informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–22
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Do People Understand the Benefit of Diversification?

Nicholas Reinholtz,[a] Philip M. Fernbach,[a] Bart de Langhe[b]

[a] Department of Marketing, Leeds School of Business, University of Colorado Boulder, Boulder, Colorado 80309; [b] Department of Marketing, Universitat Ramon Llull, ESADE, 08172 Sant Cugat des Valles, Spain

Contact: nicholas.reinholtz@colorado.edu, https://orcid.org/0000-0003-0696-1972 (NR); philip.fernbach@colorado.edu, https://orcid.org/0000-0003-1139-8619 (PMF); bart.delanghe@esade.edu, https://orcid.org/0000-0002-7913-5279 (BdL)

Received: June 12, 2018
Revised: September 26, 2019; October 6, 2020
Accepted: October 19, 2020
Published Online in Articles in Advance: March 9, 2021

https://doi.org/10.1287/mnsc.2020.3893

Abstract. Diversification—investing in imperfectly correlated assets—reduces volatility without sacrificing expected returns. Although the expected return of a diversified portfolio is the weighted average return of its constituent parts, the variance of the portfolio is less than the weighted average variance of its constituent parts. Our results suggest that very few people have correct statistical intuitions about the effects of diversification. The average person in our data sees no benefit of diversification in terms of reducing portfolio volatility. Many people, especially those low in financial literacy, believe diversification actually *increases* the volatility of a portfolio. These people seem to believe that the unpredictability of individual assets compounds when aggregated together. Additionally, most people believe diversification increases the expected return of a portfolio. Many of these people correctly link diversification with the concept of risk reduction but seem to understand risk reduction to mean greater returns on average. We show that these beliefs can lead people to construct investment portfolios that mismatch investors' risk preferences. Furthermore, these beliefs may help explain why many investors are underdiversified.

History: Accepted by Yuval Rottenstreich, decision analysis.
Funding: This project was supported by a grant from the Varieties of Understanding Project at Fordham University and The John Templeton Foundation.
Supplemental Material: The online appendix is available at https://doi.org/10.1287/mnsc.2020.3893.

Keywords: diversification • financial decision making • investing • numerical cognition • judgment and decision making

Introduction

Diversification is an essential component of effective long-term investing (Markowitz 1952, Sharpe 1964). Through diversification, an investor can reduce risk without sacrificing expected returns. Take, for example, a simple case with only two available assets (X and Y) and a finite investment horizon. Over this horizon, assume each asset has the same expected return (r) and the same risk, which can be operationalized as the variance in expected return (σ^2). Regardless of how an investor divides her investment between the two assets, the expected return is r, as the expected value of a linear combination is a linear combination of expected values (i.e., $E(aX + bY) = aE(Y) + bE(X)$). However, the risk faced by the investor depends on the relative allocation. If the investor allocates all her money to one of the assets, it will have a variance of σ^2. However, if the investor instead diversifies by splitting her investment equally between the two assets, the variance of her portfolio will be $\sigma^2 \times (1 + \rho_{XY})/2$, where ρ_{XY} is the correlation between X and Y. As long as the assets are not perfectly correlated (i.e., $\rho_{XY} \neq 1$), then diversifying reduces risk without a cost of decreased returns. For this reason, diversification has been called a *free lunch* (Campbell 2000; often attributed to Markowitz as well).

Despite the clear value of diversification and lack of cost, it is puzzling that many investors seem to be underdiversified (Kelly 1995, Barber and Odean 2000, Polkovnichenko 2005, Campbell 2006, Goetzmann and Kumar 2008, von Gaudecker 2015, Campbell et al. 2018). Past work identifies several factors that can contribute to underdiversification including preference for locally owned stocks (Cooper and Kaplanis 1994, Huberman 2001), preference for employer stock (Benartzi 2001, Mitchell and Utkus 2003), transaction costs (Brennan 1975, Merton 1987, Nieuwerburgh and Veldkamp 2010), and misapplied heuristics (Benartzi and Thaler 2001, Statman 2020). In this work, we explore a more basic question: Do people actually understand the benefit provided by diversification? Our data suggest that many people do not.

We present 13 studies that assess whether people understand the effect of diversification and whether a lack of understanding could have welfare-reducing downstream consequences. In each study, we collected a measure of financial literacy. A lack of individual financial literacy is often mentioned as a factor contributing to the global financial crisis of 2007–2008 (Klapper et al. 2012). In recent years, many countries and financial institutions have made major investments to enhance financial literacy (e.g., U.S. Financial Literacy and Education Commission, BBVA Center for Financial Education and Capability). Prior work has connected financial literacy to underdiversification directly (Guiso and Jappelli 2009, von Gaudecker 2015), and it has been shown to correlate with other desirable financial behaviors (e.g., retirement planning: Lusardi and Mitchell 2007; stock market participation: Christelis et al. 2010, Almenberg and Dreber 2015). Including a measure of financial literacy in our studies allows us to see whether, and—if so—how, people's beliefs about diversification may vary based on their financial sophistication.

Our results reveal two common biases in people's beliefs about diversification. First, the average participant in our studies sees no benefit of diversification in terms of reducing portfolio volatility. In fact, many people believe that diversification increases, rather than decreases, the volatility of a portfolio. This belief is particularly common among people low in financial literacy, which could help explain why underdiversification seems to be a greater problem for this group (von Gaudecker 2015). Second, most people believe that diversification increases the mean[1] return of a portfolio. We find this belief is prominent even among those *high* in financial literacy and seems to result from a superficial understanding of diversification and investment risk. Most people in our studies correctly answer a financial literacy question about diversification reducing risk. However, they seem to believe *reduced risk* manifests—at least in part—as increased mean returns. Although this is inconsistent with the traditional understanding of risk as variance in finance (Sharpe et al. 1998), it provides a façade of logic, as increasing the mean of a return distribution *ceteris paribus* would reduce the likelihood of losing money. More generally, people find it easier to think about the central tendency of a distribution than its dispersion, so the fact that many people think about diversification in terms of the first (versus second) moment seems reasonable through the lens of cognitive efficiency (Peterson and Beach 1967, Goldstein and Taleb 2007, Obrecht et al. 2007, de Langhe et al. 2014, Reinholtz 2015).

In Studies 1A–1E, 2A, 2B, and 3, participants make forecasts for diversified and undiversified portfolios. Most people in our large experimental sample exhibit at least one of the previously described biases (e.g., 83% in Studies 1A–1E). In the first set of studies (1A–1E), participants make forecasts using a distribution builder tool, allowing us to assess the distribution of possible outcomes they expect in terms of both expected returns (e.g., the means of the expressed distributions) and volatility (e.g., the standard deviations of the expressed distributions). Studies 2A and 2B instead use more traditional elicitation methods (e.g., point estimates and confidence intervals). Study 3 uses an incentive-compatible choice-based task. All eight studies provide consistent results.

In Studies 4A–4C, we explore the potential downstream consequences of these beliefs using a portfolio construction task. We show that many people construct a less diversified portfolio for an investor explicitly looking for risk reduction while providing a more diversified portfolio for an investor seeking higher returns. Studies 5 and 6 begin to examine the psychological processes that underlie these beliefs and behaviors.

For all studies, the initial target sample size was determined in advance. We had no prior estimate of effect size, so following Nelson (2014), we targeted a sample size in our first study that seemed reasonably large and adjusted the target sample size in later studies based on cost considerations and participant availability. In Studies 1A (50%), 1C (50%), 2B (100%), 4A (100%), and 6 (100%) we collected one additional batch of data after the initial target was reached (sample size increase reported in parentheses). All manipulations, data exclusions, and measured variables are reported. For the analyses presented in the main text, we exclude participants using consistent criteria. Analyses without exclusions are presented in the supplement and, unless otherwise noted, yield similar results both substantively and statistically. Data, code, and materials for all studies are available on Open Science Framework (OSF) (https://osf.io/hnj5y/).

Studies 1A–1E: Distribution Builder Studies

We conducted five experiments that share the same critical features and yield similar results.[2] These studies are presented here in aggregate for brevity.[3] In each study, participants made forecasts for both a diversified portfolio (stocks from 10 different companies) and an undiversified portfolio (stock from a single company) using a graphical, *balls-in-bins* tool (Sharpe et al. 2000, Delavande and Rohwedder 2008, Goldstein et al. 2008, Goldstein and Rothschild 2014, André et al. 2017, Long et al. 2018). After making forecasts for both portfolios, we collected a measure of financial literacy (Fernandes et al. 2014).

We compare the forecasts for both portfolios in terms of volatility and expected return. We further look to see whether differences between the forecasts for the two portfolios vary based on each participant's financial literacy.

Method

Participants. In total, 1,825 unique participants made forecasts for the diversified and undiversified portfolios. One hundred seventy-six participants were undergraduate business students at the University of Colorado Boulder and participated for course credit (Study 1B). The remainder (1,649) were recruited through

Amazon Mechanical Turk (AMT) and participated for monetary compensation ($1).[4] Because of technological issues, 182 of the 1,825 unique participants (identified by their AMT worker IDs) completed multiple versions of Study 1. We do not remove these participants but control for this overlap in the analysis we report. If we instead remove repeat participants, the results are essentially unchanged (see Table S2 in the online appendix). We restrict the presented analysis to participants who correctly passed an attention check question directly related to the dependent measure (identifying the order in which they evaluated the 1-stock and 10-stock portfolios) and who scored better than chance on the financial literacy measure (better than 4 of 13). Neither of these choices has a substantive effect on the results we report in these or future studies (in which we use the same exclusion criteria) but allow more accurate estimation of model parameters and effect sizes.[5] Results without data exclusions for all studies are provided in the online appendix (e.g., Table S3 for Studies 1A–1E). After the exclusions, we are left with 1,500 unique participants.

Distribution Builder. We measure participants' beliefs about future portfolio performance using a tool we call a distribution builder (see Appendix A for an example; note that Sharpe et al. 2000 use the same term for a different but related tool). Our distribution builder requires participants to assign 100 balls to different uniformly spaced bins representing possible portfolio values after one year. Participants assign balls to the different bins by clicking + or − buttons that add or subtract balls accordingly (all bins start empty). Participants are told to assign balls to bins based on how likely they think each portfolio value is and that they should assign the most balls to the bin for the portfolio value they think is most likely. They are also told that ratios matter: If they assign 20 balls to one portfolio value and 10 balls to another, it means they think the first portfolio value is twice as likely as the second. Once participants assign all 100 balls to the different outcomes, they can submit their distribution. In effect, the distribution builder yields a histogram of each participant's subjective beliefs about the future value of the stock portfolio.[6]

Prior research suggests this type of tool is capable of eliciting knowledge about distributions more effectively than more traditional measures (e.g., stated confidence ranges; Goldstein and Rothschild 2014). Consistent with this, participants seem to understand and use the distribution builder competently after a brief training tutorial, and we find consistent results across all studies.

Procedure. After consenting to participate, we gave participants a tutorial on how to use the distribution builder (see supplement for details). Before moving on to the critical measures, participants were required to demonstrate an understanding of the distribution builder tool by passing a simple test in which they translated a text-based description of a distribution to a graphical representation using the tool.

Participants then used the distribution builder tool to make forecasts for two different stock portfolios (one at a time in random order). For each portfolio, participants were given information about the composition and value of the portfolio and then asked what they thought the value of that portfolio would be in exactly one year. In each study, one portfolio consisted of stocks from 10 different companies and the other consisted of stock from a single company. Although we did not use the words *diversified* or *undiversified* in the descriptions we gave to participants, we will use those terms to describe the portfolios in our analysis. Both portfolios had the same initial value in all studies.

After participants made forecasts for all the portfolios, we administered an attention check where we asked participants to recall the order in which they evaluated the portfolios. Following this, participants completed a 13-item financial literacy measure (Fernandes et al. 2014).[7]

Results

Analysis Strategy for Distribution Builder Responses. Each response to the distribution builder task consisted of a vector with 100 elements representing the values of the bins to which each ball was assigned. We linearly transformed these values to correspond to percent changes from the initial portfolio value to allow for better interpretation of the results and aggregation across studies. We then computed distributional statistics on each of the vectors with mean (which represents the average expected value of the portfolio) and standard deviation (which represents the volatility of the portfolio) being of primary interest. These computed distributional statistics serve as the dependent variables in the following analyses. As a test of robustness, we also computed other distributional statistics including medians and implied confidence intervals.

Model Specification. We perform the presented analysis using linear regressions with standard errors clustered by study and participant and fixed effects for studies (*lfe* package for R; Gaure 2018). In the online appendix, we present a similar analysis using mixed-effect regressions (*lme4* package for R; Bates et al. 2015), which yield similar results (Table S5). We estimate regression coefficients for portfolio type (coded: +0.5 = diversified, −0.5 = undiversified), financial literacy (mean centered), and their interaction.

Overview of Results. Results for volatility (standard deviation of expressed distribution; left panel) and expected returns (mean of expressed distribution; right panel) are shown in Figure 1. This illustrates two biases in people's forecasts: First, about half of the participants expect diversification to increase the volatility of returns. Second, most participants expect diversification to increase mean returns. We describe these results in detail here.

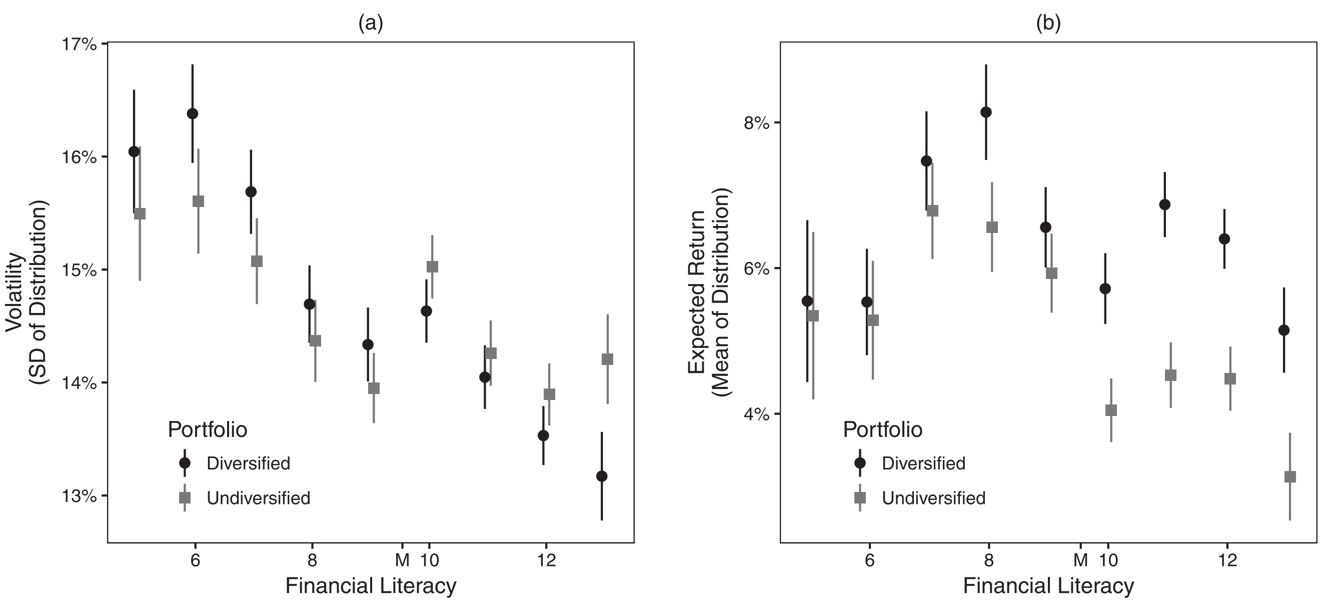
Analysis of Volatility. Because financial literacy was mean-centered before analysis, the regression coefficient for diversification can be interpreted as the main effect of going from an undiversified (1-stock) portfolio to a diversified (10-stock) portfolio. We find that on average participants do not expect diversification to reduce the volatility of returns ($\hat{\beta}_{\text{Diversified}} = 0.00\%, t = 0.00, p = 0.999$).[8] Consistent with this, 49% of participants expressed a more disperse distribution of potential returns for the diversified portfolio than the undiversified portfolio. This test—that of equal variance—is a very conservative benchmark: If the hypothetical stocks in the diversified portfolio had equal variance and had an average correlation of 0.5, the expected standard deviation of the diversified portfolio is approximately 75% of that for the undiversified portfolio.[9] Using this benchmark, 89% of participants failed to appreciate the magnitude by which diversification reduces volatility.

We also find a significant interaction between financial literacy and portfolio type ($\hat{\beta}_{\text{Interaction}} = -0.21\%$, $t = -5.32$, $p < 0.001$). These results imply that people low in financial literacy are likely to believe that diversification actually *increases* the volatility of returns. Using a *floodlight* approach (Johnson and Fay 1950, Spiller et al. 2013), we find that participants who scored less than 8.75 (of 13) on the financial literacy measure were significantly ($p < 0.05$) more likely to expect greater volatility from the diversified portfolio. For context, Fernandes et al. (2014) surveyed a nationally (United States) representative sample using the same financial literacy measure and found the average level of financial literacy to be 7.81, which suggests the modal American is likely to believe that diversification increases volatility.[10]

Analysis of Expected Returns. The diversification coefficient reveals a positive main effect ($\hat{\beta}_{\text{Diversified}} = 1.42\%$, $t = 12.94$, $p < 0.001$). This implies that on average participants expect the diversified portfolio to have a greater return—operationalized here as the arithmetic mean of possible returns—compared with the undiversified portfolio.[11] Consistent with this, 59% of participants expressed a distribution of returns with a higher mean for the diversified portfolio.

We also find a significant interaction between financial literacy and portfolio type ($\hat{\beta}_{\text{Interaction}} = 0.27\%$, $t = 3.29$, $p = 0.001$). This suggests that people *high* in financial literacy exhibit this bias to a greater degree than those low in financial literacy.

Figure 1. Studies 1A–1E: Judgments of Volatility (Standard Deviation of Expressed Return Distribution) and Expected Return (Mean of Expressed Return Distribution)



Notes. Participants are binned by their financial literacy score, and the points reflect the mean value for each bin. Error bars show model-free standard errors (standard deviation of value within each bin divided by the square root of number of observations in the bin). A plot showing linear best fit lines is available in the online appendix (Figure S2).

Discussion

A potential explanation for the observed results is that participants might believe the two portfolios are composed of qualitatively different assets.[12] However, we do not believe this explanation can fully explain the results. For example, in Study 1E, we find consistent results after explicitly telling participants that the stocks in both portfolios were *randomly selected* from the Financial Times Global 500 (see Figure S1 in online appendix for coefficient estimates). In this case, because the stocks in both portfolios are randomly selected from the same set, both portfolios should have the same expected return, and the diversified portfolio should have lower volatility.[13]

Although we originally did not collect any demographic information from participants, previous research has found that demographic variables (e.g., gender) are correlated with both financial literacy and investment behavior (Almenberg and Dreber 2015). To explore how demographic variables relate to the perceived consequences of diversification, we reached out again to all of the participants from AMT who had previously completed a version of Study 1: Four hundred eighteen (of 1,649) participants completed the follow-up study (payment = $0.25). In this sample, the median reported age was 40, the median reported income bucket was $50,000–$75,000 per year, 45% identified as female, 64% reported being college graduates, 68% reported being employed, and 61% reported participating in the stock market (owning individual stocks and/or a mutual fund/exchange traded fund (ETF)).

In short, although our measure of financial literacy is correlated with reported gender[14] ($r = -0.14$), age ($r = 0.30$), income ($r = 0.20$), education ($r = 0.33$), and market participation ($r = 0.27$), controlling for these variables (and their interactions with the diversification manipulation) does not substantively change the focal coefficient estimates from the model. Although not directly related to our conclusions, we note that women were more bullish for both portfolios (in terms of expected return) and particularly so for the diversified portfolio. Also, participants with higher levels of education indicated that diversification would reduce volatility to a greater degree. Full details of this analysis and the follow-up study are presented in the online appendix (Tables S7 and S8).

Studies 2A and 2B: Other Elicitation Methods

Across the five distribution builder studies, we found two consistent effects: First, participants—on average—did not expect diversification to lower portfolio volatility. In fact, participants with low financial literacy actually expected diversification to increase the volatility of a stock portfolio. Second, participants—on average—believed diversified portfolios would yield higher returns than undiversified portfolios. In Studies 2A and 2B, we explore the robustness of these effects using the same general procedure but different elicitation methods to obtain participants' forecasts.

Method

Participants. We recruited 400 people for Study 2A (payment = $0.50) and 360 people for Study 2B (payment = $0.50) from AMT to participate in these studies.

Procedure. In both studies, we used portfolios containing stocks from real companies to further address the potential role of inferences about qualitative differences in portfolio construction. The diversified portfolio featured stocks from four companies, with equal investment in each. The undiversified portfolio contained only one of these companies, randomly selected for each participant. The four companies used in Study 2A were as follows: DowDuPont Chemicals, Nike, Comcast, and Walmart. The four companies used in Study 2B were Facebook, PetSmart, Dow Chemicals, and Whole Foods.[15] These companies were selected because they are well known and have low historical intercorrelations.

