GFLEC
GLOBAL FINANCIAL LITERACY EXCELLENCE CENTER

Financial Education Affects Financial Knowledge and Downstream Behaviors

March 4, 2021

Tim Kaiser, Annamaria Lusardi, Lukas Menkhoff, and Carly Urban

GFLEC
GLOBAL FINANCIAL LITERACY EXCELLENCE CENTER

Main question and relevance

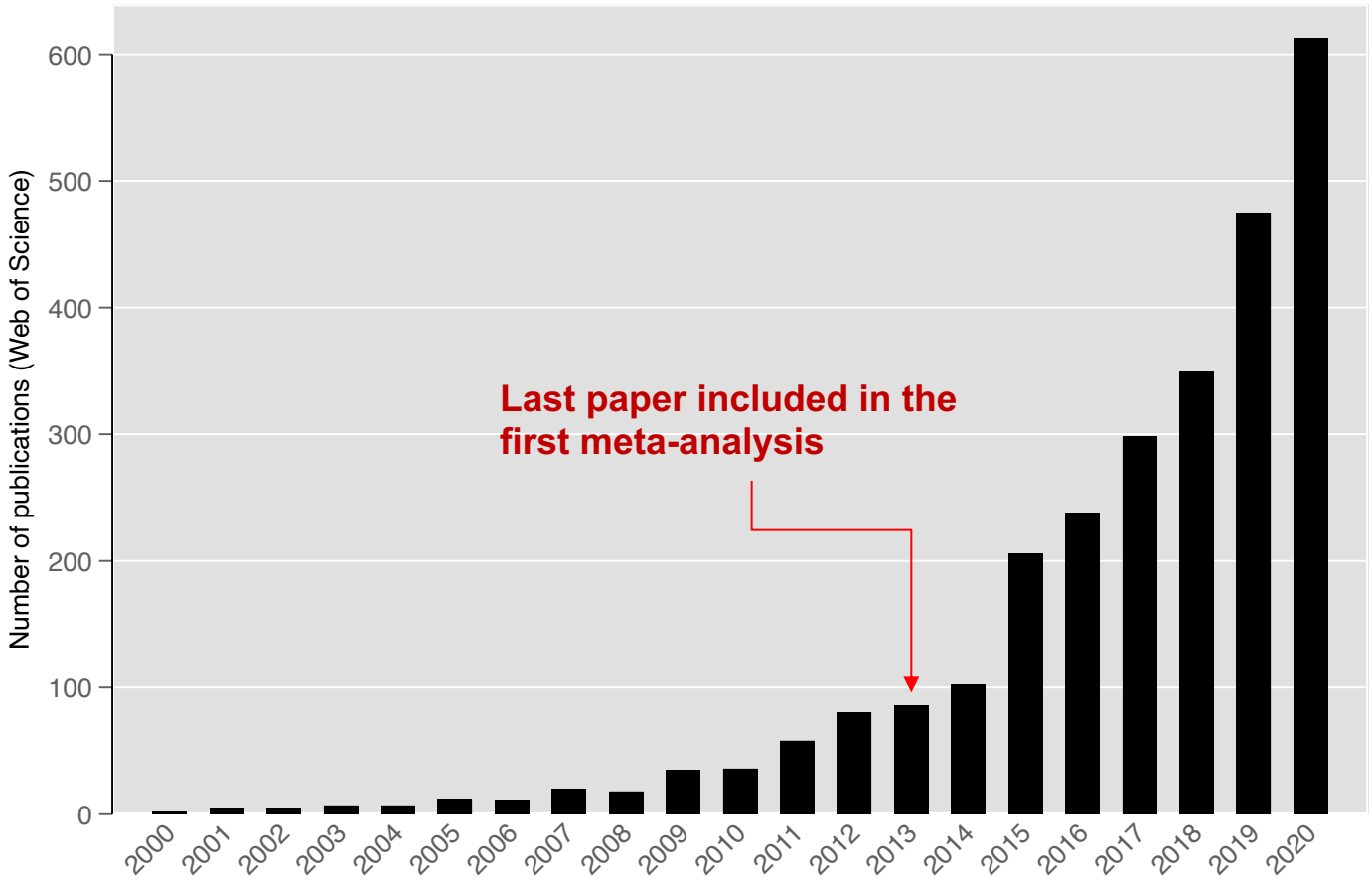
- Do financial education interventions (generally) work?
- More than 70 countries have designed or are designing national strategies for financial literacy
- Purpose of this paper is to derive empirical benchmarks for treatment effects of financial education interventions

Previous meta-analyses on financial education

- The first meta-analysis by Fernandes, Lynch, and Netemeyer (2014, *ManSci*) has been widely cited to provide evidence of ineffectiveness of financial education in general:
  - "We find that interventions to improve financial literacy explain only 0.1% of the variance in financial behaviors studied" (page 1861)
  - "Intervention effects may decay over time – the case for 'just in time financial education'."(page 1866)
- Other meta-analyses with different foci (specific outcomes and target groups) (Miller et al. 2015, Kaiser and Menkhoff 2017, 2020) have been published since, but have not moved the priors of sceptics

GFLEC
GLOBAL FINANCIAL LITERACY EXCELLENCE CENTER

Number of journal articles on “financial literacy” over time



GFLEC
GLOBAL FINANCIAL LITERACY EXCELLENCE CENTER

What we do in this paper

- (1) We take stock of the new evidence
  - Focus on RCTs, which have high internal validity
  - Include all earlier studies and more than quintuple the number of RCTs (relative to the first meta-analysis)
    - Many more studies in top economics-journals
    - Can look at different types of behavior in addition to financial knowledge

What we do (cont.)

- (2) Meticulous meta-analysis of these RCTs:
  - Account for heterogeneity in the effects of financial education treatments
  - Probe sensitivity of results to the choice of model and interpretation of results
  - Consider the power of underlying studies
  - Considering publication bias
  - Analysis of intensity and decay of effects
  - Subgroup analyses

What we do in this paper (cont.)

- (3) Calculations of the economic size of the effects and analysis of cost-effectiveness
  - What do the statistical effect sizes mean in economic terms?
  - What is the average cost of financial education and is it cost-effective?

