Does financial education impact financial behavior, and if so, when?

Tim Kaiser and Lukas Menkhoff

Abstract

In a meta-analysis of 115 experimental impact evaluations, 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 in low- and medium income countries; some target groups, such as low-income clients, or specific behaviors, such as borrowing, are difficult to influence; additionally mandatory financial education appears to be less effective. Thus, success depends crucially on increasing training intensity and offering financial education at a "teachable moment."

JEL-Classification: D 14 (personal finance), I 21 (analysis of education)
Keywords: Financial education, financial literacy, financial behavior, meta-analysis, meta-regression, impact evaluation

July 02, 2016

We would like to thank the authors who responded to our request providing their datasets or further details of their studies for their kind co-operation. Moreover, we appreciate valuable comments from seminar participants in Berlin, Kiel, and Vienna, in particular from Martin Brown, Nathan Fiala, Antonia Grohmann, Roy Kouwenberg, Jochen Kluve, Andreas Lutter, Christian Martin, Olivia Mitchell, Tom Stanley, Bertil Tungodden, and Dean Yang. Research assistance by Melanie Krüger and Iven Lützen is gratefully acknowledged.

Tim Kaiser, German Institute for Economic Research (DIW Berlin), D-10108 Berlin, Germany, and University of Kiel; tkaiser@diw.de.

Lukas Menkhoff, German Institute for Economic Research (DIW Berlin), D-10108 Berlin, Germany, and Humboldt-University Berlin; lmenkhoff@diw.de.

Does financial education impact financial behavior, and if so, when?

1 Introduction

The financial behavior of consumers and small-scale entrepreneurs has received increasing interest over the last years. 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 under-saving 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, such as 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), and participation in 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 using sufficient insurance or mitigation instruments (e.g., Cole et al. 2011; Drexler et al. 2014; Gibson et al. 2014; Sayingoza et al. 2016). All this strongly motivates providing financial education in order 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 really 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 systematic reviews of this issue, both applying a meta-analysis approach, 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?

Our main contribution is analyzing the heterogeneity around a small average positive effect of financial education. Therefore, 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 behavior and in financial literacy. Thus, we investigate the role of financial literacy for financial behavior in a unified setting. Finally, our study benefits from a rapidly rising field, as indicated by the increased number of citations of publications using the keyword "financial literacy" (see Figure 1). Beyond the number of studies, the quality is also improving due to rigorous impact evaluation methods, which allow for a more precise estimation of treatment effects.

<Figure 1 about here>

We follow established procedures of the meta-analysis approach (e.g. Lipsey and Wilson 2001; Stanley and Doucouliagos 2012). This means that we describe how we searched for relevant studies and how we chose selected studies in order to avoid biases. The result is a sample of 115 studies reporting 429 effects of financial education on the financial literacy and financial behaviors of individuals. Studies targeting entrepreneurs and measuring business outcomes exclusively are omitted by design. We only consider studies reporting about

financial education interventions, such as trainings and counseling efforts. Thus, we focus strictly on exogenous variation in financial education and neglect works analyzing exclusively the possible impact of cross-sectional (baseline) differences in financial literacy on financial behavior. Finally, we carefully code financial education interventions as we examine in detail how financial education was delivered to the target groups.

The main finding of our meta-analysis is that financial education does indeed impact financial behavior in the intended way. However, the way financial education is provided is crucial because its unconditional effectiveness is small. The average impact of financial education on all reported outcomes, the average effect size, is 0.16; the impact on financial behavior is, at 0.09, even smaller. These effects are statistically highly significant and robust, but they are clearly below the threshold value of 0.20 that characterizes "small" statistical effect sizes (see Cohen 1977). Thus, it seems important to learn from earlier work what might increase program effectiveness in the future.

Our meta-analysis results in six principle findings: (i) Increasing *financial literacy* helps. Financial education has a stronger positive impact on financial literacy (effect size of 0.27) than on behavior, while effect sizes on financial literacy are positively correlated with effect sizes on financial behavior; (ii) financial education has a positive, measurable, *impact on financial behavior*. Effectiveness is still found under rigorous evaluation methods, such as randomized experiments (RCTs); (iii) effects of financial education depend on the *target group*. First, interventions in low or lower-middle income economies appear to be generally less effective than in upper-middle and high income economies, indicating that the basic level of quality of educational institutions in the country may matter for financial education to be meaningful. Second, it also appears to be challenging to impact financial behavior as country incomes increase, probably because high baseline levels of financial literacy cause diminishing returns to additional financial education. Third, teaching low-income participants

(relative to the country mean) has less impact, especially in low and lower-middle income economies, which is an obvious challenge for policymakers targeting the poor; (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 smaller effects. By contrast, a robust 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.

Complementing these findings, the meta-analysis also provides interesting non-results because several characteristics of financial education are without systematic impact on financial behavior. This includes 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 the need to implement financial education because it can positively affect 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 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 in order 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 literacy and closely related issues. 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 studies. A quantitative meta-analysis is more rigid in approach but has the advantage that rules of procedure should be transparent, meaning that results should be fully replicable. There are just two earlier meta-analyses about financial education: The study by Miller et al. (2015) covers only 19 papers due to its extremely restrictive selection criteria. Thus, most similar to our work, is the study by Fernandes et al. (2014), which covers 90 effect sizes from financial education reported in 77 papers. Despite an overlap of 48% with their sample of studies, there are three reasons why we find new results: (i) most important is that we analyze determinants of program effectiveness in a broader way by applying respective coding. Moreover, (ii) we cover recent and mostly randomized experiments providing evidence of effective interventions and (iii) we cover additional studies focusing exclusively on financial literacy as the outcome variable.

This paper is structured in seven further sections. Section 2 introduces our meta-analytic approach. Section 3 describes our data. Section 4 provides first results of the meta-analysis, while Section 5 uses these results to explain heterogeneity of financial education treatment effects. Section 6 provides additional analyses for particular policy relevant determinants of intervention impact. Robustness tests are reported in Section 7 and Section 8 concludes with policy considerations and venues for future research.

2 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 one or multiple summary statistics reported in the primary research reports, while the explanatory variables may include characteristics of the research design, the

particular sample studied, or in case of impact evaluations, the policy intervention itself (cf. Stanley 2001, p.131). Meta-analyses are able to provide answers to two specific questions that are highly relevant in contested economic literatures (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. Ideally, all studies would operationalize and measure outcomes in the same way (i.e. in the same unit). If this was the case, 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, p.23). 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 ($\mathrm{M_C}$) (i.e. the treatment effect) groups as a proportion of the pooled standard deviation ($\mathrm{SD_p}$) of the dependent variable:

$$g = \frac{\mathrm{M}_T - \mathrm{M_C}}{\mathrm{SD_p}} \tag{1}$$

with

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

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:

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

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. 2016). Hedges' g, however, introduces minor corrections that reduce bias in the effect size estimate in cases with low sample size and when sample sizes of treatment and control groups are unequally distributed. When operationalizing effect sizes using alternative measures or converting to (partial) correlations, we do not find significant differences in results (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, including regression coefficients and t-statistics (cf. Lipsey and Wilson 2001, p.198).

In the estimation of summary effects of the literature, our main approach follows a least squares meta-regression framework as previously applied in other economic studies (e.g., Card et al. 2015). Assuming that 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

$$g_{ij} = x_{ij}\beta + \epsilon_{ji} \quad (4)$$

where $x_{ij}\beta$ is a vector of observable (exogenous) study-level covariates, such as intensity of intervention, and an intercept, and ϵ_{ji} denotes an error-term independent from $x_{ij}\beta$. We primarily discuss estimation results based on this easily interpretable meta-analytical model using ordinary least squares. We estimate our models using multiple (possibly correlated) effect sizes per study and account for heteroscedasticity by clustering standard errors at the study-level. Results are not sensitive to a set of changes in estimation strategy (see Section 7 and Appendix C).

3 Sample description

This section describes a crucial element of any meta-analysis, i.e. the way one retrieves the sample of studies selected for examination. This includes the selection of studies (Section 3.1), the extraction of effect sizes and study-level covariates (Section 3.2), and characteristics of the resulting sample (Section 3.3).

3.1 Selection of studies

We follow the established meta-analytical protocol (cf. Lipsey and Wilson 2001, p.23; Stanley 2001, p.143). This starts with systematically searching the relevant databases for the most common keywords in order to aggregate a large sample of potentially eligible studies to be included in our meta-analysis. Keywords are (i) financial literacy; (ii) financial knowledge; (iii) financial education; (iv) financial capability; and (v) combinations of these keywords with "intervention." In order to minimize publication bias and to capture the broadest sample of financial education intervention studies possible, we systematically search not only the relevant databases for published records (e.g. ISI, Business Source Premier via EBSCO Host, JStor) but also for registered trails, working papers and informal research reports (e.g. AEA RCT-registry, SSRN, RePEC, NBER, Worldbank eLibrary). All of the studies from recent

systematic accounts of the literature (Fernandes et al. 2014; Miller et al. 2015) are included in our initial pool of studies. In addition, we screen the references of 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 each of the respective working paper databases with some apparent overlap in results, as different working papers may cover the same underlying primary study. We stopped collecting articles from these databases in December 2015.

From this collection, we drop studies that do not meet the criteria of inclusion into our meta-analysis. These criteria are (i) reporting on impacts of an exogenous educational intervention designed to strengthen the participants' financial literacy and/or leading to behavioral change in the area of personal finance; (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. Consequently, econometric studies analyzing the relationship between "measured financial literacy" (Fernandes et al. 2014, p.1862) and financial behavior (i.e. cross sectional studies) are excluded. Meeting criterion (iii), we also exclude studies that rely on the estimation of intervention impacts based on observed effects in only one group (e.g., single group pre-/ and posttest studies). As a result, we only include experimental studies with sufficient information on intervention outcomes in our analysis. There are three types of experiments included in our analysis: RCTs (e.g., Drexler et al. 2014; Bruhn et al. 2016), quasi-experiments (e.g., Lührmann et al. 2015), and natural experiments (e.g., Skimmyhorn 2016). Where necessary information was partially missing, we consulted additional online resources related to the article or contacted the authors of the primary studies via e-mail.

This selection-process leads us to a final sample of 115 independent intervention studies that report 429 effect sizes. Of these, 82 studies report 260 effect sizes on financial behavior, 61 studies report 169 effect sizes on financial literacy, and 28 studies report 141 effect sizes on both financial literacy and financial behavior. The sample is comprised of 45 RCTs and 70 quasi/natural-experiments.

In comparison, Miller et al. (2015) select 19 intervention-studies for their statistical meta-analysis. Fernandes et al. (2014), with 77 studies selected, cover 90 effect sizes (15 RCTs and 75 quasi/natural-experiments) of "manipulated financial literacy" (cf. Fernandes et al. 2014, p.1863). Of their 77 papers, 55 are also part of our sample. We exclude 22 single-group pre-posttest and quasi-experimental papers because they either do not analyze education interventions (but other personal finance related programs, e.g. match incentives), report only aggregate measures of self-reported financial behavior, wellbeing or self-efficacy, or because it is not feasible to calculate a meaningful effect size statistic. In addition, we include 27 recent studies that were not previously available. Moreover, we consider another 33 studies examining the impact of financial education on financial literacy (but neglecting possible impacts on financial behavior). These modifications lead to the mentioned overlap of 48% regarding studies. We include all of these studies in our analysis, no studies are excluded because of unavailability.

3.2 Extraction of effect size estimates and study descriptors

The next step in our meta-analytic process is to extract effect size estimates from the statistical data reported in the primary studies. In addition to the coding of all possible effect sizes (g) and their standard errors of financial education treatment on financial literacy or financial behavior (cf. Section 2), we develop a coding protocol to extract potentially relevant information about the study (study descriptors) that may serve as predictor variables

explaining the variability in effect sizes. Specifically, we aim at extracting data on (i) research design and measurement of dependent variables; (ii) the intensity of education; (iii) the sample/ target group of the intervention; and (iv) the characteristics of the intervention itself, such as channel, setting and participation conditions. Coding of the included study reports was done manually by the authors of this paper and two research assistants who were trained using the guidelines provided by Lipsey and Wilson (2001, p.88). Overall intercoder reliability is high and data collection for most of the variables concerning the setting, participants, and research methodology of the primary studies was straightforward. However, key details of the underlying educational intervention are often missing or underreported in the research reports. If information is only partially missing authors were asked to provide these details via an e-mail request. The next section provides descriptive details of our final sample of 115 impact evaluation studies.

3.3 Descriptive statistics for financial education interventions

Table 1 shows the composition of our sample of financial education impact studies by the date of publication (*Panel A*) and the country in which the intervention took place (*Panel B*). Our selection of papers covers 115 independent interventions from 1999 through 2015. While the majority of interventions took place in the U.S. and other OECD countries, 20% of studies were conducted in low- or middle-income countries.

<Table 1 about here>

RCTs are rare in the early years of the literature, but the share has risen dramatically, with the majority of intervention studies conducted from 2011 onward being randomized evaluations (see Figure 2). These allow for the most precise estimation of treatment effects. This development in the literature is very favorable for meta-analyses, since it ensures a high validity of research findings reported in the primary studies, helps to clearly distinguish

between selection- and treatment effects, and leads to higher comparability across these studies.

<Figure 2 about here>

4 Results from the meta-analysis

We report the calculation of a mean effect for all studies (Section 4.1) and distributions of effect sizes across several dimensions of the dataset in this section. The effects on financial literacy and on financial behavior (Section 4.2), effects on several types of financial behavior outcomes (Section 4.3), effects as a function of research design (Section 4.4), effects as a function of publication status and quality (Section 4.5) and effects for different country groups (Section 4.6) are distinguished.

4.1 Summary effects of financial education

We discuss here the average effects of financial education on behavior and financial literacy; 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=0.156 (p=0.000, n=429). However, heterogeneity in effect sizes is considerably high, thus indicating that outcomes could be disaggregated for meaningful analyses. This further suggests distinguishing between effect sizes on financial behavior and on financial literacy.

Financial behavior. We find that the average impact of educational interventions on financial behaviors is statistically highly significant (g=0.089, p=0.000, n=260) (see Table A1 in the Appendix A). Although the coefficient of 0.089 is small in size, there exists a measurable and robust impact of financial education on various kinds of financial behavior. In comparison, Fernandes et al. (2014) estimate the summary effect of financial education on

financial behavior to be roughly g=0.066. However, the authors use averaged effect sizes per paper and weight each observation with its inverse standard error. In order to obtain a better comparison with that study, we exactly apply their method (random effects meta-regression) to our sample of studies. This provides an average (weighted) effect size of g=0.074 (p=0.000, n=82). Thus, our estimate of a summary effect for the literature is about 11% higher.

To investigate the potential source of this difference, we estimate the weighted average effect size among those recent studies that are not included in Fernandes et al. (2014). Indeed, we find that there is a larger average effect of financial education on financial behavior in this sample (g=0.12). This indicates that the new studies covered in our meta-analysis are the main source of difference. Diving deeper into this issue, we find that Fernandes et al. (2014) estimate extremely small average effect sizes for their sample of 15 RCTs. Our broader sample of randomized experiments, however, leads to a much more positive assessment. In line with this observation, the effect size of financial education on financial behavior documented in RCTs seems to increase over time (see Figure 3), indicating a positive time trend in effect sizes: a regression of effect size on year of study publication results in a statistically highly significant coefficient (b=0.015, SE=0.003). This moderate, positive time-trend is an important element in explaining our positive result about the effect of financial education on financial behavior.

<Figure 3 about here>

Financial literacy. The average impact of financial education on financial literacy is substantially higher (g=0.268, p=0.000, n=169) than the one on financial behavior. This difference is also obvious from looking at the full distribution of effect sizes. Figure 4 shows a kernel density estimate for all reported effect sizes from financial education. Obviously, the distribution for effect sizes on financial literacy is shifted to the right compared to the

distribution for financial behavior. This difference is expected because it appears that it is easier for an education measure to impact knowledge than to impact behavior. Whereas one may regard the effect on behavior as decisive in the end, the effect on literacy is also relevant. Financial literacy represents a competence, which may be applicable in various situations.

<Figure 4 about here>

Our analysis of a comprehensive sample of studies (n=61) leads to a positive assessment of the effectiveness of financial education on financial literacy. This education explains 1.8% of the variance in financial knowledge and, thus, appears only slightly less effective than educational interventions in other domains, such as math and science instruction (cf. Lipsey and Wilson 1993, p.1189). Our positive result is in remarkable contrast to Fernandes et al. (2014, p.1867), who find that financial education only explains 0.4% of the variance in financial literacy and state accordingly that, "financial education yields surprisingly weak changes in financial knowledge presumed to cause financial behavior." However, this result seems a bit fragile as it is based on only 12 studies and cannot, obviously, be replicated in our larger sample of studies (cf. Fernandes et al. 2014, p. 1867).

Relationship between financial literacy and behavior. As we have information from 27 studies about both effects on financial literacy and on financial behavior, we can analyze the importance of financial literacy for behavior. The idea is that financial literacy serves as an intermediator for the effect from education on behavior. Indeed, we find in a linear regression with standard errors clustered at the study level that the effect size on financial literacy is a statistically marginal significant predictor of effect size on financial behavior (b=0.235, SE=0.115, p =0.051). Figure 5 illustrates this result.

<Figure 5 about here>

Thus, an increase in one standard deviation unit in financial literacy scores is related to an average increase of 0.24 standard deviation units of the financial behaviors studied. This

result indicates that the impact of an educational intervention on the financial literacy of participants is an important link in the causal chain that is expected to lead to behavioral change. 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.

4.2 Effect sizes by type of financial behavior

While our analysis so far has shown that financial education interventions have larger effects on financial literacy than on financial behaviors, effect sizes for financial behaviors may vary depending on the financial behavior studied. Figure 6 shows the average effect size for seven categories of financial behaviors targeted by the educational interventions.

<Figure 6 about here>

Average effect sizes for three out of seven categories of outcomes are clearly positive and highly statistically significant at the 1%-level. Additionally, most confidence intervals for the different types of financial behaviors overlap each other, indicating that no extreme differences in impacts depending on the specific form of financial behavior targeted exist. Two things, however, are noteworthy: (i) The average effect size on "budgeting & planning" appears to be higher than the ones 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: this latter average effect size is small (g=0.02) and marginally significant.

Similarly, the average effect sizes for "open bank account" (g=0.02), "insurance" (g=0.04), and "remittances" (g=-0.05) are estimated to be small and insignificant from zero. However, it is noteworthy that these average effect sizes are calculated based on information provided in very few studies per category. Thus, of all financial behaviors studied, the effect

size on borrowing is clearly the smallest precise estimate, indicating that debt-related financial behaviors may be the most challenging to target through financial education.

Overall, these findings correspond to the results provided by Fernandes et al. (2014) and Miller et al. (2015), both also reporting average effect sizes for various financial behaviors, albeit for smaller samples of studies and effect sizes. Qualitatively our analysis confirms the observation by Miller et al. (2015) that effects on borrowing are insignificant from zero and that interventions targeting retirement savings appear to be most successful (cf. Miller et al. 2015, p.238).

4.3 Effect sizes by research methodology

As mentioned in Section 3.1, studies further differ with respect to the research methodology used to investigate the impacts of financial education. Fernandes et al. (2014, p.1865) compared observational studies ("measured financial literacy") to intervention studies ("manipulated financial literacy") and find that observational studies reported higher effect sizes than intervention studies. Additionally, within each category of their selected studies, weaker designs lead to inflated effect sizes, with 75 quasi-experimental studies showing an average effect of about g=0.068 (r=0.034), while 15 RCTs show an average (statistically insignificant) effect of only about g=0.018 (r=0.009). Thus, the comparison of effect sizes by research methodology is important in the assessment of the summary effects of this literature. Fortunately, RCTs are increasing in popularity. Our broader sample of studies covering 45 RCTs and 70 other experiments leads to a more positive assessment of the education impact than before.

Panel A of Table 2 compares average effect sizes for financial behavior and financial literacy as a function of research methodology used. When we focus on financial behaviors as outcomes, RCTs show statistically highly significant (unconditional), effect sizes. These are

only slightly smaller than quasi-experiments, indicating that the small but positive significant effects of financial education exist, even under the most rigorous empirical standards. Compared to Fernandes et al. (2014), the average effect size on financial behavior of our sample of 45 RCTs studying this relationship is 4.5 times higher than in their sample of 15 randomized experiments. RCTs also provide a significant positive effect of financial education on financial literacy, and this effect is again stronger than for financial behavior. However, here the difference between RCTs and other designs is statistically significant at the 1% level.