In Study 2A, participants were shown the descriptions for both the single-stock (undiversified; labeled Portfolio A) portfolio and the multistock (diversified; labeled Portfolio B) portfolio. We then asked participants which portfolio they thought will "have a greater value at the end of the next year" (six-point scale, anchored by Definitely Portfolio A and Definitely Portfolio B with Very Likely... and Likely... representing the medial options) and which portfolio they thought will "have a more predictable value at the end of next year?" (six-point scale with same labels as the other question). These questions were intended to tap into expected return and volatility, respectively. Although neither scale included numeric anchors for participants, we code the responses as −2.5, −1.5,...,+1.5, +2.5 for the analysis, so that an average value of zero would reflect no perceived difference between the portfolios and a positive value would reflect favoring the diversified portfolio on the given measure.

In Study 2B, participants were shown the portfolios sequentially in random order. For both portfolios, they were asked: "What do you think the value of this portfolio will be in exactly one year?" and to provide a 90% confidence interval around that prediction using language adapted from prior research (Soll and Klayman 2004; see online appendix for wording). The initial value of both portfolios was $10,000.

Following these dependent measures, participants in both studies completed the same financial literacy measure used in Studies 1A–1E.

Results

We analyze both studies with the same exclusion criteria used in Studies 1A–1E: scoring above chance on the financial literacy measure. Because Study 2B included unbounded responses, we additionally removed extreme outliers from the presented analysis (three interquartile range (IQR) criteria; Tukey 1977). This resulted in sample sizes of 364 for Study 2A and 294 for Study 2B. Results without the exclusions are presented in the online appendix.

Study 2A. As shown in Figure 2, participant responses to the scale questions are consistent with key results from Studies 1A–1E. For the volatility measure, participants' beliefs about the relative predictability of the two portfolios depended on their financial literacy (linear regression: $\hat{\beta}_{\text{FinLit}} = 0.17$, $t = 4.42$, $p < 0.001$): Participants low in financial literacy believed the diversified portfolio would have a less predictable value in one year (Johnson-Neyman point = 6.34; see left panel of Figure 2). However, unlike Studies 1A–1E, on average, participants did believe the diversified portfolio would be more predictable ($M = 0.28$, $t = 3.05$, $p = 0.002$).

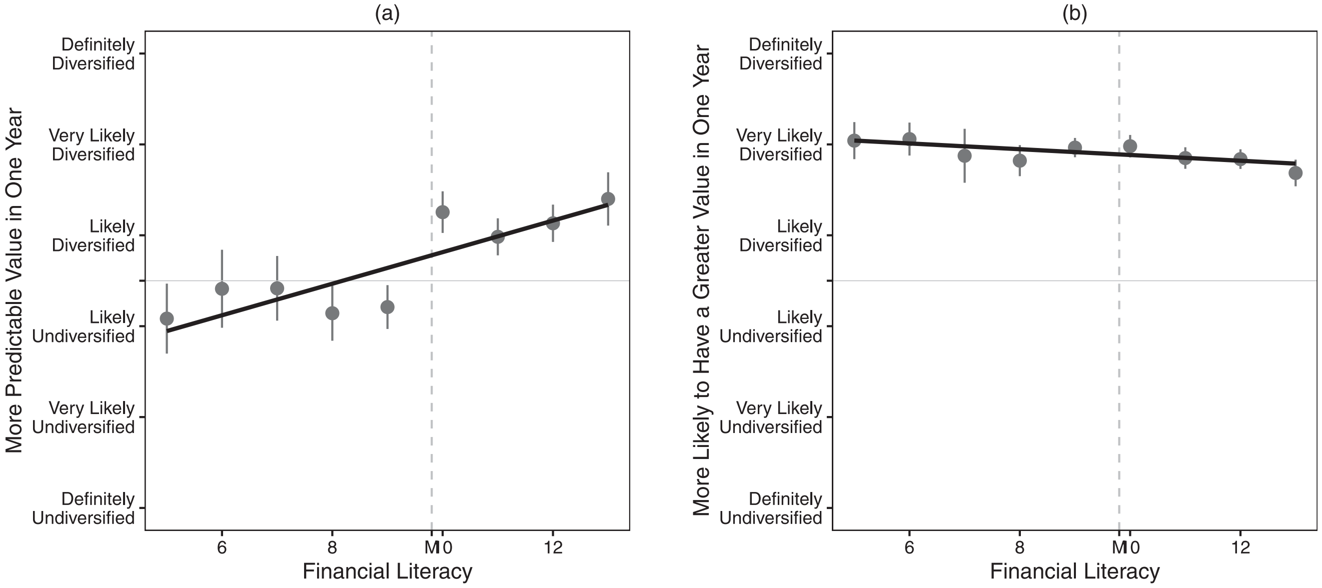
For the return measure, on average participants believed the diversified portfolio was more likely to increase in value over the next year ($M = 1.39$, $t = 28.72$, $p < 0.001$). However, unlike Studies 1A–1E, this belief was not moderated by financial literacy ($\hat{\beta}_{\text{FinLit}} = -0.03$, $t = -1.50$, $p = 0.14$; see right panel of Figure 2).

Study 2B. As shown in Figure 3, the point estimate and 90% confidence interval questions also yielded similar results to Studies 1A–1E. In this study, there was a significant main effect on the volatility (90% confidence interval) measure. On average participants gave a wider 90% confidence interval for the diversified portfolio (M = \$4,783) than the undiversified portfolio (M = \$3,865; $t = 6.72$, $p < 0.001$), which implies greater volatility for the diversified portfolio. As in previous studies, this bias was greater for those low in financial literacy (regression on difference score: $\hat{\beta}_{\text{FinLit}} = -\$168$, $t = -2.94$, $p = 0.004$; see left panel of Figure 3). For the return measure, on average, participants thought the diversified portfolio would have a higher value in one year (M = \$11,955) compared with the undiversified portfolio (M = \$11,162; $t = 7.40$, $p < 0.001$). In this study, unlike the previous, this difference was attenuated for participants high in financial literacy ($\hat{\beta}_{\text{FinLit}} = -\$125$, $t = -2.79$, $p = 0.006$; see right panel of Figure 3).

Discussion

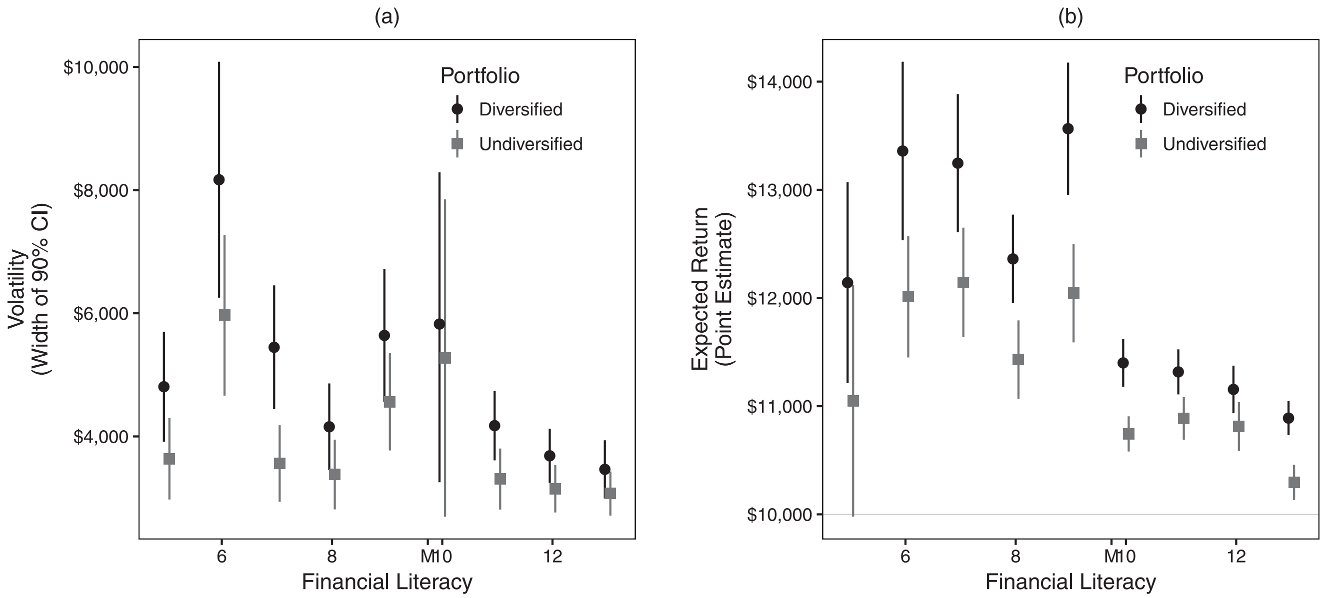
Studies 2A and 2B provide evidence that the errors we find using the distribution builder (Studies 1A–1E)—that many people believe diversification increases volatility and most people believe diversification increases expected return—are robust to other elicitation methods. We note, however, that the use of more commonplace language to elicit forecasts entails a tradeoff. Although these methods likely have the benefit of being easy to understand by participants, this comes with the cost of imprecision. A specific instantiation of this concern involves whether we are eliciting beliefs about mean versus median returns: As we describe in

Figure 2. Study 2A: Judgments About Whether the Diversified or Undiversified Portfolio Would Have a More Predictable Value (Left) and Would Be More Likely to Increase in Value (Right)



Note. Linear fits are shown in this plot. Points represent the mean judgments binned by financial literacy, with error bars indicating model-free standard errors (standard deviation of value within each bin divided by the square root of number of observations in the bin).

Figure 3. Study 2B: Volatility (Width of 90% Confidence Interval) and Return (Expressed as a Specific Number)



Notes. Participants are binned by their financial literacy score, and the points reflect the mean value for each bin. Error bars show model-free standard errors (standard deviation of value within each bin divided by the square root of number of observations in the bin). A plot showing linear best fit lines is available in the online appendix (Figure S3).

endnote 1 and show using a simulation in the online appendix, a diversified portfolio—with some reasonable assumptions—will have a higher median return than an undiversified portfolio, despite having the same mean return.

Although it is logically possible that participants are interpreting the measures in Studies 2A and 2B to be about the median of the return distribution and thus giving *correct* answers, we think this is unlikely. This would require (i) appreciating that return distributions tend to be skewed, (ii) appreciating that diversification reduces this skew (the central limit theorem), (iii) understanding means and medians and appreciating that skewness leads to a dissociation between the two, and (iv) a decision to report forecasts based on the median instead of the mean.[16]

More importantly, the median account would not apply to Studies 1A–1E, in which we observe a similar pattern of results. In these studies—because we elicit participants' beliefs about the entire distribution of possible outcomes—we explicitly calculate the expected *mean* return, which should not differ between the two portfolios. This is an important advantage of the distribution builder methodology: It allows us to identify a bias—expecting higher mean returns—that is difficult to conclusively identify using other methods. The measures in Studies 2A and 2B provide convergent evidence consistent with this bias.

Study 3: Incentive Compatibility

In Study 3, we seek to further explore the robustness of these findings in two ways: First, we simplify the context even further and compare forecasts for portfolios with either one (undiversified) or two (more diversified) stocks. Second, we incentivize participants' forecasts, and assess whether the presence of an incentive attenuates the previously documented effects.

Method

Participants. We targeted 400 participants from AMT and 399 completed the experiment (base payment = $0.75). Twenty-five participants scored below chance on the financial literacy measure and were removed from the presented analysis (results without exclusions are presented in the online appendix).

Procedure. On the first page of the survey, we told participants they would have a chance to earn a $1 bonus payment based on their answers in the study. We then described two people (A and B) who were planning to make a $10,000 investment in the stock market on August 5: Person A will invest all $10,000 in a single stock (undiversified). Person B will divide her investment between two stocks with $5,000 in each (diversified). The two stocks in Person B's portfolio were Oracle Corporation (ORCL) and Cisco Systems (CSCO).[17] The one stock in Person A's portfolio was randomly assigned (between participant) to be either Oracle Corporation or Cisco Systems.

We then asked participants two questions, similar to those asked in Study 2A. The first was intended to assess judgments of expected return: "Whose portfolio—Person A or Person B—do you think will have a *greater* value at the end of one month (on September 5th)?"

The second was intended to assess judgments of volatility: "Whose portfolio—Person A or Person B—do you think will have a *more predictable* value at the end of one month (on September 5th)?" Both choices were binary: Person A or Person B.

For each participant, we randomly assigned *one* of these two questions to be incentivized. This was clearly highlighted in bold font (see online appendix for stimuli). Participants were explicitly told how the correct answer would be calculated and were told they would receive a $1 bonus payment if they provided the correct answer.[18] Finally, participants completed the financial literacy scale used in previous studies.

Results

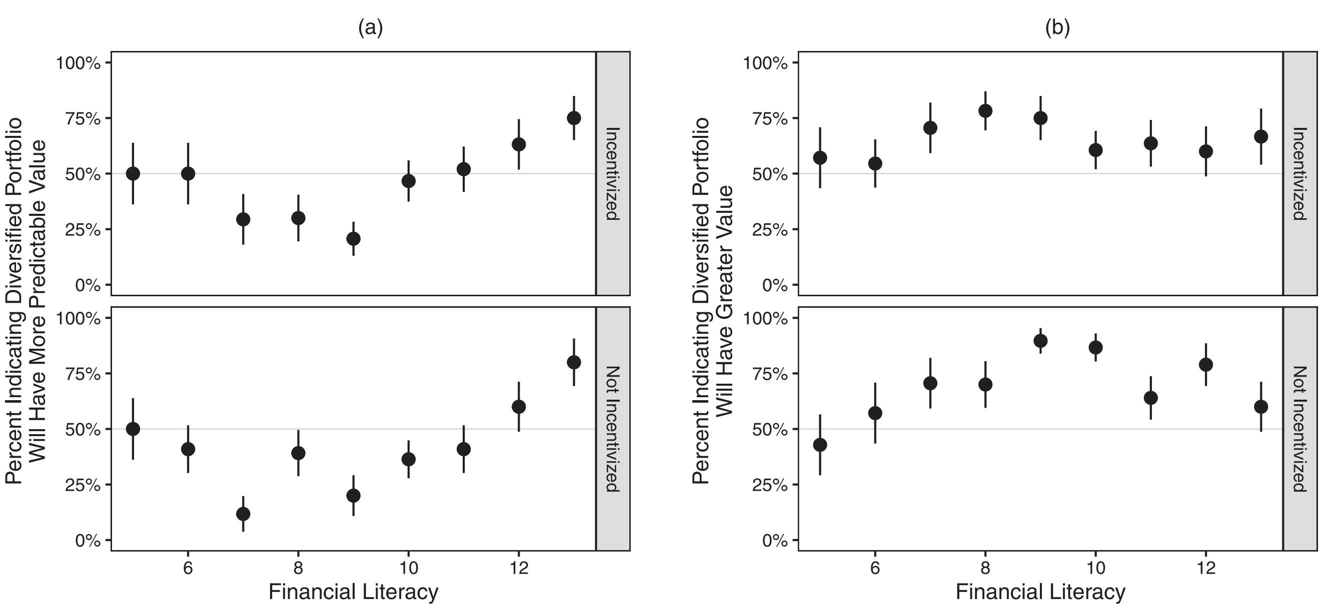
Participants on average believed the diversified portfolio would be more volatile than the undiversified portfolio: 57% of participants believed the single stock (undiversified) portfolio would have a more predictable value at the end of the month compared with the two stock (diversified) portfolio ($\chi^2 = 7.23$, $p = 0.007$). There was little evidence that the incentive affected this forecast: 55% of participants believe the single stock portfolio would have a more predictable value when the question was incentivized compared with 59% when it was not (logistic regression: $\hat{\beta}_{\text{Incentive}} = -0.09$, $z = -0.85$, $p = 0.40$). However, consistent with prior studies, the belief that a diversified portfolio will be less predictable was moderated by financial literacy (logistic regression: $\hat{\beta}_{\text{FinLit}} = 0.16$, $z = 3.54$, $p < 0.001$) as shown in the left panels of Figure 4.

Consistent with prior studies, we also find that participants on average believed the diversified portfolio would have a higher value than the undiversified portfolio: 68% of participants believed the two stock (diversified) portfolio would have a greater value at the end of the month compared with the single stock (undiversified) portfolio ($\chi^2 = 50.92$, $p < 0.001$). Again, we find little evidence that the incentive affected this forecast (incentivized = 65% versus not incentivized = 72%; logistic regression: $\hat{\beta}_{\text{Incentive}} = -0.16$, $z = -1.40$, $p = 0.16$), nor do we find evidence that financial literacy affected this forecast (logistic regression: $\hat{\beta}_{\text{FinLit}} = 0.05$, $z = 1.04$, $p = 0.30$).

Discussion of Forecasting Studies

Data from the forecasting studies (1A–1E, 2A, 2B, and 3) suggest that the average person does not have a good sense for how diversification affects the performance of a stock portfolio. In fact, the average participant in our studies does not perceive *any* benefit of diversification in terms of reducing portfolio volatility. Using different elicitation methods, we find convergent evidence that people low in financial literacy tend to believe that a diversified portfolio will have greater volatility compared with an undiversified portfolio. Additionally, the average

Figure 4. Study 3: Proportion of Participants Indicating the Diversified Portfolio Will Have a More Predictable Value (Left) and Will Have a Greater Value (Right) at the End of a Month



Notes. Plots are split (vertically) showing participants for whom the correct answer to this question was incentivized (top) and not incentivized (bottom). Participants are binned by their financial literacy score, and the points reflect the mean value for each bin. Error bars show model-free standard errors (standard deviation of value within each bin divided by the square root of number of observations in the bin).

person expects diversification to increase the mean return of portfolio. This belief seems more prominent among those *high* in financial literacy using the distribution builder elicitation method, but this moderation effect is not replicated using the other elicitation methods.

One may wonder why we include financial literacy as a focal variable in our analysis, particularly because one should expect financial errors to be more likely among those with low financial literacy. As concern over financial literacy—and effort on behalf of governments and financial institutions to improve financial literacy—has become common, we believe it is important to go beyond knowing whether people are getting questions on financial literacy questionnaires correct. Although it is true that we document erroneous beliefs in those with low financial literacy, we also document a bias in those with high financial literacy: expecting a higher mean return from diversification. On the surface, this belief may not seem problematic, as an erroneous expectation of higher returns should still lead high financial literacy investors to a desirable outcome (greater diversification). However, inflated expectations come with a cost: If you expect your diversified portfolio to always beat less-diversified investments, you will be disappointed. We have spoken with a firm that specializes in providing investors with diversified portfolios, and they see this as a consequential problem: When their diversified portfolios have lower returns than the S&P 500, they report receiving many complaints from their high net worth clients.

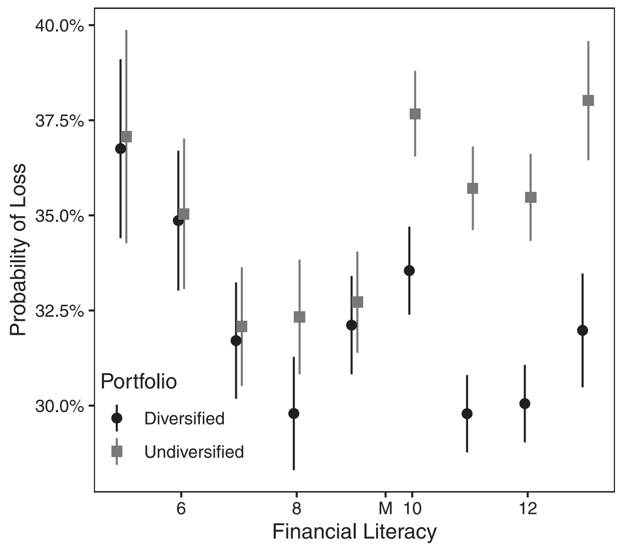
Closer inspection of the financial literacy measure we use in our studies reveals a puzzle. The measure contains a question related to diversification.[19] Most participants—82% in Studies 1A–1E—correctly answered that an investor who "spreads his money among different assets" causes "the risk of losing money" to decrease.[20] However, when asked to forecast the future value of a diversified and an undiversified portfolio moments before, many of these same participants expressed a belief that diversification increases the variance of the return distribution, thus implying greater financial risk. This may seem contradictory, but a correct belief that "diversification reduces the risk of losing money" could be consistent with participants' incorrect forecasts: A portfolio with a higher average return is less likely to lose money than a portfolio with a lower average return, holding the volatility of the two portfolios constant.

The idea that people's perceptions of risk are also influenced by the prospect of losses (versus only the variance of outcomes) is well established in the experimental literature (Kahneman and Tversky 1979; Payne et al. 1980, 1981; Lopes and Oden 1999; Payne 2005). The importance of losses has also been formalized in theoretical models of asset choice (Roy 1952, Shefrin and Statman 2000). More recently, Zeisberger and colleagues have suggested that loss probability—and not variance—is paramount in risk perceptions for financial assets (Zeisberger 2016, 2018; Holzmeister et al. 2020). Taken together, these findings provide a potential answer for the apparent puzzle: People associate risk with probability of loss, so anything that moves the probability mass in the return distribution out of the loss space is a reduction of risk. If people have learned that diversification reduces risk, the easiest way to envision this might be by shifting the entire distribution into the gain space, leading to higher expectations of returns.

