Pop-Up Fiction: A Trading App Simulation To Boost Financial Competence

Gabriele Iannotta[a], Katharina Hartinger[b], and Tommaso Agasisti[a]

[a]Politecnico di Milano, Department of Management, Economics and Industrial Engineering

[b]Johannes Gutenberg University Mainz, Chair of Public and Behavioral Economics

Working paper

Abstract

Modern investment platforms, such as trading apps, create new opportunities and challenges for financial education. We design and evaluate a simple and scalable simulation-based financial literacy intervention for young adults. Participants explore the interface of a trading app by trading fictitious assets in an incentivized 20-round game. In a randomized experiment with 485 university students, we find that the simulation significantly increases financial knowledge and improves portfolio choices two weeks after the intervention, as reflected in higher portfolio efficiency measured by a design-based Sharpe ratio.

Adding short educational pop-ups on challenges in modern trading further strengthens these effects, leading to substantially more efficient allocations also in payoff-relevant decisions. By contrast, self-reported risk attitudes remain unchanged, indicating that the intervention primarily improves how knowledge is translated into behavior rather than altering stated preferences.

1 Introduction

Improving financial behaviors is widely considered as the ultimate goal of financial education (FE) interventions. The persistence of bad financial habits highlights the need to support young adults in making better financial decisions (West and Cull, 2020). This persistence is contrasted by drastic change in the financial sector, where trading platforms have gained substantial traction in recent years. Often praised for their inclusivity by creating easily accessible low-stakes trading opportunities, trading apps, such as Robinhood and Trade Republic, have brought with them a new set of challenges for financial decision-making: These apps often encourage frequent trading and incorporate gamified elements, such as awarding badges and organizing lotteries — effectively blurring the line between trading and gambling. As Chapkovski et al. (2024) demonstrate, gamified interfaces can impair decision-making competence by fostering excessive risk-taking behaviors.

This paper makes the case that FE holds considerable promise for mitigating the dangers posed by these platforms if it is able to adjust to these new developments in the financial sector productively. Specifically, we eval-

uate different fast-paced learning-by-doing education concepts in an online experiment with undergraduate students at an Italian university. In total, 704 students were randomized after completing a pre-questionnaire, and 485 completed the full study protocol and constitute the main analytic sample. Students in the treatment arms participated either in a simulation game resembling a trading app or in the same game augmented by short educational messages focusing on non-technical skills relevant to app-based trading. We find that the *laissez-faire* learning by doing simulation approach is successful at improving financial behaviors, but significantly less effective than the augmented non-technical-skill-focused simulation game.

Although traditional research has often equated financial knowledge with overall financial competence, Ambuehl et al. (2014) caution that even when individuals acquire financial principles, they may not implement them — or may deploy them incorrectly. This concern is especially pronounced in the context of trading apps with their unique psychological and behavioral dynamics and with financial novices as an important target group of such platforms.

While quantitative skills form a fundamental component of financial capability, recent studies (Carpena and Zia, 2020; Horn et al., 2023) indicate that these skills play only a limited role in real-world financial decision-making. Instead, the key challenge lies in bridging the gap between financial knowledge and behaviors. Carpena et al. (2019) emphasize the importance of understanding the psychological and contextual barriers that prevent individuals from translating knowledge into effective action, while Yoong (2011) advocates linking FE directly to concrete financial behaviors. This perspec-

tive is further reinforced by Amagir et al. (2018), who argue for integrating behavioral economics into FE programs to more effectively target the cognitive biases and decision-making pitfalls inherent in complex financial environments.

Moreover, a set of interventions aimed at stimulating more conscious financial behaviors is grounded in active learning principles — particularly experimental learning. As noted by Kaiser and Menkhoff (2020), when educational programs are made more entertaining or personalized, their impact on financial behaviors is significantly enhanced. Lobão (2020) further contends that experimental learning activities, such as virtual market participation and trading simulations, allow participants to directly experience specific situations — both in content and context. In addition, experimental games help overcome limitations of traditional survey methods, which often capture only self-reported data. By incorporating appropriate incentives, such games encourage participants to engage fully and exert the necessary effort to solve assigned tasks (Muñoz-Murillo et al., 2020).

This study examines several core aspects of financial literacy through an innovative experimental design. It focuses on how individuals approach asset allocation by comparing self-reported intentions with actual investment decisions made in a simulation that closely mimics a modern trading app environment. For example, the Linciano et al. (2019) report reveals a widespread lack of diversification knowledge among Italian households — a finding corroborated by studies from Mouna and Jarboui (2015), Abreu and Mendes (2010), and Guiso and Jappelli (2009), which document a direct link between low financial literacy and undiversified portfolios. Beyond portfolio

diversification, the study investigates the role of risk aversion — a central but empirically ambiguous factor in financial decision-making. Existing evidence provides mixed conclusions: some studies suggest that greater financial literacy may heighten risk aversion (Amari et al., 2020; Dinç Aydemir and Aren, 2017; Litterscheidt and Streich, 2020), while others indicate that it may instead reduce it (Aren and Zengin, 2016; Bajo et al., 2015; Bayar et al., 2020; Hermansson and Jonsson, 2021). A key reason for this ambiguity lies in how risk aversion is measured. Much of the literature relies on self-reported indicators, which capture individuals' stated attitudes toward risk, but may not fully reflect how they behave when facing actual investment choices. By explicitly combining validated self-reported measures of risk attitudes with observed allocation choices made in an incentivized trading simulation, this study goes beyond standard survey-based approaches. It provides direct evidence on whether FE affects not only stated risk preferences, but also how these preferences translate into actual risk-taking under repeated investment decisions.

Furthermore, our research addresses asset allocation more broadly. EFAMA (2020) indicates that most EU citizens keep a disproportionate amount of their savings in bank deposits — a situation driven by insufficient financial literacy and pervasive risk aversion. Limited understanding of the advantages and disadvantages of different capital market instruments, combined with a predisposition to favor guaranteed savings vehicles, often leads to suboptimal investment decisions. Simultaneously, trading apps can induce risk-taking behaviors that result in over-trading and other inefficient outcomes. Consequently, it is of first-order importance to understand how FE can promote

both enhanced financial knowledge and improved decision-making competence within these new, digitally mediated, and often hard-to-regulate financial environments (with the aim of reducing excessive risk behaviors).

This paper contributes to the existing literature by (i) providing a rigorous experimental evaluation of FE interventions tailored to the trading app context and (ii) offering new insights into how FE can translate financial knowledge into effective, real-world investment behaviors. By integrating theoretical insights from behavioral finance with an innovative experimental design that captures both self-reported and observed measures, our study advances the discussion on how FE can be most effectively deployed to enhance financial decision-making in a rapidly evolving digital landscape.

2 Education Concept

We model the trading app simulation as an opportunity for participants to invest in their human capital — specifically, their FE. When applying the simple human capital framework (Becker, 1962) to a learning-by-doing simulation game, the participant weighs the costs and future benefits of training. Specifically, since participants in the treatment arms had to complete the simulation in any case, the decision made by the participant is whether or not to use the simulation as a learning opportunity and invest an appropriate amount of effort. The benefits of investing effort are twofold: an immediate monetary benefit arises from success in the simulation game itself, which determined the participant's payoff. In addition, as in a typical education decision, there are long-run returns to education. Unlike in the classical case,

these returns do not occur in the labor market but the stock market when the participants decide to use trading apps or similar digital investment interfaces. Since the investment game is a stochastic game, the decision to invest effort did not follow mechanically from the set-up, but can plausibly be seen as a deliberate (albeit potentially subconscious) choice made by the participant.

Zooming in on the design of the simulation game illuminates the outcomes on which we expect an effect as well as the underlying mechanisms: The core idea behind the trading app simulation game is a learning-by-doing approach. Rather than teaching financial literacy (only) in the context of a regular class, the simulation game allows users to actively participate in trading, observe consequences of their investment choices over multiple rounds, and gain a deeper understanding of the stochastic nature of trading. Crucially, the simulation bridges the gap between completely risk-free trading in a FE exercise — since the simulation game will affect the final payoff — and real-stakes trading on a real-world trading platform. The close link between in-simulation behavior and in-app behavior allows participants to translate new knowledge directly into behavior. Thus, we expect a significant effect of trading simulation use on financial knowledge and behaviors.

This effect should be further augmented by the pop-up treatment in the respective group, as pop-ups reinforce the pedagogical nature of the simulation task. The pop-ups are designed as brief educational prompts that increase the salience of common behavioral mechanisms relevant to investment decision-making within an app-based trading environment at moments

when participants are actively engaged with investment decisions[1] . Some pop-ups convey general investment principles, such as the importance of diversification, while others address well-documented behavioral biases that affect trading behavior, including the disposition effect and availability bias. Importantly, two of the five pop-ups focus explicitly on phenomena that are particularly salient in trading app contexts: overtrading, and herd behavior arising from platform features that encourage social comparison and the imitation of other investors' decisions. Due to their general and practical nature, the pop-ups have very limited direct relevance in the context of the simulation game: They do not disclose any additional information about the game's payoff structure (e.g., probabilities, correlations, or optimal allocations), nor do they provide round-specific recommendations. In this sense, they are not intended to function as performance guidance or "solutions" to the task. They are, not least due to their short format, invitations to dig deeper into specific aspects of financial skill and trading app use. Pop-ups bring attention to their contents in an engaging and interactive way. Recent literature

[1] A dedicated appendix reporting the full set of pop-up messages and corresponding screenshots will be included in a subsequent version of the manuscript. To give a concrete sense of their tone and informational content, we report here the text of the diversification-related pop-up: "*Are you risk-averse? Risk-neutral? Or perhaps a risk-seeker? Regardless of your risk profile, portfolio diversification is a key strategy. Why? Because diversifying — i.e., investing in multiple assets — offers a major advantage: it reduces volatility and stabilizes your returns. Be careful though: one security is not the same as another. Trading platforms offer a wide range of options, often organized by popularity or performance, but "different" in finance does not simply mean that asset A is not asset B. Rather, it means that assets are imperfectly correlated or move in opposite directions. A truly diversified portfolio is composed of assets that do not all follow the same path*".

in economics has emphasized the role of attention in decision-making and the costs associated with being attentive. Pop-ups with their element of surprise are well suited for capturing attention and are often used as marketing tools for this reason. We use the pop-ups to convey messages about non-technical aspects of using trading apps in an empathetic way. A further pop-up design feature is central to our education concept. The prompts are exogenously timed: they appear in pre-determined rounds irrespective of prior choices or performance, provided that the participant reaches the relevant round.[2] Thus, they do not provide dynamic feedback in any way, distinguishing our intervention from the literature on personalized real-time feedback or adaptive learning (Aucejo and Wong, 2025; Brade et al., 2022; Muralidharan et al., 2019; Van Klaveren et al., 2017). In contrast to, say, a personal one-on-one meeting with a financial advisor, the pop-up-augmented simulation is a cost-efficient and easily scalable intervention to increase knowledge and improve financial decision-making.

