NBER WORKING PAPER SERIES

SELECTION INTO FINANCIAL EDUCATION AND EFFECTS ON PORTFOLIO CHOICE

Irina Gemmo
Pierre-Carl Michaud
Olivia S. Mitchell

Working Paper 31682
http://www.nber.org/papers/w31682

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2023, Revised May 2024

For useful comments, we thank Bertrand Achou, David Boisclair, Claudio Daminato, Franca Glenzer, Christian Kubitza, Christian Waibel and participants at the Cherry Blossom Financial Education Institute 2023, the 13th annual meeting of the Society for Experimental Finance, and seminar participants at ETH Zurich and HEC Montreal. Further, we thank Alexa Samuel and Yong Yu for valuable research assistance and the Pension Research Council of Wharton School at the University of Pennsylvania, as well as the Retirement and Savings Institute at HEC Montréal, for generous research funding. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w31682

Selection into Financial Education and Effects on Portfolio Choice
Irina Gemmo, Pierre-Carl Michaud, and Olivia S. Mitchell
NBER Working Paper No. 31682
September 2023, Revised May 2024
JEL No. G11,G41,G53

ABSTRACT

We study the effect of a financial education intervention on portfolio choices in a unique incentivized setting that allows us to investigate selection into the intervention and treatment effect heterogeneity. After directly eliciting willingness to pay for the financial education, we find that the more financially literate, those confident they can apply the knowledge acquired in the intervention, and those expecting higher returns are willing to pay more. Using portfolio allocation tasks, we show that the financial education leads to better outcomes according to welfare metrics tailored to respondent risk aversion: there is a 20 p.p. increase in the fraction with a welfare gain, representing on average 3.2 p.p. of wealth. Those most willing to participate are those who gain the most from the education, which has important implications for the design of financial education across a variety of settings.

Irina Gemmo
Assistant Professor
Finance Department
HEC Montreal
Canada
irina.gemmo@hec.ca

Pierre-Carl Michaud
HEC Montréal
3000 Côte-Sainte-Catherine Road
Montréal QC H3T 2A7
and NBER
pierre-carl.michaud@hec.ca

Olivia S. Mitchell
University of Pennsylvania
The Wharton School
3620 Locust Walk, St 3000 SH-DH
Philadelphia, PA 19104-6302
and NBER
mitchelo@wharton.upenn.edu

1 Introduction

Financial literacy levels in developed countries are surprisingly low: many in the U.S., Europe, and Japan lack an understanding of interest rate compounding, inflation, and risk diversification (Lusardi and Mitchell, 2023); mutual fund expenses (Choi et al., 2010); and retirement planning (Lusardi and Mitchell, 2007). As a result, they earn lower returns and accumulate wealth far less effectively (Lusardi and Mitchell, 2014; Lusardi et al., 2017). Moreover, financial advice is far from a perfect substitute for financial sophistication (Hackethal and Inderst, 2013). Accordingly, enhancing financial knowledge can be critical for better financial decision making throughout the lifetime. This has become particularly clear to employers and policymakers in the wake of the Global Financial Crisis and the pandemic, leading them to devote meaningful resources to educate employees so as to reduce their financial stress, explain how to invest in their retirement plans, and possibly enhance their productivity on the job (TIAA, 2004; Miller, 2016).

Despite the clear need for enhancing financial literacy in the workplace and the general population, evidence is mixed regarding the actual effectiveness of financial education programs for adults.[1] In particular, when participation is voluntary, measured program effects may be biased if one compares outcomes for those who do take advantage of the education and those who do not. Indeed, selection into treatment has been understudied in this literature, though real-world programs inevitably involve voluntary participation with the potential for selection.[2] The average effect of the education on those treated could then

[1]See for instance Agarwal et al. (2011) and Kaiser et al. (2022) for reviews of financial education program effectiveness across a variety of programs.

[2]Evaluations of mandatory financial literacy education for high school students suffer less from selection bias, but the topics covered in school (budgeting, student loans) differ from topics commonly covered in the workplace setting (e.g., saving and investing for retirement; see Kaiser et al. (2022). Even in so-called mandatory programs for adults, there is non-random attrition or refusal to participate; see Collins (2013).

depend on who participates in the program, making it important to better understand the mechanisms producing such selection. To the best of our knowledge, there are no studies examining the effects of financial education on adults' portfolio choices which investigate how selection into the financial education shapes observed outcome measures.

Our first contribution is to develop a unique experimental setting where we elicit individuals' willingness to pay to participate in financial education, and we quantify the effect of the education on investment behavior. We conduct a large-scale incentivized experiment on a representative sample of Canadian households in which we first ask subjects to allocate an endowment across three different hypothetical assets that differ in expected return and volatility. After observing respondents' allocation decisions, we inform them that they will face the same allocation exercise again later in the survey, and that their final payouts will depend on their actual allocation choices. Next, we randomly offer some individuals the opportunity to acquire financial knowledge that can help them with the portfolio allocation task. This provides content on portfolio diversification and risk-adjusted portfolio returns, similar to that offered in other educational settings, e.g., Ambuehl et al. (2022). We elicit respondents' willingness to pay in an incentive compatible manner and tie their decisions to the actual payouts they receive upon survey completion.

To measure the causal effect of the education, we randomize who is offered the intervention, while the likelihood of participating in the intervention is directly tied to how much respondents are willing-to-pay for the intervention. Exploiting the framework of Imbens and Angrist (1994), we estimate the average effect on the treated (ATT), the relevant measure to assess how those who voluntarily participate perform as a result of the intervention, using instrumental variable techniques which we compare to difference-in-difference estimates of the same effect. Next, we study how selection impacts estimates by using the fact that, conditional on the elicited willingness-to-pay, assignment to the education is random. Using both control function and matching approaches, we compute the ATT estimates and study how the educational treatment effect varies depending on the probability of being treated. This randomization is the result of using an incentive-compatible elicitation mechanism for the willingness to pay (Becker et al., 1964).

Our second contribution is to construct preference-based outcomes, tailored to respondents' risk aversion, to measure the effect of the education on respondent welfare. We do this accounting for the fact that risk aversion is imperfectly observed for each respondent using a conventional risk aversion elicitation method (Holt and Laury, 2002). This allows us to test whether those willing to pay more are also those who benefit more from the education, a prediction that arises naturally from a rational model of financial literacy acquisition (Lusardi et al., 2020). We also assess how the intervention influences the use of heuristics for portfolio allocation, such as 1/K (splitting the allocation equally across funds), and return-chasing allocations (where respondents systematically pick the asset with the highest expected return irrespective of volatility (Thaler et al., 2001)). Finally, we evaluate the effect of the financial education on investment performance by measuring the change in the distance of portfolio allocations from the efficient frontier.

We document that the education intervention increases heterogeneity in portfolio allocations and leads participants to customize their portfolios, moving away from allocations using simple heuristics. According to our welfare metrics, those who receive the financial education also do better: there is a 20 p.p. increase in the fraction with a welfare gain, worth, on average, 3.2 p.p. of wealth. We find that those who are more willing to participate are more financially literate, are more likely to expect to be able to apply the knowledge they will receive in the intervention, and expect higher returns from the intervention. The selection analysis reveals that they indeed benefit more from the intervention. The financial education effects are positively correlated with the willingness to pay. Hence, we find substantially higher average effects for those treated than average effects in the population offered the treatment. This is consistent with a selection model where participants self-select based on the expected gain from the treatment (Lusardi et al., 2020).

In what follows, Section 2 briefly discusses relevant literature. Section 3 outlines the experiment and the outcomes we measure. Section 4 analyzes the effect of the financial education. Section 5 investigates selection and treatment effect heterogeneity, exploiting the unique features of the experiment. Finally, Section 6 concludes.

2 Related Literature

Our paper contributes to the household finance literature exploring individuals' willingness to acquire financial knowledge and the impact this knowledge has on investment outcomes. Jappelli and Padula (2013) present a two-period model where savers can acquire financial knowledge to boost the return on their savings. They predict that those with a higher propensity to save are also more willing to invest in financial knowledge; accordingly, they hypothesize a complementarity between saving and financial knowledge. Exploiting this complementarity to explain wealth inequality, Lusardi et al. (2017) calibrate a multi-period stochastic life cycle model in which savers choose between investing in financial knowledge and private consumption. In these models, financial knowledge is akin to human capital. Savers choose their investment in financial knowledge by comparing the marginal cost of investing in financial knowledge (measured in money and time) to the marginal benefit associated with greater knowledge. One key benefit is to obtain higher (risk-adjusted) returns. The authors show that this mechanism can generate substantial wealth inequality. In our approach below, we measure empirically how greater financial literacy shapes investment performance.

This framework is well-suited to help us think about the effect of financial education on financial outcomes. For instance, Lusardi et al. (2020) use this approach to generate pseudo-experimental data where some individuals receive financial education and others do not. When allowed to choose whether to participate in financial education, people naturally do so on the basis of their perceived expected gains from the program. The authors show that, in this model, the least financially savvy but with higher saving needs elect the program. Accordingly, simply comparing program participants with non-participants delivers biased estimates of program effects, leading the econometrician to over-estimate the average effect of the program on outcomes. This type of rational selection is at the root of models exploring selection bias and its impact on inequality (Heckman and Honoré, 1990).

Such rational selection can be muted or amplified by two factors. Before deciding to participate, some individuals may incorrectly perceive how the program will affect their performance. If those who would benefit most from the education underestimate the program effects (and vice-versa), selection bias may be lessened. For example, overly self-confident eligible respondents might not enroll, although they could benefit from participating in the program. Alternatively, the selection effect could be amplified if the cost of investing in knowledge, and in particular the maintenance cost of the accumulated knowledge (which can depreciate over time), is correlated with other characteristics such as cognitive skills or numeracy. In this latter instance, those with low cognitive or numeracy skills may shy away from the financial education, because their marginal cost of acquiring knowledge is high. Ultimately, the extent to which selection biases inferences is an empirical question. Nevertheless, there is little evidence in the literature on what drives participation in financial education programs and how this alters inferences about the effectiveness of such programs. Our paper tackles this issue head-on, by designing and fielding an experiment to explore selection and demonstrating how it impacts inference.

The empirical literature on the effectiveness of financial education has grown rapidly in the past two decades (Lusardi and Mitchell, 2014, 2023; Kaiser et al., 2022). While early studies used non-experimental research designs,[3] several recent studies employ randomized control trials (RCT) which allocate respondents to treatment randomly. Kaiser et al. (2022)'s meta-analysis of those RCTs concludes that there are sizeable effects of financial literacy across a range of financial decision making domains.[4] In view of the fact that compliance with assignment to treatment is imperfect, their study focuses on intent-to-treat effects, namely the effect of being assigned to treatment on outcomes, rather than average effects on the treated or average treatment effects. As a result, these and other RCTs do not reveal how those who show up at the door differ from the potential pool of eligible participants.[5]

A few studies provide a hint that selection into treatment might be non-random in the investment performance domain, but they tend to focus on financial advice rather than education. Advice differs from education, of course, as the former provides recommendations; nevertheless, one might expect that sim-

[3] See Bernheim et al. (2001) for example.