<Table 2 about here>

4.4 Effect sizes by publication status and quality

A common concern in any meta-analysis is the issue of biases arising from the aggregation of results from studies with different publication status and quality. On the one hand, researchers fear that the tendency of the scientific community to favor statistically significant positive results over insignificant non-results may lead to biased estimates favoring the rejection of the null hypothesis of a zero-effect of financial education on relevant outcomes. The standard solution in the meta-analysis literature is to include as many unpublished studies (grey literature) as possible to address this potential source of bias *a priori*.

On the other hand, economists fear that by aggregating studies of different publication status and quality, the results suffer due to the lack of empirical rigor in grey-literature primary studies. To shed light on this issue in the financial education literature, we compare average effect sizes of financial education interventions by different types of publication status and indicators of quality. *Panel B* of Table 2 compares average effect sizes on financial literacy and behavior by publication status in an academic journal. Interestingly, a bias affects

only the effect size estimates on financial literacy, as they appear to be more than twice as high in published than in unpublished papers (t=3.48). Turning to effect sizes on financial behavior, however, we observe no significant difference in average effect sizes between published and unpublished studies.

Considering indicators of study quality, we code the article influence score (ISI web of knowledge) of the respective journal (and year) for every publication and assign a value of 0 for studies available as working papers. Comparing influential (article influence score >1) with less influential (≤1) publications, we find that the publication bias for financial literacy is now insignificant (t=0.63): Moreover, influential journals tend to publish studies with smaller effect sizes on behavior than non-influential journals, although the difference is not statistically significant.

Next, we code the number of citations for each publication as reported in Google Scholar (as of December 31, 2015). The mean number of citations per article is 32.9 and we split the sample in studies cited above and below this threshold value. Again, we find no significant differences between highly cited studies and others: If anything, highly cited studies tend to report smaller average effect sizes on financial behavior than studies with few citations. Overall, we see that quality bias appears to be no issue in this literature concerned with effects on financial behavior.

We complement these examinations with conventional visual tests for publication bias in order to address the so-called file drawer problem (cf. Stanley and Doucouliagos 2012, p. 73) (i.e. insignificant results are not published at all, not even as grey literature), which can be found in the appendix (see Appendix B). Finally, note that we use alternative regression approaches in Section 7 and Appendix C that are (in principle) capable of generating unbiased estimates in the presence of publication selection.

4.5 Effect sizes by country groups

To investigate another potential source of heterogeneity, we disaggregate our data by country groups. *Panel C* of Table 2 shows effect sizes for financial literacy and financial behavior disaggregated by country groups as classified by the World Bank based on 2014 GNI per capita. Low-income economies are defined as those with a GNI per capita of $1,045 or less (lower-middle income economies are from $1,045 to $4,125; upper-middle income is from $4,126 to $12,735; and high income is greater than $12,736). We find that effect sizes on financial literacy are substantially higher in developed (high income) economies than in developing economies (low income, lower- and upper- middle income economies). This difference is statistically significant at the 1% level. Turning to effect sizes on financial behavior, this difference is statistically insignificant in this unconditional comparison but becomes significant when controlling for other relevant variables (as is shown in Section 5.2).

So far, our meta-analysis yields six important findings: (i) financial education has a significantly positive, but small, impact on financial behavior; (ii) unconditional effect sizes on financial literacy are up to four times higher than effect sizes on financial behavior; (iii) impacts on financial behavior largely have effect sizes with overlapping confidence intervals, but, borrowing behavior may be more difficult to influence than other financial behaviors; (iv) unconditional effect sizes for RCTs are smaller compared to studies relying on less rigorous experimental designs; (v) reported effect sizes on financial behavior are not systematically distorted by publication and quality biases; and (vi) the effect of financial education may be larger in higher income economies.

5 Explaining heterogeneity in treatment effects

Section 4 shows that financial education clearly has an intended effect on financial behavior, and an even stronger effect on financial literacy. However, the average effect is

accompanied by large heterogeneity that is not explained sufficiently by variations in research methodologies employed, study quality, or by the specific type of outcomes studied. Thus, we examine whether there are other systematic 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.

5.1 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: (1) the research design; (2) the intensity of education; (3) the target group of education; and (4) characteristics of the education program. We describe this full set of potential correlates that may influence the effect size of financial education interventions.

(1) 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* the treatment, generally results in higher effect sizes than focusing on the intent 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, p.15; Gertler et al. 2011, p.73). Third, the delay between financial education treatment and measurement of the effect size may negatively influence the effect size since effects of the intervention may decay over time (cf. Fernandes et al. 2014, p.1867). Additionally, effect sizes may be influenced by the sample size. Thus, we control for the inverse standard error of each effect size estimate as a proxy in order to investigate these potential relationships, since this is minimized in larger samples.

(2) A core variable of financial education interventions, which is usually reported in the papers, is the *intensity of education*, i.e. the number of hours taught. It is expected that higher intensities 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.

(3) The expectation regarding a possible relation between the *target group* of education and effectiveness of financial education – measured by its impact (effect size) on financial behavior – is as follows. Generally, learning is easier for younger people and younger people may be more open to new concepts, meaning that the *age* of the target group may have a negative relation to the effect size of financial education. In addition, various empirical studies show that financial literacy is low, especially among the *youth* (e.g. Lusardi and Mitchell 2014), indicating that financial education offered to these participants (with lower baseline scores) may also lead to higher effect sizes in contrast to "older" participants (with relatively high baseline financial literacy). Second, a gender gap in financial literacy is treated as a stylized fact in the literature (cf. Lusardi and Mitchell 2014) which lets us expect gender differences in effect sizes. 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 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.

(4) 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). Thus, we try to capture whether the provision of education came at a point that addressed immediate financial issues (such as borrowers already in default, or micro entrepreneurs borrowing to extend their business). 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).

The full set of variables motivated and described above is defined in Table 3, where descriptive statistics are also provided. Correlations between these variables and effect sizes on financial behavior are found in Table A2 of Appendix A.

<Table 3 about here>

5.2 Meta-regression models explaining effect size

This section examines determinants of financial education effectiveness using a multivariate meta-regression framework including the 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.

We start with a specification where we include more and more reasonable variables and then move to a parsimonious specification that reduces multicollinearity concerns. In order to keep the number of studies considered high, we impute average or default values for missing observations (we show in Section 7 that our results are insensitive to imputation).

The regressions consider groups of variables in four blocks, largely following the logic with which these variables were introduced in Section 5.1. Thus, we start with the methodological controls, then discuss intensity, target group, and, finally, the characteristics of the financial education program.

Method. Starting with the methodological characteristics 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 4, column 1). However, this difference is not significant. As expected, the operationalization of treatment effects as TOT-estimates leads to higher effect sizes and increasing the delay between intervention and measurement decreases effect sizes. 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. Note, that this specification explains only 4.6% of heterogeneity in effect sizes.

<Table 5 about here>

Intensity. Introducing the relationship between intensity per week and duration, column 2 of Table 4 shows that intensity has a marginally significant positive effect on financial behavior (proxied by effect sizes). Thus, controlling for all methodological characteristics, an increase of one hour of financial education per week leads to a 0.003 standard deviation unit increase in the impact on financial behaviors studied. Considering that the average weekly duration is five to six weeks and weekly intensity is only about 3.5 hours, doubling the weekly intensity to 7 hours, while keeping everything else constant at the mean, would lead to an average treatment effect around 9 percent higher than the unconditional treatment effect.

Target group. In column 3, we introduce participant characteristics as explanatory variables. Introducing these variables further reduces unexplained heterogeneity in effect sizes as the model now explains 14.7% of variance. The negative effect for increased baseline literacy and the smaller effect for financial education in low, lower-middle and upper-middle income countries are relatively strong and statistically significant at the 5% and 10% level, whereas all other coefficients are insignificant. Thus, based on this specification, which controls for research design and intensity, effect sizes in developing economies are about 34 percent smaller than in developed economies.

Characteristics of education. In column 4, we add the variables intended for capturing details of the financial education intervention. Adding these variables does impact the coefficient for participant characteristics, such as low-income clients, which increases slightly and is now statistically significant at the 5%-level. Regarding the channel, column 3 shows that no alternative channel generally appears to be more or less effective than financial education in classroom settings. Additionally, we find that mandatory financial education and implementing financial education at a "teachable moment" appear to be important when controlling for the other variables. Specifically, we find, that making financial education

mandatory decreases effect sizes by 0.065 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=0.033 (SE=0.030, p=0.269); thus, economically small and statistically insignificant from zero. In contrast, offering financial education at a teachable moment increases effect sizes by 0.124 standard deviation units. Thus, the predicted value for effect size on financial behavior would be – *ceteris paribus* – g=*0.138* (SE=0.015, p=0.000), i.e. roughly 55 percent larger than the unconditional average effect size found in the sample and statistically highly significant.

Parsimonious specification. We reduce this fully specified model further in order to arrive at a reasonable specification. Column 5 describes a reduced model that shows highly significant effects. A positive effect from increased intensity, a highly significant negative estimate of low-income samples, weaker effects for financially relatively literate participants, weaker effects in developing countries, negative effects for mandatory participation, and highly significant positive effects for financial education offered at a teachable moment. Finally, there are significant coefficients on three methodological control variables.

6 Additional analyses of policy-relevant determinants of effectiveness

Considering our results so far, there seem to be four specific determinants that impact financial behavior and, at the same time, are potentially relevant for policymakers. These are (i) improving financial literacy; (ii) considering the (*ex ante*) financial literacy and income of the sample population; (iii) paying attention to the specific type of financial behavior targeted; and (iv) considering the timing of the intervention in the sense of offering financial education at a teachable moment. For these determinants, we provide more detailed analyses. The methodological control variables are included in all regressions but not reported in the tables.

Increasing financial literacy. We have seen that financial education can impact financial literacy more than financial behavior and that effect sizes on financial literacy correlate positively with effect sizes on financial behavior. Therefore, fostering financial literacy may pay a kind of double dividend as it, first, directly contributes to improving financial behavior and, second, it also increases general competence, which may be a value in and of itself, as it is helpful for various decision making situations beyond just those measured as financial behavior.

Thus, we examine the impact of financial education on financial literacy in the same way as shown for financial behavior. Bivariate relations with respect to the methodological characteristics are qualitatively similar to those with behavior before, although there are differences in detail (see Table A2 in the Appendix).

Turning to the multiple regressions, however, only two determinants remain significant besides the set of methodological controls: the negative linear effect for the age of the participants (implying stronger effects for younger participants) and the positive coefficient for intensity per week of the intervention (see Table 5, column 1); more details are provided in Table A3 in the Appendix. This indicates that improving financial literacy may be easier when working with younger participants and that increasing intensity matters most in terms of education impact on financial knowledge. We see from these examinations that we understand the predictors of effect sizes related to the concrete teaching of financial literacy less than the relation between financial education and financial behavior.

<Table 5 about here>

Financial literacy and income of the sample population. Our results indicate that financial education effectiveness is related to several economic factors of the sample population that we would like to analyze further: effect sizes are smaller for low-income clients, for study-populations in developing economies, and for populations with higher

baseline financial literacy. In column 2 of Table 5 we introduce interaction terms for low-income clients from low/lower-income economies and upper-middle income economies (high income economies are the omitted category). We find that the two-way interaction term between low-income clients and low/lower-income economies is negative and marginally statistically significant, whereas the interaction effect between low-income clients in upper-middle income economies is positive but insignificant. We conclude that effect sizes for low-income clients are even smaller in low- and lower-income economies and that financial education, unfortunately, is least effective among the poor in poor parts of the world.

Contrasting this insight is the impression that effect sizes are smaller in populations with higher financial literacy, in particular as population financial literacy and income (GNI) are positively correlated (cf. Japelli 2010; Klapper et al. 2015). Thus, we investigate this further by again including interaction terms. Column 3 of Table 5 shows that the overall effect of a higher baseline financial literacy at the country-level is still negative. This indicates diminishing returns to additional financial education on higher baseline rates. However, the interaction effects between the population's financial literacy and the income of country groups shows that this relationship is heterogeneous: in the case of low- and lower-middle income economies, this effect is heading in the opposite direction, i.e. increased baseline financial literacy is associated with higher effect sizes on financial behavior. This is not the case in upper-middle and high-income economies, indicating that there may be diminishing returns to additional financial education for populations with relatively high baseline literacy, whereas a basic population financial literacy supports effectiveness of financial education in developing economies.

There are two overall lessons from this more detailed analyses of income and *ex ante* financial literacy: first, financial education for low-income clients is challenging, especially in

developing countries. Second, there are diminishing returns to additional financial education in financially literate populations.

Disaggregating financial behaviors. As discussed in Section 4.2 it appears to be easier to affect financial behaviors in terms of (retirement-) savings and record keeping compared to borrowing behavior. Thus, we split the sample into three categories of financial behaviors and try to replicate our findings. We reduce the choice of variables to avoid problems with degrees of freedom due to relative few observations per split.

Column 4 of Table 5 shows results for the sample split with effect sizes on borrowing behavior. This result matches our main results of the aggregated sample of effect sizes with significant positive effects from increased intensity, negative (albeit insignificant) effects for low-income target groups, and positive effects from offering financial education at a teachable moment. However, we find now insignificant differences between developing and developed economies.

Turning to effect sizes on (retirement-) saving (column 5 of Table 5), we observe that intensity is a significant predictor but timing, in the sense of offering financial education at a teachable moment, seems to matter less. The negative effects in developing economies and increasing baseline levels of financial literacy remain significant.

Finally, turning to budgeting and record keeping behavior (column 6 of Table 5) on which financial education yields the highest effects, we find that intensity is negatively related to effect sizes, albeit only marginally. However, offering financial education at a teachable moment appears to be highly important. Additionally, we find that, in this case, the negative coefficient for low-income clients is now significant, while all of the other signs and relative magnitudes of the coefficients remain the same as in our benchmark estimation (see Table 4).

Overall we find that the positive effects from increased intensity appear to be especially driven by interventions focused on saving and borrowing behavior, whereas the timing

matters most for borrowing and record keeping 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.

Implementing financial education at a teachable moment. Finally, we look further at the role of offering financial education at a teachable moment by modifying its operationalization (column 7 in Table 5). Since teachable moment is negatively correlated to financial education in schools or mandatory courses, replacing teachable moment with these alternative variables should lead to similar results. Thus, we substitute the binary indicator of offering financial education at a teachable moment with the setting variables, while keeping the indicator of mandatory participation. We find that offering financial education at the workplace is associated with significantly higher effect sizes compared to financial education in schools or other (omitted) settings. Also in line with expectations is the observation that the negative coefficient on mandatory financial education is now much larger than in the benchmark specification (Table 4), reinforcing the problematic nature of mandatory financial education. These relationships cannot be confirmed for effect sizes on financial literacy where no obvious connection between the setting and timing of the intervention and gains in financial literacy seem to exist. Overall, these results confirm that the mode and timing of education matters for its effectiveness regarding financial behavior.

7 Robustness

The robustness tests cover seven different aspects and are reported in full in Appendix C. Here, we just mention four important results regarding (i) limiting the sample of studies to RCTs only; (ii) showing results without imputing missing values; (iii) splitting the sample by country groups; and (iv) estimating our main model with inverse variance weighted least

squares. It seems important for the validity of our main findings that results hold qualitatively under all of the above described modifications.

A major concern of researchers is the heterogeneity in quality of studies underlying any meta-analysis. In this sense, 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 1 in Table 6 for the sample split of 36 RCTs covering 161 effect sizes.

<Table 6 about here>

Another major concern may refer to our procedure of imputing missing values. Even though we do this in a conservative way, one may be interested in seeing results without any modification of data. Thus, column 2 in Table 6 reports OLS meta-regression results for complete cases only. These results again confirm earlier results but show larger standard errors for some of the variables.

Moreover, for policymakers it seems interesting to know whether results generalize to all kinds of countries. Therefore, we split the sample into developing and advanced economies. While results for advanced, i.e. high income, economies are very similar to the total results, there is less explanatory power regarding developing economies. This was foretold by the negative coefficient for these countries in the general regressions, indicating that the institutional environment may hinder education effectiveness. Nevertheless, increasing intensity is still helpful. Beyond that, innovations in curricula may be needed.

Finally, we re-estimate our main regression with several alternative econometric methods (cf. Appendix C). Among these we show the result from applying weighted least squares with inverse variance weights, a strategy suggested in Stanley and Doucouliagos

(2012). Reassuringly, results are very similar to our main approach, relying on unweighted ordinary least squares.

8 Concluding policy discussion

This meta-analysis covers studies that can contribute to policy objectives, such as better financial literacy and improved financial behavior. Due to this close link to economic policy, we discuss insights with potential policy relevance in four steps, from the pure effect size to welfare considerations:

Effect size. Most important in policy terms is the finding that financial behavior can be improved through financial education measures. Of the same importance, however, is the finding that the respective coefficient on effect size is very small: this small size, i.e. a relatively small impact of financial education on intended financial behavior, provides a strong motivation to discuss possible ways of increasing the effectiveness of financial education programs.

Irrelevant determinants. Some of the determinants of successful financial education we have identified are irrelevant from a public policy point of view because they cannot be addressed meaningfully by policymakers. Most obvious in this respect is the temptation to evaluate financial education in sub-optimal ways, e.g. by evaluating its impact without the necessary methodological rigor of conducting a randomized experiment that accounts for selection effects (RCT), focusing on the treatment effects on the treated (TOT) instead of those intended to treat (ITT), and choosing to measure program effectiveness after a short delay. All this may contribute to inflated observed effect sizes without improving outcomes in reality.

A similar effect takes place by selecting the topic of education and the target group. Effects will probably be larger if one focuses on retirement-saving behavior instead of general

borrowing behavior and the effect will be larger if higher income clients are educated instead of the relatively poor. In addition, there is obvious self-selection into education measures, indicating that those who have large knowledge and behavioral deficits abstain from education. Policy may see this as a particular challenge. Simply neglecting borrowing behavior and the poor would be obviously misguided.

Relevant determinants. In this sense, there is something that policy can actively consider in order to improve effectiveness of education: The intensity and timing of the educational intervention. Increasing the intensity of financial education seems to affect both the impacts on financial literacy (in terms of knowledge gain) and on financial behavior (likelihood of behavioral change). Moreover, the timing and characteristics of education programs also matter: in this respect it is clear that there is a complex of interrelated issues contributing to success or failure. Whereas effects on *financial literacy* are bigger for the young (which would suggest implementing financial education in school) it looks like mandatory financial education is generally less effective at impacting *financial behavior*. Linked to these results, the timing matters if education is tied to concrete issues of interest to the target group, the so-called teachable moment that increases the likelihood of changes in individual financial behavior. Thus, financial education may be more effective at affecting financial behavior if offered outside of traditional educational institutions, as the age of participants does not systematically impact effects on behavior.

Considering the evidence presented in this paper, what effect sizes from financial education can potentially be realized? Given that the financial education program is evaluated by conducting a randomized experiment (RCT=1) and treatment effects are reported as intent-to-treat effects (TOT=0), while everything else (delay in measurement, study size, duration and weekly intensity, country income, population literacy, and program details) is kept at the empirical mean, the predicted effect size on financial behavior would be g=0.064 (SE=0.026,

p=0.019) for average income clients, and *g*=0.023 (*SE*=0.018, *p*=0.206) for low-income clients.

The order of magnitude improves considerably when favoring voluntary financial education as opposed to mandatory formats and offering financial education at a teachable moment. Implementing these elements increases this effectiveness to *g*=0.115 (SE= 0.029, p=0.000) for average income clients and to g=0.075 (SE=0.021, p=0.001) for low-income clients. These effect sizes are about 0.8 times larger than the unconditional average effect size for average income clients and about 2.2 times larger for low-income clients. Increasing intensity can further increase effect sizes with an effect size gain of 0.004 standard deviation units per additional weekly hour of training.

