The World Bank Economic Review, 31(3), 2017, 611–630
doi: 10.1093/wber/lhx018
Article

Does Financial Education Impact Financial Literacy and Financial Behavior, and If So, When?

Tim Kaiser and Lukas Menkhoff

Abstract

In a meta-analysis of 126 impact evaluation studies, we find that financial education significantly impacts financial behavior and, to an even larger extent, financial literacy. These results also hold for the subsample of randomized experiments (RCTs). However, intervention impacts are highly heterogeneous: financial education is less effective for low-income clients as well as in low- and lower-middle–income economies. Specific behaviors, such as the handling of debt, are more difficult to influence and mandatory financial education tentatively appears to be less effective. Thus, intervention success depends crucially on increasing education intensity and offering financial education at a "teachable moment."

JEL classification: D14, I21

The financial behavior of consumers and small-scale entrepreneurs is receiving increased interest. Evidence suggests a remarkable incidence of suboptimal individual financial decisions despite the fact that these decisions are highly relevant for individual welfare. The most prominent case of such an important financial decision in advanced economies is the amount and kind of retirement savings (cf. Duflo and Saez 2003). Studies show that undersaving is prevalent in many advanced economies and that households tend to save in inefficient ways, indicating that many may be unable to cope with the increasingly complex financial markets (e.g., Lusardi and Mitchell 2007; Choi et al. 2011; Behrman et al. 2012; van Rooij et al. 2012). This kind of behavior also stretches across other areas, including portfolio composition (Campbell 2006; Choi et al. 2010; Bucher-Koenen and Ziegelmeyer 2014; von Gaudecker 2015), excessive and overly expensive borrowing (Stango and Zinman 2009; Gathergood 2012; Agarwal and Mazumder 2013; Gerardi et al. 2013; Zinman 2015), as well as participation in

Tim Kaiser is a research associate at the University of Kiel, Germany and the German Institute for Economic Research (DIW Berlin); his email address is tkaiser@diw.de. Lukas Menkhoff (corresponding author) is the head of department of International Economics at the German Institute for Economic Research (DIW Berlin) and Professor of Economics at the Humboldt-University of Berlin; his email address is lmenkhoff@diw.de. We thank the authors who responded to our requests to provide their datasets or further details about their studies for their kind cooperation. Moreover, we appreciate valuable comments from participants at the Research in Behavioral Finance Conference 2016 in Amsterdam, the Meta-Analysis in Economics Research Network Colloquium 2016 in Conway, What Works Global Summit 2016 in London, the Conference in Behavioral Economics and Financial Literacy 2016 in Barcelona, and seminar participants in Berlin, Halle, Hamburg, Kampala, Kiel, and Vienna. In particular, we thank the editor (Eric Edmonds), three anonymous referees, Martin Brown, Nathan Fiala, Greg Fisher, Antonia Grohmann, Roy Kouwenberg, Jochen Kluve, Andreas Lutter, Christian Martin, Olivia Mitchell, Bob Reed, Anna Sokolova, Tom Stanley, Bertil Tungodden, Ludger Wössmann, and Dean Yang. Research assistance by Melanie Krüger and Iven Lützen, and financial support by DFG through CRC TRR 190 are gratefully acknowledged. An online appendix for this article can be found at *The World Bank Economic Review* website.

financial markets in general (van Rooij et al. 2011). Related problems arise in developing countries often with even more serious consequences as people are exposed to heavy shocks without having sufficient insurance or mitigation instruments (e.g., Cole et al. 2011; Drexler et al. 2014; Gibson et al. 2014; Sayinzoga et al. 2016). All this strongly motivates providing financial education to foster financial behavior.

In surprising contrast to this obvious motivation for financial education stands the lack of compelling evidence that providing financial education is an effective policy for targeting individual financial behavior (Hastings et al. 2013; Zinman 2015). Narrative literature reviews are inconclusive, either emphasizing the effectiveness of education measures (e.g., Fox et al. 2005; Lusardi and Mitchell 2014) or emphasizing the opposite (e.g., Willis 2011). Further, the two available meta-analyses of this issue do not converge in their findings: Fernandes et al. (2014) summarize overall unreliable effects of financial education, whereas Miller et al. (2015) show that education can be effective in targeting specific financial behaviors. Given this inconclusive evidence on a most important issue, what can we learn in order to explain the heterogeneity in findings and to make financial education more effective?

We go beyond the extant literature and systematically code the circumstances of financial education for our meta-analysis. This allows us to examine the determinants of a positive impact of education. Another unique characteristic of our analysis is the focus on both objectives of financial education (i.e., improvements in financial literacy and financial behavior). Hence, we investigate the role of financial literacy for financial behavior in a unified setting. Finally, our study benefits from a rapidly rising field (see fig. S1.1 in the supplemental appendix S1).

We follow the established procedures for the meta-analysis approach (e.g., Lipsey and Wilson 2001). The result is a sample of 126 studies reporting 539 effect sizes. Studies targeting entrepreneurs and exclusively measuring business outcomes (such as revenues) are omitted by design. We only consider studies reporting about interventions, such as trainings and counseling efforts. Thus, we focus strictly on exogenous variation in financial education and neglect works exclusively analyzing the possible impact of cross-sectional (baseline) differences in financial literacy on financial behavior. Finally, we carefully code interventions as we examine in detail how financial education was delivered to the target groups.

Our meta-analysis results in six principle findings: (i) increasing *financial literacy* helps. Financial education has a strong positive impact on financial literacy with an effect size of 0.26 (i.e., above the threshold value of 0.20 that characterizes "small" statistical effect sizes [see Cohen 1977]). Moreover, effects on financial literacy are positively correlated with effects on financial behavior; (ii) financial education has a positive, measurable *impact on financial behavior* with an effect size of 0.09. An effect size of 0.08 is still found under rigorous randomized experiments (RCTs); (iii) effects of financial education depend on the *target group*. First, teaching low-income participants (relative to the country mean) and target groups in low- and lower-middle–income economies has less impact, which is an obvious challenge for policymakers targeting the poor. Second, it appears to be challenging to impact financial behavior as country incomes and mean years of schooling increase, probably because high baseline levels of general education and financial literacy cause diminishing marginal returns to additional financial education; (iv) success of financial education depends on the *type of financial behavior* targeted. We provide evidence that borrowing behavior may be more difficult to impact than saving behavior by conventional financial education; (v) increasing *intensity* supports the effect of financial education; and (vi) the *characteristics* of financial education can make a difference. Making financial education mandatory is associated with deflated effect sizes. By contrast, a positive effect is associated with providing financial education at a "teachable moment" (i.e., when teaching is directly linked to decisions of immediate relevance to the target group (cf. Miller et al. 2015:13).

Complementing these findings, the meta-analysis also provides interesting non-results because several characteristics of financial education are without systematic impact on financial behavior. These include the age and gender of participants, the setting, or the choice of intervention channel through which financial education is delivered.

The findings reported above clearly motivate to implement financial education because it can positively affect financial literacy and financial behavior. However, its limited effectiveness raises two additional problems for policymakers: First, what can be done to make financial education generally more effective? Second, as a particularly obstinate aspect of the general question raised before, how can one reach those people who do not participate voluntarily? Problematic groups in this respect include low-income individuals, residents of low-income countries, and all those who do not self-select into education measures, as indicated by negative effects from mandatory courses and RCTs. For these groups, it appears that financial education needs an improved approach to be successful. More research and experience is necessary to better identify the determinants of successful financial education (e.g., Hastings et al. 2013).

Our study follows several earlier survey studies about financial education. Most of these studies have a narrative character, among them widely cited works such as Fox et al. (2005), Willis (2011), Hastings et al. (2013), and Lusardi and Mitchell (2014). This gives the authors some flexibility about selecting and interpreting the most relevant studies. A quantitative meta-analysis is more rigid in approach but has the advantages that transparent rules of procedure ensure replicable results and that quantitative relations can be derived. Overall, narrative surveys and meta-analyses complement each other.

We perform a meta-analysis because there are just two earlier systematic accounts of the financial education literature that leave much room for more research. The study by Miller et al. (2015) covers only 19 papers due to its extremely restrictive selection criteria, requiring interventions on *identical* outcomes. This limits the sample sizes to about five studies and estimates per subsample, which does not allow investigating the sources of heterogeneity.