A preview of the findings

We found that:

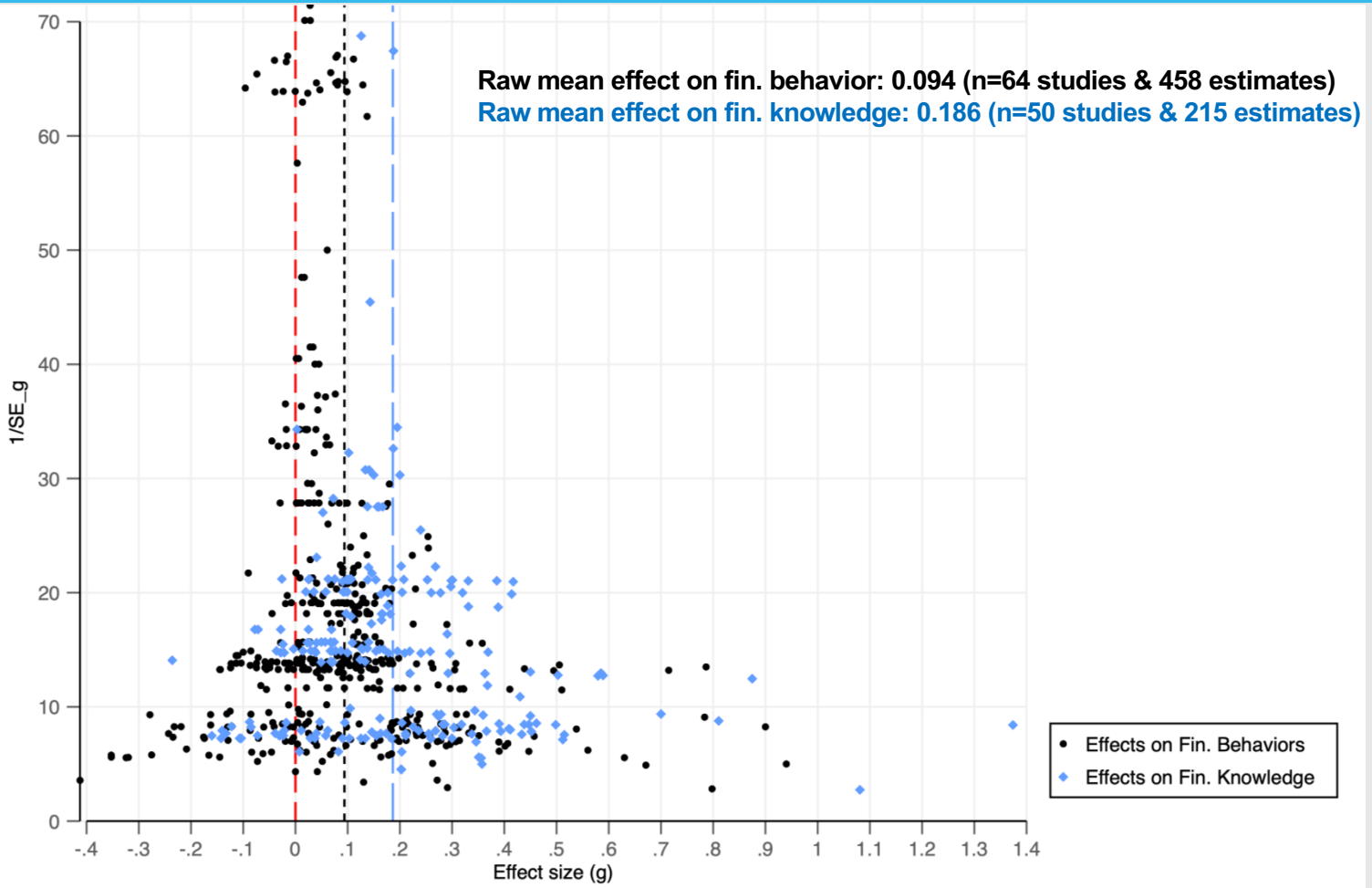
- The estimated effect of financial education is ***at least three times as large*** as the effect documented in Fernandes et al. (2014)
- Accounting for heterogeneity in true effects, effects are ***more than five times as large*** as the effects reported in Fernandes et al. (2014)
- **Results are robust** to identification of and correction for publication bias in the literature

Data

- Our study includes 76 RCTs from 33 countries with over 160,000 individuals across the life-cycle and of different socio-economic status
- Effects are measured after 30 weeks, on average, and up to more than two years.

GFLEC
GLOBAL FINANCIAL LITERACY EXCELLENCE CENTER

Raw data from 76 RCTs: Financial education treatment effects



A primer on meta-analysis

Main evidence aggregation issues:

- Outcomes and outcome units may vary across studies (→ use scale free SD units)
- Nested data: Papers may study multiple outcomes (→ consider in estimation and inference)
- Heterogeneity: Interventions vary across studies; e.g., from giving an informational brochure to time-intense education programs (→ reflect this in a model)

A primer on meta-analysis

Meta-analysis model:

- Consider a set of *j* randomized experiments, each of them reporting *i* estimates of causal treatment effects relative to a control group
- Allow different experiments to result in different treatment effects caused by the educational interventions (i.e., heterogeneity in true effects)
- Goal of this aggregation is to arrive at a "general effect" of financial education → choose weights for each observation that reflect the size of study (random sampling error) and the differences in site-specific results (heterogeneity in true effects)

Formal model

$$y_{ij} = \beta_0 + \upsilon_j + \epsilon_{ij}$$

y_{ij} is the ith treatment effect estimate within each study j.

β_0 is the mean of the distribution of true effects, i.e., the "general effect of financial education"

υ_j is a study-level random effect with $\upsilon_j \sim N(0, \tau^2)$, i.e., the true effects can vary between (but not within) studies.

