



Contents lists available at ScienceDirect

Journal of Financial Economics

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

Financial education affects financial knowledge and downstream behaviors

Tim Kaiser^a, Annamaria Lusardi^{b,*}, Lukas Menkhoff^c, Carly Urban^d

^a University of Koblenz-Landau, Landau 76829, Germany

^b The George Washington University School of Business and NBER, Washington, DC 20052, USA

^c Humboldt Universität zu Berlin, German Institute for Economic Research (DIW Berlin), and Kiel Institute for the World Economy (IfW Kiel), 10108 Berlin, Germany

^d Montana State University and Institute for Labor Studies (IZA), Bozeman, MT 59717, USA

ARTICLE INFO

Article history:

Received 26 August 2020

Revised 7 June 2021

Accepted 5 July 2021

JEL classification:

D14 (personal finance)

G53 (financial literacy)

I21 (analysis of education)

Keywords:

Financial education

Financial literacy

Financial behavior

RCT

Meta-analysis

ABSTRACT

We study the rapidly growing literature on the causal effects of financial education programs in a meta-analysis of 76 randomized experiments with a total sample size of over 160,000 individuals. Many of these experiments are published in top economics and finance journals. The evidence shows that financial education programs have, on average, positive causal treatment effects on financial knowledge and downstream financial behaviors. Treatment effects are economically meaningful in size, similar to those realized by educational interventions in other domains, and robust to accounting for publication bias in the literature. We also discuss the cost-effectiveness of financial education interventions.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

The economic importance of financial literacy is documented in a large and growing empirical literature (e.g., Collins and O'Rourke, 2010; Xu and Zia, 2012; Hastings et al., 2013; Lusardi and Mitchell, 2014; Lusardi, 2019). Consequently, the implementation of national strategies promoting financial literacy and the design of financial education policies and school mandates have become a high priority for policymakers around the

world.¹ Many of the largest economies, including most OECD member countries, as well as India and China, have implemented policies enhancing financial education to promote financial inclusion and financial stability (OECD, 2015). Together, these financial education policies seek to reach more than five billion people in more than

¹ In October 2020, the Recommendation on Financial Literacy was adopted by the OECD Council. It presents a single, comprehensive financial literacy instrument to assist governments, other public authorities, and relevant stakeholders in their effort to design, implement, and evaluate financial literacy policies. More information is provided at <https://www.oecd.org/finance/OECD-Recommendation-on-Financial-Literacy.htm>. The European Commission, as well, has focused on financial literacy in its Action Plan for the Capital Market Union. More information is provided at https://ec.europa.eu/info/business-economy-euro/growth-and-investment/capital-markets-union/capital-markets-union-2020-action-plan_en.

* Corresponding author.

E-mail addresses: kaiser@uni-landau.de (T. Kaiser), alusardi@gwu.edu (A. Lusardi), lmenkhoff@diw.de (L. Menkhoff), carly.urban@montana.edu (C. Urban).

seventy countries, and the number of countries joining this effort continues to grow.

Despite the many initiatives to foster financial literacy, the effectiveness of financial education is debated in quite fundamental ways. Much of the debate stems from the fact that the limited number of early rigorous experimental impact evaluations sometimes showed muted effects, and these early findings have contributed to the perception of mixed evidence on the effectiveness of financial education beginning with Fernandes et al. (2014). However, there has been a recent increase in empirical studies on financial education, and about a third of them have been published in top economics and finance journals. To account for this increase, we evaluate the recent empirical evidence documented in randomized experiments and provide an updated and more rigorous analysis of the existing work.

Our main finding is clear-cut: financial education in 76 randomized experiments with a total sample size of more than 160,000 individuals displays positive causal treatment effects on financial knowledge and financial behaviors. The treatment effects on financial knowledge are similar in magnitude to the average effect sizes realized by educational interventions in other domains, such as math and reading (see Hill et al., 2008; Cheung and Slavin, 2016; Fryer, 2016; Kraft, 2020). The effect sizes of financial education on financial behaviors are comparable to those realized in behavior-change interventions in the health domain (e.g., Rooney and Murray, 1996; Portnoy et al., 2008; Noar et al., 2007) or behavior-change interventions aimed at fostering energy conserving behavior (e.g., Karlin et al., 2015).

Specifically, the estimated (weighted average) treatment effect is at least three times as large as the weighted average effect documented in the first meta-analysis of the literature, which examined 13 randomized controlled trials (RCTs) (Fernandes et al., 2014). The estimated effect of financial education interventions from our meta-analysis, which accounts for the possibility of cross-study heterogeneity, is more than five times as large as the original estimate.

We interpret the effect sizes resulting from these interventions and show that they are economically significant. Our results are robust, irrespective of the model used, when restricting the sample to only those RCTs that have been published in top economics and finance journals, when restricting the sample to only those studies with adequate power to identify small treatment effects, and when employing multiple methods to account for the possibility of publication bias favoring the publication of statistically significant results (Ioannidis et al., 2017; Andrews and Kasy, 2019).

In contrast to earlier studies, we do not find differences in treatment effects for low-income individuals and the general population. We also do not find evidence to support a rapid decay in the realized treatment effects, a finding that has been heavily cited, though we do not find support for the sustainability of long-run effects either.

For completeness and to assess the external validity of the findings, we also discuss the findings from recent evaluations of financial education mandates and school financial education programs operated at scale.

This paper makes four main contributions. First, we provide the most comprehensive analysis of the burgeoning work on financial education by using the most rigorous studies: RCTs. Our analysis includes the first formal identification of and correction for publication bias in the financial education literature by estimating the conditional publication probabilities of studies with statistically insignificant results. Second, we formally account for heterogeneity in programs and consider, for example, differences in target groups, duration, and the contextual features of interventions. Third, we provide the magnitudes of the effects in terms of their economic significance and consider the per-participant program costs. Fourth, we offer a thorough discussion of topics raised in previous work, i.e., how to assess the impact of financial education and whether education decays with time. We believe that this work can provide useful guidance for those evaluating financial education programs.

The paper has eight sections. In Section 2, we explain the types of programs that constitute financial education and summarize the state of the literature. Section 3 serves as a primer on statistical meta-analyses, and we describe our method. Section 4 presents descriptive statistics of our data, while Section 5 presents the main results. In Section 6, we discuss the heterogeneity in the meta-treatment effects based on program and sample characteristics. In Section 7, we describe the economic significance of effect sizes and the cost-effectiveness associated with these effects. We conclude in "Conclusion".

2. Background

2.1. What is financial education?

While financial education is often considered one classification of intervention, the variety of financial education programs is vast. Programs vary not just in duration, intensity, and populations served but also in the interventions themselves. In this section, we discuss some "traditional" and "non-traditional" financial education interventions considered in the papers included in our meta-analysis (see Table A.1 in Appendix A).

There are examples of financial education that an outsider to the field may expect financial education to look like. One common form of financial education is workplace financial education via benefit fairs. At these fairs, benefits officers are usually available to answer questions and provide information. Duflo and Saez, 2003 study benefit fairs for a university's non-faculty employees at which individuals had the opportunity to use a computer program that analyzed each participant's unique financial situation. A second form is via school-based personal finance coursework. Financial education in this setting has been studied extensively (e.g., Bruhn et al., 2016; Alan and Ertac, 2018; Frischno, 2018; Lührmann et al., 2018), though these studies vary in duration, methods, and curriculum. For example, Frischno (2018) considers financial education in a school setting, Bruhn et al. (2016) include parental education paired with student classroom learning, and Alan and Ertac (2018) focus on the development of non-

cognitive skills that allow children to understand future consequences of contemporaneous actions.

A third form, which is intensive and more costly, is the delivery of financial education via one-on-one financial coaching (Carpena et al., 2017). A fourth form of financial education is sometimes called an “educative nudge,” where the intervention just provides information (Choi et al., 2010; Boyer et al., 2020).

While the above description captures four relatively familiar forms of financial education, there are many other variations. We highlight three studies to give readers a sampling of some of these non-traditional approaches to financial education.

First, Flory (2018) uses extension agents to provide financial information to individuals in rural Malawi. Since extension agents are a common way of disseminating new information related to health or agriculture to those in rural villages, the mode of delivery benefits from an existing and trusted relationship. After being introduced to village residents, these agents engaged in one-on-one and group meetings once every two to three weeks for a few hours per village. They were available to answer questions and provide information about savings accounts and bank services, and importantly, the intervention occurred during and just after the harvest months, when there was an opportunity to save.

Second, Seshan and Yang (2014) constructed a workshop for male migrant workers in Qatar that shaped financial education around motivational content. The workshops included educational components, such as creating a savings plan, budgeting for both the migrant worker and the family back in India, and the pros and cons of a variety of investment options. The workshops were interactive and motivational: they encouraged participants to have a positive attitude toward work and life, establish better time management, and develop a good work ethic.

Third is the financial education intervention studied in Berg and Zia (2017). It is perhaps the most scalable and cost-effective approach in all the studies included, via the inclusion of personal finance messages in mainstream television. The South African soap opera features a main character falling into a debt trap. The storyline presents clear lessons about what not to do, as well as concrete steps to help one get out of debt. At the end of each episode, viewers were shown the toll-free number for the National Debt Mediation Association, which assists those struggling with debt. This intervention provides education but also targets non-cognitive channels and works through an emotional connection to the fictional characters depicted in the series.

The examples of financial education provided highlight the heterogeneity of programs and interventions. In addition to the interventions themselves being heterogeneous, study populations vary in age, socioeconomic status, and other attributes. Interventions include 9- and 10-year-old children (Alan and Ertac, 2018), migrants in Australia and New Zealand (Gibson et al., 2014), and farmers in rural Rwanda (Sayinzoga et al., 2016). In other words, careful design of financial education programs results in specific interventions for different populations, as each face unique financial challenges.

2.2. What is the state of the literature?

As evidence from rigorous empirical studies in a given field grows, there is a need to synthesize and integrate existing findings to reach conclusions that align with the research. Traditionally, economists have relied on narrative reviews in which experts on a given literature select and discuss the most relevant findings. The advantage of such an approach is that experts who are familiar with existing studies add value by summarizing, interpreting, and linking the most convincing (i.e., internally valid) studies. Examples of widely cited narrative reviews in the financial education literature are Collins and O'Rourke (2010), Xu and Zia (2012), Hastings et al. (2013), and Lusardi and Mitchell (2014).

As the empirical literatures grow larger, however, narrative literature reviews can become difficult, since it is hard to describe a large number of empirical estimates and discuss all of the possible sources of heterogeneity in the reported findings. Meta-analyses have thus become more common in economics when aggregating findings from many studies. Diverse topics covered in recent meta-analyses in economics include microcredit expansions (Meager, 2019), the effect of going to parent-preferred schools (Beuermann and Jackson, 2018), and experimental estimates of time-preference parameters (Imai et al., 2021). Meta-analyses can complement narrative reviews when there is a sufficiently large number of well-identified studies on the same empirical research question. A meta-analysis—a systematic, quantitative literature review—is well-suited for estimation of the average effects of a given program and study of heterogeneity in reported findings (Stanley, 2001).

As noted earlier, Fernandes et al. (2014) was the first meta-analysis performed in the field of financial education. Since that work, there have been three follow-up meta-analyses on financial education programs: Miller et al. (2015), Kaiser and Menkhoff (2017, 2020). These follow-up meta-analyses present a more nuanced view of financial education interventions than Fernandes et al. (2014) by including additional studies and accounting for differences in program design and outcomes. However, each follow-up study has limitations. Miller et al. (2015) focus their statistical meta-analysis on fewer than 20 studies, seven of which are RCTs, and emphasize examining the impact differences across outcomes. Kaiser and Menkhoff (2017) explore the correlates of effective financial education interventions in (quasi-) experiments, while Kaiser and Menkhoff (2020) focus on (quasi-) experimental evaluations of financial education in schools.