Thus, the most similar study to our work is Fernandes et al. (2014), which covers 90 effect sizes from financial education reported in 77 papers. Despite an overlap of 44 percent with their sample of studies, our research differs in four crucial ways, which explains our new results: (i) most important is that we analyze determinants of program effectiveness in a broader way by applying respective coding; (ii) we consider various outcomes per study (on average about four per study) and their respective effect sizes; moreover, (iii) we cover recent and mostly randomized experiments providing evidence of effective interventions; and (iv) we cover additional studies focusing exclusively on financial literacy as the outcome variable.

This paper is structured in seven further sections. Section I introduces our meta-analytic approach. Section II describes our data. Section III provides first results of the meta-analysis, while section IV uses these results to explain heterogeneity of financial education treatment effects. Robustness tests are mentioned in section V, and section VI concludes with policy considerations and venues for future research.

I. Meta-analytic Method

Meta-analysis is a quantitative method to synthesize findings from multiple empirical studies on the same empirical research question. In a meta-analysis, the dependent variable is comprised of a summary statistics reported in the primary research reports, while the explanatory variables may include characteristics of the research design, the sample studied, or, in case of impact evaluations, the policy intervention itself (cf. Stanley 2001: 131). Meta-analyses can provide answers to two specific questions (cf. Muller 2015; Pritchett and Sandefur 2015; Vivalt 2015). First, is the combined (statistical) effect across all studies reporting effects of similar interventions on similar outcomes significantly different from zero? And, second, what explains heterogeneity in the reported findings?

In order to be able to aggregate summary statistics reported across heterogeneous studies, one must standardize these statistics into a common metric. If all studies would operationalize and measure outcomes in the same unit, meta-analysis could be performed directly using *economic* effect sizes (e.g., elasticities or marginal effects) in contrast to *statistical* effect sizes (cf. Stanley and Doucouliagos 2012). This, however, is rarely the case in a large sample of heterogeneous (quasi-)experimental impact evaluations.

Thus, we use a standard approach of coding a variable capturing intervention success and impact. Our impact measure (effect size) is the standardized mean difference (SMD) for each treatment effect estimate. We use the bias corrected standardized mean difference (Hedges' g) as our effect size measure, which is defined as the mean difference in outcomes between the treatment (M_T) and control (M_C) (i.e., the treatment effect) groups as a proportion of the pooled standard deviation (SD_p) of the dependent variable:

$$g = \frac{M_T - M_C}{SD_p} \tag{1}$$

with

$$SD_p = \sqrt{\frac{(n_T - 1)\ SD_T{}^2 + (n_C - 1)\ SD_C{}^2}{n_T{}^2 + n_C{}^2 - 2}}. \tag{2}$$

where n_T and SD_T are the sample size and standard deviation of the treatment group, and n_C and SD_C are for the control group. Additionally, we capture the standard error of each standardized mean difference (g), which is defined as:

$$SE_g = \sqrt{\frac{n_T + n_C}{n_T n_C} + \frac{g^2}{2(n_T + n_C)}} \tag{3}$$

Hedges' g informs about the size and direction of an effect in scale-free standard deviation units. This metric is only slightly different from other popular effect size measures in experimental impact evaluations, such as Cohen's d and Glass Δ (see, e.g., Banerjee et al. 2015). Hedges' g, however, introduces minor corrections that reduce bias in the effect size estimate in cases with small sample sizes and when the sample sizes of treatment and control groups are unequally distributed. Results are qualitatively robust to using alternative measures or relying on (partial) correlations (cf. Lipsey and Wilson 2001).

As a rule of thumb, Cohen (1977) suggests that effect sizes smaller than 0.20 should be considered as a "small effect"; effect sizes around 0.50 indicate a "medium effect"; while effect sizes greater than 0.80 constitute "large effects." Where pure mean comparisons, standard deviations, and sample sizes for each experimental outcome are not reported directly we exhaust all possibilities to calculate or estimate effect sizes (g) and its corresponding standard error from the range of available statistical data (cf. Lipsey and Wilson 2001).

In the estimation of summary effects of the literature, our main approach follows a full pooling least squares meta-regression framework (e.g., Card et al. 2015). Accordingly, the financial education treatment effect (g) can be explained by exogenous, observable characteristics, the impact g on an outcome i, reported in study j is expressed as a linear function

$$g_{ij} = \alpha + \ x_{ij}\beta + \epsilon_{ji} \tag{4}$$

where $x_{ij}\beta$ is a vector of observable (exogenous) study-level covariates, such as intensity of intervention, α is an intercept, and ϵ_{ji} denotes an error-term independent from $x_{ij}\beta$. We estimate our models using multiple effect sizes per study and account for heteroscedasticity by clustering standard errors at the study-level. Reassuringly, results are not sensitive to a set of changes in estimation strategy and accounting for publication selection bias (see section V and supplemental appendix S3).

II. Sample Description

This section describes the selection of studies, the extraction of effect sizes and study-level covariates, and types of financial education programs.

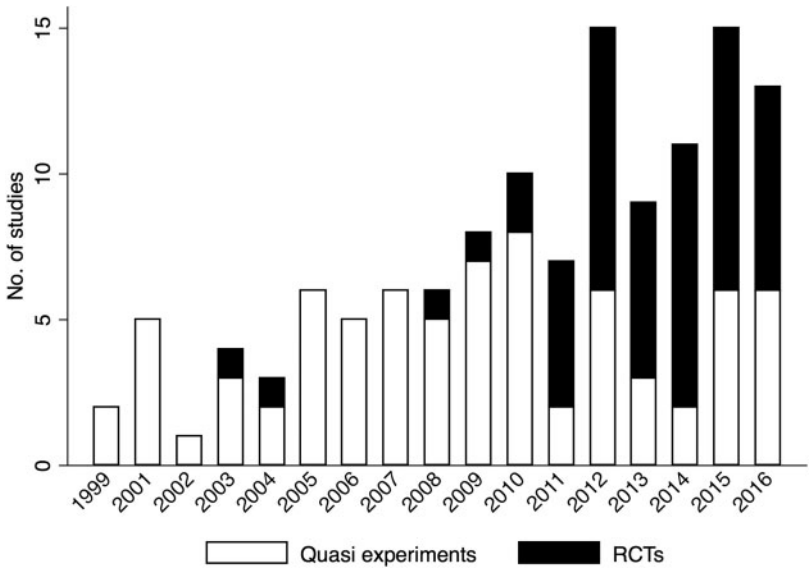
Selection of Studies

We follow the established meta-analytical protocol (cf. Lipsey and Wilson 2001, Stanley 2001). This starts with systematically searching the relevant databases, including working papers, for the following keywords: (i) financial literacy; (ii) financial knowledge; (iii) financial education; (iv) financial capability; and (v) combinations of these keywords with "intervention." Moreover, we consider all records from meta-analyses (Fernandes et al. 2014; Miller et al. 2015) and narrative literature reviews (Fox et al. 2005; Collins and O'Rourke 2010; Willis 2011; Xu and Zia 2012; Hastings et al. 2013; Blue et al. 2014; Lusardi and Mitchell 2014). This search resulted in over 500 potentially relevant published journal articles and over 600 results from working paper databases with some apparent overlap. We stopped collecting studies in October 2016 (see appendix S1).

From this collection, we drop studies that do not meet our three criteria for inclusion: (i) reporting on impacts of an exogenous educational intervention on financial literacy and/or financial behavior; (ii) providing a quantitative assessment of intervention impact that allows coding an effect size statistic (g) and its standard error; and (iii) relying on an observed counterfactual in the estimation of intervention impacts. This selection process leads to a final sample of 126 independent intervention studies that report 539 effect sizes (further details in tables S1.1 and S1.2 in the supplemental appendix S1). Of these, 90 studies report 349 effect sizes on financial behavior, and 67 studies report 190 effect sizes on financial literacy. Among these 90 plus 67 studies, there are 31 studies reporting effect sizes on both financial literacy and behavior.

RCTs are rare in the early years of the literature, but their share has risen dramatically, with the majority of studies conducted from 2011 onward being randomized evaluations (see fig. 1). This development in the literature is very favorable for meta-analyses because it ensures a high internal validity of research findings reported in the primary studies and helps to clearly distinguish between selection and treatment effects.

Figure 1. Number of Studies in Our Sample by Research Design per Year



Source: Authors' calculations based on the data source discussed in the text.

Extraction of Effect Size Estimates and Study Descriptors

As the next step, we code the effect of financial education on *financial literacy* (i.e., a measure of performance on a financial knowledge test), since knowledge development is the primary goal of financial education (Hastings et al. 2013; Lusardi and Mitchell 2014). Moreover, we code treatment effects of financial education on several *financial behaviors* (see table S1.2 in the supplemental appendix S1), such as an increase in savings after the treatment. Multiple estimates per study are considered if multiple outcomes, time-points, or treatments are reported; however, results are robust to aggregating all effects per study into one synthetic effect size. Further details about this process are described in supplemental appendix S1.