Welfare considerations. We clearly find that policy can be effective and that it can influence the degree of effectiveness; however, this does not provide a welfare consideration. Financial education is costly and policy should discuss when and where to invest scarce funds. Some argue that it may be more effective to simplify financial decisions than educate consumers. Simplifying can be achieved, for example, by limiting the kind of available products (choices), altering the choice architecture (e.g. Caroll et al. 2009), by working with nudges (e.g. Thaler and Benartzi 2004; Willis 2011), and commitment devices (e.g. Brune et al. 2016). Alternatively offering incentives (e.g. Saez 2009) or more rigid consumer financial protection policies are further policy options (cf. Campbell et al. 2011).

Another argument refers to the endogeneity of financial knowledge, which is of limited interest for many low-income consumers (e.g. Lusardi et al. 2016). In this sense, their self-selection out of education measures has some rational motivation. This rationale is fostered by the limited effectiveness of financial education measures. However, even if many individuals may not profit from education about (retirement-) savings (Lusardi et al., 2015), they may still benefit from, e.g., budgeting education.

Although our analysis does not provide exact information to facilitate concrete policy decisions, there are two arguments in favor of financial education. First, the small effect of financial education on behavior is accompanied by a much larger effect on financial literacy. This is useful as financial literacy not only supports financial behavior but can also be regarded a more general element of educational development, increasing individual competencies. Second, the average effect of financial education is small, but so is the average intensity. More than 70% of our considered studies invest no more than one day in education, indicating that these measures have small effects but also low costs.

At this stage, more research is needed to support policymakers in their decision-making. A next step in research may include going beyond the analysis of effectiveness to examining the determinants of effectiveness more thoroughly. This could include investigating combinations of these policies to identify and evaluate possible complementarities and possibilities for substitution. An important element of such analyses would be extending the impact analysis with an analysis of costs involved in the provision of specific financial education measures.

References

Agarwal, S., and Mazumder, B. (2013). Cognitive abilities and household financial decision making. *American Economic Journal: Applied Economics*, 5(1): 193–207.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Shapiro, J., Thuysbaert, B., and Udry, C. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 348(6236): 1260799-1–16.

Behrman, J.R., Mitchell, O.S., Soo, C.K., and Bravo, D. (2012). How financial literacy affects household wealth accumulation. *American Economic Review: Papers and Proceedings*, 102(3): 300–304.

Blue, L., Grootenboer, P., and Brimble, M. (2014). Financial literacy education in the curriculum: Making the grade or missing the mark? *International Review of Economics Education*, 16, Part A(0): 51–62.

Bruhn, M., Ibarra, G.L., and McKenzie, D. (2014). The minimal impact of a large-scale financial education program in Mexico City. *Journal of Development Economics*, 108: 184–189.

Bruhn, M., de Souza Leao, L., Legovini, A., Marchetti, R., and Zia, B. (2016). The impact of high school financial education: Evidence from a large-scale evaluation in Brazil. *American Economic Journal: Applied Economics,* forthcoming.

Brune, L., Giné, X., Goldberg, J. and Yang, D. (2016) Facilitating savings for agriculture: Field experimental evidence from Malawi. *Economic Development and Cultural Change*, 64: 187–220.

Bucher-Koenen, T. and Ziegelmeyer, M. (2014). Once burned, twice shy? Financial literacy and wealth losses during the financial crisis. *Review of Finance*, 18(6): 2215–2246.

Campbell, J.Y. (2006). Household finance. *Journal of Finance*, 61(4): 1553–1604.

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

Card, D., Kluve, J., and Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *Economic Journal*, 120(548): F452–F477.

Card, D., Kluve, J., and Weber, A. (2015). What works? A meta analysis of recent active labor market program evaluations. *NBER Working Paper 21431*.

Carroll, G. D., Choi, J.J., Laibson, D., Madrian, B.C, and Metrick, A. (2009). Optimal defaults and active decisions. *Quarterly Journal of Economics,* 124(4): 1639–1674.

Cho, Y. and Honorati, M. (2014). Entrepreneurship programs in developing countries: A meta regression analysis. *Labour Economics*, 28: 110–130.

Choi, J.J., Laibson, D., and Madrian, B.C. (2010). Why does the law of one price fail? An experiment on index mutual funds. *Review of Financial Studies*, 23(4): 1405–1432.

Choi, J.J., Laibson, D., and Madrian, B.C. (2011). $100 bills on the sidewalk: Suboptimal investment in 401 (k) plans. *Review of Economics and Statistics*, 93(3): 748–763.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* Routledge.

Cole, S., Sampson, T., and Zia, B. (2011). Prices or knowledge? What drives demand for financial services in emerging markets? *Journal of Finance*, 66(6): 1933–1967.

Collins, J.M. and O'Rourke, C.M. (2010). Financial education and counseling—still holding promise. *Journal of Consumer Affairs*, 44(3): 483–498.

Collins, J.M. (2013). The impacts of mandatory financial education: Evidence from a randomized field study. *Journal of Economic Behavior and Organization*, 95: 146–158.

DerSimonian, R. and Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3): 177–188.

Doi, Y., McKenzie, D., and Zia, B. (2014). Who you train matters: Identifying combined effects of financial education on migrant households. *Journal of Development Economics*, 109: 39–55.

Drexler, A., Fischer, G., and Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, 6(2): 1–31.

Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, 118(3): 815–842.

Fernandes, D., Lynch, Jr., J.G., and Netemeyer, R.G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science*, 60(8): 1861–1883.

Fox, J., Bartholomae, S., and Lee, J. (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., Goette, L., and Meier, S. (2013). Numerical ability predicts mortgage default. *Proceedings of the National Academy of Sciences*, 110(28), 11267–11271.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., and Vermeersch, C. M. (2011). *Impact evaluation in practice.* World Bank Publications, Washington D.C.

Gibson, J., McKenzie, D., and Zia, B. (2014). The impact of financial literacy training for migrants. *World Bank Economic Review*, 28(1): 130–161.

Hastings, J.S., Madrian, B.C., and Skimmyhorn, W.L. (2013). Financial literacy, financial education, and economic outcomes. *Annual Review of Economics*, 5: 347–373.

Imbens, G. W. and Wooldridge, J. M. (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–F451.

Klapper, L., Lusardi, A. and van Oudheusden (2015). Financial literacy around the world: Insights from the standard and poor's rating services global financial literacy survey. Available at: http://www.finlit.mhfi.com.

Knapp, G. and Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, 22(17): 2693–2710.

Lipsey, M.W. and Wilson, D.B. (2001). *Practical meta-analysis.* Sage, Thousand Oaks, CA.

Lusardi, A. and Mitchell, O.S. (2007). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics*, 54(1): 205–224.

Lusardi, A. and Mitchell, O.S. (2014). The economic importance of financial literacy: theory and evidence. *Journal of Economic Literature*, 52(1): 5–44.

Lusardi, A., Michaud, P.C., Mitchell, O.S. (2015). Using a life cycle model to evaluate financial literacy program effectiveness. *GFLEC Working Paper 2015-4*.

Lusardi, A., Michaud, P.C., and Mitchell, O.S. (2016). Optimal financial knowledge and wealth inequality. *Journal of Political Economy*, forthcoming.

Lührmann, M., Serra-Garcia, M., and Winter, J. (2015). Teaching teenagers in finance: Does it work? *Journal of Banking and Finance*, 54: 160–174.

Miller, M., Reichelstein, J., Salas, C., and Zia, B. (2015). Can you help someone become financially capable? A meta-analysis of the literature. *World Bank Research Observer*, 30(2): 220–246.

Muller, S. M. (2015). Causal interaction and external validity: Obstacles to the policy relevance of randomized evaluations. *World Bank Economic Review*, 29: S217–S225.

Pritchett, L. and Sandefur, J. (2015). Learning from Experiments When Context Matters. *American Economic Review: Papers and Proceedings*, 105(5):471–75.

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): 204–228.

Sayinzoga, A., Bulte, E. H., and Lensink, R. (2016). Financial literacy and financial behaviour: Experimental evidence from rural Rwanda. *Economic Journal*, forthcoming.

Skimmyhorn, W. (2016). Assessing financial education: Evidence from boot camp. *American Economic Journal: Economic Policy,* 8(2): 322–343.

Stanley, T. D. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives*, 15(3): 131–150.

Stanley, T. D. and Doucouliagos, H. (2012). *Meta-regression analysis in economics and business*, Routledge, New York, NY.

Stango, V. and Zinman, J. (2009). Exponential growth bias and household finance. *Journal of Finance*, 64(6), 2807–2849.

Thaler R.H., Benartzi S. (2004). Save more tomorrow: Using behavioral economics to increase employee saving. *Journal of Political Economy.* 112: 164–87.

van Rooij, M., Lusardi, A., and Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, 101(2): 449–472.

van Rooij, M. C., Lusardi, A., and Alessie, R.J. (2012). Financial literacy, retirement planning and household wealth. *Economic Journal*, 122(560): 449–478.

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): 429–434.

Xu, L., and Zia, B. (2012). Financial literacy around the world: an overview of the evidence with practical suggestions for the way forward. *World Bank Policy Research Working Paper 6107.*

Zinman, J. (2015). Household debt: Facts, puzzles, theories, and policies. *Annual Review of Economics*, 7: 251–276.

Table 1: Summary of financial education studies by publication date and country

	Number of studies (1)	Percent of sample (2)	
A By publication date			
1999	2	1.74	
2000	0	0.00	
2001	5	4.35	
2002	1	0.87	
2003	4	3.48	
2004	2	1.74	
2005	6	5.22	
2006	5	4.35	
2007	7	6.09	
2008	6	5.22	
2009	7	6.09	
2010	12	10.43	
2011	7	6.09	
2012	15	13.04	
2013	11	9.57	
2014	11	9.57	
2015	14	12.17	
B By country of intervention			*Income*
Australia	1	0.87	High
Bosnia and Herzegovina	1	0.87	Upper-middle
Brazil	1	0.87	Upper-middle
China	1	0.87	Upper-middle
Dominican Republic	1	0.87	Upper-middle
Germany	1	0.87	High
Ghana	1	0.87	Lower-middle
Hong Kong, China	1	0.87	High
India	8	6.96	Lower-middle
Indonesia	2	1.74	Lower-middle
Italy	6	5.22	High
Kenya	1	0.87	Lower-middle
Mexico	1	0.87	Upper-middle
New Zealand	2	1.74	High
Pakistan	1	0.87	Lower-middle
Qatar	1	0.87	High
Rwanda	1	0.87	Low
Singapore	1	0.87	High
South Africa	1	0.87	Upper-middle
Spain	1	0.87	High
Sri Lanka	1	0.87	Lower-middle
Tanzania	1	0.87	Low
USA	78	67.83	High
Uganda	1	0.87	Low
Low inc. econ.	3	2.61	
Lower-middle inc. econ.	14	12.17	
Upper-middle inc. econ.	6	5.22	
High inc. econ.	92	80.00	
Total	115	100	

Table 2: Effect sizes of financial education by research design, study quality, and country group

Outcome	Type	Studies	Obs.	ES (g)	SE	p-value	Diff. (t-value)
A Effect sizes by research design							
Fin. Literacy	RCTs	30	122	0.196	0.032	0.000	-0.257***
	Quasi-exp.	30	46	0.453	0.086	0.000	(4.611)
Fin. Behavior	RCTs	37	162	0.087	0.014	0.000	-0.005
	Quasi-exp.	46	98	0.092	0.024	0.000	(0.217)
B Effect sizes by publication status							
Fin. Literacy	Published	32	94	0.347	0.075	0.000	0.179***
	Unpublished	29	75	0.168	0.043	0.001	(3.481)
Fin. Behavior	Published	44	102	0.097	0.016	0.000	0.013
	Unpublished	38	158	0.084	0.016	0.000	(0.617)
Fin. Literacy	High influence	8	26	0.307	0.023	0.000	0.046
	Low influence	53	143	0.261	0.053	0.000	(0.632)
Fin. Behavior	High influence	23	60	0.070	0.019	0.001	-0.024
	Low influence	59	200	0.095	0.015	0.000	(0.978)
Fin. Literacy	Highly cited	10	14	0.225	0.084	0.025	-0.047
	Few citations	51	155	0.272	0.051	0.000	(0.487)
Fin. Behavior	Highly cited	34	61	0.075	0.027	0.009	-0.018
	Few citations	48	199	0.093	0.014	0.000	(0.7642)
C Effect sizes by country income group							
Fin. Literacy	High income	47	107	0.345	0.067	0.000	0.211***
	Developing	14	62	0.134	0.031	0.001	(4.009)
	- Low	*3*	*6*	*0.219*	*0.069*	*0.086*	
	- Lower-middle	*6*	*41*	*0.140*	*0.045*	*0.027*	
	- Upper-middle	*5*	*15*	*0.084*	*0.019*	*0.011*	
Fin. Behavior	High income	62	140	0.079	0.021	0.000	-0.020
	Developing	20	120	0.099	0.011	0.000	(0.983)
	- Low	*2*	*13*	*0.138*	*0.091*	*0.371*	
	- Lower-middle	*12*	*85*	*0.093*	*0.005*	*0.000*	
	- Upper-middle	*6*	*22*	*0.106*	*0.045*	*0.064*	

Notes: Average effect sizes *(g)* estimated via OLS regressions of effect sizes fitting only an intercept. Sample is split by an indicator of research design, publication status or country group. Country groups are based on the World Bank Atlas method and refer to 2014 data on GNI per capita. Low-income economies are defined as those with a GNI per capita of $1,045 or less in 2014, lower-middle income economies are defined by a GNI per capita between $1,045 and $4,125, upper-middle income economies are those with a GNI per capita between as $4,126 and $12,735, and high income economies are defined by a GNI per capita greater than $12,736. Standard errors are clustered at the study-level. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table 3: Summary statistics at the study level for independent variables

Variable	N of Obs.	Mean	Std. Dev.	Min	Max
RCT	115	0.391	0.490	0.000	1.000
TOT	104	0.442	0.499	0.000	1.000
Delay	85	70.384	235.083	0.000	1566.000
1/SE	115	53.880	185.269	2.740	1636.712
Intensity	80	11.648	16.743	0.100	108.000
Duration	66	7.200	15.603	1.000	103.000
Age	99	31.206	13.893	9.000	63.870
Youth	115	0.295	0.458	0.000	1.000
Percent female	115	55.223	17.948	0.000	100.000
Low income clients	97	0.536	0.501	0.000	1.000
Years of schooling	115	11.431	2.656	3.700	13.600
FL in population	115	50.500	11.802	24.000	66.000
Country groups					
Low-/lower-middle income economies	115	0.148	0.356	0.000	1.000
Upper middle-income economies	115	0.052	0.223	0.000	1.000
High income economies	115	0.800	0.402	0.000	1.000
Channel					
Classroom	115	0.791	0.408	0.000	1.000
Counseling	115	0.061	0.240	0.000	1.000
Online	115	0.061	0.240	0.000	1.000
Mandatory	87	0.253	0.437	0.000	1.000
Incentivized	79	0.329	0.473	0.000	1.000
Setting					
School	115	0.313	0.466	0.000	1.000
Workplace	115	0.148	0.356	0.000	1.000
Teachable moment	115	0.417	0.495	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. "Youth" is a dummy variable taking the value "1" if "Age" is smaller than 20. "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. "Channel" is a categorical variable operationalized in the form of three dummy variables: Classroom, Counseling, Online, and Other where "Other" is the default (omitted) category in the regressions. "Mandatory" is a dummy variable with "1" indicating mandatory participation in financial education and "0" if it was voluntary. "Incentivized" is a dummy variable with "1" when incentives to participate where provided and "0" if participation was unconditional on incentives. "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. "Teachable moment" is a dummy variable indicating whether the financial education intervention was offered at a teachable moment.

Table 4: Explaining heterogeneity in effect sizes on financial behavior

	(1)	(2)	(3)	(4)	(5)
RCT	-0.026	-0.025	-0.080*	-0.104***	-0.105***
	(0.029)	(0.030)	(0.041)	(0.033)	(0.034)
TOT	0.057	0.059	0.093**	0.075**	0.078**
	(0.038)	(0.038)	(0.037)	(0.031)	(0.030)
Delay	-0.000	-0.000	0.000	0.000*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
1/SE	-0.000***	-0.000***	-0.000***	-0.000**	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intensity per week		0.003*	0.003**	0.005***	0.004***
		(0.002)	(0.001)	(0.001)	(0.001)
Duration		-0.001	-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)
Age			0.001	-0.001	
			(0.001)	(0.001)	
Percent female			0.000	-0.000	
			(0.001)	(0.001)	
Low income clients			-0.037	-0.049**	-0.045**
			(0.023)	(0.024)	(0.022)
Years of schooling			-0.004	-0.001	
			(0.013)	(0.008)	
FL in population			-0.009**	-0.011***	-0.011***
			(0.003)	(0.003)	(0.003)
Country group					
Low/lower-mid.inc.econ			-0.202**	-0.319***	-0.309***
			(0.091)	(0.106)	(0.094)
Upper-mid. inc. econ.			-0.149*	-0.167*	-0.175**
			(0.082)	(0.096)	(0.085)
Channel					
a.) Classroom				0.002	
				(0.035)	
b.) Counseling				0.031	
				(0.035)	
c.) Online				0.022	
				(0.038)	
Mandatory				-0.065**	-0.052**
				(0.032)	(0.024)
Incentivized				0.003	
				(0.035)	
Teachable moment				0.124***	0.117***
				(0.026)	(0.027)
Constant	0.103***	0.093***	0.592***	0.750***	0.702***
	(0.025)	(0.028)	(0.172)	(0.206)	(0.153)
R^2	0.046	0.061	0.147	0.224	0.217
n (Studies)	82	82	82	82	82
n (Effect sizes)	260	260	260	260	260

Notes: Non-standardized coefficients from OLS regressions. Dependent variable is effect size (Hedges' g) on financial behavior. Robust standard errors clustered at the study-level in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table 5: Detailed analyses of effectiveness' determinants

	(1) FL	(2) FB	(3) FB	(4) Borrow	(5) Assets	(6) Budget	(7) FB
Intensity / week	0.010***	0.005***	0.004***	0.005**	0.006**	-0.025*	0.002
	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)	(0.012)	(0.002)
Duration	0.010***	0.005***	0.004***	0.001	0.006**	-0.019***	-0.000
	(0.003)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.001)
Youth	0.262**						
	(0.109)						
Low inc. clients		-0.026	-0.014	-0.053	-0.043	-0.161*	-0.001
		(0.028)	(0.022)	(0.035)	(0.037)	(0.089)	(0.031)
FL in pop.		-0.013***	-0.016***	-0.001	-0.013***	-0.041***	-0.009***
		(0.003)	(0.005)	(0.003)	(0.004)	(0.009)	(0.003)
Country group							
Low-/lower inc	-0.149**	-0.301***	-0.301***	0.065	-0.316**	-1.396***	-0.192**
econ.	(0.058)	(0.085)	(0.085)	(0.101)	(0.132)	(0.372)	(0.094)
Upper-mid. inc.	-0.044	-0.263**	-0.264	0.076	-0.186*	-1.201***	-0.140
econ.	(0.063)	(0.112)	(0.320)	(0.088)	(0.106)	(0.289)	(0.089)
Low/lower inc. econ			0.013**				
× Fl in pop.			(0.005)				
Upper-mid. inc econ.			-0.001				
× Fl in pop.			(0.007)				
Low/lower inc. econ		-0.089*					
× low inc. clients		(0.048)					
Upper-mid. inc econ		0.073					
× low inc. clients		(0.077)					
Setting							
a.) School							-0.008
							(0.033)
b.) Workplace							0.081*
							(0.041)
Mandatory		-0.067**	-0.067***				-0.125***
		(0.026)	(0.025)				(0.034)
Incentivized							
Teach. moment		0.107***	0.116***	0.073***	0.049	0.320***	
		(0.027)	(0.028)	(0.022)	(0.050)	(0.072)	
Constant	0.542***	0.835***	0.977***	0.093	0.824***	2.820***	0.605***
	(0.141)	(0.180)	(0.266)	(0.204)	(0.258)	(0.642)	(0.157)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.376	0.227	0.233	0.444	0.358	0.544	0.180
n (Studies)	61	82	82	25	56	16	82
n (Effect sizes)	169	260	260	68	142	35	260

Notes: Non-standardized coefficients from OLS regressions with clustered standard errors at the study-level in parentheses. All regressions include the full set of methodological controls which are not reported in the table. Dependent variable in columns (1) and (2) is effect size (Hedges' g) on financial literacy. Dependent variable in columns (3) to (7) is effect size on financial behavior. Columns (4) to (6) present results for sample splits for three categories of financial behavior. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table 6: Robustness checks

	(1) RCTs only	(2) No Imputations	(3) Dev. econ. split	(4) High inc. split	(5) WLS 1/SE weights
RCT		-0.086		-0.152***	-0.091**
		(0.053)		(0.031)	(0.037)
TOT	0.033	0.126**	-0.039	0.144***	0.102***
	(0.043)	(0.056)	(0.026)	(0.033)	(0.030)
Delay	-0.001	-0.000	0.001	0.000**	0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
1/SE	0.001	0.001*	0.001	-0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)
Intensity /week	0.005***	0.004**	0.006***	0.008**	0.003
	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)
Duration	-0.005**	-0.001	-0.001	-0.002**	-0.000
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Low inc. clients	-0.005**	-0.101**	-0.049*	0.007	-0.015
	(0.002)	(0.038)	(0.002)	(0.031)	(0.029)
FL in pop.	-0.010***	-0.014***	-0.006**	-0.018***	-0.012***
	(0.003)	(0.003)	(0.002)	(0.005)	(0.003)
Country group					
Low/lower-mid.	-0.303***	-0.391***			-0.339***
inc. econ.	(0.105)	(0.125)			(0.084)
Upper-mid. inc..	-0.185**	-0.195*			-0.186**
econ.	(0.089)	(0.103)			(0.079)
Mandatory	-0.054	-0.107**		-0.099***	-0.095**
	(0.058)	(0.048)		(0.031)	(0.041)
Teach. moment	0.083**	0.089*	-0.008	0.113***	0.088***
	(0.036)	(0.049)	(0.019)	(0.029)	(0.030)
Constant	0.635***	0.921***	0.246**	1.034***	0.755***
	(0.182)	(0.206)	(0.090)	(0.262)	(0.157)
R^2	0.177	0.195	0.076	0.361	0.308
n (Studies)	36	27	20	62	82
n (Effect Sizes)	161	147	120	140	260

Notes: Column (1) presents results from OLS meta-regressions for the subsample of RCTs only. Column (2) reports results for complete cases only. Columns (3) and (4) presents results for the sample splits for developing and developed countries respectively. These splits include only variables for which differential information from at least two studies are available. Column (5) shows results of weighted least squares estimation with inverse variance weights (cf. Stanley and Doucouliagos 2012). Standard errors clustered at the study-level in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level.

