Economics of Education Review 78 (2020) 101930

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Financial education in schools: A meta-analysis of experimental studies

Tim Kaiser[a,*], Lukas Menkhoff[b]

[a] University of Koblenz-Landau and German Institute for Economic Research (DIW Berlin), 76829 Landau, Germany
[b] Humboldt-University of Berlin and German Institute for Economic Research (DIW Berlin), 10108 Berlin, Germany

ARTICLE INFO

Keywords:
Financial behavior
Financial education
Financial literacy
Meta-analysis

JEL classification:
I21, analysis of education
A21, pre-college economic education
D14, personal finance

ABSTRACT

We study the literature on school financial education programs for children and youth via a quantitative meta-analysis of 37 (quasi-) experiments. We find that financial education treatments have, on average, sizeable impacts on financial knowledge (+0.33 SD), similar to educational interventions in other domains. Additionally, we document smaller effects on financial behaviors among students (+0.07 SD). When restricting the sample to 18 randomized experiments average effect sizes are estimated to be about 0.15 SD units on financial knowledge and 0.07 SD units on financial behaviors. These results are robust irrespective of the meta-analytic method used and when accounting for publication bias. Subgroup analyses show the beneficial effect of more intensive treatments, albeit with decreasing marginal returns.

1. Introduction

Financial education is high on the agenda of policymakers worldwide. An abundance of rigorous empirical research shows the importance of financial literacy for individual welfare (cf. Lusardi & Mitchell, 2014). Financial education policies and programs are being installed in the vast majority of OECD countries and in many of the largest emerging economies, such as China and India (see OECD, 2015). While these programs vary in size, design and coverage, many of these programs are designed to be implemented in schools. School-based financial education may be seen as a promising avenue since it allows an almost universal coverage of a cohort, mitigating previously documented low-demand of voluntary financial education later in the lifecycle (e.g., Bruhn, Ibarra & McKenzie, 2014). Moreover, providing financial education during formative years could be effective and sustainable with respect to long-term outcomes (e.g., Frisancho, 2018; Lührmann, Serra-Garcia & Winter, 2018; Lusardi, Mitchell & Curto, 2010).

We contribute to the literature – to the best of our knowledge – with the first quantitative meta-analysis focusing exclusively on the impact of school-based financial education among children and youth. The empirical basis of our meta-analysis is the complete set of those empirical studies that (i) report about impacts of financial education programs in schools among children and youth, (ii) provide a quantitative assessment of treatment effects and (iii) rely on a control group. In summary there are 37 independent (quasi-) experimental studies fulfilling the above three criteria, 18 of them are randomized experiments (RCTs). As studies mostly report impacts on a set of several outcomes, our meta-analysis relies on 177 effect size estimates, of these 70 refer to treatment effects on measures of financial knowledge and 107 refer to treatment effects on a set of financial behaviors among students.

Based on this sample of studies we find, on average, positive treatment effects, i.e., improved financial knowledge test scores and changes in financial behaviors that are typically assumed to be enhancing individual welfare (e.g., increasing personal savings). We show that these effects are statistically different from zero, that they hold for the outcomes of financial knowledge and behaviors, and that they exist also when restricting the sample to RCTs. Reassuringly, these results are robust to employing various estimation methods: the effect of financial education on knowledge is higher than on financial behavior, and the effect documented in RCTs is estimated to be smaller than in quasi-experimental studies. However, even the smallest effect size we find in our study, i.e. from financial education treatment on financial behaviors in RCTs estimated in a fixed effects meta-analysis with a correction for publication selection bias, still has a positive and significant coefficient.

We thank the authors of the primary studies who responded to requests to provide their datasets or further details about their studies for their kind cooperation. We would like to thank the editors of the special issue, Annamaria Lusardi and Carly Urban, as well as Veronica Frisancho, Jane Lopus, Frauke Peter, and two anonymous referees for their helpful comments. Financial support by the research promotion arm (Wissenschaftsförderung) of the Savings Bank Group e.V. in Germany and DFG through CRC TRR 190 is gratefully acknowledged. We thank Melanie Krüger and Iven Lützen for research assistance.

* Corresponding author.
E-mail addresses: kaiser@uni-landau.de (T. Kaiser), lmenkhoff@diw.de (L. Menkhoff).

https://doi.org/10.1016/j.econedurev.2019.101930
Received 14 November 2018; Received in revised form 14 September 2019; Accepted 20 September 2019
Available online 11 October 2019
0272-7757/

Comparing the effectiveness of financial education in schools to the larger universe of empirical studies on financial education programs (covering mostly adults) as examined in Kaiser and Menkhoff (2017), our results are consistent with the interpretation that the impact on *knowledge* is, on average, tentatively larger than in the extended sample of studies while the impact on *behavior* is rather smaller. This is a plausible result as younger people may generally have a higher capacity to learn than adults which could explain the larger average treatment effects of financial education for children and youth on financial knowledge. This finding is very similar to evidence from other domains such as math and reading, where effect sizes are largest for younger students (see Hill et al., 2008; Fryer, 2016). At the same time, the empirical evidence suggests that the motivation to incorporate financial knowledge into financial behavior is higher when financial decisions are more immediate and relevant (e.g., Kaiser & Menkhoff, 2017; Miller, Reichelstein, Salas & Zia, 2015), and this may not fully apply to younger students in schools. Additionally, changes in financial behavior among children and youth are inherently more difficult to observe (measure), since children do not engage in a lot of financial decision at this stage of the lifecycle. Thus, average statistical effect sizes on actual financial behaviors are estimated to be rather small and less certain than in the general literature.

Our main result is of high relevance for policy makers because the evidence clearly suggests that investing into the implementation of school financial education curricula does indeed impact financial knowledge, and to a smaller extent financial behaviors. This result is important, because there is a public debate questioning the effectiveness of financial education in quite fundamental ways (e.g., Fernandes, Lynch & Netemeyer, 2014; Willis, 2011).

Against the backdrop of limited public resources, we finally discuss potential determinants of effective financial education programs in schools. Unfortunately, the number and comparability of studies at hand is not large enough to generate truly granular insights in this respect. For example, potentially crucial determinants of effective programs cannot be directly assessed in this quantitative meta-analysis, such as differences in implementation quality (Urban, Schmeiser, Michael Collins & Brown, 2018), in teacher training and experience (Chetty, Friedman & Rockoff, 2014; Harris & Sass, 2011; Rockoff, 2004; Urban et al., 2018), the quality of curricula (Drexler, Fischer & Schoar, 2014), material and media (Heinberg et al., 2014; Lusardi et al., 2017), and the teaching methods employed (Kaiser & Menkhoff, 2018).

Nevertheless, and despite the limited number of 37 underlying primary studies, this meta-analysis covers all available (high-quality) evidence and thus informs about the state of the art of financial education in schools. We arrive at five results being relevant in designing and evaluating school financial education programs: research design and measurement matters in impact evaluation of financial education interventions as is known from earlier meta-analyses of the literature (Fernandes et al., 2014; Kaiser & Menkhoff, 2017) as well as meta-analyses in other educational domains (Cheung & Slavin, 2016), i.e., (i) the effect size estimated in RCTs is smaller than in other study types, such as quasi-experimental impact evaluations, and (ii) a longer delay between financial education and measurement of outcomes is associated with deflated effect size estimates. Moreover, we provide evidence of two specific results which also fit into the broader literature on educational interventions: (iii) effect sizes reported in interventions in primary schools are statistically significant and possibly larger than in secondary schools, and (iv) higher intensity of teaching increases effectiveness with declining marginal returns. Finally, and preliminary, (v) lower student to teacher ratios (class size) may be associated with a higher degree of effectiveness; however, this result has to be viewed with caution due to the small number of studies and the limited variation in class-sizes.

This meta-analysis contributes to two lines of literature, i.e., meta-analyses on the effectiveness of financial education in general, and studies examining the effectiveness of particular financial education programs in schools. Regarding the more general meta-analyses, Fernandes et al. (2014) use broader selection criteria (including observational studies) than we do and cover studies until 2013. That study fundamentally questions the effectiveness of financial education (for adults) by the argument that only observational studies in their sample show a positive correlation between financial literacy and behavior while experimental assessments show small treatment effects on financial behaviors; the effect is estimated to be insignificant in their set of 13 RCTs. The next meta-analysis by Miller et al. (2015) limits itself to just 18 studies with homogenous outcomes and shows that some financial behaviors, such as savings behavior, seem easier to be impacted than others, such as credit defaults. Kaiser and Menkhoff (2017) conduct a meta-analysis aiming for full coverage of financial education interventions; this ambition is comparable to the "manipulated-literacy sample" assembled by Fernandes et al. (2014) but data are more recent. It is found that financial education is effective, and this finding holds also for the sub-sample of rigorous RCTs. We are closest to Kaiser and Menkhoff (2017) but focus exclusively on financial education in schools and update the data by considering additional recent studies. Appendix Table A1 shows the overlap that our meta-analysis has with this most recent meta-analysis by Kaiser and Menkhoff (2017). Out of the 37 (quasi-) experiments included in our sample, 16 studies have not been included in Kaiser and Menkhoff (2017). Ten of these additional studies are randomized experiments, extending the available evidence from RCTs in schools from eight to 18 studies. The clear focus on financial education in schools results in a more homogenous sample of experimental studies which allows analyzing the potential impact of teaching intensity and its delayed effectiveness in detail. Moreover, only the school focus allows the investigation of design features specific to school financial education programs, i.e., the consideration of school types (primary, middle and high school) and the tentative consideration of class size.

Another strand of the literature uses (plausibly) exogenous variation in U.S. high school financial education mandates across federal states over time to investigate effects of exposure to financial education on financial knowledge (Tennyson & Nguyen, 2001) and financial behaviors, such as the handling of debt (i.e., reduction in defaults) and savings outcomes (see Bernheim, Garrett & Maki, 2001; Brown, Grigsby, van der Klaauw, Wen & Zafar, 2016; Cole, Paulson & Shastry, 2016; Urban et al., 2018). While this literature documents a positive effect of financial education on financial knowledge (Tennyson & Nguyen, 2001) and on savings (e.g., Bernheim et al., 2001), it provides partially conflicting results on the (long-term) effects of financial education mandates on credit-related behavior (e.g., Brown et al., 2016; Cole et al., 2016). Brown et al. (2016) reveal long-term effects of financial education on reduced debt levels and loan defaults, but Cole et al. (2016) do not find such an effect. Recently, Urban et al. (2018) show that accounting for heterogeneity in the timing and quality of policy implementation at the state-level leads to the assessment of positive effects of financial education mandates on credit outcomes among young adults. Thus, while parts of this literature document important and long-run effects of financial education on financial outcomes with high external validity, the high degree of variation in the employed research designs in these papers (e.g., the definition of policy changes varies across studies) makes the systematic integration of this literature into a meta-analysis hardly possible.[1] Thus, we only include controlled (quasi-) experiments where the treatment is closely observed by the researchers.