We make four major contributions to the literature. First, we update the dataset of treatment effects to create the most comprehensive database of financial education RCTs to date. As Fig. 1 shows, the field grew exponentially after the first meta-analysis (for which data collection ended in 2013) and has continued to grow since the most recent meta-analyses (for which data collection ended in 2016). We focus on what are considered the most rigorous sources of evidence, i.e., randomized experiments. RCTs provide more consistent internal validity than observational and quasi-experimental studies, especially

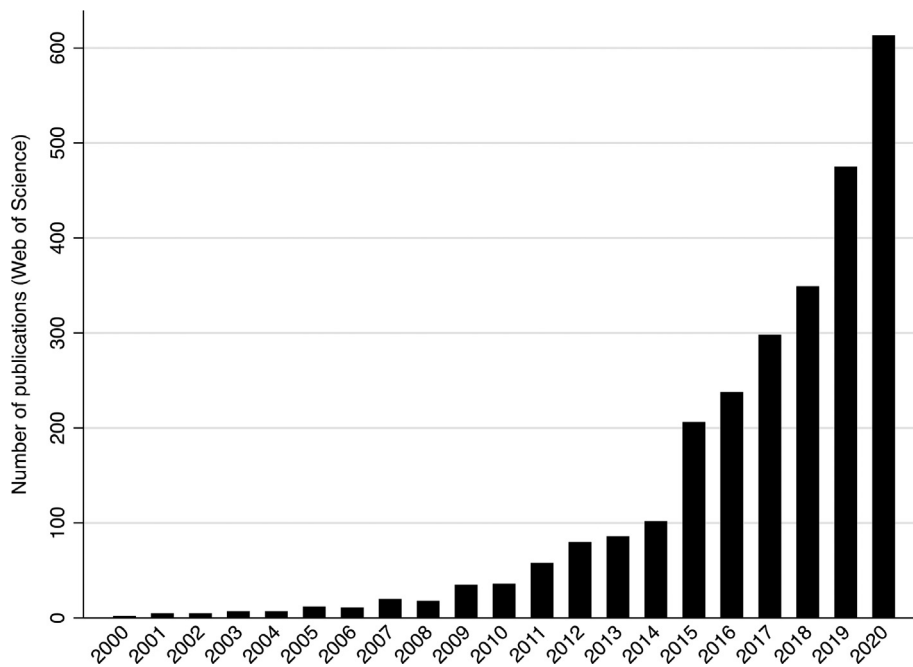


Fig. 1. Number of journal articles on financial literacy in the Web of Science per year.

Number of journal articles within the social science citation index (Web of Science) that include the term “financial literacy” in the title or the abstract. Data extracted from the Web of Science on March 3, 2021.

since there are no universally accepted instrumental variables for financial literacy, and one can debate whether non-randomized trials have used convincing empirical strategies addressing endogeneity or selection into treatment. Judging the quality of quasi-experimental studies and determining which to include or exclude from the meta-analysis gives researchers an additional degree of freedom that we wish to remove. Importantly, the number of RCTs has grown from just 13 in the first meta-analysis to 76 as of 2019. Our second contribution is an attempt to replicate the findings in the initial—and still most cited—meta-analysis (Fernandes et al., 2014). Third, we provide estimates more common in the meta-analysis literature, which account for heterogeneity in effect sizes across studies. Fourth, we carefully address the potential problem of publication bias and present novel evidence on the extent to which this mechanism is present in the financial education literature.

3. Methods

In this section, we describe our inclusion criteria for financial education papers (Section 3.1), the details we use in constructing our database of effect sizes (Section 3.2), and the specifics of our empirical model (Section 3.3).

3.1. Inclusion criteria

To draw general conclusions about a given literature, one has to conduct a systematic search of the literature and apply inclusion criteria that are defined ex ante. We conducted a search of all relevant databases for journal articles and working papers (see Appendix A for the list of

the studies and a summary of the data we extracted from those studies), ending our collection period in January 2019.² We apply three inclusion criteria: (1) studies reporting the causal effects of educational interventions designed to strengthen participants’ financial literacy and/or leading to behavior change in the area of personal finance; (2) studies using random assignment into treatment and control conditions; and (3) studies providing a quantitative assessment of an intervention impact that allows researchers to code an effect size estimate and its standard error. Where necessary information is partially missing, we consulted additional online resources related to the article or contacted the authors of the studies. We only consider the main results discussed in the text, and we do not code redundant effect sizes (e.g., effect sizes arising from other specifications of a given statistical model in the robustness section).

3.2. Constructing the database

Our analysis aggregates the treatment effects of financial education interventions into two main categories. First, we code the effect of financial education on *financial knowledge* (i.e., a measure of performance on a financial knowledge test) since improvement in knowledge is usually the primary goal of financial education (Hastings et al., 2013; Lusardi and Mitchell, 2014). We do not include self-assessments of changes in financial knowledge as an outcome, as they could be less reliable than test scores.

² This paper has gone through several revisions and the end of the collection period is when we started extracting and analyzing the data.

Second, we code the effect of financial education on financial behaviors. These behaviors can be further disaggregated into the following categories: borrowing, saving and investing, budgeting and planning, insurance, and remittances. Overall, it is useful to know which behavior is more easily impacted by financial education, and such analysis can provide relevant information to both academics and policymakers. Table A.2 in Appendix A provides an overview of the categories and definitions of outcome types.

We extract all available effect sizes per study on financial knowledge and behavioral outcomes. We include multiple estimates per study in the following cases. First, we extract multiple outcomes per study when there are estimates of treatment effects on various financial behaviors. Second, we extract multiple treatment effects per study if authors study treatment effects at different time points (i.e., short-term results vs. long-term follow-ups). Third, we include multiple estimates per experiment if more than one treatment has been randomly allocated to individuals. We only extract main treatment effects (average treatment effects) reported in the papers. We do not consider estimates reported in the “heterogeneity of treatment effects section” within papers, such as sample splits or interaction effects of binary indicators (e.g., gender, income, ability), with the treatment indicators. We aim to only consider intention-to-treat (ITT) effects unless these are not reported. If only local average treatment effects (LATE) or the treatment effect on the treated (TOT) are reported, we include these in our analysis and check for statistical differences, as described in Appendix B.³ This process leads to the inclusion of 76 independent, randomized experiments, which are described further in Section 4.

3.3. Empirical model

A major challenge in every meta-analysis lies in the heterogeneity of the studies being analyzed that must be accounted for. In the financial education literature, heterogeneity arises from several sources. In our sample, randomized experiments on financial education programs have been conducted in 33 countries with varying target groups (see Table A.3 in Appendix A). Moreover, the underlying educational interventions are diverse, ranging from provision of an informational brochure (e.g., Choi et al., 2010) to offering high-intensity classroom instruction (e.g., Bruhn et al., 2016). Additionally, outcomes are measured at different points in time and with different types of data. This heterogeneity must be accommodated in order to draw general conclusions about the literature.

When there is heterogeneity in the studies under consideration, meta-analyses require certain assumptions about the sources of variance in the observed treatment effect estimates. Consider a set of j randomized experiments, each reporting an estimate of a causal (intention to treat) treatment effect relative to a control group.⁴ Assuming no

heterogeneity in true effects implies that the observed estimates of a treatment effect are sampled from a distribution with a single true effect β_0 and variance σ^2 , as in the following meta-analysis model:

$$y_j = \beta_0 + \epsilon_j, \quad (1)$$

where y_j is an estimate of a treatment effect in the j th study, β_0 denotes the common true effect, and ϵ_j is the study level residual with $\epsilon_j \sim N(0, \sigma_j^2)$. Thus, the estimate of the common true effect is given by estimating the above model with weighted least squares using inverse variance weights ($w_j = \frac{1}{\sigma_j^2}$). While this may be a reasonable assumption for some empirical studies, such as medical trials with identical treatment, dosage, and procedures for measuring outcomes, it is not a reasonable assumption in the context of educational interventions, which tend to be quite diverse.

A more reasonable approach in an educational setting would be to assume heterogeneity between studies, hence assuming a distribution of possible true effects, allowing true effects to vary across studies with identical within-study measurement error. This is an important relaxation of the common-effect assumption; as we document in Section 2.1, programs are quite heterogeneous. The weighted average effect then does not represent a single true effect, but the mean of the distribution of true effects. Thus, the model can be written as:

$$y_j = \beta_0 + v_j + \epsilon_j \quad (2)$$

with $v_j \sim N(0, \tau^2)$ and $\epsilon_j \sim N(0, \sigma_j^2)$. τ^2 is the between-study variance in true effects that is unknown and must be estimated from the data,⁵ and σ_j is the within-study standard error of the treatment effect estimate y_j that is observed for each study j . Subsequently, weighted least squares is used to estimate β_0 with inverse variance weights defined as $w_j = (\tau^2 + \sigma_j^2)^{-1}$. Thus, instead of estimating one common effect, the goal is to estimate the mean of the distribution of true effects.

While the models in Eqn 1 and Eqn 2 have considered cases in which each study contributes one independent treatment effect estimate, this is generally not the case in the financial education literature. Instead, studies may report treatment effect estimates from multiple treatments and a common control group within studies at multiple

vert each estimate to a (bias corrected) standardized mean difference (Hedges' g), such that the treatment effect estimate y_j is standardized as $g_j = \frac{M_T - M_C}{SD_p}$ with $SD_p = \sqrt{\frac{(n_T - 1)SD_T^2 + (n_C - 1)SD_C^2}{n_T + n_C - 2}}$, i.e., the mean difference in outcomes between treatment (M_T) and control (M_C) as a proportion of the pooled standard deviation (SD_p) of the dependent variable. 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 is defined as: $SE_{g_j} = \sqrt{\frac{n_T + n_C}{n_T n_C} + \frac{g_j^2}{2(n_T + n_C)}}$.

⁵ There are several possible algorithms to estimate the between-study variance τ^2 . Our approach uses the method of moments estimator (see Harbord and Higgins, 2008), but iterative approaches such as (restricted) maximum likelihood or empirical Bayes estimation are also frequently used in meta-analyses. We show results for these alternative approaches in Tables B.3 and B.4 in Appendix B.

³ We also show results for the sample of studies reporting the ITT in Appendix B, Tables B.1 and B.2.

⁴ Because each study j may report its treatment effect estimate in a different unit (i.e., a different currency or on different scales), we con-

time points and for multiple outcomes. Therefore, we extend the model above to incorporate multiple (and potentially correlated) treatment effect estimates within studies:

$$y_{ij} = \beta_0 + \nu_j + \epsilon_{ij}, \quad (3)$$

where y_{ij} is the i th treatment effect estimate within each study j . β_0 is the mean of the distribution of true effects, ν_j is the study-level random effect with $\nu_j \sim N(0, \tau^2)$, τ^2 is the between-study variance in true effects, and $\epsilon_{ij} \sim N(0, \sigma_{ij}^2)$ is the residual of the i th treatment effect estimate within each study j . This model allows between-study heterogeneity in true effects but assumes that treatment effect estimates within studies relate to the same study-specific true effect. This means the common within-study correlation of treatment effect estimates is induced by random sampling error.

While the estimator proposed in [Hedges et al. \(2010\)](#) does not require an exact model of the within-study dependencies in true effects, [Tanner-Smith and Tipton \(2014\)](#) and [Tanner-Smith et al. \(2016\)](#) suggest that the following inverse variance weights (w_{ij}) are approximately efficient in case of a correlated effects model:

$w_{ij} = \{(\tau^2 + \frac{1}{k_j} \sum_{k=1}^{k_j} \sigma_{ik}^2)[1 + (k_j - 1)\rho]\}^{-1}$, where τ^2 is the estimated between-study variance in true effects, $(\frac{1}{k_j} \sum_{k=1}^{k_j} \sigma_{ik}^2)$ is the arithmetic mean of the within-study sampling variances (σ_{ik}^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 estimate the model with these weights and choose $\rho = 0.8$ as the default within-study correlation of estimates (see [Tanner-Smith and Tipton, 2014](#)). However, sensitivity analyses of such an assumption are easily implemented, and results for $\rho = [0, 0.9]$ in increments of 0.1 in Figs. B.1 and B.2 in Appendix B, do not show any difference in results.