3 Methodology and Research Questions

To rigorously evaluate the effectiveness of our FE interventions, we conducted an online randomized experiment with undergraduate students at Politecnico di Milano. Ethical approval was obtained through Politecnico's IRB. The experiment was not pre-registered. To ensure transparency, we provide a dedicated appendix documenting the experimental design, outcome construction, and the full set of specifications reported in the paper. We re-

[2]We find that variation in exposure due to early termination is limited.

cruited undergraduate students from Politecnico di Milano for this online experimental study. Recruitment began with an email invitation, accompanied by an informational flyer, sent to all eligible students. Participants were directed to a dedicated online platform, where they first completed a pre-questionnaire from November 22 to 26, 2023. This questionnaire established baseline levels of financial knowledge, attitudes, and self-reported behaviors. Upon submission, 704 students were randomly assigned to one of three groups, ensuring that any observed differences in outcomes could be attributed to the intervention rather than to pre-existing variations among individuals (see Figure1).

The control group (C), consisting of 180 students, received no additional intervention. Treatment Group 1 (T1), comprising 153 students, participated in an online experimental game between November 30 and December 3, 2023, which simulated a trading app environment and included only basic instructions on game mechanics and a set of financial "pills" intended to inform them about fundamental financial topics. Treatment Group 2 (T2), with 152 students, followed the same simulation schedule but, in addition to the basic information, received in-game educational pop-up messages throughout the game. These pop-ups addressed key financial behaviors and biases — such as diversification, overtrading, the disposition effect, availability bias, and herd behavior — by explicitly drawing attention to common decision-making pitfalls in app-based trading. Importantly, the pop-ups were not personalized, were not contingent on prior choices or performance, and did not provide information about the underlying payoff structure or optimal strategies. Through this design, we can evaluate both the effectiveness of

the FE program per se and its specific characteristics that contribute to the effect.

All participants completed a post-questionnaire between December 13 and 17, 2023. Each activity — whether a questionnaire or the simulation — had to be finished within a 12-hour window. The incentive scheme awarded a €5 reward for completing both questionnaires, supplemented by potential earnings from the simulation, and bank transfers were issued only upon successful completion of all required tasks. Overall, 485 students completed the full protocol (both questionnaires and, for treated students, the simulation) and constitute the main analytic sample used in the results section. Because participation requirements differed across arms (treated students also had to complete the game), attrition is an important component of the study design. For transparency, we therefore document attrition patterns and assess whether dropout is systematically related to baseline measures in Section 7. The structured timeline and randomized assignment ensure that differences in outcomes across arms can be interpreted as causal within the completion sample, while the appendix helps clarify the role of selective participation.

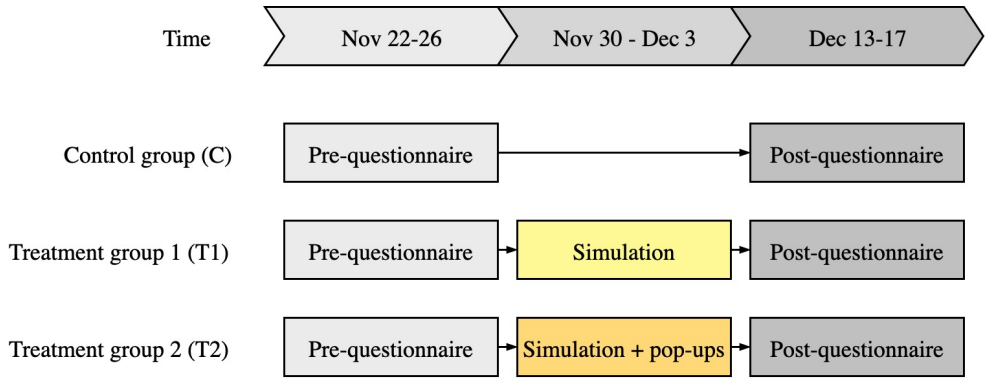


Figure 1. The experimental design and timeline

Following a standard conceptualization in the FE literature, we operationalize financial decision-making competence along three complementary dimensions: financial knowledge, financial attitudes, and financial behaviors. This structure allows us to assess whether the intervention affects what participants know, what they report about risk, and how they allocate resources under uncertainty. Across all dimensions, we combine survey-based measures with incentivized choices observed in the trading simulation.

Financial knowledge refers to an understanding of fundamental financial concepts — such as interest rates, risk, return, and the various types of financial instruments — as well as a clear awareness of the risks and dangers inherent in trading. In our study, financial knowledge is measured using a total of 14 questions. Nine of these questions, following the approach of Ungeheuer and Weber (2021), are designed to assess participants' grasp of key financial issues, ensuring that they have the essential understanding needed to make informed financial decisions. In addition, five supplementary questions evaluate participants' understanding of the content delivered through

the educational pop-ups. This combined measure allows us to capture both core financial literacy and a broader awareness of biases that can influence decision-making.

Financial attitudes are operationalized as risk attitudes, capturing how individuals perceive and approach risk in investment contexts. Stated risk attitudes are measured at both baseline and endline using the general willingness-to-take-risks item introduced by Dohmen et al. (2011), which Lönnqvist et al. (2015) identify as "the more adequate measure of individual risk attitudes for the analysis of behavior in economic (lab) experiments". Using the same item before and after the intervention allows us to assess whether FE affects reported risk tolerance in a standard and comparable format. To complement stated preferences, we derive a revealed measure of risk attitude from participants' behavior in the trading simulation. Specifically, we use the risky share, defined as the average fraction of available wealth allocated to risky assets during gameplay. The use of risky asset shares as a proxy for risk preferences is well established in both theoretical and empirical portfolio choice research: under a given investment opportunity set, the share invested in risky assets is monotonically related to risk aversion (Gomes and Michaelides, 2005), and has been widely used to infer individual risk attitudes from observed allocations (Calvet et al., 2009; Chiappori and Paiella, 2011; Guiso and Paiella, 2008). Because all participants face the same assets, payoffs, and information structure, variation in risky shares reflects differences in willingness to bear risk rather than differences in constraints or opportunities. This measurement strategy allows us to relate stated risk attitudes to revealed risk-taking in a controlled investment environment, and to assess

whether the intervention affects how reported preferences are reflected in actual allocation choices. This approach follows recent contributions emphasizing the value of integrating validated survey measures with incentivized behavioral data when studying preferences (Falk et al., 2023).

Financial behaviors are assessed using both self-reported behaviors and actual behaviors observed during the experimental simulation. Self-reported financial behaviors are captured through a dedicated asset allocation question administered at baseline, following the approach of Cavezzali et al. (2015). Participants declare their intended diversification strategies across the same set of assets featured in the simulation, based only on qualitative information about risk and return. Using these declared allocations, we compute a design-based Sharpe ratio that serves as an ex-ante measure of intended portfolio efficiency. During the simulation, actual trading behaviors — namely, portfolio allocations chosen across rounds — are recorded. Using these observed in-game allocations, we compute the same design-based Sharpe ratio as above. This yields a behavior-based measure of portfolio efficiency that is directly comparable to the questionnaire-based measure, as both rely on the same return vector and variance–covariance matrix. The two measures differ only in the source of portfolio weights (declared versus observed during gameplay), allowing us to isolate differences between intended and realized portfolio construction.

The research questions guiding our investigation are twofold. First, does participation in an experimental game that simulates a trading app environment improve participants' financial knowledge, attitudes, and behaviors compared to a control group receiving no intervention? Second, does the ad-

dition of in-game educational pop-up prompts — as provided exclusively to T2 — lead to superior financial decision-making, particularly in terms of risk attitudes and portfolio efficiency, compared to providing only basic financial information (as in T1)?

This comprehensive design allows us to isolate the effects of the educational intervention and to explore the mechanisms through which FE may enhance decision-making competence. By integrating both subjective (self-reported) and objective (behavioral) measures, we aim to generate robust evidence on how tailored FE interventions can bridge the gap between financial knowledge and effective investment behaviors.

4 The Experimental Game

The experimental game was designed to closely mimic the decision-making environment characteristic of modern trading apps, serving simultaneously as an educational intervention and a controlled laboratory for observing realistic investment behaviors. The game is fully implemented in oTree and is structured into 20 rounds, with each round representing an independent investment decision similar to what investors encounter in digital trading platforms, which in turn are representative of real-life investment decisions beyond those available in investment platforms.

At the beginning of the experiment, every participant is allocated an initial endowment of 300 points (which is equivalent to €3). In each round, the participant must decide how to distribute these points among five available investment options, here shown in Table 1. One of these options is a

zero-percent saving account that represents a risk-free alternative — offering a guaranteed, albeit neutral, return. The remaining four options are risky assets labeled A, B, C, and D. Each of these assets is characterized by a binary outcome mechanism: in each round, it can yield either a high or a low return, with the outcome determined randomly by the program. The underlying payoff probabilities that generate these outcomes (50–50 across the two scenarios) are fixed by design but not disclosed to participants, so that they are aware that each asset carries an inherent level of uncertainty but cannot infer its precise distribution or volatility. This choice ensures that the perceived riskiness of each asset relies on qualitative information rather than on calculable standard deviations, thereby reflecting the incomplete information investors face in real-world financial decisions.