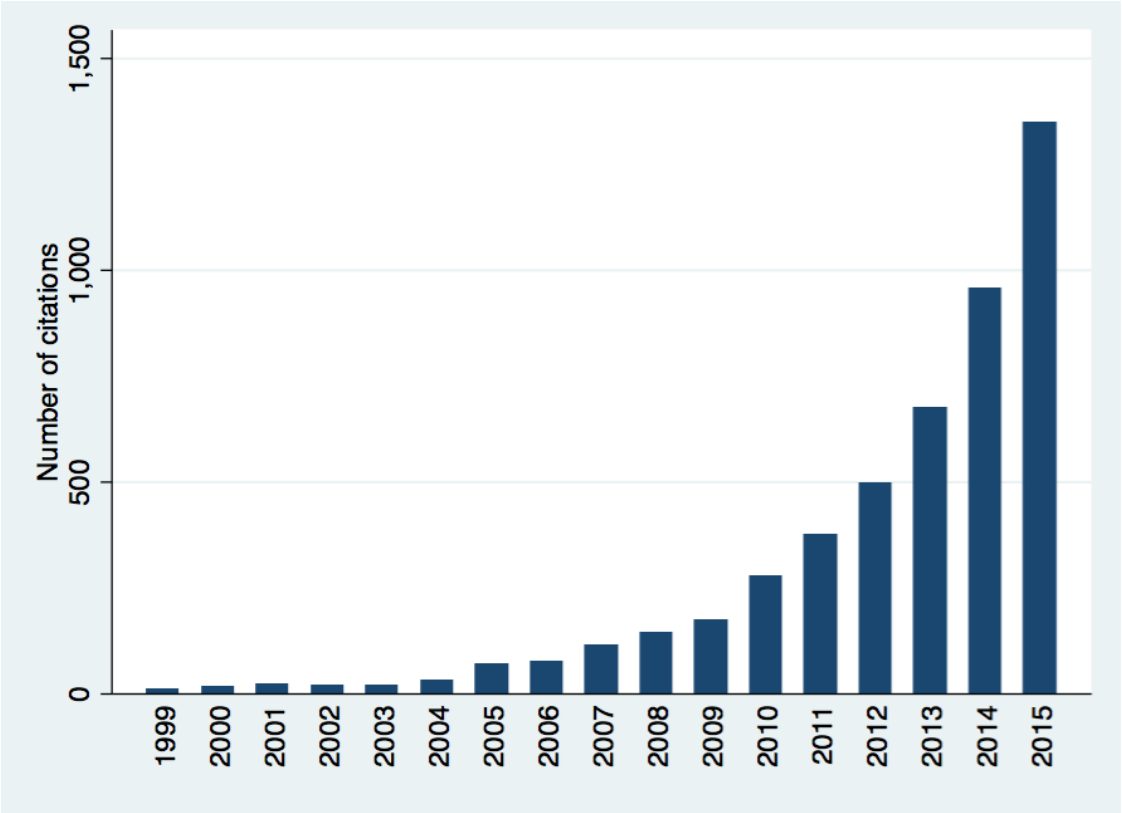

Figure 1: Citations of published items with the keyword *financial literacy* per year, source: SSCI.

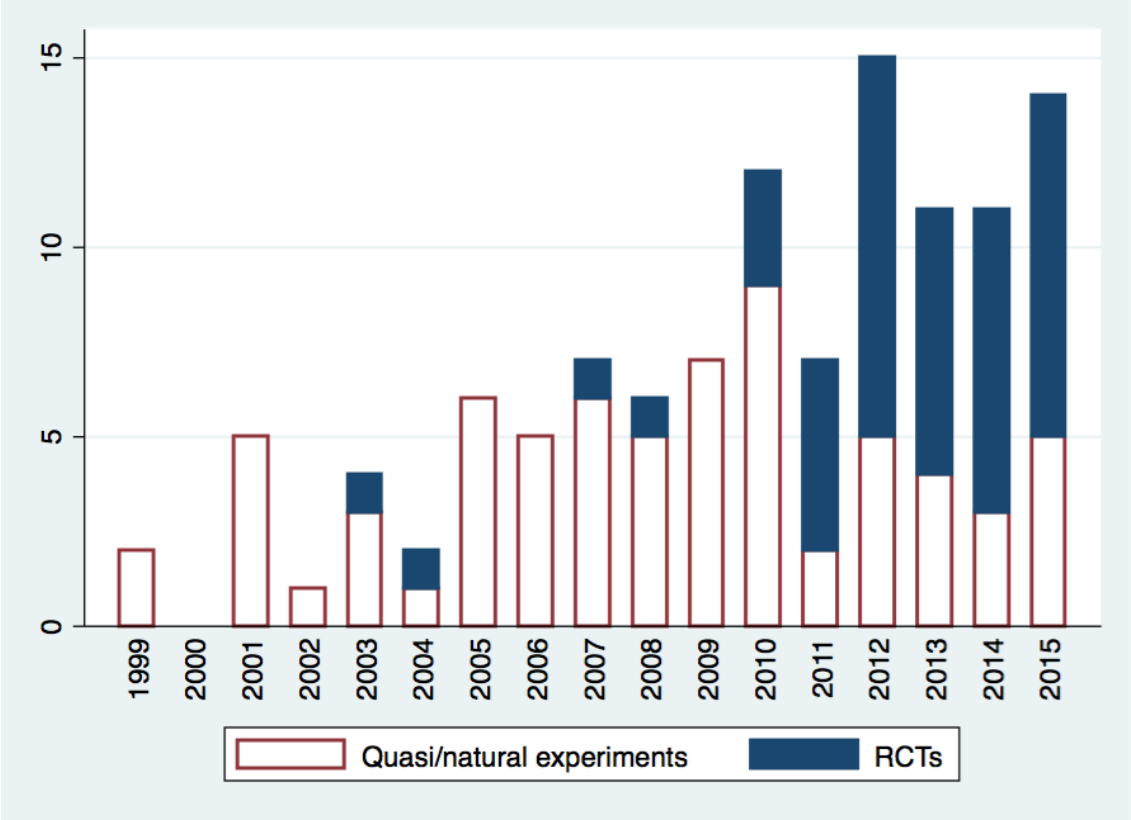

Figure 2: Number of studies in our sample by research methodology used per year

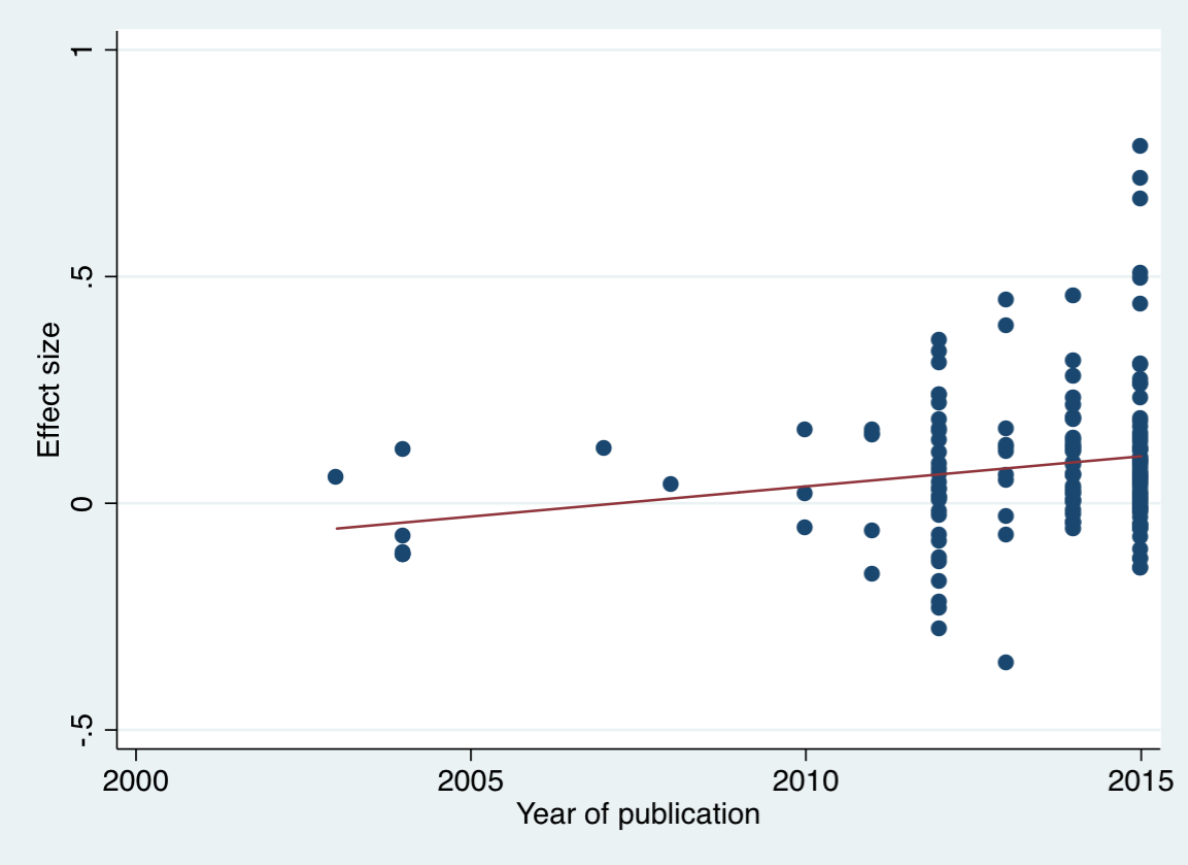

Figure 3: Effect sizes on financial behavior documented in RCTs by latest date of publication

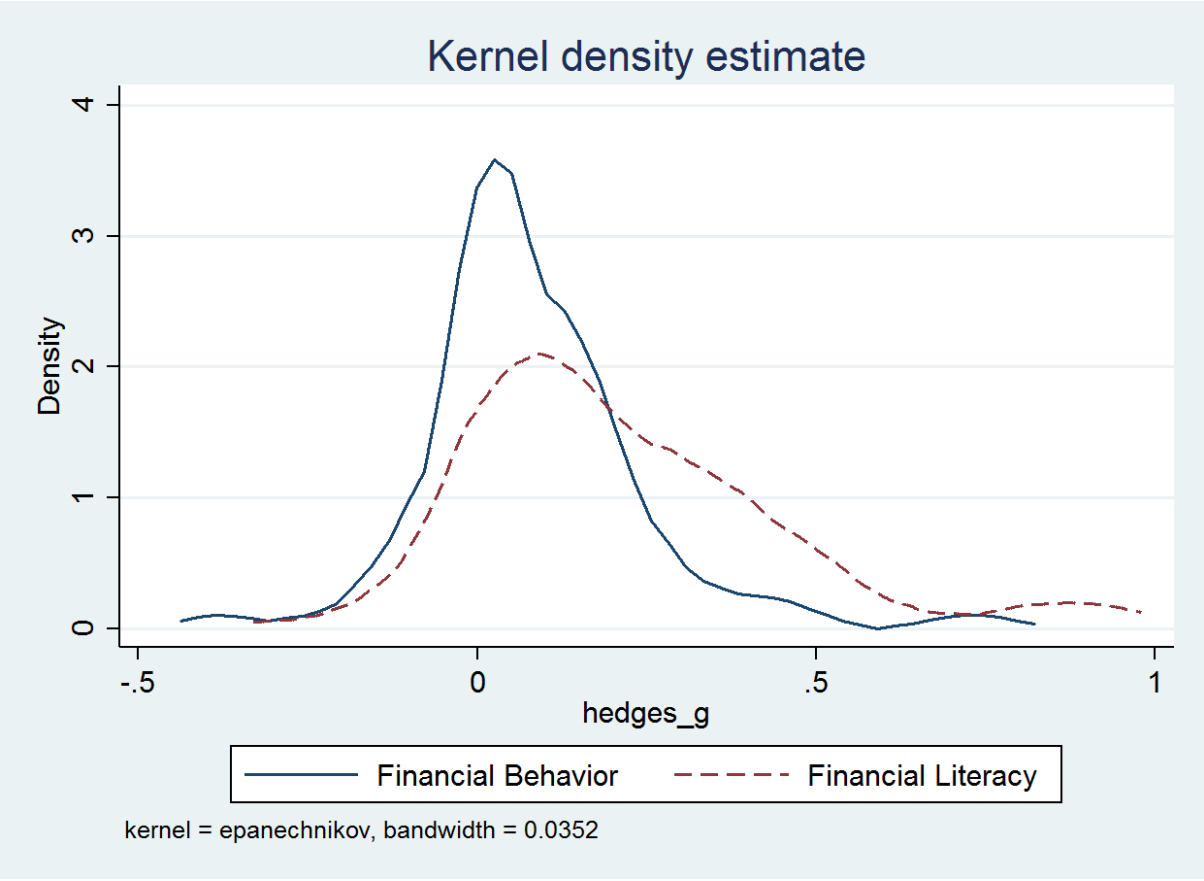

Figure 4: Kernel-density estimates of effect sizes by outcome (for Hedge's g<1)

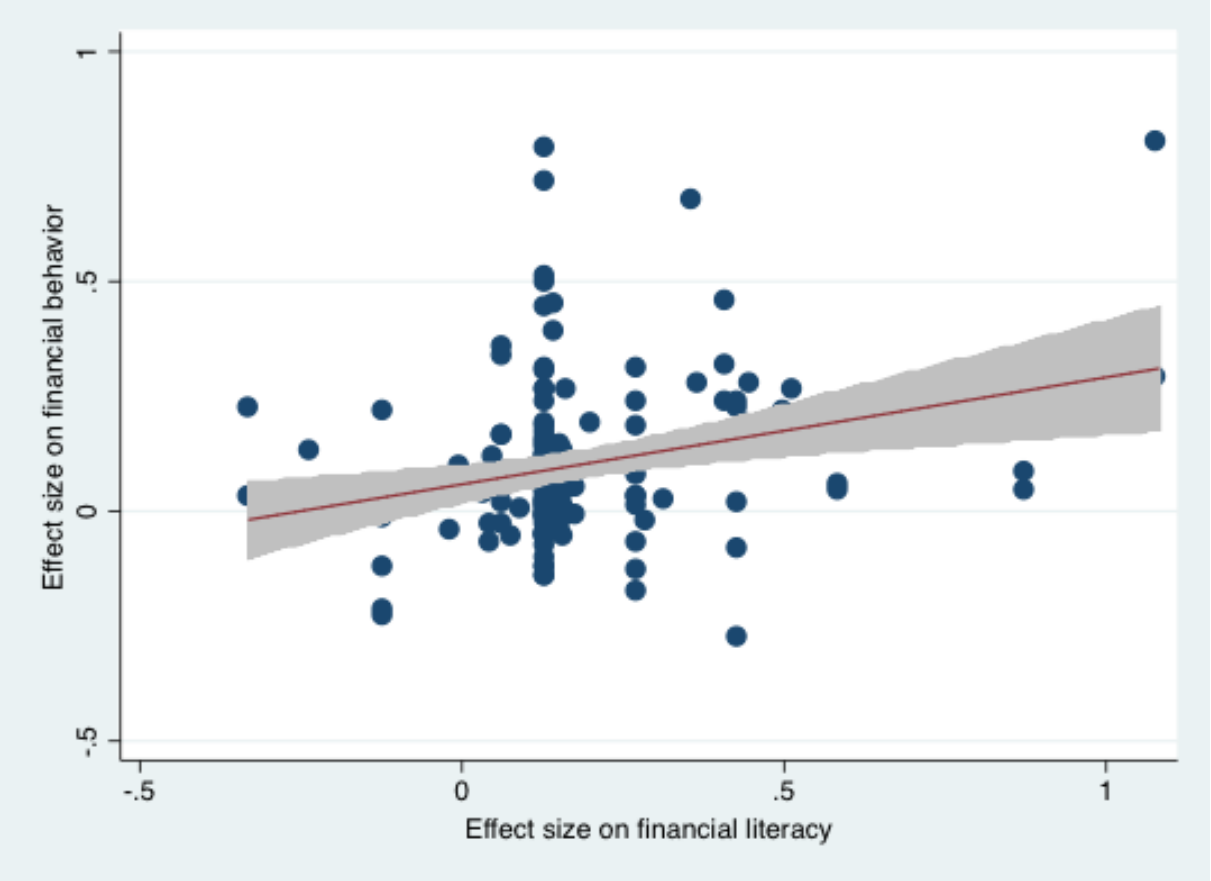

Figure 5: Relationship between ES on FL and ES on FB in a sample of 28 impact evaluation studies

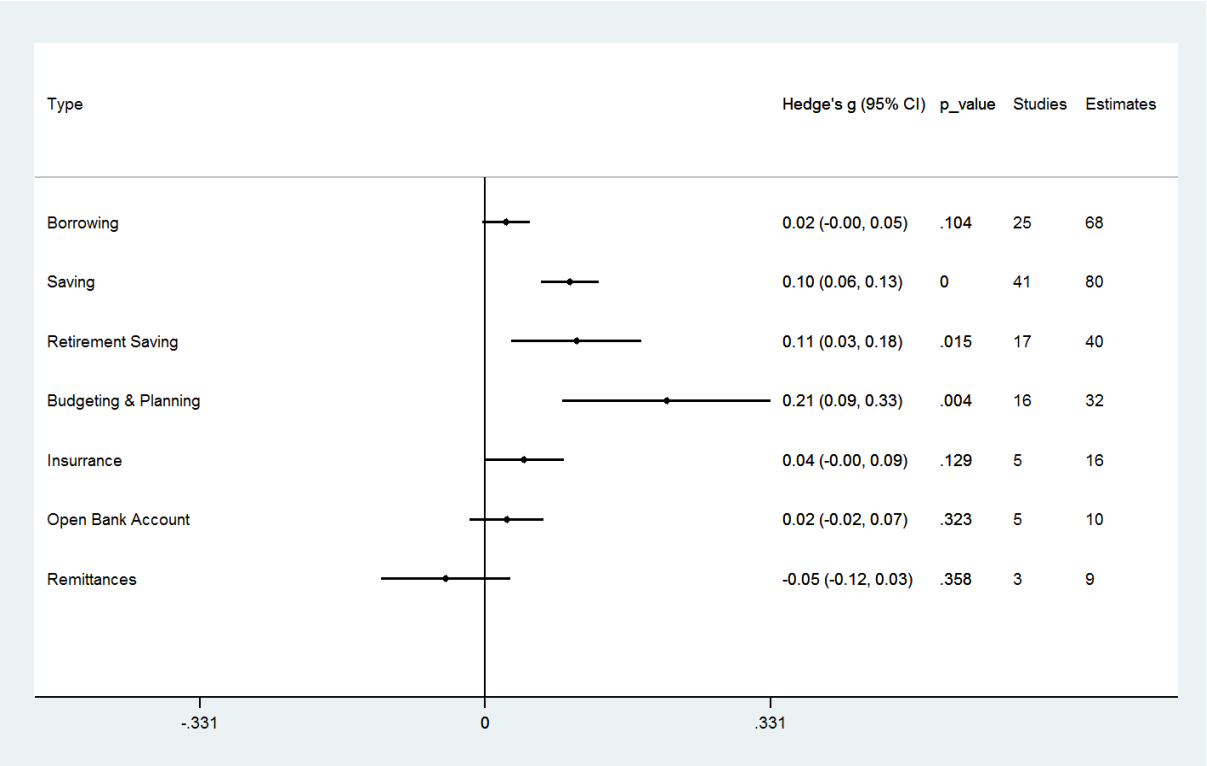

Figure 6: Forest plot of effect sizes by type of financial behavior studied

Appendix

(online appendix not intended for publication)

to accompany

"Does financial education impact financial behavior, and if so, when?"