Using this approach, we can formally investigate the importance of modeling between-study heterogeneity in treatment effects and to compare the results to a model with the common effect assumption. We use all the statistical information reported in primary studies, since the method we use can accommodate multiple estimates within studies, and thus is not dependent on creating highly aggregated measures, such as the within-study average effect sizes reported in earlier meta-analyses. To probe the robustness of our results, we estimate six alternative models, including three methods of addressing and correcting for publication bias and a consideration of the power of the underlying primary studies.

4. Data

To arrive at an unbiased estimate of the mean of the distribution of true effects of financial education programs, we compose a complete list of the randomized experiments in the financial education literature. Applying the inclusion criteria from [Section 3.1](#), we arrive at a dataset of 68 papers reporting the effects of 76 independent sam-

ple experiments. This is a much bigger sample of RCTs than considered in any previous meta-analyses.

An important part of our meta-analysis is the inclusion of many recent papers in our dataset, which enables us to provide a comprehensive and updated review of the large and rapidly increasing amount of research on this topic. The review by [Fernandes et al. \(2014\)](#) is the first paper on this topic in the literature, and it covers only 13 RCTs from which 15 observations are coded. The meta-analysis in [Miller et al. \(2015\)](#) covers seven RCTs. Of our 76 independent sample experiments, one-third have not been included in the most recent meta-analysis by [Kaiser and Menkhoff \(2017\)](#).⁶ Thus, we greatly expand on those previous studies. Table C.1 in Appendix C contains a comparison of our dataset of RCTs to these earlier meta-analyses.

From our sample of 76 independent randomized experiments, we extract 673 treatment effect estimates. Of these 76 RCTs, 64 studies report 458 treatment effects on financial behaviors and 50 studies report 215 treatment effects on financial knowledge (see Table A.2 in Appendix A). The studies vary in their choice of dependent variables, ranging from several financial behaviors to financial knowledge.

We start our analysis by showing in [Table 1](#) that the descriptive statistics suggest that financial education is, on average, effective in improving knowledge and behavior. The average effect size across all types of outcomes is 0.123 standard deviation (SD) units (SD=0.183), and the median effect size is 0.098 SD units.⁷ The minimum effect size is -0.413, and the maximum effect size is 1.374. The average standard error of the treatment effect is 0.085 (SD=0.049) and the median standard error is at 0.072. The average sample size across the 76 randomized experiments is 2136 and the median sample size is 840.

The estimates in [Table 1](#) also show that there is substantial variation in program instruction time, with the average estimate associated with a mean of 11.7 h of instruction (SD=16.3), and the median associated with 7 h of instruction. Treatment effects are estimated 30.4 weeks (7 months) after treatment, on average, with a standard deviation of 31.6 weeks (7.3 months). The median time between financial education treatment and measurement of outcomes is 25.8 weeks (5.9 months). Further, we note that nearly three-quarters (72.4%) of the treatment effect estimates target low-income individuals (income below the median country income), and 60.8% of the estimates are from programs in developing economies. Reflecting the high quality of this sample of studies, 30.8% of all estimates reported in randomized experiments appear in top economics and finance journals. The average age of study participants across all reported estimates is 33.5 years, with 7.6% of estimates focused on children (< 14 years old), 19.6% focused on youth (14–25 years old), and 72.8% focused on adults (> 25 years old).

⁶ We have been careful to update all of the papers to the latest version and include, for example, the estimates in the published version of the papers; see Table C1 in Appendix C.

⁷ Note that all effect sizes are scaled such that desirable outcomes have a positive sign. For example, we code a negative coefficient on “loan default” as a positive treatment effect (i.e., reduction in loan default) and vice versa.

Table 1

Descriptive statistics. Descriptive statistics at the extracted estimate-level, meaning we consider the total of 673 treatment effects reported in 76 RCTs.

Variable	Obs.	Mean	Median	Std. Dev.	Min.	Max.
Hedges' <i>g</i>	673	0.123	0.098	0.183	−0.413	1.374
SE (<i>g</i>)	673	0.084	0.072	0.049	0.007	0.365
Delay (in weeks)	639	30.238	25.800	31.537	0.000	143.550
Intensity (in hours)	604	11.709	7.000	16.267	0.008	108.000
Mean age (in years)	646	33.549	38.430	12.488	8.500	55.000
Children (< age 14)	673	0.076	–	–	0.000	1.000
Youth (age 14–25)	673	0.196	–	–	0.000	1.000
Adults (> age 25)	673	0.728	–	–	0.000	1.000
Low income (yes=1)	673	0.724	–	–	0.000	1.000
Developing economy (yes=1)	673	0.608	–	–	0.000	1.000
Top econ journal (yes=1)	673	0.308	–	–	0.000	1.000
Classroom	673	0.666	–	–	0.000	1.000
Online	673	0.224	–	–	0.000	1.000
Educative Nudge	673	0.037	–	–	0.000	1.000
Counseling	673	0.073	–	–	0.000	1.000

5. Results

We present the results in four steps. In [Section 5.1](#), we show the main results of our meta-analysis, including the universe of randomized experiments, and compare the results to the first meta-analysis in the literature. In [Section 5.2](#), we examine our main effects when accounting for publication bias. [Section 5.3](#) presents our main results once we include only the highest quality journals. In [Section 5.4](#), we summarize the results of comprehensive robustness exercises that are reported in full in Appendix B.

5.1. A meta-analysis of randomized experiments

We describe our findings by first plotting the universe of 673 raw effects extracted from the 76 studies against their inverse standard error (precision), which can be seen in [Fig. 2](#). We disaggregate the data and distinguish between estimated treatment effects on *financial behaviors* ($n = 458$ estimates) and *financial knowledge* ($n = 215$). The unweighted average effect on financial behaviors is 0.0937 SD units, and the unweighted average effect on financial knowledge is 0.186 SD units. With this simple analysis of the raw data, we see that financial education improves both financial knowledge and behaviors.

Next, we compare the data in our study with the results presented in earlier meta-analyses. Specifically, we estimate the weighted average effect on financial behaviors using Robust Variance Estimation in Meta-Regression with Dependent Effect Size Estimates (RVE, [Hedges et al., 2010](#)) under the common true effect assumption,⁸ as well as the random effects assumption, and compare our results in the larger sample of 64 RCTs (reporting treatment effects on financial behaviors) to their earlier accounts of the literature

(i.e., [Fernandes et al., 2014](#); [Miller et al., 2015](#); [Kaiser and Menkhoff, 2017](#)).⁹ These results are reported in [Fig. 3](#).

A few important clarifications are in order: [Fernandes et al. \(2014\)](#) estimate and standard error shown in [Fig. 3](#) is from the analysis of 15 observations of RCTs in their paper, not from our analysis of their data. We were not able to exactly replicate this estimate.¹⁰ In the replication process, we uncovered five data errors in the direct coding and classification of RCT effect sizes. The failure to replicate the estimate is a combination of those five data errors and a difference in the extraction of estimates in the original study (i.e., we generally agree with the coding of seven estimates but disagree with coding decisions regarding five estimates). In Appendix D, we describe our replication of the original result and document each coding discrepancy. We discussed the coding decisions as well as our interpretation of the data with the authors of both the original meta-analysis and individual studies that needed clarification.

Taking their published estimate at face value, [Fig. 3](#) shows that simply updating the dataset to incorporate the recent RCTs increases the effect by more than three times. Compared to the estimate of 0.018 SD units (with a 95% confidence interval (CI₉₅) from −0.008 to 0.044) reported in the first meta-analysis of the literature, the weighted average effect in this larger sample of recent RCTs is about 3.6 times higher. The new estimate of the effect size, even

⁸ Thus, we assume $\tau^2 = 0$, i.e., the weights are defined as $w_{ij} = \{(\frac{1}{k_j} \sum_{i=1}^{k_j} \sigma_{ij}^2)[1 + (k_j - 1)\rho]\}^{-1}$. Note that [Fernandes et al. \(2014\)](#) use only one observation per study by creating within-study average effect sizes, i.e., the weights in their study are defined as $w_j = \frac{1}{\sigma_j^2}$. We show results with this approach in Tables B3 and B4 of Appendix B.

⁹ We convert the correlations used as an effect size metric by [Fernandes et al. \(2014\)](#), (r) to a standardized mean difference (Cohens' d) $d = \frac{2r}{\sqrt{1-r^2}}$ and we convert the standard error using $SE_d = \sqrt{\frac{4SE_r^2}{(1-r^2)^3}}$ (cf. [Lipsey and Wilson, 2001](#)). This is justified under the assumption that the outcome measures in each group are continuous and normally distributed and that the treatment variable is a binary variable indicating treatment and control groups, i.e., a valid assumption in the context of RCTs. To arrive at the "bias corrected standardized mean difference" (Hedges' g), one may apply the following bias correction factor $\text{post } g = d \left(1 - \frac{3}{4(n_1+n_2-2)}\right)^{-1}$ (cf. [Borenstein et al., 2009](#)) but these metrics are nearly identical in the context of the financial education literature in which the average sample size is 2136 and the median sample size is 840.

¹⁰ Instead, our replication yields an estimated effect of 0.025 SDs (CI₉₅ 0.008, 0.042) relative to the result of 0.018 SDs (CI₉₅ −0.008, 0.044) in [Fernandes et al. \(2014\)](#).

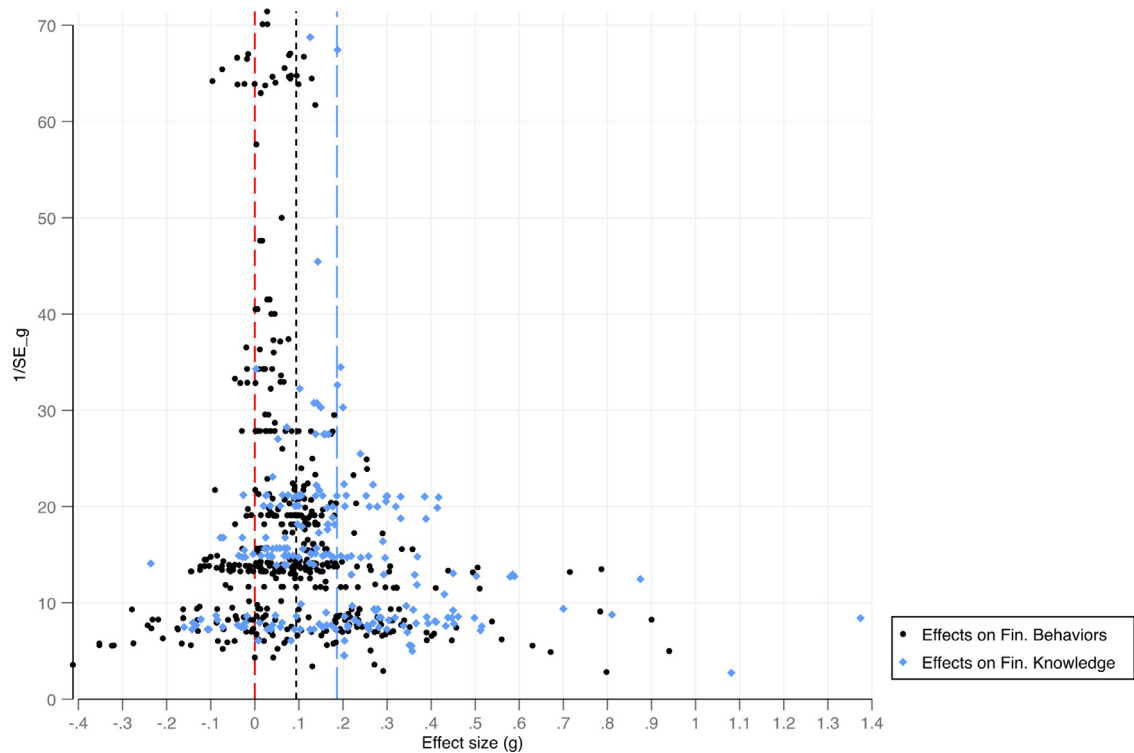


Fig. 2. Distribution of raw financial education treatment effects and their standard errors.