Types of Financial Education Programs

Our dataset includes four main types of financial education programs. First, and most frequent, are evaluations of *classroom financial education* (approximately 83 percent of all estimates) in various settings, such as schools, universities, the workplace, or specific sites such as savings groups or microfinance institutions. These studies are quasi-experiments or RCTs, in which the researcher has control over content, intensity, and survey design in order to measure specific outcomes. There is an increasing interest in the literature in multiple-treatment and cross-over designs to investigate optimal delivery strategies and potential causal mechanisms (i.e., Drexler et al. 2014; Carpena et al. 2015; Skimmyhorn 2016). These studies have high internal validity but may report site-specific effects that causally interact with unobserved features of the specific sites (cf. Muller 2015). Additionally, measurement of outcomes is typically in the short or medium run (approx. 65 percent), since long time series are usually not available. A different strand of the literature evaluating this type of program looks at classroom financial education utilizing (plausibly exogenous) variation in (mandatory) school financial education mandates (e.g., Tennyson and Nguyen 2001; Brown et al. 2016). These studies are typically quasi-experimental in nature, and, while possibly weaker in internal validity, possess high external validity, since they typically have large sample sizes and measure relatively long-run effects on behavioral outcomes, such as savings.

A second type of intervention is *online financial education* (approx. 8 percent of estimates). While similar in research design to experiments on classroom financial education, these studies usually estimate the effect of certain online modules on financial literacy and behavior and typically evaluate instructional videos or interactive applications.

The third type of financial education treatments evaluated in the literature are *individualized counseling interventions* (two percent of estimates). These have been mainly studied in the US and typically study outcomes related to the handling of (mortgage) debt.

As a fourth and last type, we identify *informational and behavioral nudges*, such as information fairs at the workplace and informational brochures (seven percent of estimates). These studies typically evaluate behavioral change in response to these low-intensity treatments. There is one study in our sample that studies the effect of a behavioral nudge in the form of "financial edutainment" in mass-media (cf. Berg and Zia 2017). This is an intervention designed to impact financial behaviors through a non-cognitive channel (as opposed to increasing financial knowledge), and the included study evaluates the impact of financial messages inserted into episodes of a popular television series in South Africa.

III. Results from Meta-analysis

We report the mean effects for all studies and then for subsamples: financial literacy and financial behavior, types of financial education programs, research designs, and different country groups.

Summary Effects of Financial Education

Here we discuss the average effects of financial education on financial literacy and financial behavior. Based thereon, we study the relation between these two outcomes. As a starting point, we note that the summary effect of financial education on all kinds of reported outcomes is estimated to be $g = 0.148$ ($p = .000$, $n = 539$). However, heterogeneity in effect sizes is high, indicating that outcomes could be disaggregated for meaningful analyses.

Financial behavior. We find that the average impact of educational interventions on financial behaviors is statistically highly significant ($g = 0.086$) (see table S1.3 in the supplemental appendix S1). The main reason that we get a more favorable result than Fernandes et al. (2014) is that we profit from a moderate, positive time trend (more details in supplemental appendix S2). To compare the magnitude of this effect size to results from health promotion on behavioral change (e.g., weight loss and nutrition in obesity studies), Portnoy et al. (2008) report in their meta-analysis of 75 RCTs an average effect size of about 0.1.

Financial literacy. The average impact of financial education on financial literacy is substantially higher ($g = 0.263$, $p = .000$, $n = 190$) than the one on financial behavior (see figure S1.2 and table S1.3 in the supplemental appendix S1). Moreover, financial education explains 1.7 percent of the variance in financial knowledge and, thus, appears to be only slightly less effective than educational interventions in other domains such as math and science instruction (cf. Fernandes et al. 2014: 1867). To put this effect size in perspective: the meta-analysis of 225 studies by Freeman et al. (2014) reports an average effect size of around 0.47 for studies evaluating student performance in response to alternatives to lecturing in undergraduate science education; however, these interventions occur in a university context and last for a full semester.

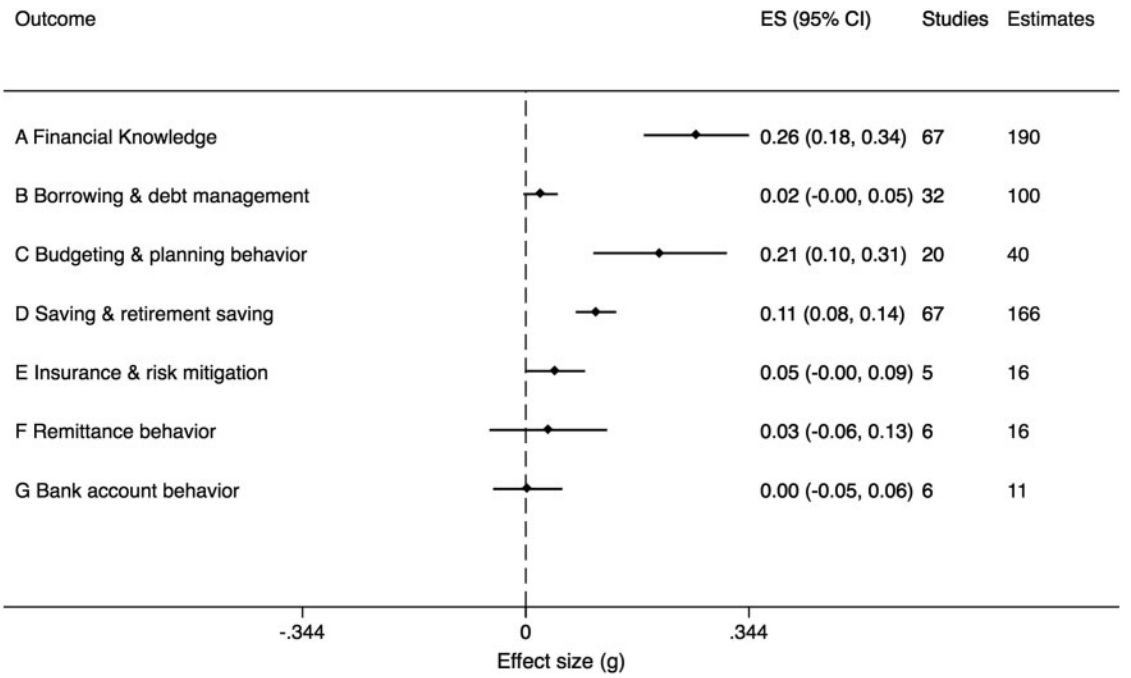
Relationship between financial literacy and behavior. The intuition is that increases in financial literacy scores are an important intermediate result in a causal chain expected to lead to behavior change (e.g., Grohmann et al. 2015; Fort et al. 2016). Indeed, for a sample of 31 studies, we find in a regression with standard errors clustered at the study-level that the effect size on financial literacy is a statistically significant predictor of effect size on financial behavior ($b = 0.230$, $p = .022$). Thus, an increase of one standard deviation unit in financial literacy scores is related to an average increase of 0.23 standard deviation units of the financial behaviors studied. However, the non-overlapping confidence intervals of these effect sizes also indicate that these two elements of the causal chain should be analyzed separately when attempting to explain the heterogeneity in effect sizes.

Effect Sizes by Type of Financial Behavior

Fig. 2 shows the average effect size for the seven categories of financial behaviors targeted by the educational interventions in our sample.

Average effect sizes for three out of seven categories of outcomes are clearly positive and highly statistically significant at the one percent level. Additionally, all confidence intervals for the different types of financial behaviors overlap each other, indicating that there are no extreme differences in impacts depending on the specific form of financial behavior targeted. In detail: (i) the average effect size on "budgeting" appears to be higher than those on downstream behaviors; and (ii) effect sizes related to saving and retirement saving appear to be higher than the average effect size of financial education on borrowing behavior; (iii) this latter average effect size is small ($g = 0.02$) and insignificant from zero; (iv) similarly, the average effect sizes for "insurance" ($g = 0.05$), "remittances" ($g = 0.03$), and "bank account behavior" ($g = 0.00$) are estimated to be small and insignificant from zero, although based on a few studies per category only. Thus, debt-related financial behaviors may be the most challenging to target through financial education (see Miller et al. 2015). Overall, these findings correspond to the

Figure 2. Forest Plot of Effect Sizes by Type of Financial Behavior Studied



Source: Authors' calculations based on the data source discussed in the text.

results provided by Fernandes et al. (2014) and Miller et al. (2015) and extend to our much larger sample.