Appendix A

Table A1: Summary of estimated financial education impacts

Outcome	Significance at 5%			Significance at 10%			Average effect size
	Negative	*Insig.*	*Positve*	*Negative*	*Insig.*	*Positve*	(SE)
A Effects on financial literacy							
Fin. literacy	1 (0.59%)	65 (38.46%)	103 (60.94%)	2 (1.19%)	54 (32.14%)	112 (66.67%)	0.268*** (0.462)
B Effects on financial behavior							
Fin. behavior	5 (1.92%)	164 (63.08%)	91 (35.00%)	13 (5.00%)	136 (52.31%)	111 (42.69%)	0.089*** (0.013)
Borrowing	2 (2.94%)	55 (80.88%)	11 (16.18%)	6 (8.82%)	46 (67.65%)	16 (23.53%)	0.024* (0.014)
Saving	2 (2.50%)	46 (57.50%)	32 (40.00%)	5 (6.25%)	37 (46.25%)	38 (47.50%)	0.098*** (0.017)
Retirement Saving	0 (0.00%)	19 (47.50%)	21 (52.50%)	0 (0.00%)	15 (37.50%)	25 (62.50%)	0.105** (0.039)
Budgeting & planning	0 (0.00%)	14 (40.00%)	21 (60.00%)	1 (2.86%)	9 (25.71%)	25 (71.43%)	0.223** (0.055)
Insurance	0 (0.00%)	13 (81.25%)	3 (18.75%)	0 (0.00%)	12 (75.00%)	4 (25.00%)	0.045 (0.045)
Open bank account	0 (0.00%)	8 (80.00%)	2 (20.00%)	0 (0.00%)	8 (80.00%)	2 (20.00%)	0.025 (0.022)
Remittance	1 (11.11%)	8 (88.89%)	0 (0.00%)	1 (11.11%)	8 (88.89%)	0 (0.00%)	-0.045 (0.038)

Notes: Average effect sizes are estimated via OLS with standard errors clustered at the study-level in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table A2: Bivariate relationships between study descriptors and effect size on fin. behavior

	(1)	(2)	(3)	(4)	(5)	(6)
	A Methodological Characteristics					
RCT	-0.005					
	(0.028)					
TOT		0.063				
		(0.039)				
Delay			-0.000***			
			(0.000)			
1/SE				-0.000***		
				(0.000)		
	B Intensity of education					
Intensity	-0.001					
	(0.001)					
Duration		-0.001				
		(0.002)				
Intensity per week			0.004*			
			(0.002)			
	C Target group of education					
Age	0.002**					
	(0.001)					
Percent female		-0.000				
		(0.001)				
Low income clients			-0.070**			
			(0.030)			
Years of schooling				-0.003		
				(0.003)		
FL in population					-0.001	
					(0.001)	
Country group						
a,) Low-income econ.						0.059
						(0.068)
b.) Lower-middle inc. econ.						0.014
						(0.021)
c.) Upper-middle inc. econ.						0.027
						(0.046)
	D Characteristics of education					
Channel						
a.) Classroom	0.044					
	(0.031)					
b.) Counseling	0.045**					
	(0.018)					
c.) Online	0.063**					
	(0.031)					
Mandatory		-0.087***				
		(0.016)				
Incentivized			0.033			
			(0.044)			
Setting						
a) School				-0.031		
				(0.022)		
b) Workplace				0.040		
				(0.056)		
Teachable moment					0.062**	
					(0.024)	

Notes: Results from OLS regressions. Intercepts and numbers of observations are not reported. Dependent variable is effect size (Hedges' g) on financial behavior. Robust standard errors clustered at the study-level in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level.

A2 Bivariate relations between financial behavior and intervention characteristics

This additional section presents exploratory results on bivariate relations between intervention characteristics and effect size on financial behavior. The discussion follows the same order as these variables have been introduced in the Section 4.1, i.e. we present results in four Panels in Table A2.

<Table A2 about here>

In *Panel A* we introduce variables that capture the methodological characteristics of the study. We find a 0.005 standard deviation unit difference in effect sizes between RCTs and weaker study designs which is statistically insignificant (column 1). However, studies operationalizing effect sizes as TOT estimates, do report effect sizes 0.063 standard deviation units higher than more conservative ITT estimates. Additionally, column 3 shows that delay in measurement of financial behavior does significantly affect effect sizes – while this effect is small with a one-week delay leading only to a increase of less than 0.0001 standard deviation units decrease in effect sizes – this relationship is highly significant. Finally, we also observe that the precision of the effect size estimate (measured by its inverse standard error) is important in the sense that estimates with smaller standard errors are associated with smaller effect sizes and vice versa. Thus, precision and sample size as methodological design characteristics are important controls in our regression analyses.

Panel B shows the relationship between intensity and effect size. Surprisingly, we find that a linear simple effect of the intensity of intervention is statistically insignificant. Actually one would expect that teaching more hours should lead to higher effect sizes. Fernandes et al. (2014) do indeed find a small but positive and significant relationship between overall intensity and effect sizes for a sample of 29 papers. However, Miller et al. (2015) fail to establish a significant link between intensity and effect size, although their results suffer from low statistical power due to the analysis of at most five studies at a time. Taking the duration

of the intervention into account, however, we find that once we measure intensity relative to the duration – we arrive at positive and marginal significant linear effects for the relationship between weekly intensity and effect sizes. Thus, an increase of one additional hour of financial education per week leads a 0.004 standard deviation units increase in effect sizes on financial behavior.

Panel C shows the bivariate relationships between the characteristics of the participants targeted by the financial education intervention and its effect sizes on financial behavior. Contrary to the hypothesis in the Section 4.1 the mean age of the participants is a statistically significant predictor of effect sizes from financial education in the sense that effect sizes seem to increase with the age of the target group (column 1). In contrast, samples the percentage share female participants (column 2) seem to be unrelated to effect sizes – indicating no differences between men and women seem to exist. Column 3 shows the dummy variable for low-income participants (relative to the country mean), which yields a negative and statistically significant coefficient as expected. However, column (4) shows that there is a insignificant relationship between the overall affluence (measured by population mean years of schooling) of the participants' country and effect size. Finally, considering the baseline financial literacy of the target groups population at the country-level, we find that there is indeed a negative relationship between higher rates of financial literacy and impacts achieved by additional financial education efforts, however this estimate is insignificant in these bivariate comparisons.

The effect of financial education may vary across the channel through which financial education is delivered. Thus, we analyze this potential source of heterogeneity in *Panel D*. Comparing a) classroom seminars, b) online courses, c) individual counselling, and d) other formats (omitted category) we find that classroom scenarios and other formats (omitted category) lead to smaller effect sizes compared to individualized counselling and online

courses. This result is intuitive for individualized counseling since the trainer to participant ratio is 1:1 and thus resembles a much costlier intervention than classroom trainings. Also it can be argued that the intensity of financial education per individual is higher in this format. However, we also find for 20 papers reporting effects of interventions using classroom-based seminars, that class size is no predictor of effect sizes on behavior (b=0.000, p>0.1). Thus it may be the case that these differences in effect sizes may be explained rather by differences in pedagogy than differences in relative input.

Moreover, our results show that there is a significant bivariate difference between mandatory or voluntary formats of financial education. It seems that those who participate voluntary benefit from higher effect sizes on financial behavior compared to those for whom participation is mandatory. This observation may be explained by differences in motivation and willingness to change individual behavior. This argument could be supported by the observation that online courses are more effective than classroom formats, too – as these formats rely heavily on self-regulated learning and require higher intrinsic motivation. Also incentivizing participation in financial education interventions does not significantly affect its outcomes, indicating that intrinsic motivation of the participants may indeed be one of the crucial factors in determining intervention success.

We also find that the timing of financial education seems to be of importance: When coding whether financial education is offered at a teachable moment (i.e. a situation where participants could immediately apply their acquired knowledge) this feature of financial education is a significant predictor of effect sizes across 82 studies reporting 260 effect sizes. In contrast the concrete setting, i.e. teaching at school, workplace or other settings (omitted category) appears to be less important in bivariate models. However, we note that the coefficient for schools is negative and workplace financial education appears to be associated

with higher impacts. This qualitative observation mirrors the impression that the timing and voluntariness of an intervention may be important.

Overall, these bivariate relations indicate that several characteristics of financial education and methodological rigor in assessing its outcomes are relevant for the measured impact, whereas other characteristics seem less important.

Table A3: Bivariate relationships between study descriptors and effect size on fin. literacy

	(1)	(2)	(3)	(4)	(5)	(6)
	A Methodological Characteristics					
RCT	0.266***					
	(0.092)					
TOT		0.264***				
		(0.059)				
Delay			0.002***			
			(0.001)			
1/SE				0.010***		
				(0.003)		
	B Intensity of education					
Intensity	-0.000					
	(0.002)					
Duration		0.000				
		(0.003)				
Intensity per week			0.004			
			(0.003)			
	C Target group of education					
Age	-0.000**					
	(0.000)					
Percent female		0.003				
		(0.003)				
Low income clients			0.062			
			(0.086)			
Years of schooling				0.030***		
				(0.011)		
FL in population					0.006***	
					(0.002)	
Country group						
a.) Low-inc. econ.						-0.125
						(0.089)
b.) Lower-middle inc. econ.						-0.205**
						(0.080)
c.) Upper-middle inc. econ.						-0.261***
						(0.070)
	D Characteristics of education					
Channel						
a.) Classroom	0.213***					
	(0.068)					
b.) Online	0.127**					
	(0.057)					
Mandatory		0.248**				
		(0.115)				
Incentivized			0.050			
			(0.096)			
Setting						
a) School				0.194**		
				(0.095)		
b) Workplace				-0.057		
				(0.042)		
Teachable moment					-0.059	
					(0.090)	

Notes: Results from OLS regressions. Intercepts and numbers of observations are not reported. Dependent variable is effect size (Hedges' g) on financial literacy. Robust standard errors clustered at the study-level in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table A4: Multivariate relationships between study descriptors and effect size on fin. literacy

	(1)	(2)	(3)	(4)	(5)
RCT	-0.288***	-0.287***	-0.122	-0.148	-0.148*
	(0.087)	(0.086)	(0.100)	(0.135)	(0.088)
TOT	0.002	-0.013	-0.084	-0.063	-0.076
	(0.089)	(0.096)	(0.081)	(0.086)	(0.094)
Delay	-0.002***	-0.003***	-0.003***	-0.002***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
1/SE	-0.011***	-0.012***	-0.016***	-0.016***	-0.014***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Intensity/week		0.004	0.014***	0.014***	0.012***
		(0.002)	(0.004)	(0.003)	(0.003)
Duration		0.005	0.006	0.007	0.005
		(0.004)	(0.005)	(0.005)	(0.004)
Age			-0.008***	-0.008***	-0.007***
			(0.003)	(0.003)	(0.002)
Percent female			0.001	0.001	
			(0.003)	(0.003)	
Low income			-0.025	-0.011	
			(0.061)	(0.090)	
Years of schooling			-0.003	-0.000	
			(0.027)	(0.029)	
FL in population			-0.006	-0.006	
			(0.007)	(0.008)	
Country group					
Low/lower-middle inc. econ.			-0.339	-0.292	-0.138**
			(0.243)	(0.240)	(0.056)
Upper-middle inc. econ.			-0.179	-0.188	-0.018
			(0.235)	(0.304)	(0.095)
Channel					
a.) Classroom				-0.062	
				(0.180)	
b.) Counseling					
c.) Online				-0.028	
				(0.145)	
Mandatory				-0.032	
				(0.149)	
Incentivized				-0.057	
				(0.100)	
Teachable moment				-0.017	
				(0.110)	
Intercept	0.703***	0.692***	1.233***	1.278***	0.881***
	(0.159)	(0.157)	(0.449)	(0.462)	(0.173)
R^2	0.310	0.320	0.376	0.380	0.367
n (studies)	61	61	61	61	61
n (effect sizes)	169	169	169	169	169

Notes: Results from OLS meta-regression with robust standard errors clustered at the study-level. Dependent variable is effect size (Hedges' g) on financial literacy. Standard errors in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level.