In Studies 1A–1E—because participants provide their beliefs about the entire distribution of possible outcomes—we can explicitly calculate their expected probability of loss for each portfolio. Across all five studies, participants believe the diversified portfolio was less likely to experience a loss (31.7%) compared with the undiversified portfolio (35.0%; $t = -7.84$, $p < 0.001$).[21] As shown in Figure 5, this difference is attenuated for participants low in financial literacy ($\hat{\beta}_{\text{Interaction}} = -0.94\%$, $t = -5.04$, $p < 0.001$) but not reversed. Thus, participants' judgments of risk seem more consistent with their beliefs about probability of loss rather than their beliefs about volatility.

We believe these results, in combination, highlight a superficial level of financial knowledge: People believe diversification reduces risk, yet many do not seem to have an understanding of risk consistent with how it is typically understood in finance (Sharpe et al. 1998) and decision theory (von Neumann and Morgenstein 1947).

Figure 5. Studies 1A–1E: Expressed Probability of Loss



Notes. Participants are binned by their financial literacy score, and the points reflect the mean value for each bin. Error bars show model-free standard errors (standard deviation of value within each bin divided by the square root of number of observations in the bin).

These different ways of understanding risk raise a question regarding financial behavior: Which will ultimately guide investment behavior? In Studies 4A–4C, we examine whether the beliefs documented in prior studies regarding portfolio performance are consistent with people's investment decisions when trying to minimize risk or maximize return.

Studies 4A and 4B: Portfolio Construction with Real Stocks

In Studies 4A–4C, we assess whether the beliefs documented in the previous studies have downstream consequences. In Studies 4A and 4B, we ask participants to construct stock portfolios for two people with different investment goals from real stocks. In Study 4C, we follow a similar procedure but instead use fictional assets to address potential confounds created by the use of real stocks.

In all three studies, each participant is asked to construct a stock portfolio for two different people: One of the two people was described as being risk averse, and the other was described as being tolerant of risk, but seeking high return. If, as our previous results suggest, people low in financial literacy believe diversification increases volatility and returns, they should create a more diversified portfolio for the investor who wants high returns compared with the one who is risk averse. Because people high in financial literacy believe diversification decreases volatility and increases returns, no clear prediction emerges: Less volatility favors giving the risk-averse investor a more diversified portfolio, but higher returns favor giving the gain-seeking investor a more diversified portfolio.

These studies differ from others in the paper in that participants are asked to choose for another. Several papers have studied whether choosing for others versus oneself changes risk preferences, and results are mixed (Reynolds et al. 2009, Eriksen and Kvaløy 2010, Chakravarty et al. 2011, Pollmann et al. 2014). Because the design of our studies is such that the choice is made for another in both conditions, any baseline differences in risk preferences because of choosing for another are present in both conditions and therefore are unlikely explanations for the effect of the manipulation.

Method

Participants. For Study 4A, we collected responses from 184 participants through AMT (target recruitment was 200, but the survey was closed after a long period of inactivity; payment = \$0.40). Ten of these participants scored below chance on the financial literacy measure and, as in previous studies, were excluded from the main analysis. For Study 4B, we collected 396 participants through AMT and removed 45 who scored below chance on the financial literacy measure (payment = \$0.60). As in previous studies, results without exclusions are presented in the online appendix.

Procedure. In both studies, we asked participants to take the role of financial advisor and help two investors with different investment goals construct stock portfolios. One of the investors was described as being tolerant of risk but seeking high return (gain-seeking investor in the analysis). The other investor was described as being intolerant of risk but accepting of lower returns (risk-averse investor in the analysis). In Study 4A, these investors were also described as being younger (gain-seeking) and older (risk-averse) to help enrich the stimuli. In Study 4B, no information about the age of the investors was provided. See the online appendix for exact stimuli.

Participants constructed portfolios for the two investors in random order. They first learned about one of the investors (either the gain-seeking or risk-averse investor), seeing a short profile about her investment goals. They then constructed a portfolio for this investor by selecting to invest in any number of companies from an available assortment. The companies were the 30 biggest American companies in 2013 according to Financial Times based on market capitalization (see online appendix for specific companies). They could select as few or as many of the companies as they wished and were told the investment would be divided equally among the companies selected. They then did the same task for the second investor. Finally, they completed the 13-item financial literacy scale used in previous studies.

Results

In both studies, we created a difference score to serve as a between-subject dependent measure by subtracting the number of companies each participant chose for the gain-seeking investor from the number of companies they chose for the risk-averse investor. Both were log-transformed to improve normality (results without transformation are similar and presented in the online appendix). Higher scores on this measure indicate that participants gave the risk-averse investor a more diversified portfolio in terms of number of companies.

Study 4A. On average, participants created a marginally *less* diversified portfolio for the risk-averse investor ($M = 5.92$ stocks) compared with the gain-seeking investor ($M = 6.32$ stocks; $t = -1.84$, $p = 0.067$) operationalized as the number of stocks over which the investment was spread.[22] More importantly, this tendency was moderated by financial literacy, such that participants lower in financial literacy had a stronger tendency to give the risk-averse investor a less diversified portfolio (regression on difference score: $\hat{\beta}_{\text{FinLit}} = 0.04$, $t = 2.87$, $p = 0.005$). The Johnson-Neyman point was 9.45, suggesting that people who scored less than this on the financial literacy were significantly more likely to exhibit this tendency. These results are illustrated in the left panel of Figure 6.

Study 4B. The results replicate those from Study 4A: On average participants created similarly diversified portfolios for the risk-averse investor ($M = 5.53$ stocks) and the gain-seeking investor ($M = 5.71$ stocks; $t = -1.05$, $p = 0.29$). However, again, this was moderated by financial literacy ($\hat{\beta}_{\text{FinLit}} = 0.06$, $t = 4.65$, $p < 0.001$). Participants scoring below 8.90 (Johnson-Neyman point) on the financial literacy measure were significantly more likely to give the risk-averse investor a less diversified portfolio compared with the gain-seeking investor. These results are illustrated in the right panel of Figure 6.

Study 4C: Portfolio Construction with Hypothetical Stock Funds

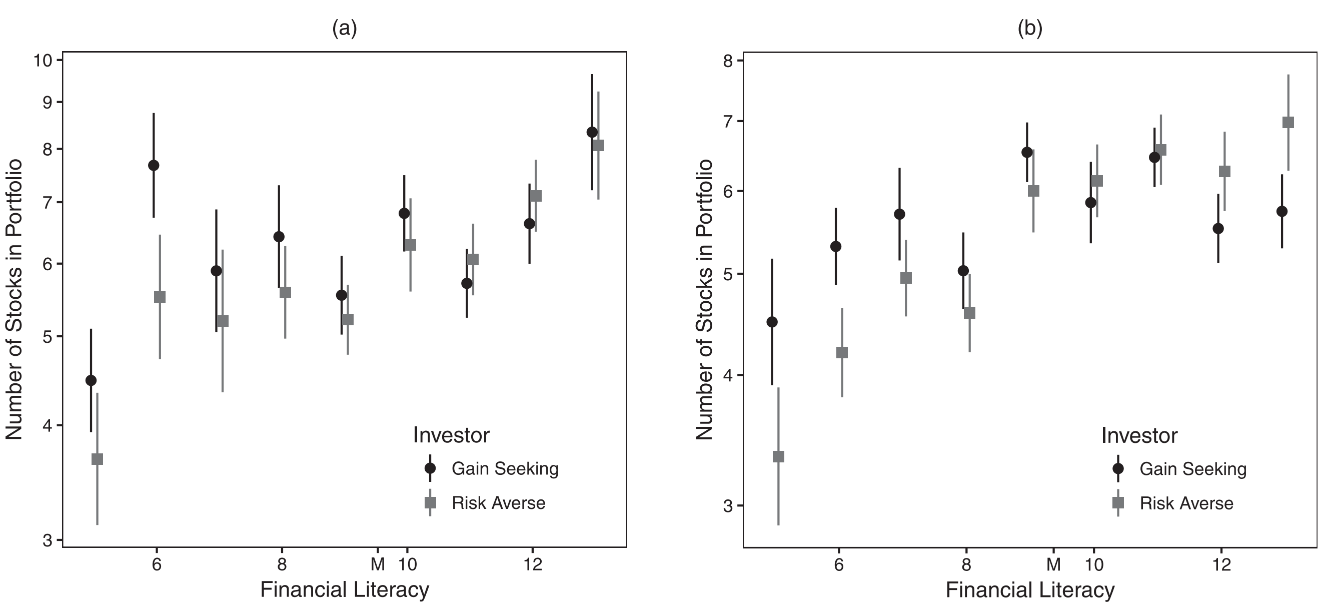
We had participants create portfolios from real stocks in Studies 4A and 4B to make the task more naturalistic for participants. However, this choice and our analysis strategy leads to a possible confound: Although participants chose fewer stocks for the risk-averse investor, those stocks may have been qualitatively different than those chosen for the gain-seeking investor in terms of historical returns.[23] In Study 4C, we instead had participants create portfolios from eight stock funds that we created to have equivalent historical volatility and expected returns. This allows us to assess whether the results in the previous studies were influenced by participant's associations with and knowledge about the real stocks. Using artificially generated funds with known generating parameters also allows us to make normative statements about participants' behavior.

Method

Participants. One hundred ninety-two undergraduate business students at the University of Colorado Boulder participated for course credit. Sixteen participants scored below chance on the financial literacy measure and are excluded from the analysis (results without exclusions are in the online appendix).

Procedure. As in Studies 4A and 4B, we asked participants to take the role of financial advisor and help two different investors (same profiles as Study 4A) construct stock portfolios to meet their investment goals. Study 4C differed from the previous studies in two ways: First, instead of constructing the portfolios with stocks from real companies, participants constructed the portfolios out of eight artificially generated stock funds, which we said were the only funds offered by the investor's brokerage firm. We generated five years of historical returns (1,250 trading days) for each of the eight funds using identical parameters. Participants could view the historical returns for each

Figure 6. Studies 4A and 4B: Number of Stocks Chosen for Each Investor (Log Transformed) by the Participant's Level of Financial Literacy



Notes. Study 4A results are shown in the left panel and Study 4B results are shown in the right panel. Participants are binned by their financial literacy score, and the points reflect the mean value for each bin. Error bars show model-free standard errors (standard deviation of value within each bin divided by the square root of number of observations in the bin). A plot showing linear best fit lines is available in the online appendix (Figure S5).

portfolio by hovering their mouse pointer over the fund's name (names were given to funds based on International Radiotelephony Spelling Alphabet; e.g., Fund Bravo). An example stimulus is shown in the online appendix (Figure S4).

Second, participants were not limited to dividing the investment equally across selected assets. Participants were told to allocate whatever percent they wished to each stock fund (using a constant sum allocation method). If they did not want to invest in a particular fund, they could allocate 0% to it.

Following the portfolio construction tasks (and a block of unrelated studies), participants completed the same 13-item financial literacy scale used in previous studies.

Results

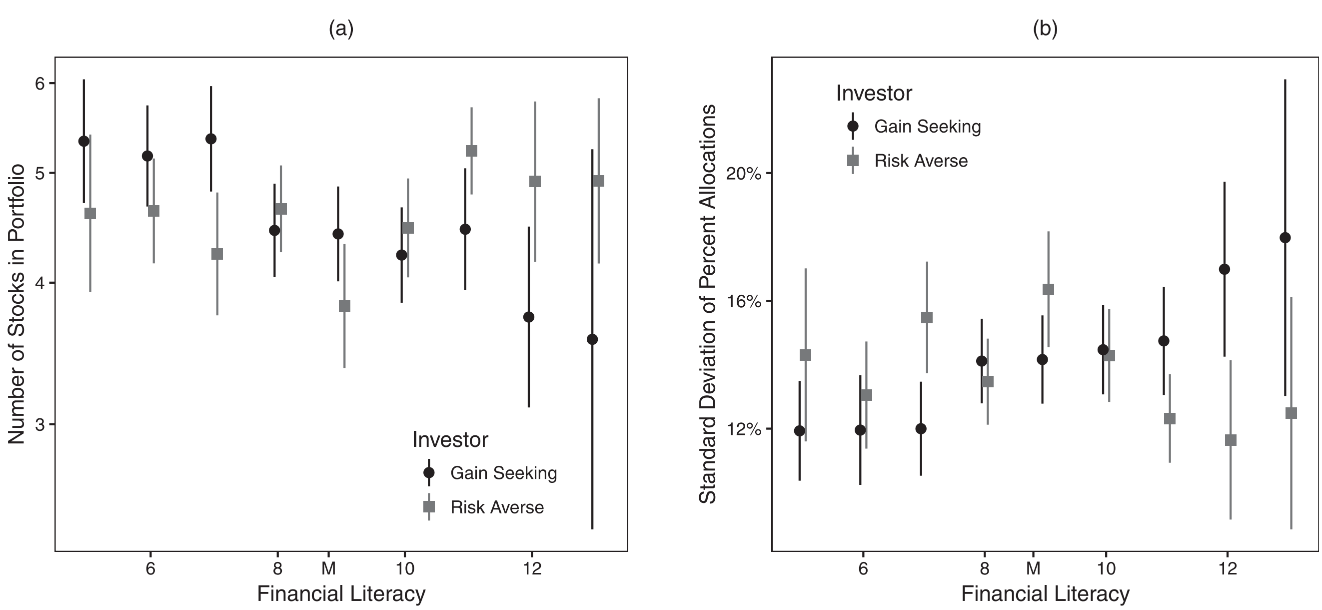
Number of Stock Funds. As a first level of analysis, we looked again at the number of assets assigned to each investor's portfolio. As in Studies 4A and 4B, we created a difference score for each participant by subtracting the number of stock funds they chose for the gain-seeking investor from the number of stock funds they chose for the risk-averse investor (both, again, log transformed to improve normality; results without transformation are available in the online appendix). On average, participants assigned a similar number of stock funds to the risk-averse investor's portfolio ($M = 4.52$ stock funds) and the gain-seeking investor's portfolio ($M = 4.54$; $t = -0.12$, $p = 0.90$). As in previous studies, this was moderated by financial literacy, such that participants with lower financial literacy were more likely to spread the risk-averse investor's portfolio over *fewer* stock funds ($\hat{\beta}_{\text{FinLit}} = 0.06$, $t = 3.37$, $p < 0.001$; Johnson-Neyman point = 7.44). This is illustrated in the left panel of Figure 7.

Deviation from Optimal Allocation. Participants were not limited to an equal allocation, so we can also examine how their portfolio deviates from an optimal allocation. Because all the stock funds are uncorrelated and have equal variance and expected returns, the optimal allocation is to spread the investment equally over the eight funds (i.e., to assign 12.5% of the investor's wealth to each fund).

We assess deviation from the optimal allocation by calculating the standard deviation in percent allocations for each portfolio for each participant. This is analogous to a diversification measure suggested by Blume and Friend (1975). The optimal allocation—12.5% in each fund—would yield 0% on this measure. The worst allocation—100% in one fund—would yield 35.56% on this measure.

On average participants gave similarly diversified portfolios to the risk-averse ($M = 14.01\%$) and gain-seeking ($M = 13.98\%$) investors ($t = 0.04$, $p = 0.97$). Consistent with the other analysis, this was moderated by financial literacy ($\hat{\beta}_{\text{FinLit}} = -0.97\%$, $t = -3.29$, $p = 0.001$). People low in financial literacy were significantly more likely to give the risk-averse investor a *less* diversified

Figure 7. Study 4C: Number of Stocks Chosen for Each Investor (Log Transformed; Left) and Level of Diversification (Standard Deviation of Investment Percent Allocations, with Lower Numbers Representing Greater Diversification; Right)



Notes. Participants are binned by their financial literacy score, and the points reflect the mean value for each bin. Error bars show model-free standard errors (standard deviation of value within each bin divided by the square root of number of observations in the bin). A plot showing linear best fit lines is available in the online appendix (Figure S7).

portfolio than the gain-seeking investor (Johnson-Neyman point = 7.31). This is illustrated in the right panel of Figure 6.

Study 5: Examining Justifications for Participants' Beliefs

In Study 5, we gather preliminary evidence for the psychological mechanisms underlying the biases documented in previous studies. We asked participants to consider a diversified or undiversified portfolio and report either (a) which one they thought would have a greater value at the end of next year or (b) which one they thought would have a more predictable value at the end of the next year. Following this, we asked the participants to explain the rationale behind their choice in as much detail as possible and coded the responses into categories.

We used this exploratory approach because there is little previous research to provide definitive hypotheses about mechanisms underlying beliefs about diversification. Additionally, in previous studies, we sometimes described portfolios consisting of actual stocks and sometimes described them just in terms of the number of stocks. In this study, we manipulated this factor to test whether it influences people's intuitions about expected returns and volatility.

Method

Participants. Four hundred participants from AMT completed the study (payment = $0.70). Thirty-three participants scored below chance on the financial literacy measure and, as in previous studies, were excluded from analysis (results without exclusions in the online appendix).

Procedure. We used a 2 (portfolio type: actual stocks or generic stocks) × 2 (question type: predictability or return) between-subjects design. Participants first saw a description of two stocks portfolios: one larger (diversified) and one consisting of a single stock (undiversified). In the actual stocks condition, the larger portfolio was described as "invested equally in four different stocks: Comcast (CMCSA), Dow-DuPont Chemicals (DWDP), Walmart (WMT), and Nike (NKE)" (the same companies used in Studies 2A). The smaller portfolio was described as invested in only one of the stocks, chosen randomly for each participant. In the generic stocks condition, the larger (smaller) portfolio was described as consisting of 10 different companies (one company) randomly selected from the Financial Times Global 500 (same approach as Study 1E).

Participants in the predictability condition were asked to indicate "Which of these two portfolios do you think will have a more predictable value at the end of the next year?" as a binary choice. Participants in the return condition were instead asked to indicate "Which of these two portfolios do you think will have a greater value at the end of the next year?" again as a binary choice. After submitting their response, participants were shown their choice, given a blank text field, and asked to "explain in as much detail as possible why you believe this to be the case." Finally, participants completed the 13-item financial literacy scale used in previous studies.

Results

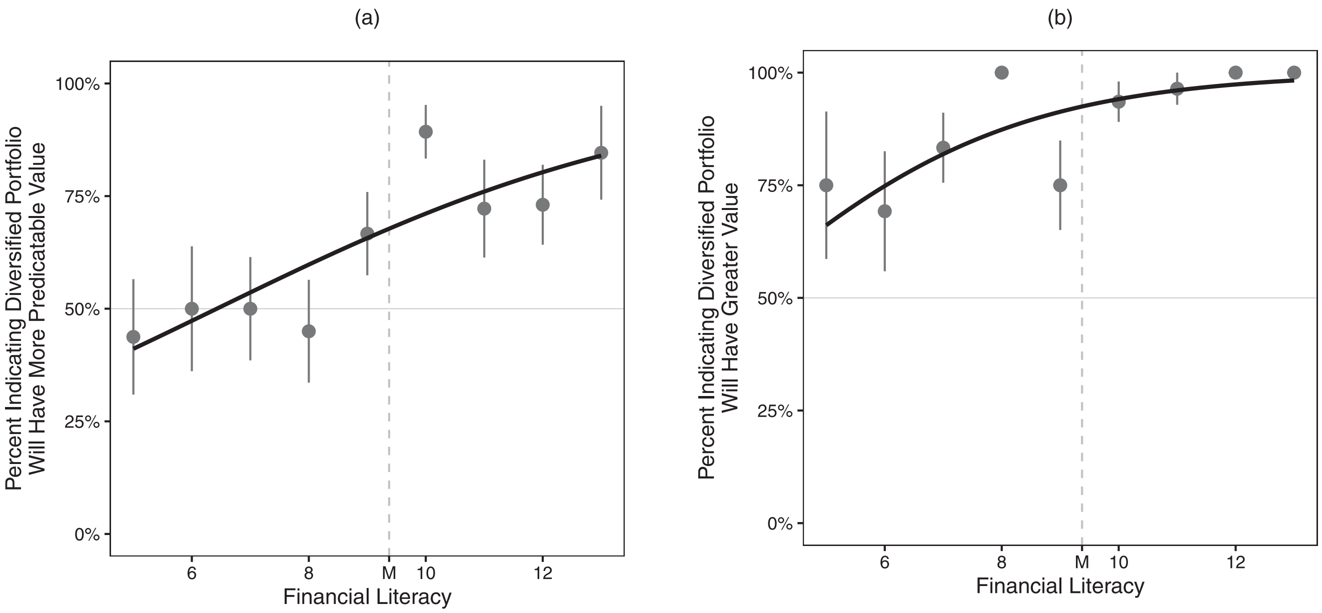
Choice Results. For the predictability question, we find 65% of participants believed the diversified portfolio would have a more predictable value at the end of the year. This result was similar for both portfolio-type conditions (actual stocks: 61%, generic stocks: 70%; difference tested using logistic regression: $z = -1.28$, $p = 0.20$). Although the proportion selecting the diversified portfolio is higher than in previous studies, we still observe a moderating role of financial literacy, such that participants low in financial literacy were more likely to select the single stock (undiversified) portfolio as being more predictable (logistic regression: $\hat{\beta}_{\text{FinLit}} = 0.25$, $z = 3.56$, $p < 0.001$). This relationship—with probabilities estimated from a logistic regression—is shown in the left panel of Figure 8.