[4] We do not focus here on school-based financial literacy programs, as young people tend not to have money to invest and hence, are less engaged with financial topics; for instance see Barua et al. (2017).

[5] See Deaton and Cartwright (2018) for a similar critique of policy decisions informed solely on the basis of evidence from RCTs.

ilar mechanisms may operate. Furthermore, some of the interventions in these studies provide both advice and education. For instance, Bhattacharya et al. (2012) conduct a field experiment where respondents were offered a treatment consisting of a mix of advice and financial education. Those most likely to need advice, based on their sub-optimal past allocations, prove to be least likely to elect the advice. This type of selection will not conform with the rational selection model unless those same investors are pessimistic about the marginal benefit from receiving the treatment, or they face a high cost of receiving the treatment. In another experiment, Hung and Yoong (2013) show that older, wealthier, and less financially literate respondents were more likely to seek financial advice.[6] In our experiment, we elicit directly the willingness-to-pay to participate in a program. This allows us to understand the determinants of *ex ante* willingness-to-pay to participate in a program. By exploiting the random assignment conditional on this willingness-to-pay we are able to test directly whether those willing to pay more are also those who benefit the most.

Interestingly, those who received the advice in the Bhattacharya et al. (2012) experiment often did not follow the recommendations provided, and therefore did not substantially change their allocations. Bhattacharya et al. (2012) undertook a before-and-after performance among those who participated which they contrasted with the before-and-after comparison for those who did not. This difference-in-difference strategy estimates an average effect on the treated which

[6] In a study of federal government employees given access to a retirement planning tool, Goda et al. (2023) found that fewer than half offered the tool used it, and those who did were more financially literate and better educated. In a study of investor demand for financial advice, Calcagno and Monticone (2015) conclude that the more financially literate investors are more likely to consult advisors in their sample of Italian customers holding at least Euro 10000 at a large national bank, over half of whom had been customers for more than two decades. Those authors did not examine how the advice obtained shaped customers' financial decisions, whereas our study includes the general population and evaluates the effect of financial education on portfolio performance. In another domain, Meier and Sprenger (2013) examined the effect of providing credit counseling to low-income households seeking free tax help.

is different from the average treatment effect when selection is present. The strategy is valid when the common trend assumption holds, i.e. in the absence of the treatment, those who did not participate serve as a good counterfactual for what would have happened to those who participated. In our setting, we show that the common trend assumption does not hold, i.e those with lower willingness to pay improve more in the second task than those with higher willingness to pay but who end up not participating due to the random assignment rule. We find that the difference-in-difference estimate of the average treatment effect on the treated is much lower than the estimate using other strategies.

To estimate the effect of advice on portfolio allocations, Hung and Yoong (2013) use a design similar to ours in terms of random assignment. They offer a treatment to a randomly selected group and keep others as a control group. Those in the treatment arm is allowed to get advice voluntarily. Advice is free and therefore there is no trade-off between the expected gain from advice vs. the cost of receiving advice. Estimation of the average treatment effect on the treated is done exploiting the randomization of eligibility as an instrument for getting the advice. We follow a similar route with two improvements. First, we use two portfolio allocation tasks, one before the treatment and one after, which allows us to estimate the effect of the treatment relative to some baseline performance. We also build welfare improvement measures exploiting differences in elicited risk aversion. Second, we ask for the willingness-to-pay which allows to estimate treatment effect heterogeneity by *ex ante* willingness to pay. The key insight is that the willingness-to-pay serves as a propensity score for estimating the distribution of the treatment effect.

3 Experiment

Our experiment was fielded in the fall of 2021 using the online panel of *Asking Canadians*, a Canadian survey organization. Of the respondent pool aged 25 to 80, 2,005 subjects were randomly selected. Respondents were paid in loyalty program rewards worth the equivalent of about \$ 66,000 in incentives (ranging from 8-78 dollars per respondent).[7] Our survey instrument consists of two modules. In the first, we collect extensive information about respondents' backgrounds and preferences, while the second module focuses on the investment experiment.

3.1 Survey Module

The first module gathered information on respondents' demographics and financials (balance sheet and income). We also elicited preferences in two domains using procedures developed in the literature: risk aversion (Holt and Laury, 2002) and ambiguity aversion (Dimmock et al., 2016). To capture subjects' cognitive ability and numeracy, we employ the cognitive reflection test introduced by Frederick (2005) and the Berlin numeracy test introduced by Cokely et al. (2012). To record subjects' financial literacy, we calculate a financial literacy score based on the Big Three questions designed by Lusardi and Mitchell (2007).[8]

A total of 2,005 respondents completed our survey; we drop 12 respondents who did not disclose their gender as we use that as a control variable in all analyses. Online Appendix Table A1 reports demographics, financial information, measures of financial knowledge, and preferences for the 1,993 respondents. Respondents are age 53 on average, almost half (44.5%) are female, 65.4% are mar-

[7] Retailers included but were not limited to the Hudson's Bay department store, Petro-Canada, VIA Rail, and Aeroplan (Air Canada).

[8] The exact formulation of the questions appears in the survey instrument posted with other information on the experiment at `https://pcmichaud.github.io/Questionnaire-Gemmo-Michaud-Mitchell.pdf`.

ried, and 61.2% have children. Half the sample (51.6%) had at least a Bachelor's degree (or more). Respondents earn an average of $84,113 in annual household income (18.9% refuse to disclose this information) and they hold an average of $248,562 in financial wealth.[9]

Nearly 27% of respondents hold domestic stocks and 19.1% hold individual stocks in plans or accounts such as Canadian Registered Retirement Savings Plans (RRSP) and Tax-Free Savings Accounts (TFSA); the latter are tax-preferred investment vehicles. A minority of respondents had traded stocks or other financial instruments (36.3%), implying that stock market experience is likely low for many respondents. Only 8.3% of respondents report having high knowledge of the stock market, and 2.5% very high knowledge.

Fewer than 12.9% (5%) of respondents in our sample assess their overall financial knowledge as high (very high), and fewer than 30% of respondents had studied economics or finance in high school. In addition to self-reported measures of sophistication, we use three scores to objectively assess individuals' actual financial knowledge levels. The Financial Literacy Score refers to the total number of correct answers to the Big Three financial literacy questions (Lusardi and Mitchell, 2007, 2023). The Cognitive Ability Score is the sum of correct answers to the cognitive reflection questions by (Frederick, 2005) that measure cognitive ability; and the Numeracy Score is the sum of correct answers to the three-question Berlin numeracy test by (Cokely et al., 2012). Overall, our respondents score relatively well on financial literacy: two-thirds (66.1%) of respondents answer all three questions correctly. The average score on cognitive skills is lower (0.97 out of 3, on average), and the average numeracy score is also low, 0.55 out of 3.

We measure risk attitudes with a multiple price list (Holt and Laury, 2002), using the point at which respondents switch to the risky lottery to infer a measure

[9]This includes assets held in Registered Retirement Savings Plans, Tax-Free Savings Accounts, individual stocks, defined contribution plans, mutual funds, and other accounts.

of their risk aversion. Fewer than 6% switch at the last choice (9). If we include those who never switch, a substantial fraction, 20.4%, exhibits high risk aversion; the median switch point is close to 5. These numbers are in line with those reported in Boyer et al. (2022) who elicited risk aversion for Canadians using an incentivized experiment.[10] Finally, we measure ambiguity aversion as in Dimmock et al. (2016), defined as the difference between the matching probability reported by the respondent and 0.5, expressed in percent. Overall, a reasonable fraction of our respondents is ambiguity averse (54%).[11]

3.2 Experimental Module

The experimental module consists of three parts: an initial portfolio allocation task (Task 1), a willingness-to-pay elicitation for financial education, and a follow-up portfolio allocation task (Task 2). The willingness-to-pay elicitation is used to determine who receives the financial education intervention. Next, we provide details on each of these tasks and the assignment mechanism to the intervention.

3.2.1 Baseline Portfolio Decision (”Allocation Task 1”)

Respondents first receive a hypothetical endowment of \$ 30 to allocate across three funds, along with information about the expected 5-year returns (μ) and volatility (σ) of each fund.[12] Further, we provide explanations about each fund's average return and volatility. To express volatility, we illustrate the probabilities of different realizations of 5-year returns for each fund.

[10] While Andersen et al. (2004) have developed a more refined risk aversion measure, they conclude (p. 383) that ”the qualitative finding that participants are generally risk averse is robust” with the conventional MPL approach.

[11] For the U.S., Dimmock et al. (2016) showed that 52% of respondents were ambiguity averse.

[12] We use a simple $\mu - \sigma$ characterization of each fund since these are fund characteristics that many respondents are likely to have heard of. Since higher order moments, or correlations, are likely to be harder to grasp, the instructions are silent regarding correlations and, in fact, the funds are assumed to be uncorrelated.

To characterize expected investment performance, we make use of the efficient frontier for these three funds. We compute the set of weights which provides the highest expected return for a given level of risk (or vice-versa) (Markowitz, 1952). A respondent picking a portfolio below the frontier can improve in terms of efficiency, for any increasing concave utility function, by moving vertically (increasing return for given volatility) or horizontally (lowering volatility for given return). Letting i denote a respondent and k an investment option, each investment option is characterized by an expected return μ_k and a standard deviation of returns σ_k. Let $\pi_{1,i,k}$ be the weight put by respondent i in Task 1 on investment option k. Given the absence of correlation across investment options (by construction), the expected return and variance of the portfolio selected are given by:

$$\mu_{1,i} = \sum_k \pi_{1,i,k}\mu_k, \quad \sigma_{1,i}^2 = \sum_k \pi_{1,i,k}^2 \sigma_k - \mu_{1,i}^2 \tag{1}$$

To characterize respondents' allocations in terms of performance, we can start with the Sharpe ratio of a given portfolio in Task 1, $S_{1,i} = \mu_{1,i}/\sigma_{1,i}$. Taking σ_i as given, we denote $\{\pi_{1,k}^*\}_{k=1,2,3}$ as the weights that maximize the portfolio's expected return. These are the weights that would bring the respondent to the efficient frontier for a given level of risk (measured by standard deviation). Let $S_{\mu,i}^*$ be the Sharpe ratio for those weights. Then, the *relative mean return loss* is defined as $RML_{1,i} = 1 - \frac{S_{1,i}}{S_{\mu,i}^*}$, which measures the relative vertical distance between a portfolio allocation and the point on the efficient frontier in the mean-variance space. We can also compute the point on the efficient frontier that minimizes the standard deviation for a given expected return (the horizontal distance in mean-variance space). This yields the relative difference in risk between the efficient frontier and what the respondent selected. Let $S_{\sigma,i}^*$ be the Sharpe ratio that minimizes the standard deviation for a given level of expected return. Then the *relative standard deviation loss* is $RSL_{1,i} = 1 - \frac{S_{1,i}}{S^*_{\sigma,i}}$.[13]

Table 1 reports summary statistics on respondents' investment performance in Allocation Task 1. The average expected return is 31.7%, ranging from 18.9% to 44.4%. We also find considerable variation in $\sigma_{1,i}$, with a mean of 26.1% and a wide range, of 7.4% to 50.2%. The relative mean loss $RML_{1,i}$ averages 3.9%, again with a wide range, from 0 to 33.1%. The relative standard deviation loss, $RSL_{1,i}$, is larger, 7.63% on average, again with a large range. We also compute the fraction of respondents who put equal weights on each of the three funds, which we label 1/K allocation behavior, an often-used heuristic (Thaler et al., 2001). We find that close to one-quarter of respondents (24.4%) use such a rule. We also report the frequency of respondents who invest their entire endowments in the fund with the highest return. This behavior, which we label return chasing, characterizes one in 10 respondents (10.8%). Overall, there is considerable heterogeneity in portfolio allocations and much scope for improvement in peoples' portfolio allocations.