Appendix B: Publication bias in the financial education literature

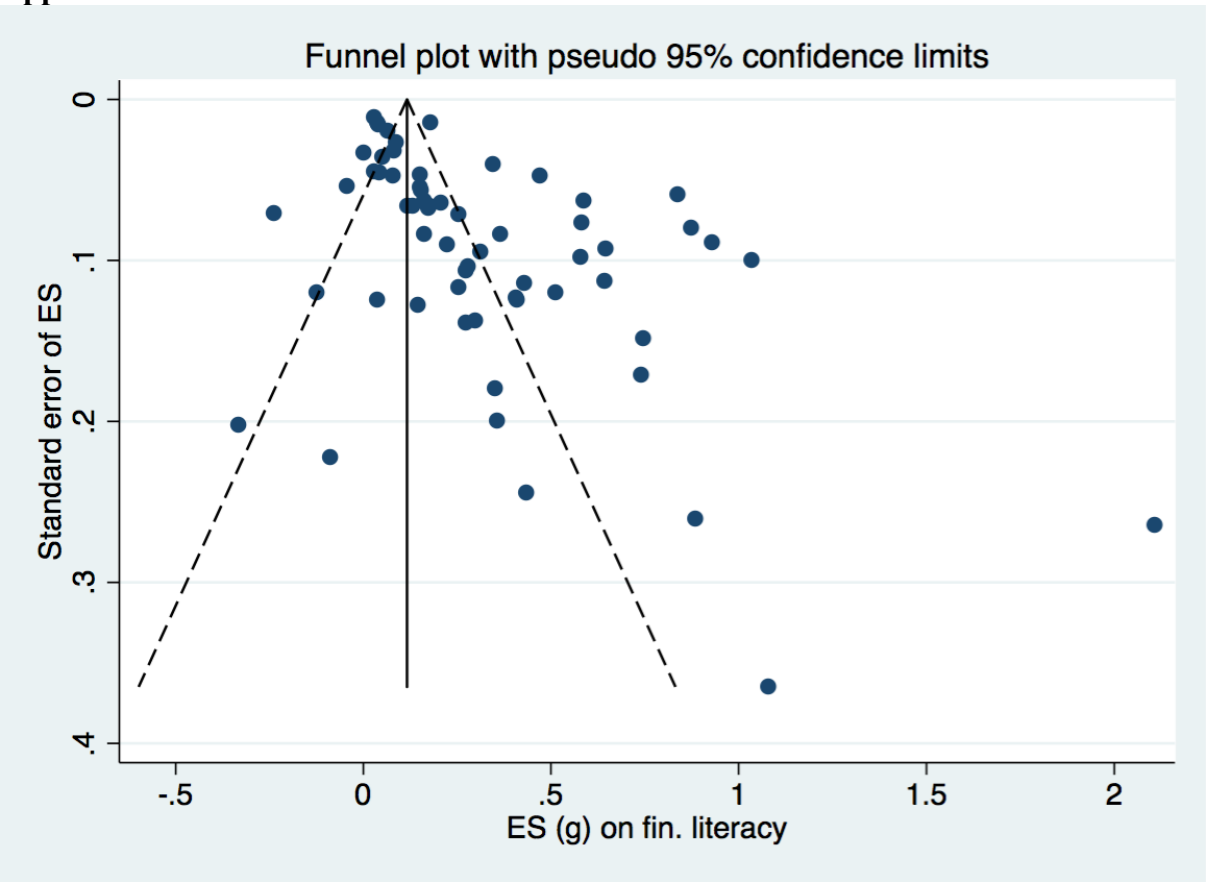

Figure B1: Funnel plot of treatment effects on financial literacy

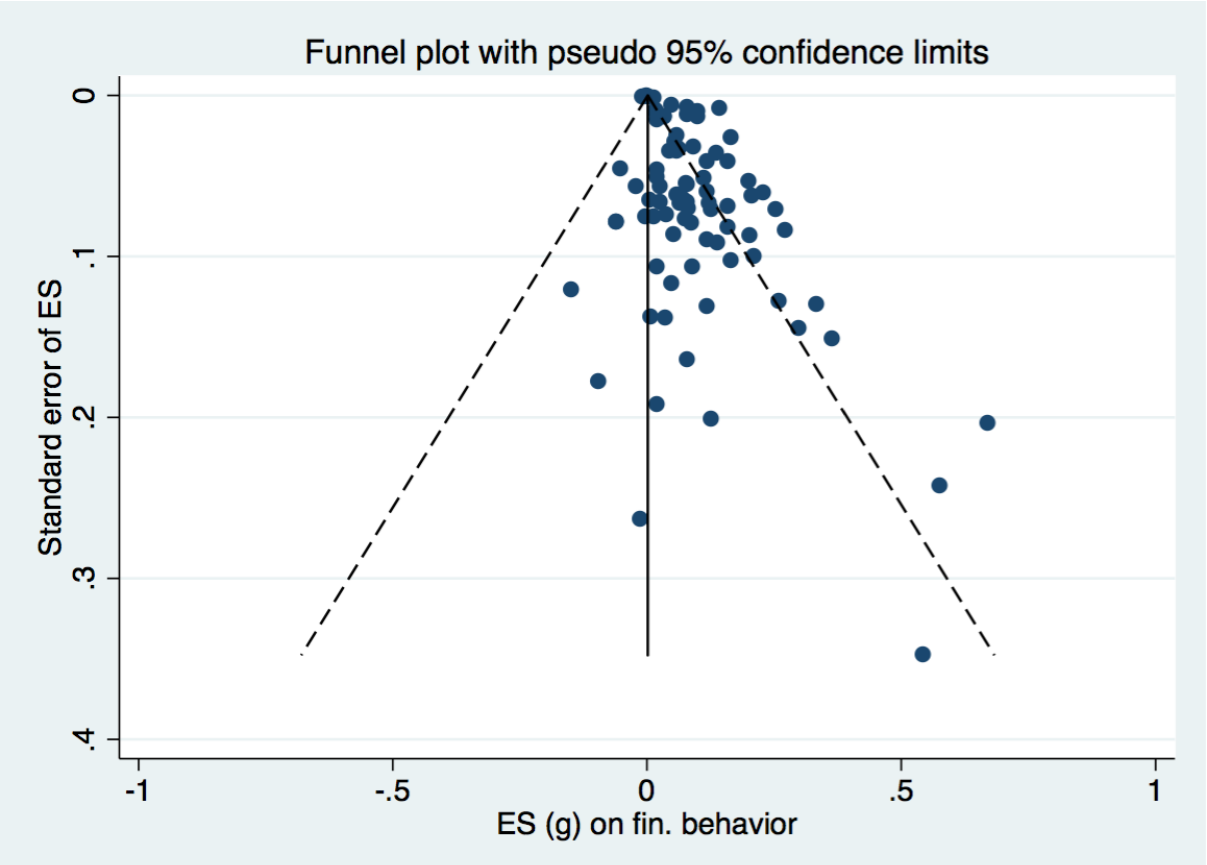

Figure B2: Funnel plot of treatment effects on financial behavior

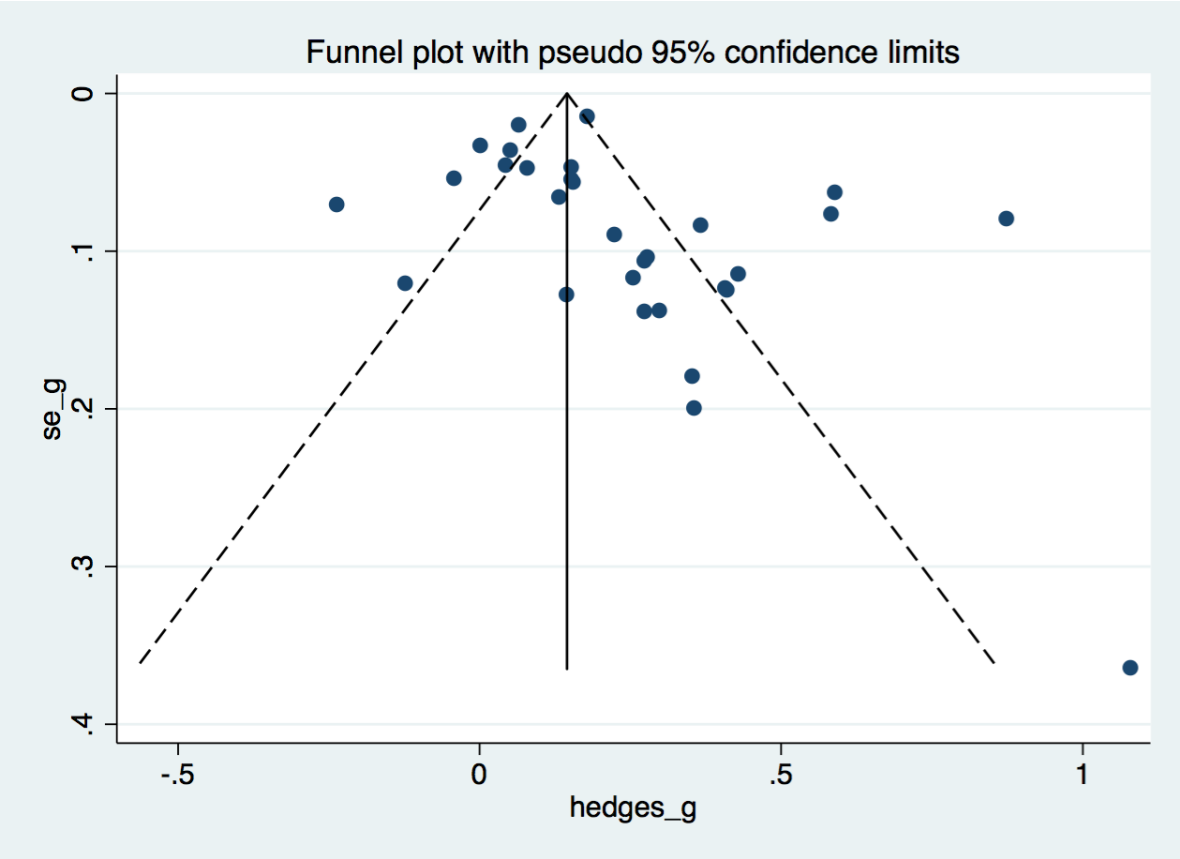

Figure B3: Funnel plot of treatment effects on financial literacy within the subsample of RCTs only

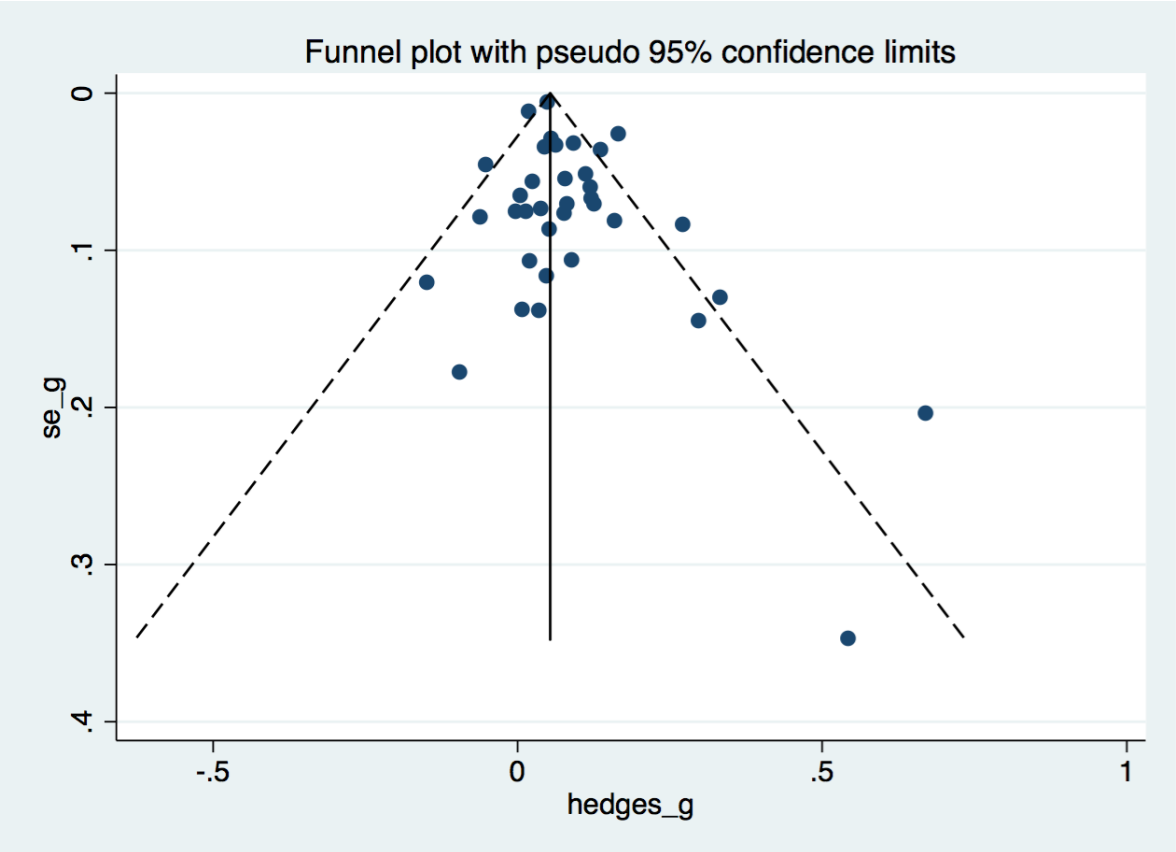

Figure A4: Funnel plot of treatment effects on financial behavior within the subsample of RCTs only

We conduct visual tests for overall publication bias, so-called funnel plots. (cf. Figures B1 and B2 in Appendix B). Precision of the estimated treatment effect should increase in larger studies. Thus, we scatter effect sizes (one average effect per study) against a proxy measure of study size (in our case standard errors of the effect size estimates). Effect estimates from small studies (larger sampling errors) should scatter more widely at the bottom of the graph, with the spread decreasing as standard errors decrease. In the absence of bias the plot resembles a symmetrical inverted funnel. Therefore, asymmetry indicates a publication bias in the sense that negative or non-results are under-represented (i.e. not published at all). Inspecting the two plots indicates that symmetry is higher for effect sizes on financial behavior than for effect sizes on literacy but both outcomes may be affected by publication biases in the sense that the overall treatment effect may suffer from a slight upward bias. This conclusion, however, requires the strong assumption that non-results are not published at all (i.e. the file drawer problem). This assumption may be more plausible for quasi- and natural experiments than for RCTs, as results from rigorous randomized experiments are likely to be published irrespective of their results. Therefore, we perform the same visual check on the subsample of RCTs only (cf. Figures B3 and B4 in Appendix B). Indeed, these plots are much more symmetric indicating that publication bias may only be an issue within the sample of non-randomized studies. As a) nearly 40 percent of our sample is comprised of RCTs, b) we control for research design in all of our regressions, and c) our main results replicate within the subsample of RCTs (cf. Section 7) we conclude that publication biases are not an issue for our analysis. However, we also test the robustness of our results using weighted least squares which is in principle a robust method in the presence of publication selection.

Appendix C: Robustness tests

Table C1: Robustness - Sample splits, missing data and operationalization of effect sizes

	(1) RCTs only	(2) No Imputations	(3) Dev. econ. split	(4) High inc. split	(5) Synthetic ES OLS
RCT		-0.086		-0.152***	-0.091**
		(0.053)		(0.031)	(0.037)
TOT	0.033	0.126**	-0.039	0.144***	0.102***
	(0.043)	(0.056)	(0.026)	(0.033)	(0.030)
Delay	-0.001	-0.000	0.001	0.000**	0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
1/SE	0.001	0.001*	0.001	-0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)
Intensity /week	0.005***	0.004**	0.006***	0.008**	0.003
	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)
Duration	-0.005**	-0.001	-0.001	-0.002**	-0.000
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Low inc. clients	-0.005**	-0.101**	-0.049*	0.007	-0.015
	(0.002)	(0.038)	(0.002)	(0.031)	(0.029)
FL in pop.	-0.010***	-0.014***	-0.006**	-0.018***	-0.012***
	(0.003)	(0.003)	(0.002)	(0.005)	(0.003)
Country group					
Low/lower-mid.	-0.303***	-0.391***			-0.339***
inc. econ.	(0.105)	(0.125)			(0.084)
Upper-mid. inc..	-0.185**	-0.195*			-0.186**
econ.	(0.089)	(0.103)			(0.079)
Mandatory	-0.054	-0.107**		-0.099***	-0.095**
	(0.058)	(0.048)		(0.031)	(0.041)
Teach. moment	0.083**	0.089*	-0.008	0.113***	0.088***
	(0.036)	(0.049)	(0.019)	(0.029)	(0.030)
Constant	0.635***	0.921***	0.246**	1.034***	0.755***
	(0.182)	(0.206)	(0.090)	(0.262)	(0.157)
R^2	0.177	0.195	0.076	0.361	0.308
n (Studies)	36	27	20	62	82
n (Effect Sizes)	161	147	120	140	82

Notes: Column (1) presents results from OLS meta-regressions for the subsample of RCTs only. Column (2) reports results for complete cases only. Columns (3) and (4) present results for the sample splits for developing and developed countries respectively. These splits include only variables for which differential information from at least two studies are available.Column (5) presents results using one synthetic effect size (average effect size across all outcomes) per study. Standard errors clustered at the study-level in parentheses (except Column 5). ***, ** and * denote significance at the 1%, 5% and 10% level.

Table C2: Robustness – Alternative meta-regression models

	(1) Probit 5%	(2) Ordered probit 10%	(3) RE GLS	(4) WLS 1/SE weights
RCT	-0.794**	-0.914***	-0.105***	-0.067***
	(0.334)	(0.301)	(0.034)	(0.024)
TOT	-0.028	0.017	0.078***	0.019
	(0.262)	(0.240)	(0.027)	(0.013)
Delay	-0.001**	-0.000***	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
1/SE	-0.000	-0.002***	-0.000	-0.000***
	(0.000)	(0.001)	(0.000)	(0.000)
Intensity /week	0.029**	0.041**	0.004**	0.005***
	(0.014)	(0.018)	(0.002)	(0.001)
Duration	0.014*	0.003	-0.001	0.001*
	(0.009)	(0.011)	(0.001)	(0.000)
Low inc. clients	-0.640***	-0.525***	-0.001	-0.036**
	(0.243)	(0.203)	(0.001)	(0.014)
FL in pop.	-0.049**	-0.048**	-0.011***	-0.008***
	(0.022)	(0.019)	(0.002)	(0.002)
Country group				
Low/lower-mid.	-1.459**	-1.394**	-0.309***	-0.214***
inc. econ.	(0.734)	(0.674)	(0.083)	(0.061)
Upper-mid. inc..	-0.648	-0.533	-0.175**	-0.111**
econ.	(0.671)	(0.565)	(0.074)	(0.054)
Mandatory	0.336	0.402	-0.052	-0.034**
	(0.338)	(0.281)	(0.036)	(0.015)
Teach. moment	0.676**	0.647***	0.117***	0.082***
	(0.283)	(0.240)	(0.029)	(0.020)
Constant cut 1		-4.967***		
		(1.106)		
Constant cut 2		-2.873***		
		(1.070)		
Constant	2.658**		0.702***	0.507***
	(1.259)		(0.140)	(0.099)
R^2			0.217	0.314
Pseudo R^2	0.119	0.114		
n (Studies)	82	82	82	82
n (Effect Sizes)	260	260	260	260

Notes: Dependent variable in columns (1) and (2) is a categorical indicator sign and significance of intervention impact. Dependent variable in columns (3) and (4) is effect size (Hedges' g) on financial behavior. Column (1) reports results from probit-regression with a binary outcome indicating whether financial education had a significantly positive effect on financial behavior at the 5%-level. Column (2) provides results for ordered probit regression with a dependent categorical variable taking the value "-1" if financial education had a significantly negative impact on financial behavior, "0" if financial education had an insignificant effect on financial behavior, and "1" if financial education had a significant positive effect on financial behavior at the 10%-level. Column (3) reports results from GLS random-effects regression. Column (4) reports results of weighted least squares estimation with inverse variance weights. Standard errors clustered at the study-level in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table C3: Random effects meta-regression on synthetic effect sizes with inverse variance weights

	(1) MM	(2) MM	(3) REML	(4) REML
RCT	-0.066**	-0.073**	-0.062**	-0.066**
	(0.030)	(0.030)	(0.026)	(0.027)
TOT	0.032	0.031	0.025	0.024
	(0.025)	(0.025)	(0.022)	(0.021)
Delay	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Intensity /week	0.003	0.004	0.003	0.004*
	(0.002)	(0.002)	(0.002)	(0.002)
Duration	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Low inc. clients	-0.016	0.003	-0.016	0.002
	(0.022)	(0.024)	(0.019)	(0.021)
FL in pop.	-0.009***	-0.011***	-0.008***	-0.011***
	(0.002)	(0.003)	(0.002)	(0.003)
Country group				
Low/lower-mid.	-0.227***	-0.594**	-0.225***	-0.577***
inc. econ.	(0.074)	(0.232)	(0.067)	(0.209)
Upper-mid. inc.	-0.127*	-0.002	-0.125**	0.021
econ.	(0.066)	(0.276)	(0.060)	(0.244)
Low/lower inc. econ		0.011*		0.010*
× Fl in pop.		(0.006)		(0.005)
Upper-mid. inc. econ.		-0.005		-0.006
× Fl in pop.		(0.007)		(0.006)
Mandatory	-0.047*	-0.053*	-0.043*	-0.049*
	(0.027)	(0.027)	(0.022)	(0.022)
Teach. moment	0.066***	0.060**	0.064***	0.056***
	(0.024)	(0.024)	(0.021)	(0.021)
Constant	0.543***	0.667***	0.533***	0.647***
	(0.140)	(0.181)	(0.127)	(0.167)
I^2	84.85%	84.14%	84.85%	84.14%
n (Studies)	82	82	82	82
n (Effect Sizes)	82	82	82	82

Notes: Results from random-effects meta-regression (DerSimonian and Larid 1986) with Knapp and Hartung (2003) adjusted standard errors. Dependent variable is effect size (Hedges' g) on financial behavior weighted by its inverse variance. Columns (1) and (2) show results for method of moments (MM) estimates. Columns (3) and (4) show results for restricted maximum likelihood (REML) estimates. Standard errors in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table C4: Effect sizes on financial behavior and delay in measurement of treatment effects

	(1) FB	(2) Short term Split	(3) Medium term split	(4) Long term split
RCT	-0.082**	0.150	-0.126*	-0.099***
	(0.035)	(0.089)	(0.067)	(0.026)
TOT	0.039	-0.259***	0.042	0.130***
	(0.029)	(0.063)	(0.050)	(0.047)
Short term	0.120**			
	(0.058)			
Medium term	-0.013			
	(0.014)			
1/SE	-0.000	-0.005**	0.000	-0.000**
	(0.000)	(0.002)	(0.000)	(0.000)
Intensity /week	0.005***	-0.054	0.004***	0.004***
	(0.001)	(0.033)	(0.001)	(0.001)
Duration	-0.001	-0.134***	-0.001	0.004*
	(0.001)	(0.024)	(0.001)	(0.002)
Low inc. clients	-0.048**	-0.366***	-0.044**	-0.012
	(0.020)	(0.113)	(0.019)	(0.028)
FL in pop.	-0.010***	0.029**	-0.007***	-0.005
	(0.002)	(0.012)	(0.002)	(0.003)
Country group				
Low/lower-mid.	-0.278***	0.230	-0.148	-0.142**
inc. econ.	(0.082)	(0.508)	(0.098)	(0.066)
Upper-mid. inc..	-0.136*		-0.054	0.110
econ.	(0.070)		(0.073)	(0.092)
Mandatory	-0.035	-0.548***	-0.037	-0.010
	(0.025)	(0.058)	(0.046)	(0.032)
Teach. moment	0.128***	0.259**	0.124**	0.119***
	(0.029)	(0.103)	(0.052)	(0.025)
Constant	0.611***	-0.306	0.439***	0.280
	-0.035	(0.560)	(0.119)	(0.177)
R^2	0.249	0.680	0.091	0.613
n (Studies)	82	13	28	42
n (Effect Sizes)	260	22	151	87

Notes: Results from OLS meta-regression with robust standard errors clustered at the study-level. Dependent variable is effect size (Hedges' g) on financial behavior. Standard errors in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level.

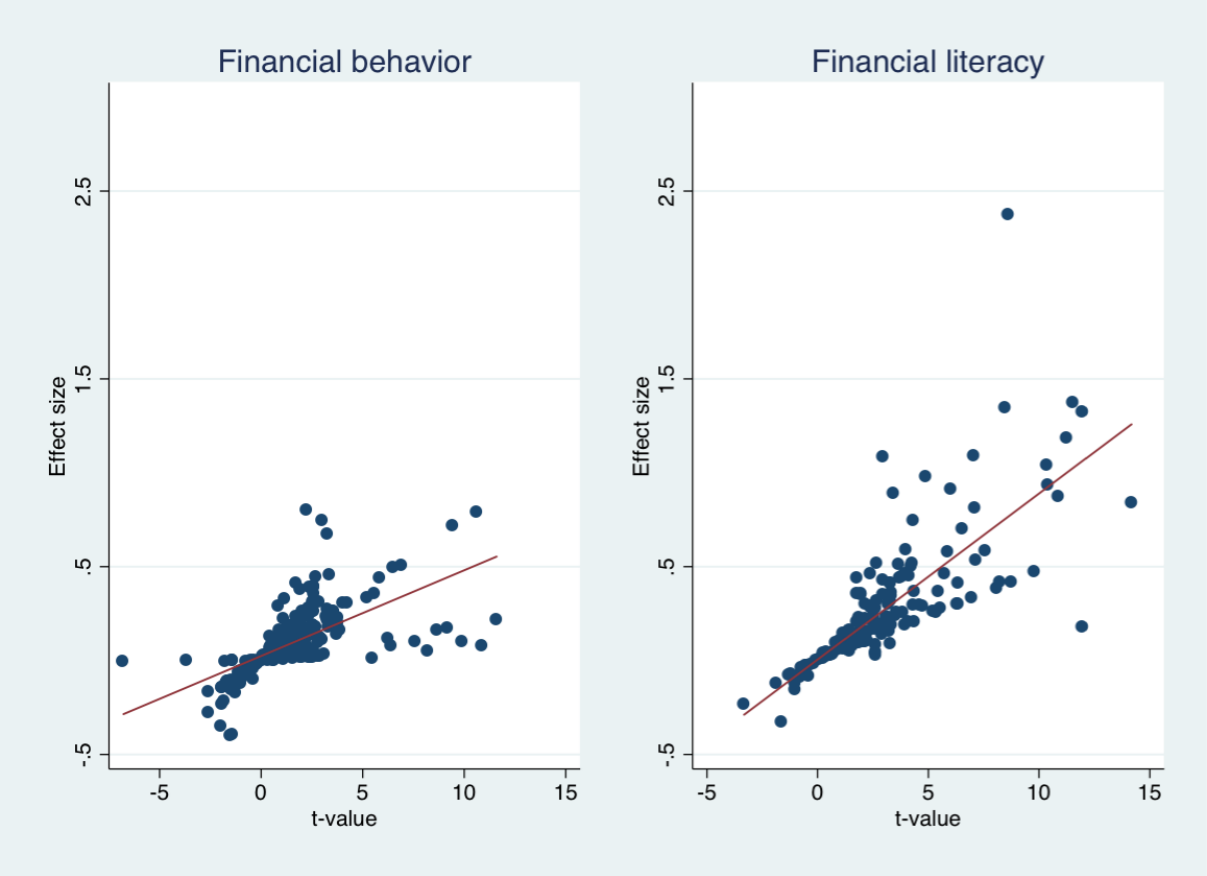

Figure C1: Relationship between statistical significance and effect size for impact estimates on financial behavior and financial literacy

This appendix contains seven kinds of robustness tests : (i) we limit the sample of studies to RCTs only, (ii) then we show results without imputing missing values, (iii) we split the sample by country groups, (iv) we estimate our main model with only one synthetic effect size per study, (v) we estimate alternative meta-regression models, (vi) we look for heterogeneous impacts depending on the delay in measurement of outcomes and finally (vii) we test a different operationalization of training intensity.