For the return question, we find results consistent with the previous studies: Most participants (90%) expected the diversified portfolio would have a greater value after one year. There was no difference between portfolio-type conditions (actual stocks: 89%, generic stocks: 91%; $z = -0.43$, $p = 0.67$). In this study, people high in financial literacy were more likely to say the diversified portfolio would have a higher return ($\hat{\beta}_{\text{FinLit}} = 0.42$, $z = 3.42$, $p < 0.001$). This relationship is shown in the right panel of Figure 8.

Open-Ended Responses. All verbatim responses are included in data file hosted on OSF. The authors read through the responses and created a categorization scheme, encompassing 11 categories. Two research assistants worked together to code responses according to the scheme. We present the full results from the coding by question-type condition and response in the online appendix (Table S9). Here we highlight a subset of the results we believe provide illumination into the effects of interest.

First, among participants who answered the predictability question, those who believed the undiversified portfolio would have a more predictable value in one year were more likely to mention aspects of the individual stocks (54%) compared with those who believed the diversified portfolio would have a more predictable value (40%). In particular, through inspection of the responses, it seems participants are thinking about how increasing the number of stocks in a portfolio increases the number of variables that

Figure 8. Study 5: Likelihood of Saying the Diversified Portfolio Will Have a More Predictable Value (Left) or a Greater Return (Right)



Notes. The vertical dashed line indicates the mean value of financial literacy in the sample (the value at this point can be interpreted as the main effect of diversification). Points represent the judgments binned by financial literacy, with error bars indicating the standard error. Dark lines represent model fits from logistic regression.

enter into the portfolio's value. For example, one participant said, "The more variables in a prediction, [...] the more likely one of those variables can go drastically off causing your prediction to be very off." This makes sense: If you have stocks from multiple industries, there are more aspects of the economy that are likely to influence at least one of the stocks you own. However, what does not seem to be appreciated is that some of these influences—in a diversified portfolio—are likely to offset each other, leading to increased predictability (i.e., decreased volatility). Other participants linked difficulty in prediction with the difficulty of knowing and aggregating the uncertainty of many assets, again not appreciating that unsystematic risk is exactly what diversification helps mitigate.[24]

Second, participants were more likely to mention diversification (or a variant thereof) when justifying their belief that a diversified portfolio would have higher *returns* (50%) compared with when justifying their belief that a diversified portfolio would have a more predictable value (33%). In the individual responses, diversification increasing returns often seemed linked to the concept of risk. In fact, a higher percentage of participants referenced risk in their justification for why a diversified portfolio should have better returns (39%) compared with why a diversified portfolio should have a more predictable value (29%). For example, one participant justified her belief that the diversified portfolio would have higher *returns* by saying it would add "[d]iversity in the portfolio instead of [putting] all eggs in one basket." Thus, mirroring previous results, people seem to recognize that diversification reduces risk, but many represent this risk reduction as increased returns rather than decreased volatility.

Study 6: Manipulating Focus to Examine the Psychological Process

Our analysis of participant responses from Study 5 suggests that both biases we document in this manuscript are likely multiply determined. However, inspection of the responses suggests common intuitions that may partially underlie each: Many of the people who think diversification increases volatility seem to be focusing on the uncertainty regarding each individual stock within the diversified portfolio and how this uncertainty aggregates when combined. Also, many of the people who think diversification increases returns seem to be thinking about the concept of diversification, albeit without the correct statistical understanding.

In this study, we assess whether cuing people to think about either the individual stocks within a portfolio or the concept of diversification will affect their forecasts for a portfolio containing multiple stocks. Based on the results from Study 5, we predict that making people think specifically about the individual stocks within the portfolio will increase their expectations for portfolio volatility. We predict that making people think specifically about the concept of diversification (and what that means) will increase their expectations for portfolio returns.

Method

Participants. We targeted 1,000 participants from AMT and successfully recorded data from 998 (payment = \$0.70). Sixty-nine participants scored below chance on the financial literacy measure and are removed from the presented analysis (results without exclusions are available in the online appendix).

Procedure. Participants were randomly assigned to one of three between-subject conditions: individual-stock focus, diversification focus, or control. All participants were first told that an investor had a stock portfolio with an equal investment in eight companies: Apple Inc., Alphabet Inc. (formerly Google), Microsoft, Amazon.com, Berkshire Hathaway, Facebook, JPMorgan Chase, and Johnson & Johnson.[25] Participants in the individual-stock focus condition were then asked to "take a moment to think about each of the individual stocks" in the portfolio and to "write a few thoughts about your expectations for each of the stocks." Participants in the diversification focus condition were instead told that "Investing in a variety of assets is one way to achieve diversification. In this sense, the investor's portfolio is more diversified than it would be if it were invested in only one stock." These participants were then asked to "write a few sentences about what it means to have a diversified portfolio." Participants in the control condition did neither of these tasks and instead advanced directly to the forecasting questions.

Following the manipulation, all participants were asked two questions about the target portfolio: Do you think this investor's portfolio will increase/decrease in value over the next year? (11-point scale with −5 = decrease substantially, 0 = no change, and +5 = increase substantially). How predictable is the value of this investor's portfolio in one year? (8-point scale with 0 = completely unpredictable and 7 = very predictable). After these two measures, participants completed the financial literacy scale used in previous studies. We did not have ex ante predictions about how financial literacy would interact with the focus manipulations, so we treat this as an exploratory factor in the analysis.

Results

Analysis of Predictability. Ratings of portfolio predictability differed by condition ($F(2, 926) = 11.30$, $p < 0.001$). Consistent with our prediction, participants in the individual-stock focus condition—those that elaborated on each of the stocks within the portfolio—rated the portfolio as less predictable than the control condition ($M_{\text{individual-stock}} = 5.06$ versus $M_{\text{control}} = 5.58$, $F(1, 926) = 22.09$, $p < 0.001$). Unexpectedly, participants in the diversification focus condition—those that elaborated on what it means to be diversified—also rated the portfolio as less predictable than the control condition ($M_{\text{diversification}} = 5.25$, $F(1, 926) = 10.01$, $p = 0.002$), and almost as low as those in the individual-stock focus condition ($F(1, 926) = 2.47$, $p = 0.12$). The interaction between condition and financial literacy was not significant ($F(2, 923) = 0.56$, $p = 0.57$) but is illustrated—along with the means by condition—in the left panel of Figure 9.

Analysis of Expected Returns. Ratings of expected returns also differed by condition ($F(2, 926) = 15.46$, $p < 0.001$). Consistent with our prediction, participants in the diversification focus condition expected the greatest increase in value from the portfolio ($M_{\text{diversification}} = +2.65$), although this was more similar to the control condition ($M_{\text{control}} = +2.45$, $F(1, 926) = 1.72$, $p = 0.19$) than the individual-stock focus condition ($M_{\text{individual-stock}} = +1.92$, $F(1, 926) = 28.75$, $p < 0.001$). The interaction between condition and financial literacy was significant in this study ($F(1, 923) = 3.22$, $p = 0.041$; right panel of Figure 9) and seems consistent with previous results: For people with high financial literacy (10 or above), the diversification focus condition becomes significantly different than the control condition (Johnson-Neyman point = 9.74). This suggests that only people high in financial literacy—those who are likely to know that diversification is a good thing—are susceptible to the diversification focus manipulation.

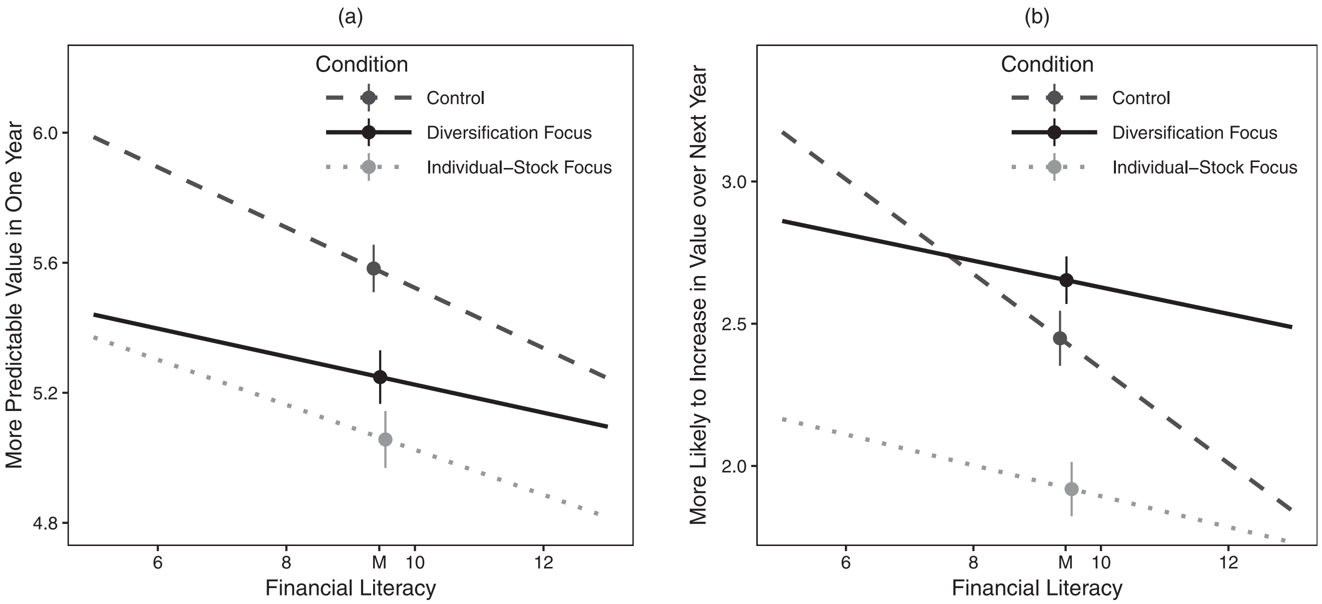
General Discussion

Across 13 studies, our results suggest that most people misunderstand the effects of diversification. People on average do not appreciate the reduction in volatility—the free lunch—that diversification can offer. This judgment bias is particularly prevalent for people low in financial literacy (Studies 1A–1E, 2A, 2B, and 3) and can lead these people to create underdiversified portfolios when they are attempting to reduce risk (Studies 4A–4C). We also find consistent evidence that people expect diversification to increase the mean return of a portfolio. People across the financial literacy spectrum exhibit this bias, and some evidence suggests it may be more pronounced for people *high* in financial literacy (Studies 1A–1E).

Implications for the Problem of Underdiversification

As highlighted by von Gaudecker (2015), the costs of underdiversification seem to be borne disproportionately by those low in financial literacy. Using survey and administrative data from the Netherlands, he concludes that "nearly all households that score high on financial literacy . . . achieve reasonable investment outcomes" (p. 489). However, households with below median financial literacy (who do not rely on external financial advice) could achieve an average of ~0.5% greater annual returns without increasing their risk through greater diversification. Our results

Figure 9. Study 6: Ratings of Portfolio Predictability (Left) and Expected Returns (Right) by Financial Literacy and Condition



Notes. Lines show linear fit by financial literacy. Points show average (mean) response by condition (located on mean financial literacy by condition) with error bars representing standard errors.

provide a mechanism through which this effect could emerge: People low in financial literacy might believe that their returns will actually be less volatile if they hold a portfolio with fewer assets, whereas people high in financial literacy appreciate the risk reduction benefit of diversification (and, perhaps, expect greater returns) and thus diversify more appropriately.

Based on the beliefs we document amongst those high in financial literacy—expecting diversification to yield lower risk and greater returns—one might expect to observe high levels of diversification amongst high financial literacy households. However, we think there are two reasons why high financial literacy investors could still be underdiversified, and—indeed—the results of von Gaudecker suggest that even many high financial literacy households could do better.

First, although high financial literacy participants in our sample do appreciate that diversification reduces volatility, it is not clear that their beliefs are well calibrated in terms of magnitude. Our back-of-the-envelope math in the discussion of Studies 1A–1E suggests that only ~10% of participants appreciate the extent to which diversification reduces risk. People with miscalibrated beliefs could remain underdiversified because they have a bad intuition of the cost-benefit calculation involved in creating a portfolio.

Second, if high financial literacy investors expect greater returns from diversification, they may frequently be disappointed in their actual returns, which could deter them from diversifying in the future. If you look at diversification decisions as a time series rather than a static judgment you may see suboptimal behavior even for high financial literacy people. This is consistent with the anecdotal evidence we mentioned earlier in the manuscript: After watching their globally diversified portfolio get outperformed by the S&P, many investors seemed to experience regret and some changed course, presumably to less diversified options.

A potential solution to underdiversification seems to be the proliferation of pooled investment funds (mutual funds, ETFs, etc.). However, 13.9% of U.S. families still own individual stocks, and the size of these investments is meaningfully large (conditional median = $25,000, conditional mean = $327,800; Brickner et al. 2017). Although much of the U.S. evidence for underdiversification is becoming dated (e.g., both Barber and Odean 2000 and Goetzmann and Kumar 2008 use discount brokerage data from 1991–1996), more recent evidence from other countries suggests the problem has not gone away (von Gaudecker 2015 uses Dutch data from 2005–2006; Campbell et al. 2018 use Indian data from 2002–2011).

Our results add to the growing literature examining lay intuitions about risk and investing (Zeisberger 2016, Long et al. 2018, Cornil et al. 2019, Merkle 2018). They can also provide novel insight into older results. For example, Benartzi (2001) found that 83.7% of Morningstar.com subscribers believed that the stock of their own employing firm was less likely to lose half of its value than the overall stock market, an effect he attributed to familiarity. Our studies suggest that a more general misunderstanding of the relationship between diversification and volatility may contribute to this result. Our results also provide a potential explanation for the correlation between financial literacy and diversification in the market (Guiso and Jappelli

2009). Finally, Goetzmann and Kumar (2008) found that personal retirement accounts tend to be *less* diversified than nonretirement accounts. In other words, people seem to be taking on more risk with the money they plan to live on during retirement. Although this may seem surprising, it is consistent with our studies: If people think that diversification leads to less predictability, they may choose to invest their retirement wealth in a small number of assets.

On the Psychology Underlying Diversification Beliefs

Although the focus of this study was documenting the existence and potential consequences of peoples' beliefs about diversification, we would be remiss to ignore the question of why these beliefs arise. In Study 5, we asked participants to self-report on the reasoning they used to justify their belief. We found that participants who believed the diversified portfolio would have a less predictable value frequently made reference to the combined unpredictability of assets in the diversified portfolio. For example, one participant wrote: "10 different companies lead to much more potential volatility and things that can happen. 10 companies compared to one means that there are 10 times the amount of possibilities that can happen to one portfolio as opposed to the other." In Study 6, we showed that asking people to think about each of the stocks in a larger portfolio led them to expect less predictability from that portfolio.

In effect, many participants do not seem to appreciate the law of large numbers and how it applies in an investment context: They seem to conflate the addition of uncertainty coming from additional individual stocks with the aggregate unpredictability of the portfolio. Although this may seem surprising to those with statistics training, the idea that averaging observations can lead to greater accuracy was actually quite slow to develop. In his book "The History of Statistics," Stigler (1986) devotes an entire chapter to this topic. Many great minds rejected the idea of averaging across observations to improve precision including Euler who stated that by combining observations in an analysis "the errors of the observations … can multiply themselves." In studying the orbits of Jupiter and Saturn, Euler refused to use all the data available to him but instead chose the observations he thought were best. Although the contexts are obviously quite different, Euler's reservations seemed to be mirrored by the intuitions of many of our participants.

As to why participants, in general, seem to think a diversified portfolio will yield a higher return, our best answer lies at the molar level (i.e., as a relationship between large and often complex constructs versus at the level of micromediation; Cook and Campbell 1979). When instructed to think about the concept of diversification, participants increased their perceptions of expected return (Study 6). This was particularly true for participants high in financial literacy, who likely have a greater familiarity with the concept of diversification. Because most participants seem to correctly link diversification with risk reduction, we are left to conclude that many believe this risk-reduction manifests as increased returns rather than decreased volatility. Although inconsistent with standard models, associating increased returns with reduced risk would be consistent with many behavioral theories of risk perception (Holzmeister et al. 2020).

One may also wonder how strongly people hold these erroneous beliefs we document. Are they deeply held and innate mistakes? Or are they the haphazard product of attempting to rationalize why diversification is *good*? We believe this distinction—although interesting theoretically—may not be of much practical consequence. We find that these beliefs persist in the face of incentives (Study 3) and influence downstream behavior (Studies 4A–4C). In the posttest for Study 1, we find that marketplace participation does not moderate these beliefs, so even people who seemingly have a large incentive to improve their understanding fail to do so.

We examined participants' self-reported confidence in their beliefs in a study reported in Appendix B. In this study, we asked participants to make binary judgments between diversified and undiversified portfolios in terms of expected return and predictability (similar to Study 3 and Study 5). We then asked participants to report their confidence in these judgments. In general, confidence was high (above the midpoint on the scale), and there were only marginal differences in confidence depending on judgment.

The psychology underlying our work also relates to past work on judgments of perceived benefits and risks. Ganzach (2000)—examining the context of financial assets—proposes a model in which judgments of risk and return for an unfamiliar asset are both based on a global attitude/preference for that asset. This leads to a negative correlation in judgments—unfamiliar assets judged to have a higher return will also be judged to have lower risk—that is inconsistent with standard assumptions about asset pricing but consistent with models of the affect heuristic (Finucane et al. 2000). Our results are partially consistent with this proposal: People high in financial literacy tend to believe that diversified portfolios are less volatile and that they yield higher returns on average. The affect heuristic may help explain why people high in financial literacy have more appropriate intuitions about diversification and risk reduction. As for people low in financial literacy, they tend to believe that diversified portfolios are more volatile and that they yield higher returns on average. This result cannot easily be traced to the affect heuristic.

Future Directions

Financial literacy has become a topic of concern for both policy makers and financial advisors, who typically use scales like we use in this study to measure the financial literacy of consumers. Our results suggest that these scales might paint an incomplete picture of financial understanding: Most of our participants get the *diversification* question correct in our measure, but—unfortunately—it does not seem that this correct answer reflects an accurate understanding of how diversification works. It seems that if governments and financial organizations really want to improve the decisions of people low in financial literacy, we need to go beyond *that* they get something wrong to *how* they get it wrong, because this can help shape both educational and informational interventions.

To this end, we partnered with the financial firm mentioned previously to design an interactive *quiz*—featuring immediate and visual feedback—to help educate the users about diversification. This quiz has currently been taken by thousands of visitors to their website. More generally, graphical methods for visualizing potential portfolio outcomes—analogous to the distribution builder—may provide a fruitful approach for successful risk communication and understanding (Goldstein et al. 2008). These could effectively be deployed by advisors and investment platforms as *just-in time* financial education—as statistical principles are easily forgotten (Fernandes et al. 2014)—and would likely communicate risk in a more behaviorally appropriate manner—illustrating things like probability of loss (Zeisberger 2018)—compared with approaches like *star ratings* (e.g., Morningstar).

Conclusion

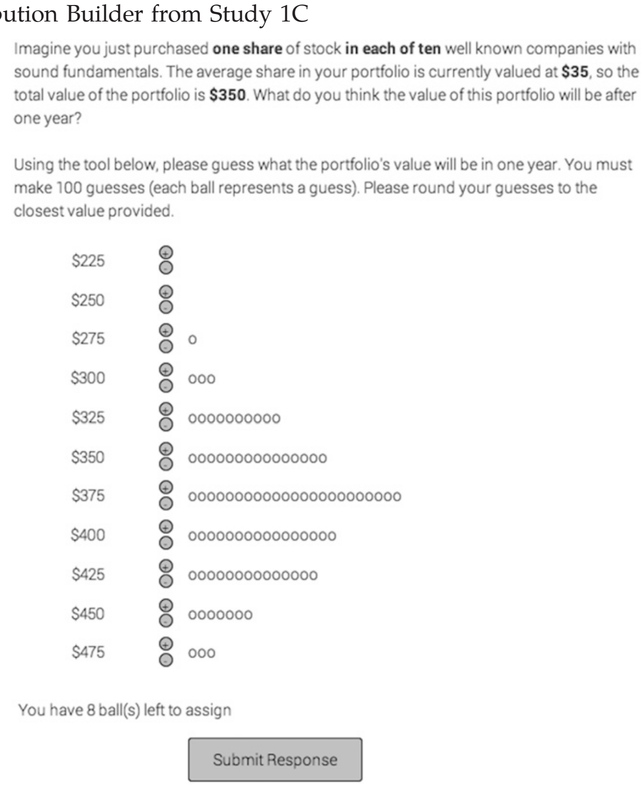
Diversification is a fundamental aspect of effective investing that, unfortunately, many people fail to sufficiently use. We document erroneous beliefs about the benefits of diversification and provide evidence for the psychology that underlies them. Our work can help explain why many people—particularly those low in financial literacy—are underdiversified.