This paper is structured into four further sections. Section 2 introduces into the method and selection criteria for considered studies. Section 3 describes the dataset, and Section 4 reports the results. Finally, Section 5 concludes by discussing these results and highlights possibilities for future research.

2. Method

Meta-analysis is a quantitative method to integrate results from multiple

[1] There are several other studies on the effects of financial education courses in school on financial knowledge and financial behavior (e.g., Peng et al., 2007; Grimes et al., 2010). These studies are observational and have varying degrees of internal validity, so that we do not include them in our meta-analysis.

empirical studies on the same empirical phenomenon (see Stanley, 2001, for an introduction). In a meta-analysis, the dependent variable is comprised of summary statistics reported in the primary research, while the explanatory variables may include, for example, characteristics of the research design, the target group, or the particular education program. Meta-analyses are helpful to address three types of general research questions about a given empirical literature: First, what is the direction and size of the (weighted) average effect of a treatment? Second, are results consistent across studies or is there a high degree of heterogeneity in reported findings (beyond measurement error)? Third, are there observable study or program characteristics that may explain part of this heterogeneity?

To be able to draw conclusions about an entire empirical literature, one has to assemble a complete representation of the literature of interest, meeting certain quality and inclusion criteria. Thus, we build on our existing database and update it using the same search strategy as described in Kaiser and Menkhoff (2017). We augment the earlier dataset with published studies on financial education in schools between October 2016 (end of collection period in Kaiser & Menkhoff, 2017) and September 2018 (end of collection period for this paper). Additionally, as our review of the larger literature on financial education included a screening of references from previous meta-analyses (Fernandes et al., 2014; Miller et al., 2015) as well as narrative reviews (Blue, Grootenboer & Brimble, 2014; Collins & O'Rourke, 2010; Fox, Bartholomae & Lee, 2005; Hastings, Madrian & Skimmyhorn, 2013; Lusardi & Mitchell, 2014; Willis, 2011; Xu & Zia, 2012) we also screen the references of more recent or more focused narrative reviews of financial education for children and youth in schools (Amagir, Groot, Maassen van den Brink & Wilschut, 2018; Collins & Odders-White, 2015; Walstad et al., 2017). We screen all of the abstracts for relevance and apply our inclusion criteria to the remaining full texts: We include papers (i) reporting on impacts of an educational intervention on financial literacy and/or financial behavior for children and/or youth in schools, (ii) providing a quantitative assessment of intervention impact that allows coding an effect size statistic (g) and its standard error, and (iii) relying on a control group in the estimation of intervention impacts. Consequently, we only include (quasi-) experimental studies with sufficient information on intervention outcomes in our analysis while neglecting single-group pre-/post comparisons, since these have a lower degree of internal validity. Where necessary information is only partially missing, we consult additional online resources related to the article or contact the authors of the primary studies directly.

In order to be able to aggregate estimated treatment effects reported across multiple studies, one must standardize these statistics into a common metric. Ideally, all of these studies would measure the outcomes of financial education identically, i.e., in the same unit. If this was the case, a meta-analysis could be performed directly on the outcomes and standardization was not necessary. In the heterogeneous body of literature on school financial education, however, standardization becomes necessary, because studies typically measure increases in financial knowledge in different ways (use different test items) or employ multiple methods or data sources to measure changes in financial behaviors. Thus, we conduct our meta-analysis using scale-free statistical effect sizes. Specifically, we compute the so-called "bias corrected standardized mean difference" (*Hedges' g*) as our effect size measure for each reported estimate within studies. This measure reports treatment effects in the form of scale-free standard deviation units.[2]

Regarding the meta-analytic method, there is a variety of models available, each making different assumptions: In the meta-analysis literature it is common to distinguish between a "fixed-effect" approach and a "random-effect" approach (see Lipsey & Wilson, 2001). Choosing a model from the "fixed-effect"-family implies that the researcher assumes the source of variance to be exclusively due to measurement error within each study. Put in other words: if each study had indefinitely large sample sizes one would be able to observe and calculate an estimate of a common true effect that every study shares. In contrast, a "random-effect" approach assumes that – in addition to within-study measurement error – there exists actual heterogeneity in the true effects between studies. Even if studies had no measurement error, it would still be possible that two studies did not share a common true effect. Most of canonical meta-analysis models (e.g., DerSimonian & Laird, 1986) from other disciplines use a random-effect approach, while meta-analyses of economic research frequently use "fixed-effect" models (e.g., Staney & Doucouliagos, 2012).

Regarding the characteristics of our sample, we argue that the degree of heterogeneity across primary studies makes it difficult to assume that there is indeed one common true effect. Rather it seems plausible to us that the diverse target groups, teaching approaches, intensities of education etc. speak in favor of a random effects approach, i.e., estimating the mean of the distribution of true effects. This is our preferred approach. Specifically, we estimate the mean of the distribution of true effects using "robust variance estimation in meta-regression with dependent effect size estimates (RVE)" (Hedges, Tipton & Johnson, 2010).[3] At the same time, we acknowledge different views on the appropriate method. Thus, we also use multiple approaches from both families of models to investigate the sensitivity of our results to the assumptions implied by each meta-analytic method.

As robustness-checks, we first estimate an unweighted average effect of financial education by relying on an ordinary least squares (OLS) model where each study contributes multiple effect sizes (see Card, Kluve & Weber, 2017, for such an approach). We account for the statistical dependency of estimates in this data-structure by clustering the standard errors at the study level. The OLS-model places equal weights on each estimate and thus represents a description about the literature, without necessarily speaking to an estimate of a possible "true effect" of financial education in the broader set of possible studies.

Second, we estimate the same model but weight each effect size estimate by its inverse standard error or the inverse variance, respectively. This unrestricted weighted least squares (WLS) estimation is advocated by Stanley and Doucouliagos (2012, 2015).

Third, we account for potential publication selection bias in the financial education literature by testing for funnel asymmetry (FAT) and estimating both "precision-effect test" and "precision-effect estimate with standard error" (PET and PEESE) models as suggested by Stanley and Doucouliagos (2012).

The unrestricted weighted least squares (WLS) models place extreme weight on larger studies, since these minimize the standard errors and variance of the estimate while assuming that each estimate relates to a single true effect. Thus, estimates from this family of models (and especially those accounting for potential publication selection bias, PET-PEESE) may serve as a conservative lower-bound of financial education treatment effects. By contrast, our preferred robust-variance estimation with dependent effect sizes explicitly models between-study heterogeneity in addition to within-study measurement error. As a consequence, smaller studies are not as strongly discounted as in the WLS-approach, since within-study

[2] Hedges' g is defined as: $g = \frac{M_T - M_C}{SD_p}$ with $SD_p = \sqrt{\frac{(n_T - 1)\, SD_T^2 + (n_C - 1)\, SD_C^2}{n_T^2 + n_C^2 - 2}}$. 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, the standard error of each standardized mean difference (g), is defined as: $SE_g = \sqrt{\frac{n_T + n_C}{n_T n_C} + \frac{g^2}{2(n_T + n_C)}}$.

[3] Formally, we estimate the following model: $y_{ij} = \beta_0 + \upsilon_j + \epsilon_{ij}$ in which y_{ij} is defined as the ith treatment effect estimate within each study j. β_0 is the mean of the distribution of true effects, υ_j is the study-level random effect with $\upsilon_j \sim N(0, \tau^2)$, τ^2 is the between study variance in true effects which is unknown and has to be estimated from the data using method of moments, and $\epsilon_{ij} \sim N(0, \sigma_{ij}^2)$ is the residual of the ith treatment effect estimate within each study j. We use the following weights to account for the correlation of estimates within studies: $w_{ij} = \frac{1}{\{(\tau^2 + \frac{1}{k_j}\sum_{k_j=1}^{k_i} \sigma_{ij}^2)[1 + (k_j - 1)\rho]\}}$, where τ^2 is the estimated between study variance in true effects, $(\frac{1}{k_j}\sum_{k_j=1}^{k_i} \sigma_{ij}^2)$ is the arithmetic mean of the within study sampling variances (σ_{ij}^2) with k_j being the number of i effect size estimates within each study j, and ρ is the assumed common within-study correlation of treatment effect estimates. We use the default $\rho = 0.8$ (Tanner-Smith & Tipton, 2014), but results are insensitive to changes in ρ.

measurement error is only one source of variance. This approach yields an estimate of the mean of the distribution of true effects in the universe of potential financial education impact evaluation studies in the presence of excess heterogeneity between studies.

In addition to estimating the average effect of financial education treatment, we are interested in exploring the determinants of effectiveness of programs reported across studies. Thus, we code observable characteristics and investigate whether these may explain some of the heterogeneity in the literature.

3. Data

The application of the reported selection criteria (see Section 2) leads to a sample of 37 independent (quasi-) experimental studies in schools reported in 35 papers published between 1978 and 2018 (these studies are listed in Appendix and an overview is provided in Table A1). The aggregate sample size of these 37 (quasi-) experiments amounts to over 115,000 students (see Table A1). The majority of papers has been published in recent years, 20 out of 37 since 2015. Out of these 37 studies, 18 are randomized experiments (RCTs) and 19 are quasi-experimental studies that employ a non-randomly selected control group. A description about the publication year of these two study types, i.e., either RCT or quasi-experimental studies, is provided in Fig. 1. It is apparent that RCTs are conducted more recently and dominate this literature since 2015 (with a 75% share of studies).

From these studies, we extract a total of 177 effect size estimates, because individual studies typically look at multiple outcomes, measure outcomes at multiple time points, or include separate effect size estimates for different school grades. In our sample, RCTs report more estimates per paper, since the 18 RCTs account for 135 estimates in our sample. The 19 quasi-experiments, in contrast, contribute 42 effect size estimates to our analysis.