Effect size (g) is the bias corrected standardized mean difference (Hedges' g). $1/SE_g$ is its inverse standard error. The number of observations in the treatment effects on financial behaviors sample is 458 effect size estimates from 64 studies. The number of observations in the treatment effects on financial knowledge sample is 215 effect size estimates from 50 studies. Thirty-eight studies report treatment effects on both types of outcomes. The mean effect size on financial behaviors is 0.094 SD units, and the mean effect size on financial knowledge is 0.186 SD units.

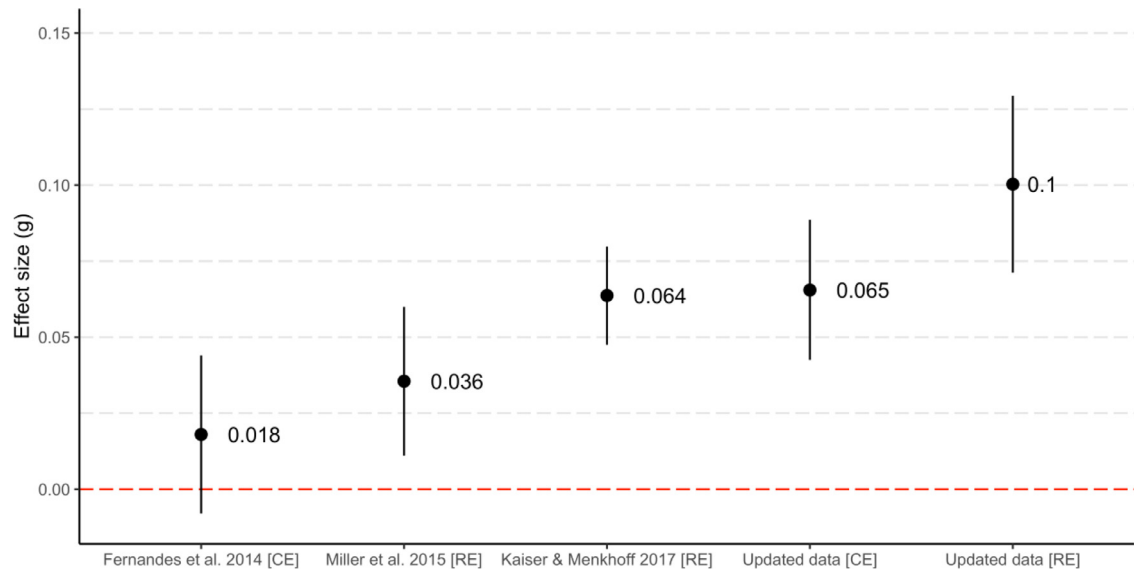


Fig. 3. Comparing the updated evidence to previous meta-analyses (treatment effects on *financial behaviors*).

Fernandes et al. (2014) report weighted least squares estimates with inverse variance weights (common-effect assumption) using 15 observations from 13 RCTs. Miller et al. (2015) use a random effects model and include results from 20 studies (13 quasi-experiments and seven RCTs). The result by Kaiser and Menkhoff (2017) is from a random effects model (RVE) using 349 observations from 90 studies (50 quasi-experiments and 40 RCTs). The results with updated data (458 treatment effect estimates from 64 RCTs) are from robust variance estimation in meta-regression with dependent effect size estimates (RVE) (Hedges et al., 2010) with $\tau^2 = 0$ in the common-effect case, and τ^2 estimated via methods of moments in the random-effects case. Dots show the point estimates, and the solid lines indicate the 95% confidence interval.

with the identical assumption of a common true effect (see Eq (1) and footnote 8), clearly rules out a null effect of financial education (0.065 SD units with CI_{95} from 0.043 to 0.089).

Because the common true effect assumption is problematic in the context of heterogeneous financial education interventions, we estimate the mean of a distribution of true effects using the model specified in Eq. (3). In addition to the already mentioned theoretical reasons to assume a distribution of true effects rather than a single true effect, we note that formal tests of heterogeneity show that 86% of the observed between-study variance can be attributed to heterogeneity in true effects and only 14% of the observed variance would have been expected to occur as a result of within-study sampling error alone (see Tables B.3 and B.4 in Appendix B).¹¹

Fig. 3 shows the result of the random effects model. In our view, this estimated mean of the distribution of financial education treatment effects is the most appropriate aggregate effect size to consider; the estimate results in a mean of 0.1003 SD units (CI_{95} from 0.071 to 0.129), and thus is significantly different from the estimate using the common true effect assumption. The effect of financial education is now approximately 5.5 times larger than the estimate reported in the first meta-analysis. Compared to the results provided by Miller et al. (2015) and Kaiser and Menkhoff (2017), both of whom report estimates with the assumption of heterogeneous effects, the effect is estimated to be significantly larger. Relative to Miller et al. (2015), the effect is about 2.8 times larger. Relative to Kaiser and Menkhoff (2017), the effect is about 65% larger. Qualitatively, this estimate is similar in magnitude to statistical effect sizes reported in meta-analyses of behavior change interventions in the health domain (e.g., Rooney and Murray, 1996; Noar et al., 2007; Portnoy et al., 2008) or energy conservation behavior domain (e.g., Karlin et al., 2015).

To summarize our main findings, evidence from our meta-analysis that incorporates an updated set of papers shows that financial education is effective, on average.

5.2. Publication bias

A potential explanation for the more favorable assessment of the mean of the distribution of treatment effects in the updated data is that the financial education literature could be subject to publication bias. Publication bias refers to the potential preference for researchers to report and journals to publish statistically significant results (Brodeur et al., 2016, 2020).

An ideal study of publication bias would include all financial education RCTs that have been conducted but never written up (i.e., all latent studies). We attempt to obtain descriptive evidence on latent studies with data from the AEA's RCT registry, searching for the terms "financial education" and "financial literacy." However, the AEA RCT registry has only been in operation since 2012, and only

14 of the studies in our sample are in the registry. Even in more recent years, researchers have failed to use the registry, making it a poor representation of the state of RCT experiments in this field of inquiry.¹² Thus, the extent to which studies are discarded due to statistical insignificance at conventional levels (sometimes called the "file-drawer problem") must be identified using alternative methods.

A visual inspection of the plot in Fig. 2 shows that both samples of estimated treatment effects (on financial knowledge and financial behaviors) resemble a roughly symmetric funnel until effect sizes of 0.5 SD units and above are reached. However, histograms of the distribution of Z-statistics (see Figs. B.3 and B.4 in Appendix B) are indicative of selective publication, since there appear to be jumps at the cut-off of values indicating statistical significance at conventional levels.

Thus, we next move to a formal investigation of publication bias using the method recently developed by Andrews and Kasy (2019). This method of identifying the extent of publication bias in the literature involves a step-function approach to estimate the conditional publication probability of treatment effect estimates with Z-statistics smaller than 1.96 (1.65) in absolute values (as opposed to Z-statistics equal to or larger than 1.96 (1.65) in absolute values), i.e., selection of statistically significant results at the 5% (10%) level. We then use this estimated conditional publication probability to re-estimate the mean of the distribution of true effects using the non-parametric estimator described in Andrews and Kasy (2019). In this model of publication bias, it is assumed that studies with different standard errors do not have systematically different estimands, which is a common assumption in meta-analyses and is automatically fulfilled with the much stronger assumption of a common true effect. Since we include multiple estimates per study, identification rests on the additional assumption that the selection of estimates is done on a case-by-case basis within papers (cf. Andrews and Kasy, 2019, p. 2786). For inference, standard errors are clustered at the study level.

Table 2 shows the estimated conditional publication probabilities for the set of estimates on *financial behaviors* and for the set of estimates on *financial knowledge*. Statistically insignificant estimates of treatment effects at the 5% level on *financial behaviors* have a publication probability of about 30%, relative to statistically significant estimates in the same set. Statistically insignificant estimates of treatment effects on *financial knowledge* have a publication probability of about 15%. Thus, there appears to be evidence of selective publication in this literature. How do these estimates compare to other empirical literatures? Relative to the estimate in Brodeur et al. (2020) who cover a universe of 145 RCTs published in 25 top economics

¹¹ A Cochran's Q-test of homogeneity (with one synthetic effect size per study) results in a Q-statistic of 464.71 ($p < 0.000$).

¹² In the subset of 33 studies with experiment end dates within our time frame — papers with formal drafts by January 2019 — only eight experiments were listed as "completed" and given a lag in the time between completing an experiment and writing up results, many of these papers would likely not have been written up within the timeframe in study. Thus, it is hard to disentangle failed experiments, time lags in formal working papers or publications, and a lack of "desirable effects" in determining why studies do or do not appear in the registry.

Table 2

Identification of and correction for publication bias in the financial education literature.

This table presents results from non-parametric identification of and correction for publication bias based on the method described in Andrews and Kasy (2019) (see Andrews and Kasy 2019, Appendix C). $\tilde{\beta}_0$ denotes the estimate of the true treatment effect in latent studies (i.e., the bias corrected treatment effect) and λ_p denotes the estimated conditional publication probability (p) based on the Z-statistic (y_{ij}/σ_{ij}) as specified in the respective column header. Columns 1 and 3 show estimates for the treatment effects on financial behaviors and financial knowledge with $p(y_{ij}/\sigma_{ij}) = \lambda_p$ if $|y_{ij}/\sigma_{ij}| < 1.96$ and $p(y_{ij}/\sigma_{ij}) = 1$ if $|y_{ij}/\sigma_{ij}| \geq 1.96$, i.e., selection on significance at the 5%-level, respectively. Columns 2 and 4 show estimates for the treatment effects on financial behaviors and financial knowledge with $p(y_{ij}/\sigma_{ij}) = \lambda_p$ if $|y_{ij}/\sigma_{ij}| < 1.65$ and $p(y_{ij}/\sigma_{ij}) = 1$ if $|y_{ij}/\sigma_{ij}| \geq 1.65$, i.e., selection on significance at the 10%-level, respectively. Standard errors (clustered at the study-level) are shown in parentheses.

(a) Treatment effects on financial behaviors				(b) Treatment effects on financial knowledge			
(1) Selection on significance (Cutoff of $ Z = 1.96$)		(2) Selection on significance (Cutoff of $ Z = 1.65$)		(3) Selection on significance (Cutoff of $ Z = 1.96$)		(4) Selection on significance (Cutoff of $ Z = 1.65$)	
$\tilde{\beta}_0$	λ_p	$\tilde{\beta}_0$	λ_p	$\tilde{\beta}_0$	λ_p	$\tilde{\beta}_0$	λ_p
0.057 (0.001)	0.303 (0.071)	0.050 (0.007)	0.256 (0.051)	0.150 (0.037)	0.150 (0.126)	0.160 (0.040)	0.250 (0.190)

journals, and arrive at a publication probability of about 50%, the degree of publication bias appears to be relatively larger (although the confidence interval of our estimate also includes the possibility of a publication probability for insignificant results of 44% for the set of studies on financial behaviors). Compared to the examples provided in Andrews and Kasy (2019), the magnitude of selection bias is substantially smaller than found in economics laboratory experiments (Camerer et al., 2016), where results statistically insignificant at conventional levels have only a 3.8% probability of being published relative to statistically significant results (cf. Andrews and Kasy, 2019). Our estimate of a conditional publication probability of 30% is similar to the literature on the effect of a minimum wage on unemployment (Belman et al., 2015) (cf. Andrews and Kasy, 2019).

