, the treatment and control groups at each location are separated by time. Note that at each base, control group members systematically precede treatment group members. Across bases though, there is overlap between treatment group members at the early adoption locations and control group members at the late adoption locations. This rollout enables me to estimate the effects of the course using location fixed effects, which allow me to compare individuals at any given base over time, and time (month-year) fixed effects, which allow me to compare individuals in any given month across locations. As a result, my reduced form estimates in equation (1) reflect the average effect of the PFMC on individual financial outcomes:

$$(1) \qquad Y_i = \alpha + \beta \cdot PFMC_i + X_i'\gamma + \varphi_j + \delta_t + \varepsilon_i.$$

In this model Y_i is a financial or labor market outcome for individual i who started AIT at location j in time period t. I suppress the j and t subscripts for clarity since I only have one observation per individual, but the data includes financial outcomes at multiple time horizons. $PFMC_i$ is the binary treatment variable that equals one if the individual attended the course and equals zero otherwise. X_i is a vector of individual characteristics that potentially affect financial decision making, including a quadratic in age, gender, race, marital status, number of dependents, education level, AFQT score, a summer enlistment indicator, enlistment term length, AIT course length, average monthly income, and the number of months the individual was deployed during the year. For the credit market outcomes, X_i also includes the credit score and the appropriate credit outcome (both with a missing indicator as appropriate) from the pre-AIT year. φ_j is a vector of location fixed effects, δ_t is a vector of time (month-year) fixed effects and ε_i is the error term.

Since the course advises soldiers to establish a budget and reduce their debt levels, I expect positive signs on all TSP outcomes and credit scores, and I expect negative signs on all other credit outcomes. However, there is the possibility that the course could increase a soldier's ability to secure better interest rates and that this could lead him to take on more debt (e.g., auto loans or credit cards) with comparable payment levels. Unless otherwise specified, I report only the main treatment effect estimates (β). I cluster the standard errors at the treatment location level ($N = 13$ clusters). Given the small number of clusters, I complete the cluster wild bootstrap procedure (Cameron, Gelbach, and Miller 2008) and provide its p-values along with the clustered p-values for reference. Overall, the clustering strategy has little effect on my results.

A. *Summary Statistics and Covariate Regressions Suggest Valid Identification*

Identification of causal estimates of the PFMC effects on financial outcomes requires that conditional on an individual's AIT location, start month, and individual characteristics, treatment assignment is unrelated to other potential determinants of the outcomes. A number of features of the PFMC implementation plan suggest a potentially valid natural experiment. First, the details of the program implementation, unannounced and staggered across locations and time, support an expectation of exogenous variation. Second, implementation dates were determined by AER Headquarters and SDCC based on discussions with local military leaders without notifying or soliciting information from individual soldiers or the US Army's Recruiting Command. So there is little reason to believe that potential enlistees had any knowledge of the PFMC or an ability to change their enlistment timing or their job based on PFMC start dates.[13] Third, selection into or out of the course seems unlikely since the eight-hour course duration is trivial when compared to the much longer (1–12 month) AIT course. Moreover, the decision to enlist is a significant career choice likely unrelated to the PFMC. I also mitigate concerns over strategic implementation by using location (base) fixed effects in all regression specifications, which remove the average effects for each location. I also note that army commanders cannot affect implementation timing based on any outcome trends, since they have no advance notice about the characteristics of their incoming recruits nor do they have any visibility on the recruits' TSP savings rates or credit outcomes.

In addition to the institutional reasons above, I complete more formal analyses in order to demonstrate plausibly exogenous variation in exposure to financial education. First, I provide raw and regression adjusted summary statistics in Table 3. The raw summary statistics reveal substantial covariate balance across control (columns 1 and 4) and treatment (columns 2 and 5) groups in both samples. Additionally, in columns 3 and 6, I provide regression adjusted estimates for each characteristic and find very few differences between the groups that are statistically significant (three characteristics in the full sample and one in the credit subsample). More importantly, univariate balance is not required for every characteristic; instead my identification assumption requires that the two groups are similar given the conditional expectation function: $E(\varepsilon_{ijt} | X_i, \varphi_j, \delta_t) = 0$. To provide further evidence in

[13] Information is based on author interviews with AER personnel and the SDCC contract leader (2011–2012). Both parties reported that the PFMC implementation schedule was driven by the ability to recruit and train instructors. In fact, neither AER nor SDCC had any data on soldier characteristics, further minimizing concerns over non-random implementation on the basis of individual characteristics.