$\epsilon_{ij} \sim N(0, \sigma_{ij}^2)$ is the residual of the ith treatment effect estimate within each study j

- We observe both y_{ij} and σ_{ij}^2 from the data
- τ^2 needs to be estimated

Formal model: Choosing the study weights

Step 1: Estimate τ^2 from the data

Step 2: Account for multiple correlated effects within studies

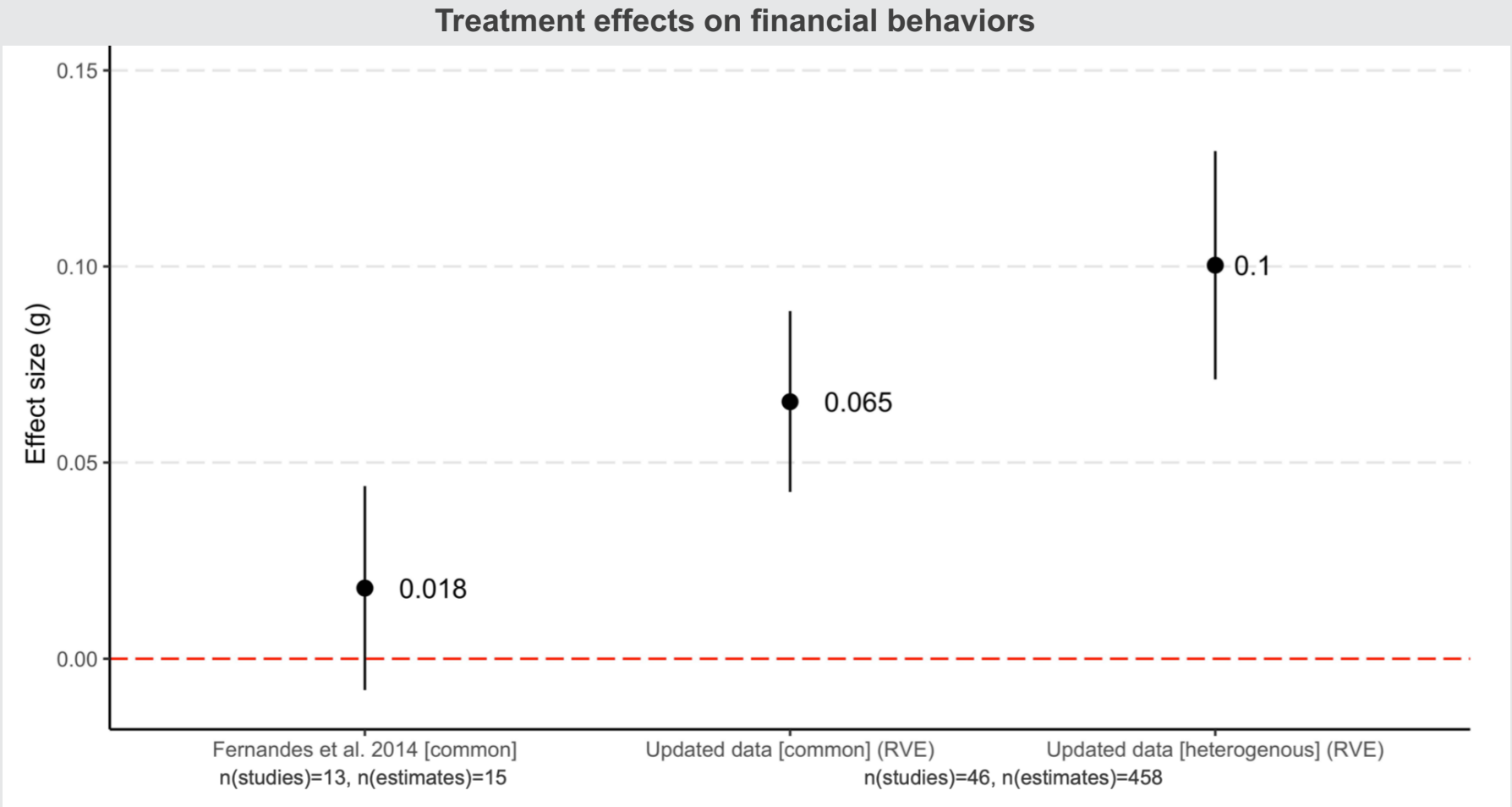
Weight: $w_{ij} = \left\{\left(\tau^2 + \frac{1}{k_j}\sum_{k_j=1}^{k_i}\sigma_{ij}^2\right)\left[1 + \left(k_j - 1\right)\rho\right]\right\}^{-1}$

Step 3: Estimate β_0 and the associated 95% confidence interval with weighted least squares