To account for some of the heterogeneity observed in Task 1, we run a set of regressions for different outcome variables using a vector of respondent characteristics as controls. Online Appendix Table A2 reports the results. Respondents with higher cognition scores select less risky allocations (with lower expected returns). Those with experience trading stocks are more likely to pick riskier portfolios (with higher expected returns) and are more likely to be return chasers. The better-educated are more likely to have larger relative return losses, both in

[13] We derive the relative mean return loss and relative standard deviation loss from the relative Sharpe ratio loss concepts introduced by Calvet et al. (2007). The authors define the relative Sharpe ratio loss as $RSRL_i = 1 - \frac{S_i}{S_B}$, where S_B is the Sharpe ratio of a common benchmark index. Since we have no benchmark index in our experiment, we instead use as a benchmark for each respondent the portfolio with the highest possible mean given her chosen standard deviation (for the RML), or the portfolio with the lowest possible standard deviation given her chosen mean (for the RSL).

Table 1: Performance on Allocation Task 1

	N	mean	sd	min	median	max
Mean_1		31.679	6.498	18.9	30.264	44.4
$\text{Standard Deviation}_1$		26.056	11.480	7.410	21.605	50.2
RML_1		3.883	5.861	0	1.375	33.086
RSL_1		7.628	11.473	0	3.365	59.852
$1/\text{K}_1$		0.244	0.430	0	0	1
Return Chasing_1		0.108	0.310	0	0	1
N	1993					

Note: This table presents summary statistics for respondents' performance in Allocation Task 1 based on the full sample. For continuous variables, we show the mean and standard deviation. For binary variables, we report a fraction. $1/\text{K}_1$ is equal to one if a respondent spreads her endowment equally over all three funds, and zero otherwise. Return chasing_1 is equal to one if a respondent invests her entire endowment in the fund with the highest expected return, and zero otherwise. RML is the relative mean loss, while RSL is the relative standard deviation loss.

terms of expected returns and risk which suggests better-educated respondents are also more likely to make mistakes (Calvet et al., 2007). Relative return losses are prevalent in the sample and they are not concentrated in particular groups. In terms of heuristics, those scoring higher on the financial literacy index are less likely to use the 1/K rule when picking their portfolios while those who thought they were very financially knowledgeable are more likely to be return chasers.

3.2.2 Willingness-to-Pay Elicitation

After respondents allocate their first endowment, they receive a second endowment of the same amount (30$). We then randomly assign respondents to one of two arms: a group to whom no financial education is offered (No Education Offered); and a second offered the financial education program (Offered Education), namely an information treatment on portfolio diversification and risk-adjusted returns. For those offered education, respondents are told that they can use this endowment to purchase an educational program that might help them make better financial decisions and help them improve in a second allocation task, in which they will invest the remaining amount of their second endowment.

To elicit individuals' willingness-to-pay for financial education, we use a Becker-DeGroot-Marschak (BDM) mechanism (Becker et al., 1964). Subjects are asked to state the maximum amount of their endowment they are willing to pay for the education, in an interval from zero to five dollars. A random number generator then determines the actual purchase price within this interval; if the price is below the respondent's elicited willingness-to-pay, she receives the program (at the price generated by the random number generator), otherwise she does not. Reporting a true willingness-to-pay is a dominant strategy with this mechanism (Becker et al., 1964).

We find that 24.5% of respondents offered the program elect not to receive it, even if they have to pay nothing for it. For those who do agree to pay, their average willingness to pay is \$2.91 (median of \$3), and fewer than 4.9% of respondents are willing to pay zero. Overall, there is a sizeable dispersion in elicited willingness-to-pay values. Using the assignment rule to treatment announced to respondents, we randomly generate a treated and an untreated group among those offered the intervention. A total of 43% ($N = 686$) receive the financial education intervention (treatment), while the remaining 57% ($N = 906$) do not.

3.2.3 Follow-up Portfolio Allocation ("Allocation Task 2")

All respondents face a follow-up portfolio task. For those offered the program, this follows the willingness-to-pay elicitation. For those assigned to the intervention (treated), it follows the intervention itself. At this point, subjects' endowments correspond to what they receive in Allocation Task 2 (\$30), minus the price paid for the financial education for those assigned to the treatment (received the intervention). Respondents allocate their (remaining) endowments across the same three funds using the same $\mu - \sigma$ representation as in Allocation Task 1. We recompute the mean and standard deviation of the portfolio chosen in Task 2 and denote those $\mu_{i,2}$ and $\sigma_{i,2}$ respectively.

To illustrate how portfolio allocations changed, Figure 1 compares combinations of expected returns and standard deviations of returns in both tasks achieved by respondents offered the financial education. The size of the markers is proportional to the number of respondents choosing the respective allocations. Respondents do change their allocations considerably: there is much more heterogeneity in allocations for Task 2 than in Task 1.

Figure 1: Portfolio Allocations in Tasks 1 and 2

Note: This figure illustrates all combinations of expected return and standard deviation of return achieved by respondents of the Offered Education group in Allocation Tasks 1 and 2. The size of the markers is proportional to the number of respondents who picked a particular allocation.

We consider several measures to characterize the change in portfolio allocations from Task 1 and 2. In particular, we measure changes in investment performance, in the use of heuristics, and in welfare metrics.

The change in relative loss in terms of mean returns and standard deviation are equal to ΔRML and ΔRSL, respectively. Both measure displacements inside the efficient frontier. A negative value denotes an improvement (a reduction of the loss). We also consider two metrics that capture portfolio allocation movement away from the heuristic rules, $\Delta 1/K$ and $\Delta Chasing$. For both, we compare the fraction of respondents who use the respective heuristic rule in Allocation Task 1 and move away from it in Allocation Task 2.[14]

Since the performance and heuristic metrics do not capture welfare changes from *better* to *worse* allocations, we can use the survey-based measure of risk aversion to construct a welfare metric. Using the multiple price list (MPL) elicitation task based on Holt and Laury (2002), we define s_i as the number of safe choices in the multiple price list elicitation task for risk aversion. For a particular number of safe choices, s_i, the indifference condition enables to estimate the bounds $(\gamma_{i,\min}, \gamma_{i,\max})$ of relative risk aversion for a utility function:

$$u(w, \gamma) = \frac{w^{1-\gamma}}{1-\gamma}.$$

Consider what we observe from Allocation Tasks 1 and 2. The respondent selects portfolio weights in Task 1 giving an expected return of $\mu_{i,1}$ and standard deviation $\sigma_{i,1}$; in Task 2, the respondent again chooses weights giving an expected return $\mu_{i,2}$ and standard deviation $\sigma_{i,2}$. Assuming normal returns, her expected

[14]The endowment for the second allocation task is heterogeneous for respondents who purchase the financial education, since it is net of the price paid. As that number may not be easily divisible by three, we allow for a small imbalance across funds to account for this (0.01 percent).

utility in each choice situation $k = 1, 2$ is:

$$EU_k(w_0, \gamma) = \int u(w_0(1 + \mu_{i,k} + \sigma_{i,k}\epsilon), \gamma)\phi(\epsilon)d\epsilon \tag{2}$$

where w_0 is the endowment, ϵ is a standard normal random variable, and $\phi(\epsilon)$ the density.[15]

For a given γ, a measure of the welfare change between the two tasks is the willingness to pay $\psi(\gamma)$ such that a positive (negative) $\psi(\gamma)$ measures a welfare gain (loss):

$$EU_1(w_0, \gamma) = EU_2(w_0(1 - \psi(\gamma)), \gamma). \tag{3}$$

We do not observe the precise value of γ, but the MPL gives us bounds. Let $f_i(\gamma|s_i, z_i)$ be the posterior distribution of γ for a respondent who answered s_i, with socio-economic characteristics z_i. Consider now the latent linear regression model:

$$\gamma_i^* = z_i\kappa + \nu_i.$$

where ν_i is a standard normal error term with standard deviation σ_γ. We observe s_i, providing bounds $(\gamma_{i,\min}, \gamma_{i,\max})$ in γ. We estimate κ and σ_γ by maximum likelihood.[16] We use a set of controls for age, gender, education, income, and wealth.[17] The density of $f_i(\gamma|s_i, z_i)$ is given by the truncated normal.[18]

[15]We can use quadrature to estimate the value of the integral in equation (2).

[16]The probability of observing s_i given z_i is given by

$$\Pr(s_i|z_i) = \Phi(\frac{\gamma_{i,\max} - z_i\kappa}{\sigma_\gamma}) - \Phi(\frac{\gamma_{i,\min} - z_i\kappa}{\sigma_\gamma})$$

[17]Parameter estimates are found in Appendix Table A3.

