Measuring Long-Run Expectations that Correlate with Investment Decisions*

Peter Haan † Chen Sun ‡ Felix Weinhardt § Georg Weizsäcker ¶

July 20, 2025

Abstract

Different methods of eliciting long-run expectations yield data that predict economic choices differently well. We ask members of a wide population sample to make a 10-year investment decision and to forecast stock market returns in one of two formats: they either predict the average of annual growth rates over the next 10 years, or they predict the total, cumulative growth that occurs over the 10-year period. Results show that total 10-year forecasts are more pessimistic than average annual forecasts, but they better predict experimental portfolio choices and real-world stock market participation.

JEL-classification: D01, D14, D84, D9
Keywords: Household finance, long-run predictions, survey experiments

*The authors thank Taisuke Imai and audience members at Humboldt-Universität zu Berlin and Lithuania Central Bank for feedback and comments, and the German Science Foundation for financial support via CRC TRR 190 (project number 280092119) and FOR 5675 (project numbers 518302089 and 545786499). The analysis was preregistered on AsPredicted #146168, `https://aspredicted.org/d755-xtfs.pdf` before the data collection.

†Freie Universität Berlin, DIW Berlin, and Berlin School of Economics, phaan@diw.de

‡Fudan University, Shanghai Institute for Mathematics and Interdisciplinary Sciences, Humboldt-Universität zu Berlin, and Berlin School of Economics, chen.sun@hu-berlin.de

§European University Viadrina, Berlin School of Economics, IZA, CESifo, and CEP at LSE, weinhardt@europa-uni.de

¶Humboldt-Universität zu Berlin and Berlin School of Economics, weizsaecker@hu-berlin.de

1 Introduction

Important economic decisions typically concern outcomes that materialize in the long run and therefore require long forecasts. Yet, long-run expectations are largely unstudied and, in particular, the search for optimal methods of eliciting long-run expectations is still in its infancy. Our paper focuses on a salient design question that arises when collecting expectations over spans of multiple years: should one ask about the total (multi-year) change, or the average change per year? Under basic assumptions of consistency, both measures of long-run expectations yield identical insights. Yet, aggregation over time is known to be a non-trivial task for most people (demonstrated, e.g., by the literature on exponential growth bias in Levy and Tasoff, 2016) such that the data resulting from the two measures may well differ in actual fact.

To assess the values of elicitation methods, we examine the correlation between the obtained expectations and the relevant choices. This approach is in line with the decision-theoretic view of economic choice: the importance of expectations lies in their role of justifying behavior – actions, but not expectations, have economic consequences and are deemed as directly welfare-relevant. Accordingly, expectations data are valuable via their predictive power for choice behavior. Our main result is that the question format indeed matters for predictive power: forecasts of total growth over a multi-year period – henceforth, data collected via the *total method* – show a correlation with actual choice behavior that is about twice as large as the corresponding correlation between choices and data collected via the *average method* that asks for expected average annual growth. This pattern appears even though the predictions under the total method are significantly more pessimistic about the stock market, compared to predictions under the average method, and therefore further from the historical growth rates.[1]

Most existing measurements of economic expectations concern the short and medium run. Many research institutions regularly elicit short- and medium-run expectations of households and firms, and including such expectations data into model estimations has become a modern tool of analysis (see the collection of surveys in Bachmann et al., 2023). In the rare cases where researchers do elicit long-run expectations, they usually use the average method. For example, the CFO Survey conducted by Duke University and the Federal Reserve Bank of Richmond elicits 10-year average annual S&P 500 returns, while the Survey of Professional Forecasters by the Federal Reserve Bank of Philadelphia elicits 10-year average annual expectations for S&P 500 returns, inflation, real GDP growth, and other economic variables. This choice of method is presumably made to enhance comparability with one-year-ahead expectations. We are not aware

[1] Consistent with previous work (Arrondel et al., 2014; Breunig et al., 2021a), all our measured expectations about long-run stock market growth are very pessimistic overall.

of any systematic methodological analysis of long-run belief elicitation, including any assessment of predictive power for investment behavior.[2] Notably, the methodic question came up very explicitly in the history of the Case-Shiller house price indices. Shiller and Thompson (2022) describe that the typical response to the total method is far below the theoretical value that would result from the average method – a qualitative result that we also observe in our data – but the authors do not compare the predictive power of the two measures: their solution is to disregard the total growth questions altogether and to use average-growth questions throughout their analysis.

Our study's expectations data concern stock market growth, not housing. It follows the growing literature on expectations in financial markets (e.g. Vissing-Jorgensen, 2004; Dominitz and Manski, 2011; Hurd et al., 2011; Merkle and Weber, 2014; Amromin and Sharpe, 2014; Breunig et al., 2021b; Laudenbach et al., 2023). Diercks and Jendoubi (2023) give an overview of methods and surveys among financial market participants, including surveys where specialist participants are asked about the more distant future, up to ten years. While this literature has a serious emphasis on the medium- and long-run expectations of experts, it only rarely collects them in household surveys.[3]