Using the estimated conditional publication probabilities to retrieve the estimate for the mean effect in the population of latent studies, we find that estimates are lower than unadjusted estimates, but still sizeable: 0.057 SD units for the set of estimates on financial behaviors and 0.15 SD units for the set of studies on financial knowledge) and significantly different from zero. Considering further tests of publication bias in Appendix B, we conclude that the reported treatment effects may be inflated by about one-third, but that bias corrected estimates generally rule out zero effects of financial education and are still about three times larger than the estimates reported in the first meta-analysis of the literature.

5.3. Journal quality

To address possible concerns regarding the internal validity and general rigor of the included experiments and to focus on what editors and reviewers have judged to be the highest quality evidence, we restrict the sample to studies published in top economics or finance journals only.¹³ We compare the estimated treatment effects on the financial behaviors of the 15 studies published in these

journals to the estimated treatment effects of the other 31 studies published in other journals and to 18 working papers. While treatment effects are estimated to be smaller in top publications, there are no statistically significant differences (see Table B.5 in Appendix B) between the average effect and the effect for just top economics or finance journals. The same is true for effect sizes on financial knowledge where eight experiments published in top general interest or top field economics journals report smaller, albeit not statistically different, effect sizes than 23 experiments published in other journals or 19 studies published as working papers.

5.4. Model sensitivity and robustness checks

We probe the robustness of our findings about the average effect of financial education programs with various sensitivity checks that are reported in full in Appendix B. These tests include (1) estimating alternative meta-analyses including models with a common effect assumption and in a Bayesian hierarchical model (as in Meager 2019), choosing different assumed within-study correlations of treatment effect estimates for the random-effects RVE approach, and estimating both common-effect and random-effects models with one observation per study relying on one composite effect size per study (inverse-variance weighted within-study average) by via several alternative algorithms; (2) investigating and correcting for publication bias via alternative methods and restricting the sample to only those studies with adequate power to identify small treatment effects (see Ioannidis et al., 2017); and (3) additional results, including the robustness of results when excluding any papers of the authors of this meta-analysis. All these sensitivity analyses and robustness checks confirm our main conclusions.

6. Understanding heterogeneity in treatment effects

To better understand the sources of heterogeneity in this literature, we further disaggregate our data into various subgroups and investigate the mean effect of financial education interventions.

¹³ These journals are: (1) *American Economic Journal: Applied Economics*, (2) *American Economic Journal: Economic Policy*, (3) *Economic Journal*, (4) *Journal of Development Economics*, (5) *Journal of the European Economic Association*, (6) *Journal of Finance*, (7) *Journal of Political Economy*, (8) *Management Science*, (9) *Quarterly Journal of Economics*, and (10) *Review of Financial Studies*. There were no publications in other top journals, such as

the *American Economic Review*, *Econometrica*, the *Journal of Financial Economics*, or the *Review of Economic Studies*.

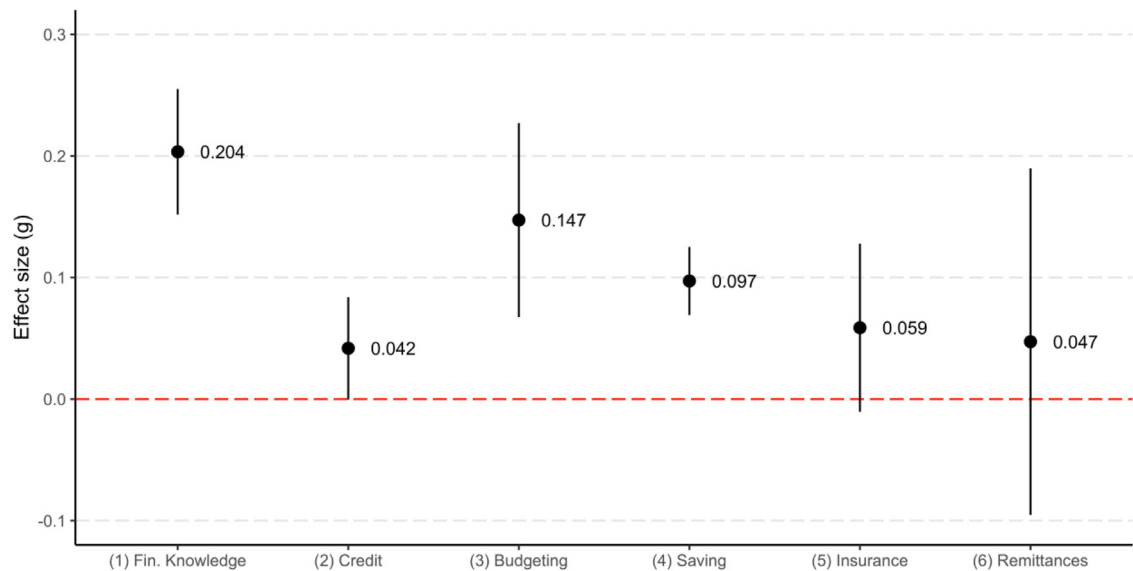


Fig. 4. Financial education treatment effects by outcome domain.

Results from robust variance estimation in meta-regression with dependent effect size estimates (RVE) (Hedges et al., 2010). The number of observations for the financial knowledge sample (1) is 215 effect size estimates within 50 studies. The number of observations for the credit behavior sample (2) is 115 within 22 studies. The number of effect size estimates for the budgeting behavior sample (3) is 55 within 23 studies. The number of observations in the saving and investing behavior (4) sample is 253 effect size estimates within 54 studies. The number of observations in the insurance behavior sample (5) is 18 effect sizes within six studies. The number of observations on remittance behavior (6) is 17 effect size estimates reported within six studies. Dots show the point estimates, and the solid lines indicate the 95% confidence interval.

6.1. Outcome domains

In addition to the effects on financial behaviors aggregated above (Fig. 3), i.e., all behaviors, we also include estimates on financial knowledge (Fig. 4). Treatment effects on *financial knowledge* are larger than the effect sizes on *financial behaviors*. This difference appears plausible because it may be easier to learn new concepts and facts than to change possibly entrenched behaviors.

In Fig. 4, we find that the mean of the distribution of true effects regarding financial knowledge is estimated to be 0.204 (CI₉₅ from 0.152 to 0.255).¹⁴ This average effect is very similar to the average effects of educational interventions in math or reading (see Hill et al., 2008; Cheung and Slavin, 2016; Fryer, 2016; Kraft, 2020).

Regarding effect sizes on *financial behaviors*, we distinguish five domains of behavior: credit, budgeting, saving (and investing), insurance, and remittances. Treatment effects on these behaviors are mostly not statistically different from each other, suggesting the adequacy of pooling across these outcomes. Still, the effect sizes across outcome domains differ, with treatment effects on *budgeting behavior* and *saving (and investing) behavior* being the largest. Additional results shown in Table B.6 in Appendix B suggest that estimates on these two behaviors are also the most robust, while the effects on other categories of fi-

ancial behaviors are less certain due to fewer studies including these outcomes (*insurance* and *remittances*) or to small average effects (*credit behaviors*).

The fact that findings are largest in magnitude for budgeting behavior and saving (and investing) behavior is not entirely surprising. Many of the papers studied specifically aim to change these behaviors. For example, financial illiteracy could be a reason people choose index funds with high fees. Choi et al. (2010) investigate high-fee funds with an intervention that provides individuals with information prior to choosing hypothetical portfolios. Other interventions in developing economies target saving as an outcome and design educational interventions to improve financial inclusion (e.g., Flory, 2018). Further, classroom financial education for school-aged children often includes a unit on budgeting (e.g., Bruhn et al., 2016). On the contrary, credit-related outcomes are not always directly covered by financial education interventions but are often included in outcome measurements. For example, Bruhn et al. (2016) find that while financial education improves budgeting and saving behavior, it increases the likelihood of making expensive consumer purchases with credit. This suggests that different financial behaviors may be subject to different (cognitive and non-cognitive) mechanisms leading to behavior change. Alan and Ertac (2018) and Lüthmann et al. (2018) demonstrate that financial education can affect patience and the quality of intertemporal decision making. Berg and Zia (2017) show that an intervention implemented into mainstream media that harnesses emotional connections affects credit behaviors without necessarily impacting the cognitive components of general financial literacy. Similarly, financial confidence and attitudes may play an important role in mediating the

¹⁴ Thus, the effect is significantly larger than the estimate reported by 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., $\sqrt{r^2} = 0.066$ or $d = 0.133$, i.e., an average effect of about 0.133 SD units.

causal effects of financial education on financial behaviors (e.g., Carpena et al., 2017; Carpena and Zia, 2020).

There are also studies in developing countries that address unique and complex financial decisions. It is important to look at those studies separately, as individuals in developing countries may face unique constraints. For example, Doi et al. (2014) study the effects of financial training on remittances and household financial management. They find no change in the use of formal channels in sending remittances, though they note that this is because nearly everyone already uses the formal channels in their setting. In addition, they find no change in the amounts or frequency of sending remittances, but they do show that the family receiving the remittances is able to better optimize their saving decisions.

Overall, our results across outcome domains are generally in line with earlier accounts of the literature (Fernandes et al., 2014; Miller et al., 2015; Kaiser and Menkhoff, 2017), and they extend to the larger set of RCTs we examine.

6.2. Sample population

We disaggregate the sample of RCTs by the characteristics of the sample population. First, we split the sample by country-level income, distinguishing between high-income and developing economies, to account for differences in resources.¹⁵ We find that the treatment effects of interventions in developing economies on financial behaviors are about 9.6% smaller than those in more affluent countries; however, this difference is not statistically significant (see Panel A(a) of Table 3). Financial knowledge treatment effects in developing economies are about 46% smaller and significantly different from the estimate in the set of studies in developed countries (see Panel B(a) of Table 3), which is in line with evidence presented in Kaiser and Menkhoff (2017). These smaller effect sizes could be due to additional barriers faced by consumers in developing countries. For example, Cole et al. (2011) note that their financial education intervention in India and Indonesia had only modest short-run effects and no long-run effects on opening a bank account. However, subsidies increased account ownership in both the short and long run. They note that costs may be a larger barrier in some emerging economies and increasing competition in financial markets or finding other ways to reduce costs may be first order. Another reason for smaller effect sizes in low-income countries could be that it is at times optimal for low-income individuals who are resource constrained to make no changes in response to additional financial literacy, as Lusardi et al. (2017) show theoretically.

We next look at the differences between low-income individuals and people with average or above aver-

age individual income (relative to the average within-country income). In contrast to the earlier studies by Fernandes et al. (2014) and Kaiser and Menkhoff (2017), which find lower effects for programs with relatively low-income participants, we do not find any significant differences between these two samples (see Panel A(b) and Panel B(b) of Table 3).

Additionally, we disaggregate our sample by participant age (see Panel A(c) and Panel B(c) of Table 3). Treatment effects on financial behaviors are smallest for children (below age 14) (0.064 SD units) relative to youth (ages 14–25) (0.120 SD units) and adults (above age 25) (0.107 SD units), while the latter difference is only marginally significant. Treatment effects on financial knowledge, on the other hand, are estimated to be largest among children (0.276 SD units) relative to youth (0.186 SD units) and adults (0.200 SD units). These differences, however, are not statistically significant due to large uncertainty around the estimate for children, which is based on 15 observations in seven studies (CI₉₅ from 0.008 to 0.545). While all interventions are designed with the age of the participants in mind, it is potentially more challenging to construct financial behavior outcome measures for minors. At the same time, the assessments designed for those under age 18 are often more substantial, as they include thoughtful school-based examinations. Further, Lusardi et al. (2017) show that those earlier in the life cycle—youth per our definition—may face challenges in responding to financial education programming due to a lack of resources and the need for borrowing to improve their human capital.