TABLE 3—INDIVIDUAL CHARACTERISTICS BY SAMPLE AND TREATMENT GROUP

	Panel A Full administrative data sample $N = 82{,}211$			Panel B Matched credit subsample $N = 33{,}178$		
Variable	No PFMC $N = 40{,}844$ **Mean** **(SD)** (1)	PFMC $N = 41{,}367$ **Mean** **(SD)** (2)	Regression adj. diff. **Coeff.** **(SE)** (3)	No PFMC $N = 16{,}740$ **Mean** **(SD)** (4)	PFMC $N = 16{,}438$ **Mean** **(SD)** (5)	Regression adj. diff. **Coeff.** **(SE)** (6)
Age, years	21.4	21.5	0.06	21.4	21.6	0.07
	(4.1)	(4.1)	(0.10)	(4.0)	(4.1)	(0.13)
Female, percent	14.9	15.9	−0.006	11.4	12.1	−0.01
	(35.6)	(36.6)	(0.011)	(31.8)	(32.7)	(0.02)
Married, percent	17.6	19.0	−0.002	17.9	19.0	0.001
	(38.1)	(39.3)	(0.008)	(38.3)	(39.3)	(0.014)
Dependents	0.4	0.5	0.037**	0.4	0.5	0.02
	(0.9)	(1.0)	(0.015)	(0.9)	(1.0)	(0.03)
Minority, percent	30.8	33.6	0.009	29.2	31.8	0.01
	(46.2)	(47.2)	(0.014)	(45.5)	(46.6)	(0.02)
< HS education, percent	28.8	24.5	−0.03***	29.8	24.9	−0.03***
	(45.3)	(43.0)	(0.009)	(45.7)	(43.3)	(0.008)
HS graduate, percent	62.6	65.9	0.042***	61.6	65.3	0.02*
	(48.4)	(47.4)	(0.008)	(48.6)	(47.6)	(0.01)
Some college, percent	6.2	6.7	−0.002	6.2	6.9	0.01
	(24.1)	(25.0)	(0.004)	(24.1)	(25.4)	(0.01)
≥ college grad, percent	2.4	2.9	−0.004	2.4	2.8	−0.004
	(15.4)	(16.9)	(0.004)	(15.3)	(16.6)	(0.004)
AFQT, percentile	55.9	56.1	−1.03	56.3	57.3	−1.64
	(19.4)	(19.8)	(1.11)	(19.3)	(19.0)	(1.14)
Joined in summer, percent	38.1	36.0	−0.06	37.1	33.5	−0.07
	(48.6)	(48.0)	(0.06)	(48.3)	(47.2)	(0.05)
Enlistment term, year	3.8	3.8	0.008	3.9	3.8	0.009
	(1.0)	(1.0)	(0.070)	(1.0)	(1.0)	(0.065)
AIT length, months	3.2	3.1	0.06	3.2	3.2	0.01
	(1.1)	(1.1)	(0.07)	(1.1)	(1.1)	(0.08)
Monthly pay, $	1,757	1,882	2.8	1,758	1,880	4.6
	(542)	(579)	(20.3)	(542)	(576)	(29.7)
Months deployed	1.2	1.5	0.21	1.2	1.6	0.27
	(2.3)	(2.7)	(0.25)	(2.3)	(2.7)	(0.23)
Prior credit score	—	—	—	557	554	−0.02
	—	—	—	(106)	(109)	(0.02)
Matched prior credit record, percent	—	—	—	52.7	56.2	1.9
	—	—	—	(49.9)	(49.6)	(4.1)
Joint test of significance:						
Partial R^2			0.0002			0.0002
p-value from F-test			0.128			0.168

Notes: Married represents formal and common law marriages for anyone ever married. Less than high school variable includes dropouts and GED holders. Mean AFQT percentiles exceed 50 due to enlistment prohibitions for low scores. Average monthly pay represents the mean base pay, subsistence pay, and housing allowance during the first year. Months deployed variable reflects the number of months that an individual received hostile fire pay during the first year. Prior credit score statistics are restricted to individuals with a pre-AIT credit score ($n = 18{,}054$). Columns 3 and 6 report the coefficients and standard errors from an OLS regression of the individual characteristics on the treatment indicator, time fixed effects, and location fixed effects. The bottom panel reports the partial R^2 values and p-values from an F-test for the joint significance of all individual characteristics (omitting high school graduate indicator and adding a quadratic term in age) from an OLS regression of equation (2) with standard errors clustered at the location level (observations = 13).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Source: DOD Data

support of this identification assumption, I use the relationship between my observable characteristics and treatment to model the relationship between unobservable characteristics and treatment in the spirit of Altonji, Elder, and Taber (2005) using equation (2):

$$PFMC_i = \rho + X_i'\sigma + \varphi_j + \delta_t + \mu_i. \tag{2}$$

In this specification, I regress my treatment variable ($PFMC_i$) on the individual characteristics and fixed effects, and I evaluate whether or not these characteristics jointly predict treatment. The F-test results of the joint significance of σ at the bottom of Table 3 suggest that the observable characteristics are jointly unrelated to treatment in both samples ($p = 0.128$ and $p = 0.168$). Importantly, these characteristics explain a trivial portion (partial R^2 values are 0.0002) of treatment variation in both samples.