Acknowledgments

The authors thank Dan Ariely, Asaf Bernstein, Dan Egan, David Gross, Nick Light, John Lynch, Matt Meister, Justin Pomerance, Brian Waters, Ed Van Wesep, Gal Zauberman, and colleagues at the University of Colorado Boulder. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of The Varieties of Understanding Project, Fordham University, or The John Templeton Foundation.

Appendix A. Example of Distribution Builder

Figure A.1. Example of Distribution Builder from Study 1C



Appendix B. Strength in Beliefs Study

Method

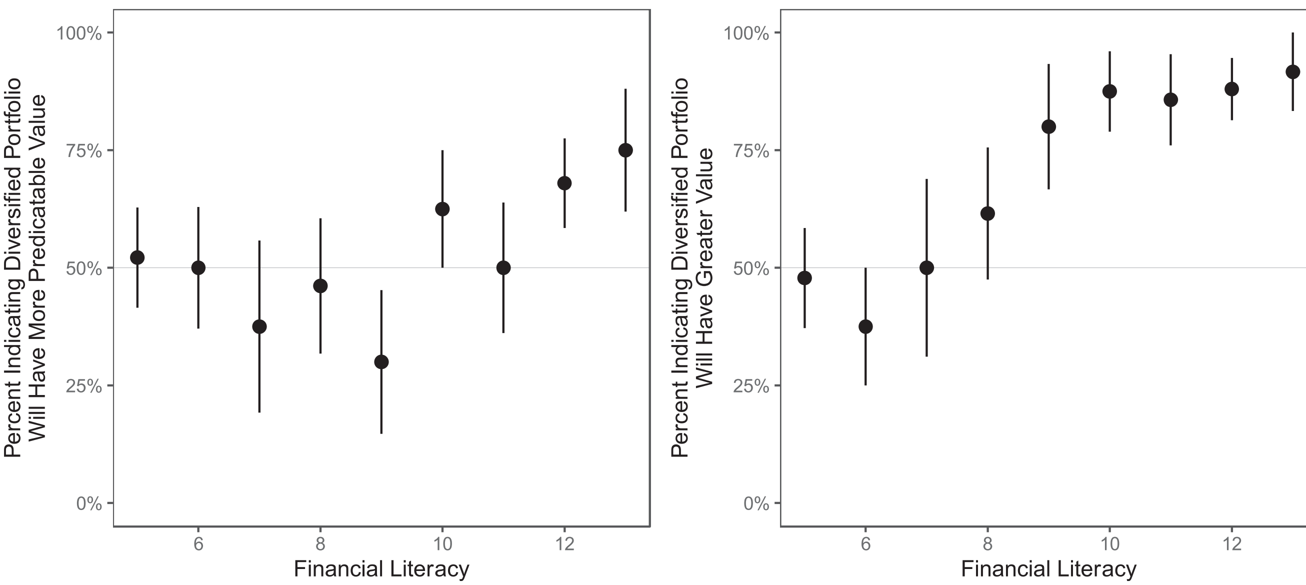
Participants. We recruited 160 people from AMT to participate in these studies (payment = $.75); 23 scored below chance on the financial literacy measure and are removed from this analysis.

Procedure. Participants were shown two stock portfolios. Portfolio A (diversified) consisted of 1 company and Portfolio B (undiversified) consisted of 10 different companies. The companies in both portfolios were described as being "well known" with "sound fundamentals." The current value of both portfolios was $10,000.

Mirroring previous studies, participants were asked to make two binary judgments: (i) which portfolio will have a greater value at the end of next year and (ii) which portfolio will have a more predictable value at the end of next year?

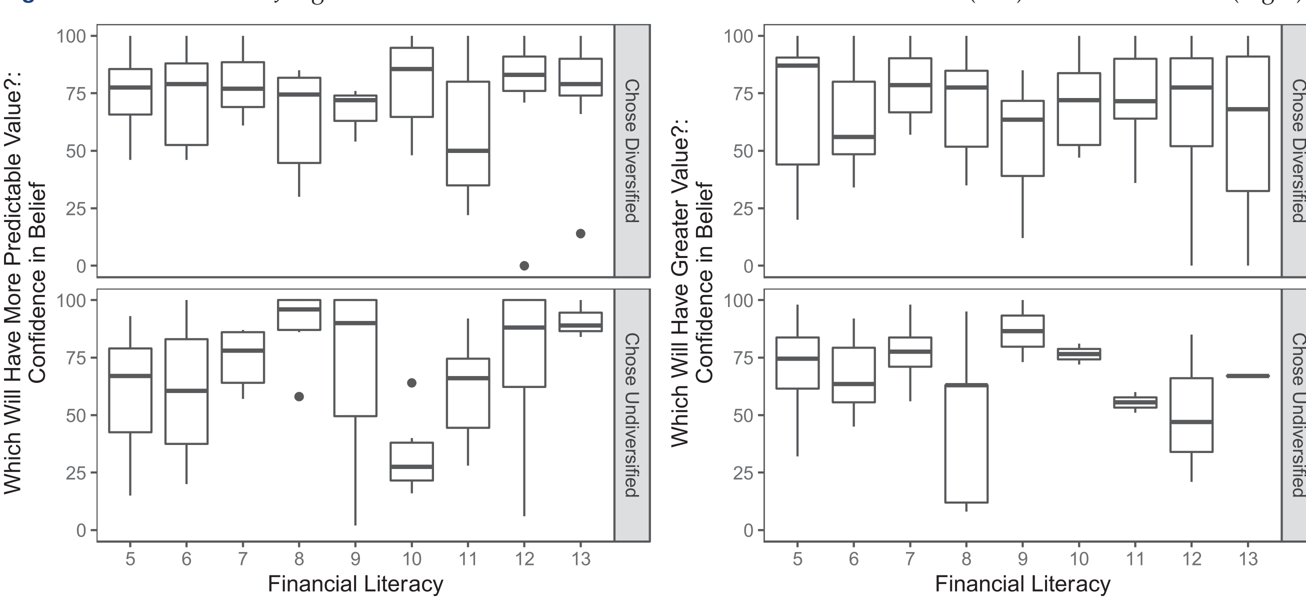
Following this, participants were shown their two judgments and asked to self-report their confidence ("When you answered this question, how confident were you in this belief?") on a scale from 0 to 100 (anchored by not confident at all and very confident).

Figure B.1. Likelihood of Saying the Diversified Portfolio Will Have a More Predictable Value (Left) or a Greater Return (Right)



Note. Points represent the judgments binned by financial literacy, with error bars indicating the standard error.

Figure B.2. Confidence in Judgments for Which Portfolio Will Have a More Predictable Value (Left) or a Greater Return (Right)



Notes. Panels reflect the response chosen. Data are binned by financial literacy.

Results

Beliefs. The judgments (shown in Figure B.1) replicate previous studies. Participants high in financial literary were more likely to indicate that the diversified portfolio would be more predictable (logistic regression: $z = 1.74$, $p = 0.083$) and have a higher return ($z = 4.51$, $p < 0.001$).

Confidence. Confidence in all judgments was high (M = 68.7 of 100, all judgments above midpoint with $p < 0.001$) and did not differ significantly within question based on judgment made (all $p > 0.10$). Boxplots of confidence, broken out by question and judgment made and binned by financial literacy, are shown in Figure B.2.

Endnotes

[1] We refer here to the simple, arithmetic mean return for a portfolio, as described in the first paragraph. We test this directly in Studies 1A–1E. We note, however, that if return distributions over a given period are right skewed (e.g., log-normal), which is likely the case given compounding, the expected *median* return from a diversified portfolio should be higher (Hughson et al. 2006). Additionally, reducing a portfolio's volatility will mechanically yield higher *log* returns (Markowitz 1976). We discuss this distinction and how it relates to our results later in the manuscript.

[2] The highlighted effects are not always *significant* in each individual study, but—as shown in Figure S1 in the online appendix—the results from individual studies are consistent and are highly significant when aggregated.

[3] We varied the descriptions of the portfolios across studies to assess both robustness and potential alternative explanations. The exact wording of the manipulations is available in the online appendix. For Study 1C, we included an exploratory third (within-subject) condition, which is outside the scope of the current study but is described for transparency in the online appendix.

[4] In these and all other AMT studies, participation was restricted to workers located in the United States. We chose to conduct many experiments in this manuscript using AMT because we wanted large sample sizes that would be difficult to obtain in a physical laboratory. Past research suggests that AMT samples are more diverse than the typical laboratory samples (albeit not perfectly representative of the U.S. population) and that data quality tends to be high (Paolacci and Chandler 2014, Goodman and Paolacci 2017, Hauser et al. 2019).

[5] Of 2,030 responses, 210 failed the attention check, 131 scored below chance on financial literacy, and 47 did both. The *chance* exclusion on financial literacy was included for two reasons. (1) It reduces skew on the x-variable because the chance cutoff for financial literacy was 1.82 SD below mean performance, whereas a perfect score was only 1.43 SD above the mean. (2) We believe that a below-chance score on financial literacy is likely indicative of inattention. These participants were more than twice as likely (26% versus 10%) to fail the attention check question. The manipulation check provides only a probabilistic check of inattention because a guessing strategy would yield a high pass rate (50% in all but one study).

[6] We also varied the wording of the distribution builder instructions across studies, which did not seem to have a substantive impact on responses. Wording for each study is provided in the online appendix.

[7] We additionally collected a measure of numeracy (Fernandes et al. 2014) in Studies 1A and 1B and measures of perceived riskiness in Studies 1A–1C. These were exploratory measures, which we discuss later in the study. We also collected a measure of perceived attractiveness in Study 1C.

[8] The results are similar for other operationalizations of volatility (middle ranges; see Table S1 in the online appendix).

[9] For reference, the historical correlation in annual returns (1987–2017) between Apple, Inc. and Microsoft Corporation is 0.40 (https://www.portfoliovisualizer.com/asset-correlations).

[10] In our sample, the mean score on the financial literacy measure was 9.54 with an SD of 2.19, so a score of 8.75 would be 0.36 SD below the mean level. Our sample average is higher than that of Fernandes et al. (2014), in part because of the exclusion criteria we use. Without exclusions, the mean level of financial literacy in our data is 8.33.

[11] Similar results are obtained using the median (see Table S1 in the online appendix).

[12] Consider, for instance, an undiversified portfolio invested entirely in a money market fund versus a portfolio invested half in a money market fund and half in the stock market. The latter will be more volatile, despite being more diversified, and will have a higher expected return.

[13] Diversification, in this case, can be conceptualized as a linear combination of identical random variables (each representing the average stock in the set).

[14] Coded: male = 0, female = 1.

[15] Study 2B was conducted before PetSmart went private, Amazon acquired Whole Foods, and Dow Chemicals merged with DuPont.

[16] In a later study (Study 5), we ask people to justify their rationale when answering a similar question and note that none of the participants reference *median* or *skew* in their text responses.

[17] These stocks were chosen because the companies are known, but we suspected participants would not have strong opinions about either company. Both also have similar historical stock performance metrics: Since 2010, the companies have had annualized returns of 9.9% (CSCO) and 9.5% (ORCL) and annualized standard deviations of 24.8% (CSCO) and 22.2% (ORCL). Their correlation (on annual returns) is 0.32 (https://www.portfoliovisualizer.com/asset-correlations).

[18] Over the month under consideration, Oracle stock had a better return (−2.3%) than Cisco (−7.6%). The diversified portfolio's return was—by construction—the average of these two values. Oracle stock was also less volatile than Cisco stock (lower variance in daily returns). Unfortunately, we missed an opportunity to make a rhetorical point here, as Oracle stock was also less volatile than the diversified portfolio over this month. In total, 211 participants earned a $1 bonus payment for correctly predicting what would happen over the next month, which we paid at the specified time.

[19] This may raise a concern about endogeneity. Removing this question from the financial literacy measure has little effect on any of the results, which we report for each study in the online appendix (e.g., Table S4 for Studies 1A–1E). We also note that we collected a numeracy measure (Fernandes et al. 2014) from participants in Studies 1A and 1B, which yields similar, but less precisely estimated, results (see Table S6 in the online appendix) as the financial literacy measure. This suggests that the biases we document relate, in part, to a general ability to reason with numbers (a component of financial literacy, $r = 0.51$ between the numeracy and financial literacy measures in our studies) and not just to factual knowledge about the financial domain (another component of financial literacy).

[20] We also explicitly asked participants about the riskiness of the two portfolios in Studies 1A–1C (as exploratory measures following the primary dependent measure). Across Studies 1A and 1B, 77% of participants said the 1-stock portfolio was riskier than the 10-stock portfolio. Similarly, in Study 1C, participants rated the 1-stock portfolio as riskier (6.97 of 9) than the 10-stock portfolio (3.96).

[21] In Studies 1C–1E, the middle bucket was centered on a 0% change. For these studies, we coded 1/2 of the balls allocated to this bucket as *losses* to increase the consistency of the measure across studies. This choice does not have a substantive impact on the difference in probability of loss between portfolios but does lower the probability

of loss for both portfolios (no balls in this bucket as losses: 29.6% versus 26.8%; $t = -5.11$, $p < 0.001$).

[22] Reported values are the exponentiated means of logged number of stocks (i.e., geometric means) in Studies 4A–4C.

[23] We examine the specific stocks chosen for the risk-averse and gain-seeking investors in Studies 4A and 4B in the online appendix (Figure S6).

[24] It is possible that participants interpreted the term *predictability* in this (and other) studies in epistemic terms (information that could—in principle—be known). For example, one participant justified a belief that the one-stock portfolio would be more predictable by saying "Because it's only 1 company, thus easier to keep track of." However, other participants linked the difficulty of obtaining information to the difficulty of *aggregating* information: "Tracking and estimating the values of 10 companies involves more analysis and uncertainty than doing so for one company." Thus, it seems like judgments of predictability are at least partially intertwined with judgments about ease of tracking and computation. Regardless, although an epistemic interpretation of predictability in this study might warrant a judgment that the larger (diversified) portfolio is less predictable, it would not explain the forecasting results from Studies 1A–1E or 2B, in which we find similar effects, nor would it explain the allocation decisions in Studies 4A–C.

[25] Although we did not inform participants how these stocks were chosen, these were the eight largest public companies by market cap at the time the survey was conducted.

References

Almenberg J, Dreber A (2015) Gender, stock market participation, and financial literacy. *Econom. Lett.* 137:140–142.

André Q, Reinholtz N, De Langhe B (2017) Variance spillover in intuitive statistical judgments. Gneezy A, Griskevicius V, Williams P, eds. *NA Advances in Consumer Research* (Association for Consumer Research, Deluth, MN)., 45:336–340.

Barber BM, Odean T (2000) Trading is hazardous to your wealth: The common stock investment performance of individual investors. *J. Finance* 55(2):773–806.

Bates D, Maechler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J. Statist. Software* 67(1):1–48.

Benartzi S (2001) Excessive extrapolation and the allocation of 401(k) accounts to company stock. *J. Finance* 56(5):1747–1764.

Benartzi S, Thaler RH (2001) Naïve diversification strategies in defined contribution savings plans. *Amer. Econom. Rev.* 91(1):79–98.

Blume ME, Friend I (1975) The asset structure of individual portfolios and some implications for utility functions. *J. Finance* 30(2):585–603.

Brennan MJ (1975) The optimal number of securities in a risky asset portfolio when there are fixed costs of transacting: Theory and some empirical results. *J. Financial Quantitative Anal.* 10(3):483–496.

Brickner J, Dettling LJ, Henriques A, Hsu JW, Jacobs L, Moore KB, Pack S, et al. (2017) Changes in U.S. family finances from 2013 to 2016: Evidence from the survey of consumer finances. *Federal Reserve Bull.* 103(3):1–41.

Campbell JY (2000) Diversification: A Bigger Free Lunch. (Winter). *Canadian Investment Review* 13(4):14–15.

Campbell JY (2006) Household finance. *J. Finance* 61(4):1553–1604.

Campbell JY, Ramadorai T, Ranish B (2018) Do the rich get richer in the stock market? Evidence from India. NBER Working Paper No. 24898, National Bureau of Economic Research, Cambridge, MA.

Chakravarty S, Harrison GW, Haruvy EE, Rutström EE (2011) Are you risk averse over other people's money? *Southern Econom. J.* 77(4):901–913.

Christelis D, Jappelli T, Padula M (2010) Cognitive ability and portfolio choice. *Eur. Econom. Rev.* 54:18–38.

Cook TD, Campbell DT (1979) *Quasi-Experimentation: Design and Analysis Issues for Field Settings* (Rand McNally, Chicago, IL).

Cooper I, Kaplanis E (1994) Home bias in equity portfolios, inflation hedging, and international capital market equilibrium. *Rev. Financial Stud.* 7(1):44–60.

Cornil Y, Hardisty D, Bart Y (2019) Easy, breezy, risky: Lay investors fail to diversify because correlated assets feel more fluent and less risky. *Organ. Behav. Human Decision Processes* 153:103–117.

de Langhe B, van Osselear SMJ, Puntoni S, McGill AL (2014) Fooled by heteroskedastic randomness: Local consistency breeds extremity in price-based quality inferences. *J. Consumer Res.* 41: 978–994.

Delavande A, Rohwedder S (2008) Eliciting subjective probabilities in Internet surveys. *Public Opinion Quart.* 72(5):866–891.

Eriksen KW, Kvaløy O (2010) Myopic investment management. *Rev. Finance* 14(3):521–542.

Fernandes D, Lynch JG Jr, Netemeyer RG (2014) Financial literacy, financial education, and downstream behavior. *Management Sci.* 60(8):1861–1883.

Finucane ML, Alhakami A, Slovic P, Johnson SM (2000) The affect heuristic in judgments of risks and benefits. *J. Behav. Decision Making* 13:1–17.

Ganzach Y (2000) Judging risk and return of financial assets. *Organ. Behav. Human Decision Processes* 83(2):353–370.

Gaure S (2018) Linear group fixed effects. R package version 2.6-2291.

Goetzmann WM, Kumar A (2008) Equity portfolio diversification. *Rev. Finance* 12(3):433–463.

Goldstein DG, Rothschild D (2014) Lay understanding of probability distributions. *Judgment Decision Making* 9(1):1–14.

Goldstein DG, Taleb NN (2007) We don't quite know what we are talking about when we talk about volatility. *J. Portfolio Management* 33(4):84–86.

Goldstein DG, Johnson EJ, Sharpe WF (2008) Choosing outcomes vs. choosing products: Consumer-focused retirement investment advice. *J. Consumer Res.* 35(3):440–456.

Goodman JK, Paolacci G (2017) Crowdsourcing consumer research. *J. Consumer Res.* 44:196–210.

Guiso L, Jappelli T (2009) Financial literacy and portfolio diversification. Working paper 212.

Hauser DJ, Paolacci G, Chandler J (2019) Common concerns with MTurk as a participant pool: Evidence and solutions. Kardes FR, Herr PM, Schwarz N, eds. *Handbook of Research Methods in Consumer Psychology* (Routledge, London), 319–337.

Holzmeister F, Huber J, Kirchler M, Lindner F, Weitzel U, Zeisberger S (2020) What drives risk perception? A global survey with financial professionals and lay people. *Management Sci.* 66(9):3977–4002.

Huberman G (2001) Familiarity breeds investment. *Rev. Financial Stud.* 14(3):659–680.

Hughson E, Stutzer M, Yung C (2006) The misuse of expected returns. *Financial Anal. J.* 62(6):88–96.

Johnson PO, Fay LC (1950) The Johnson-Neyman technique, its theory and application. *Psychometrika* 15(4):349–367.

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–291.

Kelly M (1995) All their eggs in one basket: Portfolio diversification of U.S. households. *J. Econom. Behav. Organ.* 27:87–96.

Klapper LF, Lusardi A, Panos GA (2012) Financial literacy and the financial crisis. Working paper 17930, NBER, Cambridge, MA.

Long AR, Fernbach PM, de Langhe B (2018) Circle of incompetence: Sense of understanding as an improper guide to investment risk. *J. Marketing Res.* 55(4):474–488.

Lopes LL, Oden GC (1999) The role of aspiration level in risky choice: A comparison of cumulative prospect theory and SP/A. *J. Math. Psych.* 43(2):286–313.

Lusardi A, Mitchell OS (2007) Financial literacy and retirement preparedness: Evidence and implications for financial education. *Bus. Econom.* 42(1):35–44.

Markowitz HM (1952) Portfolio selection. *J. Finance* 7(1):77–91.

Markowitz HM (1976) Investment for the long run: New evidence for an old rule. *J. Finance* 31(5):1273–1286.