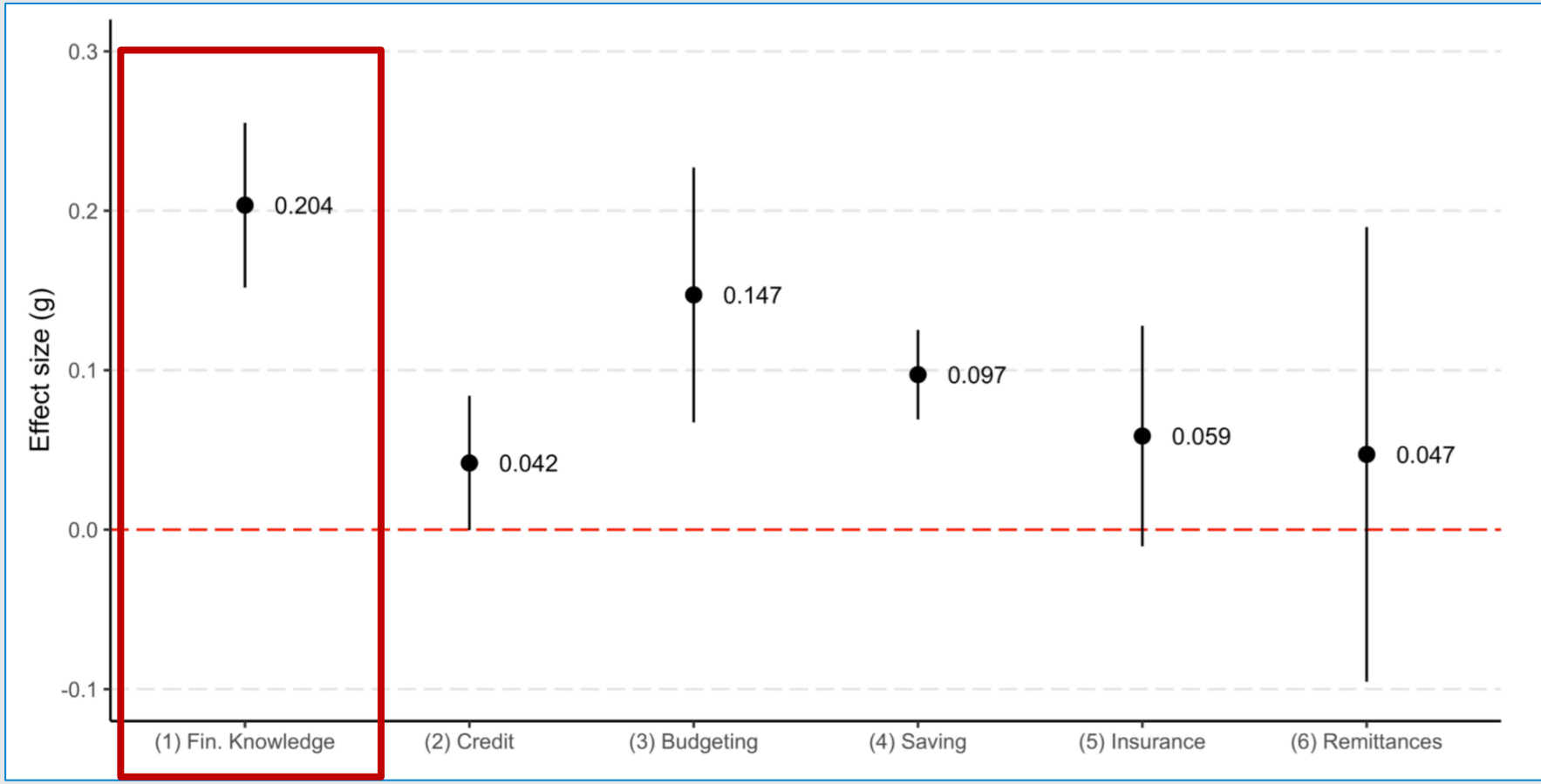
Step 4: Identify publication bias and correct for it (Andrews and Kasy 2018, *AER*)

GFLEC
GLOBAL FINANCIAL LITERACY EXCELLENCE CENTER

Comparison of the new evidence to the result in Fernandes et al. (2014)



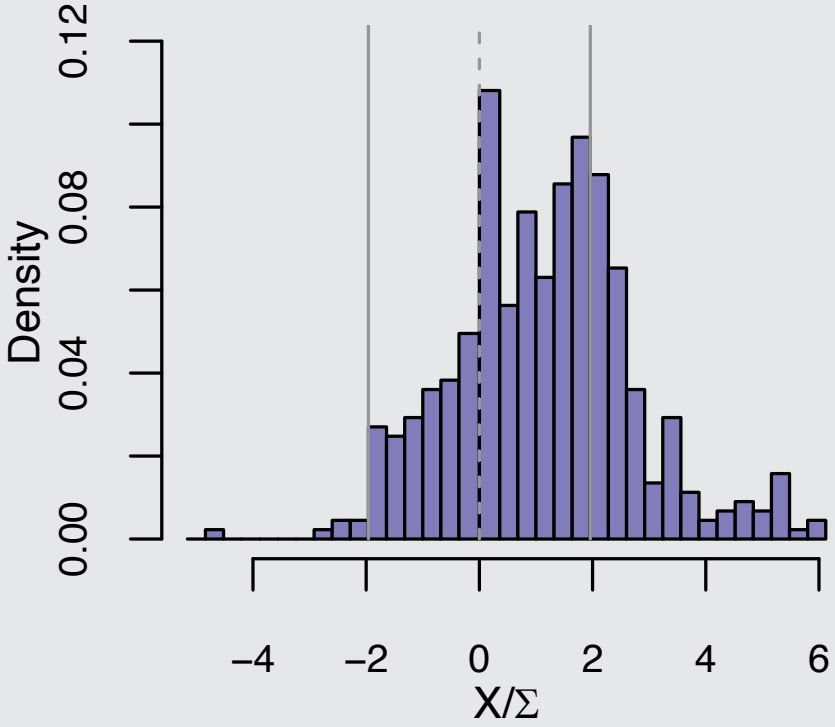
Treatment effects by outcome domain

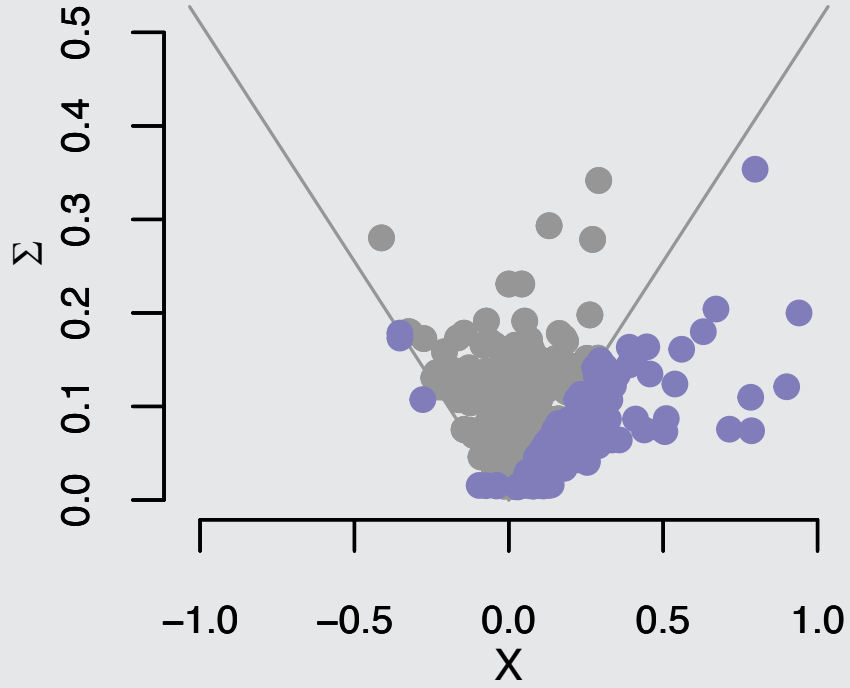


What about publication bias?