6.3. Intensity of treatment and time horizon

Next, we disaggregate the sample by intensity of financial education treatment by classifying the interventions by the time devoted to instruction. While we do not find statistically significant differences between the different splits, the results are generally suggestive of a positive relation between increased instructional time and realized treatment effects. The point estimate on financial behaviors appears to be larger in the set of studies covering more intensive treatments (i.e., equal to 20 h or more) (see Panel A(d) of Table 3) and the upper bound of the 95% CI also suggests the possibility of substantial effects on financial knowledge in the set of seven studies covering intensive classroom instruction (see Panel B(d) of Table 3).

Additionally, we tackle the topic of the potential decay of effectiveness of financial education over time. Determining how short-run effects differ from longer-run effects requires important considerations. Studies that measure longer-run outcomes often use interventions designed to affect long-run behavior. For example, Alan and Ertac (2018) use an intervention aimed at fostering a specific non-cognitive skill that will stay with children for life and show that knowledge about how to be forward-looking in making intertemporal decisions that is acquired by 9- and 10-year-old children does not fade out after three years. Some studies that explore short-run behaviors have interventions that are designed to improve a particular behavior that may not be relevant in the long run. For

¹⁵ Country groups are based on the World Bank Atlas method and refer to 2015 data on Gross National Income (GNI) per capita. Low-income economies are defined as those with a GNI per capita of \$1025 or less in 2015, lower-middle income economies are defined by a GNI per capita of between \$1026 and \$4035, upper-middle income economies are those with a GNI per capita between \$4036 and \$12,475, and high-income economies are defined by a GNI per capita greater than \$12,475.

Table 3

Financial education treatment effects by subgroups of studies and populations.

This table reports average effects of financial education treatment on financial behaviors (Panel A) and financial knowledge (Panel B) estimated via RVE. Ten studies with 34 effect size estimates (24 on financial *behaviors* and 10 on financial *knowledge*) are missing information about the delay between treatment and measurement of outcomes. Studies report treatment effects at multiple time points, so the number of studies does not add up to the total number of studies per category.

Subgroup	Effect size (g)	SE	95% CI Lower bound	95% CI Upper bound	n(Studies)	n(effects)
Panel A: Treatment effects on <i>financial behaviors</i>						
(a) By country income						
High income economies	0.1127	0.0316	0.0478	0.1777	32	129
Developing economies	0.0928	0.0130	0.0660	0.1195	32	329
(b) By respondent income						
Low-income individuals	0.0993	0.0194	0.0600	0.1387	43	367
General population	0.1035	0.0219	0.0571	0.1500	21	91
(c) By age of participants						
Children (< age 14)	0.0640	0.0186	0.0188	0.1091	9	36
Youth (age 14–25)	0.1203	0.0415	0.0250	0.2155	11	92
Adults (> age 25)	0.1068	0.0205	0.0653	0.1483	44	330
(d) By intensity of treatment						
<5 h	0.0817	0.0194	0.0407	0.1227	22	124
≥ 5 and <20 h	0.0992	0.0223	0.0533	0.1450	29	251
≥20 h	0.2319	0.0664	0.0745	0.3893	8	54
(e) By delay between treatment and measurement of outcomes						
< 6 months	0.0991	0.0169	0.0645	0.1337	34	180
≥ 6 and < 18 months	0.0901	0.0181	0.0520	0.1283	23	211
≥ 18 months	0.0653	0.0192	0.0209	0.1098	10	49
(f) By type of intervention						
Classroom	0.1064	0.0181	0.0699	0.1428	50	331
Online	0.0796	0.0336	−0.0194	0.1786	5	55
Counseling	0.1595	0.0274	−0.1887	0.5077	2	48
Educative Nudge	0.0597	0.0206	0.0055	0.1138	8	24
Panel B: Treatment effects on <i>financial knowledge</i>						
(a) By country income						
High income economies	0.2591	0.0415	0.1738	0.3443	29	135
Developing economies	0.1392	0.0218	0.0934	0.1851	21	80
(b) By respondent income						
Low-income individuals	0.2238	0.0395	0.1428	0.3049	30	120
General population	0.1835	0.0310	0.1183	0.2486	20	95
(c) By age of participants						
Children (< age 14)	0.2763	0.1098	0.0076	0.5450	7	15
Youth (age 14–25)	0.1859	0.0390	0.1015	0.2703	16	40
Adults (> age 25)	0.2001	0.0282	0.1418	0.2583	28	160
(d) By intensity of treatment						
<5 h	0.2192	0.0265	0.1638	0.2746	24	86
≥ 5 and <20 h	0.1975	0.0481	0.0968	0.2981	21	80
≥20 h	0.1925	0.0855	−0.0307	0.4157	6	9
(e) By delay between treatment and measurement of outcomes						
< 6 months	0.2305	0.0319	0.1654	0.2956	36	142
≥ 6 and < 18 months	0.1425	0.0292	0.0787	0.2064	15	56
≥ 18 months	0.1400	0.0450	−0.0518	0.2282	1	1
(f) By type of intervention						
Classroom	0.1927	0.0306	0.1306	0.2549	38	117
Online	0.2618	0.0402	0.1694	0.3542	10	96
Counseling	0.3460	0.1441	0.0636	0.6284	1	1
Educative Nudge	−0.0238	0.0646	−0.1504	0.1028	1	1

example, [Gine et al. \(2013\)](#) provide a comic that informs rural farmers in Eastern Kenya about insurance, and they then analyze short-run insurance purchase decisions, with no consideration of long-term effects. Despite this caveat, we disaggregate the sample of treatment effects within studies, considering the time span between financial education treatment and measurement of outcomes (see Panel A(e) and Panel B(e) of [Table 3](#)). We start by looking at treatment effect estimates that measure outcomes in the very short run (i.e., less than six months). The average ef-

fect of financial education on financial behaviors within this sample of 34 RCTs (180 effect sizes) is 0.099 (CI₉₅ from 0.065 to 0.134). Looking at treatment effects on financial behaviors that are measured at a time span of between 6 and 18 months (23 experiments and 211 estimates), we find that they are very similar with 0.09 SD units (CI₉₅ from 0.052 to 0.128).

Restricting the sample to even longer time spans, i.e., ten RCTs that measure effects on financial behaviors 1.5 years after treatment or longer, results in an estimated

average of 0.065 SD units (CI_{95} 0.021 to 0.110).¹⁶ These effects are reduced but are still not statistically different from the other estimates.

Regarding the decay in financial knowledge, we find significantly larger effects (0.23 SD units) in 36 RCTs that measure effects on financial knowledge in the very short run (i.e., less than 6 months) relative to RCTs with time horizons above 6 months and below 18 months (0.143 SD units). Only one study measures treatment effects on financial knowledge after 18 months with a point estimate similar to the 6–18-month horizon.

Overall, while the point estimates may be suggestive of some decay in treatment effects over time, our examinations do not find conclusive evidence, indicating that neither sustained and relatively large effects nor close to zero effects of financial education at longer time spans can be ruled out due to the limited number of studies that measure very long-run outcomes. In our set of studies measuring treatment effects on financial behaviors after 18 months, we find positive and significant effects, on average. We attribute the previous finding in [Fernandes et al. \(2014\)](#) of a relatively rapid decay to the fact that the authors chose to model this relation in a meta-regression model with four covariate variables based on a sample of only 29 observations.¹⁷ Thus, the evidence suggesting insignificant effects after time spans of more than 18 months is based on a very limited number of observations and should be viewed with caution in light of the large amount of uncertainty around this estimated effect.

6.4. Intervention type

Next, we split the sample by type of intervention. Most interventions occur in a classroom setting and effects on both financial behaviors and financial knowledge resemble the estimated meta-averages with 0.1 and 0.2 SD units, respectively. While estimates on financial behaviors for *online interventions* and *educative nudges* are smaller in the limited set of experiments on these types of interventions, estimated differences are not statistically significant. The point estimate for counseling *interventions* appears to be larger, which seems plausible due to the individual, problem-oriented and costly approach; however, the treatment effect estimates come only from two studies and the confidence interval is wide.

Treatment effects on financial knowledge do not appear to strongly depend on intervention type where *online* and *classroom interventions* show similar effects. Only two studies look at treatment effects after exposure to counseling

or an *educative nudge*; in these cases, we show the extracted raw estimates (and standard errors) from the papers. While the former effect is not significantly different from classroom or online education, the effect of exposure to the *educative nudge* is estimated to be significantly smaller than the meta-estimate for classroom or online interventions.

Overall, the evidence suggests that classroom financial education interventions are effective, on average. We can tentatively conclude that low-intensity interventions (such as information provision via *educative nudges*) are less effective while personalized interventions, such as counseling, are more effective, though more costly.

7. The economic significance of financial education

In understanding financial education interventions, as is true with any analysis of interventions, it is important to assess not just the statistical effect size but also the economic significance of the effects. In [Section 7.1](#), we discuss the choice in [Fernandes et al. \(2014\)](#) to focus on the “variance explained” as a measure of the effect size. In [Section 7.2](#), we classify our effect sizes using studies of the magnitudes of the effects in educational interventions. In [Section 7.3](#), we provide a back-of-the-envelope analysis of the cost-effectiveness of financial education interventions based on our findings. And in [Section 7.4](#), we discuss the external validity of the RCT estimates by considering recent quasi-experimental studies.

7.1. Statistical effect sizes

A main argument in [Fernandes et al. \(2014\)](#) is that even though the statistical effects of financial education on financial outcomes are positive in the overall sample, the magnitudes are relatively small. But the magnitudes of statistical effect sizes are not easily interpretable with regard to their economic importance. Specifically, the results they report may create the illusion of miniscule treatment effects by using “variance explained” (i.e., a squared correlation coefficient) as an effect size metric when, in fact, the treatment effects can be economically significant.

We illustrate this point with a simple example. Consider the median effect of education (specifically, structured pedagogy) interventions in developing countries, which is roughly 0.13 SD units (see [Evans and Yuan, 2019](#)). Translating this to the (partial) correlation results in a correlation coefficient of 0.06, which explains only 0.36% of the variance in learning outcomes. Thus, according to this criterion, this education intervention would be interpreted to be ineffective, as it “explains little of the variance.” However, [Evans and Yuan \(2019\)](#) report that this is equivalent to a sizeable effect, approximately 0.6–0.9 years of “business as usual schooling,” depending on their choice of empirical specification. They also estimate the returns to education (and specifically literacy) in Kenya and estimate the net present value of this intervention to be 1338 USD at an average annual income of 1079 USD in 2015 PPP. Reported in this way, rather than the “variance explained” metric, these effects are unlikely to be considered economically miniscule. Thus, it can be problematic

¹⁶ Restricting the sample to seven RCTs with 32 treatment effects measured two or more years after intervention results in an estimate of 0.057 SD (CI_{95} 0.001 to 0.114).

¹⁷ We also rerun their type of model (a regression of the estimated treatment effect on “linear effects of mean-centered number of hours of instructions, linear and quadratic effects of number of months between intervention and measurement of behavior, and the interaction of their linear effects” ([Fernandes et al., 2014](#), p. 1867) with our updated data (419 observations within 52 studies) and find coefficient estimates with large standard errors (i.e., insignificant coefficients) throughout (see Table B7 in Appendix B).

to solely rely upon the “variance explained” in determining the economic interpretation of statistical effect sizes.