To eliminate any potential framing effects or biases that might arise from associations with real-world securities, no actual asset names are used. Instead, assets are presented in abstract terms. It is worth noting that asset A includes a possible return of -130%, i.e., a loss that exceeds the initial amount invested. This feature is intentionally introduced to create an extreme downside scenario, ensuring that participants are confronted with the possibility of losses beyond the original stake. By doing so, the simulation provides a sharper contrast between safe and risky choices and allows us to examine decision-making under particularly adverse conditions.

Prior to the start of the game, participants are shown a graph that displays the historical returns of each asset, generated through a simulation of 15 rounds per asset. This historical performance chart is updated in real time as the game progresses. Notably, the presentation of historical returns

Table 1. Investment options

	Scenarios			
	S1	**S2**	**Risk**	**Return**
SA	0%	0%	Null	Null
Asset A	150%	-130%	Very high	Very high
Asset B	-40%	50%	Low	Low
Asset C	100%	-85%	High	High
Asset D	-20%	25%	Very low	Very low

is grounded in the practices of real trading apps, which routinely provide users with past performance data to inform their investment decisions. Together with the integration of quantitative payoff information and qualitative risk descriptions, this feature enhances the realism of the simulation by reproducing the informational asymmetries typical of retail trading platforms. Participants can thus base their decisions on the same combination of partial, performance-based cues that characterize decision-making in real-world app environments.

At the conclusion of every round, after participants have made their allocation decisions, the game randomly selects one of two possible scenarios, labeled S1 and S2. In S1, assets A and C pay the high return while assets B and D pay the low return. Conversely, in S2, the outcomes are reversed so that assets A and C yield the low return and assets B and D yield the high return. This procedure establishes a precise correlation structure among the assets. Specifically, assets A and C are perfectly positively correlated — that is, they always deliver the same outcome within any given round —

and similarly, assets B and D are perfectly positively correlated. In contrast, assets that are paired across these groups (for example, assets A and B, A and D, as well as other inter-group pairings such as asset C and D) are perfectly negatively correlated. In other words, when an asset from one group yields a high return, its perfectly negatively correlated counterpart will yield a low return. This meticulous design ensures that the investment options are interdependent in a controlled manner.

The rationale behind this correlation structure is twofold. First, it is intended to provide a more realistic picture of the complex interconnections that exist among financial assets in actual markets. In real-world scenarios, assets rarely behave in isolation, their returns often exhibit systematic patterns of co-movement, which can affect portfolio performance. Second, this design draws on well-established findings in behavioral finance. The literature documents a "natural" tendency toward naive diversification — where investors tend to distribute their investments evenly across available assets — as reported by Huberman and Jiang (2006) and Benartzi and Thaler (2001). More specifically, the "conditional 1/n diversification heuristic" described by Shefrin and Statman (2000) suggests that investors often conceptualize their portfolios as collections of narrowly framed subportfolios rather than as a single integrated portfolio. By employing an asset structure that embodies these characteristics, the game not only controls for any inherent naive diversification trend but also provides a rigorous framework for assessing whether our intervention can lead to a genuine improvement in investment management.

After each round, participants receive immediate graphical feedback that displays the evolution of their wealth over time. In T2, this feedback is

complemented by in-game educational pop-ups (see Figure 2). Pop-ups were programmed to appear for the first time at fixed points during the simulation (rounds 1, 3, 5, 8, and 11). Each pop-up was displayed in a non-skippable window for 20 seconds, ensuring that participants were necessarily exposed to the content before continuing with the game. This design choice ensured uniform exposure across participants and reduces concerns related to selective attention or differential take-up of the informational content. The diversification message, directly pertinent to the allocation task, was deliberately placed first to ensure that all participants encountered it, while the remaining four were arranged in a randomly determined sequence that was held constant across all participants. Concentrating these messages in the first half of the game increased the likelihood of exposure, given that some participants could exhaust their balance before completing all 20 rounds. Once displayed, each pop-up remained accessible for consultation in subsequent rounds, ensuring continuous availability throughout the game. This combined system of graphical feedback and educational prompts allowed participants to observe the impact of their allocation decisions and to adjust their strategies over time. Finally, to preserve incentives, one round was randomly selected at the end of the game to determine the final financial compensation, ensuring that every decision carried a direct and tangible consequence.

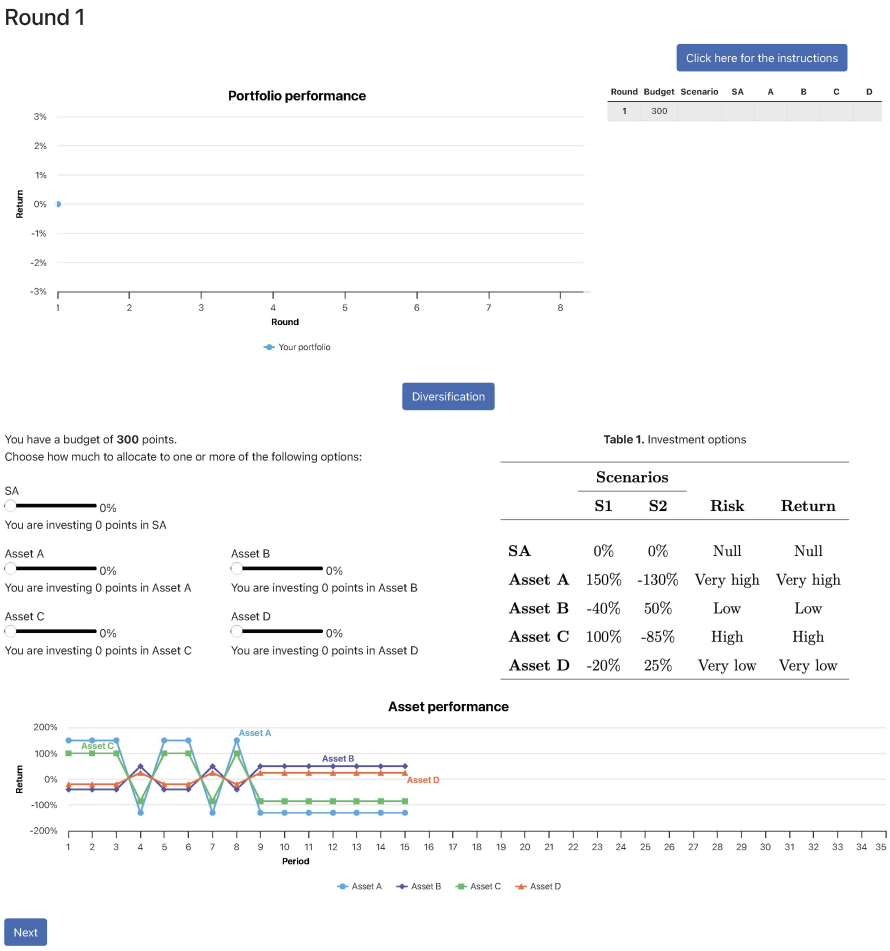


Figure 2. Graphic interface for T2

This comprehensive design, which integrates historical performance data and real-time feedback within a controlled simulation of trading app dynamics, provides a robust and realistic environment for examining the effects of our interventions on actual investment behaviors.

4.1 Measuring portfolio efficiency: The Sharpe ratio

To quantify portfolio efficiency, we rely on a Sharpe ratio constructed from participants' portfolio allocations. Importantly, the Sharpe ratio used throughout the paper is design-based rather than empirical. Expected returns and risk are not estimated from market data but are derived from the payoff structure of the experimental game, which is fixed by design and identical for all participants. This approach ensures that variation in the Sharpe ratio reflects differences in allocation choices rather than differences in the return-generating process.

For each participant, portfolio returns are computed at the round level based on the realized outcomes of the assets in that round and the participant's chosen allocation. The portfolio return used in the Sharpe ratio is the average of these realized portfolio returns over the rounds played. Portfolio risk is measured as the standard deviation of these round-level portfolio returns, capturing the volatility of outcomes implied by the participant's allocation strategy within the stochastic environment of the game. The savings account is treated as a risk-free asset with zero return and zero variance by construction.

We compute this measure in two settings. First, using the allocations declared in the questionnaire, we obtain a self-reported Sharpe ratio that reflects intended portfolio efficiency under limited and qualitative information about risk and return. Second, using the allocations observed during the simulation, we compute an in-game Sharpe ratio that captures realized portfolio efficiency under repeated choice and feedback. In both cases, the return-generating process is identical and governed by the same payoff struc-

ture.

5 Empirical Strategy and Results

Our empirical strategy employs an ordinary least squares (OLS) regression framework to identify the causal impact of the FE interventions delivered in T1 and T2 on a set of outcomes capturing financial knowledge, financial attitudes, and financial behaviors. All outcomes analyzed in this section are observed for students who completed the full study protocol (N = 485). In our baseline specification, the regression model is given by

$$Y_i = \beta_0 + \beta_1 \cdot Treat_{1_i} + \beta_2 \cdot Treat_{2_i} + \lambda \cdot X_i + \epsilon_i,$$

where Y_i denotes the outcome of interest for student i, and $Treat_{1_i}$ and $Treat_{2_i}$ are indicators for assignment to treatment groups T1 and T2. The vector X_i includes a set of baseline characteristics measured prior to treatment assignment, such as pre-test financial knowledge, pre-test risk attitudes, pre-test portfolio performance indicators, and the time required to complete the pre-test questionnaire. These variables capture heterogeneity in initial conditions and improve the precision of the estimates. In specifications focusing on in-game outcomes, we additionally incorporate information from the trading simulation logs, including measures of time spent reading instructions, average time per round, number of rounds played, and in-game earnings. These variables describe how participants interacted with the simulation environment and allow us to descriptively relate realized performance to observed patterns of engagement during gameplay. These variables are

included to characterize how participants interacted with the simulation environment and should not be interpreted as causal mechanisms.

5.1 Full-sample results

Table 3 reports descriptive statistics for the main variables across experimental arms. Pre-treatment comparability is high: pre-test financial knowledge, risk attitudes, and baseline portfolio characteristics are statistically indistinguishable across groups, consistent with successful randomization. Differences emerge primarily in post-test outcomes. Post-test financial knowledge increases markedly for treated students, with the largest gains observed in T2. While post-test self-reported risk attitudes remain stable across arms, portfolio-based outcomes reveal meaningful improvements in efficiency. In particular, volatility computed from post-test declared allocations declines for treated students, whereas differences in post-test returns are small and statistically insignificant. As a result, post-test self-reported Sharpe ratios are higher in both treatment groups, with the largest gains again observed in T2.