<Table C1 about here>

RCTs only. As a first robustness check we estimate the OLS meta-regression models for a reduced sample of studies including only RCTs. This tests whether our results are driven by misspecification bias in non-randomized studies. Column 1 of Table C1 reports results for this subsample of 36 financial education studies. We find that the magnitude and direction of the estimated coefficients and their standard errors for our main explanatory variables correspond with the full sample and our main results entirely replicate under this approach.

Conservative handling of missing data. Next, we turn to estimations of complete cases only, in order to test the robustness of our results using imputed default categories or mean values for missing observations. Column 2 in Table C1 reports OLS meta-regression results for complete cases only. These results again correspond to the results presented in Table 5 but show larger standard errors for some of the variables, however, turning none of the main explanatory variables insignificant This result strongly supports the conclusions drawn from estimations with a large number of studies in the sample.

Splits by country group. Next, we split our sample of studies into studies reporting on interventions in developing countries (Column 3) and studies reporting on interventions in high income economies (Column 4). We include only variables into the regressions where at least more than one study contributes to each of the subgroups. Thus, we exclude the variables “RCT” and “mandatory” in the split of interventions in developing countries, since

only one study in this subsample is a quasi-experiment and only one study reports on effects from mandatory financial education. We include all of the relevant variables in the high income split, because there is sufficient variation for all of them.

Turning to the regression results we note that most of our main results replicate well for both splits, including significant estimates for intensity and smaller effect sizes with increasing baseline financial literacy in the population. Two differences are noteworthy however, first the coefficient for a teachable moment appears to be insignificant in the developing country subsample, indicating that the small effect sizes reported in developing country interventions may be especially difficult to address even with well-timed interventions. However, caution in the presence of low sample size per subgroup is advised. Second, the coefficient for low-income clients is now insignificant in the high income economies split, indicating that this result may also be driven by weak results of interventions in developing countries targeting the poor. Overall we conclude from these brief checks, that our overall conclusion of weak effects in developing countries seems to be underscored. It also appears to be difficult to increase these effect by modifying financial education program details other than increased input in form of more effective teaching hours.

One synthetic effect size per study. Much of the meta-analysis literature in other fields than economics uses effect size models where each study contributes only one synthetic effect size to the meta-regression analysis. There are different options to provide such a single effect. Some suggest only using the most robust results in a primary study (cf. Cho and Honorati 2014, p. 119). The textbook literature on meta-analysis, however, tends to recommend creating a synthetic effect size per study by using the average effect across multiple outcomes (cf. Lipsey and Wilson 2001).

We follow this approach here for the purpose of robustness exercises, but we point at the major disadvantage that effects heading in opposite directions within one study may be

cancelling each other out. Column 5 of Table C1 shows results for such an approach. The signs and magnitudes of our coefficients are very similar to the model with multiple non-synthetic effect sizes per study and standard errors clustered at the study-level. However, in the estimation based on this sample, the standard errors increase which leads to insignificant covariates in two cases: intensity per week and low-income clients. Since this approach works with much less information than would be available we conclude that qualitatively this check also confirms our main findings derived from the larger sample of available effect sizes.

Regression approaches. Here we discuss the use of alternative statistical regression models in the estimation of predictors of intervention impact. In column 1 of Table C2 we apply a probit-regression on an indicator variable of statistically significant effect estimates (at the 5%-level). This is a departure from earlier analyses because we now neglect the size of effects but only consider their statistical significance. This approach is meaningful when the estimated effect sizes and their t-values are correlated (i.e. not driven by differences in sample size alone) (cf. Figure C1).

Following the approach applied by Card et al. (2010) and Cho and Honorati (2014) we code the sign and significance for each impact estimate reported in the primary studies. This indicator of intervention success has the advantage that it is easily interpretable and neutral to the unit of the outcome variable. However, it only captures the direction and significance of an effect, unlike the standardized mean difference which preserves its magnitude (cf. Stanley and Doucouliagos 2012, p. 6). Using this approach, we construct a binary dependent variable taking the value 0 if the primary study impact estimate t-statistic is smaller than 1.96 and taking the value 1 if $t \geq 1.96$. Additionally, we extend this approach and construct an ordered categorical variable which can take three values of -1 if $t \leq -1.64$, 0 if $t \geq -1.64$ and $t \leq 1.64$, and 1 if $t \geq 1.64$. Thus, we distinguish between significant negative, insignificant, and

significant positive estimates at the 10%-level because there are hardly negative estimates at the 5%-level (see Table A1 in the Appendix).

We observe that the sign and significance of the logged odds correspond with the model using a continuous measure of effect size reported in Table 4, column 5. However, estimated standard errors differ slightly, as the coefficients for TOT and mandatory are now insignificant – probably resulting from reduced variance in the dependent variable in comparison to the use of continuous effect sizes.

<Table C2 about here>

In column 2 we extend this approach and estimate an ordered probit model where the dependent variable consists of three ordered categories which distinguish between significant negative, insignificant and significant positive estimates at the 10%-level of financial education impact. This leads to a very similar assessment of predictor sign and magnitude as in our benchmark model in Table 5, column 5 but again slightly different estimates for the standard errors.

So far, we have discussed models where we assume that heterogeneity can be explained by observable characteristics. However, if one assumes that the between-study heterogeneity cannot readily be explained by the observable characteristics included, $x_{ij}\beta$, one has to incorporate unobservable characteristics through random effects into the model (cf. Cho and Honorati 2014). Thus, including an effect capturing unobservable characteristics of the study, the meta-analytic model is defined as:

$$g_{ij} = x_{ij}\beta + \theta_{ij} + \epsilon_{ij} \qquad (5)$$

where g_{ij} is the impact (continuous effect size) of a financial education intervention on outcome *i* reported in study *j*, $x_{ij}\beta$ is a vector of observable covariates, θ_{ij} is a random effect of unobservable study characteristics and ϵ_{ij} is an error term independent of $x_{ij}\beta$ and θ_{ij}. The results in column 3 show coefficients from a GLS random effects regression based on the

assumptions discussed in equation 5. This estimation almost entirely matches the results of the benchmark model shown in Table 4, column 5 with the expected increased standard errors for low-income clients and mandatory financial education.

Next, we turn to an alternative unrestricted weighted least squares approach. In column 4 we weight each effect size with its inverse standard error (1/SE). The results show that our results, again, largely match the results of the ordinary least squares estimations. Overall, the pattern in sign and magnitude (including most standard errors) of our main explanatory variables can be confirmed under various statistical models.

Random effects meta-regression with weighted effect sizes. Finally, we check whether weighting each synthetic effect size estimated by its precision and controlling for unobservables affects our results. This is common in meta-analyses in other disciplines (such as clinical trials) and thus serves as a further check of the sensitivity of our results to the estimation strategy. This approach assigns weights for each study based on the inverse variance of the within study measurement error plus the between study variance (tau squared) ($w_i = \frac{1}{V_{Yi}+\tau^2}$). Thus we define our meta-analytic model as

$$g_i = x_i\beta + u_i + \epsilon_i \quad (6)$$

where

$$u_i \sim N(0,\ \tau^{\ 2}) \quad (7)$$

and

$$\epsilon_i \sim N(0, \sigma_i^2) \quad (8)$$

Here g_i is defined as the effect size estimate of study *i*, $\sigma_i$is the corresponding standard error, $\tau^2$is the between study variance in true effects, and $x_i\beta$ is a vector of study level covariates (including an intercept). We estimate this model using either method of moments (DerSimonian and Laird 1986) or alternatively restricted maximum likelihood.

Table 8 shows our preferred specification with and without the interaction effects between population financial literacy and country-income indicators for the two different estimators of these random-effects meta-regression models. Column 1 reports results without these interactions using method of moments. Here we find that all of our results are similar with the exception of increased standard errors, especially for low-income clients and intensity per week which are now statistically insignificant. The results are basically unchanged when interaction terms between population financial literacy and country-income indicators are introduced (column 2). Turning to an alternative estimator (restricted maximum likelihood) we find that these results are again nearly identical with the exception that intensity per week is now again marginally significant in presence of the interaction effects (column 4). Overall of our main results replicate well using either of these approaches, including the interaction-effects of interest.

<Table C3 about here>

Heterogeneous impacts depending on delay in measurement. In order to check for heterogeneous impacts depending on the considered time-frame we conduct two tests. First, we model the relationship between delay in measurement and effect size on financial behavior outcomes in a non-linear fashion by creating a categorical variable that distinguishes between short term (less than one month, approx. 10% of estimates), medium term (less than one year, approx. 68% of estimates) and long-term (longer than one year, approx. 22% of estimates) effects on financial behavior. Column 1 of Table A4 shows that short term effects tend to be higher than medium- or long term effects on financial behavior which is in line with the present literature (cf. Fernandes et al. 2014; Lusardi et al. 2015b). Splitting the sample according to these three time-frames we do not find that any predictor is strikingly different compared to the aggregate sample with the exception of the split comprising short term

effects reported in 13 studies only, where signs and significance of predictors change – but this estimation suffers from degrees of freedom problems.

Intensity. Since the intensity of financial education supports its effectiveness, we check which aspect of intensity of education drives our results. Using only the total number of hours taught as a linear predictor of effect size (and neglecting the duration of the intervention), we find that intensity does not predict effect sizes on financial behavior (available on request). This result remains the same in several variants of variable and model specifications (e.g. including polynomial forms of intensity, interaction effects between delay and intensity, and centering) and holds when effect sizes on financial literacy are regressed on this linear predictor. Thus, the intensity relative to the duration of the intervention appears to matter most for the impact on financial behavior. This finding seems to have practical implications, since it favors education with higher relative intensity, i.e. trainings with relatively more hours per week.

Appendix D: Overview of studies included in the statistical meta-analysis

Study	Country	Research design	Target group		Intevention	
			Mean age	*Low-income*	*Channel*	*Teach. moment*
Agarwal et al. 2009	USA	Natural exp.	-	Yes	Counseling	Yes
Agarwal et al. 2010	USA	Natural exp.	-	Yes	Counseling	Yes
Ambuhel et al. 2014	USA	RCT	29	Yes	Online	No
Asarta et al. 2014	USA	Quasi exp.	15	-	Classroom	No
Barcellos et al. 2012	USA	RCT	52	No	Online	No
Baron-Donovan et al. 2005	USA	Quasi exp.	44	No	Classroom	No
Barua et al. 2012	Singapore	RCT	37	Yes	Classroom	Yes
Batty et al. 2015	USA	RCT	9	Yes	Classroom	No
Bauer et al. 2011	USA	Quasi exp.	-	Yes	Classroom	No
Bayer et al. 2009	USA	Natural exp.	-	No	Classroom	Yes
Becchetti et al. 2013	Italy	RCT	18	-	Classroom	No
Berg and Zia 2013	South Africa	RCT	20	Yes	Mass Media	No
Bell et al. 2009	USA	Quasi exp.	22	No	Classroom	No
Bernheim and Garrett 2003	USA	Natural exp.	39	No	Classroom	Yes
Bernheim et al. 2001	USA	Natural exp.	40	No	Classroom	No
Berry et al. 2015	Ghana	RCT	11	-	Classroom	Yes
Bjorvatn and Tungodden 2010	Tanzania	RCT	39	-	Classroom	Yes
Brown et al. 2014	USA	Natural exp.	28	-	Classroom	No
Brugiavini et al. 2015	Italy	RCT	23	No	Classroom	No
Bruhn and Zia 2013	Bosnia and Herzegovina	RCT	28	Yes	Classroom	Yes
Bruhn et al. 2014	Mexico	RCT	33	-	Classroom	No

Bruhn et al. 2013	Brazil	RCT	16	Yes	Classroom	No
Butt et al. 2008	USA	Quasi exp.	12	No	Classroom	No
Calderone et al. 2013	India	RCT	45	Yes	Classroom (video)	Yes
Carlin and Robinson 2012	USA	Quasi exp.	16	No	Classroom	No
Carpena et al. 2011	India	RCT	39	Yes	Classroom	Yes
Carpena et al. 2015	India	RCT	39	Yes	Classroom + Counseling	Yes
Chen and Heath 2012	USA	Quasi exp.	9	-	Classroom	No
Choi et al. 2005	USA	Natural exp.	40	No	Classroom	No
Choi et al. 2010	USA	RCT	31	No	Info. nudge	No
Choi et al. 2011	USA	Natural exp.	64	No	Info. nudge	No
Clancy et al. 2001	USA	Natural exp.	36	Yes	Classroom	Yes
Clark et al. 2006	USA	Quasi exp.	54	No	Classroom	Yes
Clark et al. 2015	USA	Quasi exp.	44	No	Online	No
Clark et al. 2014	USA	RCT	35	No	Info. nudge	Yes
Clark et al. 2010	USA	Quasi exp-	57	No	Classroom	Yes
Cole and Shastry 2010	USA	Natural exp.	-	No	Classroom	No
Cole et al. 2013	India	RCT	48	Yes	Counseling	Yes
Cole et al. 2014	USA	Natural exp.	17	Yes	Classroom	No
Cole et al. 2011	Indonesia	RCT	41	Yes	Classroom	Yes
Collins 2013	USA	RCT	39	Yes	Classroom	No
Custers 2011	India	RCT	34	Yes	Classroom	Yes
Danes and Haberman 2004	USA	Quasi exp.	15	No	Classroom	No
Danes et al. 1999	USA	USA	15	No	Classroom	No
De Mel et al. 2011	Sri Lanka	Quasi exp.	41	-	Classroom	Yes

DeLaune et al. 2010	USA	Quasi exp.	18	No	Classroom	No
Ding et al. 2008	USA	Natural exp.	-	Yes	Counseling	Yes
Doi et al. 2014	Indonesia	RCT	44	Yes	Classroom	Yes
Dolvin and Templeton 2006	USA	Quasi exp.	46	No	Classroom	Yes
Drexler et al. 2014	Dominican Republic	RCT	41	Yes	Classroom	Yes
Duflo and Saez 2003	USA	RCT	38	No	Info. nudge	Yes
Elliehausen et al. 2007	USA	Natural exp.	41	No	Counseling	No
ETI 2008	USA	Quasi exp.	14	-	Classroom	No
Field et al. 2010	India	RCT	32	Yes	Classroom	yes
Garman et al. 1999	USA	Quasi exp.	43	No	Classroom	yes
Gaurav et al. 2011	India	RCT	50	Yes	Classroom	Yes
Gibson et al. 2014	New Zealand / Australia	RCT	-	Yes	Classroom	Yes
Gill and Bhattacharya 2015	USA	Quasi exp.	17	Yes	Classroom	No
Gine and Mansuri 2014	Pakistan	RCT	38	Yes	Classroom	Yes
Gine et al. 2013	Kenya	RCT	49	Yes	Edu. materials	Yes
Go et al. 2012	USA	Quasi exp.	9	Yes	Classroom	No
Goda et al. 2014	USA	Quasi exp.	45	No	Info. nudge	No
Goldsmith and Goldsmith 2006	USA	Quasi exp.	19	No	Classroom	No
Grimes et al. 2010	USA	Natural exp.	51	No	Classroom	No
Grinstein-Weiss et al. 2015	USA	Natural exp.	36	Yes	Classroom	Yes
Han et al. 2009	USA	RCT	41	Yes	Classroom	Yes
Hartaska and Gonzalez-Vega 2005	USA	Natural exp.	-	Yes	Counseling	Yes

Hartaska and Gonzalez-Vega 2006	USA	Natural exp.	35	No	Counseling	Yes
Harter and Harter 2009	USA	Quasi exp.	-	Yes	Classroom	No
Harter and Harter 2010	USA	Quasi exp.	17	No	Classroom	No
Haynes et al- 2011	USA	RCT	55	Yes	Online	No
Haynes-Bordas et al. 2008	USA	Quasi exp.	38	Yes	Classroom	Yes
Heinberg et al. 2014	USA	RCT	35	No	Online	No
Hershey et al. 2003	USA	RCT	34	Yes	Classroom	No
Hirad and Zorn 2001	USA	Natural exp.	-	Yes	Mixed	Yes
Hospido et al. 2015	Spain	Quasi exp.	15	-	Classroom	No
Jamison et al. 2014	Uganda	RCT	24	No	Classroom	Yes
Kimball and Shumway 2010	USA	Natural exp.	50	No	Mixed	Yes
Loke et al. 2015	USA	Quasi exp.	15	Yes	Classroom	Yes
Lusardi 2002	USA	Natural exp.	-	-	Classroom	Yes
Lusardi 2005	USA	Natural exp.	55	No	Classroom	No
Lusardi and Mitchell 2007	USA	Natural exp.	53	No	Classroom	No
Lusardi et al. 2014	USA	RCT	50	No	Online	No
Lührmann et al. 2015	Germany	Quasi exp.	14	Yes	Classroom	No
Maki 2004	USA	Natural exp.	40	No	Classroom	No
Mandell 2006	USA	Quasi exp.	12	-	Classroom	No
Mandell 2009a	USA	Quasi exp.	-	-	Classroom	No
Mandell 2009b	USA	Quasi exp.	13	-	Classroom	No
Mandell and Schmid-Klein 2009	USA	Quasi exp.	16	-	Classroom	No

Mills et al. 2004	USA	RCT	36	Yes	Classroom	No
Muller 2003	USA	Natural exp.	-	No	Classroom	No
Pang 2010	Hong Kong, China	Quasi exp.	19	-	Classroom	No
Peng et al. 2010	USA	Natural exp.	35	No	Classroom	Yes
Quercia and Spader 2008	USA	Natural exp.	30	Yes	Classroom	No
Reich and Berman 2015	USA	RCT	30	Yes	Classroom	Yes
Romagnoli and Trifildis 2013	Italy	Quasi exp.	14	No	Classroom	No
Sanders et al. 2007	USA	Quasi exp.	35	Yes	Classroom	Yes
Sarr et al. 2012	India	RCT	38	Yes	Classroom	Yes
Sayinzoga et al. 2015	Rwanda	RCT	40	Yes	Classroom	Yes
Schreiner et al. 2001	USA	Natural exp.	-	Yes	Classroom	Yes
Seshan and Yang 2014	Qatar	RCT	40	Yes	Classroom	Yes
Skimmyhorn 2015	USA	Natural exp.	21	Yes	Classroom	No
Song 2012	China	RCT	45	No	Info. nudge	No
Tennyson and Nguyen 2001	USA	Natural exp.	17	Yes	Classroom	No
Vacroe et al. 2005	USA	Quasi exp.	17	-	Classroom	No
Walstad et al. 2010	USA	Quasi exp.	18	No	Classroom	No
Wiener et al. 2005	USA	Quasi exp.	39	No	Classroom	Yes
Xiao et al. 2012	USA	Natural exp.	18	No	Classroom	No
Yetter and Suiter 2015	USA	RCT	24	Yes	Classroom	No

Appendix E: References for studies included in the statistical meta-analysis

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., and Evanoff, D.D. (2009). Do financial counseling mandates improve mortgage choice and performance? Evidence from a natural experiment. *Federal Reserve Bank of Chicago Working Paper* 2009–07.

