How does financial literacy – or lack thereof – influence risk-taking behaviour, overconfidence and portfolio construction among Norwegian investors?

Master Thesis

by

Henrik Bekkevold and Valiant Evers

MSc in Finance

BI Norwegian Business School

Oslo, June 2026

Disclaimer: This thesis is a part of the MSc programme at BI Norwegian Business School. The school takes no responsibility for the methods used, results found, or conclusions drawn.

Acknowledgements

We would like to express our sincere gratitude to our supervisor, Ignacio. His constructive feedback, academic guidance, and valuable insights provided us with support throughout the thesis process. His comments significantly strengthened both the theoretical framing and empirical strategy of this study.

We are also grateful to the investor communities and respondents who participated in the survey. Their time and engagement made this research possible.

Finally, we thank our family and friends for their support and encouragement during the completion of this thesis.

Abstract

This thesis examines the link between financial literacy, risk-taking, overconfidence, and portfolio construction among Norwegian retail investors. Using survey data from 746 investors, we measure objective literacy, subjective confidence, and various behavioural and portfolio outcomes. We apply multivariate regression and mediation analysis to assess direct and indirect relationships.

We find that higher financial literacy is linked with wider portfolio diversification and greater risk tolerance. It also improves calibration, shown by a smaller gap between the confidence and actual knowledge, though this depends on how overconfidence is measured. Mediation analysis shows that literacy directly increases risk-taking via competence but indirectly reduces it by lowering overconfidence. Improved calibration partly offsets the direct effect, so the overall link between literacy and risk-taking remains positive. However, this result is not robust in all checks, and mediation analysis shows, rather than confirms, the proposed causal path.

Overall, the evidence suggests that financial literacy is associated with better-informed and better-calibrated risk-taking, rather than confidence-driven excess risk.

Contents

Acknowledgements **1**

Abstract **2**

1 Introduction and motivation **6**

1.1 Research relevance 6
1.2 Research question 7
1.3 Empirical approach and scope 7
1.4 Contribution and roadmap 8

2 Literature review **9**

2.1 What we know and why it matters here 9
2.2 Financial literacy and portfolio choice 9
2.3 Overconfidence and investor behaviour 10
2.4 Calibration: the gap between confidence and knowledge 11
2.5 What this thesis adds 11
2.6 Identification challenges and empirical implications 12
2.6.1 Reverse Causality 12
2.6.2 Omitted Variable Bias 12
2.6.3 Selection Bias 12
2.6.4 Scope of Claims and Interpretation 12

3 Methodology and testable hypotheses **13**

3.1 Theoretical Framework 13
3.1.1 Financial Literacy as Information Precision 13
3.1.2 Overconfidence as Belief Distortion 13
3.1.3 Interaction between Literacy and Overconfidence 14
3.2 Econometric Model 14
3.2.1 Baseline Specification 14
3.2.2 Mediation Model 15
3.3 Testable hypotheses 16
3.4 Endogeneity and Limitations 17
3.4.1 Reverse Causality 17
3.4.2 Omitted Variable Bias 18
3.4.3 Measurement Error 18
3.4.4 Mechanical Correlation in Gap-Based Overconfidence Measures 18
3.4.5 Mediation and Sequential Exogeneity 18
3.4.6 External Validity 18

3.4.7 Summary of Interpretation 19

4 Data **19**

4.1 Data source and Recruitment 19

4.2 Data Cleaning and Sample Restrictions 19

4.2.1 Initial processing 19

4.2.2 Category harmonisation 20

4.2.3 Treatment of missing values and outliers 20

4.2.4 Participation variable 20

4.3 Variable construction 20

4.3.1 Objective financial literacy 20

4.3.2 Subjective confidence and overconfidence 21

4.3.3 Risk-taking behaviour 22

4.3.4 Portfolio construction outcomes 22

4.3.5 Control variables 23

4.4 Descriptive statistics and correlations 24

4.5 Limitations 25

5 Results and analysis **25**

5.1 Sample Characteristics and Preliminary Evidence 25

5.2 Financial literacy and diversification (H2) 26

5.3 Financial literacy and risk-taking (H3) 27

5.4 Financial literacy and overconfidence (H4) 27

5.5 Overconfidence and risk-taking (H5) 28

5.6 Mediation Analysis: Quantifying the Dual-Channel Structure (H6) . . . 28

5.7 Overall Interpretation of Findings 31

5.8 Robustness Checks 31

5.8.1 Subgroup Analyses 32

5.8.2 Ceiling Effects Sensitivity 32

5.8.3 Outlier Sensitivity 33

5.8.4 Specification Robustness 33

5.8.5 Overall Robustness Assessment 34

6 Conclusion **34**

6.1 Main Findings 34

6.2 Contribution to the literature 35

6.3 Economic and Policy Implications 35

6.4 Limitations 36

6.5 Final Answer to the Research Question 36

Appendix A Survey Instrument **40**

Appendix B Complete Regression Tables **49**

Appendix C Robustness Check Summary **58**

Appendix D Additional Figures **59**

1 Introduction and motivation

Financial literacy plays a central role in shaping how households allocate capital, manage risk, and construct portfolios. In an investment context, understanding core concepts such as compounding, inflation, fees, and diversification is essential for translating long-run objectives into coherent portfolio choices. Although Northern Europe ranks among the highest in global literacy surveys (GFLEC, 2016), substantial within-country heterogeneity remains. Even in high-literacy environments, many individuals enter financial markets with incomplete mental models of risk and limited intuition for diversification and costs.

At the same time, retail participation in Norwegian equity markets has expanded sharply in recent years. The number of private shareholders increased markedly during and after the pandemic period, surpassing 600,000 by 2025 (Kaupr, 2024; E24, 2025). AksjeNorge (2024) characterises this period as a structural break from earlier trends, with unusually large inflows of new retail investors.

The coexistence of rising retail participation and heterogeneous financial knowledge raises an important question: how does financial literacy shape investor behaviour? In particular, does literacy influence not only portfolio structure and risk-taking, but also behavioural distortions such as overconfidence? Behavioural finance suggests that miscalibration between subjective confidence and objective knowledge can distort perceived risk and portfolio choices. Overconfidence may therefore function not merely as an outcome of interest, but as a mechanism linking literacy to investment behaviour.

1.1 Research relevance

This thesis contributes to the intersection of household finance and behavioural finance by jointly examining financial literacy, overconfidence, and portfolio outcomes within a unified empirical framework. Rather than treating literacy and behavioural biases as separate determinants of investor behaviour, the analysis evaluates whether overconfidence operates as a mechanism linking knowledge to risk-taking and portfolio structure.

From a practical perspective, the question is relevant because common retail investment mistakes correspond closely to the constructs studied here. Under-diversification, fee neglect, and risk choices misaligned with actual knowledge can all be interpreted as manifestations of informational frictions or miscalibration. If financial literacy systematically influences these outcomes, either directly or through behavioural channels, then interventions aimed at improving investor education and decision environments may enhance portfolio quality and welfare.

From a scientific perspective, Norway provides an informative setting. It is a financially developed market with high baseline participation and documented growth in retail ownership. Studying literacy effects in such an environment offers a relatively conservative test of the literacy–behaviour relationship, as variation occurs within an already investor-active population. At the same time, the thesis remains transparent about the limits of generalising beyond the sampled communities.

1.2 Research question

The thesis addresses:

> **How does financial literacy – or lack thereof – influence risk-taking behaviour, overconfidence, and portfolio construction among Norwegian investors?**

This question is operationalised through three interrelated empirical dimensions.

Risk-taking: Is objective financial literacy associated with differences in stated or elicited risk tolerance?

Calibration/overconfidence: Do more knowledgeable investors exhibit smaller gaps between subjective confidence and objective performance?

Portfolio construction: Is financial literacy associated with portfolio structure outcomes such as diversification, measured as the breadth of asset classes held, and related allocation proxies?

A central feature of the design is that overconfidence is analysed not only as an outcome of literacy, but also as a potential mediating mechanism linking knowledge to risk-taking behaviour. This allows the analysis to distinguish between competence-driven and confidence-driven channels of investor decision-making.

1.3 Empirical approach and scope

The empirical analysis is based on a structured survey of Norwegian retail investors recruited from online investor communities. The survey combines three core components: an objective financial literacy battery covering fundamental concepts such as interest, inflation, diversification, fees, and percentage reasoning; measures of subjective financial competence; and multiple behavioural outcomes capturing risk-taking and portfolio construction.

Empirically, we estimate multivariate regression models linking behavioural and portfolio outcomes to financial literacy while controlling for standard demographic characteristics, including age, gender, education, and income. Model specifications are adapted to

the nature of each dependent variable, and overconfidence is evaluated as a potential mediating channel using sequential regression models with bootstrap inference for the indirect effect.

The recruitment strategy targets active investor communities and therefore entails self-selection. The findings should consequently be interpreted as evidence on behavioural variation within an investor-active segment rather than as population-wide causal estimates for Norway as a whole. As a consequence of this sampling strategy, all respondents in the cleaned dataset report active investment or plans to invest, leaving no variation in market participation. The market participation hypothesis (H1) is therefore untestable within this sample and is excluded from the empirical analysis.

1.4 Contribution and roadmap

This thesis makes three main contributions.

First, it integrates the financial literacy and overconfidence literatures within a unified empirical framework. Rather than analysing informational frictions and behavioural biases separately, the study evaluates whether overconfidence mediates the relationship between literacy and risk-taking, thereby identifying distinct competence and calibration channels.

Second, it clearly distinguishes between objective knowledge and subjective confidence and quantifies miscalibration using both a gap-based measure and a residual-based robustness specification. This dual approach strengthens the identification of behavioural mechanisms and addresses mechanical correlation concerns present in gap-based measures.

Third, it provides evidence from a Norwegian retail-investor setting characterised by high baseline literacy and rapid growth in market participation. Studying literacy effects in such an environment offers a relatively conservative test of behavioural relationships within an investor-active population.

The remainder of the thesis is structured as follows. Chapter 2 reviews the relevant literature on financial literacy, overconfidence, and household portfolio choice. Chapter 3 develops the theoretical framework and presents testable hypotheses and econometric models. Chapter 4 describes the data, sampling frame, and variable construction. Chapter 5 presents the empirical results and mediation analysis. Chapter 6 concludes and discusses implications and avenues for future research.

2 Literature review

2.1 What we know and why it matters here

This thesis sits at the intersection of two well-developed but largely separate literatures: household finance research on financial literacy and portfolio choice, and behavioural finance research on overconfidence and investor behaviour. The central question, whether overconfidence acts as a mechanism through which literacy shapes portfolio outcomes, has received surprisingly little direct attention. This review builds the case for why such a mechanism is theoretically plausible, what the existing evidence implies about each link in the chain, and where the open questions lie.

A consistent finding across both literatures is that household portfolios deviate systematically from normative theory. Investors hold underdiversified portfolios, trade too frequently, and take on risk that does not align with their circumstances. Two broad explanations have been proposed: informational frictions (investors simply do not know enough) and belief distortions (investors know things but are systematically wrong about their own knowledge or signal quality). Financial literacy and overconfidence are the primary empirical proxies for these two mechanisms respectively, and understanding how they interact is essential for explaining observed behaviour.

2.2 Financial literacy and portfolio choice

The most direct evidence on financial literacy and portfolio outcomes comes from Van Rooij et al. (2011), who use Dutch household survey data and an instrumental variable strategy to show that financial literacy is strongly positively associated with stock market participation, even after controlling for income, wealth, and education. The authors argue that literacy reduces the fixed informational costs of entry, making participation rational for a broader set of households. Their findings have since been replicated across multiple countries and settings, establishing the literacy–participation link as one of the more robust findings in household finance.

Portfolio structure conditional on participation is examined by Guiso and Jappelli (2008), who use Italian household data to show that more financially literate investors hold more diversified portfolios and make more efficient asset allocation decisions. This result is intuitive: understanding covariance and the mechanics of diversification is a prerequisite for acting on them. Importantly, literacy effects on diversification persist after controlling for wealth and education, suggesting that knowledge carries independent explanatory power beyond socioeconomic status.

Abreu and Mendes (2012), working with Portuguese retail investor data, extend the picture by linking information processing and investor beliefs to trading behaviour.

Their findings suggest that how investors interpret and act on financial information, not just whether they have access to it, shapes outcomes. This points toward subjective confidence and belief formation as important mediating variables, bridging the literacy literature and the overconfidence literature.

Taken together, these studies establish a clear empirical pattern: literacy reduces informational frictions, improves portfolio structure, and appears to support more sophisticated financial decision-making. What they leave largely unaddressed is whether literacy also shapes risk-taking directly, and whether any such relationship operates through behavioural channels such as miscalibration.