5.1.1 Financial knowledge

Turning first to financial knowledge, Table 4 shows that both T1 and T2 significantly increase post-test knowledge scores. Relative to the control group mean of 0.57, the estimated coefficients imply gains of approximately 8% for students in T1 and 16% for students in T2, corresponding to 0.24 SD and 0.53 SD, respectively. These effects are precisely estimated and remain stable when conditioning on pre-treatment covariates, including

Table 2. Variable definitions (full sample)

Variable	Description	Value/range
Individual characteristics		
Pre-test (Post-test) trading app use_No	Whether the student had never used a trading app and did not know what it was before (after) the intervention	1; otherwise
Pre-test (Post-test) trading app use_No, but	Whether the student had never used a trading app but knew what it was before (after) the intervention	1; otherwise
Pre-test (Post-test) trading app use_Yes	Whether the student had used a trading app before (after) the intervention	1; otherwise
Financial knowledge		
Pre-test score	The student's financial knowledge score before the intervention	from 0.11 to 0.88
Post-test score	The student's financial knowledge score after the intervention	from 0.11 to 1.00
Financial attitudes		
Pre-test (Post-test) risk attitude	The student's self-reported risk attitude before (after) the intervention	from 2 to 10
In-game risk attitude	Average share of the portfolio allocated to risky assets during the simulation	from 0.1425 to 1
Financial behaviors		
Pre-test return	Portfolio return based on the pre-test asset allocation	from 0.00375 to 0.10
Post-test return	Portfolio return based on the post-test asset allocation	from 0.00250 to 0.10
Pre-test volatility	Portfolio standard deviation based on the pre-test asset allocation	from 0.00375 to 1.40
Post-test volatility	Portfolio standard deviation based on the post-test asset allocation	from 0.0025 to 1.40
Pre-test SR	Portfolio Sharpe ratio based on the pre-test asset allocation	from 0.0714 to 7
Post-test SR	Portfolio Sharpe ratio based on the post-test asset allocation	from 0.0714 to 21.00
In-game SR	Portfolio Sharpe ratio computed from in-game asset allocations	from 0.0768 to 9.91
Pre-test SA	Percentage of the portfolio allocated to the savings account in the pre-test	from 0 to 0.9
Post-test SA	Percentage of the portfolio allocated to the savings account in the post-test	from 0 to 0.95
Experimental game data		
Pre-test time	Time (in seconds) the student took to complete the pre-test questionnaire	from 180 to 19,920
Time on instructions	Time (in seconds) spent reading the game instructions	from 5 to 3,227
Time per round	Average time (in seconds) spent per round during the simulation	from 8.90 to 775.15
Rounds played	Total number of rounds completed during the simulation	from 1 to 20
In-game earnings	Total earnings obtained during the simulation (in euros)	from 0 to 23.02

Table 3. Descriptive statistics (full sample)

Variable	C	T1	T2	p-value
Pre-test score (w/ biases)	0.61 (0.15)	0.61 (0.16)	0.61 (0.14)	>0.9
Pre-test score	0.58 (0.16)	0.57 (0.17)	0.58 (0.15)	0.9
Post-test score (w/ biases)	0.60 (0.15)	0.61 (0.14)	0.70 (0.14)	<0.001
Post-test score	0.57 (0.17)	0.61 (0.16)	0.66 (0.16)	<0.001
Pre-test risk attitude	5.89 (1.66)	5.93 (1.76)	5.90 (1.67)	>0.9
Post-test risk attitude	6.04 (1.64)	6.08 (1.50)	5.89 (1.66)	0.5
Pre-test return	0.048 (0.014)	0.047 (0.016)	0.045 (0.015)	0.038
Pre-test volatility	0.23 (0.19)	0.25 (0.22)	0.21 (0.19)	0.4
Pre-test SR	0.56 (1.10)	0.56 (1.06)	0.57 (1.04)	0.7
Post-test return	0.049 (0.013)	0.046 (0.015)	0.046 (0.015)	0.10
Post-test volatility	0.26 (0.20)	0.24 (0.22)	0.22 (0.20)	0.010
Post-test SR	0.43 (0.86)	0.80 (2.05)	1.06 (2.79)	0.11
Pre-test SA	0.20 (0.20)	0.23 (0.21)	0.25 (0.21)	0.049
Post-test SA	0.20 (0.19)	0.23 (0.22)	0.22 (0.20)	0.3
Pre-test time (s)	1,916 (2,794)	1,413 (929)	1,606 (1,588)	0.5
Pre-test trading app use				0.4
No	40 (22%)	30 (20%)	24 (16%)	
No, but	91 (51%)	82 (54%)	93 (61%)	
Yes	49 (27%)	41 (27%)	35 (23%)	
Post-test trading app use				<0.001
No	34 (19%)	5 (3.3%)	7 (4.6%)	
No, but	100 (56%)	95 (62%)	105 (69%)	
Yes	46 (26%)	53 (35%)	40 (26%)	
In-game risk attitude (risky share)	–	0.75 (0.23)	0.79 (0.21)	0.20
In-game SR	–	0.69 (1.13)	1.05 (1.72)	0.005
Time on instructions (s)	–	426 (466)	509 (651)	0.9
Time per round (s)	–	84 (90)	89 (69)	0.054
Rounds played	–	18.7 (3.7)	19.1 (3.3)	0.15
In-game earnings (€)	–	4.3 (3.2)	4.8 (4.0)	0.6
N	180	153	152	

Note: Continuous variables are reported as means with standard deviations in parentheses. Categorical variables are reported as frequencies and percentages. Post-test variables are highlighted in light gray. "SA" stands for "savings account", "SR" stands for "Sharpe ratio". "No, but" corresponds to the answer "no, but I know what it is" to the question "Have you ever used a trading platform?", while "No" corresponds to "no, and I have no idea what it is". *p*-values for continuous variables are based on Kruskal–Wallis rank-sum tests (Wilcoxon rank-sum tests for variables below the dashed line), while *p*-values for categorical variables are based on Pearson's chi-squared tests. Variables below the dashed line are computed only for treated students (T1 and T2).

baseline knowledge, risk attitudes, and prior exposure to trading apps.

5.1.2 Financial attitudes

We next turn to financial attitudes, captured through both stated and revealed measures of risk preferences. We begin with self-reported risk attitudes elicited via the Dohmen et al. (2011) item. Table 5 shows that neither T1 nor T2 has a statistically significant effect on post-test stated risk attitudes.

We then examine revealed risk attitudes observed during the trading simulation, measured by the share of wealth allocated to risky assets. Table 6 shows that students in T2 allocate a significantly higher fraction of their portfolio to risky assets than those in T1. In the specification including pretreatment controls, the estimated coefficient corresponds to an increase of about 0.045 in the risky share, equivalent to roughly a 5-6% rise relative to the T1 mean. Including in-game log measures yields a similar estimate, while providing additional context on how students interacted with the simulation. Overall, these results suggest that while stated risk preferences remain unchanged, the augmented simulation affects how students translate those preferences into actual, payoff-relevant portfolio choices.

5.1.3 Financial behaviors

Finally, we turn to financial behaviors, captured through portfolio efficiency measures derived from both declared and realized allocations. We begin with self-reported behavior. Table 7 documents that both treatments significantly increase the Sharpe ratio computed from post-test declared allo-

Table 4. Treatment effect on post-test score

	Dependent variable:	
	Post-test score	
	(1)	(2)
T1	0.039**	0.048***
	(0.018)	(0.015)
T2	0.086***	0.090***
	(0.018)	(0.015)
Pre-test score		0.492***
		(0.042)
Pre-test time		0.005
		(0.011)
Pre-test trading app use_No, but		0.018
		(0.017)
Pre-test trading app use_Yes		0.030
		(0.020)
Pre-test risk attitude		−0.010**
		(0.004)
Pre-test SR		0.007
		(0.007)
Pre-test SA		−0.076**
		(0.033)
Observations	485	485
R^2	0.046	0.325
Adjusted R^2	0.042	0.313
Residual Std. Error	0.162 (df = 482)	0.137 (df = 475)

Note: *p<0.1; **p<0.05; ***p<0.01. "Pre-test time" and "Pre-test SR" were logged due to their right-skewed distributions. "T1" stands for "treatment group 1", "T2" stands for "treatment group 2", "SR" stands for "Sharpe ratio", and "SA" stands for "savings account".

Table 5. Treatment effect on post-test risk attitude

	Dependent variable:	
	Post-test risk attitude	
	(1)	(2)
T1	0.053	0.041
	(0.176)	(0.124)
T2	−0.124	−0.086
	(0.176)	(0.125)
Pre-test score		−0.791**
		(0.339)
Pre-test time		0.145*
		(0.086)
Pre-test trading app use_No, but		0.125
		(0.136)
Pre-test trading app use_Yes		0.443***
		(0.162)
Pre-test risk attitude		0.619***
		(0.032)
Pre-test SR		−0.132**
		(0.056)
Pre-test SA		−0.743***
		(0.268)
Observations	485	485
R^2	0.002	0.518
Adjusted R^2	−0.002	0.509
Residual Std. Error	1.600 (df = 482)	1.119 (df = 475)

Note: *p<0.1; **p<0.05; ***p<0.01. "Pre-test time" and "Pre-test SR" were logged due to their right-skewed distributions. "T1" stands for "treatment group 1", "T2" stands for "treatment group 2", "SR" stands for "Sharpe ratio", and "SA" stands for "savings account".