With regard to outcome types, we consider two main families of outcomes: (i) treatment effects on *financial knowledge* (i.e., performance on a standardized financial knowledge test), and (ii) treatment effects on *financial behaviors* and their antecedents (for example an increase in savings or an observed financial decision in an incentivized experimental task) (see Table A2 in the Appendix for an overview and definition of the included outcomes). Not all of the included studies report treatment effects on both outcome families: The dataset includes information from 31 studies (70 effect size estimates) on *financial knowledge*. Out of these, 14 are RCTs which report 41 effect size estimates. Information on impact on *financial behaviors* comes from 22 studies (107 effect size estimates). Out of these, 16 are RCTs and account for 94 effect size estimates. Thus, 16 studies report on both types of outcomes.

For each effect size estimate we code a number of characteristics in order to analyze, in Section 4, potential determinants of effectiveness. These

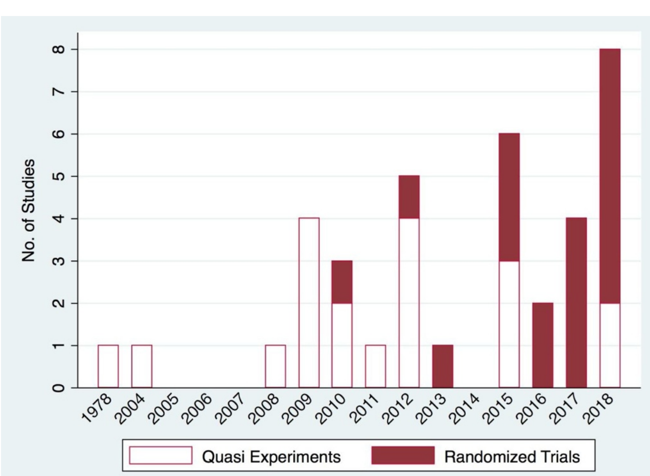

Fig. 1. Number of included studies by research design per year.

characteristics fall into three groups, i.e., (i) research design, (ii) characteristics of the target group, and (iii) design elements of the education program.

The mean of the extracted effect sizes ($n = 177$) is 0.162 with a standard deviation of 0.251 and values between −0.236 and 1.321 (see Table 1). Among all effects we distinguish between the outcome types of financial knowledge and financial behaviors.

Regarding the (i) research design we code, as mentioned already, whether the study is a RCT or a quasi-experiment and the standard errors of the effect sizes. Moreover, for 166 of 177 effect sizes, we have information about the average delay between treatment and measurement of potential effects (mean of 17.17 weeks).

Coming to (ii) characteristics of the target group, we code the country where the study takes place, and studies provide information about school grades, so that we can group into elementary, middle and high school students, covering 24%, 49% and 27% of observations, respectively. Note that, some studies omit continuous measures of age, so that we only include grades as a proxy of age. Also, information about gender-composition of the sample or the social status of parents (such as their income) is not always available.

The last group of characteristics covers (iii) the design elements of the educational program. While we have information about the intensity of education, which is 19.9 h on average, and, for a sub-sample of 138 effect size estimates, about the average class size of 26 students per class, there is a lack of systematic information regarding the content of curricula, the quality of materials and media such as textbooks, the quality of teachers or program implementation, details about previous teacher training, and the teaching method employed (i.e., lecture or active learning). Thus, unfortunately, these latter characteristics cannot be considered in a quantitative meta-analysis as long as studies do not document enough detailed information to capture these differences.

4. Results

We present results in three steps: first, main results are shown (Section 4.1), then the possibility of publication bias and small-study effects is discussed (Section 4.2), and finally, potential determinants of the effectiveness of financial education in schools are examined (Section 4.3).

4.1. Summary effects

The summary effects of financial education in schools are estimated separately for the outcome types of financial knowledge and financial behavior. It is known from the literature (e.g., Kaiser & Menkhoff, 2017) and seems to be intuitive that educational effects on knowledge are larger than on behavior. This is indeed the finding from this meta-analysis as shown in Figs. 2 and 3: the estimated mean of the distribution of true effects on financial knowledge is about 0.33 standard deviation units (SDs), based on 70 effect size estimates from 31 studies, while the average effect on behavior is about 0.07 SDs, based on 107 effect size estimates from 22 studies.

Table 1
Descriptive statistics at the estimate-level.

Variable	Obs.	Mean	Median	Std. dev.	Min.	Max.
Hedges g	177	0.162	0.100	0.252	−0.236	1.321
SE	177	0.065	0.050	0.059	0.013	0.372
SE^2	177	0.008	0.003	0.016	0.000	0.139
RCT	177	0.763	–	–	0	1
High income economy	177	0.519	–	–	0	1
Delay (in weeks)	166	17.17	4.35	30.620	0	132.675
Intensity (in hours)	174	19.90	7.5	36.536	0	150
Elementary school	177	0.239	–	–	0	1
Middle school	177	0.490	–	–	0	1
High school	177	0.271	–	–	0	1
Class size (no. of students)	138	25.808	25	7.323	7	35

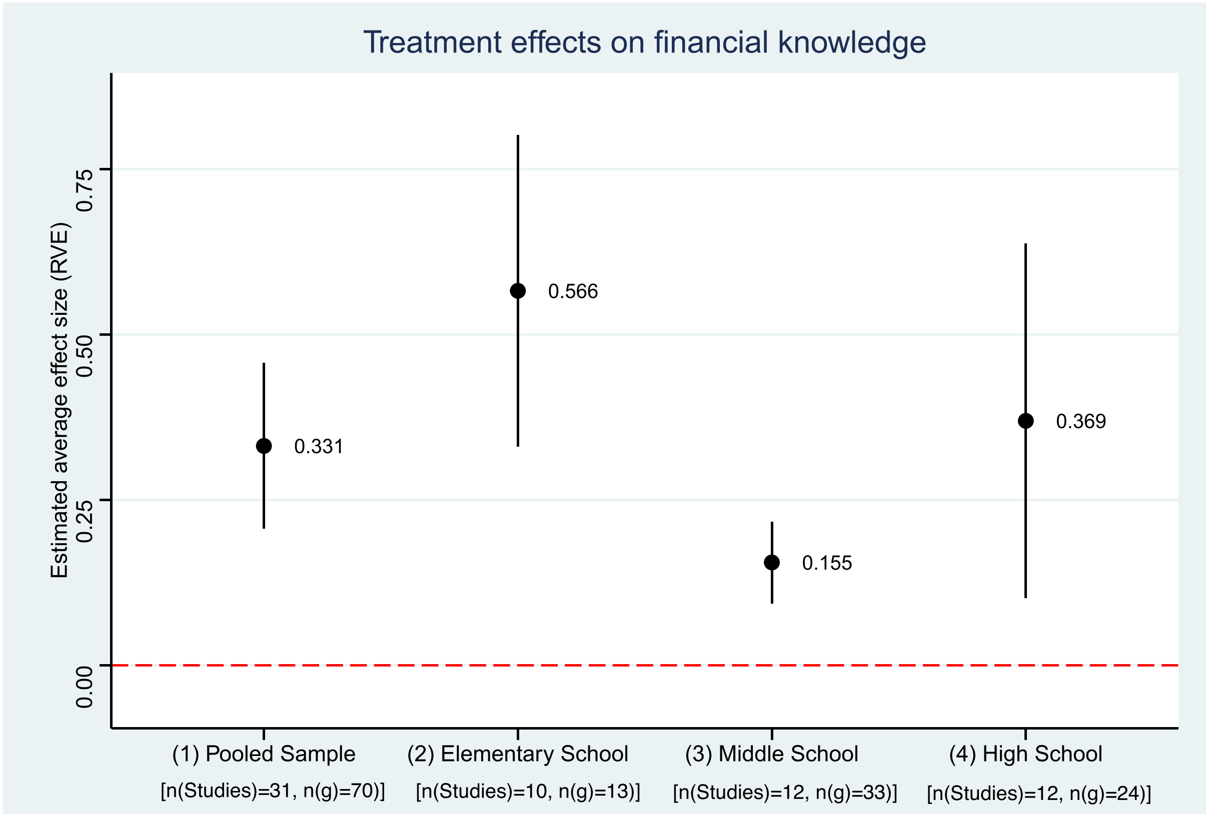

Fig. 2. Treatment effects on financial knowledge (RVE).
Notes: The figure shows the (weighted) average effects and 95% Cis estimated by RVE.

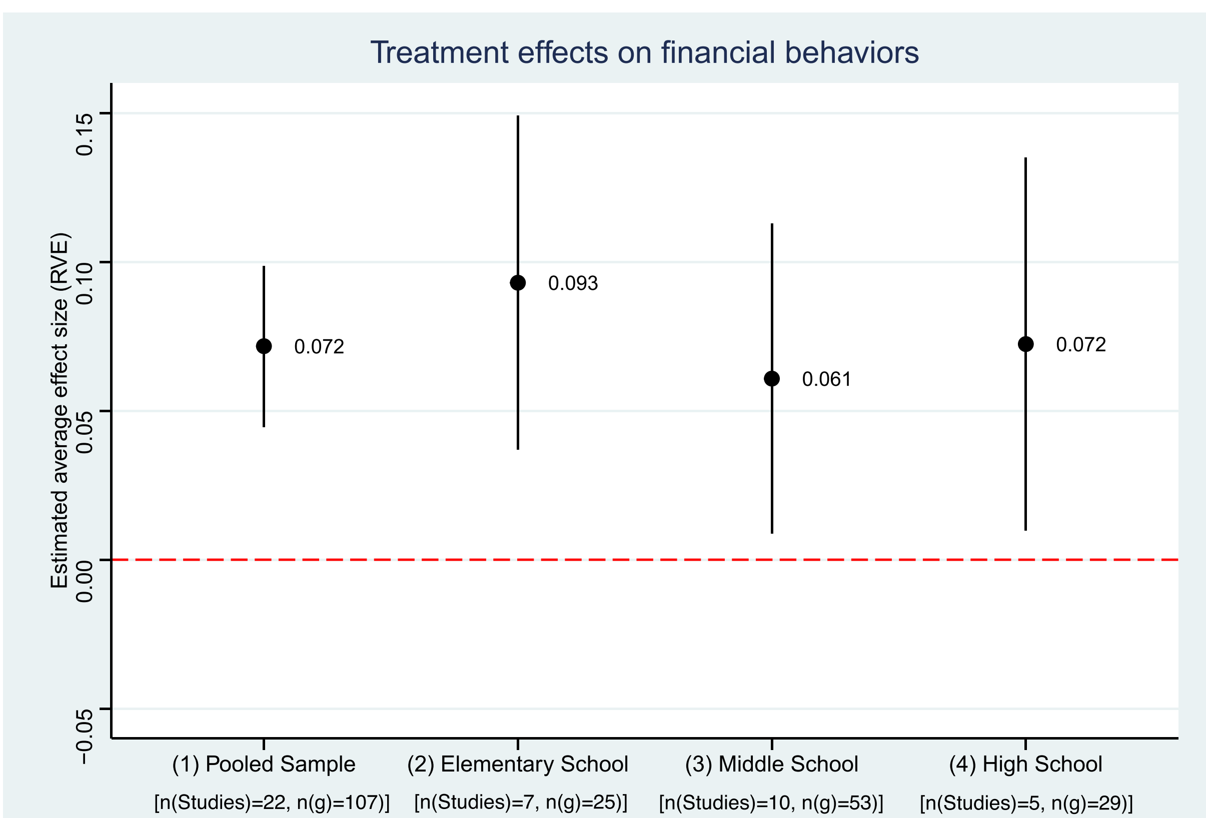

Fig. 3. Treatment effects on financial behaviors (RVE).
Notes: The figure shows the (weighted) average effects and 95% Cis estimated by RVE.