[18]The truncated normal is given by

$$f_i(\gamma|s_i, z_i) = \frac{\frac{1}{\sigma_\gamma}\phi\left(\frac{\gamma - z_i\kappa}{\sigma_\gamma}\right)}{\Phi(\frac{\gamma_{i,\max} - z_i\kappa}{\sigma_\gamma}) - \Phi(\frac{\gamma_{i,\min} - z_i\kappa}{\sigma_\gamma})}$$

We compute the probability that ψ is positive, ψ^+, using:

$$\psi_i^+ = \Pr(\psi(\gamma)|s_i, z_i) = \int I(\psi_i(\gamma) > 0) f_i(\gamma|s_i, z_i) d\gamma \tag{4}$$

A measure equal to one indicates an unambiguous improvement, while a measure equal to zero indicates no improvement. An in between value gives us a degree of certainty about the improvement. When we compute this for all respondents, we find that 64% of respondents have an unambiguous welfare change using this metric. Figure A1 in the Appendix provides a cumulative distribution function of this measure. To measure the magnitude of the welfare change, one can also compute the expected value of ψ:

$$\overline{\psi}_i = \int \psi_i(\gamma) f_i(\gamma|s_i) d\gamma.$$

Table 2 presents summary statistics of these portfolio performance measures comparing Allocation Tasks 2 versus 1. Of the 486 respondents who previously spread their endowment equally across all three funds in Allocation Task 1, 46.1% changed their allocations in the second task. Of the 215 respondents who put their entire endowments in the fund with the highest expected return in Task 1, 43.3% adjusted this behavior in Task 2. Both of our portfolio efficiency measures in Allocation Task 2, the vertical distance to the efficiency frontier (RML) and the horizontal distance to the efficiency frontier (RSL), are smaller, on average, relative to the average values achieved in Task 1. Finally, Table 2 shows that 47% of the respondents had an increase in welfare, though the average welfare gain was small, 0.2%. There is substantial heterogeneity in the welfare change.

for $\gamma \in (\gamma_{i,\min}, \gamma_{i,\max})$ and zero elsewhere.

Table 2: Performance on Allocation Task 2

	count	mean	sd	min	p50	max
Δ RML	1993	-0.335	5.816	-33.086	0	32.911
Δ RSL	1993	-0.478	11.827	-59.852	0	59.467
Δ 1/K	486	0.461	0.499	0	0	1
Δ Chasing	215	0.433	0.497	0	0	1
ψ^+	1993	0.470	0.411	0	0.420	1
$\overline{\psi}$	1993	0.002	0.103	-0.839	0	0.715
N	1993					

Note: The performance improvement measures are defined as follows: Δ RML = RML_2 - RML_1; Δ RSL = RSL_2 - RSL_1; Δ 1/K = 1-1/K_2 if 1/K_1==1; Δ Chasing = 1-Return chasing_2 if Return chasing_1==1; and ψ^+ is the probability of a welfare gain, given the uncertainty about risk aversion γ while $\overline{\psi}$ is the expected welfare change.

3.3 The Educational Program

The financial education targets two important concepts related to portfolio choice: diversification and risk-adjusted returns. The intervention consists of several screens displayed to respondents, where the first defines portfolio allocation, and subsequent screens discuss the value of diversification and risk-adjusted returns. To that end, we first illustrate a hypothetical investment opportunity consisting of three different funds with the same expected return and standard deviation (referred to as variability). The education illustrates verbally and graphically that a portfolio's standard deviation decreases when an endowment is spread equally across the three funds, relative to investing everything in a single fund, while the expected return is unchanged. We then relate this decrease in variability to the term diversification.

In the next step, the education focuses on the concept of risk-adjusted returns. To that end, we present a second hypothetical investment opportunity consisting of three funds with different expected returns and standard deviations. The instructions suggest that subjects first build a portfolio by spreading the endowment equally across the three funds, and then they discuss how subjects can increase the portfolio's expected return while keeping the standard deviation unchanged. To achieve this, we suggest that subjects calculate risk-adjusted returns of each fund by dividing its expected return by its standard deviation, and then allocating more money to funds with higher risk-adjusted returns.

4 Effect of the Intervention

We seek to estimate the effect of the financial education on outcomes. Let y_i be some measure of change in portfolio allocation from Task 1 to 2. The experimental design is such that we have three groups. First, we randomized offering the intervention. Let q_i be equal to one if a respondent is offered the education, and zero if not. Second, we allowed those who were offered the intervention to receive it (being treated). Respondents can influence the likelihood of participation by stating a willingness to pay. Let $d_i = 1$ if the respondent is provided the education, and zero if not. Hence, we have three groups: not offered (and not treated), offered and treated, and finally offered and untreated.

In terms of potential outcomes depending on treatment status, $y_{i,0}$ is the outcome if untreated and $y_{i,1}$ if treated. Only one of these outcomes is observed for each respondent, $y_i = d_i y_{i,1} + (1 - d_i) y_{i,0}$. For those not offered the treatment, $y_i = y_{i,0}$. Hence, the treatment effect $\Delta_i = y_{i,1} - y_{i,0}$ for each respondent is not observed. Features of the distribution of Δ_i can, however, be uncovered.

A first building block is to define the effect of offering the program (intent-to-treat) by

$$ITT = E[y_i | q_i = 1] - E[y_i | q_i = 0].$$

Given the randomization of q_i, this can be estimated directly from comparing the distributions of y_i among those offered the program and the control group and obtain $\widehat{ITT} = E_N[y_i|q_i = 1] - E_N[y_i|q_i = 0]$ where E_N denotes the sample analogue of the conditional expectation E. Given the randomization, this effect does not suffer from selection. Kaiser et al. (2022) survey intent-to-treat estimates of the effect of financial education.

Second, we can define the effect of the intervention on those who were treated,

$$ATT = E[\Delta_i|d_i = 1] = E[y_{i,1} - y_{i,0}|d_i = 1].$$

This parameter is of interest if one seeks to evaluate the effect that the intervention had on those who received it. A last parameter of interest is the average treatment effect:

$$ATE = E[\Delta_i] = E[y_{i,1} - y_{i,0}].$$

This parameter is useful when thinking of the expected effect of taking a random individual in the population and giving her the intervention.

Since participating, d_i, is voluntary, the estimate of the ATE by comparison of those treated and untreated may be biased. This would occur if for example treatment was based on $d_i = I(y_{i,1} - y_{i,0} > 0)$. In this case d_i is not independent of Δ_i.

Imbens and Angrist (1994) show that we can exploit randomization of the intent-to-treat, offering the program. We can estimate the local average treatment effect with a Wald IV-estimator:

$$\widehat{LATE} = \frac{E_N[y_i|q_i = 1] - E_N[y_i|q_i = 0]}{E_N[d_i|q_i = 1] - E_N[d_i|q_i = 0]}.$$

This is the effect of the treatment for those who are induced to participate by offering the program. Since $E[d_i|q_i = 0] = 0$ and $E[d_i|q_i = 1] = \Pr(d_i = 1|q_i = 1)$ with sample analogue $P_N(d_i = 1|q_i = 1)$, the IV estimator becomes:

$$\widehat{LATE} = \frac{E_N[y_i|q_i = 1] - E_N[y_i|q_i = 0]}{P_N(d_i = 1|q_i = 1)}$$

which is the ITT divided by the probability of being treated conditional on being offered the treatment. The IV estimator can be implemented by two-stage least squares controlling for a set of controls X_i. This estimate coincides with the average effect of the treatment on the treated because part of the sample is (randomly) excluded from being able to participate in the treatment. The expected outcome among those offered the treatment is:

$$E[y_i|q_i = 1] = E[y_{i,0}|q_i = 1] + \Pr(d_i = 1|q_i = 1)E[y_{i,1} - y_{i,0}|q_i = 1, d_i = 1]$$

while among those not offered treatment we have: $E[y_i|q_i = 0] = E[y_{i,0}|q_i = 0]$. Given random assignment of q_i, $E[y_{i,0}|q_i = 1] = E[y_{i,0}|q_i = 0] = E[y_i|q_i = 0]$. It follows that:

$$E[y_{i,1} - y_{i,0}|q_i = 1, d_i = 1] = \frac{E[y_i|q_i = 1] - E[y_i|q_i = 0]}{\Pr(d_i = 1|q_i = 1)}.$$

The LATE estimator is also the ATT estimator in this particular case. Denote this estimator $\widehat{ATT}_{IV}$.

While selection makes estimating of ATE difficult by comparing treated and untreated respondents, one can potentially recover the ATT. The difference-in-differences estimator can be used, as it compares the change in outcomes across treated and untreated groups. Since our outcomes are already estimated in changes (from Task 1 to 2), the difference-in difference-estimator is given by:

$$\widehat{ATT}_{DD} = E_N[y_i|d_i = 1] - E_N[y_i|d_i = 0]$$

This estimator can be implemented by OLS and augmented with controls X_i. The key assumption for this estimator to deliver the ATT is that the counterfactual change in outcomes for those treated is well captured by the change in outcomes for those untreated. Any violation of the parallel trend assumption will bias the ATT estimate arising from the difference-in-differences estimator.

Table 3 provides estimates of the effect of the intervention using different estimators.

Table 3: Effect of the Financial Education Intervention on Outcomes

	Performance metrics		Heuristic metrics		Welfare metrics	
	Δ RML	Δ RSL	Δ 1/K	Δ Chasing	ψ^+	$\overline{\psi}$
change not offered	-0.036	-0.321	0.253	0.3	0.406	-0.008
change offered	-0.410	-0.517	0.489	0.473	0.486	0.005
Diff.-in-Diff. ITT	-0.374	-0.196	0.236	0.173	0.080	0.014
	(0.325)	(0.661)	(0.057)	(0.079)	(0.023)	(0.006)
change not treated	-0.155	-0.335	0.277	0.307	0.439	0.002
change treated	-0.747	-0.758	0.931	0.734	0.548	0.009
Diff.-in-Diff. ATT	-0.592	-0.422	0.659	0.427	0.109	0.007
	(0.302)	(0.601)	(0.042)	(0.073)	(0.021)	(0.005)
IV ATT	-0.868	-0.456	0.717	0.445	0.185	0.032
	(0.753)	(1.533)	(0.142)	(0.189)	(0.053)	(0.013)
IV ATT (w. controls)	-0.996	-0.550	0.662	0.466	0.204	0.034
	(0.766)	(1.561)	(0.159)	(0.161)	(0.053)	(0.014)

Note: This table reports estimates of the effect of the financial education on several outcomes. We consider Δ RML, the change in the relative mean loss; Δ RSL, the change in the relative standard deviation loss; Δ 1/K, or the fraction moving away from 1/K allocations in Task 2, Δ Chasing or the fraction moving away from chasing the highest return in allocation Task 2, and finally the two welfare metrics, ψ^+ and $\overline{\psi}$, which are respectively the probability of a welfare gain and the expected welfare change. The first three lines report the average change in outcomes for those not offered ($q_i = 0$) and offered ($q_i = 1$) and the resulting difference-in-difference ITT, estimated by regression. The same comparisons are done for those treated and untreated resulting in a difference-in-difference ATT estimate. The last two groups of estimates estimate the ATT by IV using q_i as as an instrument for d_i with and without controls X_i. Standard errors are reported in parentheses.