Despite this evidence, there remains the possibility of differential secular trends across locations explaining my results. I address this concern in a number of ways. First, the analyses above suggest that for a large and rich set of individual characteristics, there do not appear to be important differences between the treatment and control groups, which seems unlikely if there are important secular trends at work. Second, while I do not observe any pre-AIT outcome data for the TSP (since eligibility coincides with treatment) or labor market outcomes, I analyze pre-AIT credit outcome data in online Appendix Table 12. The results show substantial baseline balance, and the joint tests of significance again confirm that the observable characteristics (demographic and baseline credit) are jointly unrelated to treatment, which further suggests no trending over the years surrounding implementation. Finally and most importantly, I complete robustness checks for all regression specifications that include unique linear time trends by location to account for the possibility of outcome trends that vary by location. The results (online Appendix Tables 7–11) are very similar to the main specifications, and if anything, suggest slightly larger PFMC effects. Taken together, these points suggest that it is unlikely that differential secular trends can account for my findings.

IV. Empirical Evidence Suggests Important Effects from the PFMC

Table 4 presents estimates for the course effects on the probability of matching to a credit bureau record (column 1), the probability of having any active credit accounts (column 2), and an individual's credit score (column 3), in Year 1 (panel A) and Year 2 (panel B). The column 1 results provide additional evidence to support my experimental validity by demonstrating that treatment group assignment is unrelated to having a matching credit bureau file. Using the 95 percent confidence interval I can rule out effects of decreases of 2 percentage points (pp) and increases of 1.5pp. The column 2 results show that the course has no meaningful impact on having active credit in either year ($p = 0.949$ and $p = 0.987$), which is unsurprising as the course taught responsible credit use and not credit market avoidance. Once again, the estimates are relatively precise and I can rule out effects larger than 2pp in either direction. The column 3

Table 4—PFMC Effects on Credit Record Matching, Activity, and Score

Outcome	Probability (matched record) (1)	Probability (active credit) (2)	Credit score (3)
Panel A. Year 1			
Control mean	0.851	0.904	581
PFMC effect	−0.003	−0.0007	−0.16
SE	(0.009)	(0.0109)	(2.98)
Cluster p-value	0.723	0.949	0.956
Wild bootstrap p-value	0.780	0.932	0.932
Observations	39,486	33,178	29,843
R^2	0.366	0.103	0.371
Panel B. Year 2			
Control mean	0.881	0.941	587
PFMC effect	0.054	0.0001	−3.97
SE	(0.032)	(0.0114)	(3.47)
Cluster p-value	0.111	0.987	0.274
Wild bootstrap p-value	0.182	0.990	0.398
Observations	39,486	24,235	21,960
R^2	0.371	0.591	0.044

Notes: The table reports LPM and OLS estimates of equation (1). The regressions in column 1 include the treatment effect indicator (PFMC), location fixed effects, and year-month fixed effects. The regressions in columns 2 and 3 also include a quadratic in age, number of dependents, indicators for female, married, minority, summer entry and education levels (high school graduate is omitted), AFQT score, enlistment term, average monthly pay in the first year, AIT length, the number of months deployed in the year, and the credit outcome in the base year. The active credit sample is restricted to those with matched records. The credit score sample is restricted to those with matched and active records. Credit outcomes are measured relative to the month an individual finished AIT. Standard errors are clustered at the AIT location level (clusters = 13). I present p-values for the clustered standard errors and 1,000 iterations of the cluster wild bootstrap procedure.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Source: DOD and Credit Bureau data

results reveal no effects on individuals' credit scores in either year ($p = 0.956$ and $p = 0.274$). The confidence intervals rule out score changes greater than 6 points in either direction, a relatively precise estimate of a negligible effect since most lenders price based on score bins that are usually in the range of 30 points or more. Since the results below provide evidence of account balance (Table 5) and adverse status reductions (Table 6), credit scores might be unaffected on average if these factors affect the bureau's score calculation in opposite ways. The panel B results for the credit outcomes in Year 2 are also statistically insignificant. The estimates of negligible effects are less precise, but enable me to rule out probability changes greater than 12pp and 3pp, and score changes of more than 11 points (columns 1–3, respectively).

Table 5 presents estimates for the course effects on the probability of having several types of account balances. The column 1 results suggest that the course reduces the probability of having any positive balance by 6.3pp, which represents a 9 percent reduction based on the control mean of 67.6 percent, and the result is statistically significant ($p = 0.024$). The column 2–4 results provide suggestive evidence that the course reduces the probability of carrying balances for all three types of credit

TABLE 5—PFMC EFFECTS ON THE PROBABILITY OF CREDIT BALANCES

	Any aggregate balance (1)	Any credit card balance (2)	Any automobile balance (3)	Any finance loan balance (4)
Panel A. Year 1				
Control mean	0.676	0.550	0.269	0.143
PFMC effect	−0.063**	−0.045	−0.020	−0.046**
SE	(0.025)	(0.027)	(0.016)	(0.021)
Cluster p-value	0.024	0.113	0.221	0.048
Wild bootstrap p-value	0.018	0.138	0.298	0.106
Observations	33,178	33,178	33,178	33,178
R^2	0.067	0.061	0.178	0.092
Panel B. Year 2				
Control mean	0.712	0.555	0.354	0.218
PFMC effect	−0.016	0.004	−0.017	0.013
SE	(0.020)	(0.032)	(0.046)	(0.014)
Cluster p-value	0.429	0.897	0.712	0.357
Wild bootstrap p-value	0.566	0.896	0.842	0.440
Observations	23,269	23,269	23,269	23,269
R^2	0.043	0.044	0.086	0.056