- Publication bias refers to the problem of authors (or journal editors) favoring the publication (selection) of statistically significant results
- Leaving this selection unadressed can lead to a biased assesment of a mean effect in a given literature
- Andrews and Kasy (2018, *AER*) develop a method for identifying and correcting publication bias using a step function approach

Results of Andrews and Kasy (2018) approach to this literature





Results of Andrews and Kasy (2018) approach to this literature

Table 2: Identification of and correction for publitcation bias in the financial education literature

(a) Treatment effects on *financial behaviors*				(b) Treatment effects on *financial knowlege*			
(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 $Z = 1.65$)	
$\bar{\beta}_0$	λ_p	$\bar{\beta}_0$	λ_p	$\bar{\beta}_0$	λ_p	$\bar{\beta}_0$	λ_p
0.057	0.303	0.050	0.256	0.150	0.150	0.160	0.250
(0.001)	(0.071)	(0.007)	(0.051)	(0.037)	(0.126)	(0.040)	(0.190)

Notes: This table presents results from non-parametric identification of and correction for publication bias based on the method described in Andrews and Kasy (2018) (see Andrews and Kasy 2018, Appendix C). $\bar{\beta}_0$denotes the estimate of the true treatment effect in latent studies (i.e., the bias corrected treatmen effect) and $\lambda_p$denotes the estimated conditional publication probability (p) based on the Z-statistic (y_{ij}/σ_{ij}) as specified in the repective 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, repectively. Columns (2) and (4) show estimates for 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, repectively. Standard errors (clustered at the study-level) are shown in parentheses.

How large are the effects?

- Effects of financial education on *financial knowledge* are comparable to studies on math and reading (Hill et al. 2008; Cheung and Slavin 2016; Fryer 2016).
- Effects of financial education on *financial behaviors* are comparable to meta-analyses of behavior change interventions in other domains
  - anti-smoking (Rooney & Murray 1996)
  - tailored printed health interventions (Noar et al. 2017)
  - energy conservation (Karlin et al. 2015)

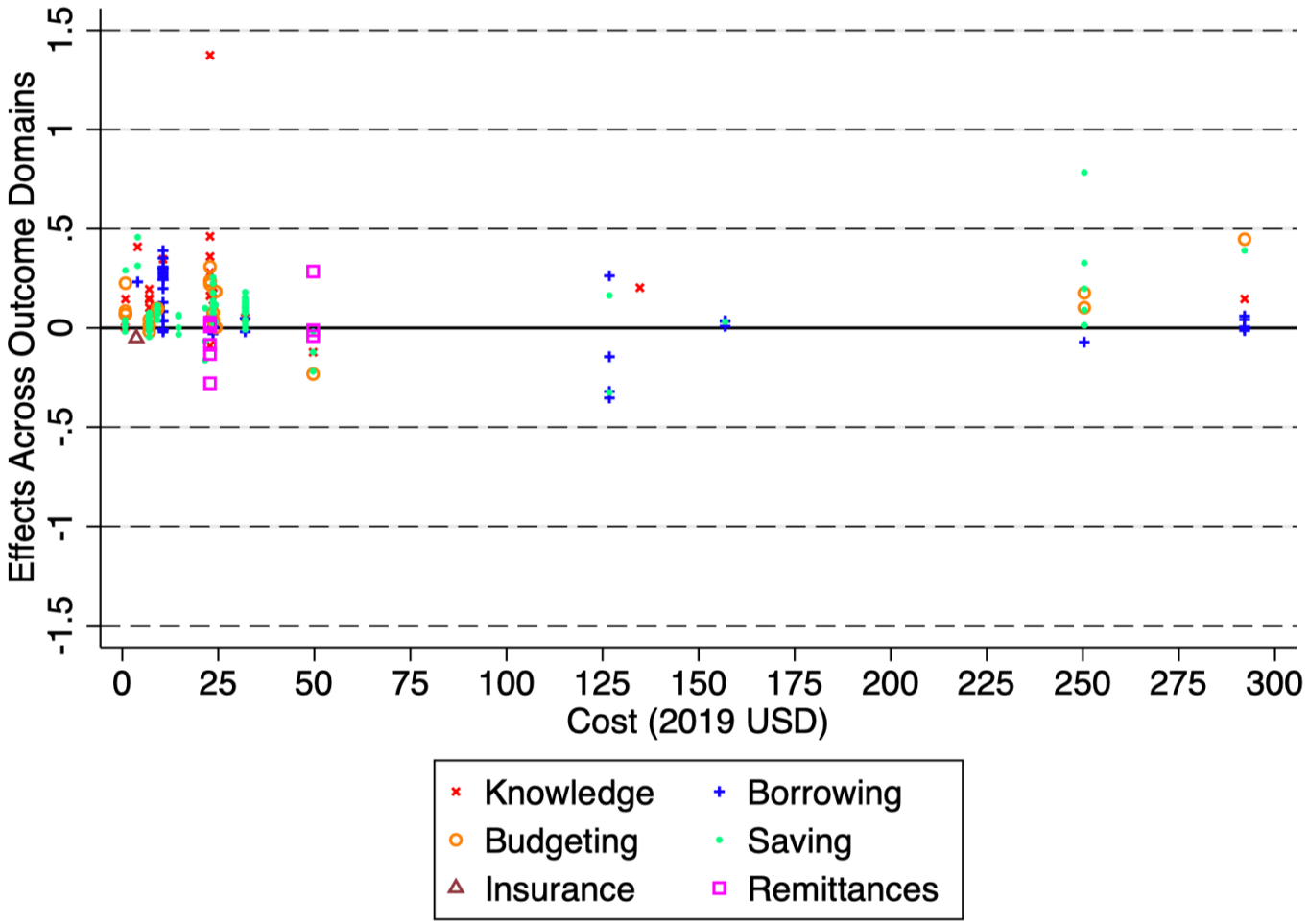
A scheme for interpreting effect sizes from causal studies (Kraft 2018)

		Cost-Effectiveness Ratio (ES/Cost)		
		Cost Per Pupil		
		Low (< $500)	Moderate ($500 to <$4,000)	High ($4,000 or >)
Effect Size	Small (<.05)	Small ES / Low Cost	Small ES / Moderate Cost	Small ES / High Cost
	Medium (.05 to <.20)	Medium ES / Low Cost	Medium ES / Moderate Cost	Medium ES / High Cost
	Large (.20 or >)	Large ES / Low Cost	Large ES / Moderate Cost	Large ES / High Cost

Notes: ES = Effect Size

(Kraft 2018, p. 20)

Costs and effect sizes of financial education interventions



Are interventions cost-effective?