7.2. Interpreting treatment effects in the education literature

Recent studies in education interventions compare effect sizes across heterogeneous treatments, populations, and outcomes—as we are doing in our analysis—and we turn to that work to get some guidance in interpreting effects. [Kraft \(2020\)](#) suggests five key considerations in determining whether or not programs are effective. First, one should make sure only studies with a causal interpretation (e.g., RCTs) are included in “effect sizes.” Second, one should expect effects to be larger when the outcome is easier to change; this is particularly relevant if the intervention is designed to change the specific outcome. Third, one should take into account heterogeneous effects on different populations. Fourth, one should always consider costs per participant. A small effect size can have a large return on investment if the per participant cost is low. Fifth, one should consider whether the program is easily scalable. We have followed these recommendations.

With these five points in mind, [Kraft \(2020\)](#) further points to a scheme for assessing the effect of education interventions with academic outcomes (i.e., test scores) as the main outcome of interest. He suggests that effects larger than 0.20 standard deviations are “large”; effects between 0.05 and 0.20 standard deviations are “medium”; and effects under 0.05 standard deviations are “small.” This classification is roughly consistent with [Hedges and Hedberg \(2007\)](#), [Bloom et al. \(2008\)](#), and the [What Works Clearinghouse \(2014\)](#). Our effects on financial knowledge ([Fig. 4](#)) show an effect size of roughly 0.203, consistent with an education intervention having a large effect on test scores.

[Kraft \(2020\)](#) also notes that it is more difficult to affect long-run outcomes that are not directly addressed in the intervention. It is, thus, not surprising that effects on financial behavior are more modest than effects on financial knowledge. Even so, these effects are classified as “medium” in magnitude in his interpretation of effect sizes realized in RCTs. This type of comparison is very helpful, as it can be hard to interpret the size of effects by simply looking at the estimates.

7.3. Cost-effectiveness

While understanding effect sizes in standard deviation units is more consistent across educational interventions and a more intuitive metric than “variance explained,” a discussion of effect sizes is incomplete without quantifying costs, as also noted in [Kraft \(2020\)](#). Unfortunately, only 20 papers within the 76 that we examine include a discussion of cost. If we conduct a meta-analysis with only these papers, we find that the estimated treatment effects are smaller in the set of studies reporting costs than in the fully aggregated sample. In [Fig. B.5](#) in Appendix B, we report the results of a regression of a binary indicator of reporting costs on sample and experiment characteristics to examine which studies report costs. The only notable difference is that studies reporting costs are more likely to

involve low-income samples. Since we see no difference in effect sizes based on whether the intervention was targeted to low-income populations, we cannot precisely say what is driving the difference in effect sizes with respect to studies reporting costs.

We report the average costs by study in 2019 US dollars in Table A.1 in Appendix A. Averaging across all studies reporting costs, the mean and median per participant costs are \$60.40 and \$22.90, respectively. Using the [Kraft \(2020\)](#) scheme with respect to effect sizes, an average cost of \$60 per participant would be classified as a “low cost” educational intervention. It could be that studies reporting costs have, on average, lower costs than those that do not report costs. If that is the case, the costs are understated, as are the benefits since effect sizes are smaller in the reporting sample. Several studies mention their interventions had “minimal costs” but do not report a number; we do not include these studies in the cost estimates. Some programs may have costs that are difficult to quantify. Other programs may be difficult to scale. For example, [Calderone et al. \(2018\)](#) report a \$25 per person cost and \$39 per person benefit for a financial education program in India. However, they state that the program is still too costly for a large company to implement at scale. While some studies pass a cost-benefit analysis on the surface, there may be other barriers prohibiting implementation.

Overall, our cost-effectiveness ratio is \$60.40 per person for one-fifth of a standard deviation improvement in outcomes. [Fig. 5](#) displays the cost and effect size by outcome domain for each study. There are two direct takeaways from the figure. First, most effect sizes lie above the zero line but below 0.5 standard deviations. The effects below the zero line largely reflect papers that study the impact of financial education on remittances (e.g., switching to a cheaper financial product when transferring money across countries). Second, there does not appear to be a linear relation between costs and effect sizes. [Fig. B.6](#) in Appendix B displays the effect sizes by cost for each outcome domain, depicting 95% confidence intervals for each estimate.

To make the discussion more salient, we use one paper, which evaluates a large-scale randomized control trial in Peruvian schools ([Frisancho, 2018](#)), that clearly spells out the costs. That paper reports a cost per pupil of \$4.80 USD and that a \$1 increase in spending on the program yields a 3.3-point improvement in the PISA financial literacy assessment. Since this study represents financial education within a year-long class and average and median interventions in the sample are only 12 and 7 h, respectively, it is likely that the average effect across studies corresponds to lower costs. [Frisancho \(2018\)](#) also shows that the course does not detract from performance in other courses, limiting opportunity costs.

Our back-of-the-envelope estimate is conservative in that it does not consider positive externalities of the program. For example, [Frisancho \(2018\)](#) documents that in addition to improved student outcomes, teachers’ financial literacy and credit scores increase. Further, [Bruhn et al. \(2016\)](#) document positive “trickle up” effects for parents. Thus, financial education programs may have externalities beyond the target group, such as affect-

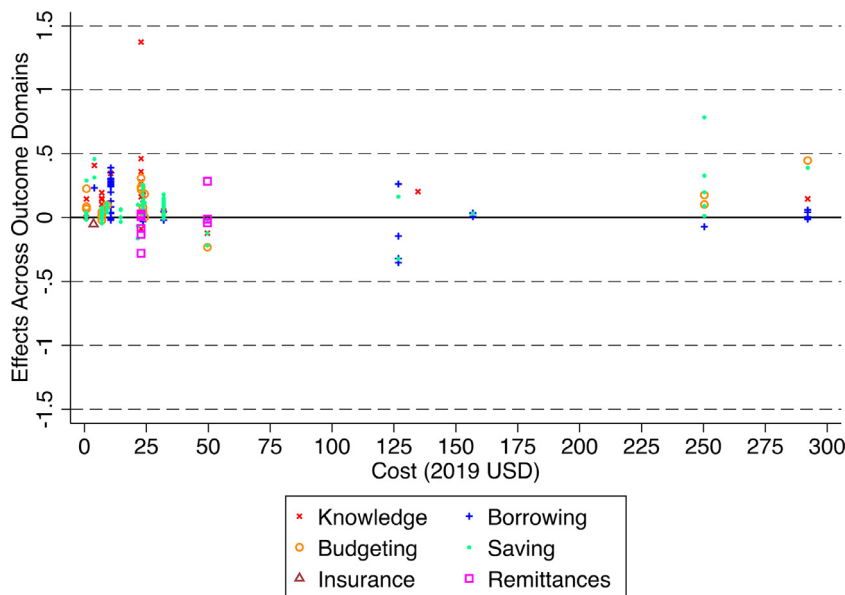


Fig. 5. Cost of intervention and treatment effects.

The graph depicts the cost and treatment effects (in standard deviations) for each outcome domain among the 20 experiments that report costs. Each data point is an effect size for an outcome studied. Fig. B6 in Appendix B provides a graph for each outcome domain that contains standard errors of the estimates.

ing the behaviors of teachers, parents, and possibly peers (Haliassos et al., 2019).

7.4. External validity

While a benefit of only including RCTs is that there is little debate regarding their internal validity, it is more common to study long-term effects in quasi-experimental settings. There is mounting quasi-experimental evidence that requiring US high school students to complete financial education prior to graduating improves long-term financial behaviors. These studies use a difference-in-difference strategy, comparing students who would have graduated just before and just after the requirement was in place within a state with a requirement, as well as across states with and without requirements over the same period.¹⁸

High school personal finance graduation requirements, which include standalone courses and personal finance standards incorporated into another required class or curriculum, show that financial education reduces non-student debt (Brown et al., 2016), increases credit scores (Brown et al., 2016; Urban et al., 2020), reduces default rates (Brown et al., 2016; Urban et al., 2020), shifts student loan borrowing from high-interest to low-interest methods

(Stoddard and Urban, 2020), increases student loan repayment rates (Mangrum, 2021), reduces payday loan borrowing for young adults (Harvey, 2019), and increases bank account ownership for those with only a high school education (Harvey, 2020). These studies also confirm the findings in the meta-analysis.

Additionally, calibrations of theoretical models suggest large estimates for the effects of financial knowledge on behavior, such as retirement savings (Lusardi et al., 2017).

Conclusion

Our analysis of the research on financial education using the most rigorous evaluation methods has three main findings. First, financial education treatment effects from RCTs have, on average, positive effects on financial knowledge and behaviors. This result is very robust: it holds up to accounting for publication bias, including only adequately powered studies, looking only at studies published in top economics and finance journals, and accounting for heterogeneity across studies. Financial education interventions have sizable effects on both financial knowledge (0.15–0.2 SD units) and financial behaviors (0.06–0.1 SD units). Thus, the treatment effects on financial knowledge are quite similar to or even larger in magnitude than the average effect sizes realized by educational interventions in other domains, such as math and reading (see Hill et al., 2008; Cheung and Slavin, 2016; Fryer, 2016; Kraft, 2020) and the effect sizes on financial behaviors are comparable to those realized in behavior-change interventions in the health domain (e.g., Rooney and Murray, 1996; Noar et al., 2007; Portnoy et al., 2008) or behavior-change interventions aimed at fostering energy conserving behavior (e.g., Karlin et al., 2015). Our findings are in stark contrast to

¹⁸ Cole et al. (2016) use this method but study “personal finance mandates” between 1957 and 1982, which often did not comprise course requirements but instead brought a representative from a bank to give a one-off lecture. The authors document no effects of the education on investment or credit management behaviors. This contrasted with Bernheim et al. (2001), who find that these same mandates improve investment behaviors, though they did not include state-level fixed effects in their analysis.

the findings presented in the first meta-analysis of the financial education literature (Fernandes et al., 2014). How can we interpret these differences in findings? While we are unable to precisely replicate the original estimate on RCTs presented in that meta-analysis (see Appendix D), we observe that the number of recent RCTs added to the database is driving the more positive result of financial education treatment effects on financial knowledge and behaviors. Additionally, we show that explicitly accounting for heterogeneity in studies and programs is crucial in assessing the average impact of financial education.

Second, while the results may be suggestive of some decay in treatment effects over time, there is no evidence to support or refute the decay of financial education treatment effects six months or more after the intervention. Since only six studies in our sample look at impacts 24 months beyond the intervention, we cannot rule out that this effect is statistically different from short-run effects. Because the literature is characterized by very few longer-term impact assessments, evidence on the sustainability of effects is inconclusive. What we can say, however, is that we do not find evidence for dramatic decay up to six months after the intervention.

Third, we document that the estimates of statistical effect sizes are economically significant. We further document that many of the financial education interventions studied in randomized experiments are cost-effective. This finding is crucial since the discussion of the effectiveness of financial education has focused on statistical effect sizes without considering their economic interpretation or the costs associated with financial education.

The evidence in this meta-analysis represents financial education interventions from 33 countries and six continents across various lifespans. In the analysis, we account for heterogeneity across interventions and identify the extent to which published results in this literature are subject to publication bias. However, there are still some limitations. Since few RCTs study long-run effects, it is hard to determine the long-run impacts of these interventions. The same is true for the quality of the data used to study changes in financial behaviors: few studies can link their experiments to administrative data, so the usual caveats of having to rely on self-reported survey data also apply to these studies. Future research should aim to collect longer-run administrative data or follow up with original participants from earlier field experiments. Finally, we encourage more researchers to report on the costs of their programs, in order to provide policymakers with an estimate of cost-effectiveness.

Research data

Replication data and code available at Mendeley Data (doi: 10.17632/svc3t8v3hs.2).

Funding

This work was supported by DFG through CRC TRR 190.

Acknowledgments

G. William Schwert was the editor for this article. We thank the anonymous referees, participants of the 5th Cherry Blossom Financial Education Institute in Washington, DC, and Michael Collins, Andrea Hasler, Rachael Meager, Olivia Mitchell, and Pierre-Carl Michaud for many helpful comments. We also thank Shawn Cole, Daniel Fernandes, Xavier Gine, John Lynch, Richard Netemeyer, and Bilal Zia for providing details about their studies.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jfineco.2021.09.022.