Merkle C (2018) The curious case of negative volatility. *J. Financial Marketing* 40:92–108.

Merton RC (1987) A simple model of capital market equilibrium with incomplete information. *J. Finance* 42(3):483–510.

Mitchell OS, Utkus SP (2003) The role of company stock in defined contribution plans. Mitchell OS, Smetters K, eds. *The Pension Challenge: Risk Transfers and Retirement Income Security* (Oxford University Press, Oxford, UK), 33–70.

Nelson LD (2014) False-positives, p-hacking, statistical power, and evidential value [Power Point Slides]. Presented at Berkeley Initiative for Transparency in the Social Sciences Summer Institute, Berkeley, CA. Accessed December 28, 2020, http://web.archive.org/web/20190929020751.

Nieuwerburgh SV, Veldkamp L (2010) Information acquisition and underdiversification. *Rev. Econom. Stud.* 77(2):779–805.

Obrecht NA, Chapman GB, Gelman R (2007) Intuitive t tests: Lay use of statistical information. *Psych. Bull. Rev.* 14(6):1147–1152.

Paolacci G, Chandler J (2014) Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Direction Psych. Sci.* 23(3):184–188.

Payne J (2005) It is whether you win or lose: The importance of overall probabilities of winning or losing in risky choice. *J. Risk Uncertainty* 30(1):5–19.

Payne J, Laughhunn D, Crum R (1980) Translation of games and aspiration level effects in risky choice behavior. *Management Sci.* 26(10):1039–1060.

Payne J, Laughhunn D, Crum R (1981) Further tests of aspiration level effects in risky choice behavior. *Management Sci.* 27(8):953–958.

Peterson CR, Beach LR (1967) Man as an intuitive statistician. *Psych. Bull.* 68(1):29–46.

Polkovnichenko V (2005) Household portfolio diversification: A case for rank dependent preferences. *Rev. Financial Stud.* 18:1467–1502.

Pollmann MMH, Potters J, Trautmann ST (2014) Risk taking by agents: The role of ex-ante and ex-post accountability. *Econom. Lett.* 123(3):387–390.

Reinholtz N (2015) *Persistence in Consumer Search* (Columbia University Academic Commons, New York).

Reynolds DB, Joseph J, Sherwood R (2009) Risky shift vs. cautious shift: Determining differences in risk taking between private and public management decision making. *J. Bus. Econom. Res.* 7(1):63–77.

Roy A (1952) Safety first and the holding of assets. *Econometrica* 20(3):431–449.

Sharpe WF (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Finance* 19(3):425–442.

Sharpe WF, Alexander GJ, Bailey JW (1998) *Investments*, 6th ed. (Prentice Hall, Upper Saddle River, NJ).

Sharpe WF, Goldstein DG, Blythe PW (2000) The distribution builder: A tool for inferring investor preferences. Accessed December 28, 2020, https://web.archive.org/web/20191111034611/https://web.stanford.edu/~wfsharpe/art/qpaper/qpaper.pdf.

Shefrin H, Statman M (2000) Behavioral portfolio theory. *J. Financial Quantitative Anal.* 35(2):127–151.

Soll JB, Klayman J (2004) Overconfidence in interval estimates. *J. Experiment. Psych. Learning Memory Cognition* 30(2):299–314.

Spiller SA, Fitzsimons GJ, Lynch JG Jr, McClelland GH (2013) Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. *J. Consumer Res.* 50(2): 277–288.

Statman M (2020) Five myths about stock diversification. *Wall Street Journal* (August 8), https://www.wsj.com/articles/5-myths-about-stock-diversification-11596896778).

Stigler SM (1986) *The History of Statistics* (Harvard University Press, Cambridge, MA).

Tukey JW (1977) *Exploratory Data Analysis* (Addison-Wesley, Reading, MA).

von Gaudecker HM (2015) How does household portfolio diversification vary with financial literacy and financial advice?. *J. Finance* 489–507.

von Neumann J, Morgenstein O (1947) *Theory of Games and Economic Behavior* (Princeton University Press, Princeton, NJ).

Zeisberger S (2016) Do people care about loss probabilities? Accessed May 15, 2020, https://ssrn.com/abstract=2169394.

Zeisberger S (2018) What is risk? How investors perceive risk in return distributions. Accessed December 11, 2020, https://ssrn.com/abstract=2811636.

Do People Understand the Benefit of Diversification?

ONLINE SUPPLEMENT

Table of Contents

Studies 4A–4C

Study 5

Study 6

Simulation: How Diversification Can Effect Simple (Mean) and Median Returns

Studies 1A–1E: Descriptions of the Portfolios Used in Each Study

Study 1A

Undiversified. Imagine **you just invested $15,000** in the stock market. You invested all of your $15,000 in **a single company** with sound fundamentals. Using the tool below, please indicate your expectations for the value of your portfolio **in exactly one year**.

Diversified. Imagine **you just invested $15,000** in the stock market. You divided your $15,000 investment evenly **between ten different companies** with sound fundamentals. Using the tool below, please indicate your expectations for the value of your portfolio **in exactly one year**.

Study 1B

Undiversified. Imagine you just purchased one share of stock in a well-known company with sound fundamentals. The **share is currently valued at $350**, so the total value of your portfolio is currently **$350**. Using the tool below, please indicate your expectations for the value of your portfolio **in exactly one year**.

Diversified. Imagine you just purchased one share of stock in each of ten different well-known companies with sound fundamentals. The **average value of each share is $35**, so the total value of your portfolio is currently **$350**. Using the tool below, please indicate your expectations for the value of your portfolio **in exactly one year**.

Study 1C

Undiversified. Imagine you just purchased **ten shares** of stock in a well-known company with sound fundamentals. The shares are currently valued at **$35**, so the total value of the portfolio is **$350**. What do you think the value of this portfolio will be after one year? Using the tool below, please guess what the portfolio's value will be in one year. You must make 100 guesses (each ball represents a guess). Please round your guesses to the closest value provided.

Diversified. Imagine you just purchased **one share** of stock **in each of ten** well-known companies with sound fundamentals. The average share in your portfolio is currently valued at **$35**, so the total value of the portfolio is **$350**. What do you think the value of this portfolio will be after one year? Using the tool below, please guess what the portfolio's value will be in one year. You must make 100 guesses (each ball represents a guess). Please round your guesses to the closest value provided.

*Undiversified (***Exploratory Condition, Not Used in Reported Analysis***)*. Imagine you just purchased **one share** of stock in a well-known company with sound fundamentals. The share is currently valued at **$350**. What do you think the value of this stock will be after one year? Using the tool below, please guess what the stock's value will be in one year. You must make 100 guesses (each ball represents a guess). Please round your guesses to the closest value provided.

Study 1D

Undiversified. Imagine you just purchased stock from **one company** (the company is well known and has sound fundamentals). The current value of your shares is **$3500**. Using the tool below, please guess what this portfolio's value will be in one year. You must make 100 guesses (each ball represents a guess). Please round your guesses to the closest value provided.

Diversified. Imagine you just purchased stock from **ten *different* companies** (the companies are well known and have sound fundamentals). The current value of your shares is **$3500**. Using the tool below, please guess what this portfolio's value will be in one year. You must make 100 guesses (each ball represents a guess). Please round your guesses to the closest value provided.

Study 1E

Undiversified. Imagine a stock portfolio consisting of **one company** *randomly selected* from the Financial Times Global 500 (list of the most valuable companies in the world). The current value of your shares is **$3500**. Using the tool below, please guess what this portfolio's value will be in one year. You must make 100 guesses (each ball represents a guess). Please round your guesses to the closest value provided.

Diversified. Imagine a stock portfolio consisting of **ten *different* company** *randomly selected* from the Financial Times Global 500 (list of the most valuable companies in the world). The current value of your shares is **$3500**. Using the tool below, please guess what this portfolio's value will be in one year. You must make 100 guesses (each ball represents a guess). Please round your guesses to the closest value provided.

Studies 1A–1E: Distribution Builder Training Given to Participants

In this study, we will ask you some questions about your expectations for uncertain events.

We will ask you to answer these questions using a tool we created.

The tool, shown below, requires you to assign "balls" to different possible outcomes. You should assign these balls based on how likely you think each outcome is. You should assign the most balls to the outcome you think is most likely and the fewest to the outcome you think is least likely.

In the example below, you can see there are 7 balls assigned to "Outcome A" and 3 balls assigned to "Outcome B." This would mean that you think "Outcome A" is more likely than "Outcome B." More specifically, it means you think there is a 7 out of 10 chance of "Outcome A" and only a 3 out of 10 chance of "Outcome B."

Outcome A + - ooooooo

Outcome B + - ooo

On the next page you will practice using this tool.

To practice using our tool, imagine there are 3 possible outcomes: 'Outcome A,' 'Outcome B,' and 'Outcome C.'

Please use the tool below to indicate the following: you think 'Outcome C' is most likely and 'Outcome B' is least likely. Use the plus and minus buttons beside each outcome to add or subtract balls from that outcome.

Remember: you should assign more balls to the outcomes you think are more likely.

You cannot advance to the next screen until you (1) assign all 10 balls to the possible outcomes and (2) correctly indicate that 'Outcome C' is the most likely and 'Outcome B' is the least likely.

Outcome A + -

Outcome B + -

Outcome C + -

You have 10 ball(s) left to assign

Submit Response

Studies 1A–1E: Language Used to Elicit Distribution Builder Responses

Study 1A and 1B:
Using the tool below, please indicate your expectations for the value of your portfolio in exactly one year.

Studies 1C–1E:
What do you think the value of this portfolio will be after one year? Using the tool below, please guess what the portfolio's value will be in one year. You must make 100 guesses (each ball represents a guess). Please round your guesses to the closest value provided.

Studies 1A–1E: Initial Value of Portfolios and Values of Bins for Distribution Builder

Study 1A
Initial Value = $15,000
Bins = [“less than $11,000”, “$11,000-$12,000”, “$12,000-$13,000”, “$13,000-$14,000”, “$14,000-$15,000”, “$15,000-$16,000”, “$16,000-$17,000”, “$17,000-$18,000”, “$18,000-$19,000”, “greater than $19,000”]

Note: For the analysis, the middle value was used for each bin. For example, a ball placed in the “$15,000-16,000” bin would be given the value of $15,500. For the extreme bins, we assumed a value that would lead to equally spaced bins (e.g., for “greater than $19,000” we used $19,500). The reported analyses are insensitive to other reasonable transformations. A similar approach was used for the ranges in Study 1B.

Study 1B
Initial Value = $350
Bins = [“less than $250”, “$250-$275”, “$275-$300”, “$300-$325”, “$325-$350”, “$350-$375”, “$375-$400”, “$400-$425”, “$425-$450”, “greater than $450”]

Study 1C
Initial Value = $350
Bins = [“$225”, “$250”, “$275”, “$300”, “$325”, “$350”, “$375”, “$400”, “$425”, “$450”, “$475”]

Study 1D and 1E
Initial Value = $3,500
Bins = [“$2,250”, “$2,500”, “$2,750”, “$3,000”, “$3,250”, “$3,500”, “$3,750”, “$4,000”, “$4,250”, “$4,500”, “$4,750”]

Studies 1A–1E: Regression Coefficients from Each Individual Study

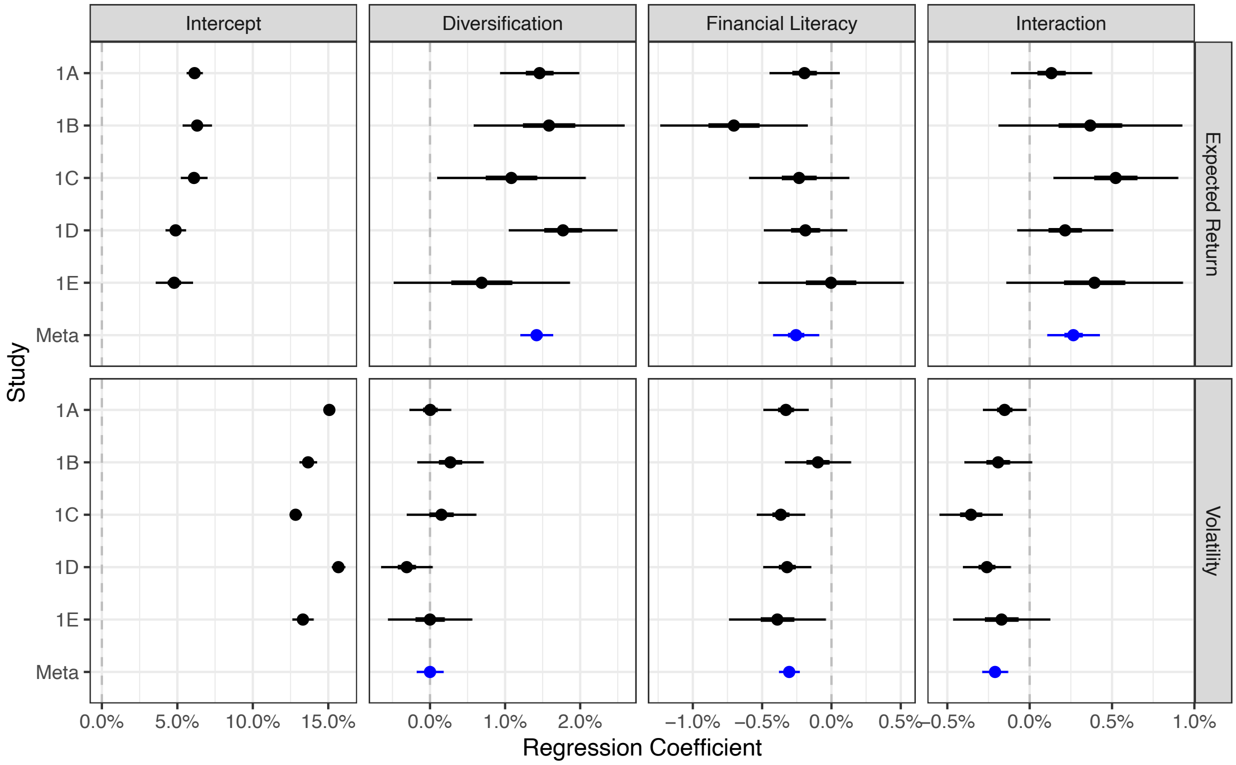


Figure S1. Regression coefficients for Studies 1A–1E. Expected return is operationalized as the mean of each response. Volatility is operationalized as the standard deviation of each response. Thick lines represent 50% confidence intervals and thin lines represent 95% confidence intervals. Financial literacy is centered and diversification is coded .5 = diversified, -.5 = not diversified. Individual study models are fit with cluster-robust errors for participants. The "Meta" model is the cluster-robust standard error model described in the main text (clustering standard errors on studies and participants with fixed effects for studies). Participants who failed the attention check or scored below chance on financial literacy are excluded.

Studies 1A–1E: Results for Other Operationalizations of Expected Return and Volatility

Table S1. Regression coefficients for a meta-analysis of Studies 1A–1E with different operationalizations of expected returns (mean of response, median of response) and volatility (standard deviation of response, middle 50% range / IQR of response, middle 75% range of response). Financial literacy is centered (SD = 2.19) and diversification is coded .5 = diversified, -.5 = not diversified. Coefficients are scaled to reflect the percent of the initial portfolio value. Models are fit with cluster-robust errors for studies and participants and fixed effects for studies. Participants who failed the attention check or scored below chance on financial literacy are excluded. Model-based standard errors shown in brackets.

	Dependent variable:				
	MEAN	MEDIAN	SD	RANGE50	RANGE75
	(1)	(2)	(3)	(4)	(5)
Diversification	1.419***	1.733***	.0001	.273	.110
	(.110)	(.121)	(.090)	(.172)	(.205)
Financial Literacy	−.256**	−.316***	−.305***	−.714***	−.867***
	(.084)	(.085)	(.037)	(.101)	(.109)
Interaction	.265**	.265**	−.210***	−.367***	−.598***
	(.081)	(.096)	(.040)	(.075)	(.180)
Observations	3,284	3,284	3,284	3,284	3,284

Note: *p<0.05; **p<0.01; ***p<0.001

Studies 1A–1E: Results Excluding Repeat Participants

Table S2. Same models as Table S1, but only including responses from the first study a participant completed (i.e., excluding, ex-post, repeat participants).

	Dependent variable:				
	MEAN	MEDIAN	SD	RANGE50	RANGE75
	(1)	(2)	(3)	(4)	(5)
Diversification	1.357^{***}	1.657^{***}	.005	.222	.146
	(.121)	(.108)	(.075)	(.169)	(.192)
Financial Literacy	$-.278^{*}$	$-.348^{***}$	$-.304^{***}$	$-.721^{***}$	$-.875^{***}$
	(.114)	(.100)	(.049)	(.113)	(.130)
Interaction	$.323^{**}$	$.327^{**}$	$-.221^{***}$	$-.385^{***}$	$-.627^{***}$
	(.102)	(.116)	(.033)	(.056)	(.153)
Observations	2,960	2,960	2,960	2,960	2,960

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Studies 1A–1E: Results without Any Participant Exclusions

Table S3. Same models as Table S1, but with no participant exclusions.

	Dependent variable:				
	MEAN	MEDIAN	SD	RANGE50	RANGE75
	(1)	(2)	(3)	(4)	(5)
Diversification	1.147***	1.419***	.045	.256*	.106
	(.121)	(.165)	(.128)	(.129)	(.234)
Financial Literacy	−.125	−.150	−.245***	−.602***	−.767***
	(.072)	(.079)	(.036)	(.049)	(.085)
Interaction	.279***	.288***	−.093***	−.115	−.203**
	(.027)	(.049)	(.023)	(.063)	(.072)
Observations	4,060	4,060	4,060	4,060	4,060

Note: *p<0.05; **p<0.01; ***p<0.001

Studies 1A–1E: Results with Diversification Question Removed from Fin. Lit. Measure

Table S4. Same models as Table S1, but excluding a diversification focused question from the financial literacy measure (question #5: "When an investor spreads his money among different assets, does the risk of losing a lot of money increase, decrease, or stay the same?"; correct answer = decrease). No participants are excluded.

	Dependent variable:				
	MEAN	MEDIAN	SD	RANGE50	RANGE75
	(1)	(2)	(3)	(4)	(5)
Diversification	1.147***	1.419***	.045	.256*	.106
	(.119)	(.163)	(.129)	(.129)	(.235)
Financial Literacy	−.128	−.157	−.259***	−.630***	−.806***
	(.083)	(.089)	(.043)	(.056)	(.100)
Interaction	.281***	.289***	−.102***	−.116	−.223**
	(.024)	(.048)	(.024)	(.072)	(.080)
Observations	4,060	4,060	4,060	4,060	4,060

Note: *p<0.05; **p<0.01; ***p<0.001

Studies 1A–1E: Results using Mixed-Effect Regressions

Table S5. Regression coefficients for a meta-analysis of Studies 1A–1E with different operationalizations of expected returns (mean of response, median of response) and volatility (standard deviation of response, middle 50% range / IQR of response, middle 75% range of response). Financial literacy is centered and diversification is coded .5 = diversified, -.5 = not diversified. Coefficients are scaled to reflect the percent of the initial portfolio value. Models are fit using a mixed-effects approach: Random intercepts are included for participants and studies. A random slope *by participant* is included for diversification. And random slopes *by study* are included for financial literacy, diversification, and their interaction. Participants who failed the attention check or scored below chance on financial literacy are excluded.

	Dependent variable:				
	MEAN	MEDIAN	SD	RANGE50	RANGE75
	(1)	(2)	(3)	(4)	(5)
Diversification	1.426***	1.724***	.027	.209	.139
	(.189)	(.227)	(.123)	(.295)	(.320)
Financial Literacy	−.222*	−.271**	−.300***	−.696***	−.863***
	(.087)	(.104)	(.048)	(.150)	(.136)
Interaction	.271**	.290*	−.222***	−.362**	−.635***
	(.089)	(.122)	(.049)	(.119)	(.173)
(Intercept)	5.763***	6.901***	14.148***	19.112***	32.679***
	(.406)	(.441)	(.581)	(.903)	(1.477)
Observations	3,284	3,284	3,284	3,284	3,284

Note: *p<0.05; **p<0.01; ***p<0.001

Studies 1A–1E: Results using Numeracy Measure

Table S6. Same models as Table S1, but with numeracy (centered; SD = 2.02) instead of financial literacy. The numeracy measure was only collected in Studies 1A and 1B. (Because of the data limitations, for these models SEs are only clustered by participant and not by study.)