- Using Kraft's (2019) scale of educational interventions, effects are "medium/large."
- Average intervention has low cost per participant (mean costs are $60.40 and median costs are $22.90)
- With the data we have, for "medium effect sizes," Kraft's educational intervention scale would say average cost per participant of $60 implies "low cost."

GFLEC
GLOBAL FINANCIAL LITERACY EXCELLENCE CENTER

Subgroup analyses

Subgroup	Effect size (g)	SE	95% CI Lower bound	95% CI Upper bound	n(Studies)	n(effects)
Panel A: Treatment effects on *financial behaviors*						
(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 to 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 type of publication						
Top econ. journals	0.0833	0.0235	0.0325	0.1342	15	161
Other publications	0.1075	0.0183	0.0704	0.1445	49	297
(e) By delay between treatment and measurement of outcomes						
Delay of < 6 months	0.0991	0.0169	0.0645	0.1337	34	180
Delay of ≥ 6 months	0.0710	0.0137	0.0425	0.0995	28	260
Delay of ≥ 12 months	0.0878	0.0200	0.0450	0.1308	18	134
Delay of ≥ 18 months	0.0653	0.0192	0.0209	0.1098	10	49
Delay of ≥ 24 months	0.0574	0.0225	0.0013	0.1136	7	32

Findings among sub-groups (1st block of the table)

- No significant differences between high-income and developing economies (effects on behavior)
- No significant differences between low-income individuals and general population
- No differences across publications (if in top econ journals or not)
- Financial education “works” for all age groups

GFLEC
GLOBAL FINANCIAL LITERACY EXCELLENCE CENTER

Do the effects decay over time? (2nd block of the table)

- Different from the initial meta-analysis (Fernandes et al 2014), we find **no evidence to support or refute decay** of effects 6 months or more after the intervention.
- Note that their *prediction* was based on a very small sample of studies.
- The effect on financial knowledge is estimated to be positive after more than one year in 5 studies.
- The effect on behavior is estimated to be positive after more than two years after intervention in 7 studies

Main takeaways

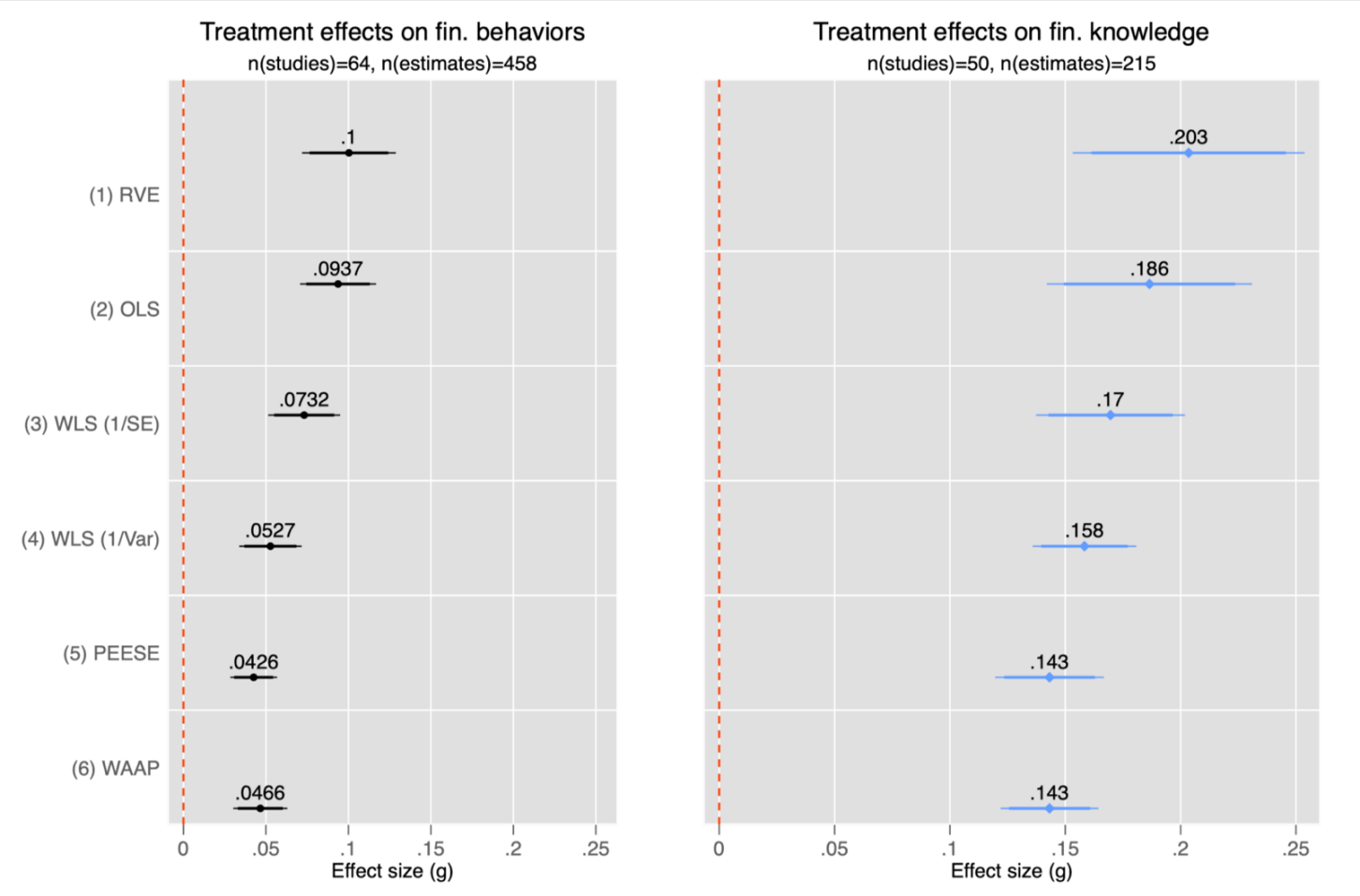
1) Financial education works! Recent work shows clear evidence of positive effects of financial education on financial behaviors (+knowledge)
   - Statistical effect size is at three times as large as the effect in Fernandes et al. (2014)
   - It may be up to five times as large (when allowing for between-study heterogeneity in true effects)
   - Robust to a lot of different approaches to meta-analysis and even when accounting for publication selection for statistical significance
2) Policy recommendations should be based on "economic effect sizes", not statistical effect sizes
3) No evidence of "rapid decay" but no evidence against it either