Effect Sizes by Type of Financial Education Intervention

We form subsamples by the main types of financial education interventions, as discussed in section II. First, we compare classroom financial education to three types of non-classroom delivery channels (online financial education, counseling, and informational/behavioral nudges). Second, we distinguish between financial education at school and two non-school settings (workplace and other settings). Panel A of table 1 shows results split by outcomes on financial literacy and financial behavior. While in-person classroom trainings appear to be (unconditionally) more effective than non-classroom delivery channels in increasing financial knowledge, we observe no statistically significant difference regarding impacts on financial behavior. Turning to the intervention setting, it appears that interventions in schools are more effective at increasing financial literacy but yield marginally significant smaller treatment effects on financial behavior. However, we note that these relations are obviously partially confounded with several other relevant variables (e.g., the age of the participants, the delay in measurement, and research design), which indicates the importance of an examination in a multivariate setting (cf. section IV).

Effect Sizes by Research Design

Regarding research design, Fernandes et al. (2014) find that weaker research designs lead to inflated effect sizes. Thus panel B of table 1 compares average effect sizes as a function of research design. When we focus on financial behaviors as outcomes, RCTs show statistically highly significant (unconditional) effect sizes of 0.081. These are only slightly smaller than for quasi-experiments with 0.093, indicating that the small but positive significant effects of financial education exist even under the most rigorous empirical standards.

Table 1. Effect Sizes of Financial Education by Intervention Type, Research Design, and Country Groups

Outcome	Type	Studies	Obs.	ES (g)	SE_g	*p*-value	Diff. (t-value)
A. Effect sizes by intervention channel & setting							
Fin. literacy	Classroom	58	135	0.294	0.054	.000	0.106**
	Non-classroom	9	55	0.188	0.039	.001	(2.015)
	– Online	5	41	0.217	0.060	.018	
	– Counseling	0					
	– Nudge	4	14	0.103	0.045	.108	
Fin. behavior	Classroom	70	317	0.084	0.013	.000	−0.014
	Non-classroom	20	32	0.098	0.020	.000	(0.452)
	– Online	11	18	0.085	0.034	.031	
	– Counseling	7	8	0.095	0.030	.020	
	– Nudge	2	6	0.140	0.007	.031	
Fin. literacy	School	35	62	0.373	0.076	.000	0.163***
	Non-school	32	128	0.210	0.035	.000	(3.273)
	– Workplace	1	1	0.164	0.063		
	– Other	31	127	0.210	0.035	.000	
Fin. behavior	School	27	90	0.057	0.014	.000	−0.039*
	Non-school	63	259	0.096	0.014	.000	(1.96)
	– Workplace	17	47	0.121	0.049	.023	
	– Other	46	212	0.090	0.015	.000	
B. Effect sizes by research design							
Fin. literacy	RCTs	33	135	0.209	0.033	.000	−0.185***
	Quasi-exp.	34	55	0.394	0.083	.000	(−3.638)
Fin. behavior	RCTs	40	227	0.081	0.015	.000	−0.012
	Quasi-exp.	50	122	0.093	0.022	.000	(−0.661)
C. Effect sizes by country group							
Fin. literacy	High income	53	123	0.328	0.058	.000	0.183***
	Developing	14	67	0.145	0.031	.000	(3.787)
	– Low	3	6	0.219	0.069	.086	
	– Lower-middle	6	44	0.155	0.047	.023	
	– Upper-middle	5	17	0.092	0.023	.017	
Fin. behavior	High income	66	168	0.071	0.019	.000	−0.027
	Developing	24	181	0.098	0.014	.000	(−1.512)
	– Low	6	39	0.161	0.038	.009	
	– Lower-middle	12	90	0.091	0.008	.000	
	– Upper-middle	6	52	0.06	0.023	.045	

Notes: Average effect sizes *(g)* estimated via OLS regressions of effect sizes fitting only an intercept. Sample is split by an indicator of intervention type, research design, or country group. "Channel" is a categorical variable operationalized in the form of four dummy variables: Classroom, Counseling, Online, and "Nudge" where "Nudge" is the default (omitted) category in the regressions. "Setting" is a categorical variable operationalized through three dummy variables: School, Workplace and Other where Other is the omitted category in the meta-regression analyses. Country groups are based on the World Bank Atlas method and refer to 2015 data on GNI per capita. Low-income economies are defined as those with a GNI per capita of $1,025 or less in 2015, lower-middle–income economies are defined by a GNI per capita between $1,026 and $4,035, upper-middle income economies are those with a GNI per capita between $4,036 and $12,475, and high-income economies are defined by a GNI per capita greater than $12,475. Standard errors are clustered at the study level. ***, **, and * denote significance at the one percent, five percent, and ten percent level.

Source: Authors' analysis based on data sources discussed in the text.

RCTs also provide a significant positive effect of financial education on financial literacy with 0.209. Here the difference to other designs (effect size of 0.394) is significant at the one percent level.

Effect Sizes by Country Groups

To investigate another potential source of heterogeneity, we disaggregate our data by country groups. Panel C of table 1 shows effect sizes by country groups as classified by the World Bank based on 2015

GNI per capita. We find that effect sizes on *financial literacy* are significantly higher in developed (high income) economies ($g = 0.328$) than in developing economies (low income, lower- and upper-middle–income economies, $g = 0.145$). Turning to effect sizes on *financial behavior*, this difference is statistically insignificant in this unconditional comparison but differences between country groups become more nuanced and statistically significant when controlling for other relevant variables (see section IV).

IV. Explaining Heterogeneity in Financial Education Treatment Effects

Section III shows that the average effect size of financial education is accompanied by large heterogeneity. Thus, we examine whether there are factors explaining this heterogeneity. This will also suggest directions that future financial education policies might take in order to increase their impact on financial behavior.

Potential Correlates of Effect Size

The effectiveness of financial education is potentially influenced by the peculiarities of the specific intervention. Based on prior literature, we group these characteristics into four categories: (i) the research design; (ii) the intensity of education; (iii) the target group of education; and (iv) the characteristics of the education program.

(i) Regarding the *research design* of a financial education study, we expect the method of investigation (i.e., RCT vs. less rigorous designs) to be relevant. Second, the concrete measurement of an effect will influence the estimated size of impact. It is known that focusing on treatment on the treated (TOT) (i.e., measuring a treatment effect on the population who actually *received or attended* the treatment) generally results in higher effect sizes than focusing on the intention to treat (ITT) effect (i.e., the population who was in principle *assigned* to treatment). However, ITT may be more relevant for policy (cf. Imbens and Wooldridge 2009; Gertler et al. 2011). Third, the delay between financial education treatment and measurement of the effect may negatively influence the effect size since effects of the intervention may decay over time (cf. Fernandes et al. 2014). Additionally, we control for the precision of effect size estimation by the inverse standard error (or the (squared) standard error, see supplemental appendix S3). All these variables are defined in table 2, which also provides descriptive statistics.

(ii) A core variable of financial education interventions is the *intensity of education* (i.e., the number of hours taught). It is expected that higher intensity will support the effect. However, the time-frame over which the financial education intervention is delivered to the target group may also be of importance. We expect differences between high intensity and low intensity relative to the duration. Thus, we code the *hours of financial education per week* (i.e., intensity per week) and the *duration* of the intervention in weeks to investigate this issue.