2.3 Overconfidence and investor behaviour

The behavioural finance literature offers a parallel body of evidence showing that biased beliefs, specifically overconfidence in one's own knowledge or signal precision, are a first-order determinant of investor behaviour. Barber and Odean (2000) provide the canonical result: using a large U.S. brokerage dataset, they show that individual investors trade excessively and systematically underperform after transaction costs, a pattern they attribute to overconfidence-driven overestimation of private signal quality. The mechanism is that overconfident investors believe their information is more valuable than it is, leading them to trade when they should not.

Barber and Odean (2001) extend this framework to gender, documenting that men trade significantly more than women and achieve lower net returns. They interpret this as evidence that overconfidence is more prevalent among male investors, a finding that has proven robust across multiple replications. For the present thesis, the gender result is relevant not as a primary finding but as a motivation for including gender controls and for remaining cautious about demographic confounding in the literacy–risk relationship.

More directly related to the present thesis, Broihanne et al. (2014) examine overconfidence among finance professionals and find that miscalibration is associated with higher risk-taking and willingness to bear losses. Baek and Cho (2022) provide further evidence that overconfidence is positively associated with risky investment choices. Taken together, these results establish a robust empirical link between miscalibration and risk exposure: investors who overestimate their competence take on more risk.

Broekema and Kramer (2021) examine the intersection of overconfidence and financial advice seeking, showing that overconfident investors are less likely to seek professional advice, which in turn may amplify under-diversification. This finding is relevant here because it connects the overconfidence literature to portfolio structure outcomes, suggesting that miscalibration matters not only for risk-taking but also for how investors

construct their portfolios.

Kramer (2016) provides an important direct link between literacy and confidence, showing that financial literacy is associated with higher confidence in financial decisions. The relationship is non-linear, with moderate literacy associated with the strongest confidence effects. This finding motivates the joint analysis of literacy and calibration in the present thesis and underscores the importance of measuring whether confidence increases are proportional to actual knowledge gains or represent miscalibration.

2.4 Calibration: the gap between confidence and knowledge

A key issue in the overconfidence literature is calibration – the extent to which subjective confidence aligns with actual knowledge or performance. Well-calibrated investors accurately assess their own competence, overconfident investors overestimate it, and underconfident investors underestimate it. Calibration is distinct from the level of confidence per se: an investor can be both highly confident and well-calibrated if their objective knowledge is also high.

For the literacy–behaviour relationship, calibration is important because it determines whether higher literacy translates into accurately expanded confidence or into excess confidence. If literacy improves calibration, then more knowledgeable investors become neither overconfident nor underconfident, but better-aligned. If literacy fails to fully update confidence beliefs, some degree of systematic miscalibration persists even among knowledgeable investors.

The evidence on whether financial literacy improves calibration is limited and mixed. Kramer (2016) finds that literacy is associated with higher confidence, but the relationship is non-linear, with moderate literacy producing the strongest confidence effects. Importantly, some studies in the broader overconfidence literature suggest that domain expertise can actually increase miscalibration rather than reduce it, as experts become more certain without proportional accuracy gains (see Moore and Healy, 2008, for a review). Whether this pattern extends to financial literacy specifically is an open question. The Norwegian setting, with high average literacy and potential ceiling effects, provides a useful additional test of this mechanism.

2.5 What this thesis adds

Against this background, the present thesis makes a specific contribution by integrating the two literatures within a unified mediation framework. Rather than treating literacy and overconfidence as separate determinants of portfolio outcomes, we examine whether overconfidence operates as a mechanism linking financial knowledge to risk-taking and

portfolio structure. This framing allows the analysis to distinguish between direct competence effects and indirect calibration effects.

The key contribution is therefore not primarily to provide new evidence on any individual link, such as the literacy–participation, overconfidence–trading, or literacy–confidence relationship. Rather, it is to test whether these relationships form a coherent behavioural pathway in which literacy shapes risk-taking partly by improving calibration. While individual links in this chain have been studied extensively, we are not aware of published work that formally tests overconfidence as a mediator of the literacy–risk relationship in a retail investor sample. However, given the breadth of the household finance literature, we cannot rule out that similar frameworks have been applied in unpublished or non-English-language research.

2.6 Identification challenges and empirical implications

Several identification challenges are relevant to the existing literature and motivate the empirical strategy of this thesis.

2.6.1 Reverse Causality

Many studies using cross-sectional survey data cannot rule out that investment experience increases financial literacy rather than literacy driving portfolio choice. This concern applies to the present study as well, and is addressed through transparent reporting of limitations rather than instrumental variable strategies (see Section 3.4).

2.6.2 Omitted Variable Bias

Cognitive ability, numeracy, risk preferences, and interest in finance are correlated with both literacy and portfolio outcomes. While standard demographic controls are included, they are imperfect proxies for these underlying dimensions.

2.6.3 Selection Bias

Recruitment through investor communities implies that the sample is more financially engaged than the general population. This is a deliberate design choice that focuses the analysis on within-investor behavioural variation but limits external validity.

2.6.4 Scope of Claims and Interpretation

Given these challenges, the empirical analysis is designed to estimate conditional associations rather than causal effects. All results should be interpreted within this framework.

3 Methodology and testable hypotheses

3.1 Theoretical Framework

3.1.1 Financial Literacy as Information Precision

The theoretical framework conceptualises financial literacy as a component of investors' information sets that improves precision in estimating asset returns and risks. Within a standard mean–variance framework, an investor i chooses portfolio risk θ_i to maximise expected utility subject to a perceived return distribution. The key insight is that financial literacy ℓ_i affects the investor's ability to form accurate beliefs about asset returns μ and variance σ^2.

Specifically, define perceived return estimation error as:

$$\epsilon_i = \mu^* - \hat{\mu}_i \tag{1}$$

where μ^* is the true expected return and $\hat{\mu}_i$ is the investor's estimate. Higher literacy reduces the variance of this estimation error, i.e., $\mathrm{Var}(\epsilon_i)$ decreases in ℓ_i. This creates a direct *competence effect*: investors with lower estimation error are more confident in identifying positive risk–return trade-offs, increasing optimal risky-asset allocation θ_i^*.

3.1.2 Overconfidence as Belief Distortion

Overconfidence operates through a separate channel: it distorts investors' beliefs about their own ability rather than about asset returns directly. Formally, define overconfidence as the systematic upward bias in self-assessed competence relative to objective performance:

$$\text{Overconfidence}_i = z(\text{Confidence}_i) - z(\ell_i) \tag{2}$$

where $z(\cdot)$ denotes standardisation. Positive values indicate that subjective confidence exceeds objective knowledge.

Within the mean–variance framework, overconfidence distorts perception in two related ways. First, the investor places excessive weight on perceived signal quality when forming expectations about returns. Second, by overstating precision of information, the investor effectively underestimates risk. Both effects increase optimal risky-asset allocation relative to the full-information benchmark. Consequently, overconfidence generates higher risk exposure even when objective financial knowledge is limited.

3.1.3 Interaction between Literacy and Overconfidence

Financial literacy and overconfidence interact through distinct but related behavioural channels. On the one hand, higher literacy improves calibration by reducing misperceptions of risk and narrowing the gap between subjective confidence and objective knowledge. On the other hand, literacy may also increase risk-taking directly by improving understanding of the risk–return trade-off and the benefits of diversification.

These mechanisms imply that the total effect of literacy on risk-taking can be decomposed into a direct and an indirect component:

$$\text{Total Effect} = \text{Direct Effect} + \text{Indirect Effect} \tag{3}$$

The indirect effect operates through overconfidence. If higher literacy reduces overconfidence, and overconfidence in turn increases risk-taking, then the indirect effect of literacy on risk-taking is negative. When the direct effect is positive and the indirect effect is negative, the overall relationship exhibits a *suppression structure*. The observed total effect of literacy on risk-taking may therefore understate the magnitude of the competence channel once overconfidence is controlled for.

This theoretical decomposition motivates the empirical mediation framework developed in Section 3.2.

3.2 Econometric Model

3.2.1 Baseline Specification

To estimate the total effect of financial literacy on behavioural and portfolio outcomes, we estimate the following baseline specification:

$$Y_i = \alpha + \beta FL_i + \boldsymbol{\gamma}'\mathbf{X}_i + \varepsilon_i \tag{4}$$

where Y_i denotes the outcome variable for individual i, FL_i represents the standardized financial literacy measure, and $\mathbf{X}_i$ is a vector of demographic control variables including age, gender, education, and income. The error term is denoted by ε_i.

Financial literacy is standardized to facilitate interpretation and comparability across models. Specifically, the financial literacy variable is defined as the Z-score of the respondent's raw literacy score:

$$FL_i = \frac{FL_i^{raw} - \mu_{FL}}{\sigma_{FL}} \tag{5}$$

where FL_i^{raw} denotes the respondent's raw financial literacy score, while μ_{FL} and σ_{FL} represent the sample mean and standard deviation of the literacy measure.

The dependent variable Y_i varies across hypotheses:

- Risk-taking (`Risk_Index` or `Risk_Z`) when testing H3
- Diversification (`Div_Score`) when testing H2
- Overconfidence measures when testing H4

The functional form of the model is adapted to the nature of the dependent variable. Continuous outcomes such as risk-taking and overconfidence are estimated using ordinary least squares (OLS) with heteroskedasticity-robust standard errors. Count outcomes such as diversification are estimated using Poisson regression. The coefficient β captures the total effect of financial literacy on the outcome variable prior to introducing mediation channels.

3.2.2 Mediation Model

To examine whether overconfidence mediates the relationship between financial literacy and risk-taking, we estimate a standard two-equation mediation framework following Baron and Kenny (1986).

First, we estimate the effect of financial literacy on overconfidence (the a-path):

$$OC_i = \alpha_1 + a\,FL_i + \boldsymbol{\gamma}_1'\mathbf{X}_i + u_i \tag{6}$$

where OC_i denotes the overconfidence measure for individual i, FL_i represents the standardized financial literacy score, and $\mathbf{X}_i$ is the vector of demographic control variables defined in Equation (4). The coefficient a captures the effect of financial literacy on miscalibration.

Second, we estimate the joint effect of financial literacy and overconfidence on risk-taking (the b- and c'-paths):

$$Risk_i = \alpha_2 + c'FL_i + b\,OC_i + \boldsymbol{\gamma}_2'\mathbf{X}_i + v_i \tag{7}$$

where $Risk_i$ denotes the risk-taking outcome for individual i. The coefficient b measures the effect of overconfidence on risk-taking, while c' represents the direct effect of financial literacy on risk-taking after accounting for the behavioural channel.

The indirect behavioural effect is given by:

$$\text{Indirect Effect} = a \times b \tag{8}$$

A negative value of $a \times b$ indicates that financial literacy reduces risk-taking indirectly by improving calibration. The total effect can be decomposed as:

$$\text{Total Effect} = c' + (a \times b) \tag{9}$$

Inference for the indirect effect is conducted using nonparametric bootstrap confidence intervals based on 5,000 replications, following Preacher and Hayes (2008), as the sampling distribution of the product term $a \times b$ is generally non-normal.

3.3 Testable hypotheses

Grounded in the theoretical framework above, we test the following hypotheses.

H1: Market Participation (untestable)

A positive association between literacy and market participation is expected based on prior literature. However, all respondents in the cleaned sample report current or planned investment activity, leaving no variation in the participation variable. H1 is therefore retained for theoretical completeness but excluded from all empirical analysis.

H2: Diversification

Higher financial literacy improves understanding of covariance and diversification benefits. We expect a positive association between literacy and portfolio diversification.

Expected sign: $\hat{\beta} > 0$

H3: Risk-Taking

The theoretical prediction for risk-taking is ambiguous. On the one hand, higher literacy may increase optimal risky-asset allocation through improved understanding of the risk premium. On the other hand, better calibration may reduce unwarranted risk-taking. We therefore test whether literacy is significantly associated with risk-taking without imposing a directional prior.

Expected sign: $\hat{\beta} \neq 0$

H4: Calibration

Higher financial literacy improves belief accuracy and reduces overconfidence. Accordingly, we expect literacy to be negatively associated with the overconfidence gap.

Expected sign: $\hat{a} < 0$

H5: Behavioural Channel

Overconfidence increases perceived signal precision and leads to higher risk-taking. We expect a positive association between overconfidence and risk-taking.

Expected sign: $\hat{b} > 0$

H6: Mediation

Overconfidence mediates the relationship between literacy and risk-taking. Specifically, if literacy reduces overconfidence ($a < 0$) and overconfidence increases risk-taking ($b > 0$), the indirect effect $a \times b$ is negative. We test whether the indirect effect differs significantly from zero using bootstrap inference.

3.4 Endogeneity and Limitations

This section discusses key identification challenges and limitations that constrain the interpretation of the empirical results. Given the cross-sectional survey design, the estimates should be interpreted as conditional associations rather than structural causal effects.