GFLEC
GLOBAL FINANCIAL LITERACY
EXCELLENCE CENTER

Additional slides

GFLEC
GLOBAL FINANCIAL LITERACY EXCELLENCE CENTER

Results are robust to choosing lots of different models and also when correcting for publication selection bias



Descriptive statistcs

Table 1: Descriptive statistics

Variable	Obs.	Mean	Median	Std. Dev.	Min.	Max.
Hedges' *g*	677	0.123	0.098	0.183	-0.413	1.374
SE (*g*)	677	0.084	0.072	0.049	0.007	0.365
Time span (in weeks)	643	30.239	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)	650	33.480	38.300	12.480	8.500	55.000
Children (< age 14)	677	0.075	-	-	0.000	1.000
Youth (age 14 to 25)	677	0.201	-	-	0.000	1.000
Adults (> age 25)	677	0.724	-	-	0.000	1.000
Low income (yes=1)	677	0.725	-	-	0.000	1.000
Developing economy (yes=1)	677	0.604	-	-	0.000	1.000
Top econ journal (yes=1)	677	0.267	-	-	0.000	1.000

Note: Descriptive statistics at the extracted estimate-level, i.e., we consider the total of 677 treatment effects reported in 76 RCTs.

Rapid decay in effects?

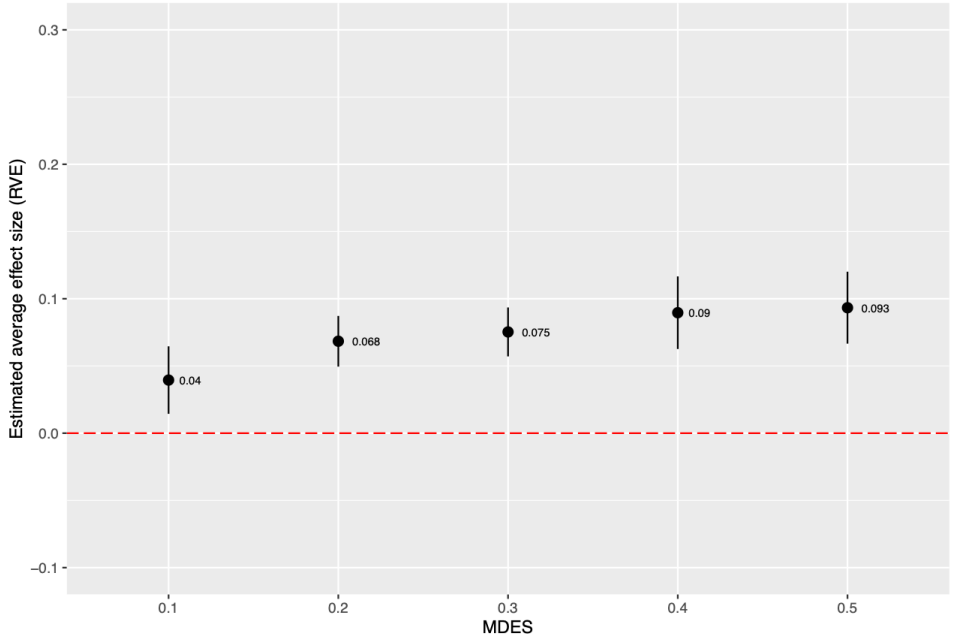
	(1) Effect size (g)
Intensity	0.0043
	(0.0024)
Intensity× Intensity	-0.0000
	(0.0000)
Delay	-0.0018
	(0.0052)
Delay × Delay	-0.0000
	(0.0002)
Intensity × Delay	-0.0001
	(0.0003)
n (Studies)	52
n (Effect sizes)	419

Note: This table reruns the main analysis of the result presented in Figure 4 in Fernandes et al. (2014) with updated data. Intensity is (mean-centered) number of hours of instruction, Delay is delay between treatment and measurement of outcomes in months. Results from RVE (random-effects assumption). Robust standard errors in parentheses. Assumed $\rho = 0.8$. Estimated τ^2=0.0111.

- Standard errors for the coefficients are very large, so there is a lot of uncertainty around this prediction.

Power

Figure B6: Power in the financial behavior sample



Notes: Average effect size of treatment effects on financial behaviors (from RVE) within the set of studies with the respective MDES. Minimum detectable effect size (MDES) at $\alpha = 0.05$ and $1 - \beta = 0.8$. The number of observations for the sample with a MDES of 0.1 is 60 effect sizes within 7 studies. For MDES=0.2, the sample size is 198 effect size estimates within 31 studies. For MDES=0.3, the sample size is 326 effect sizes in 45 studies. For MDES=0.4, it is 402 effect sizes within 53 studies. For MDES=0.5, it is 443 effect size estimates within 60 studies. The mean MDES in the entire sample is 0.23 SD units. The median MDES in the entire sample of effect sizes is 0.2 SD units (Carpena et al. 2017). The smallest MDES is 0.04 SD units (Frisancho 2018). The largest MDES is 1 SD unit (Reich and Berman 2015). Dots show the point estimate, and the solid lines indicate the 95% confidence interval.