(iii) The expectation regarding a possible relation between the *target group* of education and effectiveness of financial education is as follows. Generally, learning is easier for younger people, younger people may be more open to new concepts and their baseline financial literacy scores are low (e.g., Lusardi and Mitchell 2014), meaning that the *age* of the target group may be negatively related to the effect size of financial education. Second, a gender gap in financial literacy is treated as a stylized fact in the literature (cf. Lusardi and Mitchell 2014) which may also translate to gender differences in treatment effects. Thus we include the *percentage of women* in the sample. Third, it is expected that the acquaintance of the target group with an educational environment may be helpful. As a proxy for such openness to education, we take the *income* of the target group relative to the overall population. Fourth, we expect that the overall institutional level of education should support domain-specific educational efforts (Jappelli 2010). As a proxy for this potential relationship, we take a country's population mean *years of schooling* as reported by the United Nations Development Program Human Development Reports. Additionally, we augment our data with *country-level financial literacy* data from a 2015 global financial literacy

Table 2. Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
A. Descriptive statistics at the study-level					
RCT	126	0.405	0.493	0.000	1.000
TOT	115	0.452	0.500	0.000	1.000
Delay	93	82.231	273.613	0.000	1566
1/SE	126	57.535	210.450	2.480	1636.712
Intensity	87	11.211	14.929	0.100	87.000
Duration	76	7.341	14.150	1.000	103.000
Age	109	30.717	14.120	9.000	63.870
Percent female	123	54.011	18.493	0.000	100.000
Low income clients	102	0.529	0.502	0.000	1.000
Years of schooling	126	11.270	2.843	3.200	13.600
FL in population	124	50.419	11.658	24.000	66.000
Mandatory	96	0.292	0.457	0.000	1.000
Incentivized	86	0.314	0.467	0.000	1.000
Teachable moment	126	0.397	0.491	0.000	1.000
B. Descriptive statistics at the estimate-level					
RCT	539	0.672	0.470	0.000	1.000
TOT	510	0.282	0.451	0.000	1.000
Delay	463	93.742	292.025	0.000	1566.000
1/SE	539	41.260	124.389	2.740	957.167
Intensity	451	15.384	23.444	0.100	144.000
Duration	434	7.908	14.236	1.000	103.000
Age	494	31.814	11.720	9.000	63.870
Percent female	525	52.923	18.200	0.000	100.000
Low income clients	451	0.681	0.467	0.000	1.000
Years of schooling	539	9.890	3.463	3.200	13.600
FL in population	523	44.170	14.668	24.000	66.000
Mandatory	480	0.240	0.427	0.000	1.000
Incentivized	445	0.247	0.432	0.000	1.000
Teachable moment	539	0.479	0.500	0.000	1.000

Notes: "RCT" is a dummy variable with "1" if selection into treatment was conducted through randomization and "0" otherwise (such as matched designs). "TOT" is a dummy variable with "1" if the effect size estimate is derived from the treatment effect on the treated and "0" if it is derived from the ITT estimate. "Delay" is a continuous variable indicating the delay between treatment and measurement of outcomes in weeks. "1/SE" is the inverse standard error for each effect size estimate. "Intensity" is the total number of hours of financial education exposure to the treated. "Duration" indicated the time-frame of financial education in weeks. "Age" is the mean age of the sample in years. "Percent Female" is the relative frequency of female participants in the sample in percent. "Low income" is a dummy variable with "1" if the mean annual income per capita of the sample is below the country average income per capita. "Mandatory" is a dummy variable with "1" indicating mandatory participation in financial education and "0" voluntary participation. "Incentivized" is a dummy variable with "1" when incentives to participate where provided and "0" if participation was unconditional on incentives. "Teachable moment" is a dummy variable indicating whether the financial education intervention was offered at a teachable moment.

Source: Authors' analysis based on data sources discussed in the text.

survey (Klapper et al. 2015). We hypothesize that financial education interventions may yield higher effects when the population baseline financial literacy is lower, indicating more room for improvement through education. Finally, as a control variable we code the country of intervention according to the World Bank country group classifications.

(iv) Regarding the *characteristics of the education program*, it seems interesting whether the *channel* (i.e., classroom, online, individual counseling, etc.) is important in explaining education effectiveness, since these formats come with different trainer to participant ratios and may rely on different pedagogical approaches to financial education. It may be that willingness to learn and change financial behavior is lower when financial education is *mandatory* (cf. Collins 2013) or motivation to participate in

financial education is not intrinsic but driven by *incentives* provided by the offering institution. Lastly, these characteristics may be correlated with specific *settings* (i.e., at school or at the workplace).

Next, and going further in this direction, it is coded whether participants are educated at a *teachable moment* (i.e., that they have the possibility to apply their knowledge in a concrete case of interest to them, e.g., Doi et al. 2014; Miller et al. 2015). Thus, we capture whether the education addressed immediate financial issues (such as borrowers already in default, or migrants confronted with deciding through which channel remittances are sent). Alternatively, financial education was generic and offered at an unspecific moment, as is often the case in large scale financial education programs (e.g., Bruhn et al. 2014).

Meta-Regression Models Explaining Intervention Impacts

This section examines determinants of financial education effectiveness using a multivariate meta-regression framework including the above discussed potential correlates as right-hand side variables. Our procedure is motivated by economic and econometric considerations. From an economic point of view, we aim for including all variables that have a substantial theoretical foundation. From an econometric viewpoint, the specification should be parsimonious, especially in the presence of a relatively small sample size of studies.

Thus, we start with a specification where we include all reasonable and available variables (table 3, column 1). In order to keep the number of studies considered high, we impute average or default values for missing observations (we show in supplemental appendix S3 that our main results are insensitive to imputation). The discussion considers groups of variables in four blocks, following their introduction in section IV.

Research design. Starting with the research design of the underlying primary studies, we find that RCTs report—*ceteris paribus*—slightly smaller effect sizes than non-RCTs, which is in line with earlier presumptions (see table 1, panel B). However, now this difference is statistically significant (see column 1 of table 3). As expected, the operationalization of treatment effects as TOT-estimates leads to inflated effect size estimates. Apparently, the delay between intervention and measurement of outcomes does not seem to be systematically related to effect sizes in this estimation (cf. supplemental appendix S3 for an alternative approach and investigation of heterogeneous treatment effects depending on delay in measurement). In addition, estimates with large inverse standard errors are associated with smaller effect sizes, indicating that larger and more precise studies report smaller effect sizes overall. However, this coefficient is small in size and insignificant.

Intensity. Turning to the relationship between intensity per week and duration, column 1 of table 3 shows that intensity has a significant positive effect on treatment effects on financial behavior. Thus, an increase of one hour of financial education per week leads to a 0.004 standard deviation unit increase in the impact on financial behaviors studied. Considering that the average weekly duration is in this subsample is roughly nine weeks and weekly intensity is only about four hours, doubling the weekly intensity to eight hours while keeping everything else constant at the mean, would lead to an average treatment effect around 14 percent higher than the empirical mean predicted treatment effect in this fully specified model.

Target group. Among participant characteristics, age and gender are not significant explanatory variables. However, the coefficient on "low income clients" is highly significant and negative, indicating that these individuals are more difficult to educate. Regarding increasing mean years of schooling at the country level, returns to additional financial education appear to diminish. This is in line with results from two studies in very different contexts (Europe and India) that report higher treatment effects for lower-educated individuals and diminishing returns to financial education upon higher baseline levels of education (cf. Cole et al. 2011; Fort et al. 2016). Similarly, the coefficient for baseline financial literacy

Table 3. Explaining Heterogeneity in Effect Sizes on Financial Behavior

	(1) All	(2) All	(3) RCTs	(4) Low inc. econ	(5) High / middle inc. econ	(6) Low income clients
RCT	−0.070**	−0.068**		−0.209**	−0.079**	−0.066**
	(0.027)	(0.028)		(0.091)	(0.036)	(0.032)
TOT	0.079***	0.068**	0.012	−0.016	0.076**	0.031
	(0.027)	(0.027)	(0.040)	(0.066)	(0.035)	(0.032)
Delay	0.000	0.000	−0.001**	−0.001**	0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
1/SE	−0.000	−0.000	0.000	−0.003	−0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)	(0.000)
Intensity / week	0.004**	0.004***	0.007***	0.004**	0.003	0.004***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.003)	(0.001)
Duration	−0.000	−0.000	−0.001	−0.001	−0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
Age	−0.001					
	(0.001)					
Percent female	−0.000					
	(0.001)					
Low income clients	−0.065***	−0.055***	−0.074***	−0.042**	−0.048**	
	(0.020)	(0.017)	(0.024)	(0.019)	(0.021)	
Years of schooling	−0.016***	−0.019***	−0.016**	−0.026***	−0.025***	−0.011*
	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.006)
FL in population	−0.003					
	(0.002)					
Country group						
a) Low/lower-mid. inc. econ.	−0.129*	−0.093**	−0.092**			−0.059
	(0.073)	(0.036)	(0.042)			(0.042)
b) Upper-mid. inc. econ.	0.000					
	(0.060)					
Channel						
a) Classroom	−0.003					
	(0.028)					
b) Counseling	−0.018					
	(0.033)					
c) Online	−0.028					
	(0.028)					
Setting						
a) School	0.022					
	(0.023)					
b) Workplace	0.041					
	(0.036)					
Mandatory	−0.074***	−0.051**	−0.078*	−0.015	−0.065**	−0.052
	(0.024)	(0.023)	(0.044)	(0.042)	(0.025)	(0.033)
Incentivized	−0.012					
	(0.029)					
Teachable moment	0.079***	0.064**	0.016	0.025	0.069**	0.072**
	(0.021)	(0.026)	(0.035)	(0.026)	(0.029)	(0.032)
Constant	0.477***	0.332***	0.338***	0.514***	0.406***	0.188*
	(0.157)	(0.079)	(0.095)	(0.110)	(0.114)	(0.095)
R^2	0.210	0.183	0.149	0.170	0.204	0.109
n (Studies)	90	90	40	18	72	44
n (Effect sizes)	349	349	227	129	220	234

Notes: Non-standardized coefficients from OLS regressions. Dependent variable in columns (1) and (2) is effect size (Hedges' g) on financial behavior in the full sample of studies reporting on financial behavior as an outcome. Column (3) shows results for RCTs only. Column (4) and (5) show results for financial behavior split by country groups. Column (6) limits the sample to classroom trainings only. Robust standard errors clustered at the study-level in parentheses. ***, **, and * denote significance at the one percent, five percent, and ten percent level.