3.4.1 Reverse Causality

A primary concern is reverse causality. While the theoretical framework models financial literacy as influencing risk-taking, overconfidence, and portfolio construction, the direction of causality may also operate in the opposite direction. Investors who actively participate in financial markets may accumulate knowledge through experience ("learning by doing"), thereby increasing their measured financial literacy.

This concern is particularly relevant in the present sample, which consists of individuals recruited from investor communities and therefore likely to be relatively active market participants. The empirical models control for observable demographic proxies, including age, education, income, and gender. Notably, investment experience is not included as a control variable despite being a likely confounder of both literacy and risk-taking. Investors with more years of market participation may score higher on literacy questions through experiential learning and may simultaneously exhibit higher risk tolerance

through survivorship or habituation. The omission of experience means that the estimated literacy coefficients may partly capture experience-driven variation. Future specifications incorporating experience as a control or moderator would help disentangle these channels.

3.4.2 Omitted Variable Bias

Unobserved characteristics such as cognitive ability, numeracy, patience, intrinsic interest in finance, and baseline risk preferences are likely to correlate with both financial literacy and portfolio outcomes. Without an instrumental variable strategy or experimental variation in literacy, omitted variable bias cannot be ruled out.

3.4.3 Measurement Error

Financial literacy is measured using a survey-based battery of questions. While widely used in the literature, such instruments capture a limited subset of financial knowledge. Portfolio construction and allocation measures are also self-reported, which introduces recall error, rounding, and potential social desirability bias.

3.4.4 Mechanical Correlation in Gap-Based Overconfidence Measures

When overconfidence is defined as the gap between subjective confidence and objective literacy, a mechanical dependence arises between literacy and the overconfidence measure. This creates potential multicollinearity and interpretation challenges when both variables enter the same regression specification. To address this issue, the analysis employs a residual-based robustness measure of overconfidence, defined as the component of subjective confidence unexplained by objective literacy.

3.4.5 Mediation and Sequential Exogeneity

The mediation analysis assumes that financial literacy affects overconfidence, which in turn affects risk-taking. In a cross-sectional setting, sequential exogeneity cannot be formally verified. The mediation results should therefore be interpreted as evidence consistent with, but not definitive proof of, the proposed behavioural mechanism.

3.4.6 External Validity

The recruitment strategy targets online investor communities and thus involves self-selection. The findings describe behavioural variation within an investor-active segment rather than providing population-wide estimates.

3.4.7 Summary of Interpretation

Taken together, these limitations require that all results be interpreted as conditional associations rather than causal effects. The identification of structural literacy effects awaits experimental or quasi-experimental research designs.

4 Data

4.1 Data source and Recruitment

The empirical analysis is based on original survey data collected from Norwegian retail investors during October 2025–February 2026. The survey was distributed through online investor communities and broker-related forums where private investors are active. Distribution channels included:

- Reddit forums (*r/aksjer*, *r/TollbugataBets*, *r/personligokonomi*)
- Nordnet Shareville investor groups
- Facebook investment communities

The recruitment strategy provides direct access to an active investor population. However, it implies selection into a specific subculture of online-engaged, predominantly young (64% aged 18–34) and male (84%) retail investors. Communities such as *r/TollbugataBets* may attract particularly risk-tolerant individuals, which could inflate baseline risk-taking levels in the sample. The study therefore focuses on behavioural variation within this investor segment rather than estimating population-wide parameters, and caution is warranted when comparing descriptive statistics (e.g., the high mean risk index of 5.22/7) to broader populations. The raw dataset consists of **1,015 responses**.

4.2 Data Cleaning and Sample Restrictions

4.2.1 Initial processing

Before constructing variables, several cleaning steps were applied to ensure internal consistency and data quality. The following observations were excluded:

1. **Age restriction.** Respondents reporting being under 18 were removed.
2. **Attention check failure.** Respondents who failed either instructed-response attention check were excluded.
3. **Incomplete key responses.** Observations missing core variables were excluded on a model-by-model basis using listwise deletion.

After applying these restrictions, the final baseline sample consists of:

$$N = 746$$

For models involving risk-taking and overconfidence, the effective sample size ranges between 710 and 713 observations due to item-level non-response.

4.2.2 Category harmonisation

To avoid sparse cells in regression models, selected demographic categories were collapsed:

- Age grouped into: 18–34, 35–54, 55+
- Education grouped into: Low vs. High
- Income grouped into: Low, Mid, High, Prefer Not
- Gender collapsed into: Male (baseline), Female, Other

4.2.3 Treatment of missing values and outliers

No extreme outliers were removed. Financial literacy scores are naturally bounded (0–6), and risk measures are constructed from Likert scales (1–7). Missing observations are handled via listwise deletion within each regression model. All regressions use heteroskedasticity-robust standard errors (HC1).

4.2.4 Participation variable

All respondents in the cleaned dataset report either current investment activity or plans to invest. As a result, the participation variable exhibits no cross-sectional variation. Consequently, the market participation hypothesis (H1) cannot be estimated within this sample.

4.3 Variable construction

4.3.1 Objective financial literacy

Objective financial literacy is measured using six multiple-choice questions covering compound interest, inflation, the risk–return trade-off, diversification, investment fees, and percentage reasoning. Each item is coded as

$$Correct_{ik} \in \{0, 1\},$$

where $Correct_{ik}$ indicates whether respondent i answered question k correctly.

The aggregate literacy score is defined as

$$FL_i^{raw} = \sum_{k=1}^{6} Correct_{ik}, \tag{10}$$

which ranges from 0 to 6. In the sample, the mean literacy score is 5.63 with a standard deviation of 0.73.

For regression analysis, financial literacy is standardized to facilitate interpretation and comparability across models:

$$FL_i = \frac{FL_i^{raw} - \mu_{FL}}{\sigma_{FL}}, \tag{11}$$

where μ_{FL} and σ_{FL} denote the sample mean and standard deviation of the raw literacy score.

4.3.2 Subjective confidence and overconfidence

Subjective confidence is constructed using two survey measures:

- Self-rated financial knowledge (0–6 scale)
- Self-ranking relative to other investors (percentile category)

Both measures are standardized and averaged to form a composite confidence index, denoted $Conf_i$.

Overconfidence is operationalized using two alternative approaches.

Gap measure

Overconfidence is defined as the difference between standardized confidence and standardized objective literacy:

$$OC_i^{gap} = z(Conf_i) - z(FL_i^{raw}). \tag{12}$$

Positive values indicate subjective confidence exceeding objective financial knowledge.

Residual measure (robustness)

Following common approaches in the behavioural finance literature, overconfidence is alternatively defined as the residual from the regression

$$Conf_i = \alpha + \beta FL_i^{raw} + \varepsilon_i. \tag{13}$$

The residual component represents confidence unexplained by objective literacy:

$$OC_i^{res} = \hat{\varepsilon}_i.$$

4.3.3 Risk-taking behaviour

Risk-taking is measured using six Likert-scale survey items (1–7), with higher values indicating greater willingness to take financial risk. Selected items are reverse-coded to ensure directional consistency. The composite risk index is constructed as

$$Risk_i = \frac{1}{K} \sum_{k=1}^{K} Item_{ik}, \tag{14}$$

where $Item_{ik}$ denotes the response of individual i to item k and K is the number of valid responses. Respondents must answer at least four items to be included in the index.

Internal consistency of the scale is below the conventional 0.70 threshold (Nunnally, 1978) (Cronbach's $\alpha = 0.665$), placing it in the questionable-to-acceptable range. While values in this range are not uncommon for short behavioural scales measuring heterogeneous risk attitudes, the modest reliability implies that measurement error may attenuate estimated coefficients. Results involving the risk index should be interpreted with this caveat in mind. For standardized interpretation, a Z-score transformation is also used:

$$Risk_i^Z = z(Risk_i).$$

4.3.4 Portfolio construction outcomes

Portfolio construction outcomes are captured using several measures of investment behaviour, including:

- The number of asset classes held (diversification score)
- Participation in higher-risk instruments (e.g., cryptocurrencies or derivatives)

- Approximate exposure to risky assets

Diversification is proxied by counting the number of selected asset categories held by the respondent. This measure captures portfolio breadth but treats all asset classes equally; adding cash and adding derivatives both contribute one unit, despite very different implications for portfolio sophistication and risk. Alternative measures, such as concentration indices or distinctions between risky and safe asset classes, could provide additional insight but are beyond the scope of the present instrument. The mean diversification score in the sample is 5.90 with a standard deviation of 2.46.

4.3.5 Control variables

All regression models include standard demographic controls, denoted collectively by X_i. These include age (collapsed categories), gender, education level, and income category. Including these variables helps reduce omitted-variable bias and ensures comparability with prior household-finance research.

4.4 Descriptive statistics and correlations

Table 1: Descriptive Statistics

Variable	N	Mean	SD	Min	25%	50%	75%	Max
Panel A: Financial Literacy								
FL^{raw}	746	5.630	0.734	0.000	5.000	6.000	6.000	6.000
FL_i	746	0.000	1.001	−7.671	−0.858	0.504	0.504	0.504
Panel B: Overconfidence								
$Conf_i$	710	0.000	0.921	−2.731	−0.594	0.291	0.703	1.956
OC_i^{gap}	710	−0.047	1.170	−3.473	−0.701	−0.188	0.661	3.852
OC_i^{res}	710	0.000	0.890	−2.849	−0.712	0.172	0.585	2.071
Panel C: Risk-Taking								
$Risk_i$	713	5.219	0.846	1.500	4.667	5.167	5.833	7.000
$Risk_i^Z$	713	0.000	1.001	−4.399	−0.653	−0.062	0.727	2.107
Panel D: Portfolio Diversification								
Div_i	746	5.897	2.459	0.000	4.000	6.000	8.000	13.000

Notes: FL^{raw} (dataset variable `FinLit_Sum`) is the raw count of correct answers across six financial literacy items (range: 0–6). FL_i (`Financial_Literacy_Z`) is the standardized literacy measure used in the regression models. $Conf_i$ is the composite subjective confidence index constructed as the average of two standardized self-assessment measures. OC_i^{gap} (`Overconfidence_Gap`) is defined as z(Confidence) − z(Literacy), where positive values indicate overconfidence relative to objective knowledge. OC_i^{res} (`Overconfidence_Residual`) is the residual from regressing subjective confidence on literacy. $Risk_i$ (`Risk_Index`) is the mean of six 7-point Likert risk-tolerance items (three reverse-coded); Cronbach's $\alpha = 0.665$. Div_i (`Div_Score`) counts the number of distinct asset classes held or planned (range: 0–13). N varies across variables due to item-level non-response; regression models use listwise deletion.

Key descriptive patterns are as follows. Financial literacy is high relative to international survey benchmarks. The risk index mean (5.22) suggests relatively high self-reported risk tolerance. Literacy positively correlates with diversification ($r \approx 0.19$) and risk-taking ($r \approx 0.12$). Literacy negatively correlates with the overconfidence gap ($r \approx -0.56$). Variance Inflation Factors in regression models are below 2, indicating no multicollinearity concerns.

Table 2: Pearson Correlation Matrix

	FL^{raw}	$Conf_i$	OC_i^{gap}	OC_i^{res}	$Risk_i$	Div_i
FL^{raw}	1.000	0.257	−0.563	0.000	0.124	0.194
$Conf_i$	0.257	1.000	0.654	0.966	0.307	0.137
OC_i^{gap}	−0.563	0.654	1.000	0.827	0.166	0.034
OC_i^{res}	0.000	0.966	0.827	1.000	0.285	0.113
$Risk_i$	0.124	0.307	0.166	0.285	1.000	0.000
Div_i	0.194	0.137	0.034	0.113	0.000	1.000

Notes: The table reports Pearson correlation coefficients between the main variables used in the analysis. FL^{raw} (dataset variable `FinLit_Sum`) denotes the raw financial literacy score. $Conf_i$ (`Subj_Conf_Index`) is the standardized subjective confidence index. OC_i^{gap} (`Overconf_Gap`) represents the overconfidence gap defined as z(Confidence) − z(Literacy). OC_i^{res} (`Overconf_Resid`) is the residual-based overconfidence measure obtained from regressing subjective confidence on literacy. $Risk_i$ (`Risk_Index`) denotes the composite risk-taking index, and Div_i (`Div_Score`) measures portfolio diversification as the number of asset classes held.

4.5 Limitations

Two data-specific limitations should be noted. First, the participation variable exhibits no cross-sectional variation (Section 4.2.4), preventing estimation of H1. Second, the literacy distribution is concentrated near the upper bound (mean 5.63/6), which may attenuate estimated effects. Broader identification limitations are discussed in Section 3.4.

5 Results and analysis

5.1 Sample Characteristics and Preliminary Evidence

The final analysis sample consists of 746 respondents after applying eligibility and attention filters. As noted in Sections 1.3 and 4.2.4, H1 cannot be estimated due to universal participation in this sample and is excluded from all results.