Bayesian hierarchical analysis

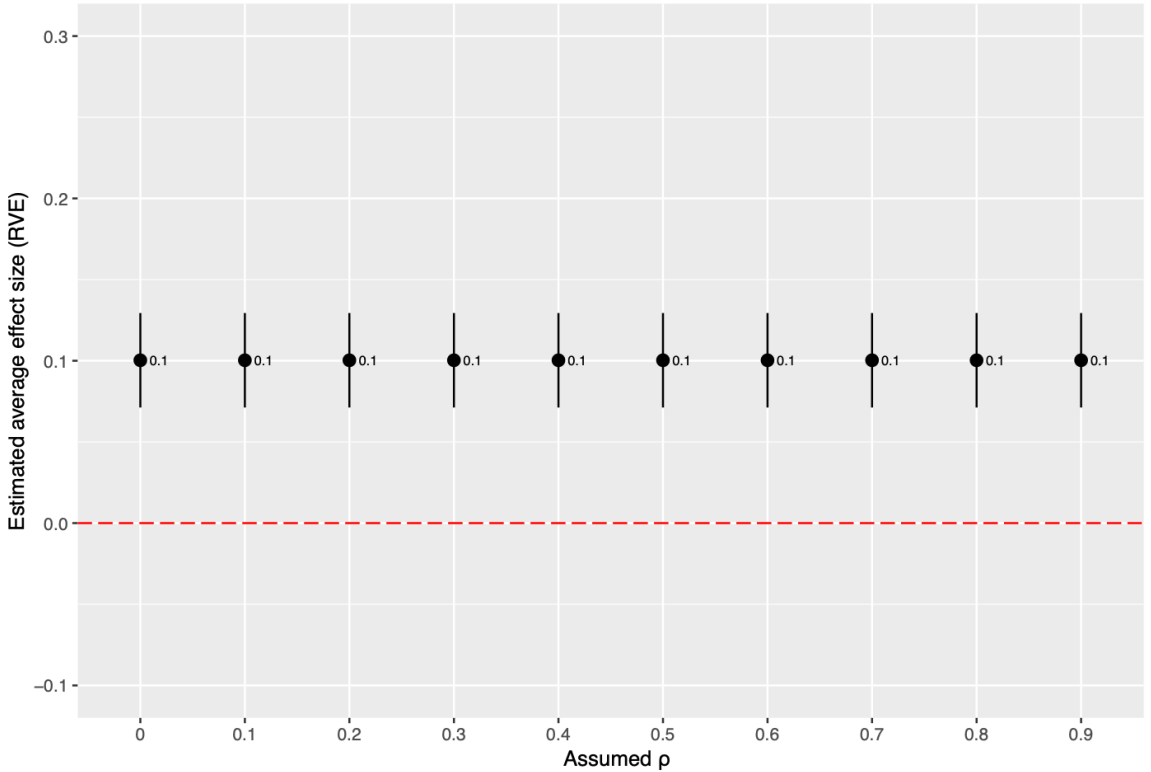
Table B5: Results from Bayesian Hierarchical Models (BHM)

Outcome	Model	Posterior mean [95% uncertainty interval]	Hyper SD [95% uncertainty interval]	I^2 [95% uncertainty interval]
Financial behaviors (n= 64 studies)	Partial pooling	0.090 [0.067; 0.117]	0.081 [0.057; 0.111]	36% [22%; 52%]
	Full pooling	0.055 [0.05; 0.059]	-	-
Financial knowledge (n=50 studies)	Partial pooling	0.210 [0.159; 0.264]	0.160 [0.12; 0.21]	67% [54% to 77%]
	Full pooling	0.159 [0.145; 0.174]	-	-

Notes: Results from fitting Bayesian Hierarchical Models in *Stan* using the R package *baggr* by Wiecek and Meager (2020). The "partial pooling" model fits a Rubin model whereas the "full pooling" model assumes no heterogeneity in true effects by definition (see discussion of the common effect assumption in the main text). All estimations rely on synthetic effect sizes (one summary treatment effect estimate per study) and use the default (i.e., very weak) Gaussian priors which assume that treatment effects are small unless the data provides evidence to the contrary.

Sensitivity to choice of within study correlation of effects

Figure B4: (Non-)Sensitivity of RVE estimate to the choice of ρ (treatment effects on financial behaviors)



Notes: Figure shows results from (random effects) RVE for different choices of assumed ρ.

Economic vs. statistical significance

- Fernandes et al. (2014) effect size measure creates the illusion of miniscule effects, when they can be economically significant.
  - "variance explained" is a misleading concept

- Consider the following example:
  - Median effect of structured pedagogy interventions in developing countries = 0.13 SD units. (Evans et al. 2019)
  - In the Fernandes et al. (2014) metric: this intervention explains 0.36% of the variance in learning outcomes.
    - ➢ Seems small?
  - Evans et al. (2019) report that this effect = ~0.6 years of "business as usual schooling"
  - In separate analysis they estimate the returns to literacy in Kenya. The net present value of this intervention is 1,338 USD at an average annual income of 1,079 USD in 2015 PPP.
    - ➢ Economically, this effect appears to be large.

External Validity

- There are concerns that RCTs may have limited external validity.
- This study increases the number of individuals in the interventions from Fernandes, Lynch, and Netermeyer (2014) from 23,000 to over 140,000.
  - But what about scale?
- Findings are consistent with recent work studying post-2000 state-mandated financial education in U.S. high schools that relies upon quasi-experimental research. (Brown et Al, 2016; Harvey, 2019; Urban et Al, 2018; Stoddard and Urban, 2019)
- Findings also consistent with large-scale RCTs, such as the school-based RCTs (e.g., Frisancho (2018))

GFLEC

GLOBAL FINANCIAL LITERACY
EXCELLENCE CENTER

(202) 994-7148 | gflec@gwu.edu | www.gflec.org