Table 6. Treatment effect on in-game risk attitude

	Dependent variable:		
	In-game risk attitude		
	(1)	(2)	(3)
T2	0.038	0.044**	0.046**
	(0.025)	(0.022)	(0.022)
Pre-test score		0.063	0.072
		(0.074)	(0.074)
Pre-test time		−0.019	−0.016
		(0.021)	(0.021)
Pre-test trading app use_No, but		0.080**	0.085***
		(0.031)	(0.031)
Pre-test trading app use_Yes		0.042	0.036
		(0.037)	(0.037)
Pre-test risk attitude		0.002	0.001
		(0.007)	(0.007)
Pre-test SR		0.009	0.013
		(0.012)	(0.012)
Pre-test SA		−0.489***	−0.472***
		(0.056)	(0.056)
Time on instructions			−0.001
			(0.009)
Time per round			0.001
			(0.017)
Rounds played			−0.009***
			(0.003)
In-game earnings			0.001
			(0.003)
Observations	305	305	305
R^2	0.008	0.264	0.284
Adjusted R^2	0.004	0.244	0.255
Residual Std. Error	0.219 (df = 303)	0.191 (df = 296)	0.190 (df = 292)

Note: *p<0.1; **p<0.05; ***p<0.01. "In-game SR", "Pre-test time", "Time on instructions", "Time per round", and "Pre-test SR" were logged due to their right-skewed distributions. "T2" stands for "treatment group 2", "SR" stands for "Sharpe ratio", and "SA" stands for "savings account".

cations. Relative to the control group mean of 0.43, the estimated coefficients correspond to gains of roughly 27% for T1 and 36% for T2. Baseline portfolio quality is a strong predictor of post-test efficiency, indicating persistence in allocation skills.

We then examine realized behavior during the trading simulation, focusing on differences between the two treatment groups. Table 8 shows that students in T2 achieve significantly higher in-game Sharpe ratios than those in T1. In the baseline specification without covariates, the T2 coefficient implies a sizable improvement in realized portfolio efficiency. Conditioning on pre-treatment characteristics yields a similar estimate, corresponding to an increase of approximately 36% relative to T1.

Augmenting the specification with in-game process measures sheds light on underlying mechanisms. Both the number of rounds played and the average time per round are strongly and positively associated with the in-game Sharpe ratio, while the treatment coefficient attenuates in magnitude. This pattern suggests that more sustained engagement with the simulation task — rather than treatment assignment alone — is a key driver of improved realized performance. Taken together, these results highlight that the augmented simulation not only improves intended portfolio choices, but also translates into more efficient behavior under repeated, payoff-relevant decision-making.

5.2 Demographic analysis

Analyses leveraging individual-level demographics and academic background are performed on the subset of participants who consented to data processing (317 students for post-test knowledge and self-reported SR; 215

Table 7. Treatment effect on post-test self-reported SR

	Dependent variable:	
	Post-test self-reported SR	
	(1)	(2)
T1	0.218*	0.236**
	(0.116)	(0.106)
T2	0.343***	0.307***
	(0.116)	(0.106)
Pre-test score		0.623**
		(0.290)
Pre-test time		0.001
		(0.073)
Pre-test trading app use_No, but		0.256**
		(0.116)
Pre-test trading app use_Yes		0.085
		(0.138)
Pre-test risk attitude		−0.032
		(0.027)
Pre-test SR		0.445***
		(0.048)
Pre-test SA		−0.113
		(0.229)
Observations	485	485
R^2	0.018	0.208
Adjusted R^2	0.014	0.193
Residual Std. Error	1.057 (df = 482)	0.956 (df = 475)

Note: *p<0.1; **p<0.05; ***p<0.01. "Post-test self-reported SR", "Pre-test time", and "Pre-test SR" were logged due to their right-skewed distributions. "T1" stands for "treatment group 1", "T2" stands for "treatment group 2", "SR" stands for "Sharpe ratio", and "SA" stands for "savings account".

Table 8. Treatment effect on in-game SR

	Dependent variable:		
	In-game SR		
	(1)	(2)	(3)
T2	0.340***	0.308**	0.251**
	(0.125)	(0.123)	(0.118)
Pre-test score		0.601	0.454
		(0.412)	(0.393)
Pre-test time		0.098	0.019
		(0.116)	(0.111)
Pre-test trading app use_No, but		0.169	0.122
		(0.173)	(0.165)
Pre-test trading app use_Yes		0.120	0.226
		(0.207)	(0.198)
Pre-test risk attitude		−0.017	−0.007
		(0.039)	(0.037)
Pre-test SR		0.243***	0.201***
		(0.068)	(0.065)
Pre-test SA		−0.001	−0.003
		(0.003)	(0.003)
Time on instructions			0.049
			(0.047)
Time per round			0.168*
			(0.088)
Rounds played			0.088***
			(0.017)
In-game earnings			0.021
			(0.017)
Observations	305	305	305
R^2	0.024	0.092	0.189
Adjusted R^2	0.021	0.067	0.155
Residual Std. Error	1.084 (df = 303)	1.058 (df = 296)	1.007 (df = 292)

Note: *p<0.1; **p<0.05; ***p<0.01. "In-game SR", "Pre-test time", "Time on instructions", "Time per round", and "Pre-test SR" were logged due to their right-skewed distributions. "T2" stands for "treatment group 2", "SR" stands for "Sharpe ratio", and "SA" stands for "savings account".

treated students for in-game SR). Balance diagnostics in Table 10 indicate that randomization remains credible in the consented subsamples: age, gender composition, GPA, and ISEE distribution are statistically comparable across arms; a small imbalance in pre-test SA, while significant, is minor in magnitude and controlled for in regressions.

When augmenting the post-test knowledge specification with student-level demographics (Table 11), the main pattern is unchanged. T1 yields a positive but modest gain (about +0.16 SD), while T2 delivers a larger and precisely estimated improvement (about +0.44 SD). In the fully controlled specification, the estimated coefficients correspond to +0.039 for T1 and +0.075 for T2. Relative to the control-group mean in the consented sample (0.59), this implies increases of roughly 7% and 13%, respectively. Demographic and background covariates are generally muted, while baseline knowledge remains a strong predictor of post-test performance.

Results for stated risk attitudes are consistent with the full-sample analysis. Table 12 shows no statistically significant effect of either treatment on post-test self-reported risk attitudes once demographics are included. By contrast, post-test stated risk attitudes display substantial persistence, with baseline risk attitudes strongly predicting endline values. Prior exposure to trading apps is positively associated with stated risk tolerance, while other demographic and socioeconomic controls do not meaningfully shift the pattern.

Turning to revealed risk attitudes during the simulation, Table 13 indicates that students in T2 allocate a significantly larger share of wealth to risky assets than students in T1. The estimated effect is stable across speci-

Table 9. Variable definitions (consented subsample)

Variable	Description	Value/range
Individual characteristics		
Female	Whether the student is female	1; otherwise
Age	The student's age in years	from 18 to 37
ISEE_Scholarship	Whether the student received a scholarship in the past academic year	1; otherwise
ISEE_Mid-low	Whether the student's ISEE falls within brackets 1–8	1; otherwise
ISEE_Max	Whether the student's ISEE is in the highest (ninth) bracket	1; otherwise
Non-scientific high school	Whether the student obtained a diploma from a non-scientific high school	1; otherwise
High-dropout major	Whether the student is enrolled in a major whose dropout rate exceeds the median	1; otherwise
Pre-test (Post-test) trading app use_No	Whether the student had never used a trading app and did not know what it was before (after) the intervention	1; otherwise
Pre-test (Post-test) trading app use_No, but	Whether the student had never used a trading app but knew what it was before (after) the intervention	1; otherwise
Pre-test (Post-test) trading app use_Yes	Whether the student had used a trading app before (after) the intervention	1; otherwise
Academic performance		
GPA	The student's GPA on a 30-point Italian scale	from 18 to 30
Financial knowledge		
Pre-test score	The student's financial knowledge score before the intervention	from 0.11 to 0.88
Post-test score	The student's financial knowledge score after the intervention	from 0.11 to 1.00
Financial attitudes		
Pre-test (Post-test) risk attitude	The student's self-reported risk attitude before (after) the intervention	from 2 to 10
Financial behaviors		
Pre-test return	Portfolio return based on the pre-test asset allocation	from 0.00625 to 0.10
Post-test return	Portfolio return based on the post-test asset allocation	from 0.0075 to 0.10
Pre-test volatility	Portfolio standard deviation based on the pre-test asset allocation	from 0.00375 to 1.40
Post-test volatility	Portfolio standard deviation based on the post-test asset allocation	from 0.0025 to 1.40
Pre-test SR	Portfolio Sharpe ratio based on the pre-test asset allocation	from 0.0714 to 6.20
Post-test SR	Portfolio Sharpe ratio based on the post-test asset allocation	from 0.0714 to 21.00
Pre-test SA	Percentage of the portfolio allocated to the savings account in the pre-test	from 0 to 0.9
Post-test SA	Percentage of the portfolio allocated to the savings account in the post-test	from 0 to 0.85
Experimental game data		
Pre-test time	Time (in seconds) the student took to complete the pre-test questionnaire	from 180 to 19,920
Time on instructions	Time (in seconds) spent reading the game instructions	from 5 to 3,227
Time per round	Average time (in seconds) spent per round during the simulation	from 12.65 to 506.80
Rounds played	Total number of rounds completed during the simulation	from 2 to 20
In-game earnings	Total earnings obtained during the simulation (in euros)	from 0 to 21.43

Table 10. Descriptive statistics (consented subsample)