Notes: See Table 4 for data and specification details. Standard errors are clustered at the AIT location level (clusters = 13).
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

(credit cards, automobile leases/loans, and finance loans). However, the probability effects are only significant for finance loans (4.6pp, 32 percent, $p = 0.048$). The finance loan results (column 4) are sensitive to the small number of clusters as the results are only marginally statistically significant ($p = 0.106$) using the cluster wild bootstrap. The 95 percent confidence intervals for the credit card (column 2) and automobile (column 3) probabilities rule out increases of 0.72 pp and 1.06 pp respectively. The panel B results suggest that course effects do not persist into Year 2, as none of the estimates are statistically or economically significant.

Table 6 analyzes the account balances and shows even more substantial effects in Year 1. The column 1 results suggest that the course, on average, reduced aggregate account balances by \$608 (12 percent, $p = 0.046$). The course reduced credit card balances by \$121 (12 percent, $p = 0.058$) and finance loan balances by \$123 (30 percent, $p = 0.069$). The automobile balance results suggest 9 percent reductions but are not statistically significant. The confidence interval enables me to rule out balance changes larger than −\$809 and \$171. Note that none of the results persist into Year 2 as the panel B results are all statistically insignificant. In this case, the estimates are less precise and I cannot rule out large increases or decreases (e.g., −\$2,117 or \$2,009 for aggregate balances) using the 95 percent confidence intervals for these outcomes.[14]

[14] In online Appendix Table 9, I analyze the same credit account balances using a specification that includes unique linear trends by location. The automobile and finance balances are very similar and remain statistically insignificant. The aggregate balance point estimate is a positive \$145 but not statistically significant. The credit

TABLE 6—PFMC EFFECTS ON CREDIT BALANCES

	Aggregate balance (1)	Credit card balance (2)	Automobile balance (3)	Finance loan balance (4)
Panel A. Year 1				
Control mean	5,006	974	3,532	405
PFMC effect	−608**	−121*	−319	−123*
SE	(273)	(58.2)	(250)	(62.1)
Cluster p-value	0.046	0.058	0.225	0.069
Wild bootstrap p-value	0.064	0.030	0.290	0.090
Observations	33,178	33,178	33,178	33,178
R^2	0.223	0.216	0.190	0.211
Panel B. Year 2				
Control mean	6,465	1,087	4,689	590
PFMC effect	−202	185	−396	18
SE	(617)	(119)	(555)	(67.7)
Cluster p-value	0.748	0.144	0.489	0.785
Wild bootstrap p-value	0.832	0.258	0.772	0.834
Observations	23,269	23,269	23,269	23,269
R^2	0.107	0.111	0.088	0.082

Notes: See Table 4 for data and specification details. Balances are winsorized at the first and ninety-ninth percentiles. Standard errors are clustered at the AIT location level (clusters = 13).
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table 7 shows that the course generated meaningful reductions in adverse credit account outcomes. In Year 1 the course reduced the probability of any adverse legal actions (e.g., liens and judgments) by 1.2pp (20 percent, $p = 0.009$) and the actual number of actions by 0.09 (41 percent, $p = 0.004$). The course reduced the probability of having any accounts 60 days past due by 3.1pp (12 percent, $p = 0.004$) and the actual number of accounts in this status by 0.06 (16 percent, $p = 0.005$). The course effects for accounts 120 days past due (columns 5 and 6) are suggestive but not statistically significant. My confidence intervals enable me to rule out probability of these accounts increasing by more than 1pp as well as actual account number increases greater than 0.01. The panel B results for Year 2 are again only suggestive. All of the point estimates remain negative, but most are statistically insignificant. The course appears to reduce the probability of adverse legal actions by 1.3pp (20 percent, $p = 0.025$), though this result appears sensitive to the standard error computation as the cluster wild bootstrap result is marginally insignificant ($p = 0.114$). This Year 2 result may also be unsurprising since adverse actions may remain on an individual's credit report for longer periods of time. The similarity of the point estimate across years is encouraging but does not suggest increasing effects over time. For the other outcomes I can rule out any meaningful increases

card balance point estimate is larger (−\$355), of the opposite sign, and statistically significant. This suggests that the course effects may not persist, that individuals might regress in their credit card behaviors, and that recurrent financial education may be required.