Source: Authors' analysis based on data sources discussed in the text.

in the population is also negative, albeit statistically insignificant. While these results suggest declining marginal returns to financial education, the negative effect for low- and lower-middle–income economies—and also the above-mentioned coefficient on low-income clients—shows a countervailing influence from challenging groups or country circumstances.

Characteristics of education. Regarding the channel variables, column 1 shows that no alternative channel appears to be generally more or less effective than financial education in classroom settings or informational nudges (omitted category). The same is true for the setting of the intervention where school and workplace settings are not systematically different from other settings. However, mandatory financial education and implementing financial education at a "teachable moment" appear to be important. Specifically, we find, that making financial education mandatory decreases effect sizes by 0.074 standard deviation units: The predicted value for effect size on financial behavior in mandatory formats with everything else kept equal at the (empirical) mean would be only $g = 0.030$ ($SE = 0.020$, $p = .134$); thus, economically small and statistically insignificant from zero. In contrast, offering financial education at a teachable moment increases effect sizes by 0.079 standard deviation units. Thus, the predicted value for effect size on financial behavior would be *ceteris paribus* $g = 0.124$ ($SE = 0.014$, $p = .000$) (i.e., statistically highly significant), roughly 48 percent larger than the unconditional average effect size found in the sample and about 45 percent larger conditional on the empirical means for all other covariates in this full model.

Parsimonious specification. We reduce the above discussed fully specified model by keeping the variables on research design and intensity but otherwise eliminating the insignificant variables. Column 2 of table 3 describes the resulting reduced model that confirms the fully specified regression results from column 1. There are just some smaller changes in the estimated standard errors that occur at a few variables. This indicates that it is justified to rely on the parsimonious specification, in particular when we analyze subsamples with a much smaller number of observations in the following.

Meta-Regression Models for Subsamples

Given the large degree of heterogeneity across the 90 studies and their underlying financial education programs, we move to an analysis of more homogenous subsamples.

RCTs only. Many will agree that RCTs fulfill the most rigorous requirements, implying that results limited to this subsample of studies are indeed reliable. We do not prefer this procedure because many observations are lost. Nevertheless, it is reassuring that results qualitatively hold, as shown in column 3 of table 3 for the subsample of 40 RCTs covering 227 effect sizes. However, while the negative coefficient for mandatory courses remains to be large in magnitude and statistically (marginally) significant, the coefficient for teachable moment loses explanatory power in this estimation.

Interventions in low- and lower-middle–income economies. This subsample covers 18 studies that report 129 effect sizes (see column 4 of table 3). Again, all coefficients have the same sign and similar magnitude as in our parsimonious specification (column 2 in table 3), but differences in standard errors arise. While intensity of the intervention remains a strong predictor and low-income clients in low-income economies also benefit significantly less from financial education, mandatory formats and timing in the sense of offering financial education at a teachable moment appear less predictive of treatment effects.

Interventions in upper-middle and high-income economies. Turning to the 72 studies that examine financial education in more affluent economies (column 5 of table 3), we find that results again are qualitatively very similar to the pooled analysis in column 2. Here, the opposing coefficients for mandatory formats and offering financial education at a teachable moment are statistically significant at the five percent level, indicating that these effects may be primarily driven by interventions in middle- or high-income economies.

Interventions for low-income individuals. Examining the subsample of 44 studies focusing on low-income individuals results in a similar picture arising. Effects appear to be higher with increased training intensity and offering financial education at a teachable moment. However, country-level years of schooling and country income are now only marginally significant and insignificant covariates, respectively. Additionally, the coefficient for mandatory courses still has the same sign and similar magnitude, but is estimated with a larger standard error.

Disaggregating financial behaviors and financial behaviors by target group. As discussed in section III, it appears to be easier to affect financial behaviors in terms of (retirement) savings and budgeting compared to borrowing behavior. Thus, we disaggregate the sample into three categories of financial behaviors and search for potentially heterogeneous effects of our main explanatory variables. We reduce the choice of variables for some subsamples to avoid problems with degrees of freedom due to relative few observations.

Column 1 of table 4 shows results for the subsample of 32 studies reporting effect sizes on borrowing behavior. This result matches our main results of the aggregated sample of effect sizes (column 2 of table 3) with significant positive effects from increased intensity, negative effects for low-income target groups, and countries, negative effects from making financial education mandatory and positive effects from offering financial education at a teachable moment. Column 2 of table 4 shows results for the subsample of 20 studies that focus on borrowing as the outcome and have low-income clients as the target group. Again, results are nearly identical. However, the delay in measurement is now a marginally significant predictor: effect sizes in this sample seem to diminish as time between intervention and measurement of outcomes increases. Hence, treatment effects on debt related behaviors among low-income individuals may be shorter-lived.

Turning to effect sizes reported in 67 studies on (retirement) saving (column 3 of table 4), we observe that the relevant variables from our benchmark model (column 2 of table 3) remain significant predictors. However, voluntary versus mandatory formats seem to be unrelated to effectiveness. Column 4 of table 4 shows the results on savings and retirement savings for low-income individuals reported in 31 studies. Signs and magnitude are similar to the benchmark estimation, but the only coefficients estimated with a small standard error are intensity per week and the teachable moment. Thus, qualitative results hold, but effect sizes on saving behavior for low-income individuals may be difficult to impact through the considered covariates.

Turning to the subsample of 20 studies on budgeting and record keeping behavior (column 5 of table 4), on which financial education yields the largest effects, we find that intensity is not significantly related to effect size. Additionally, all of the other signs and relative magnitudes of the coefficients remain similar to our benchmark estimation; however, with increased standard errors due to only 20 studies and 40 observations. Completing this exercise, we now examine determinants of treatment effects for the subsample of studies reporting on budgeting outcomes for low-income clients (column 6 of table 4). There are 11 studies in this subsample reporting 27 estimates. Again, qualitative results are similar and intensity now, again, is a marginally significant predictor of effect sizes on budgeting behavior.

Overall, we find that the positive effects from increased intensity appear to be driven by interventions focused on (retirement) saving and borrowing behavior, whereas the timing and voluntary participation matter, especially for borrowing behavior. Thus, the financial behavior that is hardest to impact (borrowing) needs special effort in the sense of increased intensity and timing the financial education intervention at a teachable moment.