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., and Evanoff, D.D. (2010). Learning to cope: Voluntary financial education and loan performance during a housing crisis. *American Economic Review: Papers and Proceedings* 100: 495–500.

Ambuehl, S., Bernheim, B.D., and Lusardi, L. (2014). The effect of financial education on the quality of decision making. *NBER Working Paper 20618.*

Asarta, C.J., Hill, A.T., and Meszaros, B.T. (2014). The features and effectiveness of the keys to financial success curriculum. *International Review of Economics Education*, 16, Part A(0): 39–50.

Barcellos, S.H., Smith, J.P., Yoong, J.K., and Carvalho, L. (2012). Barriers to immigrant use of financial services. The role of language skills, U.S. experience, and return migration expectations. *Financial Literacy Center Working Paper*, *Wr-923-SSA. Dartmouth College and the Wharton School.*

Baron-Donovan, C., Wiener, R.L., Gross, K., and Block-Lieb, S. (2005). Financial literacy teacher training: A multiple-measure evaluation. *Journal of Financial Counseling and Planning*, 16(2): 63–75.

Barua, R., Shastry, G.K., and Yang, D. (2012). Evaluating the effect of peer-based financial education on savings and remittances for foreign domestic workers in Singapore. *Working Paper. Singapoore Management University, Wellesley College, and University of Michigan.*

Batty, M., Collins, J.M., and Odders-White, E. (2015). Experimental evidence on the effects of financial education on elementary school students' knowledge, behavior, and attitudes. *Journal of Consumer Affairs*, 49(1): 69–96.

Bauer, J.W., Son, S., Hur, J., Anderson-Porich, S., Heins, R.H., Petersen, C., Hooper, S., Marczak, M., Olson, P.D., and Wiik, N.B. (2011). Dollar works 2: Impact evaluation report. *University of Minnesota, Extension. St Paul, MN.*

Bayer, P.J., Bernheim, B.D., and Scholz, J.K. (2009). The effects of financial education in the workplace: Evidence from a survey of employers. *Economic Inquiry*, 47(4): 605–624.

Becchetti, L., Caiazza, S., and Coviello, D. (2013). Financial education and investment attitudes in high schools: Evidence from a randomized experiment. *Applied Financial Economics*, 23(10): 817–836.

Berg, G. and Zia, B. (2013). Harnessing emotional connections to improve financial decisions. Evaluating the impact of financial education in mainstream media. *World Bank Policy Research Working Paper 6407.*

Bell, C., Gorin, D., and Hogarth, J.M. (2009). Does financial education affect soldiers' financial behaviors? *Networks Financial Institute Working Paper 2009-WP-08, Indiana State University.*

Bernheim, B.D.. and Garrett, D.M. (2003). The effects of financial education in the workplace: Evidence from a survey of households. *Journal of Public Economics*, 87(7–8): 1487–1519.

Bernheim, B.D., Garrett, D.M., and Maki, D.M. (2001). Education and saving: the long-term effects of high school financial curriculum mandates. *Journal of Public Economics*, 80(3): 435–465.

Berry, J., Karlan, D., and Pradhan, M. (2015). The impact of financial education for youth in Ghana. *NBER Working Paper 21068*.

Bjorvatn, K, and Tungodden, B. (2010). Teaching business in Tanzania: Evaluating participation and performance. *Journal of the European Economic Association*, 8 (2-3): 561–570.

Brown, M., Van der Klaauw, W., Wen, J., and Zafar, B. (2014). Financial education and the debt behavior of the young. *Federal Reserve Bank of New York Staff Report No. 634.*

Brugiavini, A., Cavapozzi, D., Padula, M., and Pettinicchi, Y. (2015). Financial education, literacy and investment attitudes. *SAFE Working Paper No. 86. University of Venice and SAFE-Center, University of Frankfurt.*

Bruhn, M. and Zia, B. (2013). Stimulating managerial capital in emerging markets: The impact of business training for young entrepreneurs. *Journal of Development Effectiveness*, 5(2): 232–266.

Bruhn, M., Ibarra, G.L. and McKenzie, D. (2014). The minimal impact of a large-scale financial education program in Mexico city. *Journal of Development Economics*, 108: 184–189.

Bruhn, M., de Souza Leao, L., Legovini, A., Marchetti, R., and Zia, B. (2013). The impact of high school financial education. Experimental evidence from Brazil. *World Bank Policy Research Working Paper 6723.*

Butt, N.M., Haessler, S.J., and Schug, M.C. (2008). An incentives-based approach to implementing financial fitness for life in the Milwaukee public schools. *Journal of Private Enterprise*, 24(Fall 2008): 165–173.

Calderone, M., Fiala, N., Mulaj, F. Sadhu, S., and Sarr, L. (2013). When can financial education affect savings behavior? Evidence from a randomized experiment among low income clients of branchless banking in India. *Working Paper, German Institute for Economic Research, Berlin and World Bank, Washington D.C.*

Carlin, B.I. and Robinson, D.T. (2012). What does financial literacy training teach us? *Journal of Economic Education*, 43(3): 235–247.

Carpena, F., Cole, S., Shapiro, J., and Zia, B. (2011). Unpacking the causal chain of financial literacy. *World Bank Policy Research Working Paper 5798.*

Carpena, F., Cole, S., Shapiro, J., and Zia, B (2015). The ABCs of financial education. Experimental evidence on attitudes, behavior, and cognitive biases. *World Bank Policy Research Working Paper 7413.*

Chen, W. and Heath, J.A. (2012). The efficacy of financial education in the early grades. In Laney, J.D. and Lucey, T.A., editors, *Reframing Financial Literacy: Exploring the Value of Social Currency*, 189–207. Information Age Publishing. Charlotte, N.C.

Choi, J.J., Laibson, D., and Madrian, B.C. (2005). Are empowerment and education enough? Underdiversification in 401 (k) plans. *Brookings Papers on Economic Activity*, 2005(2): 151–213.

Choi, J.J., Laibson, D., and Madrian, B.C. (2010). Why does the law of one price fail? An experiment on index mutual funds. *Review of Financial Studies*, 23(4): 1405–1432.

Choi, J.J., Laibson, D., and Madrian, B.C. (2011). $100 bills on the sidewalk: Suboptimal investment in 401 (k) plans. *Review of Economics and Statistics*, 93(3): 748–763.

Clancy, M., Grinstein-Weiss, M., and Schreiner, M. (2001). Financial education and savings outcomes in individual development accounts. *Working Paper, Washington University in St. Louis Center for Social Development, St. Louis, MO.*

Clark, R.L., d'Ambrosio, M.B., McDermed, A.A., and Sawant, K. (2006). Retirement plans and saving decisions: The role of information and education. *Journal of Pension Economics and Finance*, 5: 45–67.

Clark, R., Lusardi, A., and Mitchell, O.S. (2015). Employee financial literacy and retirement plan behavior: A case study. *NBER Working Paper 21461.*

Clark, R.L., Maki, J.A., and Morrill, M.S. (2014). Can simple informational nudges increase employee participation in a 401(k) plan? *Southern Economic Journal*, 80(3): 677–701.

Clark, R.L., Morrill, M.S., and Allen, S.G. (2010). Employer-provided retirement planning programs. In Clark, R.L. and Mitchell, O.S., editors, *Reorienting Retirement Risk Management*, 36–64. Oxford University Press.

Cole, S. and Shastry, G.K. (2010). Is high school the right time to teach savings behavior? The effect of financial education and mathematics courses on savings. *Harvard Business School Working Paper.*

Cole, S., Gine, X., Tobacman, J., Topalova, P., Townsend, R., and Vickery, J. (2013). Barriers to household risk management: Evidence from India. *American Economic Journal: Applied Economics*, 5(1): 104–135.

Cole, S., Paulson, A., and Shastry, G.K. (2014), High school curriculum and financial outcomes: The impact of mandated personal finance and mathematics courses. *Harvard Business School Working Paper 13-064.*

Cole, S., Sampson, T., and Zia, B. (2011). Prices or knowledge? What drives demand for financial services in emerging markets? *Journal of Finance*, 66(6): 1933–1967.

Collins, J.M. (2013). The impacts of mandatory financial education: Evidence from a randomized field study. *Journal of Economic Behavior and Organization*, 95: 146–158.

Custers, A. (2011). Furthering financial literacy: Experimental evidence from a financial literacy program for microfinance clients in Bhopal, India. *LSE International Development Working Paper 11-113*, London.

Danes, S.M. and Haberman, H. (2004). Evaluation of the NEFE high school financial planning program: 2003-2004. *National Endowment for Financial Education Project Report, University of Minnesota.*

Danes, S.M, Huddleston-Casas, C., and Boyce, L. (1999). Financial planning curriculum for teens: Impact evaluation. *Financial Counseling and Planning*, 10(1): 26-39.

De Mel, S., McKenzie, D., and Woodruff, C. (2011). Getting credit to high return microentrepreneurs: The results of an information intervention. *World Bank Economic Review*, 25(3): 456–485.

DeLaune, L.D., Rakow, J.S., and Rakow, K. (2010). Teaching financial literacy in a co-curricular service-learning model. *Journal of Accounting Education*, 28(2): 103–113.

Ding, L., Quercia, R.G., and Ratcliffe, J. (2008). Post-purchase counseling and default resolutions among low- and moderate-income borrowers. *Journal of Real Estate Research*, 30(3): 315–344.

Doi, Y., McKenzie, D., and Zia, B. (2014). Who you train matters: Identifying combined effects of financial education on migrant households. *Journal of Development Economics*, 109(0): 39–55.

Dolvin, S. and Templeton, W.K. (2006), Financial education and asset allocation. *Financial Services Review*, 15: 133–149.

Drexler, A., Fischer, G., and Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, 6(2): 1–31.

Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, 118(3): 815–842.

Elliehausen, G., Lundquist, C.E., and Staten, M.E. (2007). The impact of credit counseling on subsequent borrower behavior. *Journal of Consumer Affairs*, 41(1): 1–28.

Evaluation and Training Institute ETI (2008). JA finance park. Final report. Los Angeles: ETI.

Field, E., Jayachandran, S., and Pande, R. (2010). Do traditional institutions constrain female entrepreneurship? A field experiment on business training in india. *American Economic Review: Papers and Proceedings*, 100(2): 125–29.

Garman, E.T., Kim, J., Kratzer, C.Y., Bruce, H.B., and Joo, S. (1999). Workplace financial education improves personal Financial Wellness. *Financial Counseling and Planning,* 10(1): 79–88.

Gaurav, S., Cole, S., and Tobacman, J. (2011). Marketing complex financial products in emerging markets: Evidence from rainfall insurance in India. *Journal of Marketing Research*, 48(SPL): S150–S162.

Gibson, J., McKenzie, D., and Zia, B. (2014). The impact of financial literacy training for migrants. *World Bank Economic Review*, 28(1): 130–161.

Gill, A. and Bhattacharya, R. (2015). Integration of a financial literacy curriculum in a high school economics class: Implications of varying the input mix from an experiment. *Journal of Consumer Affairs*, 49 (2): 472–487.

Gine, X. and Mansuri, G. (2014). Money or ideas? A field experiment on constraints to entrepreneurship in rural Pakistan. *World Bank Policy Research Working Paper 6959*.

Gine, X., Karlan, D., and Ngatia, M. (2013). Social networks, financial literacy and index insurance. *World Bank, Washington, DC.*

Go, C.G., Varcoe, K., Eng, T., Pho, W., and Choi, L. (2012). Money savvy youth: Evaluating the effectiveness of financial education for fourth and fifth graders. *Federal Reserve Bank of San Francisco Working Paper 2012-02.*

Goda, G.S., Manchester, C.F., and Sojourner, A.J. (2014). What will my account really be worth? Experimental evidence on how retirement income projections affect saving. *Journal of Public Economics*, 119: 80–92.

Goldsmith, R.E. and Goldsmith, E.B. (2006). The effects of investment education on gender differences in financial knowledge. *Journal of Personal Finance*, 5(2): 55–69.

Grimes, P.W., Rogers, K.E., and Smith, R.C. (2010). High school economic education and access to financial services. *Journal of Consumer Affairs*, 44(2): 317–335.

Grinstein-Weiss, M., Guo, S., Reinertson, V., and Russel, B. (2015). Financial education and savings outcomes for low-income IDA participants: Does age make a difference? *Journal of Consumer Affairs*, 49(1): 156–185.

Han, C.-K., Grinstein-Weiss, M., and Sherraden, M. (2009). Assets beyond savings in individual development accounts. *Social Service Review*, 83(2): 221–244.

Hartarska, V. and Gonzalez-Vega, C. (2005). Credit counseling and mortgage termination by low-income households. *Journal of Real Estate Finance and Economics*, 30(3): 227–243.

Hartarska, V. and Gonzalez-Vega, C. (2006). Evidence on the effect of credit counseling on mortgage loan default by low-income households. *Journal of Housing Economics*, 15(1): 63–79.

Harter, C.L. and Harter, J.F. (2009). Assesing the effectiveness of Financial Fitness for life in eastern Kentucky. *Journal of Applied Economics and Policy*, 28: 20–33.

Harter, C.L. and Harter, J.F. (2010). Is financial literacy improved by participating in a stock market game? *Journal for Economic Educators*, 10(1): 21–32.

Haynes, D.C., Haynes, G., and Weinert, C. (2011). Outcomes of on-line financial education for chronically ill rural women. *Journal of Financial Counseling and Planning*, 22(1): 3–17.

Haynes-Bordas, R., Kiss, D., and Yilmazer, T. (2008). Effectiveness of financial education on financial management behavior and account usage: Evidence from a 'second chance' program. *Journal of Family and Economic Issues*, 29(3): 362–390.

Heinberg, A., Hung, A.A., Kapteyn, A., Lusardi, A., Samek, A.S., and Yoong, J. (2014). Five steps to planning success. Experimental evidence from U.S. households. *Oxford Review of Economic Policy,* 30(4): 697-724.

Hershey, D.A., Mowen, J.C., and Jacobs-Lawson, J.M. (2003). An experimental comparison of retirement planning intervention seminars. *Educational Gerontology*, 29(4): 339–359.

Hirad, A. and Zorn, P.M. (2001). A little knowledge is a good thing: Empirical evidence of the effectiveness of pre-purchase homeownership counseling. *Low-Income Homeownership Working Paper Series LIHO-01.4, JCHS, Harvard University.*

Hospido, L., Villanueva, E., and Zamarro, G. (2015). Finance for all: The impact of financial literacy training in compulsory secondary education in Spain. *Banco de España Working Paper 1502*, Madrid.

Jamison, J.C., Karlan, D, and Zinman, J. (2014). Financial education and acces to savings accounts: Complements or substitutes? Evidence from Ugandan youth clubs. *NBER Working Paper 20135.*

Kimball, M.S. and Shumway, T. (2010). Investor sophistication and the home bias, diversification, and employer stock puzzles. *Working Paper, University of Michigan.*

Loke, V., Choi, L., and Libby, M. (2015). Increasing youth financial capability: An evaluation of the mypath savings initiative. *Journal of Consumer Affairs*, 49(1): 97–126.

Lusardi, A. (2002). Preparing for retirement: The importance of planning costs. *National Tax Association Proceedings–2002*: 148–154.

Lusardi, A. (2005). Financial education and the saving behavior of African American and Hispanic households, *Report*, *US Department of Labor, Employee Benefits Security Administration.*

Lusardi, A. and Mitchell, O.S. (2007). Financial literacy and retirement planning: New evidence from the Rand American life panel. *Michigan Retirement Research Center Research Paper No. WP 2007-157.*

Lusardi, A., Samek, A.S., Kapteyn, A., Glinert, L., Hung, A., and Heinberg, A. (2015). Visual tools and narratives: New ways to improve financial literacy. *Forthcoming in Journal of Pension Economics and Finance.*

Lührmann, M., Serra-Garcia, M., and Winter, J. (2015). Teaching teenagers in finance: Does it work? *Journal of Banking and Finance*, 54: 160–174.

Maki, D. (2004), Financial education and private pensions, in Gale, W., Shoven, J. and Warshowsky, M. (editors.), *Private Pensions and Public Policies*, Washington, DC: Brookings Institution Press: 126-139.

Mandell, L. (2006). Teaching young dogs old tricks:The effectiveness of financial literacy intervention in pre-high school grades. P*aper Presented at the Academy of Financial Services 2006 Annual Conference Salt Lake City.*

Mandell, L. (2009a). Starting younger: Evidence supporting the effectiveness of personal financial education for pre-high school students. *Working Paper, Aspen Institute and University of Washington.*

Mandell, L. (2009b). The impact of financial education in high school and college on financial literacy and subsequent financial decision making. *Paper presented at the American Economic Association Meetings, San Francisco, CA*.

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

Mills, G., Patterson, R. Orr, L., and DeMarco, D. (2004). *Evaluation of the american dream demonstration: Final evaluation report*, Cambridge, MA.

Muller, L. (2003). Does retirement education teach people to save pension distributions? *Social Security Bulletin,* 64 (4), 48–65.

Pang, M.F. (2010). Boosting financial literacy: benefits from learning study. *Instructional Science,* 38(6): 659–677.

Peng, T.-C.M., Bartholomae, S., Fox, J.J., and Cravener, G. (2007). The impact of personal finance education delivered in high school and college courses. *Journal of Family and Economic Issues*, 28(2): 265–284.

Quercia, R. and Spader, J. (2008). Does homeownership counseling affect the prepayment and default behavior of affordable mortgage borrowers? *Journal of Policy Analysis and Management*, 27(2): 304–325.

Reich, C.M. and Berman, J.S. (2015). Do financial literacy classes help? An experimental assessment in a low-income population. *Journal of Social Service Research*, 41(2): 193–203.

Romagnoli, A. and Trifilidis, M. (2013). Does financial education at school work? Evidence from Italy. *Banco D'Italia Occasional Papers N. 155.*

Sanders, C.K., Weaver, T.L., and Schnabel, M. (2007) Economic education for battered women. An evaluation of outcomes. *Affilia: Journal of Women and Social Work*, 22(3): 240–254.

Sarr, L., Sadhu, S., and Fiala, N. (2012). Bringing the bank to the doorstep: Does financial education influence savings behavior among the poor? Evidence from a randomized financial literacy program in India. *Working Paper. Institute for Financial Management and Research, Centre for Micro Finance, German Instiute for Economic Research and the World Bank, Washington, D.C.*

Sayinzoga, A., Bulte, E. H., and Lensink, R. (2015). Financial literacy and financial behaviour: Experimental evidence from rural Rwanda. *Economic Journal*, forthcoming.

Schreiner, M., Sherraden, M., Clancy, M., Johnson, L., Curley, J., Grinstein-Weiss, M., Zhan, M., and Beverly, S. (2001). Savings and asset accumulation in individual development accounts. *Working paper, Center for Social Development, Washington University.*

Seshan, G. and Yang, D. (2014). Motivating migrants: A field experiment on financial decision-making in transnational households. *NBER Working Paper 19805.*

Skimmyhorn, W. (2015). Assessing financial education: Evidence from boot camp. *American Economic Journal: Economic Policy,* forthcoming.

Song, C. (2012). Financial illiteracy and pension contributions: A field experiment on compound interest in China. *Unpublished Manuscript.*

Tennyson, S. and Nguyen, C. (2001). State curriculum mandates and student knowledge of personal finance. *Journal of Consumer Affairs*, 35(2): 241–262.

Varcoe, K.P., Martin, A., Devitto, Z, and Go, C. (2005). Using a financial education curriculum for teens. *Journal of Financial Counseling and Planning*, 16(1): 63–71.

Walstad, W.B., Rebeck, K., and MacDonald, R.A. (2010). The effects of financial education on the financial knowledge of high school students. *Journal of Consumer Affairs*, 44(2): 336–357.

Wiener, R.L., Baron-Donovan, C., Gross, K., and Block-Lieb, S. (2005). Debtor education, financial literacy, and pending bankruptcy legislation. *Behavioral Sciences and the Law*, 23(3): 347–366.

Xiao, J.J., Serido, J., and Shim, S. (2012). Financial education, financial knowledge and risky credit behavior of college students. In: Lamdin, D.J. (editor) *Consumer Knowledge and Financial Decisions: Lifespan Perspectives*. New York: Springer: 113–128.

Yetter, E.A. and Suiter, M. (2015). Financial literacy in the community college classroom: A curriculum intervention study. *Federal Reserve Bank of St. Louis Working Paper 2015-001.*