In terms of intent-to-treat, offering the intervention increases the proportion who gain in terms of welfare by 8 percentage points and the average welfare by

1.4 percentage point. Offering the intervention leads to reductions in relative loss (both mean and variance) but these estimates are imprecisely estimated. Those offered are more likely to move away from 1/K allocations and return chasing allocations. Among those offered the treatment, a simple comparison of the change in outcomes at the mean between those treated and untreated is a difference-in-difference ATT estimator. Hence, the proportion with a welfare gain is larger. However, we find an imprecisely estimated average welfare gain. The same contrast also reveals larger reductions in relative losses among those treated and a reduction in the proportion selecting 1/K and return chasing allocations. IV ATT estimates indicate that the increase in the proportion with a welfare gain is nearly double (18.5 p.p. compared to 10.9 pp.) among those treated, and a 5 times larger average gain (3.2 p.p. compared to 0.7 p.p.). The pattern is the same for other performance measures. Adding controls has little effect on these estimates. The difference between ATT estimates coming from the IV and difference-in-difference strategies suggests that there is heterogeneity in the effect of the treated - that the common trend assumption does not hold. In this second case, the ATT is biased using a difference-difference strategy. We investigate next the nature of selection using the unique setup of the experiment.

5 Selection and Heterogeneity

Our experiment gives a natural measure of a respondent's intention to receive financial education, which allows us to study selection into the treatment. The probability of getting the treatment is a direct function of the elicited willingness-to-pay, w_i. The probability of being assigned to the intervention is $\Pr(d_i = 1|w_i) = w_i/w_{\max}$, where $w_{\max}$ is the maximum price that could be paid for the education ($ 5). Effectively, respondents can opt out of the education by refusing it even if provided for free, and they can opt in with probability one if they report the maximum willingness-to-pay. The key innovation of our setting is that, given w_i, the treatment is randomly delivered. This provides a laboratory to study the relationship between the perceived gain from receiving the intervention and realized gain.

Who is willing to pay more? Close to half, 46.2% of respondents offered the program, indicate that they expect to be able to apply the information received, while 19.3% report that they do not know whether they will. Almost half (46.7%) of respondents believe that their return in Allocation Task 2 would be higher than in Task 1 if they receive the program, while 26.9% report that they do not know whether they would do better.

To understand what factors influence respondents' willingness to pay for the financial education program, we next assess who rejects the education even if available at no cost (extensive margin), as well as the factors shaping how much respondents are willing to pay for the intervention (intensive margin). The marginal effects on the extensive margin (from a Logit regression) appear in Column 1 of Table 4; estimated coefficients on the intensive margin appear in Column 2. Column 3 combines both margins with a dependent variable equal to 0 if the respondent rejects the program, and equal to the willingness to pay if the respondent provides one.

Respondents must trade off their willingness to pay against their expected benefit from the educational program. Accordingly, it is reasonable to find that respondents rationally base their demand for financial education on their expectation about whether they expect to be able to apply the program information conveyed, and whether they expect that the knowledge will boost their return in Allocation Task 2. Respondents who expect to be able to apply the financial education are 6.7 percentage points less likely to refuse the program, and they are

Table 4: Regression Estimates of Factors Associated with Willingness to Pay for Financial Education

	(1) Reject program		(2) Willingness to pay (>= 0)		(3) Willingness to pay	
Apply information: yes	-0.067	(0.025)	0.399	(0.111)	0.508	(0.110)
Apply information: dk	0.045	(0.026)	0.091	(0.149)	-0.083	(0.133)
Exp. higher return: yes	-0.129	(0.026)	0.338	(0.119)	0.588	(0.117)
Exp. higher return: dk	0.017	(0.025)	0.057	(0.142)	-0.072	(0.129)
Female	-0.029	(0.020)	0.073	(0.094)	0.139	(0.091)
College or some university	0.042	(0.030)	-0.111	(0.142)	-0.188	(0.135)
Bachelor degree or more	0.058	(0.030)	-0.221	(0.141)	-0.307	(0.135)
ln(Household income)	0.020	(0.006)	-0.017	(0.021)	-0.060	(0.022)
Household income missing	0.128	(0.022)	-0.154	(0.140)	-0.573	(0.117)
Financial wealth	-0.005	(0.003)	0.012	(0.010)	0.019	(0.010)
Financial Literacy Score	-0.045	(0.013)	-0.029	(0.073)	0.150	(0.063)
Cognitive Ability Score	0.017	(0.011)	0.046	(0.051)	0.012	(0.050)
Numeracy Score	-0.055	(0.015)	-0.066	(0.058)	0.061	(0.059)
Financial knowledge: high	0.010	(0.037)	-0.498	(0.154)	-0.446	(0.153)
Financial knowledge: very high	0.039	(0.051)	-0.517	(0.247)	-0.483	(0.236)
St. market knowledge: high	-0.002	(0.046)	-0.109	(0.194)	-0.107	(0.192)
St. market knowledge: very high	0.054	(0.071)	-0.141	(0.354)	-0.250	(0.336)
Has traded stocks	-0.047	(0.024)	0.008	(0.100)	0.141	(0.100)
Has studied economics	0.016	(0.022)	0.149	(0.099)	0.090	(0.096)
Mean_1	-0.057	(0.036)	0.022	(0.155)	0.104	(0.151)
$\text{Standard Deviation}_1$	0.031	(0.020)	-0.008	(0.086)	-0.052	(0.084)
RML_1	-0.034	(0.020)	-0.020	(0.084)	0.038	(0.082)
RSL_1	0.006	(0.004)	0.007	(0.017)	-0.006	(0.016)
$1/\text{K}_1$	0.092	(0.022)	-0.144	(0.116)	-0.401	(0.107)
Return Chasing_1	0.074	(0.049)	-0.171	(0.223)	-0.405	(0.216)
_cons			2.314	(2.927)	-0.735	(2.846)
Mean	0.245		2.909		2.196	
N	1592		1202		1592	
chi2	426.706					
r2			0.080		0.200	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table is based on the "Offered Education" sample, i.e. respondents with $q_i = 1$. Reject program is an indicator variable equal to one if a respondent indicates that she does not want the financial education. Willingness to pay takes the value of 0 if she indicates that she does not want the education; otherwise it takes the value the respondent is willing to pay for it. Willingness to pay (>= 0) indicates the respondent's stated willingness to pay for the education if she elects to receive it. Column 1 reports marginal effects from a Logit regression. Columns 2-3 report OLS coefficient estimates. All columns also control for the respondent's region of residence, ownership of individual stocks, ownership of domestic stocks, marital status, children, number of HH members, age, and preferences such as risk aversion, ambiguity aversion, and patience.

willing to pay more for it, compared to respondents who do not think that they can apply it. Analogously, respondents who expect to obtain a higher return in Task 2 if they receive the education are 12.9 percentage points less likely to refuse the treatment and are also willing to pay more for it, than their counterparts.[19] Interestingly, respondents with high levels of revealed sophistication - measured by their financial literacy and numeracy scores - are less likely to reject the educational program, while self-reported financial knowledge is negatively related to respondent willingness to pay for the intervention.[20] Also, higher household income and greater formal education are associated with a lower willingness to receive the program, while respondents experienced in trading stocks are 4.7 percentage points less likely to reject the education.

Respondents who spread their endowments equally across all funds in the first allocation task are 8.3 percentage points more likely to reject the program, and they are also less willing to pay for the intervention, compared to others. Beyond this, we find no evidence that respondents' performance in Allocation Task 1 affected their willingness to receive the financial education that might improve their performance in Allocation Task 2.[21] The Becker et al. (1964) market revelation mechanism allows us to elicit respondents' willingness to pay for the education on the condition that respondents understood the mechanism sufficiently well for

[19]A higher probability of rejecting the education, even if offered at no cost, likely reflects a respondent's opportunity cost of time associated with receiving the program. Kim et al. (2016) show, in a theoretical setting, that acquiring financial knowledge can be sub-optimal for certain individuals, given opportunity costs of time.

[20]Note that these results hold even though we control for respondents' expectations about their ability to apply the program information conveyed. These findings are in line with the results from a survey study by Anderson and Robinson (2017). The authors report a negative relationship between participants' overestimation of their own financial literacy (the difference between participants' perceived financial literacy and participants' revealed financial literacy) with the willingness to accept free financial advice in the context of the social media platform LinkedIn.

[21]Since some of the performance measures in the first allocation task are correlated, we also run our analysis on each individual performance measure separately, without including the others. The results are qualitatively similar to results presented in Table 4.

it to work.[22] To this end, during the experiment, respondents are provided with information on the mechanism and provided examples. Furthermore we include a control question to test whether respondents understand the process, and over half responds correctly.[23]

As noted above, one important driver of respondents' willingness to pay was their anticipated benefits from the intervention. That is, respondents are willing to pay more for the intervention if they expect to be able to apply the new information to Allocation Task 2, and if they expect the return from the second task to be higher than that earned on Task 1. While we do not observe how individuals form these expectations, we can explore the association between respondent expectations about the benefits of the financial education with their measured socio-demographic characteristics, cognitive ability, and performance in Allocation Task 1. Results in Online Appendix Table A4 show that women are less confident than men in their ability to apply the new information, and they frequently respond that they do not know if they have that ability. Financial literacy is positively related to the perceived ability to apply the program information and to earn a higher return in Allocation Task 2, conditional on receiving it.

Interestingly, high self-reported financial knowledge is positively and significantly associated with peoples' anticipated ability to apply the education. Respondents with past stock trading experience are 7.4 percentage points more confident that they can transform the information acquired into a higher return

[22]Note that even if respondents did not understand the mechanism, they may still have stated their true willingness to pay based on intuition, but we cannot test if this was the case.