Variable	C	T1	T2	p-value
Female	0.25 (0.43)	0.27 (0.44)	0.28 (0.45)	0.8
Age	20.42 (1.48)	20.79 (2.29)	20.55 (1.53)	0.7
GPA	25.11 (2.75)	24.92 (2.81)	24.85 (2.92)	0.7
ISEE				0.032
Scholarship	36 (35%)	22 (21%)	23 (21%)	
Mid-low	33 (32%)	37 (35%)	50 (45%)	
Max	33 (32%)	46 (44%)	37 (34%)	
Non-scientific high school	0.31 (0.47)	0.21 (0.41)	0.25 (0.43)	0.2
High-dropout major	0.44 (0.50)	0.39 (0.49)	0.36 (0.48)	0.5
Pre-test score	0.59 (0.15)	0.56 (0.18)	0.58 (0.16)	0.5
Post-test score	0.59 (0.15)	0.62 (0.16)	0.66 (0.16)	0.009
Pre-test risk attitude	5.76 (1.64)	5.81 (1.76)	5.74 (1.60)	> 0.9
Post-test risk attitude	5.99 (1.58)	6.00 (1.55)	5.79 (1.53)	0.5
Pre-test return	0.049 (0.014)	0.046 (0.014)	0.044 (0.015)	0.007
Pre-test volatility	0.24 (0.17)	0.23 (0.18)	0.21 (0.19)	0.3
Pre-test SR	0.44 (0.76)	0.58 (1.12)	0.58 (1.11)	0.9
Post-test return	0.051 (0.014)	0.046 (0.014)	0.044 (0.014)	0.009
Post-test volatility	0.28 (0.21)	0.22 (0.19)	0.19 (0.19)	0.004
Post-test SR	0.45 (0.87)	0.90 (2.35)	1.21 (3.20)	0.10
Pre-test SA	0.19 (0.20)	0.24 (0.20)	0.27 (0.21)	0.005
Post-test SA	0.17 (0.17)	0.24 (0.21)	0.24 (0.22)	0.038
Pre-test time (s)	1,864 (2,990)	1,401 (976)	1,665 (1,826)	0.7
Pre-test trading app use				0.7
No	23 (23%)	21 (20%)	21 (19%)	
No, but	52 (51%)	54 (51%)	65 (59%)	
Yes	27 (26%)	30 (29%)	24 (22%)	
Post-test trading app use				0.002
No	16 (16%)	3 (2.9%)	5 (4.5%)	
No, but	60 (59%)	64 (61%)	77 (70%)	
Yes	26 (25%)	38 (36%)	28 (25%)	
N	102	105	110	

Note: Continuous variables are reported as means and standard deviations. Dummy variables are reported as frequencies and percentages. Post-test variables are highlighted in light gray. "SR" stands for "Sharpe ratio", "SA" stands for "savings account", "No, but" corresponds to the answer "no, but I know what it is" to the question "Have you ever used a trading platform?", while "No" corresponds to "no, and I have no idea what it is". "ISEE" refers to the socioeconomic status (via parental ISEE brackets). The p-values for continuous variables come from a Kruskal-Wallis rank sum test, whereas the p-values for categorical variables come from a Pearson's Chi-squared test.

Table 11. Treatment effect on post-test score (consented subsample)

	Dependent variable:	
	Post-test score	
	(1)	(2)
T1	0.024	0.039**
	(0.022)	(0.020)
T2	0.068***	0.075***
	(0.022)	(0.020)
Female		0.010
		(0.020)
Age		0.002
		(0.004)
GPA		0.004
		(0.003)
ISEE_Scholarship		−0.019
		(0.021)
ISEE_Mid-low		−0.005
		(0.018)
Non-scientific HS		−0.015
		(0.019)
High-dropout major		0.011
		(0.017)
Pre-test score		0.451***
		(0.051)
Pre-test time		0.011
		(0.013)
Pre-test trading app use_No, but		0.021
		(0.021)
Pre-test trading app use_Yes		0.011
		(0.026)
Pre-test risk attitude		−0.005
		(0.005)
Pre-test SR		0.009
		(0.009)
Pre-test SA		−0.095**
		(0.041)
Observations	317	317
R^2	0.031	0.313
Adjusted R^2	0.025	0.276
Residual Std. Error	0.158 (df = 314)	0.137 (df = 300)

Note: *p<0.1; **p<0.05; ***p<0.01. "Pre-test time" and "Pre-test SR" were logged due to their right-skewed distributions. "T1" stands for "treatment group 1", "T2" stands for "treatment group 2", "SR" stands for "Sharpe ratio", and "SA" stands for "savings account".

fications and amounts to about +0.056 in the model with demographic and baseline controls. Academic performance is positively related to risk-taking in the game, while students in high-dropout majors allocate a lower risky share. Baseline savings account allocations are strongly negatively associated with in-game risk-taking, consistent with persistence in safe-versus-risky investment orientation. When adding in-game process measures, the treatment effect remains, and the number of rounds played is weakly negatively related to the risky share, suggesting that participants who stay longer in the task may not necessarily do so by systematically increasing risk exposure.

For self-reported portfolio efficiency (Table 14), treatment effects remain positive and economically meaningful after controlling for student demographics. In the fully specified model, the estimated coefficients are approximately +0.274 for T1 and +0.358 for T2, indicating that the augmented intervention continues to outperform the baseline simulation also within the consented subsample. Baseline portfolio efficiency is a strong predictor of post-test SR, consistent with persistence in allocation skills. By contrast, demographic covariates do not display systematic or robust associations with self-reported portfolio efficiency.

For realized portfolio efficiency during the trading simulation (Table 15), the T2 premium persists and remains statistically significant across specifications. In the model with baseline and demographic controls, the estimated T2 coefficient is about +0.377, and it attenuates to roughly +0.341 when in-game process measures are added. Engagement remains strongly related to performance: the number of rounds completed is again a robust positive correlate of in-game SR, while most demographic controls are small

Table 12. Treatment effect on post-test risk attitude (consented subsample)

	Dependent variable:	
	Post-test risk attitude	
	(1)	(2)
T1	0.010	0.012
	(0.216)	(0.168)
T2	−0.199	−0.106
	(0.213)	(0.167)
Female		−0.136
		(0.174)
Age		0.015
		(0.037)
GPA		−0.001
		(0.024)
ISEE_Scholarship		0.158
		(0.179)
ISEE_Mid-low		−0.134
		(0.155)
Non-scientific HS		−0.113
		(0.159)
High-dropout major		−0.064
		(0.141)
Pre-test score		−0.623
		(0.434)
Pre-test time		0.237**
		(0.113)
Pre-test trading app use_No, but		0.250
		(0.179)
Pre-test trading app use_Yes		0.460**
		(0.219)
Pre-test risk attitude		0.569***
		(0.043)
Pre-test SR		−0.065
		(0.077)
Pre-test SA		−0.506
		(0.349)
Observations	317	317
R^2	0.004	0.463
Adjusted R^2	−0.002	0.434
Residual Std. Error	1.553 (df = 314)	1.167 (df = 300)

Note: *p<0.1; **p<0.05; ***p<0.01. "Pre-test time" and "Pre-test SR" were logged due to their right-skewed distributions. "T1" stands for "treatment group 1", "T2" stands for "treatment group 2", "SR" stands for "Sharpe ratio", and "SA" stands for "savings account".

Table 13. Treatment effect on in-game risk attitude (consented subsample)

	Dependent variable:		
	In-game risk attitude		
	(1)	(2)	(3)
T2	0.040	0.056**	0.057**
	(0.030)	(0.026)	(0.026)
Female		0.023	0.020
		(0.034)	(0.034)
Age		−0.001	0.0003
		(0.007)	(0.007)
GPA		0.010**	0.009*
		(0.005)	(0.005)
ISEE_Scholarship		0.010	0.012
		(0.036)	(0.036)
ISEE_Mid-low		0.0002	0.0002
		(0.037)	(0.037)
Non-scientific HS		0.024	0.021
		(0.033)	(0.033)
High-dropout major		−0.068**	−0.065**
		(0.028)	(0.028)
Pre-test score		0.023	0.032
		(0.083)	(0.083)
Pre-test time		−0.003	−0.001
		(0.023)	(0.023)
Pre-test trading app use_No, but		0.080**	0.084**
		(0.036)	(0.036)
Pre-test trading app use_Yes		0.056	0.052
		(0.044)	(0.045)
Pre-test risk attitude		−0.001	−0.001
		(0.008)	(0.008)
Pre-test SR		0.011	0.014
		(0.014)	(0.015)
Pre-test SA		−0.538***	−0.527***
		(0.066)	(0.067)
Time on instructions			−0.002
			(0.010)
Time per round			0.0004
			(0.020)
Rounds played			−0.007*
			(0.004)
In-game earnings			0.002
			(0.004)
Observations	215	215	215
R^2	0.008	0.348	0.361
Adjusted R^2	0.004	0.298	0.297
Residual Std. Error	0.220 (df = 213)	0.185 (df = 199)	0.185 (df = 195)

Note: *p<0.1; **p<0.05; ***p<0.01. "In-game SR", "Pre-test time", "Time on instructions", "Time per round", and "Pre-test SR" were logged due to their right-skewed distributions. "T2" stands for "treatment group 2", "SR" stands for "Sharpe ratio", and "SA" stands for "savings account".

Table 14. Treatment effect on post-test self-reported SR (consented subsample)

	Dependent variable:	
	Post-test self-reported SR	
	(1)	(2)
T1	0.271*	0.274*
	(0.153)	(0.145)
T2	0.388**	0.358**
	(0.151)	(0.144)
Female		−0.200
		(0.150)
Age		−0.011
		(0.032)
GPA		−0.005
		(0.021)
ISEE_Scholarship		−0.0004
		(0.155)
ISEE_Mid-low		0.008
		(0.134)
Non-scientific HS		0.173
		(0.137)
High-dropout Major		−0.152
		(0.122)
Pre-test score		0.411
		(0.375)
Pre-test time		0.028
		(0.098)
Pre-test trading app_No, but		0.275*
		(0.155)
Pre-test trading app_Yes		0.103
		(0.189)
Pre-test return		−0.055
		(0.037)
Pre-test SR		0.454***
		(0.066)
Pre-test SA		−0.316
		(0.301)
Observations	317	317
R^2	0.021	0.217
Adjusted R^2	0.015	0.175
Residual Std. Error	1.102 (df = 314)	1.008 (df = 300)

Note: *p<0.1; **p<0.05; ***p<0.01. "Post-test self reported SR", "Pre-test time", and "Pre-test SR" were logged due to their right-skewed distributions. "T1" stands for "treatment group 1", "T2" stands for "treatment group 2", "SR" stands for "Sharpe ratio", and "SA" stands for "savings account".

and imprecisely estimated. As in the full sample, baseline portfolio efficiency predicts realized performance, suggesting that the intervention builds on pre-existing heterogeneity in allocation ability but does not depend on a specific demographic subgroup.

To further assess whether the intervention's effects vary across individuals, we explored treatment heterogeneity along three dimensions that have been consistently linked to financial learning and behavioral outcomes: gender, academic performance, and socioeconomic status (SES). These moderators are among the most frequently examined in empirical FE studies, reflecting persistent disparities in baseline financial literacy and potential differences in responsiveness to instruction. Nonetheless, we label these analyses as exploratory (see Transparency Appendix).