Descriptive statistics are presented in Table 1. The average objective financial literacy score (FL^{raw}) is 5.63 out of 6 ($SD = 0.73$), with a median of 6 and an interquartile range from 5 to 6. This indicates that a large share of respondents cluster near the upper bound of the literacy scale.

The standardized financial literacy measure (FL_i) is centered at zero by construction and exhibits unit variance. Subjective confidence ($Conf_i$) also displays substantial variation ($SD = 0.92$), while the overconfidence gap measure (OC_i^{gap}) has a standard

deviation of 1.17, indicating meaningful dispersion in calibration despite the high average literacy level.

The risk index ($Risk_i$) has a mean of 5.22 on a 1–7 scale ($SD = 0.85$), with the 25th percentile at 4.67 and the 75th percentile at 5.83. This suggests moderately high but not extreme risk tolerance within the sample. Internal consistency of the risk scale is modest (Cronbach's $\alpha = 0.665$), which is below the conventional 0.70 threshold but not unusual for short financial risk tolerance instruments. This implies some measurement noise, which may attenuate estimated associations.

Diversification (Div_i), measured as the number of asset categories held, has a mean of 5.90 ($SD = 2.46$), with values ranging from 0 to 13. The interquartile range from 4 to 8 indicates substantial heterogeneity in portfolio breadth even within an investor-active sample.

Pairwise correlations are reported in Table 2. Financial literacy (FL_i) is positively correlated with diversification ($r = 0.194$) and risk-taking ($r = 0.124$), providing preliminary support for H2 and H3. Literacy is strongly negatively correlated with the overconfidence gap ($r = -0.563$), consistent with H4 and indicating that more knowledgeable investors are substantially better calibrated. Subjective confidence ($Conf_i$) is positively correlated with risk-taking ($r = 0.307$), while overconfidence (OC_i^{gap}) is also positively related to risk ($r = 0.166$), providing initial support for the behavioural channel examined in H5.

Importantly, the relatively strong correlation between subjective confidence and the gap-based overconfidence measure ($r = 0.654$) underscores the need for the residual-based robustness specification discussed earlier. Variance inflation factors (VIFs) are below 2 in all multivariate models, indicating no multicollinearity concerns.

5.2 Financial literacy and diversification (H2)

To test H2, diversification, measured as the number of asset classes held (`Div_Score`), is estimated using a Poisson regression model with heteroskedasticity-robust standard errors. Given the count nature of the dependent variable, the Poisson specification provides an appropriate functional form for modeling expected diversification levels.

Financial literacy is positively and statistically significantly associated with diversification ($\hat{\beta} = 0.0305$, $p = 0.043$). Interpreting the coefficient in exponential form, a one standard deviation increase in financial literacy is associated with approximately a 3 percent increase in the expected number of asset classes held, holding controls constant.

In economic terms, relative to the sample mean of 5.90 asset categories, this corresponds to an increase of roughly 0.18 additional asset classes. While modest in magnitude, the

effect represents a systematic shift toward broader portfolio breadth within an already investor-active sample.

The result is consistent with the informational friction mechanism documented by Guiso and Jappelli (2008) and Abreu and Mendes (2012), suggesting that greater financial literacy improves understanding of diversification benefits and reduces under-diversification.

Conclusion for H2: H2 is supported. Higher financial literacy is associated with broader portfolio diversification.

5.3 Financial literacy and risk-taking (H3)

Risk-taking ($Risk_i$) is estimated using ordinary least squares (OLS) with heteroskedasticity-robust (HC1) standard errors. Financial literacy (FL_i) is positively and statistically significantly associated with the risk index ($\hat{\beta} = 0.099$, $p = 0.008$). This estimate is based on the full sample with all demographic controls ($N = 683$). The total effect estimated within the mediation subsample ($N = 710$) is $\hat{\beta} = 0.122$ ($p = 0.004$); the difference reflects variation in sample composition due to item-level missingness in control variables.

Using standardized coefficients, a one standard deviation increase in financial literacy is associated with a 0.12 standard deviation increase in risk-taking ($\hat{\beta}_{std} = 0.117$, $p = 0.008$). In the original scale, this corresponds to approximately a 0.10 point increase on the 1–7 risk tolerance scale relative to a sample mean of 5.22.

Although the magnitude is moderate, the effect reflects a systematic upward shift in stated risk tolerance within an already investor-active population. This association is consistent with the interpretation that more financially literate investors have a stronger understanding of long-run risk–return trade-offs and are therefore more comfortable accepting risk exposure.

Conclusion for H3: H3 is supported. Financial literacy is positively associated with risk-taking behaviour.

5.4 Financial literacy and overconfidence (H4)

Overconfidence is measured as the difference between standardized subjective confidence and objective financial literacy, denoted OC_i^{gap}. By construction, positive values indicate that subjective confidence exceeds actual knowledge, reflecting miscalibration.

Financial literacy (FL_i) is strongly and negatively associated with the overconfidence gap ($\hat{\beta} = -0.779$, $p < 0.001$). In standardized terms, a one standard deviation increase

in literacy is associated with a 0.78 standard deviation reduction in the overconfidence gap, indicating a substantial improvement in calibration.

To address the mechanical component explicitly, we also estimate models using the residual-based overconfidence measure (OC_i^{res}). Under this specification, the coefficient on literacy becomes smaller and statistically insignificant ($p = 0.122$). This suggests that literacy primarily improves calibration relative to objective knowledge rather than independently reducing raw subjective confidence.

Conclusion for H4: H4 is supported when overconfidence is measured as a calibration gap ($p < 0.001$). However, under the residual-based specification, which removes the mechanical correlation between literacy and the gap measure, the coefficient becomes smaller and statistically insignificant ($p = 0.122$). The calibration finding is therefore sensitive to operationalisation. Readers should note that the gap measure contains a built-in negative dependence on literacy, and the residual measure may provide a more conservative test. On balance, the evidence is suggestive of a calibration effect but not unambiguous.

5.5 Overconfidence and risk-taking (H5)

Overconfidence is positively and statistically significantly associated with risk-taking ($\hat{\beta} = 0.125$, $p < 0.001$). Using standardized variables in the mediation framework, a one standard deviation increase in overconfidence is associated with approximately a 0.17 standard deviation increase in risk-taking. Given the standard deviation of the risk index (0.85), this corresponds to an increase of roughly 0.15 points on the 1–7 risk tolerance scale.

This finding aligns with Barber and Odean (2001) and Broihanne et al. (2014), who document that overconfident or miscalibrated investors underestimate risk and allocate more heavily toward risky assets.

Conclusion for H5: H5 is supported. Overconfidence is strongly associated with greater risk-taking behaviour, consistent with a belief distortion channel.

5.6 Mediation Analysis: Quantifying the Dual-Channel Structure (H6)

Interpretive note. The mediation framework assumes a specific causal ordering: literacy affects overconfidence, which in turn affects risk-taking. With cross-sectional data, this ordering cannot be empirically verified. Alternative sequences, for instance, that risk-taking experience shapes both literacy and calibration simultaneously, are equally consistent with the observed correlations. The mediation results should therefore be interpreted as evidence that the data are consistent with the proposed mechanism, not

as confirmation of causal mediation. The terminology of "direct" and "indirect" effects follows standard mediation conventions but refers to statistical decompositions rather than established causal pathways.

Path Coefficients

The mediation framework is estimated following the procedure outlined in Section 3.2.2. All path coefficients are statistically significant. The a-path (literacy $\rightarrow$ overconfidence) is negative ($\hat{a} < 0$), consistent with H4. The b-path (overconfidence $\rightarrow$ risk-taking) is positive ($\hat{b} > 0$), consistent with H5. The direct effect (c'-path) is $\hat{c}' = 0.351$ ($p < 0.001$).

Note: Minor differences between path coefficients reported here and the baseline H3/H4 estimates in Sections 5.3–5.4 reflect differences in effective sample size due to listwise deletion. The mediation sample ($N = 710$) includes all observations with complete data on literacy, overconfidence, and risk-taking, while the baseline models additionally require complete demographic controls ($N = 683$).

Indirect Effect

The indirect effect is $a \times b = -0.229$ ($p < 0.001$, bootstrap 95% CI excludes zero). This confirms mediation. The CI width represents approximately 57% of the point estimate (0.130/0.229), which is typical for well-powered mediation analyses with large samples ($N = 710$) and stable relationships.

Suppression Pattern

The negative proportion mediated (-187.3%) is characteristic of *suppressor variables* in the mediation literature (MacKinnon et al., 2000). This pattern arises when: (i) the mediator is correlated with both X and Y; (ii) the mediator suppresses irrelevant variance in the $X \rightarrow Y$ relationship; and (iii) controlling for the mediator reveals a stronger direct effect.

In our context, overconfidence acts as a suppressor: controlling for it removes the calibration channel (which dampens risk-taking), thereby revealing the true strength of the competence channel. Without this control, the two opposing forces mask each other.

Interpretation

The mediation analysis provides strong support for H6. The estimates are consistent with two distinct channels linking literacy to risk-taking:

Competence Channel (Direct Effect) Higher literacy is associated with greater risk-taking after controlling for overconfidence ($\hat{\beta} = 0.351$), consistent with improved understanding of risk–return trade-offs.

Calibration Channel (Indirect Effect) Higher literacy is associated with lower overconfidence, which in turn is associated with lower risk-taking ($\hat{\beta} = -0.229$).

Net Result The positive direct association dominates the negative indirect association, yielding a net positive literacy–risk relationship. However, since these estimates derive from cross-sectional data, the decomposition reflects statistical patterns consistent with the proposed mechanism rather than confirmed causal pathways. That said, the net effect ($\hat{\beta} = 0.122$) is substantially smaller than the direct effect ($\hat{\beta} = 0.351$) because the calibration mechanism dampens approximately 65% of the direct effect [$\frac{0.351-0.122}{0.351} \approx 0.65$].

Practical Significance

The total literacy–risk effect ($\hat{\beta} = 0.122$) implies that a one standard deviation increase in literacy is associated with a 0.122 SD increase in risk tolerance. On the original 1–7 risk scale (SD = 0.85), this translates to approximately 0.10 points on the risk index. This is smaller than the gender effect (males score $\approx$ 0.25 points higher) but comparable to the effect of a 10-year age difference.

Connection to Theoretical Framework.

The empirical results map closely onto the theoretical structure developed in Section 3.1. The competence channel ($\hat{c}' = 0.351$) corresponds to the information precision mechanism discussed in Section 3.1.1, while the calibration channel ($a \times b = -0.229$) corresponds to the belief distortion mechanism in Section 3.1.2. The suppression pattern, in which the two channels partially offset each other, validates the interaction mechanism proposed in Section 3.1.3.

5.7 Overall Interpretation of Findings

Table 3: Summary of Main Regression Results

Hypothesis	Model	Predictor	Coefficient	Std. Error	N	R^2
H2: Diversification	Poisson	FL_i	0.030**	(0.015)	683	—
H3: Risk-taking	OLS	FL_i	0.099***	(0.037)	683	0.031
H4: Overconfidence (Gap)	OLS	FL_i	−0.779***	(0.039)	683	0.357
H5: OC → Risk	OLS	OC_i^{gap}	0.125***	(0.028)	683	0.047

Notes: FL_i denotes the standardized financial literacy measure (dataset variable `Financial_Literacy_Z`). OC_i^{gap} (`Overconfidence_Gap`) is the overconfidence gap defined as the standardized difference between subjective confidence and objective literacy. Standard errors are reported in parentheses. R^2 is reported for OLS models; "—" for the Poisson specification (H2), which does not yield a conventional R^2. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 3 summarises the main regression results across hypotheses. Three core findings emerge.

First, financial literacy is positively associated with portfolio diversification. Consistent with H2, more knowledgeable investors hold broader portfolios, suggesting that informational capacity reduces under-diversification even within an already active investor sample.

Second, financial literacy is positively associated with risk-taking. The magnitude is moderate but statistically robust, indicating that more literate investors report higher willingness to bear risk. Importantly, this association does not appear to reflect naïve risk-taking.

Third, financial literacy is associated with substantially better calibration when overconfidence is measured as a gap between standardised confidence and knowledge. However, this finding is sensitive to operationalisation: the residual-based measure, which removes the mechanical component, does not yield a statistically significant result. The calibration effect should therefore be regarded as indicative rather than definitive.

These three findings are consistent with the competing mechanisms identified in the mediation analysis (Section 5.6).

5.8 Robustness Checks

To ensure findings are not driven by specific modeling choices, sample characteristics, or influential observations, we conduct five sets of robustness checks. Table 4 presents

results for H3 (literacy → risk-taking) and H4 (literacy → calibration) across alternative specifications.