Next, we disaggregate the sample of studies with regard to the age of students. We find that treatment effects on financial knowledge are estimated to be highest among interventions in elementary schools (0.57 SDs) relative to interventions in middle school (0.16 SDs) and high schools (0.37 SD), as shown in Fig. 2. However, only the difference between elementary schools and middle schools is estimated to be statistically significant, while the 95 percent confidence interval of high school interventions includes the estimate for elementary school interventions. Note, that the result on middle-schools may be partly driven by the fact that the largest RCTs in the sample are with children in this age group (Bruhn et al., 2016; Frisancho, 2018) and it is known from meta-analyses of educational interventions that smaller studies may report inflated effect sizes (Cheung & Slavin, 2016). Nevertheless, it appears to be reassuring that effect sizes reported in studies with elementary school children appear to be substantial in magnitude (lower bound of 0.33 SDs). Turning to the effect on financial behavior (see Fig. 3), we find no significant differences between the different age groups, but the estimated average effects follow a similar pattern with the treatment effect in elementary schools being estimated to be larger than effect sizes in middle or high schools. Thus, the evidence suggests that increases in financial knowledge and changes in financial behavior can be achieved irrespective of the age of the students.

Next, we compare these effect sizes to earlier findings in the literature and to effect sizes realized in interventions from other educational domains. In the financial education literature, Fernandes et al. (2014) are the first to apply a quantitative meta-analysis. They arrive at small (weighted) average effect sizes on financial behavior for interventions (about $g = 0.02$ for 15 estimates from RCTs and about $g = 0.07$ including 75 estimates from quasi-experiments).[4] Thus, the effect on behavior among students is higher or at least near identical to the findings from the limited number of early experiments on adults. Regarding treatment effects on financial knowledge, Fernandes et al. (2014) state that 12 papers in their sample report an average effect of about 0.13 SD units.[5] This result, however, is an obvious contrast to the results of our meta-analysis on students where the estimated average effect is more than twice as large. Thus, the assertion that "[…] financial education yields surprisingly weak changes in financial knowledge […]" (Fernandes et al., 2014, p.1867) does not hold in this sample of studies on children and youth and may be seen as a particular result of the sample studied by Fernandes et al. (2014).

The second meta-analysis in the (adult) financial education literature uses a slightly different approach comparing only studies that measure effects on identical outcomes (Miller et al., 2015). This study does not quantify effects on financial knowledge but provides estimates on various financial behaviors reported in studies on adult financial education programs.

Third, the most-recent meta-analysis covering the largest number of interventions provides evidence of an average effect of about 0.2 SD units on financial knowledge, and about 0.09 on financial behaviors in a sample including many studies on adult financial education programs (Kaiser & Menkhoff, 2017). Thus, effect sizes on financial knowledge appear to be larger for programs that focus exclusively on children and youth than for adults. In contrast, effect sizes that measure changes in financial behavior appear to be slightly smaller for children and youth than for adults. However, these differences in programs for children and youth versus adults are not statistically significant.

How do these effect sizes compare to learning that takes place in other domains? Comparing effect sizes across disciplines and research questions is difficult, however, there exist some normative and empirical benchmarks with regard to learning outcomes in school: Hill et al. (2008) provide examples of effect sizes on reading and mathematics achievement. They document typical knowledge gains from year to year in school (in the absence of a particular intervention), achievement gaps with regard to specific subgroups, as well as a summary of effect sizes realized by interventions in these domains. If one compares their descriptive evidence to the result of our synthesis, financial education has near identical effect sizes on average, as reported in 76 meta-analyses of various educational interventions (0.22 to 0.27 SD units) (cf. Hill et al., 2008, p.176).

To make another empirical comparison: The average effect size realized by financial education appears to be of similar magnitude as the estimated increase in learning in mathematics in the transition from grade 9 to 10 (0.25 SD) or of similar size as the increase in reading achievement occurring in the transition from grade 7 to 8 (0.26 SD) (Hill et al., 2008, p.173). Thus, one can argue that these knowledge gains are indeed of high practical significance.

4.2. Robustness checks

To probe the robustness of our findings, we first restrict the sample of studies to randomized experiments only and also consider alternative models in the estimation of treatment average effects.

4.2.1. Restricting the sample to RCTs

The first row of Fig. 4 shows results for our preferred random-effects model for the full sample of studies and the disaggregated set of randomized experiments. Treatment effects on financial knowledge reported in RCTs are estimated to be much smaller than in the sample of quasi-experimental studies. The weighted average effect in RCTs results in 0.19 SD units. This finding confirms the results on adult financial education programs studied in Kaiser and Menkhoff (2017) as well as findings from other educational domains where non-randomized impact evaluations also appear to report inflated effect size estimates (Cheung & Slavin, 2016). While the effect reported in RCTs appears to be approximately 42 percent smaller, this effect size is still relatively large in magnitude and practically relevant. Compared to effects reported in randomized experiments in math and reading, the treatment effect on financial knowledge (0.19 SD) would fall in the range of the 70th to 80th percentile of all treatment effects reported in 242 studies (Kraft, 2018). Fryer (2016) reports an average effect of all kinds of school-based interventions on math and reading test scores of 0.05 and 0.07, respectively. Turning to the effects on financial behaviors, we find that both study types show very similar results which are not statistically significant from each other.

4.2.2. Simple average (OLS)

As an alternative strategy, we estimate the unweighted average effect in a simple OLS framework and cluster the standard errors at the study-level to account for the nested structure of the data. We find that the estimate is very similar to the more sophisticated random-effects model.

4.2.3. Fixed effect models (WLS)

Next, we weight each observation by its inverse standard error (WLS 1/SE_g) or its inverse variance (1/Var_g), respectively. Thus, this model assumes one common true effect and strongly discounts relatively smaller studies due to their larger sampling errors. In the full set of studies, this assumption leads to a significantly deflated estimate of 0.22 or 0.17 SD units (versus 0.33 SD units) on financial knowledge, and 0.05 or 0.04 SD units (versus 0.07 SD units) on financial behaviors. Within the sample of RCTs, however, these WLS-models do not arrive at statistically significantly different estimates relative to the random-effects model and the 95 percent confidence intervals are considerably tighter than in the OLS or random-effects case.

4.2.4. Publication selection bias

Publication selection bias refers to the potential behavior of researchers and journal editors to favor statistically significant results and

[4] Note, that Fernandes et al. (2014) use partial correlations (r) as their effect sizes measure. We transform these to standardized mean differences $d = \frac{2r}{\sqrt{1-r^2}}$ and apply the bias correction factor to arrive at (g) ex post.

[5] See Fernandes et al. (2014), p. 1867: "In 12 papers reporting effects of interventions on both measured literacy (knowledge) and some downstream financial behavior, the interventions explained only 0.44% of the variance in financial knowledge", i.e. $\sqrt{r^2} = 0.066$.

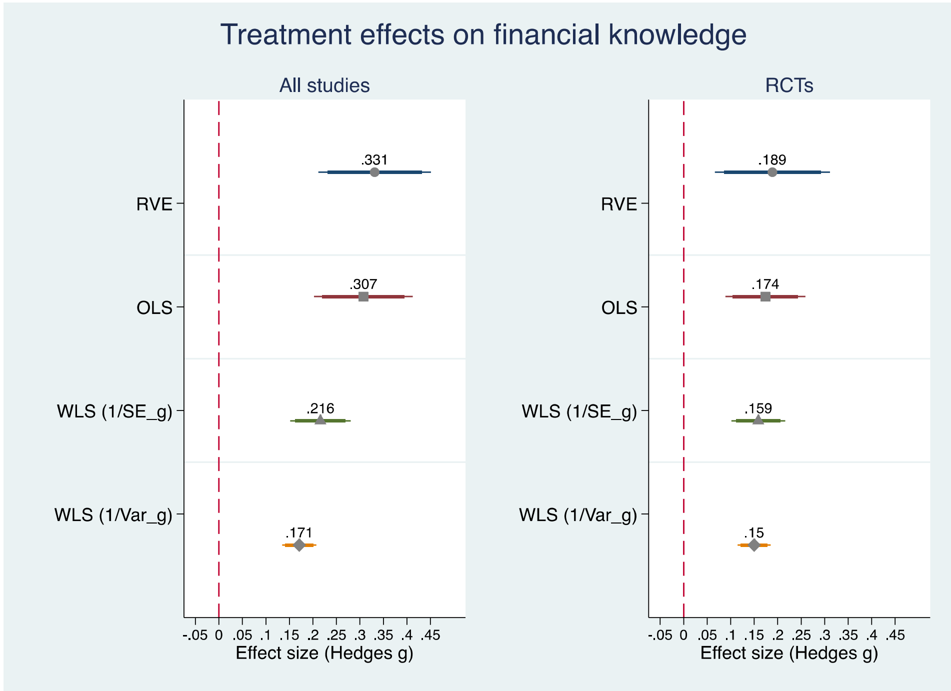

Fig. 4. Robustness exercises (financial knowledge).
Notes: The figure shows the (weighted) average effects and 90% and 95% Cis estimated by the different meta-analysis models. Number of observations for all (31) studies is $n = 70$ effect size estimates. Number of observations for the 14 RCTs is $n = 41$ effect size estimates.