Existing evidence suggests that gender gaps in financial literacy emerge early and can persist despite instruction, with females often showing lower baseline knowledge and weaker behavioral improvements Kalwij et al. (2019); Sconti (2022); Torma et al. (2023). Academic performance is also commonly associated with learning and retention, although moderation by achievement is not consistently observed across settings (Becchetti et al., 2013; Kalmi and Rahko, 2022; Torma et al., 2023). Finally, SES — proxied by parental resources such as education or income — may shape both exposure to financial concepts and opportunities to apply them; accordingly, prior evaluations report mixed evidence, with some finding larger gains among low-SES participants and others finding muted or null heterogeneity (Amagir et al., 2022; Bover et al., 2024; Iterbeke et al., 2022; Sconti, 2022). Building on this literature, we focus on a parsimonious set of interactions between treatment as-

Table 15. Treatment effect on in-game SR (consented subsample)

	Dependent variable:		
	In-game SR		
	(1)	(2)	(3)
T2	0.397**	0.377**	0.341**
	(0.153)	(0.153)	(0.148)
Female		0.165	0.178
		(0.198)	(0.190)
Age		0.005	−0.006
		(0.040)	(0.039)
GPA		−0.038	−0.028
		(0.027)	(0.026)
ISEE_Scholarship		−0.019	−0.073
		(0.210)	(0.202)
ISEE_Mid-low		−0.078	−0.093
		(0.214)	(0.206)
Non-scientific HS		−0.316	−0.245
		(0.193)	(0.185)
High-dropout Major		−0.242	−0.224
		(0.163)	(0.157)
Pre-test score		0.472	0.266
		(0.485)	(0.466)
Pre-test time		0.064	0.014
		(0.135)	(0.131)
Pre-test trading app use_No, but		0.196	0.130
		(0.212)	(0.204)
Pre-test trading app use_Yes		−0.040	0.083
		(0.258)	(0.249)
Pre-test risk attitude		0.013	0.013
		(0.049)	(0.047)
Pre-test SR		0.250***	0.212***
		(0.084)	(0.081)
Pre-test SA		−0.402	−0.574
		(0.387)	(0.371)
Time on instructions			0.054
			(0.058)
Time per round			0.180
			(0.114)
Rounds played			0.092***
			(0.022)
In-game earnings			0.029
			(0.023)
Observations	215	215	215
R^2	0.031	0.148	0.241
Adjusted R^2	0.026	0.083	0.166
Residual Std. Error	1.115 (df = 213)	1.082 (df = 199)	1.032 (df = 195)

Note: *p<0.1; **p<0.05; ***p<0.01. "Post-test self-reported SR", "Pre-test time", "Time on instructions", "Time per round", and "Pre-test SR" were logged due to their right-skewed distributions. "T1" stands for "treatment group 1", "T2" stands for "treatment group 2", "SR" stands for "Sharpe ratio", and "SA" stands for "savings account".

signment and three pre-treatment characteristics — female, high GPA, and ISEE — to test whether T1 and T2 yield systematically different impacts across groups.

Across outcomes capturing financial knowledge, financial attitudes, and financial behaviors, we do not find systematic or robust heterogeneity in treatment effects. Interaction estimates are generally small, do not follow a consistent pattern across measures, and are sensitive to specification. Taken together, these results suggest that the improvements documented in the main effects are broadly shared across the student population rather than concentrated in specific demographic or academic subgroups. Given the exploratory nature of this analysis and the limited power to detect interaction effects, we do not emphasize subgroup-specific findings in the main text. All heterogeneity specifications will be reported in the Appendix in a forthcoming version of the paper.

6 Conclusion

This paper studies the effects of two short, simulation-based FE interventions that differ in the degree of behavioral scaffolding embedded in the learning environment. Using a randomized design, we show that both interventions improve students' financial knowledge and portfolio choices relative to a no-intervention control group, while the augmented simulation consistently outperforms the baseline version.

In terms of financial knowledge, both treatment arms generate statistically significant gains. Relative to the control group mean, estimated effects

correspond to increases of roughly 8% for the baseline simulation and 15–16% for the augmented version, with these differences remaining stable when conditioning on pre-treatment characteristics. These results confirm that even brief, app-like interventions can deliver meaningful learning gains.

Turning to financial attitudes, we find no evidence that either intervention affects self-reported risk preferences as measured by the standard survey item. In contrast, revealed risk attitudes observed during the trading simulation do respond to treatment intensity. Students assigned to the augmented simulation allocate a higher share of their portfolio to risky assets than those in the baseline treatment, with estimated effects on the order of 5–6% relative to the baseline mean. The resulting gap between stated and revealed measures indicates that the intervention primarily shapes how students act on their preferences rather than how they report them. This interpretation aligns with evidence showing that self-reported risk attitudes tend to be stable and sensitive to elicitation, whereas behavior in incentivized and context-rich environments is more responsive to changes in decision-making frameworks and informational cues (Chapman et al., 2025; Falk et al., 2023).

The strongest effects emerge for financial behaviors captured through portfolio efficiency. Both treatments substantially increase the Sharpe ratio computed from post-test declared allocations, with gains of approximately 27% for the baseline simulation and over 36% for the augmented version relative to the control group. Differences between the two treatments become even more pronounced when focusing on realized behavior during the trading game. Students exposed to the augmented simulation achieve in-game Sharpe ratios that are about 35% higher than those of students in the base-

line treatment, indicating a sizeable improvement in payoff-relevant decision quality.

Importantly, augmenting the analysis with detailed log data from the simulation shows that these differences are partly mediated by engagement. Measures such as the number of rounds played and time spent per round are strongly associated with realized portfolio efficiency, while treatment coefficients attenuate once these process variables are included. This pattern suggests that the augmented intervention improves outcomes not by changing the return-generating environment, but by fostering more sustained and effective interaction with it.

Overall, the findings contribute to the FE literature in three ways. First, they provide causal evidence that a fast, light-touch trading simulation can generate sizable gains in financial knowledge and portfolio efficiency. Second, they show that short, behaviorally informed prompts embedded in a realistic decision environment can further improve realized investment performance, even in the absence of changes in stated risk preferences. Third, by combining self-reported intentions with realized, incentivized behavior under a common return-generating structure, the study offers a tighter link between financial competence and actual decision-making than is typical in FE evaluations.

Several limitations point to directions for future research. External validity is bounded by a single institutional setting and a relatively homogeneous sample. Outcomes are measured over a short horizon, preventing conclusions about persistence. Attrition differs across experimental arms due to differential participation requirements, and while baseline competence does not predict dropout, selection on unobservables cannot be ruled out. Future

work should examine longer-run effects, replicate the intervention in more diverse populations, and investigate mechanisms more directly, including how attention and engagement interact with behavioral prompts over time.

Overall, the evidence suggests that well-designed, behaviorally informed simulations can do more than transmit financial concepts: they can meaningfully improve how novice investors allocate risk in environments that increasingly shape everyday financial decision-making.

7 Transparency Appendix

To facilitate transparency and replicability en lieu of a pre-registration, this appendix documents and discusses details regarding the choices made by the researchers the experimental design and outcome measures.

7.1 General Set-up of the Experiment

The core idea of this study is straightforward in that we evaluate two FE interventions in comparison to a control group that does not receive an intervention: the baseline intervention consists of a novel simulation game, the augmented intervention integrates educational pop-ups into the simulation game. As is typical when comparing a basic intervention to a more comprehensive one, we expect both interventions to have positive but the comprehensive intervention — i.e., the simulation training with educational pop-ups — to have a larger effect. Otherwise there would have been no reason to augment the baseline treatment in the first place, since the treatments are identical in terms of scalability, costs, and nearly identical with respect

to duration.

7.2 Sample and Attrition

A further transparency-relevant aspect concerns study completion and attrition. After the pre-questionnaire, 704 students were randomized, while 485 completed the full study protocol and constitute the main analytic sample. Attrition occurred after random assignment and differed across experimental arms: dropout rates amount to 23.1% in the control group, 35.3% in T1, and 34.9% in T2 (Table 16, Panel A). A Pearson's chi-squared test rejects equality of dropout rates across groups ($p = 0.005$), reflecting the fact that treated students faced higher participation requirements due to the simulation task.

To assess whether attrition is systematically related to baseline competence, we compare pre-test financial knowledge between completers and dropouts. As shown in Panel B of Table 16, baseline knowledge is statistically indistinguishable across the two groups (Welch t-test, $p = 0.48$). This result is corroborated by a multivariate logit model reported in Panel C, where dropout is regressed on treatment assignment and pre-test financial knowledge: while assignment to T1 and T2 significantly increases the probability of dropout, baseline financial knowledge does not predict attrition.

Taken together, these results indicate that attrition is differential across experimental arms but not systematically related to pre-treatment financial knowledge. While estimates throughout the paper are therefore conditional on study completion, the absence of selection on observed baseline competence mitigates concerns that treatment effects are mechanically driven by differential attrition.

Table 16. Attrition analysis

	C	**T1**	**T2**
Panel A: Dropout rates			
Number of students (pre-test)	234	235	235
Number of dropouts	54	83	82
Dropout rate (%)	23.1	35.3	34.9
Pearson's χ^2 test (p-value)		0.005	
Panel B: Baseline financial knowledge			
Pre-test score (non-dropouts)		0.577 (0.12)	
Pre-test score (dropouts)		0.584 (0.13)	
Welch two-sample t-test (p-value)		0.48	
Panel C: Logit model for dropout			
Treatment group T1		0.600**	
		(0.207)	
Treatment group T2		0.579**	
		(0.207)	
Pre-test score		0.337	
		(0.548)	
Observations		704	

Note: **p<0.05. Panel A reports dropout rates by experimental arm; the reported p-value refers to a Pearson's chi-squared test of equality of dropout rates across groups. Panel B reports mean pre-test financial knowledge scores (standard deviations in parentheses) for students who completed the study and those who dropped out. Panel C reports coefficients from a logit regression where the dependent variable equals one if the student dropped out (robust standard errors in parentheses).