V. Robustness

The robustness tests cover eight different aspects and are reported in full in supplemental appendix S3. All of them confirm our qualitative findings. Here, we just mention these tests: (i) testing the average

Table 4. Explaining Heterogeneity in Effect Sizes for Subsamples by Type of Financial Behavior and Target Group

	(1) Borrow	(2) Borrow× low inc. clients	(3) Save	(4) Save × low inc. clients	(5) Budget	(6) Budget× low inc. clients
RCT	−0.136***	−0.100***	−0.002	−0.035		
	(0.022)	(0.026)	(0.045)	(0.058)		
TOT	0.089**	0.106**	0.090	0.074		
	(0.033)	(0.039)	(0.054)	(0.079)		
Delay	−0.000	−0.000*	0.000	−0.000	−0.001	−0.019
	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.012)
1/SE	0.000	0.001**	−0.000	0.000	−0.003*	−0.007
	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.005)
Intensity / week	0.003**	0.003**	0.003*	0.004**	0.037	0.595*
	(0.001)	(0.001)	(0.002)	(0.002)	(0.031)	(0.308)
Duration	−0.000	−0.000	−0.001	0.000	−0.000	0.017
	(0.000)	(0.000)	(0.001)	(0.001)	(0.003)	(0.014)
Low income clients	−0.043**		−0.050**			
	(0.019)		(0.022)			
Years of schooling	−0.023***	−0.023***	−0.018***	−0.011	−0.020*	0.017
	(0.006)	(0.008)	(0.007)	(0.011)	(0.011)	(0.022)
Low/lower-mid. inc. econ.	−0.178***	−0.199***	−0.142***	−0.102		
	(0.052)	(0.067)	(0.045)	(0.066)		
Mandatory	−0.069**	−0.120***	−0.025	−0.010		
	(0.032)	(0.039)	(0.031)	(0.049)		
Teachable moment	0.100***	0.087***	0.084**	0.114*		
	(0.025)	(0.026)	(0.036)	(0.065)		
Constant	0.375***	0.326**	0.305***	0.147	0.361**	−0.685
	(0.087)	(0.114)	(0.091)	(0.165)	(0.134)	(0.524)
R^2	0.473	0.394	0.194	0.147	0.206	0.359
n (Studies)	32	20	67	31	20	11
n (Effect sizes)	100	73	166	91	40	27

Notes: Non-standardized coefficients from OLS regressions with clustered standard errors at the study-level in parentheses. We only include right hand side variables where differential information from at least two studies is available in the regressions. ***, **, and * denote significance at the one percent, five percent, and ten percent level.

Source: Authors' analysis based on data sources discussed in the text.

treatment effect with several alternative meta-regression models; (ii) repeating the parsimonious benchmark model without imputing missing values; (iii) running this model for studies about the US only; (iv) running this benchmark model with classroom studies only; (v) running this model with equal weight per study by either calculating one synthetic effect size per study or weighting effect sizes accordingly; (vi) running the benchmark specification with different empirical approaches; (vii) analyzing the influence of delay on effects; and (viii) testing a different definition of training intensity. Additionally, we further examine publication bias and possible heterogeneity in study quality in supplemental appendix S4 and use alternative econometric techniques that account for publication selection bias in supplemental appendix S3.

VI. Concluding Policy Discussion

This meta-analysis covers studies that potentially contribute to realizing policy objectives, such as improved financial literacy and changes in individual financial behavior. Due to this close link to economic policy, we discuss insights that have potential policy relevance in three steps:

General policy lessons. (i) The most important policy lesson from our research is that financial education can be effective. However, the field of financial education is not developed enough that established standards could be followed "blindly," rather the process of designing interventions needs careful attention due to large heterogeneity across program types and individual studies.

(ii) Interventions targeting improvements in financial literacy are quite successful as they achieve effectiveness similar to comparable education interventions in other domains. As financial literacy education basically aims at improving financial knowledge and awareness, it seems evidentiary that it works well in the classroom and at school (see e.g. Bruhn et al. 2016). Improved financial literacy also has an indirect positive effect on financial behavior, although this indirect effect is small so that changes in financial behavior should also be addressed directly.

(iii) Education interventions targeting financial behavior have desired effects on average. Although these effects are economically rather small, they are statistically robust. Impacts on financial behavior are higher if the intensity of education is increased and if financial education is offered at a teachable moment. The effects are smaller if "problematic" groups are addressed, such as low-income clients.

Policy lessons for subgroups. As the universe of studies covers widely diverse financial education interventions, we draw three lessons for more homogeneous groups: (i) regarding the country groups, education effects seem to be somewhat lower in *low- and lower-middle–income countries*. This is probably due to the disadvantageous institutional circumstances in these countries. A relative advantage in these countries, however, is that the general level of education (mean years of schooling in the population) is comparatively low so that marginal returns to additional domain-specific education are high. The lower opportunity costs of education may be a reason why mandatory participation conditions, such as school based programs, are less problematic and offering financial education at teachable moment appears to be of lesser importance in these countries.

(ii) While problematic target groups, such as *low-income clients*, are more difficult to educate in general, the determinants of effective financial education are not different from the general population. If there is a difference, it appears that a teachable moment is relatively important, indicating that there is a particular need to get the attention of this target group.

(iii) Regarding the *outcomes* of financial education, improving debt related behavior is, on average, hardly successful. At the same time, mistakes can be rather consequential and the structure of many significant determinants is the same as for other financial behaviors, such that the general lessons may translate to this specific case; however, it needs much more input to reach economically significant results. Moreover, there is variation across studies revealing clear success cases, which suggests that it is useful to go down to the study level and learn from best practices. The effects on improving savings or budgeting behavior are much larger in magnitude than on borrowing.

Research on open policy issues. In order to improve financial education policies in the future we see three areas of urgent research: (i) we need quite generally *more reliable evidence on the effectiveness* of financial education interventions. Almost two-thirds of the evidence comes from the US, indicating that there are large gaps of evaluation elsewhere.

(ii) Regarding the documentation of impact evaluations within published reports, it would be very desirable to provide *more information about study and program characteristics* (see Miller et al. 2015). A straight-forward example is the quality of teacher training or implementation, which can make a crucial difference but is unknown in almost all studies (Brown et al. 2016). The same applies to the ways in which the curriculum is structured and implemented (see Drexler et al. 2014 as a notable exception).

(iii) Finally, in order to come closer to *welfare assessments*, information in two directions is needed: first, information about program costs is frequently missing. Thus, in terms of welfare, positive education effects could be balanced with the true costs of the intervention (see also Lusardi et al. 2016). Second, the discussion of effectiveness of financial education policy should also consider principal

alternatives to financial education in general. Such alternatives include limiting the kind of available products (choices), altering the choice architecture (e.g., Carroll et al. 2009), working with nudges (e.g., Thaler and Benartzi 2004; Willis 2011), considering the promotion of commitment devices (e.g., Brune et al. 2016), offering incentives (e.g., Saez 2009), or implementing more rigid consumer financial protection policies (cf. Campbell et al. 2011).

There are two arguments in favor of implementing financial education. First, the small average effect comes with low average intensity. More than 70 percent of our considered studies invest no more than two days in education, indicating that these measures may have only small effects, but also low costs. Second, the average small effect of financial education is accompanied by large heterogeneity, indicating that those offering financial education measures can still learn from best practice experiences, a development that is ongoing as evidenced by time trend of slowly increasing effectiveness documented in rigorous impact evaluation studies.

References

Agarwal, S., and B. Mazumder. 2013. "Cognitive Abilities and Household Financial Decision Making." *American Economic Journal: Applied Economics* 5 (1): 193–207.

Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Parienté, J. Shapiro, B. Thuysbaert, and C. Udry. 2015. "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries." *Science* 348 (6236): 1260799.

Behrman, J.R., O.S. Mitchell, C.K. Soo, and D. Bravo. 2012. "How Financial Literacy Affects Household Wealth Accumulation." *American Economic Review: Papers and Proceedings* 102 (3): 300–4.

Berg, G., and B. Zia. 2017. "Harnessing Emotional Connections to Improve Financial Decisions. Evaluating the Impact of Financial Education in Mainstream Media." *Journal of the European Economic Association.* https://doi.org/10.1093/jeea/jvw021.

Blue, L., P. Grootenboer, and M. Brimble. 2014. "Financial Literacy Education in the Curriculum: Making the Grade or Missing the Mark?" *International Review of Economics Education* 16 (Part A): 51–62.

Brown, M., J. Grigsby, W. van der Klaauw, J. Wen, and B. Zafar. 2016. "Financial Education and the Debt Behavior of the Young." *The Review of Financial Studies* 29 (9): 2490–522.

Brune, L., X. Giné, J. Goldberg, and D. Yang. 2016. "Facilitating Savings for Agriculture: Field Experimental Evidence from Malawi." *Economic Development and Cultural Change* 64: 187–220.

Bruhn, M., L. de Souza Leao, A. Legovini, R. Marchetti, and B. Zia. 2016. "The Impact of High School Financial Education: Evidence from a Large-Scale Evaluation in Brazil." *American Economic Journal: Applied Economics* 8 (4): 256–95.

Bruhn, M., G.L. Ibarra, and D. McKenzie. 2014. "The Minimal Impact of a Large-Scale Financial Education Program in Mexico City." *Journal of Development Economics* 108: 184–9.

Bucher-Koenen, T., and M. Ziegelmeyer. 2014. "Once Burned, Twice Shy? Financial Literacy and Wealth Losses During the Financial Crisis." *Review of Finance* 18 (6): 2215–46.