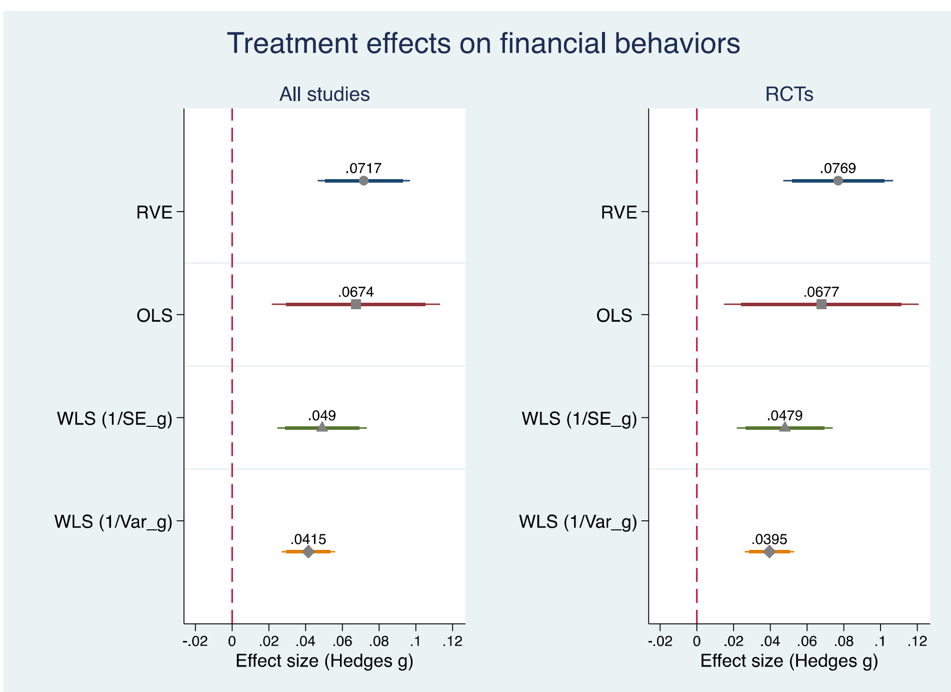

Fig. 5. Robustness exercises (financial behaviors).
Notes: The figure shows the (weighted) average effects and 90% and 95% Cis estimated by the different meta-analysis models. Number of observations for all (22) studies is $n = 107$ effect size estimates. Number of observations for the 16 RCTs is $n = 94$ effect size estimates.

not reporting estimates which do not pass tests for statistical significance. Given a single true empirical effect (which may be questioned due to the heterogeneity of treatments), the standard error of this estimate should be orthogonal to the reported effect sizes in a given literature. If this is not the case, we observe so-called funnel asymmetry. This tendency to underreport "undesired" estimates with large standard errors (especially in small studies) can lead to a biased assessment of the (weighted) average effect of a given literature. In the following, we test whether such a mechanism can be observed in the literature on school financial education.

In the presence of "publication selection", researchers and editors may favor the publication of empirical estimates that pass tests for conventional levels of statistical significance. When such a mechanism is present, the reported effect is (ceteris paribus) correlated with its

Table 2
Testing for publication selection bias.

Panel A: All studies	(1) FK Unadjusted	(2) FK FAT-PET	(3) FK PEESE	(4) FB Unadjusted	(5) FB FAT-PET	(6) FB PEESE
SE		2.493***			1.067	
		(0.588)			(0.709)	
SE^2			11.136***			8.963
			(2.916)			(5.291)
Average effect	0.216***	0.075**	0.147***	0.049***	0.017	0.036***
	(0.032)	(0.032)	(0.015)	(0.012)	(0.018)	(0.006)
R^2		0.235	0.136		0.094	0.038
n (Studies)	31	31	31	22	22	22
n (Effect sizes)	70	70	70	107	107	107
Panel B: RCTs						
SE		0.298			1.180	
		(0.869)			(0.833)	
SE^2			5.627			9.912
			(7.596)			(6.082)
Average effect	0.159***	0.155***	0.142***	0.048***	0.015	0.034***
	(0.027)	(0.047)	(0.015)	(0.012)	(0.021)	(0.006)
R^2		0.005	0.020		0.111	0.043
n (Studies)	14	14	14	16	16	16
n (Effect sizes)	41	41	41	94	94	94

Notes: Dependent variable is effect size (Hedges g). Robust standard errors clustered at the study-level in parentheses. ***, **, and * denote significance at the one percent, five percent, and ten percent level.

standard error (Stanley & Doucouliagos, 2012, p.61). The intuition of this method is to "correct" the estimate of the average empirical effect (the intercept of a given meta-regression model). In order to arrive at an estimate of a genuine empirical effect Stanley and Doucouliagos (2012) suggest including the standard error (PET) or the variance (PEESE) as a predictor of effect sizes and estimate the model by employing an unrestricted weighted least squares procedure using inverse variance weights.

Table 2 shows results from these tests for publication selection bias and its correction, where Panel A considers all studies while Panel B considers RCTs only. Thus the ordering and estimated models are the same in both panels, which differ only with respect to the sample. Column 1 of Table 2 shows the unadjusted (weighted average) effect on financial knowledge. In the next step, column 2 introduces the standard error of each estimate as a regressor (funnel asymmetry test) (FAT) and precision effect testing (PET)) and indicates funnel asymmetry regarding the reported effects on financial knowledge. Thus, column 3 applies the correction proposed by Stanley and Doucouliagos (2012) and includes the variance of each estimate and weighting each effect size estimate with its inverse variance. Applying this correction still leads to a statistically significant estimated (weighted) average effect of financial education on financial knowledge (0.147 SDs). Turning to effect sizes on financial behavior, column (4) repeats the WLS result from Table 2 for comparison. The PET estimate (column 5) suggests that there may be no empirical effect (just selection) while the PEESE estimate arrives at a significant effect of about 0.036 SD units and an insignificant coefficient on the variance of the estimate.

The results for the sub-sample of RCTs shown in Panel B of Table 2 qualitatively confirms the results for all studies. However, it is interesting to note that – regarding financial knowledge (see the first three columns) – there is no publication selection bias in RCTs. Also the results on financial behaviors are not conclusive and do not indicate a strong publication bias, as in the larger set of studies.

We conclude from these examinations, that even when correcting for potential publication selection bias, the positive effects on financial knowledge remain statistically and economically significant. The small average positive effect on financial behavior, however, is less certain – as already suggested by the small effect size estimated in the other meta-analysis models. It may be noted that most of the literature, to which we have sometimes compared our results, does not apply these corrections for potential publication bias, and that this correction does not seem to be necessary if findings are based on RCTs which appear to be published regardless of the p-value of estimaes.

4.3. Heterogeneity in treatment effects

In the following we show results for subsamples of included studies. As discussed in Section 3 above, available studies either do not report all variables we are interested in or do not provide enough variation in our limited sample of studies, so that the group of variables to be considered is to a large extent driven by data availability. Thus, we primarily consider four types of variables: (i) the country per capita income of the study setting, (ii), the delay between financial education treatment and measurement of outcomes (iii), the intensity of financial education in hours taught and (iv) the class size of the respective financial education interventions.

4.3.1. Country per capita income

First, we split the sample between school interventions in high income economies and developing economies. We find that effect sizes on financial knowledge are significantly larger in developed economies (0.39 SD) relative to developing economies (0.14 SD). Treatment effects on financial behaviors, however, do not appear to be systematically different with regard to the country income.

4.3.2. Delay in measurement

Next, we study another potential determinant of estimated effectiveness, i.e., the effect of delayed measurement on the estimated size of the treatment effect. It may be expected that learning effects typically decay over time because people forget what they had learnt. This has been shown in the context of financial education by Fernandes et al. (2014); we demonstrate this effect also for our sample of studies covering only financial education in schools and show results for five groups of increasing delays in Table 3 and 4. While effects on knowledge appear to be larger (0.43 SD units) immediately after treatment (up to one month), the effect decays with increased delay between treatment and measurement of outcomes. While there are very few studies that measure effects of financial education with long time horizons after the treatment, the estimated effect is significant up to 3 months after treatment. The longer-run effects (after more than 12 months) are estimated to be positive but less certain, since only two studies exist that measure outcomes at such long delays. This result also arises with respect to changes in financial behaviors (see Table 4), however, estimated effects are not a strong function of delay. Since only three studies report on effects after six months, however, the long-term impact is uncertain. The 95-percent confidence intervals can neither rule out zero or relatively large effects (0.3 on financial knowledge and 0.4 on financial behavior) at delays of more than 12 months. Appendix Fig. A1 shows predicted values from an unrestricted weighted least squares meta-regression (using inverse variance weights) of effect size on delay and squared delay while controlling for the variance of the treatment effect estimate and both linear and quadratic intensity. Treatment effects on financial knowledge are predicted to be significantly larger than zero up to 7 months after the intervention took place. The long-term effect of financial education, again, is predicted to be uncertain as very few studies report on long time horizons after the treatment. While the point estimates remain positive even at a delay of 80 weeks, the extreme degree of uncertainty can be recognized from the large confidence bands. Thus, we cannot say that financial education in school is effective with a delayed measurement of 7 months and longer, but we also cannot say the opposite. The 95% confidence intervals at long delays (i.e., over 30–40 weeks) can neither rule out zero-effects nor an increase of effectiveness at longer time horizons, i.e., the “long-term” effect on financial behavior may be effectively zero or over 10% of a standard deviation (0.1 SDs) after 80 or more weeks. Unless the literature provides more long-term assessments of financial education programs, this relationship will remain unclear.

4.3.3. *Intensity*

Next, we investigate the effect of higher intensity (hours taught) on the estimated treatment effects. It has been hypothesized by Miller et al. (2015) that the effect increases with intensity and that this increase declines with intensity, reflecting declining marginal returns of education. While very brief interventions (up to one hour) show merely effect sizes of 0.14 SD units, higher intensity interventions (up to 10 h) show effect size estimates of up to 0.4 SD units (see Table 3). While there are strong marginal gains from increasing intensity from one to 15 h, there appear to be declining marginal returns to increased intensity, since even the consideration of the most intensive interventions (up to 90 or 150 h) does not increase the estimated weighted average effect size beyond this threshold. Modeling this relationship in a meta-regression with linear intensity and quadratic intensity as predictors (while controlling for the variance of the treatment effect estimate) also results in predicted values showing declining marginal returns to increased intensity (cf. Fig. A2 in the Appendix). Treatment effects on financial behaviors (see Table 4), in contrast, appear to be largely independent from intensity in school settings. While very brief interventions (less than one to five hours) produce insignificant effects, the marginal gains of increasing intensity appear less strong than with regard to increases in financial knowledge.