Table 4: Robustness Checks: Key Coefficients Across Specifications

Specification	**N**	**H2 Div.**	**H3 Risk**	**H4 OC Gap**	**H5 OC→Risk**
Baseline	683	0.030*	0.099**	−0.779***	0.125***
Gender: Male	628	—	0.126**	−0.747***	—
Age: Young (18-34)	475	—	0.139**	−0.732***	—
Age: Middle (35-54)	232	—	0.030	−0.871***	—
Exclude Literacy=6	195	—	0.042	−0.787***	—
Exclude Outliers (Top 5%)	711	—	0.114***	−0.766***	—
No Controls	710	—	0.115***	−0.719***	—

Notes: Cells report the coefficient on the key predictor (`Financial_Literacy_Z` for H2–H4; `Overconfidence_Gap` for H5). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. `--` indicates the check was not run for that hypothesis. All models use HC1 robust standard errors.

5.8.1 Subgroup Analyses

Results are generally consistent across demographic subgroups, though with some heterogeneity.

Gender: Among male investors (84% of sample, $N = 628$), the literacy–risk association is slightly stronger ($\hat{\beta} = 0.126$, $p = 0.003$) than in the full sample. The female subsample ($N \approx 55$, 7% of sample) is too small for reliable subgroup analysis. The calibration effect (H4) among males is highly consistent with the full sample ($\hat{\beta} = -0.747$, $p < 0.001$).

Age: The literacy–risk association exhibits substantial heterogeneity by age. Among young investors (18–34), the effect is strong and significant ($\hat{\beta} = 0.139$, $p = 0.002$). However, among middle-aged investors (35–54), the association becomes non-significant ($\hat{\beta} = 0.030$, $p = 0.646$). In contrast, the calibration effect (H4) is remarkably stable across all age groups (ranging from $\hat{\beta} = -0.732$ to -0.871, all $p < 0.001$), suggesting that literacy improves self-assessment accuracy universally.

5.8.2 Ceiling Effects Sensitivity

Excluding respondents with perfect literacy scores (6/6, $n = 276$ removed) yields H3: $\hat{\beta} = 0.042$ ($p = 0.605$, not significant) and H4: $\hat{\beta} = -0.787$ ($p < 0.001$). The calibration result remains virtually identical, but the risk-taking association becomes non-significant in this restricted sample.

This finding warrants careful interpretation. The literacy–risk association (H3) appears to be driven primarily by respondents at the upper end of the literacy distribution,

where the distinction between scoring 5 versus 6 out of 6 accounts for most of the observable variation. Once this group is excluded, the remaining sample ($N = 195$) exhibits insufficient variation in literacy to identify a statistically reliable effect.

Two interpretations are consistent with this pattern. First, the effect may be genuine but concentrated at the high end of the literacy distribution, meaning that the final increment of financial knowledge has a disproportionate impact on risk-taking. Second, the association may partly reflect ceiling-induced measurement artefacts rather than a true literacy–risk relationship across the full distribution. Given the cross-sectional design, these interpretations cannot be distinguished empirically. Readers should therefore treat the H3 result as suggestive rather than definitive. The practical implication is that the literacy–risk association documented in this thesis may reflect a threshold effect at the top of the distribution (scoring 6 versus 5) rather than a continuous gradient across literacy levels. Future research using instruments with greater discriminatory power at high literacy levels, such as advanced financial literacy batteries (Van Rooij et al., 2011), would help resolve this ambiguity. This limitation does not affect H4, which remains robust to ceiling exclusions.

5.8.3 Outlier Sensitivity

Excluding the 5% most influential observations based on Cook's distance (threshold $>$ 95th percentile) yields H3: $\hat{\beta} = 0.114$ ($p < 0.001$) and H4: $\hat{\beta} = -0.766$ ($p < 0.001$). Both results are nearly identical to baseline (H3: 0.099; H4: -0.779), indicating that findings are not driven by unusual respondents.

5.8.4 Specification Robustness

Removing all control variables yields H3: $\hat{\beta} = 0.115$ ($p < 0.001$) and H4: $\hat{\beta} = -0.719$ ($p < 0.001$). Both coefficients are slightly *larger* than in the baseline controlled specification, confirming that demographic controls attenuate but do not eliminate literacy effects.

Multicollinearity Assessment

VIFs for all predictors in the mediation model are below 2.0 (see Appendix B, Table B.6a), well below conventional thresholds (VIF $>$ 5 or 10). This confirms that the correlation between `Financial_Literacy_Z` and `Overconfidence_Gap` ($r = -0.563$), while substantial, does not create problematic multicollinearity.

5.8.5 Overall Robustness Assessment

H3 (Risk-Taking): Significant in 5 of 7 specifications. The coefficient ranges from 0.030 to 0.139, with most estimates clustered around $\hat{\beta} \approx 0.10$–0.13.

H4 (Calibration): Significant in all 7 specifications at $p < 0.001$. The coefficient is remarkably stable, ranging from -0.719 to -0.871, with most estimates near $\hat{\beta} \approx -0.75$ to -0.78.

Conclusion: The calibration effect is robust across all specifications, while the risk-taking association holds in most but shows age-related heterogeneity. Both core findings survive outlier exclusion and alternative specifications, supporting the two-channel interpretation.

6 Conclusion

This thesis examined how financial literacy influences risk-taking behaviour, overconfidence, and portfolio construction among Norwegian retail investors. The objective was to determine whether literacy affects investor behaviour directly and whether overconfidence operates as a behavioural mechanism linking knowledge to portfolio outcomes.

6.1 Main Findings

The results yield three central conclusions.

First, financial literacy is systematically associated with improved portfolio structure. More knowledgeable investors hold broader portfolios, consistent with standard portfolio theory and prior empirical evidence. Even within an already investor-active sample, differences in literacy translate into measurable differences in diversification.

Second, literacy is positively associated with risk-taking in most specifications. However, this association does not survive the exclusion of respondents with perfect literacy scores, suggesting it may be concentrated at the upper boundary of the literacy distribution rather than reflecting a continuous gradient. The finding should be treated as suggestive, particularly given the limited variation in the literacy measure.

Third, financial literacy is associated with improved calibration, though this finding is sensitive to how overconfidence is measured (see Section 5.4).

Taken together, the mediation analysis reveals two opposing pathways. Literacy increases risk-taking directly through a competence channel, while simultaneously dampening

risk-taking indirectly by reducing overconfidence. The competence effect dominates, implying that the net association between literacy and risk-taking remains positive.

Overall, the evidence suggests that financial literacy does not simply encourage greater risk exposure. Rather, it appears to promote more informed and better-calibrated risk-taking, improving both portfolio structure and belief accuracy.

6.2 Contribution to the literature

The findings contribute to the household finance and behavioural finance literature in three ways.

First, they provide further evidence that informational frictions help explain under-diversification, even within a financially developed and investor-active setting.

Second, the results reinforce the empirical distinction between objective financial literacy and overconfidence. Higher literacy is not associated with inflated subjective confidence; rather, it improves calibration by narrowing the gap between perceived and actual knowledge. This supports the view that competence and confidence are separate behavioural constructs with distinct effects.

Third, and most importantly, the thesis integrates the literacy and overconfidence literatures within a unified mediation framework. This decomposition into competence and calibration effects, documented in Section 5.6, provides a more precise interpretation of the literacy–risk relationship than models treating these mechanisms independently.

6.3 Economic and Policy Implications

The findings suggest that improving financial literacy may enhance portfolio quality without necessarily encouraging excessive speculative behaviour. Although literacy is associated with higher risk-taking, this increase appears to reflect improved understanding of risk–return trade-offs rather than inflated confidence. At the same time, literacy reduces miscalibration, dampening confidence-driven excess risk.

For policymakers and financial institutions, the results suggest that financial education initiatives should extend beyond factual knowledge acquisition. Interventions that also improve self-assessment accuracy and calibration may be particularly valuable. Tools that provide feedback on actual versus perceived knowledge, clearer communication of risk metrics, and structured decision environments may help reinforce both competence and calibration.

6.4 Limitations

The study is subject to several limitations discussed in Section 3.4, including cross-sectional design constraints, self-selection from investor communities, self-reported measures, and gender imbalance (84% male). Of particular note, the literacy distribution is heavily concentrated near the upper bound (mean 5.63/6), and the literacy–risk association (H3) does not survive robustness checks that exclude respondents with perfect scores. This limits confidence in the generalisability of the H3 finding beyond high-literacy investor populations. These constraints mean findings should be interpreted as conditional associations within an investor-active population. Future research could address these limitations through randomized financial education interventions, quasi-experimental designs exploiting policy variation, instrumental variable strategies, or longitudinal panel data enabling within-person estimation.

6.5 Final Answer to the Research Question

How does financial literacy influence risk-taking behaviour, overconfidence, and portfolio construction among Norwegian investors?

The evidence indicates that higher financial literacy is associated with broader portfolio diversification, greater willingness to bear risk, and substantially improved calibration. More knowledgeable investors not only take more risk on average, but do so with a closer alignment between perceived and actual competence.

Mediation analysis reveals two competing channels. Financial literacy increases risk-taking directly through a competence effect, while indirectly dampening risk-taking by reducing overconfidence. Although improved calibration partially offsets the positive direct effect, the net association between literacy and risk-taking remains positive.

Overall, financial literacy appears to shape both the level and the quality of risk-taking. It influences not only how much risk investors accept, but also whether that risk reflects informed decision-making rather than miscalibrated confidence.

References

Abreu, M., & Mendes, V. (2012). Information, overconfidence and trading: Do the sources of information matter? *Journal of Economic Psychology*, *33*(4), 868–881. https://doi.org/10.1016/j.joep.2012.04.003

AksjeNorge. (2024). *Aksjestatistikk 2. kvartal 2024*. https://aksjenorge.no/wp-content/uploads/2024/07/Aksjestatistikk-2-Kvartal-2024-AksjeNorge.pdf

Baek, H. Y., & Cho, D. D. (2022). Overconfidence and risky investment choices. *Economics Bulletin*, *42*(4), 2267–2278. http://www.accessecon.com/Pubs/EB/2022/Volume42/EB-22-V42-I4-P187.pdf

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, *55*(2), 773–806. https://doi.org/10.1111/0022-1082.00226

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, *116*(1), 261–292. https://doi.org/10.1162/003355301556400

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Broekema, S. P. M., & Kramer, M. M. (2021). Overconfidence, financial advice seeking and household portfolio under-diversification. *Journal of Risk and Financial Management*, *14*(11), 553. https://doi.org/10.3390/jrfm14110553

Broihanne, M.-H., Merli, M., & Roger, P. (2014). Overconfidence, risk perception and the risk-taking behavior of finance professionals. *Finance Research Letters*, *11*(2), 64–73. https://doi.org/10.1016/j.frl.2013.11.002

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522–550. https://doi.org/10.1111/j.1542-4774.2011.01015.x

E24. (2025, July 7). *Slik investerte nordmenn på børs i første halvår*. https://e24.no/boers-og-finans/i/93aJBq/slik-investerte-nordmenn-paa-boers-i-foerste-halvaar

Global Financial Literacy Excellence Center. (2016). *Financial literacy around the world: Insights from the Standard & Poor's Ratings Services Global Financial Literacy Survey*. https://gflec.org/wp-content/uploads/2015/11/3313-Finlit_Report_FINAL-5.11.16.pdf

Grable, J., & Lytton, R. H. (1999). Financial risk tolerance revisited: The development of a risk assessment instrument. *Financial Services Review*, *8*(3), 163–181. https://doi.org/10.1016/S1057-0810(99)00041-4

Guiso, L., & Jappelli, T. (2008). *Financial literacy and portfolio diversification* (EUI Working Papers, ECO 2008/31). European University Institute. https://www.eief.it/files/2008/11/guiso.pdf

Kaupr. (2024, January 10). *New record in number of Norwegians owning shares*. https://www.kaupr.io/en/news/new-record-in-number-of-norwegians-owning-shares

Kramer, M. M. (2016). Financial literacy, confidence and financial advice seeking. *Journal of Economic Behavior & Organization*, *131*, 198–217. https://doi.org/10.1016/j.jebo.2016.08.016

Lusardi, A., & Mitchell, O. S. (2011). Financial literacy around the world: An overview. *Journal of Pension Economics and Finance*, *10*(4), 497–508. https://doi.org/10.1017/S1474747211000448

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, *1*(4), 173–181. https://doi.org/10.1023/A:1026595011371

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, *115*(2), 502–517. https://doi.org/10.1037/0033-295X.115.2.502

Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). McGraw-Hill.

Norsk Gjeldsinformasjon. (n.d.). *Årsrapport og statistikk over usikret gjeld*. Retrieved December 14, 2025, from https://www.norskgjeld.no/

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891. https://doi.org/10.3758/BRM.40.3.879

Statistics Norway. (n.d.). *Inntekt, formue og skatt for husholdninger*. Retrieved December 14, 2025, from https://www.ssb.no/

Van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market

participation. *Journal of Financial Economics*, *101*(2), 449–472. https://doi.org/10.1016/j.jfineco.2011.03.006

A Survey Instrument

This appendix presents the complete survey instrument administered to Norwegian retail investors. The survey was distributed online through investor communities and social media platforms between October 2025 and February 2026. All questions were presented in Norwegian; English translations are provided in brackets where relevant for international readers.