Campbell, J.Y. 2006. "Household Finance." *Journal of Finance* 61 (4): 1553–604.

Campbell, J.Y., H.E. Jackson, B.C. Madrian, and P. Tufano. 2011. "Consumer Financial Protection." *Journal of Economic Perspectives* 25 (1): 91–114.

Card, D., J. Kluve, and A. Weber. 2015. "What Works? A Meta-Analysis of Recent Active Labor Market Program Evaluations." NBER Working Paper No. 21431, National Bureau of Economic Research, Cambridge, MA.

Carpena, F., S. Cole, J. Shapiro, and B. Zia. 2015. "The ABCs of Financial Education. Experimental Evidence on Attitudes, Behavior, and Cognitive Biases." World Bank Policy Research Working Paper No. 7413. World Bank, Washington, DC.

Carroll, G.D., J.J. Choi, D. Laibson, B.C. Madrian, and A. Metrick. 2009. "Optimal Defaults and Active Decisions." *Quarterly Journal of Economics* 124 (4): 1639–74.

Choi, J.J., D. Laibson, and B.C. Madrian. 2010. "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds." *The Review of Financial Studies* 23 (4): 1405–32.

———. 2011. "$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans." *Review of Economics and Statistics* 93 (3): 748–63.
Cohen, J. 1977. *Statistical Power Analysis for the Behavioral Sciences*. (Rev. Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
Cole, S., T. Sampson, and B. Zia. 2011. "Prices or Knowledge? What Drives Demand for Financial Services in Emerging Markets?" *Journal of Finance* 66 (6): 1933–67.
Collins, J.M. 2013. "The Impacts of Mandatory Financial Education: Evidence from a Randomized Field Study." *Journal of Economic Behavior and Organization* 95: 146–58.
Collins, J.M., and C.M. O'Rourke. 2010. "Financial Education and Counseling—Still Holding Promise." *Journal of Consumer Affairs* 44 (3): 483–98.
Doi, Y., D. McKenzie, and B. Zia. 2014. "Who You Train Matters: Identifying Combined Effects of Financial Education on Migrant Households." *Journal of Development Economics* 109: 39–55.
Drexler, A., G. Fischer, and A. Schoar. 2014. "Keeping It Simple: Financial Literacy and Rules of Thumb." *American Economic Journal: Applied Economics* 6 (2): 1–31.
Duflo, E., and E. 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, D., J.G. Lynch Jr, and R.G. Netemeyer. 2014. "Financial Literacy, Financial Education, and Downstream Financial Behaviors." *Management Science* 60 (8): 1861–83.
Freeman, S., S.L. Eddy, M. McDonough, M.K. Smith, N. Okoroafor, H. Jordt, and M.P. Wenderoth. 2014. "Active Learning Increases Student Performance in Science, Engineering, and Mathematics." *Proceedings of the National Academy of Sciences* 111 (23): 8410–5.
Fort, M., F. Manaresi, and S. Trucchi. 2016. "Adult Financial Literacy and Households' Financial Assets: The Role of Bank Information Policies." *Economic Policy* 31 (88): 743–82.
Fox, J., S. Bartholomae, and J. Lee. 2005. "Building the Case for Financial Education." *Journal of Consumer Affairs* 39 (1): 195–214.
Gathergood, J. 2012. "Self-Control, Financial Literacy and Consumer Over-Indebtedness." *Journal of Economic Psychology* 33 (3): 590–602.
Gerardi, K., L. Goette, and S. Meier. 2013. "Numerical Ability Predicts Mortgage Default." *Proceedings of the National Academy of Sciences* 110 (28): 11267–71.
Gertler, P.J., S. Martinez, P. Premand, L.B. Rawlings, and C.M. Vermeersch. 2011. *Impact Evaluation in Practice*. Washington, DC: World Bank Publications.
Gibson, J., D. McKenzie, and B. Zia. 2014. "The Impact of Financial Literacy Training for Migrants." *World Bank Economic Review* 28 (1): 130–61.
Grohmann, A., R. Kouwenberg, and L. Menkhoff. 2015. "Childhood Roots of Financial Literacy." *Journal of Economic Psychology* 51: 114–33.
Hastings, J.S., B.C. Madrian, and W.L. Skimmyhorn. 2013. "Financial Literacy, Financial Education, and Economic Outcomes." *Annual Review of Economics* 5: 347–73.
Imbens, G.W., and J.M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47 (1): 5–86.
Jappelli, T. 2010. "Economic Literacy: An International Comparison." *Economic Journal* 120 (548): F429–51.
Klapper, L., A. Lusardi, and P. van Oudheusden. 2015. "Financial Literacy Around the World: Insights from the Standard and Poor's Rating Services Global Financial Literacy Survey." Accessed August 7, 2017. http://gflec.org/initiatives/sp-global-finlit-survey/
Lipsey, M.W., and D.B. Wilson. 2001. *Practical Meta-Analysis*. Thousand Oaks, CA: Sage.
Lusardi, A., and O.S. Mitchell. 2007. "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth." *Journal of Monetary Economics* 54 (1): 205–24.
———. 2014. "The Economic Importance of Financial Literacy: Theory and Evidence." *Journal of Economic Literature* 52 (1): 5–44.
Lusardi, A., P.-C. Michaud, and O.S. Mitchell. 2016. "Optimal Financial Knowledge and Wealth Inequality." *Journal of Political Economy* 125 (2): 431–77.
Miller, M., J. Reichelstein, C. Salas, and B. Zia. 2015. "Can You Help Someone Become Financially Capable? A Meta-Analysis of the Literature." *World Bank Research Observer* 30 (2): 220–46.

Muller, S.M. 2015. "Causal Interaction and External Validity: Obstacles to the Policy Relevance of Randomized Evaluations." *World Bank Economic Review* 29 (Suppl 1): S217–25.

Portnoy, D.B., L.A. Scott-Sheldon, B.T. Johnson, and M.P. Carey. 2008. "Computer-Delivered Interventions for Health Promotion and Behavioral Risk Reduction: A Meta-Analysis of 75 Randomized Controlled Trials, 1988–2007." *Preventive Medicine* 47 (1): 3–16.

Pritchett, L., and J. Sandefur. 2015. "Learning from Experiments When Context Matters." *American Economic Review: Papers and Proceedings* 105 (5): 471–5.

Saez, E. 2009. "Details Matter: The Impact of Presentation and Information on the Take-Up of Financial Incentives for Retirement Saving." *American Economic Journal: Economic Policy* 1 (1): 204–28.

Sayinzoga, A., E.H. Bulte, and R. Lensink. 2016. "Financial Literacy and Financial Behaviour: Experimental Evidence from Rural Rwanda." *Economic Journal* 126 (594): 1571–99.

Skimmyhorn, W. 2016. "Assessing Financial Education: Evidence from Boot Camp." *American Economic Journal: Economic Policy* 8 (2): 322–43.

Stango, V., and J. Zinman. 2009. "Exponential Growth Bias and Household Finance." *Journal of Finance* 64 (6): 2807–49.

Stanley, T.D. 2001. "Wheat from Chaff: Meta-Analysis As Quantitative Literature Review." *Journal of Economic Perspectives* 15 (3): 131–50.

Stanley, T.D., and H. Doucouliagos. 2012. *Meta-Regression Analysis in Economics and Business*. New York: Routledge.

Tennyson, S., and C. Nguyen. 2001. "State Curriculum Mandates and Student Knowledge of Personal Finance." *Journal of Consumer Affairs* 35 (2): 241–62.

Thaler, R.H., and S. Benartzi. 2004. "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy* 112: 164–87.

van Rooij, M., A. Lusardi, and R. Alessie. 2011. "Financial Literacy and Stock Market Participation." *Journal of Financial Economics* 101 (2): 449–72.

———. 2012. "Financial Literacy, Retirement Planning and Household Wealth." *Economic Journal* 122 (560): 449–78.

Vivalt, E. 2015. "Heterogeneous Treatment Effects in Impact Evaluation." *American Economic Review: Papers and Proceedings* 105 (5): 467–70.

von Gaudecker, H.-M. 2015. "How Does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice?" *Journal of Finance* 70 (2): 489–507.

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

Xu, L., and B. Zia. 2012. "Financial Literacy Around the World: An Overview of the Evidence with Practical Suggestions for the Way Forward." World Bank Policy Research Working Paper No. 6107. World Bank, Washington, DC.

Zinman, J. 2015. "Household Debt: Facts, Puzzles, Theories, and Policies." *Annual Review of Economics* 7: 251–76.