4.3.4. *Class size*

Finally, we study the relationship between the student to teacher ratio (class size) and estimated treatment effects, covered towards the bottom of Table 3. Effects on financial knowledge are estimated to be larger in smaller groups of students, if one looks at the effect size and goes from class size below 15 up to class sizes of more than 20 students. Beyond this, i.e., whether class sizes are larger than 25 or more, there is obvious decline in the effect size. However, the relationship may only be regarded as suggestive because most estimates are not significantly different from each other. Effects of class-size on effect sizes on behavior (Table 4) are relatively inconclusive with very large confidence intervals for the few studies with relatively large class-sizes.

Table 3
Heterogeneity in treatment effects on *financial knowledge*.

Subgroup	Effect size (g)	SE	95% CI Lower bound	95% CI Upper bound	n(Studies)	n(effects)
(a) By country income						
High income economies	0.3881	0.0721	0.2394	0.5368	26	56
Developing economies	0.1376	0.0495	−0.0005	0.2757	5	14
(b) By delay between treatment and measurement of outcomes						
Delay of < 1 month	0.4322	0.0976	0.0976	0.6397	17	30
Delay of ≥ 1 month	0.2050	0.0521	0.0886	0.3214	12	30
Delay of ≥ 3 months	0.1032	0.0298	0.0216	0.1848	6	14
Delay of ≥ 6 months	0.0591	0.0582	−0.4073	0.5255	3	4
Delay of ≥ 12 months	0.1249	0.0138	−0.0501	0.3000	2	2
(c) By intensity of treatment						
Intensity ≤ 1 h	0.1431	0.0148	0.0871	0.1991	5	10
Intensity ≤ 5 h	0.1683	0.0285	0.1025	0.2341	11	22
Intensity ≤ 10 h	0.4035	0.0851	0.2235	0.5836	18	45
Intensity ≤ 15 h	0.3890	0.0721	0.2391	0.5389	23	54
Intensity ≤ 20 h	0.3467	0.0694	0.2038	0.4895	27	63
Intensity ≤ 25 h	0.3467	0.0694	0.2038	0.4895	27	63
Intensity ≤ 30 h	0.3467	0.0694	0.2038	0.4895	27	63
Intensity ≤ 35 *h*3	0.3467	0.0694	0.2038	0.4895	27	63
Intensity ≤ 40 h	0.3336	0.0663	0.1972	0.4699	27	63
Intensity ≤ 90 h	0.3323	0.0636	0.2018	0.4629	29	65
Intensity ≤ 150 h	0.3235	0.0616	0.1972	0.4499	30	67
(e) By class size						
Class size < 15	0.4054	0.1448	−1.4342	2.2449	2	3
Class size ≥ 15	0.2739	0.0715	0.1227	0.4251	19	46
Class size ≥ 20	0.2087	0.0764	0.0398	0.3775	12	34
Class size ≥ 25	0.1378	0.0480	0.0250	0.2506	9	29
Class size ≥ 30	0.1666	0.0318	0.0765	0.2567	5	19
Class size ≥ 35	0.1428	0.0144	0.0608	0.2248	3	9

Notes: Results from RVE (Tanner-Smith & Tipton, 2014).

Table 4
Heterogeneity in treatment effects on *financial behaviors*.

Subgroup	Effect size (g)	SE	95% CI Lower bound	95% CI Upper bound	n(Studies)	n(effects)
(a) By country income						
High income economies	0.0642	0.0165	0.0279	0.1004	15	36
Developing economies	0.0923	0.0269	0.0247	0.1600	7	71
(b) By delay between treatment and measurement of outcomes						
Delay of $<$ 1 month	0.0689	0.0139	0.0364	0.1014	11	48
Delay of $\geq$ 1 month	0.0760	0.0271	0.0146	0.1373	11	58
Delay of $\geq$ 3 months	0.0671	0.0241	0.0045	0.1297	6	40
Delay of $\geq$ 6 months	0.0720	0.0211	−0.0206	0.1646	3	20
Delay of $\geq$ 12 months	0.0633	0.0288	−0.3028	0.4294	2	17
(c) By intensity of treatment						
Intensity $\leq$ 1 h	0.0656	0.0305	−0.0343	0.1655	5	8
Intensity $\leq$ 5 h	0.0429	0.0214	−0.0092	0.0949	10	25
Intensity $\leq$ 10 h	0.0543	0.0181	0.0125	0.0962	13	48
Intensity $\leq$ 15 h	0.0583	0.0156	0.0239	0.0928	15	56
Intensity $\leq$ 20 h	0.0693	0.0121	0.0435	0.0951	20	83
Intensity $\leq$ 25 h	0.0693	0.0121	0.0435	0.0951	20	83
Intensity $\leq$ 30 h	0.0693	0.0121	0.0435	0.0951	20	83
Intensity $\leq$ 35 $h3$	0.0664	0.0115	0.0421	0.0906	21	93
Intensity $\leq$ 40 h	0.0720	0.0129	0.0449	0.0991	22	96
Intensity $\leq$ 90 h	0.0720	0.0129	0.0449	0.0991	22	96
Intensity $\leq$ 150 h	0.0717	0.0128	0.0446	0.0987	22	107
(d) By class size						
Class size $<$ 15	–	–	–	–	–	–
Class size $\geq$ 15	0.0517	0.0145	0.0198	0.0836	14	86
Class size $\geq$ 20	0.0652	0.0168	0.0260	0.1044	10	76
Class size $\geq$ 25	0.0640	0.0243	0.0019	0.1261	7	70
Class size $\geq$ 30	0.0730	0.0462	−0.0916	0.2376	4	24
Class size $\geq$ 35	0.0730	0.0462	−0.0916	0.2376	4	24

Notes: Results from RVE (Tanner-Smith & Tipton, 2014).

5. Policy conclusions

We start this concluding section with the caveat that the number and heterogeneity of available studies allows drawing conclusions only with caution. This said, we present our conclusions in the following from a policy perspective: What can policy makers learn from the meta-analysis being presented so far, which elements could be integrated into an effective program, which elements may be added beyond the scope of this study, and what does this imply for the discussion of principal alternatives?

Meta-analysis lessons: The main lesson is that financial education seems to be quite successful in increasing financial knowledge among school students. This result is robust irrespective of the meta-analytic model and when accounting for potential publication selection bias in the financial education literature. In particular, financial education in schools has a statistically and economically significant effect also when the most rigorous type of impact evaluation is conducted, i.e., in the sub-sample of RCTs. This also holds if the intended outcome is a change in financial behavior, however, the degree of effectiveness is smaller. When compared to all kinds of financial education (whether in schools or not), the effects on knowledge are possibly larger while those on behavior tend to be relatively smaller (Kaiser & Menkhoff, 2017), although the differences to effect sizes in the general literature are not statistically significant.

Design elements of effective financial education in schools: There are few potential determinants of effectiveness which can be directly tested in this paper due to a lack of better data. The two potential determinants where enough information is available are the grade (elementary vs. middle school vs. high school) and the intensity of education. While we find that effectiveness is highest at elementary schools, this does not imply that financial education should necessarily be limited to these early ages. Rather, this finding mirrors the results of meta-analyses in other educational domains (Hill et al., 2008; Fryer 2016) and it seems encouraging that financial education is successful in increasing knowledge and changing behavior irrespective of student age. Also, regarding the positive impact of increased intensity, the consequence is not as straight forward as it may look like, i.e., to make programs as comprehensive as possible. We rather suggest thinking about a format with limited content that is taught for up to 20–40 h which translates into roughly one or two teaching hours per a half year of schooling. However, the desirable intensity also depends on the comprehensiveness and duration of the program.

Finally, it seems advisable to think about reducing class sizes, although we would need more research in this respect to be sure about a recommendation and the cost-effectiveness of such an approach. In addition, there are further insights from the general literature on financial education programs which may be applied to schools as well.

Design elements not covered in this meta-analysis: One important element which could not been tested here is the impact of a so-called ‘teachable moment’. It has been shown for studies covering such an effect that the additional positive impact on financial behaviors may be in the order of magnitude of 0.05 to 0.07 standard deviations and thus quite sizable (Kaiser & Menkhoff, 2017). This suggests considering proper teachable moments during the process of life-long financial education. Moreover, there is evidence that education that is more entertaining or personalized has more impact on financial behavior (Berg & Zia, 2017; Carpena, Cole, Shapiro & Zia, 2017). Finally, it appears that those programs that employ design elements resembling 'active learning' (e.g., simulations and experimental learning) may yield higher effect sizes (see Amagir et al., 2018; Kaiser & Menkhoff, 2018). All these are elements which may contribute to increasing effectiveness of financial education.

*Assessment relative to alternatives:*The finding of successful financial education in schools is a necessary but not sufficient condition that respective programs should be implemented. Opposing positions either emphasize to regulate the financial sector in a way that financial education becomes less necessary or favor a more general education in economics, mathematics or statistics over more narrow financial education (see Brown et al., 2016). A discussion of these positions is beyond the scope of this paper. However, we have shown that financial education impacts financial knowledge and behavior and evidence suggests that financial education can also have significantly positive externalities, such as positive effects on the financial knowledge of parents (Bruhn et al., 2016) and of teachers (Frisancho, 2018). Additionally, recent experiments show that financial education has an impact on intertemporal decision making among children and youth, leading to more consistent and more patient intertemporal choices (see Alan & Ertac, 2018; Bover, Hospido & Villanueva, 2018; Frisancho, 2018; Lührmann et al., 2018; Migheli & Moscarola, 2017). Thus, financial education provided early in the life-cycle may have beneficial outcomes with regard to debt taking or long-term savings and may reach even beyond the financial domain. Thus, financial education improves the understanding of financial affairs but seems to have broader welfare implications, similar to other forms of education.