A.1 Financial Literacy Test (Objective Knowledge Assessment)

Respondents completed six multiple-choice questions measuring financial literacy. Each correct answer was awarded one point, yielding a total score ranging from 0 to 6. Correct answers are marked with ✓ below.

Question 4: Compound Interest

Q4. Renters rente [Compound Interest]

"Du setter inn 100 kr på en sparekonto til 6% rente. Hvor mye har du etter to år? (ingen uttak i mellomtiden)"

[Translation: You deposit 100 NOK in a savings account at 6% interest. How much do you have after two years? (no withdrawals in the meantime)]

- 106 kr [106 NOK]
- 112 kr [112 NOK]
- 112,36 kr [112.36 NOK] ✓ **Correct**
- 130 kr [130 NOK]
- Vet ikke [Don't know]

Scoring: 1 point if "112,36 kr" selected, 0 otherwise.
Concept tested: Understanding of compound interest and exponential growth.

Question 5: Inflation and Real Returns

Q5. Kjøpekraft [Purchasing Power]

"Hvis kontoen din gir 1% rente og inflasjonen er 3%, hvordan endrer kjøpekraften seg etter ett år?"

[Translation: If your account gives 1% interest and inflation is 3%, how does your purchasing power change after one year?]

- Den øker [It increases]

- Den synker [It decreases] ✓ **Correct**
- Forblir den samme [Stays the same]
- Vet ikke [Don't know]

Scoring: 1 point if "Den synker" selected, 0 otherwise.
Concept tested: Inflation eroding real returns.

Question 6: Risk and Diversification Principle

Q6. Diversifisering [Diversification]

"Bedøm utsagnet: 'Å kjøpe aksjer i ett selskap er vanligvis mer risikabelt enn å kjøpe et bredt aksjefond"'

[Translation: Evaluate the statement: "Buying shares in a single company is usually more risky than buying a broad equity fund"]

- Sant [True] ✓ **Correct**
- Usant [False]
- Vet ikke [Don't know]

Scoring: 1 point if "Sant" selected, 0 otherwise.
Concept tested: Understanding of diversification and single-stock risk.

Question 7: Portfolio Risk Assessment

Q7. Risiko [Risk]

"En investering med høy forventet avkastning har vanligvis. . . "

[Translation: An investment with high expected return usually has. . .]

- Lav risiko [Low risk]
- Høy risiko [High risk] ✓ **Correct**
- Samme risiko som andre [The same risk as others]
- Vet ikke [Don't know]

Scoring: 1 point if "Høy risiko" selected, 0 otherwise.
Concept tested: Risk-return trade-off.

Question 9: Investment Fees and Net Returns

Q9. Gebyrer [Fees]

"To fond forventes å gi samme avkastning før kostnader. Fond A har 1,0% i årlig kostnad, fond B har 0,2%. Hvilket fond gir høyest avkastning etter kostnader?"

[Translation: Two funds are expected to deliver the same return before fees. Fund A has a 1.0% annual fee, Fund B has 0.2%. Which fund gives the highest return after fees?]

- Fond A [Fund A]
- Fond B [Fund B] ✓ **Correct**
- Likt [The same]
- Vet ikke [Don't know]

Scoring: 1 point if "Fond B" selected, 0 otherwise.
Concept tested: Understanding of fee impact on investment returns.

Question 10: Percentage and Recovery

Q10. Prosent [Percentage]

"En aksje faller 50% i verdi ett år. Hvor mye må den stige året etter for å komme tilbake til start?"

[Translation: A stock falls 50% in value one year. How much must it rise the following year to return to its starting value?]

- 50%
- 100% ✓ **Correct**
- 75%
- Vet ikke [Don't know]

Scoring: 1 point if "100%" selected, 0 otherwise.
Concept tested: Understanding of percentage losses and asymmetric recovery.

Literacy Score Construction

Total Score: Sum of correct answers across all six questions (range: 0–6).

Standardized Score: The literacy sum was standardized (mean = 0, SD = 1) to create `Financial_Literacy_Z` for use in regression analyses.

Missing Data Handling: Respondents who answered fewer than 5 of the 6 questions were excluded from literacy score calculation.

A.2 Subjective Confidence Measures

Two items assessed subjective confidence in financial knowledge. Responses were standardized and averaged to create the `Subjective_Confidence_Index`.

Question 13: Comparative Confidence

Q13. Relativ kompetanse [Relative Competence]

"Sammenlignet med andre norske sparere og investorer, hvor vil du rangere din kunnskap?"

[Translation: Compared to other Norwegian savers and investors, where would you rank your knowledge?]

- Nederste 10% [Bottom 10%] → Coded as 10
- 10–30% → Coded as 20
- 31–50% → Coded as 40
- 51–70% → Coded as 60
- 71–90% → Coded as 80
- Øverste 10% [Top 10%] → Coded as 90

Scoring: Responses were coded to percentile midpoints and standardized.

Question 14: Self-Rated Knowledge

Q14. Selvvurdert kunnskap [Self-Rated Knowledge]

"På en skala fra 0–6, hvor godt vil du rangere din egen finansielle kunnskap? Hvor 0 = Ingen kunnskap, og 6 = Ekspert"

[Translation: On a scale from 0 (no knowledge) to 6 (expert), how would you rate your own financial knowledge?]

Response Scale: 0 – 1 – 2 – 3 – 4 – 5 – 6
Scoring: Numeric response extracted and standardized.

Overconfidence Gap Construction

$$\texttt{Overconfidence_Gap} = \texttt{Subjective_Confidence_Z} - \texttt{Financial_Literacy_Z} \tag{15}$$

Where both components are standardized (mean = 0, SD = 1).

Interpretation:

- Positive values indicate overconfidence (subjective confidence exceeds objective knowledge)
- Negative values indicate underconfidence
- Zero indicates perfect calibration

A.3 Risk-Taking Scale

Six items assessed willingness to bear financial risk on a 7-point Likert scale (1 = Helt uenig [Strongly Disagree], 7 = Helt enig [Strongly Agree]). Items marked **(R)** were reverse-coded before averaging.

Question 11: Risk Attitudes

Q11_1. "Jeg er villig til å ta økonomisk risiko."

[Translation: I am willing to take financial risk.]
Scale: 1 (Helt uenig) – 7 (Helt enig)

Q11_2. (R) "Jeg foretrekker lav risiko selv om avkastningen blir lavere."

[Translation: I prefer low risk even if the return is lower.]
Reverse-coded: Response = 8− Original Response

Q11_3. "Jeg er komfortabel med midlertidige verditap hvis langsiktig avkastning er høyere."

[Translation: I am comfortable with temporary losses if long-term return is higher.]

Q11_4. (R) "Jeg blir stresset av svingninger i investeringene mine."

[Translation: I get stressed by fluctuations in my investments.]
Reverse-coded: Response = 8− Original Response

Q11_5. "Jeg aksepterer mulighet for tap for å oppnå høyere avkastning."

[Translation: I accept the possibility of loss to achieve higher returns.]

Q11_6. (R) "Jeg angrer ofte på investeringsbeslutninger når markedet svinger."

[Translation: I often regret investment decisions when the market fluctuates.]
Reverse-coded: Response = 8− Original Response

Risk Index Construction

1. Reverse-code items Q11_2, Q11_4, and Q11_6
2. Average all six items (range: 1–7)
3. Require at least 4 of 6 items answered to compute score
4. Standardize to create `Risk_Z` (mean = 0, SD = 1) for regression analyses

Internal Consistency: Cronbach's $\alpha = 0.665$ (calculated on complete cases), indicating reasonable reliability.

A.4 Portfolio Diversification

Question 15: Asset Class Holdings

Q15. Plassering av penger [Placement of Money]

"Hvor har du plassert penger? (flere valg mulig)"

[Translation: Where have you placed money? (multiple selections possible)]

- Kontanter (Cash, Brukskonto o.l.) [Cash (cash, current account, etc.)]
- Sparekonto [Savings account]
- BSU (Boligsparing for ungdom) [BSU (housing savings for young people)]
- Enkeltaksjer / aksjer direkte [Individual stocks / direct shares]
- Aksjesparekonto (ASK) [Equity savings account (ASK)]
- Indeksfond/ETF [Index funds/ETF]
- Aktivt forvaltede aksjefond [Actively managed equity funds]
- Obligasjonsfond [Bond funds]
- Kombinasjons- /balanserte fond [Combination/balanced funds]
- IPS / egenspart pensjon [IPS / individual pension savings]
- Eiendom – egen bolig [Real estate – primary residence]
- Eiendom – sekundær bolig/utleie [Real estate – secondary property/rental]
- Råvarer/edelmetaller (Gull, sølv etc.) [Commodities/precious metals (gold, silver, etc.)]
- Kryptovaluta [Cryptocurrency]
- Derivater (opsjoner/CFD) [Derivatives (options/CFD)]
- Annet (vennligst spesifiser) [Other (please specify)]

Scoring: Count of selected asset classes = `Diversification_Score` (theoretical range: 0–16; observed range: 0–13).
Interpretation: Higher scores indicate broader portfolio diversification across asset classes.

A.5 Market Participation

Question 3: Investment Activity

Q3. Investeringsaktivitet [Investment Activity]

"Har du investert de siste 12 måneder eller planlegger du å starte i løpet av de neste 12 måneder?"

[Translation: Have you invested in the last 12 months or do you plan to start within the next 12 months?]

- Ja, jeg investerer i dag [Yes, I invest today]
- Ikke nå, men planlegger å starte [Not now, but planning to start]
- Nei [No]

Scoring:

- "Ja" OR "Ikke nå..." → Coded as 1 (Participant)
- "Nei" → Coded as 0 (Non-participant)

Note: 100% of respondents in the final analysis sample ($N = 746$) were classified as participants, consistent with recruitment through investor communities. This universal participation precludes testing H1.

A.6 Demographic and Control Variables

Question 2: Age

Q2. Hvor gammel er du? [How old are you?]

Response options: Under 18 (excluded), 18–24, 25–34, 35–44, 45–54, 55–64, 65+.

Coding: Collapsed into three categories:

- **18–34:** Ages 18–24 and 25–34 combined
- **35–54:** Ages 35–44 and 45–54 combined
- **55+:** Ages 55–64 and 65+ combined

Baseline category in regressions: 18–34

Question 23: Gender

Q23. Hva er ditt kjønn? [What is your gender?]

Response options: Kvinne [Female], Mann [Male], Annet [Other], Ønsker ikke å oppgi [Prefer not to say].

Coding: Male, Female, Other ("Annet" and "Ønsker ikke å oppgi" combined).
Baseline category: Male
Sample composition: Approximately 84% male.

Question 24: Education

Q24. Hva er ditt høyeste fullførte utdanning? [What is your highest completed education?]

Response options: Grunnskole [Primary school], Videregående skole [High school], Bachelor [Bachelor's degree], Master [Master's degree], PhD (Doktorgrad), Ønsker ikke å oppgi [Prefer not to say].

Coding: Collapsed into two categories:

- **Low:** Grunnskole, Videregående skole
- **High:** Bachelor, Master, PhD

Baseline category: High.
Note: Respondents selecting "Ønsker ikke å oppgi" ($n = 8$) were coded as missing.

Question 25: Income

Q25. Omtrent hvor stor er din brutto årsinntekt (kr)? [Approximately how large is your gross annual income (NOK)?]

Response options: Under 300 000 kr; 300 000–499 999 kr; 500 000–749 999 kr; 750 000–999 999 kr; 1 000 000+ kr; Ønsker ikke å oppgi.

Coding: Collapsed into three categories:

- **Low:** Under 300 000 kr
- **Mid:** 300 000–749 999 kr
- **High:** 750 000 kr or more
- **Prefer Not:** Separate category

Baseline category: Low

A.7 Data Quality Control

Question 8: Attention Check (Literacy Block)

Q8. For å vise at du følger med, velg "Usant" her

[Translation: To show that you are paying attention, select "False" here.]

- Sant [True]
- Usant [False] ✓ **Correct response**
- Vet ikke [Don't know]

Data Quality Rule: Respondents who failed this attention check were excluded.

Question 29: Attention Check (Demographics Block)

Q29. For å vise at du følger med, velg alternativ 2 på denne linjen

[Translation: To show that you are paying attention, select option 2 on this line.]

- Alternativ 1 [Option 1]
- Alternativ 2 [Option 2] ✓ **Correct response**
- Vet ikke [Don't know]

Data Quality Rule: Respondents who failed this attention check were excluded. Combined with Q8, a total of $n = 254$ respondents were excluded for failing one or both attention checks.