	Dependent variable:				
	MEAN	MEDIAN	SD	RANGE50	RANGE75
	(1)	(2)	(3)	(4)	(5)
Diversification	.957*	1.175*	.103	−.145	.280
	(.394)	(.478)	(.186)	(.458)	(.558)
Numeracy	.002	−.007	−.190*	−.753***	−.596*
	(.174)	(.202)	(.093)	(.189)	(.257)
Interaction	.293	.270	−.122	−.175	−.336
	(.178)	(.227)	(.075)	(.201)	(.241)
Observations	782	782	782	782	782

Note: *p<0.05; **p<0.01; ***p<0.001

Studies 1A–1E: Plots of Linear Best Fit Lines

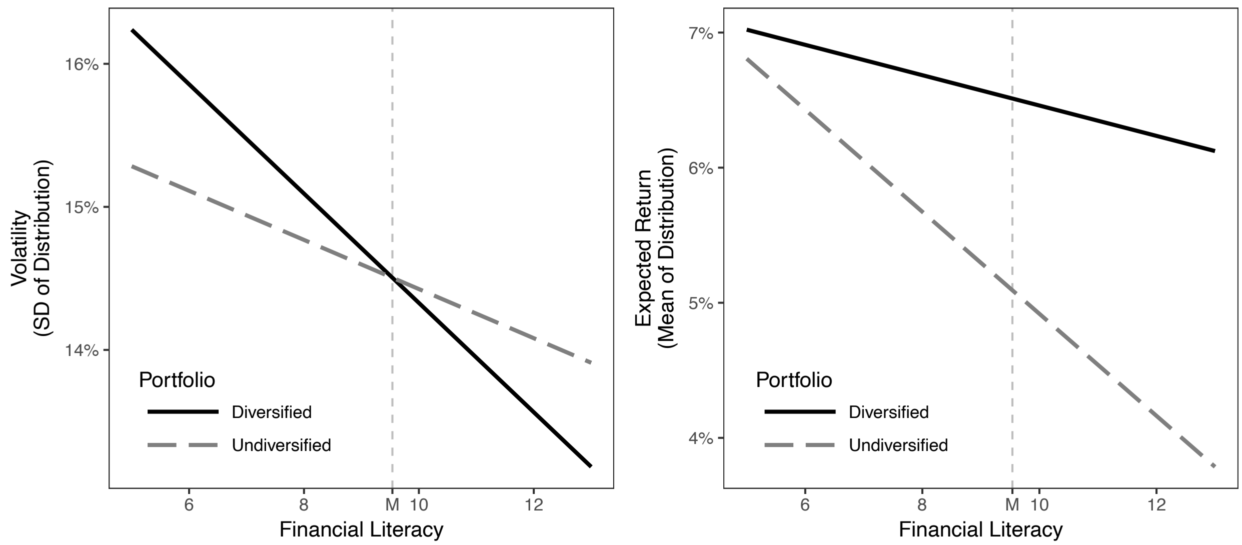


Figure S2. Volatility (standard deviation of expressed return distribution) and expected return (mean of expressed return distribution) by the participant's level of financial literacy in Studies 1A–1E. Linear fits are shown in this plot. The vertical dashed line indicates the mean value of financial literacy in the sample (the difference between the lines at this point can be interpreted as the main effect of diversification).

Follow-Up Study Details

We attempted to recruit all AMT participants who completed a version of Study 1 (messaging done via TurkPrime and the participants' AMT worker ID). Four hundred and eighteen (out of 1649) participants completed the follow-up study. Participants were asked the following questions, based on Almenberg and Dreber (2015):

1) Are you invested in the stock market? For example, do you own any individual stocks? Or are you invested in a mutual fund / ETF?

- no (coded: 0)
- yes (coded: 1)

2) In general, are you a person who is prepared to take risks? Or do you try to avoid taking risks?

- not risk taking at all = 1
- 2
- …
- 9
- very risk taking = 10

3) What is your gender?

- male (coded: 0)
- female (coded: 1)
- other / prefer not to respond (coded: NA)

4) What is your yearly (pre-tax) household income?

- $25,000 or less (coded: 1)
- $25,001–$50,000
- $50,001–$75,000
- $75,001–$100,000
- more than $100,000 (coded: 5)

5) What is your highest level of education?

- some high school (coded: 1)
- high school degree / GED
- some college
- 2 year college degree
- 4 year college degree
- some graduate school
- graduate degree (coded: 5)

6) How old are you? [open response]

7) What best describes your current employment status?

- employed
- self employed
- unemployed
- retired
- student

Follow-Up Study Robustness Analysis

We ran a series of regression models (fixed effects for studies, clustering on participants and studies) to assess the degree to which variables measured in the follow-up study could be confounding the results presented in the main analysis.

The following two tables use data from participants who completed the follow-up study and whose initial responses meet the previously inclusion criteria. We note that some participants did not answer the gender/sex question and are thus excluded. Also, an error prevented us from recording an answer to the "in market" question for some of the participants, who are thus excluded from that model. In the regression models below, we mean center all variables (within the sample; coding scheme shown on previous page) and treat the covariate measures as linear (for ease of presentation), although similar results are obtained using other coding schemes.

Table S7 shows coefficient estimates for models predicting expected return (the mean of the forecast distribution). Table S8 shows coefficient estimates for models predicting volatility (the standard deviation of the forecast distribution). The first model in each table is the base model (on the subset of the data) and the following models add the control variables (and their interaction with the diversification manipulation).

Table S7. Regression coefficients from a regression predicting expected return (operationalized as the mean of the expressed distribution) using data from Study 1A–1E and the follow-up study (fixed effects for studies, standard errors clustered on participants and studies). Predictor variables are mean centered and treated linearly. Measures and coding scheme are described on p. 16.

	Dependent variable:						
	MEAN						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Diversification	1.735***	1.782***	1.741***	1.735***	1.735***	1.735***	1.799***
	(.232)	(.213)	(.223)	(.232)	(.233)	(.216)	(.202)
Financial Literacy	−.490**	−.544***	−.431**	−.549***	−.528**	−.450*	−.427*
	(.169)	(.145)	(.155)	(.149)	(.171)	(.197)	(.167)
In Market?		.418					.176
		(.326)					(.421)
Gender			1.787**				1.957***
			(.578)				(.521)
Income				.462**			.216
				(.145)			(.184)
Education					.169		.002
					(.268)		(.288)
Age						−.023	−.039
						(.026)	(.032)
Div. x Fin. Lit.	.186**	.142	.200**	.184	.177**	.246**	.244
	(.067)	(.119)	(.061)	(.099)	(.058)	(.086)	(.139)
Div. x In Market?		.173					.204
		(.553)					(.654)
Div. x Gender			.342				.661*
			(.336)				(.306)
Div. x Income				.014			−.160
				(.320)			(.443)
Div. x Education					.039		.026
					(.152)		(.222)
Div. x Age						−.036	−.040
						(.022)	(.036)
Observations	832	776	826	832	832	832	770

Note: *p<0.05; **p<0.01; ***p<0.001

Table S8. Regression coefficients from a regression predicting volatility (operationalized as the standard deviation of the expressed distribution) using data from Study 1A–1E and the follow-up study (fixed effects for studies, standard errors clustered on participants and studies). Predictor variables are mean centered and treated linearly. Measures and coding scheme are described on p. 16.

	Dependent variable:						
	SD						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Diversification	−.176 (.240)	−.151 (.235)	−.149 (.240)	−.176 (.239)	−.176 (.239)	−.176 (.245)	−.126 (.243)
Financial Literacy	−.342** (.122)	−.341** (.115)	−.349** (.119)	−.327** (.108)	−.342** (.126)	−.348** (.131)	−.350** (.123)
In Market?		−.258 (.309)					−.037 (.423)
Gender			−.205 (.275)				−.129 (.353)
Income				−.114 (.159)			−.210 (.186)
Education					.004 (.087)		.062 (.135)
Age						.004 (.023)	.002 (.024)
Div. x Fin. Lit.	−.146** (.055)	−.172* (.083)	−.140* (.057)	−.141* (.061)	−.119* (.051)	−.167* (.066)	−.163 (.085)
Div. x In Market?		.262 (.529)					.430 (.492)
Div. x Gender			.002 (.118)				−.014 (.146)
Div. x Income				−.034 (.102)			.018 (.107)
Div. x Education					−.120 (.117)		−.189* (.087)
Div. x Age						.013 (.015)	.015 (.016)
Observations	832	776	826	832	832	832	770

Note: *p<0.05; **p<0.01; ***p<0.001

Study 2A: Robustness Analysis

Results without Any Participant Exclusions

Without any participant exclusions, we observe the same effects with similar statistical significance. There was a small difference between the portfolios on the predictability measure ($M = .24$, $t = 2.78$, $p = .006$) and a moderating effect of financial literacy ($\beta_{\text{FinLit}} = .12$, $t = 4.26$, $p < .001$). There was a significant difference on the expected return measure between the two portfolios, with people expecting the diversified portfolio to increase in value more than the undiversified portfolio ($M = 1.34$, $t = 26.78$, $p < .001$) and an insignificant moderating effect of financial literacy ($\beta_{\text{FinLit}} = .02$, $t = 1.04$, $p = .30$).

Results with Diversification Question (#5) Removed from Financial Literacy Measure

Analysis without participant exclusions and without the diversification question in the financial literacy measure is redundant for the main effects (see previous section) and yields similar results for the financial literacy analysis on the predictability measure ($\beta_{\text{FinLit}} = .13$, $t = 4.11$, $p < .001$) and the increase in value measure ($\beta_{\text{FinLit}} = .01$, $t = .78$, $p = .44$).

Analysis of Individual Stocks

The undiversified portfolio had only one stock and it was randomly chosen from the four in the diversified portfolio. We can thus further examine the results by individual stocks:

For the predictability measure, an ANOVA using financial literacy and the specific stock (comprising the undiversified portfolio) as predictor variables suggests that the individual stocks were not perceived to be different in terms of predictability relative to the diversified (four stock) portfolio ($F(3, 356) = .10$, $p = .96$). Nor was there evidence of an interaction between specific stock and financial literacy ($F(3, 356) = .82$, $p = .48$), suggesting the moderating role of financial literacy on the difference in predictability between the portfolios did not differ based on the specific stock.

For the expected return measure, an ANOVA using financial literacy and the specific stock (comprising the undiversified portfolio) as predictor variables suggests that the individual stocks were not perceived to be different in terms of expected return relative to the diversified (four stock) portfolio ($F(3, 356) = .21$, $p = .89$). There was also no interaction between specific stock and financial literacy ($F(3, 356) = 1.80$, $p = .15$).

Study 2B: Robustness Analysis

Results without Any Participant Exclusions

We can examine the main effects in the model using non-parametric tests without excluding participants. Wilcoxon signed rank tests on the full data set are consistent with the results reported in the manuscript: People gave larger 90% confidence intervals for the diversified portfolio ($V = 14074$, $p < .001$) and expected a higher return for the diversified portfolio ($V = 39412$, $p < .001$).

To examine the interactions, we can use a similar approach by converting the dependent variables (difference in 90% confidence interval ranges and difference in expected returns) to ranks and running the same OLS models as reported in the paper. This yields similar conclusions to the reported results in the paper: An interaction on the volatility measure such that those low in financial literacy are more likely to give bigger 90% confidence intervals for the diversified portfolio ($\beta_{\text{FinLit}} = -5.77$, $t = -3.18$, $p = .002$). An interaction on the expected return measure such that those high in financial literacy expect the difference in returns to be smaller (but still favoring the diversified portfolio; $\beta_{\text{FinLit}} = -4.20$, $t = -2.30$, $p = .022$).

Results with Diversification Question (#5) Removed from Financial Literacy Measure

Analysis without participant exclusions and without the diversification question in the financial literacy measure is redundant for the main effects (see previous section) and yields similar results for the financial literacy analysis on the predictability measure ($\beta_{\text{FinLit}} = -5.91$, $t = -3.01$, $p = .003$) and the increase in value measure ($\beta_{\text{FinLit}} = -4.33$, $t = -2.19$, $p = .029$).

Analysis of Individual Stocks

The undiversified portfolio had only one stock and it was randomly chosen from the four in the diversified portfolio. We can thus further examine the results by individual stocks:

For the volatility (90% confidence interval) measure, an ANOVA using financial literacy and the specific stock (comprising the undiversified portfolio) as predictor variables suggests that the individual stocks were perceived to be different in terms of predictability ($F(3, 286) = 2.50$, $p = .060$). People gave the widest confidence intervals for PetSmart compared to the diversified portfolio ($M = \$1347$) and the smallest for Facebook ($M = \$380$). However, all stocks were given wider confidence intervals than the diversified portfolio ($|t|$s > 1.49, ps $< .14$; all but Facebook significant at $p < .01$). There was no interaction between specific stock and financial literacy ($F(3, 286) = .98$, $p = .40$), suggesting the moderating role of financial literacy on the difference in 90% confidence intervals between the portfolios did not differ based on the specific stock.

For the “increase in value” measure, an ANOVA using financial literacy and the specific stock (comprising the undiversified portfolio) as predictor variables suggests that the individual stocks were perceived to be different in terms of expected return ($F(3, 286) = 3.45$ $p = .017$). Facebook was perceived to be the closest (in terms of expected return) to the diversified portfolio ($M = \$351$) and PetSmart was perceived to be the furthest ($M = \$1{,}136$). However, the diversified portfolio was expected to outperform *each* of the individual stocks (ts > 1.65, ps $< .11$; all but Facebook significant at $p < .01$). There was again no interaction between specific stock and financial literacy ($F(3, 286) = 1.48$, $p = .22$).

Study 2B: Plots of Linear Best Fit Lines

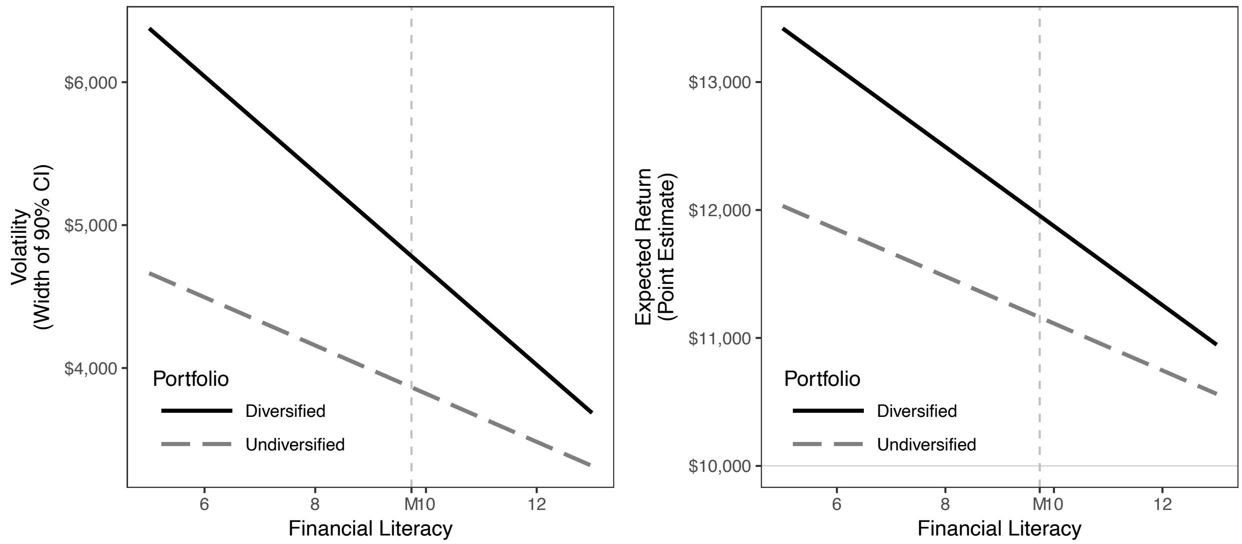


Figure S3. Width of 90% confidence intervals (volatility; left panel) and point estimates for expected return (right panel). Linear fits are shown in this plot. The vertical dashed line indicates the mean value of financial literacy in the sample.

Study 2B: Wording for Confidence Interval Question

Please give a range such that you think there is a 90% chance the value of this portfolio in exactly one year will be somewhere in this range. In other words, provide a range such that you would expect to be wrong only about one out of ten times.

Study 3: Experimental Stimuli

One the first screen participants were told:

“This study includes an opportunity to win a **$1 bonus payment**. Please pay attention to the instructions to maximize your chance of winning this bonus!”

The portfolios were described:

Imagine two people who are going to invest in the stock market. Both people will invest $10,000 on August 5th, 2019.

Person A will invest all $10,000 in Cisco Systems (CSCO) [Oracle Corporation (ORCL)] stock.

Person B will divide her investment between two different stocks. She will invest $5,000 in Oracle Corporation (ORCL) stock and $5,000 in Cisco Systems (CSCO) stock.

Expected Value Question (incentive wording in brackets):

Whose portfolio--Person A or Person B--do you think will have a **greater** value at the end of one month (on September 5th, 2019)?

We will determine the "correct" answer by assessing and comparing the values of each portfolio at the end of the month. [**If you answer this question correctly, we will award you with a $1 bonus payment**. Bonus payments for correct answers will be made on Sept. 2nd.]

Volatility Question (incentive wording in brackets):

Whose portfolio--Person A or Person B--do you think will have a **more predictable** value at the end of one month (on September 5th, 2019)?

We will determine the "correct" answer by calculating the standard deviation of daily portfolio returns over the month. We will call the portfolio with a *smaller* standard deviation in daily portfolio returns the portfolio with the more predictable value. [**If you answer this question correctly, we will award you with a $1 bonus payment**. Bonus payments for correct answers will be made on Sept. 2nd.]

Study 3: Robustness Analysis

Results without Any Participant Exclusions

Fifty-seven percent of participants indicated the single stock (undiversified) portfolio would have a more predictable value ($\chi^2 = 7.23$, $p = .007$). There was no effect of the incentive on participants' forecasts (logistic regression: $\beta_{\text{Incentive}} = -.08$, $z = -.75$, $p = .45$). Participants low in financial literacy had a greater tendency to indicate the single stock (undiversified) portfolio would have a more predictable value (logistic regression: $\beta_{\text{FinLit}} = .12$, $z = 3.03$, $p = .002$).

Sixty-eight percent of participants indicated the diversified portfolio would have a greater value ($\chi^2 = 52.69$, $p < .001$). This was not moderated by incentive (logistic regression: $\beta_{\text{Incentive}} = -.12$, $z = -1.15$, $p = .25$) or financial literacy (logistic regression: $\beta_{\text{FinLit}} = .04$, $z = 1.03$, $p = .30$).

Results with Diversification Question (#5) Removed from Financial Literacy Measure

Analysis without participant exclusions and without the diversification question in the financial literacy measure is redundant for the main effects (see previous section) and yields similar results for the financial literacy analysis on the predictability measure ($\beta_{\text{FinLit}} = .13$, $z = 3.10$, $p = .002$) and the increase in value measure ($\beta_{\text{FinLit}} = .04$, $z = .99$, $p = .32$).

Analysis of Individual Stocks

When the single stock (undiversified portfolio) was Cisco, 61% of participants thought it would have a more predictable value than the diversified portfolio vs. 53% for Oracle.

When the single stock (undiversified portfolio) was Cisco, 72% of participants thought the diversified portfolio would have a greater return vs. 65% for Oracle.

Studies 4A and 4C: Description of Investors

Gain-Seeking Investor Profile:

[Stock Picture of Younger Woman]

Name: Janelle Thompson

Age: 27

Investment Goals: Janelle has been working at her job for 5 years. She has saved up some money and wants to invest it in a portfolio of stocks that will yield a high return. Janelle is willing to tolerate unpredictability and volatility from her investments. She just wants her investments to make money!

Risk-Averse Investor Profile:

[Stock Picture of Older Woman]

Name: Doris Westward

Age: 63

Investment Goals: Doris is about to retire and wants predictability from her investments. After retirement, Doris plans to withdraw some money from her portfolio each year for living expenses. She doesn't need her portfolio to make a lot of money, she just wants a stable source of income in retirement.

Study 4B: Description of Investors

Gain-Seeking Investor Profile:

Name: Ms. R

Investment Goals: Ms. R wants to invest in a portfolio of stocks that will yield a high return. She is willing to tolerate unpredictability and volatility from her investments. She just wants her investments to make money!

Risk-Averse Investor Profile:

Name: Ms. S

Investment Goals: Ms. S wants predictability from her investments. She doesn't need her portfolio to make a lot of money, she just wants stable returns.

Studies 4A and 4B: Available Stocks

Apple, Google, Amazon, Microsoft, Oracle, Verizon, Cisco, Qualcomm, Comcast, Intel, Pfizer, Philip Morris, CitiGroup, IBM, Merck, Berkshire Hathaway, Wells Fargo, Exxon, Chevron, AT&T, Bank of America, JP Morgan Chase, Home Depot, Walt Disney, PepsiCo, Walmart, Procter and Gamble, General Electric, Johnson and Johnson, and Coca-Cola.