Overall, academic research alone cannot answer the policy question whether financial education in schools should be introduced at all or the extent to which it should be implemented. What can be said, however, given the current evidence, is that financial education is as effective as education in other domains and that effect sizes are substantial in magnitude at around 20–40 h of total instruction. Despite this encouraging situation, we want to emphasize that more could be done in order to increase effectiveness of financial education and that more thorough documentation of such efforts within empirical studies would be crucial to gain deeper insights in future narrative reviews or meta-analyses. While most studies report necessary information to arrive at standardized effect sizes, details about the financial education programs that may be driving a large part of the residual heterogeneity in treatment effects are often missing from published reports. Thus, we encourage authors to include more information about the underlying programs, especially regarding the implementation (e.g., information about the teachers, curricula, media, and the cost of the intervention) (cf. Miller et al., 2015; Appendix).

Appendix. : List of studies included in the meta-analysis

Alan, S. and Ertac, S. (2018). Fostering Patience in the Classroom: Results from Randomized Educational Intervention. *Journal of Political Economy*, 126(5), 1865–1911.

Angel, S. (2018). Smart Tools? A Randomized Controlled Trial on the Impact of Three Different Media Tools on Personal Finance. *Journal of Behavioral and Experimental Economics*, 74, 104–111.

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.

Batty, M., Collins, M., O'Rourke, C. and Elizabeth, O. (2017). Evaluating Experiential Financial Capability Education. A Field Study of My Classroom Economy. *Working Paper.*

Becchetti, L. and Pisani, F. (2012). Financial Education on Secondary School Students: The Randomized Experiment Revisited. *Facolta di Economia di Forli, Working Paper No. 98.*

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.

Berry, J., Karlan, D., and Pradhan, M. (2018). The Impact of Financial Education for Youth in Ghana. *World Development*, 102, 71–89.

Bover, O., Hospido, L., and Villanueva, E. (2018). The Impact of High School Financial Education on Financial Knowledge and Choices: Evidence from a Randomized Trial in Spain. *IZA Discussion Papers 11,265.*

Bruhn, M., Leão, L. de S., 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*, 8(4), 256–295.

Carlin, B.I. and Robinson, D.T. (2012). What Does Financial Literacy Training Teach Us? *Journal of Economic Education*, 43(3), 235–247.

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.

Frisancho, V. (2018). The Impact of School-Based Financial Education on High School Students and their Teachers: Experimental Evidence from Peru. Inter-American Development Bank Working Paper No. 871.

Furtado, I., Legovini, A., and Piza, C. (2017). How Early One Should Start Financial Education: Evidence from a Large-Scale Experiment. *World Bank, DIME Financial and PSD Program in Brief.*

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.

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.*

Grody, A. D., Grody, D., Kromann, E., and Sutliff, J. (2008). A Financial Literacy and Financial Services Program for Elementary School Grades-results of a Pilot Study. *Working Paper.*

Harter, C. L. and Harter, J. F. (2009). Assessing 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.

Hinojosa, T., Miller, S., Swanlund A, Hallberg K, Brown, M., and O'Brien, B. (2010). The Impact of the Stock Market Game on Financial Literacy and Mathematics Achievement. Results from a National Randomized Controlled Trial. *Working Paper. Evanston, IL: Society for Research on Educational Effectiveness.*

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.*

Kalmi, P. (2018). The Effects of Financial Education: Evidence from Finnish Lower Secondary Schools. *Economic Notes*, 47, 353–386.

Kalwij, A.S., Alessie, R., Dinkova, M., Schonewille, G., van der Schors, A., and van der Werf, M. (2017). The Effects of Financial Education on Financial Literacy and Savings Behavior: Evidence from a Controlled Field Experiment in Dutch Primary Schools. *Working Papers 17–05, Utrecht School of Economics.*

Langrehr, F. W. (1979). Consumer Education: Does It Change Students' Competencies and Attitudes? *Journal of Consumer Affairs*, 13(1), 41–53.

Lührmann, M., Serra-Garcia, M., and Winter, J. (2015). Teaching Teenagers in Finance: Does It Work? *Journal of Banking and Finance*, 54, 160–174.

Lührmann, M., Serra-Garcia, M., and Winter, J. (2018). The Impact of Financial Education on Adolescents' Intertemporal Choices. *American Economic Journal: Economic Policy*, 10(3), 309–332.

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.

Migheli, M. and Coda Moscarola, F. (2017). Gender Differences in Financial Education: Evidence from Primary School. *De Economist*, 165(3), 321–347.

Schug, M. C. and Hagedorn, E. A. (2004). Can Charter Schools Improve Financial and Economic Education? The Case of Milwaukee Urban League Academy of Business and Economics. *Journal of Private Enterprise,* 20, 100–111.

Shephard, D. D., Kaneza, Y. V., and Moclair, P. (2017). What Curriculum? Which Methods? A Cluster Randomized Controlled Trial of Social and Financial Education in Rwanda. *Children and Youth Services Review*, 82, 310–320.

Sherraden, M. S., Johnson, L., Guo, B., and Elliott, W. (2011). Financial Capability in Children: Effects of Participation in a School-based Financial Education and Savings Program. *Journal of Family and Economic Issues*, 32(3), 385–399.

Supanantaroek, S., Lensink, R., and Hansen, N. (2016). The Impact of Social and Financial Education on Savings Attitudes and Behavior Among Primary School Children in Uganda. *Evaluation Review, 41*(6), 511–541.

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.

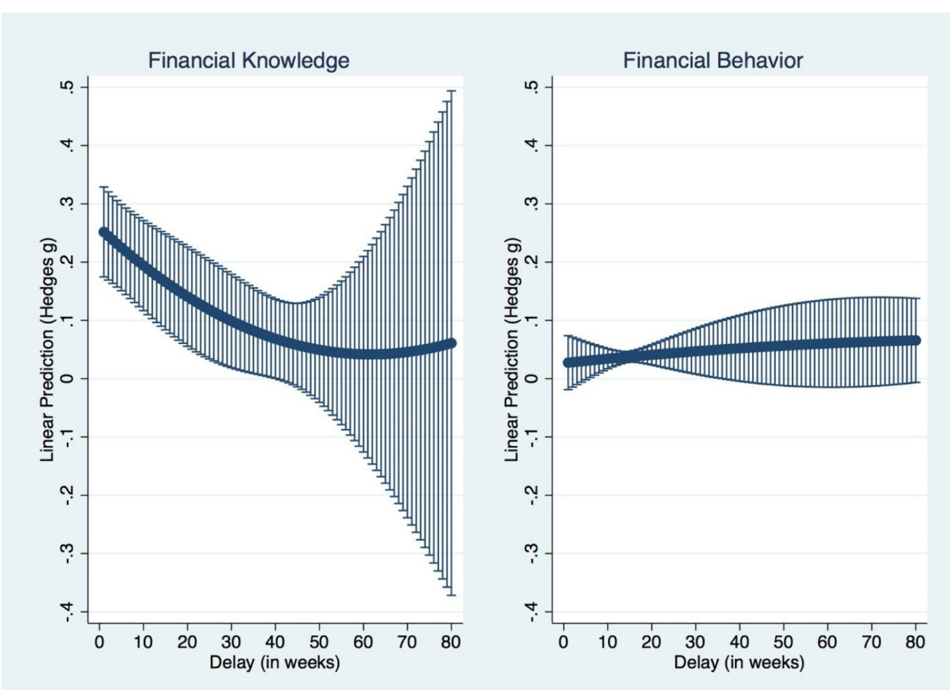

Fig. A1. Decreasing effect size with increasing delay of measurement.
Notes: These figures show the effect size of financial education treatments as a function of delay between treatment and measurement of outcomes (at average empirical intensity and controlling for the variance of the estimate in a unrestricted WLS regression with inverse variance weights (PEESE)). Delay is measured in weeks. The shaded areas cover the 95% confidence upper- and lower bounds.

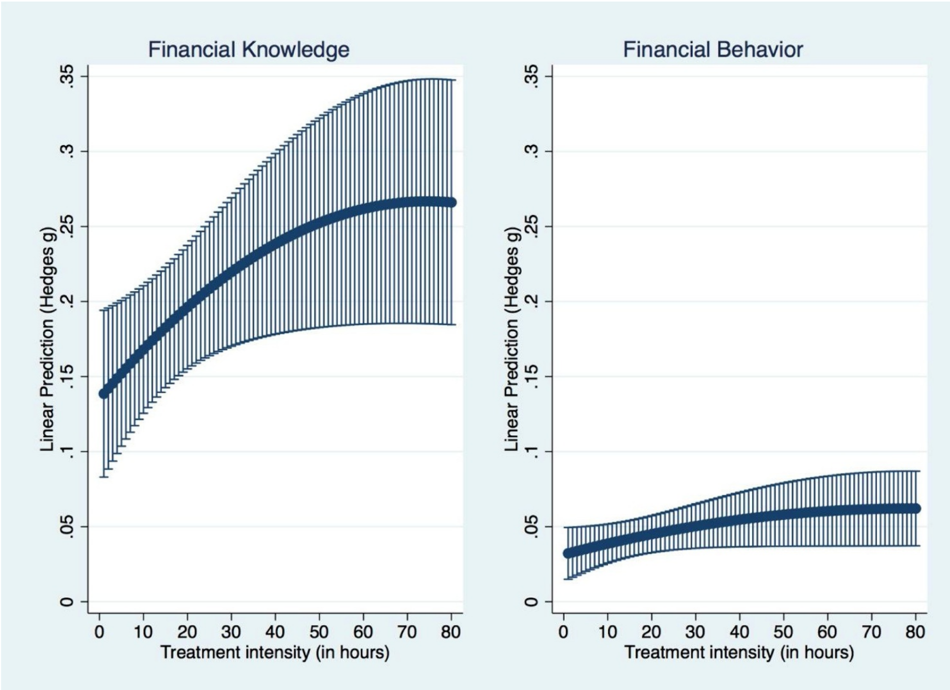

Fig. A2. Decreasing marginal returns to increased intensity.
Notes: These figures show the effect size of financial education treatments as a function of treatment intensity (controlling for the variance of the estimate in a unrestricted WLS regression with inverse variance weights (PEESE)). Intensity is measured in hours. The shaded areas cover the 95% confidence upper- and lower bounds.

Table A1
Overview of included experiments.