[23]The control question first shows respondents a hypothetical stated willingness to pay for the program as well as a hypothetical price. Next they are asked to state whether or not they would receive the program in this case, and if so, what price they would have to pay. As a robustness check, we split the analysis sample into a group which responds correctly to the control question (54.15%), and another which does not (45.85%). We repeat our analyses on the determinants of willingness to pay on these two sub-samples, and though we lose power when we split the group, our results still hold qualitatively for both sub-samples.

in Allocation Task 2 (relative to the returns earned in Allocation Task 1). Respondents who studied economics and finance in high school are, respectively, 7.5 and 6 percentage points more likely to believe that they can apply the information and that it would lead to a higher return. Finally, performance in Allocation Task 1 is associated with people's beliefs about whether the program would help them achieve a higher return in Allocation Task 2. Respondents with a higher standard deviation in Task 1 are more likely to respond "don't know" to this question.[24]

Let $p_i = p(w_i)$ be the probability of receiving the intervention given that it is offered. It depends on w_i in a deterministic way and is known, $p_i = w_i/5$. Given that treatment is assigned randomly, conditional on w_i, or $p_i(w_i)$, we can make the missing at random assumption, $y_{i,0}, y_{i,1} \perp d_i|p(w_i)$. This assumption, akin to allowing for selection on observables, states that the treatment is randomly assigned conditional on the propensity score p_i. Using this assumption, we can implement an estimator of the ATT using a matching estimator. Let ω_{ij} be weight attached to a respondent j in the untreated group for a comparison with respondent i who was treated. For example, one could consider the K nearest-neighbor (KNN) estimator to construct weights based on the propensity score p_i. A matching estimator of the ATT is given by:

$$\widehat{ATT}_{KNN} = \frac{1}{N_T} \sum_{i:d_i=1} \left(y_{i,1} - \sum_{j:d_j=0} \omega_{ij} y_{j,0} \right)$$

where N_T is the number of treated respondents. An alternative is to use a

[24]Since some of the performance measures in the Allocation Task 1 were correlated, we also re-ran our analyses on each individual performance measure separately, excluding the others. These results do not differ qualitatively from those appearing in Appendix Table A4.

control function approach by estimating a conditional expectation:

$$E[y_i|d_i, p_i, X_i] = \beta_0 + \beta_1 d_i + m(p_i) + X_i\beta_X$$

where $m(p_i)$ is some possibly non-linear function of p_i. The estimate $\hat{\beta}_1$ is an ATT estimate, assuming a constant treatment effect.

The balancing condition for propensity matching states that $d_i \perp X_i|p_i$. We can verify this, given that we know p_i, by regressing d_i on p_i (or w_i) and X_i and testing that coefficients on X_i are all zero. We do not reject this condition with p-value of 0.55. The required overlap condition for estimating an ATT is that $\Pr(d_i = 1|w_i) < 1$. Among those offered the treatment, 13% report a willingness-to-pay of 5, and we drop those observations.

We estimate the ATT using different estimators exploiting the missing-at-random assumption. First, we consider 5 and 10 nearest-neighbors matching estimators using propensity scores. We also estimate the ATT by OLS with a linear control for w_i, adding X_i as controls and finally we estimate a non-linear specification in w_i (cubic) with controls X_i. Table 5 reports the estimates.

One interesting result is that matching estimates ATT for measures of welfare are in line with those obtained by IV. Using 5 nearest neighbors, those who participate in the education have a 27 percentage point higher probability of a welfare gain, and a 3.7% higher average welfare gain. The effects are similar with 10 neighbors. The estimates with a control function approach are smaller than matching or IV estimates, but larger than difference-in-difference estimates reported earlier. For 1/K and return chasing behavior, we find ATT metrics which are largely in line with IV ATT estimates. Finally, we find mainly negative effects on relative mean and standard deviation loss, although most estimates are not statistically significant.

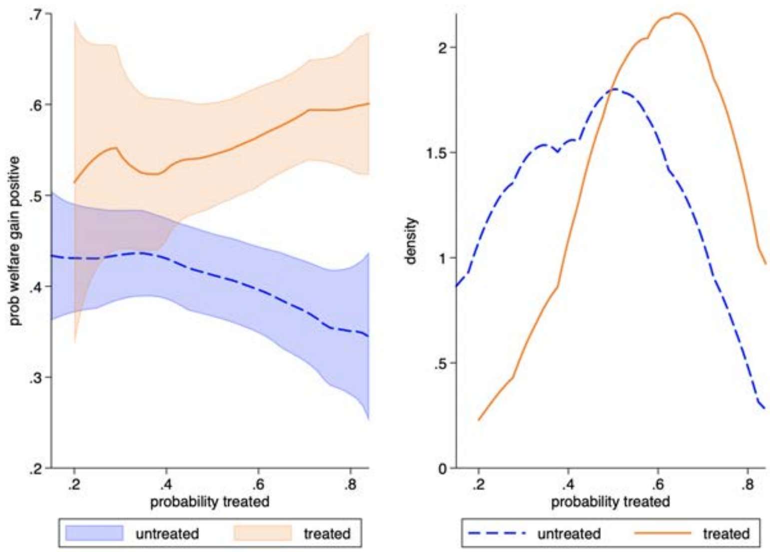
Table 5: Matching and Control Function Estimates of the Average Effect of the Intervention on those who Participated

	Performance metrics		Heuristic metrics		Welfare metrics	
	Δ RML	Δ RSL	Δ 1/K	Δ Chasing	ψ^+	$\overline{\psi}$
KNN matching (5)	-0.216	0.210	0.522	0.257	0.272	0.037
	(1.061)	(2.119)	(0.102)	(0.159)	(0.067)	(0.014)
KNN matching (10)	-0.663	-0.959	0.511	0.317	0.244	0.024
	(0.717)	(1.476)	(0.079)	(0.141)	(0.049)	(0.010)
Control Function	-0.818	-1.076	0.558	0.293	0.155	0.014
	(0.451)	(0.899)	(0.063)	(0.128)	(0.031)	(0.008)
Control Fct. + X	-0.842	-1.101	0.572	0.309	0.153	0.013
	(0.456)	(0.910)	(0.067)	(0.145)	(0.032)	(0.008)
NL Control Fct.+X	-0.840	-1.083	0.574	0.301	0.153	0.013
	(0.457)	(0.912)	(0.067)	(0.145)	(0.032)	(0.008)

Notes: This table reports estimates of the effect of the intervention on a set of outcomes using (weighted) comparisons of treated and untreated respondents (in the "Offered Education" sample, N = 1592): Δ RML, the change in the relative mean loss; Δ RSL, the change in the relative standard deviation loss; Δ 1/K, the fraction moving away from 1/K allocations in Task 2 ; Δ Chasing, the fraction moving away from chasing the highest return in allocation Task 2; and the two welfare metrics, ψ^+, the probability of a welfare gain, and $\overline{\psi}$, the expected welfare change. The first two rows report ATT estimates using propensity score matching with 5 and 10 nearest neighbors. The next set of estimates (rows 3 to 5) report estimates using a control function specification. In row 3, only a linear control for p_i (the propensity score) is included. In row 4, controls X_i are included. Finally in row 5, we add a cubic polynomial in p. Standard errors are reported in parentheses.

To understand the differences across ATT estimates, we also investigate how the treatment effect varies along the distribution of p_i. To this end, we estimate a kernel local polynomial regression of ψ^+ on p_i separately for those treated and untreated. Since there is little common support at the extremes of the distribution, we focus on p above 0.15 and below 0.85. Figure 2 reports the estimates along with 95% confidence intervals. We also plot the (kernel) density of the propensity score for those treated and untreated but offered the program.

Figure 2: Welfare Gains by the Probability of being Treated among the Treated and Untreated Respondents



Note: The figure on the left panel shows a non-parametric local polynomial regression of the probability of a welfare gain from Allocation Tasks 1 to 2 on the probability of receiving the financial education (willingness to pay divided by 5) for both the treated and untreated respondents offered the treatment. 95% confidence intervals are reported. The right panel displays the kernel density estimate of the probability of being treated for both those treated and not treated.

The distance between the two non-parametric regression lines (in the left panel of Figure 2) provides a measure of the treatment effect along the distribution of the propensity score. Clearly, the largest effects are observed for those with a high willingness to pay (propensity score); these individuals are also the most likely to participate in the education, given the selection mechanism. This is consistent with finding substantial ATT estimates, as those who are treated are disproportionately those with higher willingness to pay. It also suggests a positive correlation between willingness to pay and actual benefits from the intervention, consistent with a selection model where participants choose to participate based on (their expectation of) the treatment effect Δ_i (as in Lusardi and Mitchell

(2023)).

The difference-in-difference estimator yielded lower effects of the financial education on ψ^+. To understand why, one must assess the validity of the common trend assumption. From Figure 2, one can see that the group most likely to be treated has a counterfactual $y_{i,0}$, which is lower than that of the group less likely to be treated. Accordingly, the common trend assumption clearly does not hold. When one uses the untreated as a control group, the estimate of the counterfactual outcome for those treated is higher than the actual counterfactual outcome for that group. Hence, the difference-in-difference estimate is lower than the IV estimate.

In sum, those who participate in the financial education benefit from the intervention in terms of welfare, moving away from using the simple investment heuristics. Those who have the most to gain do participate in the intervention. They are more able to apply what they learn, are more financially literate, and expect to earn the most in Allocation Task 2 after receiving the education. Those who express a low willingness to pay but receive the education anyway, due to the randomized nature of the treatment, also experience improvements in outcomes, although gains are more modest. The average treatment effect, or the effect of giving a random person the offered education, is significantly smaller than the effect for those who are more interested in participating in the financial education program.

6 Conclusions

Evaluating the impact of financial education faces two important challenges. First, individuals who participate in voluntary programs such as those offered by employers and governments are likely different from the general target population. Hence, the effect of the education on those electing to participate may differ from the likely effect among non-participants. This matters for evaluating programs, and also for thinking of the potential benefits of boosting participation via recruiting efforts. Accordingly, one key contribution of this paper is to devise a novel experimental setting where we elicit individuals' willingness to pay to participate in financial education, and we quantify the effect of the program on investment behavior while accounting for selection using the potential outcomes framework proposed by Imbens and Angrist (1994). While the setting is experimental, the educational content on portfolio diversification and risk-adjusted portfolio returns is similar to what is used to teach portfolio allocation, and we use real financial stakes so the results are as relevant as they can be for real world interventions, given financial constraints.

A second challenge to evaluating the effect of financial education is defining outcomes of interest. Ideally, we seek to measure the impact on participant welfare, but this is not always easy to do. Therefore, another contribution of our paper is to develop preference-based outcomes tailored to individuals' risk aversion, so we can estimate the effect of the education on respondent welfare, accounting for imperfect elicitation of risk aversion. This allows us to test whether those willing to pay more are also those who benefit more from the intervention, a prediction that arises naturally from models with endogenous accumulation of financial knowledge (Lusardi et al., 2020). We also measure how the educational intervention influences the use of heuristics in portfolio allocation.

We find that the financial education leads to more heterogeneity in portfolio allocations and more customized portfolios. Using the proposed welfare metrics, those receiving the education do better: there is a 20 p.p. rise in the fraction with a welfare improvement, amounting to 3.2 p.p. of wealth, on average. Importantly for understanding the effect of the treatment on those receiving it, we show that those more willing to participate are more financially literate, more likely to expect to apply the knowledge received, and anticipate higher returns from the intervention. In line with these expectations, we document that participants do, indeed, benefit more from the education, and the effect is positively correlated with willingness to pay. Hence, we find substantially higher average effects for those receiving the education, compared to the average effect across the population offered it. This is consistent with a selection model where participants self-select based on their expected gain from the treatment (Lusardi et al., 2020).