Table 7—PFMC Effects on Adverse Credit Outcomes

	Any adverse legal action (1)	Number adverse legal actions (2)	Any trades 60 days past due (3)	Number trades 60 days past due (4)	Any trades 120 days past due (5)	Number trades 120 days past due (6)
Panel A. Year 1						
Control mean	0.051	0.194	0.247	0.378	0.144	0.204
PFMC effect	−0.012***	−0.087***	−0.031***	−0.061***	−0.013	−0.026
SE	(0.004)	(0.025)	(0.009)	(0.018)	(0.009)	(0.018)
Cluster p-value	0.009	0.004	0.004	0.005	0.185	0.182
Wild bootstrap p-value	0.044	0.004	0.022	0.028	0.230	0.294
Observations	33,178	33,178	33,178	33,178	33,178	33,178
R^2	0.233	0.443	0.219	0.295	0.250	0.291
Panel B. Year 2						
Control mean	0.051	0.269	0.356	0.530	0.188	0.261
PFMC effect	−0.013**	−0.119	−0.005	−0.036	−0.011	−0.025
SE	(0.005)	(0.081)	(0.010)	(0.025)	(0.009)	(0.018)
Cluster p-value	0.025	0.170	0.654	0.170	0.240	0.196
Wild bootstrap p-value	0.114	0.272	0.730	0.244	0.316	0.302
Observations	23,269	23,269	23,269	23,269	23,269	23,269
R^2	0.136	0.245	0.103	0.145	0.152	0.192

Notes: See Table 4 for data and specification details. All number outcomes are winsorized at the first and ninety-ninth percentiles. Standard errors are clustered at the AIT location level (clusters = 13).

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

in delinquencies (i.e., increases of 0.04 adverse actions, 0.02pp probability of an account 60 days past due, 0.01 accounts 60 days past due, 0.01pp probability of an account 120 days past due and 0.01 accounts 120 days past due) in the second year.

Table 8 presents PFMC effect estimates on TSP contributions in Years 1–4. The panel A results suggest that the PFMC has large effects on the probability of contributing to the TSP in all four years. The course increased participation by 15pp in Year 1 (125 percent, $p = 0.015$), 13pp in Year 2 (89 percent, $p = 0.014$), 12pp in Year 3 (71 percent, $p = 0.026$), and 8pp in Year 4 (47 percent, $p = 0.063$). The panel B results suggest that the course increased average monthly contributions by \$19.93 in Year 1 (115 percent, $p = 0.029$) and by \$14.02 in Year 2 (49 percent, $p = 0.038$). The effects in Years 3–4 (\$9.75 and \$7.17) remain positive, but they are statistically insignificant.[15] Using my 95 percent confidence intervals I can rule out contribution decreases greater than \$2 in Year 3 and decreases greater than \$6 in Year 4. The large but diminishing effects result from a "catch up" effect for the control group and not contribution decreases by the treatment group.[16] Online Appendix Table 13 provides robustness checks for the functional form of the TSP

[15] In unpublished results, I analyze the PFMC effects on the TSP saving distributions for Years 1–4. The positive effects are statistically significant through at least the ninetieth percentile of the contribution distributions in Years 1–2.

[16] In unpublished results, I visually analyzed the longitudinal contributions of both groups and found that control group members slowly but steadily increased their probability of contribution and average contributions. As

Table 8—PFMC Effects on the Thrift Savings Plan Outcomes in Years 1–4

	Probability (participation) in Year 1	Probability (participation) in Year 2	Probability (participation) in Year 3	Probability (participation) in Year 4
	(1)	(2)	(3)	(4)
Panel A. Probability of TSP participation				
Control mean	0.120	0.151	0.162	0.173
PFMC effect	0.150**	0.134**	0.116**	0.082*
SE	(0.053)	(0.047)	(0.046)	(0.040)
Cluster p-value	0.015	0.014	0.026	0.063
Wild bootstrap p-value	0.012	0.018	0.042	0.098
Observations	82,211	70,785	59,765	44,946
R^2	0.104	0.082	0.074	0.068
	Average contribution in Year 1	Average contribution in Year 2	Average contribution in Year 3	Average contribution in Year 4
Panel B. Average monthly TSP contributions				
Control mean	17.27	28.51	28.90	30.26
PFMC effect	19.93**	14.02**	9.75	7.17
SE	(8.06)	(5.98)	(6.19)	(6.58)
Cluster p-value	0.029	0.037	0.141	0.297
Wild bootstrap p-value	0.010	0.020	0.214	0.600
Observations	82,211	70,785	59,765	44,946
R^2	0.097	0.061	0.061	0.066

Notes: The table reports LPM (panel A) and OLS (panel B) estimates of equation (1). All regressions include the treatment effect indicator (PFMC), quadratic in age, number of dependents, indicators for female, married, minority, summer entry and education levels (high school graduate is omitted), AFQT score, enlistment term, average monthly pay in the first year, AIT length, the number of months deployed in the year, location fixed effects, and year-month fixed effects. The average contribution outcomes are winsorized at the first and ninety-ninth percentiles. The regressions in each column are limited to those still serving in each year. The TSP outcomes are measured relative to the month an individual started AIT. Standard errors are clustered at the AIT location level (clusters = 13).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

estimates. The logit (panel A) marginal effect estimates are nearly identical to the main estimates. The panel B (Tobit) estimates are larger than the main results, as expected when there is observation pooling at the outcome value of zero.