Study 4C: Example of Stock Fund Stimuli

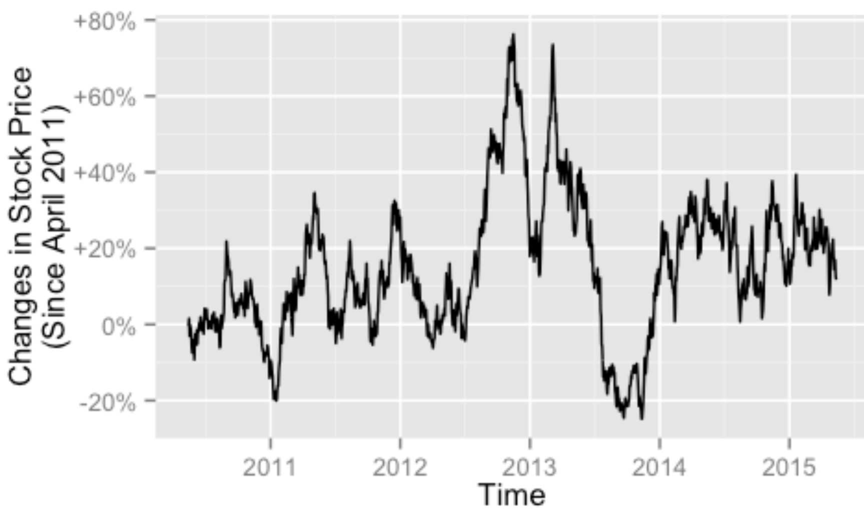


Figure S4. Example of the “historical returns” a participant would see if they hovered their mouse pointer over the name of one of the funds.

Studies 4A and 4B: Robustness Analysis

Results without Any Participant Exclusions

Results without exclusions replicate those with exclusions: In Study 3A, the average participant gave a more diversified portfolio to the gain-seeking investor (M = 6.12 stocks) compared to the risk-averse investor (M = 5.76; t = -1.86, p = .064). This was moderated by financial literacy (β_{FinLit} = .03, t = 2.28, p = .024). In Study 3B, the average participant gave a more diversified portfolio to the gain-seeking investor (M = 5.53 stocks) compared to the risk-averse investor (M = 5.27; t = -1.75, p = .074). This was moderated by financial literacy (β_{FinLit} = .04, t = 4.35, p < .001).

Results with Diversification Question (#5) Removed from Financial Literacy Measure

Analysis without participant exclusions and without the diversification question in the financial literacy measure yields nearly identical results. The main effects reported in the above section are unchanged, as they do not involved the financial literacy measure. The moderating results of financial literacy without the diversification question are the same (at the level of precision reported—the differences are obscured by rounding) in both Study 3A (β_{FinLit} = .03, t = 2.28, p = .024) and 3B (β_{FinLit} = .04, t = 4.35, p < .001).

Results without Log Transformation

In the paper and previous analysis, we log transformed the number of stocks assigned to each portfolio, because the distributions were right skewed across participants. Similar, but slightly less significant, results are obtained without this transformation: In Study 3A, the average participant gave a more diversified portfolio to the gain-seeking investor (M = 7.25 stocks) compared to the risk-averse investor (M = 6.83; t = -1.47, p = .14). This was moderated by financial literacy (β_{FinLit} = .25, t = 1.95, p = .053). In Study 3B, there was no difference—on average—between the gain-seeking investor (M = 6.55 stocks) compared to the risk-averse investor (M = 6.64) in terms of diversification (t = .35, p = .72). However, the moderating effect of financial literacy was still strong (β_{FinLit} = .40, t = 3.70, p < .001).

Studies 4A and 4B: Plots of Linear Best Fit Lines

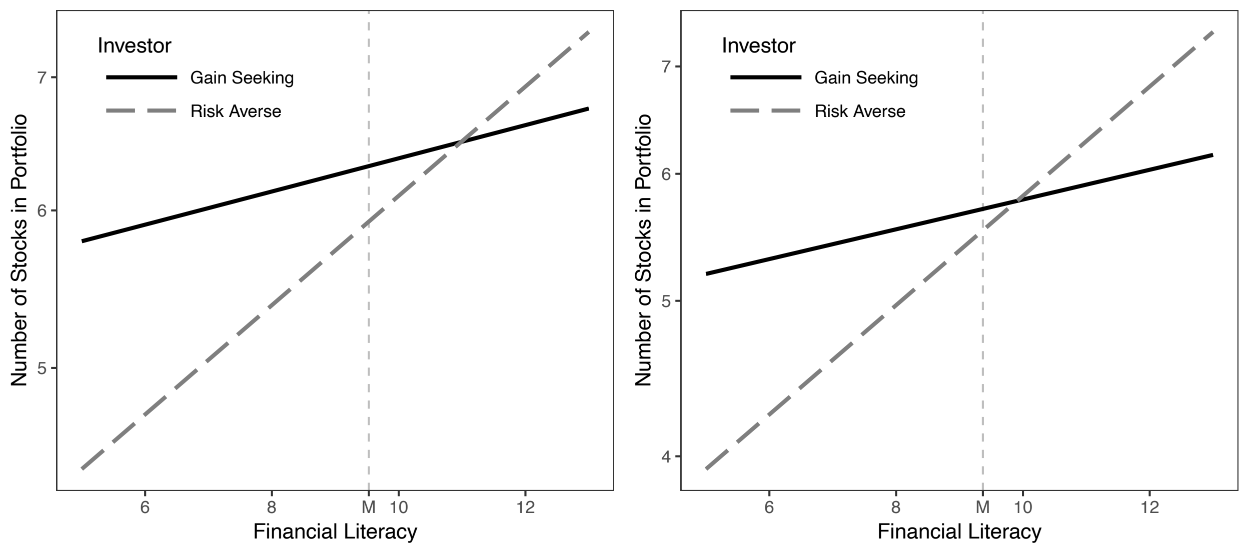


Figure S5. Number of stocks assigned to each investor in Studies 4A (left panel) and 4B (right panel) by financial literacy. Linear fits are shown (regressions on log number of stocks). The vertical dashed line indicates the mean value of financial literacy in the sample (the difference between the lines at this point can be interpreted as the main effect of diversification).

Studies 4A and 4B: Analysis of Individual Stocks

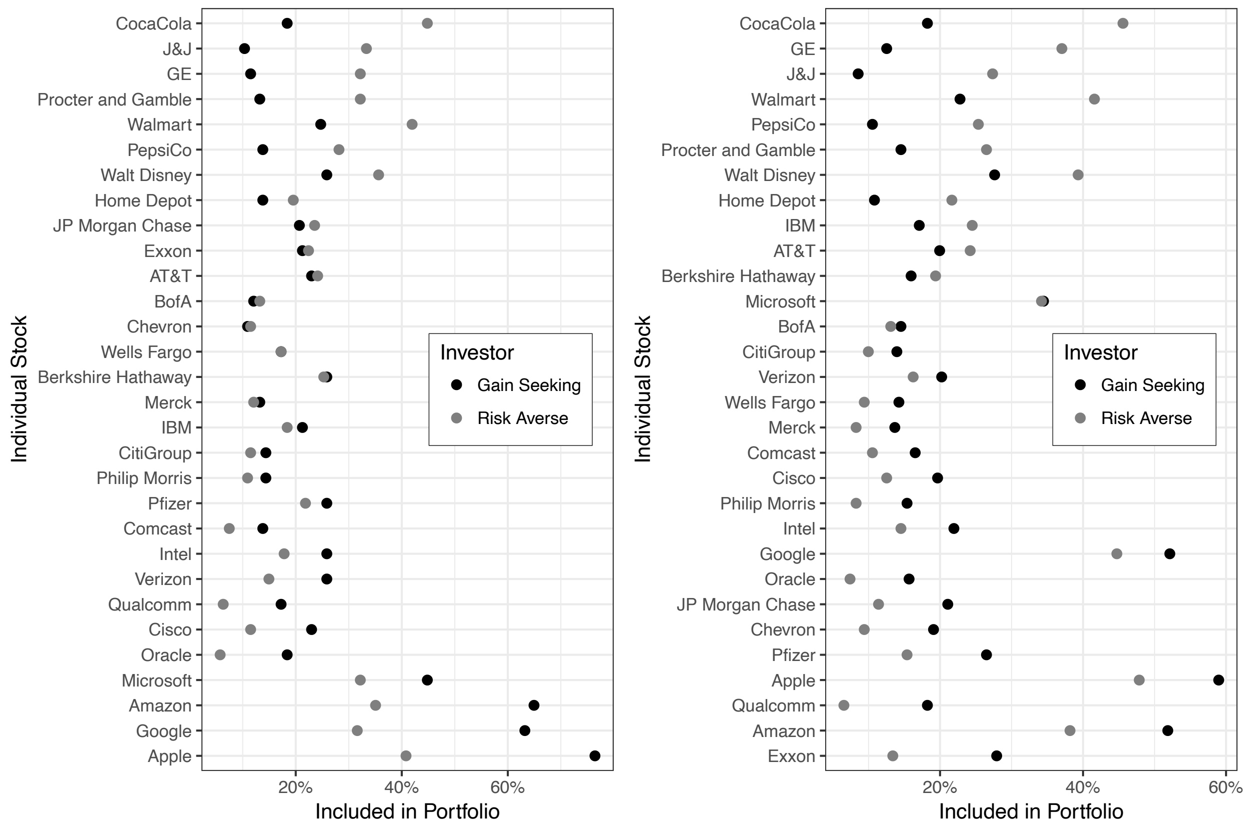


Figure S6. Analysis of portfolio composition in Studies 4A (left panel) and 4B (right panel). Dots represent the proportion of portfolios the given stock was included in for each of the two investors. Companies are ordered by the absolute difference between the two investors: Stocks more likely to be included in the risk-averse investor's portfolio are at the top and stocks more likely to be included in the gain-seeking investor's portfolio are at the bottom.

Study 4C: Robustness Analysis

Results without Any Participant Exclusions

Results without exclusions replicate those with exclusions: The average participant did not differ in the number of stocks assigned to the gain-seeking investor's portfolio (M = 4.49 stock funds) compared to the risk-averse investor's portfolio (M = 4.39 stock funds; t = -.57, p = .57). But, this was moderated by financial literacy such that low financial literacy participants included fewer stocks in the risk-averse investor's portfolio (β_{FinLit} = .05, t = 3.74, p < .001). Similarly, there was no difference for the average participant on the standard deviation measure between the gain-seeking investor's portfolio (M = 14.13%) and the risk-averse investor's portfolio (M = 14.36%; t = .40, p = .69). But, again, this was moderated by financial literacy such that the low financial literacy participants gave the risk-averse investor a less diversified portfolio (β_{FinLit} = -.79%, t = -3.48, p < .001).

Results with Diversification Question (#5) Removed from Financial Literacy Measure

Analysis without participant exclusions and without the diversification question in the financial literacy measure yields nearly identical results. The main effects reported in the above section are unchanged, as they do not involve the financial literacy measure. The moderating results of financial literacy without the diversification question for the number of stocks (β_{FinLit} = .06, t = 3.79, p < .001) and the standard deviation measure (β_{FinLit} = -.84%, t = -3.52, p < .001) are similar to the above analysis.

Results without Log Transformation

In the paper and previous analysis, we log transformed the number of stocks assigned to each portfolio, because the distributions were right skewed across participants. Similar results are obtained without this transformation: The average participant did not differ in the number of stocks assigned to the gain-seeking investor's portfolio (M = 5.09 stock funds) compared to the risk-averse investor's portfolio (M = 5.07 stock funds; t = -.11, p = .91). But, this was moderated by financial literacy such that low financial literacy participants included fewer stocks in the risk-averse investor's portfolio (β_{FinLit} = .25, t = 3.42, p < .001).

Study 4C: Plots of Linear Best Fit Lines

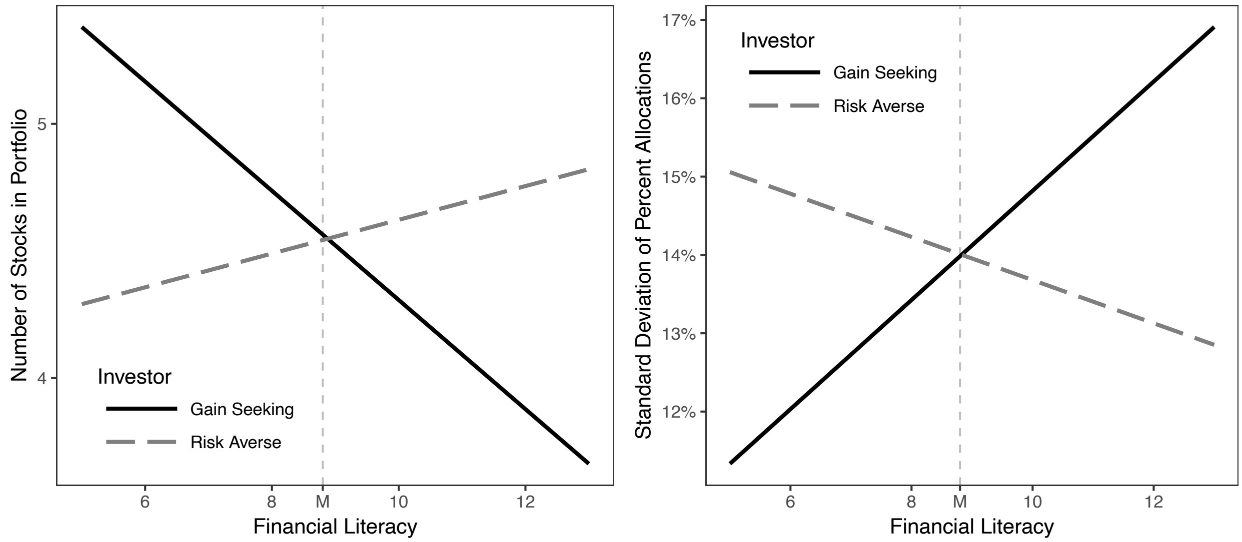


Figure S7. Number of stocks assigned to each investor (left panel; linear fits on log number of stocks) and the standard deviation of percent allocations (right panel; a measure of diversification, with higher values indicating better diversification). The vertical dashed line indicates the mean value of financial literacy in the sample (the difference between the lines at this point can be interpreted as the main effect of diversification).

Study 5: Results from Coding

Table S11. Proportion of participant responses coded into each category, by question type and response.

	Better Return		More Predictable	
	Not Div.	Div.	Not Div.	Div.
"Diversify"	0	0.50	0.05	0.33
Ind. Stocks	0.28	0.46	0.54	0.39
More Vars.	0	0	0.76	0
Gains	0.33	0.44	0.06	0.25
Losses	0.11	0.34	0.05	0.24
"Risk"	0.22	0.39	0.08	0.29
Correct Int.	0.06	0.19	0	0.33
"Eggs in Basket"	0	0.09	0	0.04
Fin. Advice	0.06	0.05	0	0.02
Probability	0.06	0.28	0.06	0.18
Extremity	0	0.11	0.03	0.12

Notes: Results are separated by question type and the participant's response. For example, the "More Predictable"-"Div." column indicates the proportion of responses coded into each category for participants who were asked which portfolio would have a more predictable value at the end of the year and chose the diversified (larger) portfolio. Codes representing each category described below.

1) "Diversify": Did the response mention the term diversification, diversify, etc.?
2) Ind. Stocks: Did the response mention something about the performance of specific/individual stocks?
3) More Vars.: Did the response talk about the difficulty of aggregating information across many stocks?
4) Gains: Did the response reference outcomes in the gain domain?
5) Losses: Did the response reference outcomes in the loss domain?
6) "Risk": Did the response mention the term risk, riskiness, etc.?
7) Correct Int.: Did the response indicate a correct intuition about diversification?
8) "Eggs in Basket": Did the response mention the phrase "eggs in a basket" or a variant thereof?
9) Fin. Advice: Did the response mention financial advice, best practices, etc.?
10) Probability: Did the response mention probability, chance, etc.?
11) Extremity: Did the response reference the extremity of outcomes?

NOTE: All text responses and coding are available in the S5 data file hosted on OSF (https://osf.io/hnj5y/?view_only=6d865bb6f5ad4315a70d2277af68ae9b).

Study 5: Robustness Analysis

Results without Any Participant Exclusions

Sixty-six percent of participants who answered the predictability question indicated that the diversified portfolio would have a more predictable value in one year. People higher in financial literacy were more likely to say the diversified portfolio was more predictable (logistic regression: $\beta_{\text{FinLit}} = .17$, $z = 3.26$, $p = .001$). Eighty-nine percent of participants who answered the return question indicated the diversified portfolio would have a higher return. People higher in financial literacy were more likely to say the diversified portfolio would have a greater return (logistic regression: $\beta_{\text{FinLit}} = .31$, $z = 3.99$, $p < .001$).

Results with Diversification Question (#5) Removed from Financial Literacy Measure

Analysis without participant exclusions and without the diversification question in the financial literacy yield similar (predictability measure: $\beta_{\text{FinLit}} = .19$, $z = 3.24$, $p = .001$; return measure: $\beta_{\text{FinLit}} = .32$, $z = 3.85$, $p < .001$).

Study 6: Results without Any Participant Exclusions

For the predictability measure, an ANOVA revealed significant differences by condition ($F(2, 995) = 9.19$, $p < .001$). The condition in which participants elaborated on the individual stocks led to the lowest ratings of predictability ($M = 5.08$). This was significantly lower than the control condition ($M = 5.55$; $F(1, 995) = 18.00$, $p < .001$), but not significantly different from the condition in which people elaborated on the concept of diversification ($M = 5.25$; $F(1, 995) = 2.07$, $p = .15$).

For the return measure, an ANOVA revealed significant differences by condition ($F(2, 995) = 13.74$, $p < .001$). The condition in which people elaborated on the concept of diversification led to the highest expectations of return ($M = +2.61$). This was significantly different from the condition in which people elaborated on the individual stocks ($M = +1.92$; $F(1, 995) = 25.83$, $p < .001$), but not significantly different from the control condition ($M = +2.41$; $F(1, 995) = 1.83$, $p = .18$).

Simulation: How Diversification Can Affect Simple (Mean) and Median Returns

To provide an intuition for how diversification can reduce median returns without affecting simple (mean) returns, we conducted a simple simulation.

We simulated 1,000,000 returns for two portfolios. One portfolio ("undiversified") contained a single asset. The other portfolio ("diversified") contained ten (uncorrelated) assets, with equal allocations in each. Each asset's return came from the same, log-normal (right skewed) generating distribution: Return = $e^{N(.13,\ .30)} - 1$.

Both portfolios had the same mean return (~19%), but the portfolios differed in the median return (diversified = 19% vs. undiversified = 14%).

We plot the return distributions, with lines indicated the medians, in Figure S8 below.

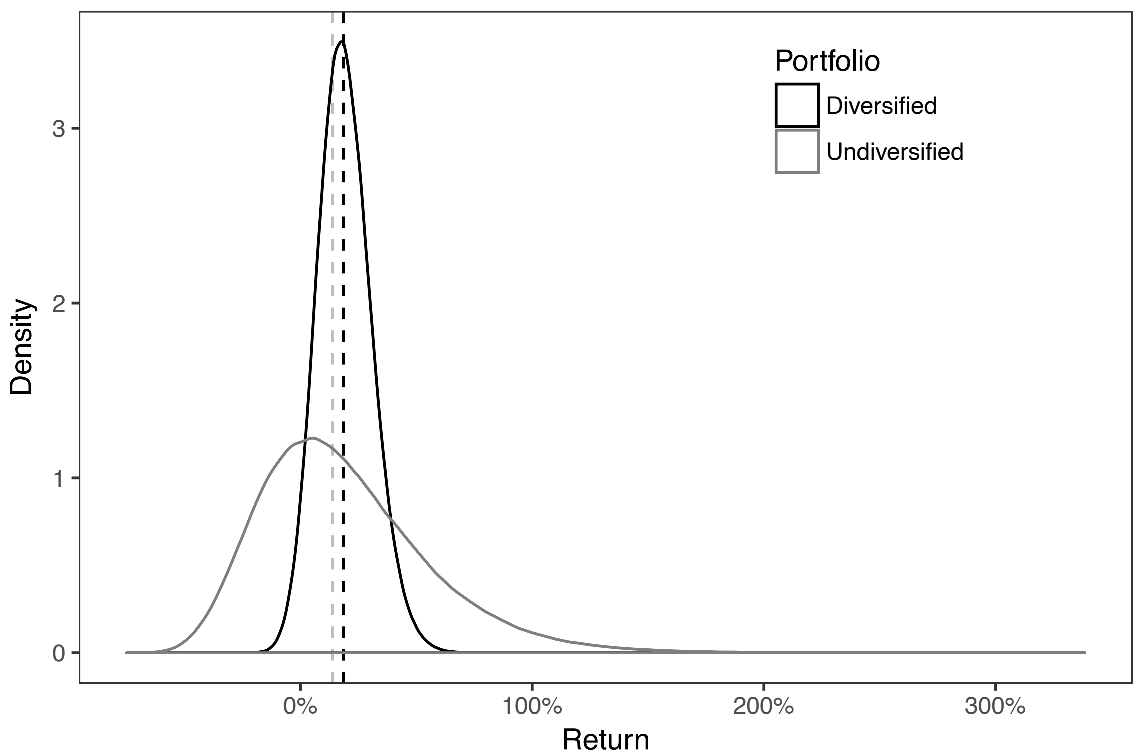


Figure S8. Simulated return distributions for a diversified (10 stock) and undiversified (1 stock) portfolio. Dashed lines show the medians of each distribution. The mean of both distributions is the same (minus simulation noise) and very similar to the median of the diversified distribution.