7.3 Outcomes and Specifications

In terms of outcomes, we use three different types of outcomes: financial knowledge test scores, and financial decision-making outcomes captured through both self-reported and incentivized portfolio-based measures. Intended behavior is measured via a portfolio allocation task that mirrors the structure of the experimental game. Based on declared allocations, we compute a design-based Sharpe ratio that reflects ex-ante portfolio efficiency, with expected returns and risk imputed from the game's known payoff structure rather than estimated from data. Behavior-based outcomes are derived from actual allocation choices made during the simulation. Using these observed in-game allocations, we compute the same design-based Sharpe ratio, relying on the identical return vector and variance–covariance matrix. As a result, self-reported and in-game measures are directly comparable, differing only in the source of portfolio weights (declared intentions versus realized behavior), not in the underlying return-generating process. All measures are well established and follow the standard of the FE literature. The contribution of this paper lies within the design and evaluation of the intervention, not the choice of outcome measure. As self-reported and incentivized outcomes are available for comparisons between distinct groups, we do not specify one main outcome category: for comparisons with the control group, we have to rely on self-reported outcomes, as the incentivized outcomes are not available for the control group by definition. However, whenever we compare the two different treatment groups, we can use the incentivized measures in addition, since both groups played the payoff-relevant simulation game. This choice of outcome variables was predetermined by the experimental design.

Exploratively, we also look at in-game risk attitudes measured by the share of funds invested in risky assets.

In terms of covariates, we choose covariates that improve the estimation by adding explanatory power, or that are of general interest in the literature on financial decision-making. However, we are aware that in a clean experimental design, covariates are not a requirement, but may rather raise concerns due to the degree of freedom they afford the researchers. Crucially, all results hold when we remove all control variables — each regression table in the paper provides the respective specification without controls.

In addition, Table 2 and Table 9 list all survey measures that were elicited for full transparency. Section 5.2 discusses treatment effect heterogeneity along selected individual characteristics. These analyses are explicitly exploratory since they beyond the simple comparison of treatment group averages, and for some of these heterogeneities there is no strong ex-ante hypothesis for why the treatment(s) should have stronger effects for certain subgroups of students. A dedicated appendix reporting the full set of heterogeneity specifications will be included in a subsequent version of the manuscript, in line with the discussion in the main text.

References

Abreu, M. and Mendes, V. (2010). Financial literacy and portfolio diversification. *Quantitative finance*, 10(5):515–528.

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

Amagir, A., van den Brink, H. M., Groot, W., and Wilschut, A. (2022). Savewise: The impact of a real-life financial education program for ninth grade students in the netherlands. *Journal of Behavioral and Experimental Finance*, 33:100605.

Amari, M., Salhi, B., and Jarboui, A. (2020). Evaluating the effects of sociodemographic characteristics and financial education on saving behavior. *International Journal of Sociology and Social Policy*, 40(11/12):1423–1438.

Ambuehl, S., Bernheim, B. D., and Lusardi, A. (2014). *Effect of financial education on the quality of decision making*, volume 20618. National Bureau of Economic Research.

Aren, S. and Zengin, A. N. (2016). Influence of financial literacy and risk perception on choice of investment. *Procedia-Social and Behavioral Sciences*, 235:656–663.

Aucejo, E. M. and Wong, K. (2025). The effect of feedback on student performance. *Journal of Public Economics*, 241:105274.

Bajo, E., Barbi, M., and Sandri, S. (2015). Financial literacy, households' investment behavior, and risk propensity. *Journal of Financial Management, Markets and Institutions*, 3(1):157–174.

Bayar, Y., Sezgin, H. F., Öztürk, Ö. F., and Şaşmaz, M. Ü. (2020). Financial literacy and financial risk tolerance of individual investors: Multinomial logistic regression approach. *Sage Open*, 10(3):2158244020945717.

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

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2):9–49.

Benartzi, S. and Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American economic review*, 91(1):79–98.

Bover, O., Hospido, L., and Villanueva, E. (2024). The impact of high school financial education on financial knowledge and saving choices: Evidence from a randomized trial in spain. *Journal of Human Resources*.

Brade, R., Himmler, O., and Jäckle, R. (2022). Relative performance feedback and the effects of being above average—field experiment and replication. *Economics of Education Review*, 89:102268.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2009). Measuring the financial sophistication of households. *American economic review*, 99(2):393–398.

Carpena, F., Cole, S., Shapiro, J., and Zia, B. (2019). The abcs of financial education: Experimental evidence on attitudes, behavior, and cognitive biases. *Management Science*, 65(1):346–369.

Carpena, F. and Zia, B. (2020). The causal mechanism of financial education: Evidence from mediation analysis. *Journal of Economic Behavior &*

Organization, 177:143–184.

Cavezzali, E., Gardenal, G., and Rigoni, U. (2015). Risk taking behaviour and diversification strategies: Do financial literacy and financial education play a role? *Journal of Financial Management, Markets and Institutions*, 3(1):121–156.

Chapkovski, P., Khapko, M., and Zoican, M. (2024). Trading gamification and investor behavior. *Management Science*.

Chapman, J., Ortoleva, P., Snowberg, E., Yariv, L., and Camerer, C. (2025). Reassessing qualitative self-assessments and experimental validation. Technical report, National Bureau of Economic Research.

Chiappori, P.-A. and Paiella, M. (2011). Relative risk aversion is constant: Evidence from panel data. *Journal of the European Economic Association*, 9(6):1021–1052.

Dinç Aydemir, S. and Aren, S. (2017). Do the effects of individual factors on financial risk-taking behavior diversify with financial literacy? *Kybernetes*, 46(10):1706–1734.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the european economic association*, 9(3):522–550.

EFAMA (2020). Household participation in capital markets.

Falk, A., Becker, A., Dohmen, T., Huffman, D., and Sunde, U. (2023). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *Management Science*, 69(4):1935–1950.

Gomes, F. and Michaelides, A. (2005). Optimal life-cycle asset alloca-

tion: Understanding the empirical evidence. *The Journal of Finance*, 60(2):869–904.

Guiso, L. and Jappelli, T. (2009). Financial Literacy and Portfolio Diversification. CSEF Working Papers 212, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy.

Guiso, L. and Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic association*, 6(6):1109–1150.

Hermansson, C. and Jonsson, S. (2021). The impact of financial literacy and financial interest on risk tolerance. *Journal of Behavioral and Experimental Finance*, 29:100450.

Horn, S., Jamison, J. C., Karlan, D., and Zinman, J. (2023). Five-year impacts of group-based financial education and savings promotion for ugandan youth. *Review of Economics and Statistics*, pages 1–53.

Huberman, G. and Jiang, W. (2006). Offering versus choice in 401 (k) plans: Equity exposure and number of funds. *The Journal of Finance*, 61(2):763–801.

Iterbeke, K., Schelfhout, W., and De Witte, K. (2022). The role of students' interests during computer-assisted learning: A field experiment. *Computers in Human Behavior*, 130:107168.

Kaiser, T. and Menkhoff, L. (2020). Financial education in schools: A meta-analysis of experimental studies. *Economics of education review*, 78:101930.

Kalmi, P. and Rahko, J. (2022). The effects of game-based financial education: New survey evidence from lower-secondary school students in finland. *The Journal of economic educaTion*, 53(2):109–125.

Kalwij, A., Alessie, R., Dinkova, M., Schonewille, G., Van der Schors, A., and Van der Werf, M. (2019). The effects of financial education on financial literacy and savings behavior: Evidence from a controlled field experiment in dutch primary schools. *Journal of Consumer Affairs*, 53(3):699–730.

Linciano, N., Costa, D., Gentile, M., and Soccorso, P. (2019). Report on financial investments of italian households. *Behavioural Attitudes and Approaches 2019 Survey, Consob Statistics and analyses, Rome*.

Litterscheidt, R. and Streich, D. J. (2020). Financial education and digital asset management: What's in the black box? *Journal of Behavioral and Experimental Economics*, 87:101573.

Lobão, J. (2020). What to do in financial markets? preferences and incoherences of future investors. *Corvinus Journal of Sociology and Social Policy*, 11(2):3–22.

Lönnqvist, J.-E., Verkasalo, M., Walkowitz, G., and Wichardt, P. C. (2015). Measuring individual risk attitudes in the lab: Task or ask? an empirical comparison. *Journal of Economic Behavior & Organization*, 119:254–266.

Mouna, A. and Jarboui, A. (2015). Financial literacy and portfolio diversification: an observation from the tunisian stock market. *International Journal of Bank Marketing*, 33(6):808–822.

Muñoz-Murillo, M., Álvarez-Franco, P. B., and Restrepo-Tobón, D. A. (2020). The role of cognitive abilities on financial literacy: New experimental evidence. *Journal of Behavioral and Experimental Economics*, 84:101482.

Muralidharan, K., Singh, A., and Ganimian, A. J. (2019). Disrupting education? experimental evidence on technology-aided instruction in india. *American Economic Review*, 109(4):1426–60.

Sconti, A. (2022). Digital vs. in-person financial education: What works best for generation z? *Journal of Economic Behavior & Organization*, 194:300–318.

Shefrin, H. and Statman, M. (2000). Behavioral portfolio theory. *Journal of financial and quantitative analysis*, 35(2):127–151.

Torma, J., Barbić, D., and Ivanov, M. (2023). Analyzing the effects of financial education on financial literacy and financial behaviour: A randomized field experiment in croatia. *The South East European Journal of Economics and Business*, 18(2):63–86.

Ungeheuer, M. and Weber, M. (2021). The perception of dependence, investment decisions, and stock prices. *The Journal of Finance*, 76(2):797–844.

Van Klaveren, C., Vonk, S., and Cornelisz, I. (2017). The effect of adaptive versus static practicing on student learning-evidence from a randomized field experiment. *Economics of Education Review*, 58:175–187.

West, T. and Cull, M. (2020). Future expectations and financial satisfaction. *Economic Papers: A journal of applied economics and policy*, 39(4):318–335.

Yoong, J. (2011). Financial illiteracy and stock market participation: Evidence from the rand american life panel. *Financial literacy: Implications for retirement security and the financial marketplace*, 76:39.