I find no course effects on the military labor market outcomes (i.e., adverse separations, rapid promotions, being offered re-enlistment, or re-enlisting conditional on eligibility). See online Appendix Table 2 for the results. The estimates are economically small (0.81pp [0.4 percent], 0.10pp [2 percent], −1.25pp [2 percent], and 1.48pp [2 percent]) and statistically insignificant ($p = 0.404, p = 0.881, p = 0.382$, and $p = 0.306$ respectively). Using the 95 percent confidence intervals, I can rule out separation increases greater than 2.7pp, rapid promotion decreases greater than 1.1pp, decreases in re-enlistment offerings greater than 4.0pp and reenlistment (conditional on eligibility) decreases greater than 4.2pp.

suggestive evidence, see the control group mean increases in both panels of Table 8. The stickiness of the treatment group's initial contributions may be one factor in these patterns.

TABLE 9.—HETEROGENEOUS PFMC EFFECTS ON CREDIT OUTCOMES IN YEAR 1

	Main specification (1)	Previous credit activity (2)	≥ Median baseline credit score (3)	Any baseline 60-day delinquency (4)	Any baseline 120-day delinquency (5)	≥ Median AFQT score (6)
Panel A. Credit card balances						
Control mean	5,006	5,006	5,006	5,006	5,006	5,006
PFMC effect	−609**	−368	−608**	−535*	−561*	−504
SE	(273)	(293)	(273)	(279)	(279)	(285)
PFMC × Previous activity		−437***				
SE		(114)				
PFMC × High baseline credit score			−3.31			
SE			(168)			
PFMC × Baseline 60-day delinquency				−507***		
SE				(130)		
PFMC × Baseline 120-day delinquency					−438***	
SE					(99)	
PFMC × High AFQT score						−227*
SE						(120)
Observations	33,178	33,178	33,178	33,178	33,178	33,178
R^2	0.223	0.223	0.223	0.223	0.223	0.223
Panel B. Probability of any adverse legal action						
Control mean	0.051	0.051	0.051	0.051	0.051	0.051
PFMC effect	−0.011***	−0.012***	−0.011**	−0.015***	−0.015***	−0.009*
SE	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)
PFMC × Previous activity		0.001				
SE		(0.001)				
PFMC × High baseline credit score			0.0002			
SE			(0.0054)			
PFMC × Baseline 60-day delinquency				0.027***		
SE				(0.008)		
PFMC × Baseline 120-day delinquency					0.035***	
SE					(0.008)	
PFMC × High AFQT score						−0.005
SE						(0.004)
Observations	33,178	33,178	33,178	33,178	33,178	33,178
R^2	0.233	0.233	0.233	0.234	0.234	0.233

Notes: See Table 4 for data and specification details. I modify the main specification to include the interaction terms identified in each column. Both outcomes are winsorized at the first and ninety-ninth percentiles. Standard errors are clustered at the AIT location level (clusters = 13).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

A. *PFMC Effects Differ for Some Groups*

Table 9 presents heterogeneous treatment estimates for the course on credit card balances and the probability of adverse legal actions in Year 1 by several individual characteristics of interest: those with previous credit market activity, those with above average baseline credit scores, those with any baseline credit delinquencies (any accounts 60 or 120 days past due), and those with above average cognitive ability (using median AFQT scores). For credit card balances (panel A), the course appears to have larger effects for those with baseline credit activity (column 2), as they reduce their balances by an additional $437 relative to those treated without this experience. This may suggest that market experience motivates additional learning and application of course concepts. It may also support linking financial education to financial product use and/or providing more "just-in-time" financial education as recommended by Fernandes, Lynch, and Netemeyer (2014). The column 3 results suggest no differential course effects for those with higher baseline credit scores. The column 4 and 5 results suggest that individuals with baseline account delinquencies appear to reduce their balances by more than treated individuals without these statuses by $507 for those with 60-day-past-due accounts and by $438 for those with 120 day-past-due accounts. These results suggest that education may appeal most to those with an immediate need for help and motivation to learn. The column 6 results suggest that those with higher levels of human capital benefit more from the course, reducing their credit balances by an additional $227 relative to those treated who have lower AFQT scores.

For the probability of adverse legal actions (panel B), the results suggest no differential treatment effects for those with previous activity (column 2), those with high baseline credit scores (column 3), or those with higher AFQT scores (column 6). However, the results suggest that treated individuals with baseline delinquencies (columns 4 and 5) are more likely to have legal adverse actions than treated individuals without these baseline problems. The relationship here may be mechanical as the delinquencies turn into adverse actions in due course. This suggests that financial education is no replacement for financial counseling and dispute resolution for those with existing problems.