The implications of our results are threefold. First, we conclude that those more likely to participate in the financial education do gain more from the treatment. Since the education provided is simple and inexpensive, yet it helps participants substantially, our results suggest that a cost-benefit analysis of a financial education program akin to this one would yield substantial benefits to those participating. Currently, many employees covered by defined contribution retirement plans must select between investment options, so interventions similar to ours could substantially improve outcomes for those participating. Extending participation to cover more employees would also yield better-constructed portfolios, but the benefits are likely to be lower for the group least interested in participating.

Second, we confirm that the education provided does not appear to select negatively on potential benefits, thus missing the target group who benefits the most. Prior research in the investment domain yields conflicting results on the nature of selection, yet most of those studies focused on financial advice rather than education and used non-representative respondent populations (Bhattacharya et al., 2012; Goda et al., 2023; Hung and Yoong, 2013; Meier and Sprenger, 2013). This is important, as targeting the wrong participants might erroneously suggest that financial education is ineffective. Additionally, while a more complex and lengthy program might yield different conclusions, our unique setting featuring a broader set of respondents, along with randomization of the intent-to-treat and treatment assignment, confirms strong evidence of effects in the investment domain which grow with peoples' willingness to pay for the financial education. Those receiving the program are more likely to be confident they can apply the knowledge received and perceive higher returns from the intervention. Crucial to these findings is the measurement of outcomes of interest.

Third, our findings are relevant to the optimal amount of financial education provided. There is substantial willingness-to-pay for financial education and potential welfare benefits may outweigh the cost of providing such education. While we do not assess the price sensitivity of demand for financial education, our framework can inform optimal financial education subsidies in the workplace, as well as public provision of financial education. Going further, a similar experiment could be conducted to infer the value of financial advice, perhaps even jointly with education, so as to determine whether these two are substitutes or complements.

References

Agarwal, S., G. Amromin, I. Ben-David, S. Chomsisengphet, and D. D. Evanoff (2011): "Financial Counseling, Financial Literacy, and Household Decision Making," in *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*, ed. by O. S. Mitchell and A. Lusardi, Oxford University Press, New York, 181–205.

Ambuehl, S., B. D. Bernheim, and A. Lusardi (2022): "Evaluating Deliberative Competence: A Simple Method with an Application to Financial Choice," *American Economic Review*, 112, 3584–3626.

Andersen, S., G. Harrison, M. Lau, and E. Rutstrom (2004): "Elicitation using Multiple Price List Formats," *Experimental Economics*, 9, 383–405.

Anderson, A. Baker, F. and D. Robinson (2017): "Precautionary savings, retirement planning and misperceptions of financial literacy," *Journal of Financial Economics*, 126, 383–398.

Barua, R., B. Koh, and O. Mitchell (2017): "Does Financial Education Enhance Financial Preparedness? Evidence from a Field Experiment in Singapore," *Journal of Pension Economics and Finance*, 17, 254–277.

Becker, G. M., M. H. DeGroot, and J. Marschak (1964): "Measuring Utility by a Single-Response Sequential Method," *Behavioral Science*, 9, 226–232.

Bernheim, B. D., D. M. Garrett, and D. M. Maki (2001): "Education and Saving: The Long-term Effects of High School Financial Curriculum Mandates," *Journal of Public Economics*, 80, 435–465.

Bhattacharya, U., A. Hackethal, S. Kaesler, B. Loos, and S. Meyer (2012): "Is Unbiased Financial Advice to Retail Investors Sufficient? Answers from a Large Field Study," *The Review of Financial Studies*, 25, 975–1032.

Boyer, M. M., P. d'Astous, and P.-C. Michaud (2022): "Tax-Preferred Savings Vehicles: Can Financial Education Improve Asset Location Decisions?" *The Review of Economics and Statistics*, 104, 541–556.

Calcagno, R. and C. Monticone (2015): "Financial literacy and the demand for financial advice," *Journal of Banking & Finance*, 50, 363–380.

Calvet, L. E., J. Y. Campbell, and P. Sodini (2007): "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 115, 707–747.

Choi, J. J., D. Laibson, and B. C. Madrian (2010): "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds," *The Review of Financial Studies*, 23, 1405–1432.

Cokely, E. T., M. Galesic, E. Schulz, S. Ghazal, and R. Garcia-Retamero (2012): "Measuring Risk Literacy: The Berlin Numeracy Test," *Judgement and Decision Making*, 7, 25–47.

Collins, J. M. (2013): "The Impacts of Mandatory Financial Education: Evidence from a Randomized Field Study," *Journal of Economic Behavior and Organization*, 95, 146–158.

Deaton, A. and N. Cartwright (2018): "Understanding and misunderstanding randomized controlled trials," *Social Science Medicine*, 210, 2–21, randomized Controlled Trials and Evidence-based Policy: A Multidisciplinary Dialogue.

Dimmock, S. G., R. Kouwenberg, O. S. Mitchell, and K. Peijnenburg (2016): "Ambiguity Aversion and Household Portfolio Choice Puzzles: Empirical Evidence," *Journal of Financial Economics*, 119, 559–577.

Frederick, S. (2005): "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19, 25–42.

Goda, S., M. Levy, C. Manchester, A. Sojourner, J. Tasoff, and J. Xiao (2023): "Are Retirement Planning Tools Substitutes or Complements to Financial Capability?" *Journal of Economic Behavior Organization*, 214, 561–573.

Hackethal, A. and R. Inderst (2013): "How to Make the Market for Financial Advice Work," in *The Market for Retirement Financial Advice*, ed. by O. S. Mitchell and K. Smetters, Oxford: Oxford University Press, chap. 10, 213–228.

Heckman, J. J. and B. E. Honoré (1990): "The Empirical Content of the Roy Model," *Econometrica*, 58, 1121–1149.

Holt, C. A. and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92, 1644–1655.

Hung, A. A. and J. K. Yoong (2013): "Asking for Help - Survey and Experimental Evidence on Financial Advice and Behavior Change," in *The Market for Retirement Financial Advice*, ed. by O. S. Mitchell and K. Smetters, Oxford: Oxford University Press, chap. 9, 128–212.

Imbens, G. and J. Angrist (1994): "Identifcation and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475.

Jappelli, T. and M. Padula (2013): "Investment in Financial Literacy and Saving Decisions," *Journal of Banking & Finance*, 37, 2779–2792.

Kaiser, T., A. Lusardi, L. Menkhoff, and C. Urban (2022): "Financial Education Affects Financial Knowledge and Downstream Behaviors," *Journal of Financial Economics*, 145, 255–272.

Kim, H. H., R. Maurer, and O. S. Mitchell (2016): "Time is Money: Rational Life Cycle Inertia and the Delegation of Investment Management," *Journal of Financial Economics*, 121, 427–447.

Lusardi, A., P.-C. Michaud, and O. S. Mitchell (2017): "Optimal Financial Knowledge and Wealth Inequality," *Journal of Political Economy*, 125, 431–477.

——— (2020): "Assessing the Impact of Financial Education Programs: A Quantitative Model," *Economics of Education Review*, 78.

Lusardi, A. and O. S. Mitchell (2007): "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth," *Journal of Monetary Economics*, 54, 205–224.

——— (2011): "Financial Literacy Around the World: an Overview," *Journal of Pension Economics & Finance*, 10, 497–508.

——— (2014): "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 52, 5–44.

——— (2023): "The Importance of Financial Literacy: Opening a New Field," *Journal of Economic Perspectives*, 37, 467–475.

Markowitz, H. (1952): "Portfolio Selection," *The Journal of Finance*, 7, 77–91.

Meier, S. and C. D. Sprenger (2013): "Discounting Financial Literacy: Time Preferences and Participation in Financial Education Programs," *Journal of Economic Behavior & Organization*, 95, 159–174.

Miller, S. (2016): "Employees' Financial Issues Affect Their Job Performance," *SHRM.com*.

Thaler, R. H., S. Benartzi, et al. (2001): "Naive Diversification Strategies in Defined Contribution Saving Plans," *American Economic Review*, 91, 79–98.

TIAA (2004): "Connecting Mental and Financial Wellbeing," *TIAA Institute Insights.*

A Online Appendix Tables

Table A1: Respondent Characteristics

	Mean	SD	Median
Demographics			
Female	0.445	0.497	0
Age	52.950	14.128	54
Married or common-law	0.654	0.476	1
Has children	0.612	0.487	1
Number of household members	2.125	1.168	2
Education			
College or some university	0.347	0.476	0
Bachelor degree or more	0.516	0.500	1
Financials			
ln(Household income) (log of \$ '000)	4.035	1.975	4.605
Household income missing	0.189	0.392	0
Financial wealth (\$ 00'000)	2.486	4.890	0.5
Ownership of individual stocks	0.191	0.393	0
Ownership of domestic stocks	0.268	0.443	0
Sophistication			
Financial Literacy Score	2.513	0.776	3
Cognitive Ability Score	0.966	1.056	1
Numeracy Score	0.554	0.859	0
Financial knowledge: high (self-reported)	0.129	0.336	0
Financial knowledge: very high (self-reported)	0.050	0.218	0
Stock market knowledge: high (self-reported)	0.083	0.276	0
Stock market knowledge: very high (self-reported)	0.025	0.156	0
Has traded stocks	0.363	0.481	0
Has studied economics or finance in high school	0.298	0.457	0
Preferences			
Risk averse: 2	0.013	0.113	0
Risk averse: 3	0.047	0.212	0
Risk averse: 4	0.180	0.384	0
Risk averse: 5	0.218	0.413	0
Risk averse: 6	0.140	0.347	0
Risk averse: 7	0.090	0.287	0
Risk averse: 8	0.052	0.222	0
Risk averse: 9	0.204	0.403	0
Impatient: 2	0.607	0.489	1
Impatient: 3	0.135	0.342	0
Impatient: 4	0.037	0.189	0
Ambiguity averse	7.903	20.080	3

Note: This table presents summary statistics on control variables for the full sample. For continuous variables, we show mean and standard deviation; for binary variables we show the share. Household income missing =1 if a respondent refused to provide information on household income, 0 otherwise. We report the log of annual household income and impute missing values of this variable with the sample's mean income. Financial wealth is the sum of wealth held in RRSPs, TFSAs, defined contribution plans, and other accounts. Financial Literacy Score is the sum of correct answers to Big Three questions measuring financial literacy (Lusardi and Mitchell, 2007, 2011), Cognitive Ability Score is the sum of correct answers to the three question cognitive reflection test (Frederick, 2005), and Numeracy Score is the sum of correct answers to the 3 question Berlin numeracy test (Cokely et al., 2012). Indicators of risk aversion report where in the multiple price list respondents switched to the riskier lottery. A higher switching point suggests higher risk aversion. N = 1993.