	Study	Country	RCT	Included in KM (2017)	Students (mean) age	Sample size	Outcomes coded
1	Alan and Ertac (2018)	Turkey	Yes	No	3rd and 4th grade (elementary school)	1970	D
2	Angel (2018)	Austria	Yes	No	18	296	A, D
3	Batty et al. (2015) [independent sample 1]	USA	Yes	Yes	Elementary school (4th and 5th graders)	703	A, C, D
4	Batty et al. (2015) [independent sample 2]	USA	Yes	Yes	Elementary school (4th and 5th graders)	277	A, C, D
5	Batty et al. (2017)	USA	Yes	No	9	1972	A, C, D
6	Becchetti and Pisani (2012)	Italy	Yes	No	High School	3820	A
7	Becchetti et al. (2013)	Italy	Yes	Yes	High School	1063	A, D
8	Berry et al. (2018)	Ghana	Yes	Yes (2015 WP)	11	5400	A, B, D
9	Bover et al. (2018)	Spain	Yes	No	15	3070	A, D
10	Bruhn et al. (2016)	Brazil	Yes	Yes	16	25,000	A, B, C, D
11	Carlin and Robinson (2012)	USA	No	Yes	16	1672	B, C, D, E
12	Chen and Heath (2012) [independent sample 1]	USA	No	Yes	NA (elementary)	1244	A
13	Chen and Heath (2012) [independent sample 2]	USA	No	Yes	NA (middle)	155	A
14	Frisancho (2018)	Peru	Yes	No	15	25,980	A, C, D
15	Furtado et al. (2017)	Brazil	Yes	No	12	14,655	A, D
16	Gill and Bhattacharya (2015)	USA	No	Yes	High School	159	A
17	Go et al. (2012)	USA	No	Yes	9 (4th and 5th graders)	403	A, C, D
18	Grody et al. (2008)	USA	No	No	Elementary school	31	A
19	Harter and Harter (2009)	USA	No	Yes	NA (Elementary, Middle, and High School)	2438	A
20	Harter and Harter (2010)	USA	No	Yes	17	730	A
21	Hinojosa et al. (2010)	USA	Yes	No	9 / 15	8594	A
22	Hospido et al. (2015)	Spain	No	Yes	15	1223	A
23	Kalmi (2018) [independent sample 1]	Finland	No	No	15	2386	A, D
24	Kalmi (2018) [independent sample 2]	Finland	No	No	15	2085	A, D
25	Kajwij et al. (2017)	Netherlands	Yes	No	10	1816	A, D
26	Lührmann et al. (2015)	Germany	No	Yes	14 (7th and 8th grade)	770	
27	Lührmann et al. (2018)	Germany	Yes	No	14 (7th and 8th grade)	914	A, D
28	Langrehr (1979)	USA	No	No	High School	110	A
29	Migheli and Moscarola (2017)	Italy	Yes	No	8 to 9 (Elementary School)	213	D
30	Mandell (2009a)	USA	No	Yes	High School	1279	D
31	Mandell (2009b)	USA	No	Yes	High School	1030	A
32	Mandell and Schmid-Klein (2009)	USA	No	Yes	High School	79	A
33	Schug and Hagedorn (2004)	USA	No	Yes	Middle School	109	A
34	Shephard et al. (2017)	Rwanda	Yes	No	15	1750	A, C, D
35	Sherraden et al. (2011)	USA	No	Yes	Elementary School	93	A
36	Supanataroek et al. (2016)	Uganda	Yes	Yes	13	1746	C, D
37	Walstad et al. (2010)	USA	No	Yes	High School	800	A

Table A2
Outcome definitions.

	Outcome category		Definition
A	*Financial knowledge* (+)		Raw score on financial knowledge test Indicator of scoring above a defined threshold Indicator of solving a test item correctly
B	*Credit behavior*		
	1)	Reduction of delinquencies within certain time frame (+)	Binary indicator
	2)	Lower cost of credit / interest rate (+)	Sum of real interest amount or interest rate and (if applicable) cost of fees
	3)	Any debt (−) / (+) (depending on intervention goal)	Binary indicator
	4)	Borrowing index (+)	Study-specific index of survey items to measure borrowing amount, frequency, and repayment
C	*Budgeting behavior*		
	1)	Having a written budget (+)	Binary indicator
	2)	Having a financial plan or long-term aspirations (+)	Binary indicator
	3)	Seeking information before making financial decisions (+)	Binary indicator
	4)	Self-rating of adherence to budget (+)	Study-specific scale
D	*Saving & retirement saving behavior*		
	1)	Amount of savings (+)	Continuous measure (or log) of savings amount (in currency or number of valuable assets) or categorical variable indicating amount within range Savings relative to income
	2)	Savings rate or savings within timeframe (+)	Amount over defined time-frame
	3)	Savings index (+)	Study-specific index of survey items designed to measure savings amount and frequency
	4)	Any savings (+)	Binary indicator
	5)	Has formal bank (savings) account (+)	Binary indicator
	6)	Amount saved in allocation task (+)	Continuous measure of amount saved in allocation task
	7)	Amount allocated to delayed payment date in experimental elicitation task (+)	Continuous measure of amount delayed to be paid out at a later date within an experimental elicitation task
E	*Insurance behavior*		
	1)	Any formal insurance (hypothetical task) (+)	Binary indicator

References

Alan, S., & Ertac, S. (2018). Fostering patience in the classroom: Results from randomized educational intervention. *Journal of Political Economy, 126*(5), 1865–1911.

Amagir, A., Groot, W., Maassen van den Brink, H., & Wilschut, A. (2018). A review of financial-literacy education programs for children and adolescents. *Citizenship, Social and Economics Education, 17*(1), 56–80.

Berg, G., & Zia, B. (2017). Harnessing emotional connections to improve financial decisions. evaluating the impact of financial education in mainstream media. *Journal of the European Economic Association, 15*(5), 1025–1055.

Bernheim, B. D., Garrett, D. M., & 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.

Blue, L., Grootenboer, P., & 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.

Bover, O., Hospido, L., & Villanueva, E. (2018). The impact of high school financial education on financial knowledge and choices: Evidence from a randomized trial in spain. *IZA Discussion Papers 11265*.

Brown, M., Grigsby, J., van der Klaauw, W., Wen, J., & Zafar, B. (2016). Financial education and the debt behavior of the young. *Review of Financial Studies, 29*(9), 2490–2522.

Bruhn, M., Ibarra, G. L., & 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., Leão, L., de, S., Legovini, A., Marchetti, R., & Zia, B. (2016). The impact of high school financial education: Evidence from a large-scale evaluation in Brazil. *American Economic Journal: Applied Economics, 8*(4), 256–295.

Card, D., Kluve, J., & Weber, A. (2017). What works? A meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association, 16*(3), 894–931.

Carpena, F., Cole, S., Shapiro, J., & Zia, B. (2017). The ABCs of financial education. Experimental evidence on attitudes, behavior, and cognitive biases. Management Science,forthcoming, 10.1287/mnsc.2017.2819.

Chetty, Raj, Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*(9), 2633–2679.

Cheung, A. C., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher, 45*(5), 283–292.

Cole, S., Paulson, A., & Shastry, G. K. (2016). High school curriculum and financial outcomes: The impact of mandated personal finance and mathematics courses. *Journal of Human Resources, 51*(3), 656–698.

Collins, J. M., & Odders-White, E. (2015). A framework for developing and testing financial capability education programs targeted to elementary schools. *Journal of Economic Education, 46*(1), 105–120.

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

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

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

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

Fox, J., Bartholomae, S., & Lee, J. (2005). Building the case for financial education. *Journal of Consumer Affairs, 39*(1), 195–214.

Frisancho, V. (2018). The impact of school-based financial education on high school students and their teachers: Experimental evidence from Peru. *Inter-American Development Bank Working Paper No. 871*.

Fryer, R. G. (2016). The production of human capital in developed countries: evidence from 196 randomized field experiments. *NBER Working Paper No. 22130*.

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

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics, 95*(7–8), 798–812.

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

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65.

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

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical Benchmarks for Interpreting Effect Sizes in Research. *Child Development Perspectives, 2*(3), 172–177.

Kaiser, T., & Menkhoff, L. (2017). Does financial education impact financial behavior, and if so, when? *World Bank Economic Review, 31*(3), 611–630.

Kaiser, T., & Menkhoff, L. (2018). Active learning fosters financial behavior: Experimental evidence. *DIW Discussion Paper No. 1743*.

Kraft, M. A. (2018). Interpreting Effect Sizes of Education Interventions. *Brown University Working Paper*.

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

Lührmann, M., Serra-Garcia, M., & Winter, J. (2018). The impact of financial education on adolescents' intertemporal choices. *American Economic Journal: Economic Policy, 10*(3), 309–332.

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

Lusardi, A., Mitchell, O. S., & Curto, V. (2010). Financial literacy among the young. *Journal of Consumer Affairs, 44*(2), 358–380.

Lusardi, A., Samek, A. S., Kapteyn, A., Glinert, L., Hung, A., & Heinberg, A. (2017). Visual tools and narratives: New ways to improve financial literacy. *Journal of Pension Economics and Finance, 16*(3), 297–323.

Migheli, M., & Coda Moscarola, F. (2017). Gender differences in financial education: Evidence from primary school. *De Economist, 165*(3), 321–347.

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

OECD (2015). National strategies for financial education. OECD/INFE Policy Handbook, https://www.oecd.org/finance/National-Strategies-Financial-Education-Policy-Handbook.pdf.
Peng, T.-C. M., Bartholomae, S., Fox, J. J., & 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.
Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review: Papers and Proceedings, 94*(2), 247–252.
Stanley, T. D. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives, 15*(3), 131–150.
Stanley, T. D., & Doucouliagos, H. (2012). *Meta-Regression analysis in economics and business.* New York: Routledge.
Stanley, T. D., & Doucouliagos, H. (2015). Neither fixed nor random: Weighted least squares meta-analysis. *Statistics in Medicine, 34*(13), 2115–2127.
Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in stata and SPSS. *Research Synthesis Methods, 5*(1), 13–30.
Tennyson, S., & Nguyen, C. (2001). State curriculum mandates and student knowledge of personal finance. *Journal of Consumer Affairs, 35*(2), 241–262.
Urban, C., Schmeiser, M., Michael Collins, J., & Brown, A. (2018). The effects of high school personal financial education policies on financial behavior. *Economics of Education Review*. https://doi.org/10.1016/j.econedurev.2018.03.006 forthcoming.
Walstad, W., Urban, C., Asarta, C. J., Breitbach, E., Bosshardt, W., & Heath, J. (2017). Perspectives on evaluation in financial education: Landscape, issues, and studies. *Journal of Economic Education, 48*(2), 93–112.
Willis, L. E. (2011). The financial education fallacy. *American Economic Review: Papers and Proceedings, 101*(3), 429–434.
Xu, L., & 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*.