A.8 Survey Administration

Platform: Qualtrics online survey software
Survey Length: Approximately 8–12 minutes
Incentive: Prize draw of 1 000 NOK (via Vipps); voluntary participation
Language: All questions administered in Norwegian
Data Collection Period: October 2025 – February 2026
Total Responses: 1,015 individuals initiated the survey
Analysis Sample: 746 respondents after exclusions

Ethical Approval: The study was conducted in accordance with BI Norwegian Business School's research ethics guidelines. Respondents provided informed consent before participating and were informed that participation was voluntary and anonymous.

References for Survey Design

The financial literacy questions were adapted from:

- Lusardi, A., & Mitchell, O. S. (2011). Financial literacy around the world: An overview. *Journal of Pension Economics and Finance*, *10*(4), 497–508.
- Van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, *101*(2), 449–472.

The risk tolerance scale draws on items from:

- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522–550.
- Grable, J., & Lytton, R. H. (1999). Financial risk tolerance revisited: The development of a risk assessment instrument. *Financial Services Review*, *8*(3), 163–181.

B Complete Regression Tables

B.1 H2: Financial Literacy and Diversification

Table 5: Diversification Model (Poisson Regression)

Poisson regression, robust SE (HC1)	/ *DV: Diversification Score (0–11 asset classes)*				/ $N = 683$		
Variable	**Coeff.**	**SE**	z	p	**CI Low**	**CI High**	**Sig.**
Intercept	1.746	0.039	44.990	<0.001	1.670	1.822	***
Age: 35–54 (ref: 18–34)	−0.012	0.031	−0.396	0.692	−0.073	0.048	
Age: 55+ (ref: 18–34)	0.077	0.058	1.319	0.187	−0.037	0.191	
Gender: Female (ref: Male)	−0.010	0.045	−0.214	0.831	−0.098	0.079	
Gender: Other (ref: Male)	−0.129	0.211	−0.611	0.542	−0.542	0.284	
Education: Low (ref: High)	−0.048	0.030	−1.588	0.112	−0.108	0.011	
Income: High (ref: Low)	0.140	0.042	3.351	0.001	0.058	0.221	***
Income: Mid (ref: Low)	0.074	0.041	1.799	0.072	−0.007	0.155	*
Income: Prefer Not to Say (ref: Low)	0.121	0.084	1.432	0.152	−0.045	0.286	
Financial Literacy (Z-score)	**0.031**	**0.015**	**2.028**	**0.043**	**0.001**	**0.060**	**

Notes: Coefficients are log-incidence rate ratios from a Poisson regression. Exponentiate to obtain incidence rate ratios (IRR). OLS with heteroskedasticity-robust standard errors (HC1). Reference categories: Age 18–34, Male, Education High, Income Low. Financial Literacy Z-score standardised (mean = 0, SD = 1). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.2 H3: Financial Literacy and Risk-Taking

Table 6: Risk-Taking Model (Unstandardized)

OLS regression, robust SE (HC1) / *DV: Risk Index (1–7 scale)* / $N = 683$							
Variable	**Coeff.**	**SE**	t	p	**CI Low**	**CI High**	**Sig.**
Intercept	5.163	0.092	56.097	<0.001	4.983	5.343	***
Age: 35–54 (ref: 18–34)	−0.118	0.079	−1.502	0.133	−0.272	0.036	
Age: 55+ (ref: 18–34)	−0.042	0.148	−0.283	0.777	−0.333	0.249	
Gender: Female (ref: Male)	−0.254	0.117	−2.172	0.030	−0.484	−0.025	**
Gender: Other (ref: Male)	−0.230	0.459	−0.501	0.617	−1.129	0.670	
Education: Low (ref: High)	0.028	0.073	0.382	0.703	−0.116	0.172	
Income: High (ref: Low)	0.149	0.103	1.445	0.148	−0.053	0.351	
Income: Mid (ref: Low)	0.102	0.098	1.033	0.302	−0.091	0.294	
Income: Prefer Not to Say (ref: Low)	−0.188	0.182	−1.031	0.303	−0.546	0.170	
Financial Literacy (Z-score)	**0.099**	**0.037**	**2.673**	**0.008**	**0.026**	**0.172**	*****
Model diagnostics							
N	683						
R^2	0.0313						
Adj. R^2	0.0184						
F-statistic	0.0099						
Residual Std. Error	0.8448						

Notes: OLS with heteroskedasticity-robust standard errors (HC1). DV is the Risk Index (mean of 6 Likert items, 1–7 scale; items Q11_2, Q11_4, Q11_6 reverse-coded). Reference categories: Age 18–34, Male, Education High, Income Low. Financial Literacy Z-score standardised (mean = 0, SD = 1). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.3 H4: Financial Literacy and Overconfidence

Table 7: Overconfidence Gap Model (Primary)

OLS, robust SE (HC1) / *DV: Overconfidence Gap* / $N = 683$							
Variable	**Coeff.**	**SE**	z	p	**CI Low**	**CI High**	**Sig.**
Intercept	0.061	0.111	0.550	0.582	−0.157	0.279	
Age: 35–54 (ref: 18–34)	−0.165	0.089	−1.862	0.063	−0.339	0.009	*
Age: 55+ (ref: 18–34)	−0.013	0.163	−0.077	0.939	−0.332	0.307	
Gender: Female (ref: Male)	−0.399	0.139	−2.878	0.004	−0.671	−0.127	***
Gender: Other (ref: Male)	−0.093	0.846	−0.110	0.913	−1.751	1.565	
Education: Low (ref: High)	−0.352	0.087	−4.047	<0.001	−0.522	−0.181	***
Income: High (ref: Low)	0.218	0.123	1.768	0.077	−0.024	0.460	*
Income: Mid (ref: Low)	0.104	0.114	0.915	0.360	−0.119	0.327	
Income: Prefer Not to Say (ref: Low)	0.414	0.260	1.594	0.111	−0.095	0.924	
Financial Literacy (Z-score)	**−0.779**	**0.039**	**−20.120**	**<0.001**	**−0.847**	**−0.695**	*****
Model diagnostics							
N	683						
R^2	0.357						
Adj. R^2	0.349						
F-statistic	45.114 ($p < 0.001$)						
Residual Std. Error	0.945						

Notes: DV is the Overconfidence Gap (standardised confidence index minus Financial Literacy Z-score). OLS with heteroskedasticity-robust standard errors (HC1). Reference categories: Age 18–34, Male, Education High, Income Low. Financial Literacy Z-score standardised (mean = 0, SD = 1). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 8: Overconfidence Residual Model (Robustness)

OLS, robust SE (HC1) / *DV: Overconfidence Residual* / $N = 683$						
Variable	**Coeff.**	**Std. Err.**	z	p**-value**	**CI Low**	**CI High**
Intercept	0.069	0.102	0.671	0.502	−0.132	0.269
Age: 35–54 (ref: 18–34)	−0.152	0.082	−1.862	0.063*	−0.312	0.008
Age: 55+ (ref: 18–34)	−0.012	0.150	−0.077	0.939	−0.305	0.282
Gender: Female (ref: Male)	−0.368	0.128	−2.878	0.004***	−0.618	−0.117
Gender: Other (ref: Male)	−0.086	0.779	−0.110	0.913	−1.612	1.441
Education: Low (ref: High)	−0.324	0.080	−4.047	<0.001***	−0.481	−0.167
Income: High (ref: Low)	0.201	0.114	1.768	0.077*	−0.022	0.424
Income: Mid (ref: Low)	0.096	0.105	0.915	0.360	−0.109	0.301
Income: Prefer Not to Say (ref: Low)	0.381	0.239	1.594	0.111	−0.088	0.851
Financial Literacy (Z-score)	**−0.053**	**0.036**	**−1.481**	**0.139**	**−0.123**	**0.017**
Model diagnostics						
N	683					
R^2	0.056					
Adj. R^2	0.043					
F-statistic	4.139 ($p < 0.001$)					
Residual Std. Error	0.870					

Notes: DV is the Overconfidence Residual (residual from regressing subjective confidence on literacy). OLS with heteroskedasticity-robust standard errors (HC1). Reference categories: Age 18–34, Male, Education High, Income Low. Financial Literacy Z-score standardised (mean = 0, SD = 1). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.4 H5: Overconfidence and Risk-Taking

Table 9: Overconfidence → Risk-Taking Model (H5)

OLS, robust SE (HC1) / *DV: Risk Index (1–7)* / $N = 683$							
Variable	**Coeff.**	**Std. Err.**	z	p**-value**	**CI Low**	**CI High**	**Sig.**
Intercept	5.190	0.094	55.312	<0.001	5.006	5.374	***
Age: 35–54 (ref: 18–34)	−0.109	0.077	−1.404	0.160	−0.261	0.043	
Age: 55+ (ref: 18–34)	−0.026	0.143	−0.182	0.856	−0.306	0.254	
Gender: Female (ref: Male)	−0.305	0.113	−2.710	0.007	−0.526	−0.084	***
Gender: Other (ref: Male)	−0.501	0.297	−1.685	0.092	−1.083	0.082	*
Education: Low (ref: High)	0.019	0.074	0.256	0.798	−0.125	0.163	
Income: High (ref: Low)	0.139	0.104	1.329	0.184	−0.066	0.343	
Income: Mid (ref: Low)	0.085	0.100	0.848	0.396	−0.111	0.280	
Income: Prefer Not to Say (ref: Low)	−0.221	0.189	−1.166	0.244	−0.592	0.150	
Overconfidence Gap	**0.125**	**0.028**	**4.445**	**<0.001**	**0.070**	**0.179**	*****
Model diagnostics							
N	683						
R^2	0.047						
Adj. R^2	0.034						
F-statistic	3.472 ($p = 0.0003$)						
Residual Std. Error	0.838						

Notes: DV is the Risk Index (1–7). Key predictor is the Overconfidence Gap. OLS with heteroskedasticity-robust standard errors (HC1). Reference categories: Age 18–34, Male, Education High, Income Low. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.5 H6: Mediation Analysis

Table 10: Mediation a-path: Literacy → Overconfidence (H6)

OLS, robust SE (HC1) / *DV: Overconfidence Gap* / $N = 683$							
Variable	**Coeff.**	**Std. Err.**	z	p**-value**	**CI Low**	**CI High**	**Sig.**
Intercept	0.061	0.111	0.550	0.582	−0.157	0.279	
Age: 35–54 (ref: 18–34)	−0.165	0.089	−1.862	0.063	−0.339	0.009	*
Age: 55+ (ref: 18–34)	−0.013	0.163	−0.077	0.939	−0.332	0.307	
Gender: Female (ref: Male)	−0.399	0.139	−2.878	0.004	−0.671	−0.127	***
Gender: Other (ref: Male)	−0.093	0.846	−0.110	0.913	−1.751	1.565	
Education: Low (ref: High)	−0.352	0.087	−4.047	<0.001	−0.522	−0.181	***
Income: High (ref: Low)	0.218	0.123	1.768	0.077	−0.024	0.460	*
Income: Mid (ref: Low)	0.104	0.114	0.915	0.360	−0.119	0.327	
Income: Prefer Not to Say (ref: Low)	0.414	0.260	1.594	0.111	−0.095	0.924	
Financial Literacy (Z-score)	**−0.779**	**0.039**	**−20.120**	**<0.001**	**−0.847**	**−0.695**	*****
Model diagnostics							
N	683						
R^2	0.357						
Adj. R^2	0.349						
F-statistic	45.114 ($p < 0.001$)						
Residual Std. Error	0.945						

Notes: This is the a-path in the mediation model (H6); same specification as Table B.3a. Model diagnostics are therefore identical to Table B.3a ($R^2 = 0.357$). OLS with heteroskedasticity-robust standard errors (HC1). Reference categories: Age 18–34, Male, Education High, Income Low. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 11: Mediation b+c'-path: Full Mediation Model (H6)

OLS, robust SE (HC1) / *DV: Risk Index* / $N = 683$							
Variable	**Coeff.**	**Std. Err.**	z	p**-value**	**CI Low**	**CI High**	**Sig.**
Intercept	5.156	0.088	58.738	<0.001	4.984	5.328	***
Age: 35–54 (ref: 18–34)	−0.069	0.075	−0.917	0.359	−0.216	0.078	
Age: 55+ (ref: 18–34)	−0.018	0.137	−0.128	0.898	−0.285	0.250	
Gender: Female (ref: Male)	−0.153	0.112	−1.365	0.172	−0.372	0.066	
Gender: Other (ref: Male)	−0.203	0.259	−0.783	0.434	−0.709	0.304	
Education: Low (ref: High)	0.109	0.070	1.551	0.121	−0.029	0.246	
Income: High (ref: Low)	0.089	0.100	0.895	0.371	−0.106	0.285	
Income: Mid (ref: Low)	0.063	0.093	0.678	0.498	−0.120	0.246	
Income: Prefer Not to Say (ref: Low)	−0.304	0.193	−1.580	0.114	−0.682	0.073	
Financial Literacy (Z-score)	**0.296**	**0.045**	**6.533**	**<0.001**	**0.207**	**0.384**	*******
Overconfidence Gap	**0.253**	**0.036**	**7.013**	**<0.001**	**0.182**	**0.324**	*******
Model diagnostics							
N	683						
R^2	0.104						
Adj. R^2	0.091						
F-statistic	See full model output						
Residual Std. Error	0.804						