Table A2: Factors Associated with Performance on Allocation Task 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$Mean_1$	Standard $Deviation_1$	Sharpe $Ratio_1$	RML_1	RSL_1	$1/K_1$	Return $Chasing_1$
Female	0.002	0.332	0.014	0.359	0.286	-0.022	0.026
	(0.311)	(0.552)	(0.020)	(0.284)	(0.558)	(0.020)	(0.015)
College or some university	-0.218	0.023	0.002	0.530	0.979	-0.059	-0.005
	(0.462)	(0.821)	(0.029)	(0.423)	(0.830)	(0.028)	(0.022)
Bachelor degree or more	-0.781	-0.575	0.039	1.097	1.631	-0.034	0.004
	(0.463)	(0.822)	(0.029)	(0.423)	(0.830)	(0.028)	(0.022)
ln(Household income)	-0.178	-0.224	0.008	0.139	0.227	-0.005	-0.006
	(0.075)	(0.133)	(0.005)	(0.068)	(0.134)	(0.005)	(0.003)
Financial wealth	-0.064	-0.111	0.005	0.006	0.014	-0.003	-0.000
	(0.034)	(0.061)	(0.002)	(0.031)	(0.061)	(0.003)	(0.002)
Financial Literacy Score	0.006	0.163	0.001	0.229	0.484	-0.065	-0.001
	(0.217)	(0.385)	(0.014)	(0.198)	(0.389)	(0.012)	(0.010)
Cognitive Ability Score	-0.570	-0.909	0.042	0.136	0.102	-0.006	-0.001
	(0.171)	(0.303)	(0.011)	(0.156)	(0.307)	(0.011)	(0.008)
Numeracy Score	-0.324	-0.606	0.027	-0.046	-0.074	-0.021	-0.000
	(0.201)	(0.357)	(0.013)	(0.184)	(0.361)	(0.014)	(0.010)
Financial knowledge: high	0.684	1.334	-0.003	0.010	-0.564	-0.053	0.053
	(0.523)	(0.928)	(0.033)	(0.478)	(0.938)	(0.037)	(0.023)
Financial knowledge: very high	0.525	1.041	-0.047	0.098	0.210	-0.023	0.004
	(0.837)	(1.487)	(0.053)	(0.766)	(1.503)	(0.056)	(0.041)
St. market knowledge: high	-0.033	0.313	-0.012	0.544	0.993	0.002	0.004
	(0.648)	(1.150)	(0.041)	(0.592)	(1.162)	(0.047)	(0.029)
St. market knowledge: very high	0.462	-0.171	0.022	-1.533	-3.101	0.020	0.009
	(1.130)	(2.007)	(0.072)	(1.033)	(2.027)	(0.078)	(0.052)
Has traded stocks	0.712	1.243	-0.010	-0.161	-0.559	-0.052	0.045
	(0.346)	(0.615)	(0.022)	(0.317)	(0.621)	(0.023)	(0.016)
Has studied economics	0.011	0.237	-0.024	0.365	0.843	-0.003	-0.011
	(0.328)	(0.582)	(0.021)	(0.300)	(0.588)	(0.022)	(0.016)
Mean	31.679	26.056	1.374	3.883	7.628	0.244	0.108
r2	0.055	0.045	0.054	0.029	0.024		
chi2						177.675	54.232

Standard errors in parentheses

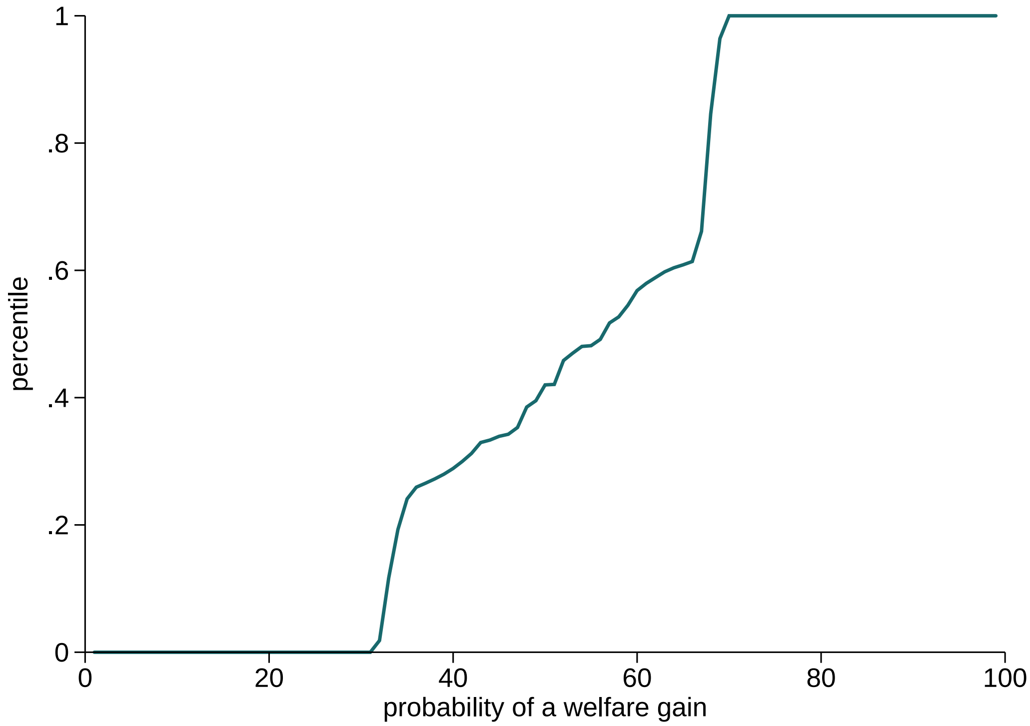
Note: This table is based on the full sample. The dependent variables are defined in the table notes of Table 1. Columns 1-5 report OLS coefficient estimates. Columns 6-7 report marginal effects from Logit regressions. All regressions control for region, ownership of individual stocks, ownership of domestic stocks, marital status, children, number of HH members, age, and preferences such as risk and ambiguity aversion as well as patience; we also include a control for having missing income information. N = 1993.

Table A3: Estimates of the Latent Model for Risk Aversion from Multiple Price List Elicitation

	point estimate	standard error
age (z)	0.019	0.026
female	0.093	0.048
married	-0.050	0.054
some college	-0.044	0.075
college	-0.101	0.074
any kids	0.035	0.054
log hh income	1e-4	1.2e-2
income missing	0.055	0.065
log fin. wealth	-0.036	0.010
constant	0.805	0.094
σ_γ	1.016	0.018

Note: This table reports maximum likelihood estimates of the latent model for risk aversion, $\gamma_i^* = z_i\kappa + \nu_i$ where ν_i is normally distributed with standard deviation σ_γ. Standard errors are estimated using the inverse of the numerical hessian. N = 1993.

Figure A1: Cumulative Distribution of the Probability of a Welfare Gain ψ^+



Note: This figure reports the cumulative distribution function of the probability of a welfare gain from task 1 to 2.

Table A4: Regression Estimates of Factors Associated with Respondent Expectations about the Financial Education Program

	(1)	(2)	(3)	(4)
	Apply information: yes	Apply information: dk	Exp. higher return: yes	Exp. higher return: dk
Female	-0.078	0.049	-0.027	0.024
	(0.025)	(0.020)	(0.025)	(0.022)
College or some university	0.051	-0.027	0.036	-0.026
	(0.038)	(0.028)	(0.038)	(0.032)
Bachelor degree or higher	0.051	-0.043	0.038	-0.059
	(0.038)	(0.029)	(0.038)	(0.032)
ln(Household income)	0.010	0.007	0.006	0.022
	(0.006)	(0.006)	(0.006)	(0.007)
Household income missing	-0.181	0.128	-0.136	0.130
	(0.032)	(0.022)	(0.032)	(0.025)
Financial wealth	0.001	-0.001	-0.001	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)
Financial Literacy Score	0.071	-0.036	0.065	-0.058
	(0.018)	(0.013)	(0.018)	(0.014)
Cognitive Ability Score	0.007	0.007	0.023	0.001
	(0.013)	(0.012)	(0.014)	(0.013)
Numeracy Score	0.046	-0.024	0.041	-0.008
	(0.016)	(0.015)	(0.016)	(0.015)
Financial knowledge: high	0.176	-0.020	0.067	-0.070
	(0.043)	(0.039)	(0.043)	(0.043)
Financial knowledge: very high	0.015	-0.061	-0.040	-0.106
	(0.065)	(0.063)	(0.066)	(0.067)
St. market knowledge: high	-0.104	0.005	-0.103	0.040
	(0.054)	(0.049)	(0.054)	(0.054)
St. market knowledge: very high	-0.094	0.011	-0.110	0.146
	(0.092)	(0.089)	(0.094)	(0.087)
Has traded stocks	0.030	-0.039	0.074	-0.070
	(0.027)	(0.024)	(0.027)	(0.026)
Has studied economics	0.075	-0.076	0.060	-0.047
	(0.026)	(0.023)	(0.026)	(0.025)
Mean_1	0.022	-0.005	0.082	-0.131
	(0.049)	(0.042)	(0.050)	(0.048)
$\text{Standard Deviation}_1$	-0.005	-0.008	-0.038	0.058
	(0.025)	(0.021)	(0.025)	(0.023)
Sharpe Ratio_1	0.183	-0.231	0.195	-0.336
	(0.168)	(0.147)	(0.170)	(0.164)
RML_1	-0.009	0.008	0.014	-0.043
	(0.023)	(0.020)	(0.024)	(0.021)
RSL_1	0.011	-0.005	0.011	-0.005
	(0.005)	(0.005)	(0.005)	(0.005)
$1/\text{K}_1$	-0.029	-0.032	-0.073	0.024
	(0.038)	(0.031)	(0.038)	(0.033)
Return Chasing_1	-0.102	0.148	-0.145	0.179
	(0.094)	(0.085)	(0.094)	(0.091)
Mean	0.462	0.193	0.467	0.269
chi2	272.890	164.537	256.316	219.204

Standard errors in parentheses

Note: This table reports marginal effects from Logit regressions for the analysis sample. *Apply information: yes* and *Apply information: dk* are dummy variables equal to one if the participant responded with "yes" or "don't know" to the question "Do you think you will be able to apply the financial information provided to your investment decision in Allocation Task 2, later in this survey?". The dummy variables *Exp. higher return: yes* and *Exp. higher return: dk* equal one if the participant responded with "yes" or "don't know" to the question "Do you expect your total return from Allocation Task 2 to be higher than the total return from Allocation Task 1, if you acquire additional financial information?". All regressions also control for region, ownership of individual stocks, ownership of domestic stocks, marital status, children, number of HH members, age, and preferences such as risk aversion, ambiguity aversion, and patience.