References

- Alan, S., Ertac, S., 2018. Fostering patience in the classroom: results from randomized educational intervention. *J. Polit. Econ.* 126 (5), 1865–1911.
- Andrews, I., Kasy, M., 2019. Identification of and correction for publication bias. *Am. Econ. Rev.* 109 (8), 2766–2794.
- Belman, D., Wolfson, P., Nawakitphaitoon, K., 2015. Who is affected by the minimum wage? *Ind. Relat. A J. Econ. Soc.* 54 (4), 582–621.
- Berg, G., Zia, B., 2017. Harnessing emotional connections to improve financial decisions: evaluating the impact of financial education in mainstream media. *J. Eur. Econ. Assoc.* 15 (5), 1025–1055.
- Beuermann, D.W., Jackson, C.K., 2018. The short and long-run effects of attending the schools that parents prefer. NBER Working Paper No. 24920. DOI: 10.3386/w24920.
- Bernheim, Douglas, et al., 2001. Education and saving: The long-term effects of high school financial curriculum mandates. *Journal of Public Economics* 80 (3), 435–465. doi:10.1016/S0047-2727(00)00120-1.
- Bloom, H.S., Hill, C.J., Black, A.R., Lipsey, M.W., 2008. Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *J. Res. Educ. Eff.* 1 (4), 289–328.
- Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2009. *Introduction to Meta-Analysis*. Wiley, Chichester, UK doi:10.1002/9780470743386.
- Boyer, M., d'Astous, P., Michaud, P.C., 2020. Tax-sheltered retirement accounts: can financial education improve decisions? *Rev. Econ. Stat.* doi:10.1162/rest_a_00973.
- Brodeur, A., Lé, M., Sangnier, M., Zylberberg, Y., 2016. Star Wars: the empirics strike back. *Am. Econ. J. Appl. Econ.* 8 (1), 1–32.
- Brodeur, A., Cook, N., Heyes, A., 2020. Methods matter: p-hacking and publication bias in causal analysis in economics. *Am. Econ. Rev.* 110 (11), 3634–3660.
- Bruhn, M., de Souza Leao, L., Legovini, A., Marchetti, R., Zia, B., 2016. The impact of high school financial education: evidence from a large-scale evaluation in Brazil. *Am. Econ. J. Appl. Econ.* 8 (4), 256–295.
- Brown, M., Grigsby, J., van der Klaauw, W., Wen, J., Zafar, B., 2016. Financial education and the debt behavior of the young. *Rev. Financ. Stud.* 29 (9), 2490–2522.
- Calderone, M., Fiala, N., Mulaj, F., Sadhu, S., Sarr, L., 2018. Financial education and savings behavior: evidence from a randomized experiment among low-income clients of branchless banking in India. *Econ. Dev. Cult. Chang.* 66 (4), 793–825.
- Camerer, C.F., Dreber, A., Forsell, E., Ho, T.H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmeld, A., Chan, T., Heikensten, E., 2016. Evaluating replicability of laboratory experiments in economics. *Science* 351 (6280), 1433–1436.
- Carpena, F., Cole, S., Shapiro, J., Zia, B., 2017. The ABCs of financial education. *Experimental evidence on attitudes, behavior, and cognitive biases*. *Manag. Sci.* 65 (1), 346–369.
- Carpena, Fenella, Zia, Bilal, 2020. The causal mechanism of financial education: Evidence from mediation analysis. *Journal of Economic Behavior & Organization* 177, 143–184. doi:10.1016/j.jebo.2020.05.001.
- Cheung, A., Slavin, R., 2016. How methodological features affect effect sizes in education. *Educ. Res.* 45 (5), 283–292.
- Choi, J.J., Laibson, D., Madrian, B.C., 2010. Why does the law of one price fail? An experiment on index mutual funds. *Rev. Financ. Stud.* 23 (4), 1405–1432.

- Cole, Shawn, et al., 2016. High School Curriculum and Financial Outcomes: The Impact of Mandated Personal Finance and Mathematics Courses. *Journal of Human Resources* 51, 656–698. doi:10.3368/jhr.51.3.0113-5410R1.
- Cole, S., Sampson, T., Zia, B., 2011. Prices or knowledge? What drives demand for financial services in emerging markets? *J. Financ.* 66 (6), 1933–1967.
- Collins, J.M., O'Rourke, C.M., 2010. Financial education and counseling – still holding promise. *J. Consum. Aff.* 44 (3), 483–498.
- Doi, Y., McKenzie, D., Zia, B., 2014. Who you train matters: identifying combined effects of financial education on migrant households. *J. Dev. Econ.* 109, 39–55.
- Evans, D., Yuan, F., 2019. Equivalent years of schooling: a metric to communicate learning gains in concrete terms. World Bank Policy Research Working Paper No. 8752, World Bank, Washington, DC.
- Duflo, Esther, Saez, Emmanuel, 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., Netemeyer, R.G., 2014. Financial literacy, financial education, and downstream financial behaviors. *Manag. J. Sci.* 60 (8), 1861–1883.
- Flory, J.A., 2018. Formal finance and informal safety nets of the poor: evidence from a savings field experiment. *J. Dev. Econ.* 135, 517–533.
- 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. DOI: 10.3386/w22130.
- Gibson, J., McKenzie, D., Zia, B., 2014. The impact of financial literacy training for migrants. *World Bank Econ. Rev.* 28 (1), 130–161.
- Gine, X., Karlan, D., Ngatia, M., 2013. Social Networks, Financial Literacy and Index Insurance. World Bank, Washington, DC.
- Haliassos, M., Jansson, T., Karabulut, Y., 2019. Financial literacy externalities. *Rev. Financ. Stud.* 33 (2), 950–989.
- Harbord, R.M., Higgins, J.P.T., 2008. Meta-regression in Stata. *Stata J.* 8 (4), 493–519.
- Harvey, M., 2019. Impact of financial education mandates on young consumers' use of alternative financial services. *J. Consum. Aff.* 53 (3), 731–769.
- Harvey, M., 2020. Does state-mandated high school financial education affect savings by low-income households? Working paper. https://static1.squarespace.com/static/5c4d314bb27e3999d515a9e4/t/5e0a1b2841180e2960023175/1577720633380/Harvey_FinEd_Savings_Working+Paper_v20191230.pdf.
- Hastings, J.S., Madrian, B.C., Skimmyhorn, W.L., 2013. Financial literacy, financial education, and economic outcomes. *Annu. Rev. Econ.* 5, 347–373.
- Hedges, L.V., Hedberg, E.C., 2007. Intraclass correlation values for planning group-randomized trials in education. *Educ. Eval. Policy Anal.* 29 (1), 60–87.
- Hedges, L.V., Tipton, E., Johnson, M.C., 2010. Robust variance estimation in meta-regression with dependent effect size estimates. *Res. Synth. Methods* 1 (1), 39–65.
- Hill, C.J., Bloom, H.S., Black, A.R., Lipsey, M.W., 2008. Empirical benchmarks for interpreting effect sizes in research. *Child Dev. Perspect.* 2 (3), 172–177.
- Imai, T., Rutter, T.A., Camerer, C.F., 2021. Meta-analysis of present-bias estimation using convex time budgets. *Econ. J.* 131 (636), 1788–1814.
- Ioannidis, J., Stanley, T.D., Doucouliagos, H., 2017. The power of bias in economics research. *Econ. J.* 127 (605), F236–F265.
- Kaiser, T., Menkhoff, L., 2017. Does financial education impact financial behavior, and if so, when? *World Bank Econ. Rev.* 31 (3), 611–630.
- Kaiser, T., Menkhoff, L., 2020. Financial education in schools: a meta-analysis of experimental studies. *Econ. Educ. Rev.* 78, 101930.
- Karlin, B., Zinger, J.F., Ford, R., 2015. The effects of feedback on energy conservation: a meta-analysis. *Psychol. Bull.* 141 (6), 1205–1227.
- Kraft, M.A., 2020. Interpreting effect sizes of education interventions. *Educ. Res.* 49 (4), 241–253.
- Lipsey, M.W., Wilson, D.B., 2001. *Practical Meta-Analysis*. Sage, Thousand Oaks.
- Lüthmann, M., Serra-Garcia, M., Winter, J., 2018. The impact of financial education on adolescents' intertemporal choices. *Am. Econ. J. Econ. Policy* 10 (3), 309–332.
- Lusardi, A., 2019. Financial literacy and the need for financial education: evidence and implications. *Swiss J. Econ. Stat.* 155 (1). doi:10.1186/s41937-019-0027-5.
- Lusardi, A., Mitchell, O.S., 2014. The economic importance of financial literacy: theory and evidence. *J. Econ. Lit.* 52 (1), 5–44.
- Lusardi, A., Michaud, P.C., Mitchell, O.S., 2017. Optimal financial knowledge and wealth inequality. *J. Polit. Econ.* 125 (2), 431–477.
- Mangrum, D., 2021. Personal finance education mandates and student loan repayment Working Paper, 10.2139/ssrn.3349384.
- Meager, R., 2019. Understanding the average impact of microcredit expansions: a Bayesian hierarchical analysis of seven randomized experiments. *Am. Econ. J. Appl. Econ.* 11 (1), 57–91.
- Miller, M., Reichelstein, J., Salas, C., Zia, B., 2015. Can you help someone become financially capable? A meta-analysis of the literature. *World Bank Res. Obs.* 30 (2), 220–246.
- Noar, S.M., Benac, C.N., Harris, M.S., 2007. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol. Bull.* 133 (4), 673–693.
- OECD, 2015. National Strategies for Financial Education. OECD/INFE Policy Handbook. <https://www.oecd.org/finance/National-Strategies-Financial-Education-Policy-Handbook.pdf>.
- Portnoy, D.B., Scott-Sheldon, L.A., Johnson, B.T., Carey, M.P., 2008. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials. *Prev. Med.* 47 (1), 3–16.
- Rooney, B.L., Murray, D.M., 1996. A meta-analysis of smoking prevention programs after adjustment for errors in the unit of analysis. *Health Educ. Q.* 23 (1), 48–64.
- Sayinzoga, A., Bulte, E.H., Lensink, R., 2016. Financial literacy and financial behaviour: experimental evidence from rural Rwanda. *Econ. J.* 126 (594), 1571–1599.
- Seshan, G., Yang, D., 2014. Motivating migrants: a field experiment on financial decision-making in transnational households. *J. Dev. Econ.* 108, 119–127.
- Stanley, T.D., 2001. Wheat from chaff: meta-analysis as quantitative literature review. *J. Econ. Perspect.* 15 (3), 131–150.
- Stoddard, C., Urban, C., 2020. The effects of state-mandated financial education on college financing behaviors. *J. Money Credit Bank.* 52, 747–776.
- Tanner-Smith, E.E., Tipton, E., 2014. Robust variance estimation with dependent effect sizes: practical considerations including a software tutorial in STATA and SPSS. *Res. Synth. Methods* 5 (1), 13–30.
- Tanner-Smith, E.E., Tipton, E., Polanin, J.R., 2016. Handling complex meta-analytic data structures using robust variance estimates: a tutorial in R. *J. Dev. Life Course Criminol.* 2 (1), 85–112.
- Urban, C., Schmeiser, M., Collins, J.M., Brown, A., 2020. The effects of high school personal financial education policies on financial behavior. *Econ. Educ. Rev.* 78, 101786.
- What Works Clearinghouse, 2014. WWC Procedures and Standards Handbook (Version 3.0). <https://eric.ed.gov/?id=ED544775>.
- Xu, L., Zia, B., 2012. Financial Literacy around the World : An Overview of the Evidence with Practical Suggestions for the Way Forward. Policy Research Working Paper; No. 6107. World Bank, Washington, DC.