Notes: Financial Literacy Z-score represents the c'-path (direct effect controlling for overconfidence). Overconfidence Gap represents the b-path (effect of overconfidence controlling for literacy). R^2 estimated from standardised β coefficients and pairwise correlations ($\hat{\beta}^{std}_{FL} \times r_{FL,Risk} + \hat{\beta}^{std}_{OC} \times r_{OC,Risk} \approx 0.104$); verify against statsmodels `rsquared` output. OLS with heteroskedasticity-robust standard errors (HC1). Reference categories: Age 18–34, Male, Education High, Income Low. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 12: Mediation c-path: Total Effect of Literacy on Risk (H6)

OLS, robust SE (HC1) / *DV: Risk Index* / $N = 683$

Variable	**Coeff.**	**Std. Err.**	z	p**-value**	**CI Low**	**CI High**	**Sig.**
Intercept	5.172	0.091	56.713	<0.001	4.993	5.351	***
Age: 35–54 (ref: 18–34)	−0.111	0.079	−1.406	0.160	−0.265	0.044	
Age: 55+ (ref: 18–34)	−0.021	0.146	−0.142	0.887	−0.307	0.266	
Gender: Female (ref: Male)	−0.254	0.115	−2.198	0.028	−0.480	−0.027	**
Gender: Other (ref: Male)	−0.226	0.462	−0.489	0.625	−1.132	0.680	
Education: Low (ref: High)	0.020	0.073	0.267	0.789	−0.124	0.163	
Income: High (ref: Low)	0.145	0.102	1.413	0.158	−0.056	0.345	
Income: Mid (ref: Low)	0.090	0.097	0.920	0.358	−0.101	0.281	
Income: Prefer Not to Say (ref: Low)	−0.199	0.183	−1.092	0.275	−0.557	0.158	
Financial Literacy (Z-score)	**0.100**	**0.037**	**2.730**	**0.006**	**0.028**	**0.172**	*****

Model diagnostics	
N	683
R^2	0.031
Adj. R^2	0.018
F-statistic	2.437 ($p = 0.0099$)
Residual Std. Error	0.845

Notes: Total effect of literacy on risk without controlling for overconfidence (c-path). Smaller than the direct effect $c' = 0.296$, confirming a suppression structure. OLS with heteroskedasticity-robust standard errors (HC1). Reference categories: Age 18–34, Male, Education High, Income Low. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 13: Indirect Effect — Bootstrap Summary (H6)

5,000 bootstrap samples / $N = 683$ / *Primary: Overconfidence Gap*

Specification	**Point Est.**	**Std. Err.**	**95% CI**	**Excl. Zero**	**Sig.**
Primary: Risk Index (unstandardised)	−0.195	—	[−0.255, −0.138]	YES	***
Primary: Risk Z (standardised)	−0.230	—	[−0.302, −0.163]	YES	***
Robustness: Risk Index (Residual)	−0.014	—	[−0.035, 0.006]	NO	
Robustness: Risk Z (Residual)	−0.017	—	[−0.041, 0.007]	NO	

Notes: Bootstrap percentile confidence intervals based on 5,000 samples. Primary specification uses the Overconfidence Gap; robustness uses the Overconfidence Residual. A negative indirect effect indicates a suppression mechanism: literacy reduces overconfidence, which would otherwise increase risk-taking. *** $p < 0.01$.

B.6 Additional Model Specifications

Table 14: Variance Inflation Factors (Multicollinearity Check)

Gap Model		**Residual Model**	
Variable	**VIF**	**Variable**	**VIF**
Income: High (ref: Low)	1.845	Income: High (ref: Low)	1.848
Age: 35–54 (ref: 18–34)	1.787	Age: 35–54 (ref: 18–34)	1.787
Financial Literacy (Z-score)	1.623	Income: Mid (ref: Low)	1.395
Overconfidence Gap	1.557	Education: Low (ref: High)	1.302
Income: Mid (ref: Low)	1.394	Gender: Female (ref: Male)	1.148
Education: Low (ref: High)	1.304	Age: 55+ (ref: 18–34)	1.131
Gender: Female (ref: Male)	1.148	Income: Prefer Not to Say (ref: Low)	1.077
Age: 55+ (ref: 18–34)	1.131	Financial Literacy (Z-score)	1.059
Income: Prefer Not to Say (ref: Low)	1.076	Overconfidence Residual	1.058
Gender: Other (ref: Male)	1.038	Gender: Other (ref: Male)	1.038

Notes: Variance Inflation Factors (VIF). Values above 10 indicate severe multicollinearity, while values above 5 warrant attention. All reported VIF values are well below these thresholds, indicating no multicollinearity concerns in either model specification.

Table 15: Pearson Correlation Matrix (Key Variables)

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
(1) Financial Literacy (Sum)	1.000					
(2) Subjective Confidence	0.263	1.000				
(3) Overconfidence Gap	−0.561	0.651	1.000			
(4) Overconfidence Residual	0.000	0.965	0.828	1.000		
(5) Risk Index	0.125	0.300	0.160	0.277	1.000	
(6) Diversification Score	0.107	0.141	0.037	0.117	0.010	1.000
N (pairwise)	746	710	710	710	713	746

Notes: Pearson correlations. The number of observations varies due to pairwise deletion. All correlations are computed using available cases.

B.7 Summary Statistics

Table 16: Descriptive Statistics for All Variables

Variable	N	**Mean**	**SD**	**Min**	**25%**	**Median**	**75%**	**Max**
Financial Literacy Score (0–6)	746	5.627	0.738	0.000	5.000	6.000	6.000	6.000
Financial Literacy (Z-score)	746	0.000	1.001	−7.634	−0.850	0.507	0.507	0.507
Subjective Confidence Index	710	0.000	0.921	−2.702	−0.581	0.297	0.708	1.951
Overconfidence Gap	710	−0.047	1.167	−3.441	−0.691	−0.184	0.666	3.843
Overconfidence Residual	710	0.000	0.889	−2.823	−0.702	0.176	0.587	2.078
Risk Index (1–7)	713	5.221	0.847	1.500	4.667	5.167	5.833	7.000
Risk Index (Z-score)	713	0.000	1.001	−4.395	−0.655	−0.064	0.723	2.101
Diversification Score (0–11)	746	6.889	2.457	1.000	5.000	7.000	9.000	15.000

Notes: Descriptive statistics computed on available cases (listwise deletion per variable). Z-scores are standardised to mean = 0 and SD = 1. Risk Index is the mean of six Likert items (Q11). Diversification Score is the count of asset classes held (0–11). The number of observations varies across variables due to item-level non-response.

C Robustness Check Summary

Table 17: Robustness Checks: H3 (Risk-Taking) and H4 (Calibration)

	H3: Risk-Taking		**H4: Calibration**		
	Fin. Literacy → Risk		*Fin. Literacy → OC Gap*		
Specification	β	p**-value**	β	p**-value**	N
Baseline	0.099***	0.008	−0.779***	<0.001	683
Gender: Male	0.126***	0.003	−0.747***	<0.001	628
Age: Young (18–34)	0.139***	0.002	−0.732***	<0.001	475
Age: Middle (35–54)	0.030	0.646	−0.871***	<0.001	232
Exclude Literacy = 6/6	0.042	0.605	−0.787***	<0.001	195
Exclude Outliers (Top 5%)	0.114***	0.001	−0.766***	<0.001	711
No Controls	0.115***	0.001	−0.719***	<0.001	710

Notes: Baseline model includes controls for age, gender, education, and income. Robust standard errors (HC1). Subgroup models exclude the stratifying variable from the control set. Outliers identified via Cook's distance (top 5%). Perfect literacy scores (6/6) excluded in the ceiling-effects sensitivity test. Coefficients are standardised β. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

D Additional Figures

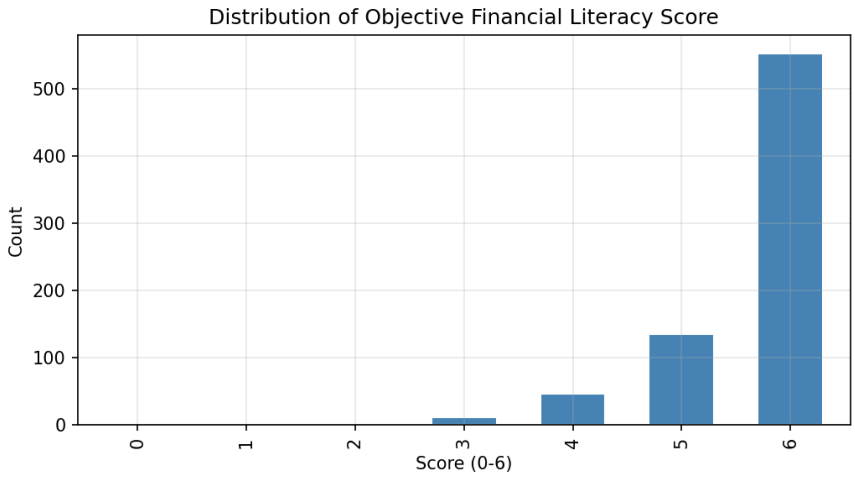

Figure 1: Distribution of Financial Literacy Scores

Note: $N = 746$. Scores range from 0–6. Mean = 5.63, showing concentration at the upper bound (ceiling effects).

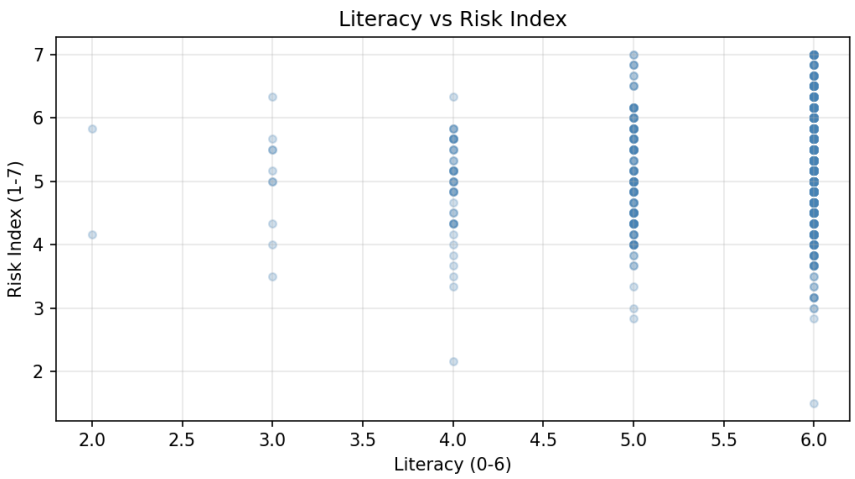

Figure 2: Financial Literacy vs. Risk-Taking

Note: Scatterplot shows a positive association ($r = 0.124$, $p < 0.001$). Each point represents one respondent.

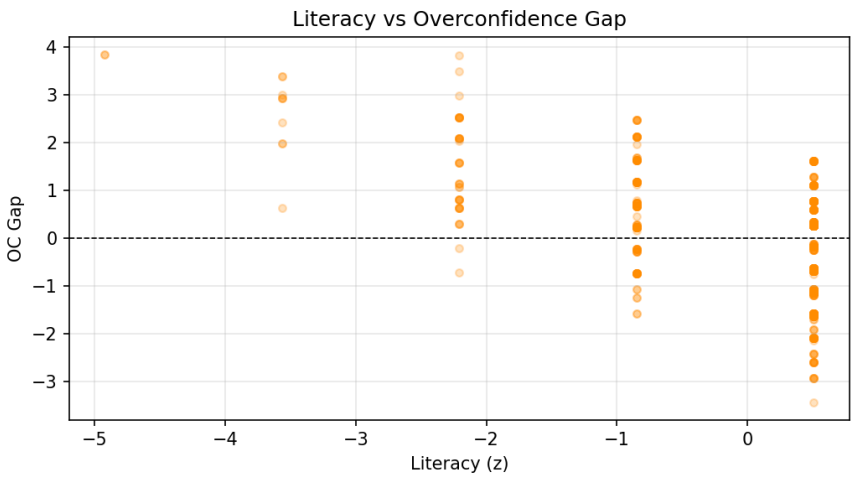

Figure 3: Financial Literacy vs. Overconfidence Gap

Note: Scatterplot shows a strong negative association ($r = -0.563$, $p < 0.001$). Horizontal line at zero represents perfect calibration.