V. Discussion and Lessons Learned from the PFMC

A. *Benchmarking and Interpreting the PFMC Results*

The PFMC effects on credit market outcomes are important but limited in duration. The course had no significant effects on the most routine outcomes (credit scores and having active credit), but it caused substantial reductions (7–28 percent) in the probability of having credit balances and similar effects (9–30 percent reductions) on balances themselves. The course's effects in reducing the probability of adverse credit outcomes (12–20 percent) and number of adverse outcomes (10–41 percent) are also large and strongly suggest welfare improvements among those treated. These results stand in contrast to the most recent meta-analyses (Miller et al. 2015) suggesting limited effects for financial education on adverse (e.g., defaults) financial outcomes. However, virtually none of the effects persist into the second year, and so continued education might be required to sustain financial gains.

In related work, financial education has been linked to 10–20 percent changes in desired financial behaviors including self-reported accounting behaviors (Drexler, Fischer, and Schoar 2014) and rainfall insurance purchases (Gaurav, Cole, and Tobacman 2011). Relative to these imperfect benchmarks, the PFMC appears to be about as successful as other programs in improving financial decision making and more successful than previous programs that generated no measurable effects (e.g., Choi, Laibson, and Madrian 2011). Short education programs might also be compared to information interventions, which have generated increased demand for better schools by 23 percent (Hastings and Weinstein 2008) and shown the potential to mitigate consumption losses by up to 1 percent (Stango and Zinman 2011).

The observed retirement savings effects in this study are much larger, increasing TSP participation from 12 percent to 24 percent, and persist through at least two years. Overall, the bundled intervention, which combined education and enrollment assistance, substantially increased retirement savings for treated members. The two-year differences amount to a retirement account balance difference of over $4,200 under conservative assumptions. While there is little experimental evidence on financial education's effects on retirement savings, Lusardi and Mitchell (2007) estimate that workplace education correlates with 18 percent increases in wealth; Duflo and Saez (2003) find that exposure to an employee benefit fair increases tax deferred account saving by 3–4 percent; and Cole and Shastry (2013) find that exposure to additional high school math courses increases investment income by 3–11 percent for women. My results further support Miller et al. (2014), who suggest that financial education can have meaningful effects on savings outcomes.

My larger effect estimates are unsurprising, as the PFMC combined education and enrollment assistance. Given this bundling, I also compare my effect magnitudes to some choice architecture interventions. Madrian and Shea (2001) find that automatic enrollment increases 401(k) participation by 103 percent and Carroll et al. (2009) find that an active decision enrollment regime increases 401(k) participation by 68 percent. So combining education and assistance appears to achieve results as large as other leading policies. Since some organizations that provide financial education lack the authority to change policy defaults (e.g., nonprofits) and other organizations that have the authority have chosen not to implement opt-out defaults (e.g., the Department of Defense for service members), this bundling strategy represents an attractive policy choice for increasing retirement savings and is a worthy addition to the growing choice architecture menu.

The combined effects of the PFMC across financial domains provide reason for additional optimism. The absence of any intra-budget transfer evidence, in which individuals could have financed retirement savings with credit spending, is noteworthy. The Year 2 results (more saving and comparable debt levels) are encouraging and the Year 1 results (more saving and less credit use) are doubly indicative that the course had its intended effects.

Finally, the PFMC has no statistically significant effects on the military labor market outcomes (separations, promotions, re-enlistment eligibility and choices), consistent with the economics literature that, to my knowledge, includes no findings on the causal effects of financial education on labor market decisions. Additional research might inform employers on the potential returns to financial education.

B. *Cost Comparisons*

I estimate that the course costs approximately $22 per soldier. While behavioral interventions may be cheaper methods for increasing retirement savings, the PFMC's broader curriculum, which includes other beneficial content, makes the course a reasonably inexpensive alternative. However, the PFMC's effects on credit decisions appear to be short-lived (significant in Year 1 but not in Year 2). If students forget about credit concepts, just as they might forget about geometry after high school, then financial education might need to be repeated frequently and at a higher total cost.

C. *Validity of PFMC Effect Estimates*

Several institutional factors suggest that my results are likely lower bound estimates among this population. These include course absences, training delays among the control group, interactions between control and treatment members after AIT (e.g., as roommates or friends), routine counseling or assistance by military leaders for soldiers facing financial problems, and control group members' voluntary attendance at other army financial courses. Diminishing returns among the treatment group attending more training and/or any "John Henry" effects among control group members who seek to "catch up" will mitigate positive findings. But other courses do not explain my effects since they were not initiated concurrently with the PFMC. If the other courses are complements to the PFMC, then my estimates could be upward biased, but this seems unlikely given the substantial overlap in course content.

Still, some external validity concerns suggest that these estimates may be difficult to replicate in other settings. These include the mandatory nature of the course, the environmental influences of role models and students who are used to receiving instructions and taking orders, and course timing (enlistees are young, new to the labor force, and often living alone for the first time). In addition, while individuals could not plausibly select into the military for the PFMC, they may be selecting into the military for career goals that include securing a better financial future, making these individuals "better compliers" than the average individual. As a result, my findings are most usefully applied to other groups of new workers (e.g., other service members, those in apprenticeship or union programs, and new public sector employees).

VI. Conclusion

This study estimates the effects of financial education and enrollment assistance using a natural experiment in the US Army. I find that Personal Financial Management Course attendance reduces credit account balances, delinquencies, and adverse legal actions in the first year after the course but not the second. The course and its bundled enrollment assistance substantially increase retirement savings levels for at least two years. The course and these financial gains do not effect military labor market outcomes.

While the program did not employ experimental variation in its methods or content, I briefly offer some potential explanations for its success. First, the course has a targeted curriculum that covers the most relevant topics for the students. Second, the course is well-timed in reaching individuals who are increasingly responsible for their financial welfare. Third, the course's tailored advice (e.g., use credit to purchase assets not consumables) seems better suited to this group than broad principles (e.g., how to complete a net-present-value analysis) in maintaining their interest and attention. Finally, in some cases the course combines teaching with assistance, generating actionable education.

While this study demonstrates that financial education and enrollment assistance can affect short-term financial outcomes, several issues warrant further research. First, there may be potential improvements in the curriculum design and teaching methods. Second, we still know very little about the effect mechanisms (e.g., knowledge, rules of thumb, time preferences, or peers). While isolating these will be difficult, program administrators can learn a great deal if they commit to experimental approaches. Recently proposed reforms to military compensation and expanded financial education in the Department of Defense may present such opportunities.

REFERENCES

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–84.

Army Emergency Relief and Department of Defense. 2003. "Memorandum of Understanding." Unpublished.

Bell, Casey, Dan Gorin, and Jeanne Hogarth. 2009. "Financial Education: Does It Work and How Do We Know? Research Findings from a Study of Financial Education Among Soldiers." *Community Investments* 21 (2): 15–16.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90 (3): 414–27.

Carrell, Scott, and Jonathan Zinman. 2014. "In Harm's Way? Payday Loan Access and Military Personnel Performance." *Review of Financial Studies* 27 (9): 2805–40.

Carroll, Gabriel D., James J. Choi, David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2009. "Optimal Defaults and Active Decisions." *Quarterly Journal of Economics* 124 (4): 1639–74.

Choi, James J., David Laibson, and Brigitte C. Madrian. 2011. "$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans." *Review of Economics and Statistics* 93 (3): 748–63.

Cole, Shawn, Anna Paulson, and Gauri Kartini Shastry. 2013. "High School Curriculum and Financial Outcomes: The Impact of Mandated Personal Finance and Mathematics Courses." Harvard Business School Working Paper 13-064.

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.

Duflo, Esther, and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics* 118 (3): 815–42.

Fernandes, Daniel, John G. Lynch, Jr., and Richard G. Netemeyer. 2014. "Financial Literacy, Financial Education, and Downstream Financial Behaviors." *Management Science* 60 (8): 1861–83.

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* 48 (SPL): S150–162.

Government Accountability Office. 2012. *Financial Literacy: Enhancing the Effectiveness of the Federal Government's Role*. Washington, DC: Government Printing Office.

Hastings, Justine S., Brigitte C. Madrian, and William L. Skimmyhorn. 2013. "Financial Literacy, Financial Education and Economic Outcomes." *Annual Review of Economics* 5: 347–73.

Hastings, Justine S., and Jeffrey M. Weinstein. 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *Quarterly Journal of Economics* 123 (4): 1373–1414.

Lusardi, Annamaria. 2004. "Savings and the Effectiveness of Financial Education," In *Pension Design and Structure: New Lessons from Behavioral Finance*, edited by Olivia S. Mitchell and Stephen P. Utkus, 157–84. New York: Oxford University Press.

Lusardi, Annamaria, and Olivia S. Mitchell. 2007. "Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education." *Business Economics* 42: 35–44.

Madrian, Brigitte C., and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116 (4): 1149–87.

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 6745.

Mullainathan, Sendhil, and Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means So Much*. New York: Times Books.

Schreiner, Mark, Margaret Clancy, and Michael Sherraden. 2002. "Saving Performance in the American Dream Demonstration: A National Demonstration of Individual Development Accounts." http://www.usc.edu/dept/chepa/IDApays/publications/ADDReport2002.pdf.

Skimmyhorn, William. 2016. "Assessing Financial Education: Evidence from Boot Camp: Dataset." *American Economic Journal: Economic Policy*. http:dx.doi.org/10.1257/pol.20140283.

Stango, Victor, and Jonathan Zinman. 2011. "Fuzzy Math, Disclosure Regulation, and Market Outcomes: Evidence from Truth-in-Lending Reform." *Review of Financial Studies* 24 (2): 506–34.

Willis, Lauren E. 2011. "The Financial Education Fallacy." *American Economic Review* 101 (3): 429–34.