We are, of course, not the first to emphasize the importance of correlations with investments as a measure of the methodic value of expectations data. A prominent previous emphasis was given by Brunnermeier et al. (2021). Many other studies, however, aim at maximizing the accuracy of forecasts of relevant economic variables (Knüppel and Vladu, 2016; Snowberg et al., 2013; Grothe and Meyler, 2018; Ang et al., 2007; Faust and Wright, 2009; Aiolfi et al., 2010; Clements, 2015; Berge, 2018) and do not ask what type of data is most predictive of investments. Here, we hasten to point out that while expectations are shown to correlate with economic choice (e.g. Hurd et al., 2011; Arrondel et al., 2014; Merkle and Weber, 2014; Amromin and Sharpe, 2014; Knüpfer et al., 2017; Breunig et al., 2021b; Zimpelmann, 2021; Laudenbach et al., 2023), this correlation is not very strong (Ameriks et al., 2020; Giglio et al., 2021; Meeuwis et al., 2022). In our view, this is not a reason to ignore the issue of correlation – rather, it only invites more questions about the right method of asking for expectations.[4]

To test which elicitation method is more informative about the investment behavior of indi-

[2]Some surveys use a variant of the average method, by eliciting single-year expectations for a distant future period. For example, the Consumer Expectations Survey conducted by the European Central Bank asks about inflation between 2 and 3 years from now.

[3]An exception is Breunig et al. (2021a).

[4]Bailey et al. (2018) is a good example of an alternative empirical strategy for demonstrating fairly strong causal effects of beliefs. Using elaborate sets of regressions, the authors show that real-estate market behavior is influenced by experiences of people in one's own social network, and they rule our alternative explanations. The strategy does not employ quantitative statements of expectations.

viduals, we use the experimental set-up in Breunig et al. (2021b) and design an online survey experiment with an incentivized financial decision. We ask members of the U.S. population to make a simplified portfolio choice with a 10-year investment horizon. In addition, we randomly split the sample into two different treatment groups and elicit expectations about the 10-year stock market return. The first treatment group is asked using the average method, and the second group using the total method. We then use a regression analysis to test which elicitation measure correlates more strongly with the incentivized investment choice of the individual. We include various robustness checks: using both point and probabilistic measures of expectations, annualizing total expectations in both geometric and arithmetical ways, performing both between-subject and within-subject comparisons, and using real-world stock market participation as the outcome variable in addition to the simplified portfolio choice from our survey experiment. We also provide tests for heterogeneous subgroups (all pre-registered) differing in gender, age, education, financial literacy, and stock investment experience.

We find that while expectations elicited with the total method are very pessimistic – on average 76% lower than those elicited with the average method – they are a far better predictor of individuals' portfolio choices, with correlations roughly twice as high. This finding is robust regardless of whether expectations are elicited by asking for the mean or the probability, and regardless of whether the total expectation is annualized in a geometric or arithmetical way. In addition, we find that the differences between average annual expectations and total expectations in their correlation with investment behavior are smaller for men, older people, and the financially literate, who are often considered to behave more like *homo economicus*. The results show that higher accuracy of expectations does not imply greater economic relevance; asking for the total price change over a long period yields expectations that are less accurate but more predictive of investment choices than eliciting the average price change per year.

The paper is organized as follows. Section 2 introduces the design of our survey experiment and describes the data. Section 3 presents the findings from our comparisons of expectation levels and their correlations with investment choices. Section 4 concludes.

2 Design and Data

2.1 Design

For our main empirical analysis, we use a between-subject design in which we randomly assign individuals into two treatment groups and compare their responses. The experimental survey is

organized in six stages (see Table 1).

Table 1: Experimental Design

Treatments	Treatment A		Treatment T	
Stage 1	Portfolio choice			
Stage 2	Average method		Total method	
Stage 3	Uncertainty of average expectation	Year-by-year expectations	Uncertainty of total expectation	
Stage 4	Total method		Average method	
Stage 5	Uncertainty of total expectation		Uncertainty of average expectation	Year-by-year expectations
Stage 6	Financial literacy, risk attitudes, demographics, financial experience			

In the portfolio choice task of Stage 1, each participant invests \$50,000 over a period of 10 years. They can choose between two assets, "Stocks" and "Bond". The "Bond" is described as a US treasury bond that is safe and guarantees that the value of the investment is unchanged, net of inflation. The "Stocks" are described as a bundle of stocks with gains and losses that move one-to-one with the stock market index Standard and Poor's 500 (S&P 500), also net of inflation. The task is incentivized by paying their investment outcomes at a smaller scale. In particular, the outcome of the stock investment is based on a randomly drawn past return for a corresponding real-world investment.[5] This task is adapted from Breunig et al. (2021b) who show that such experimental investments correlate well with beliefs about the stock market return for a one-year investment, and also with real-life investments.[6]

Including the portfolio choice already in Stage 1, before any other data collection, ensures that the treatment variations in the subsequent expectation questions do not affect the investment choices. In Stage 2, we randomly split the sample and elicit expectations about the performance of the S&P 500 over the next 10 years. For Group A – the participants in Treatment A – we elicit long-run expectations using the *average method*, while for Group T, we use the *total method*. The wording for the two questions in the survey reads as follows (including the underlining):

Average method: What do you think is the percentage change that the S&P 500 makes

[5]The precise wording is: "To determine the gain or loss of your possible investment in "Stocks", we consider all gains and losses of the S&P 500 between 1958 and 2022 and randomly select a 10-year time period from this range."

[6]Relative to Breunig et al. (2021b), we made the following changes in experimental design and procedures: a US-based asset instead of a German-based one; a horizon of 10 years instead of 1 year; omission of exogenous shifts of asset returns; inclusion of a simple neutralization of all inflation effects; asking for belief uncertainty in a separate question, instead of using a histogram elicitation; implementation in an online survey instead of a personal interview or laboratory setting; use of USD instead of EUR; scaling down by a factor of 25,000 instead of 2,000.

per year, during the next 10 years on average? Please select a direction first, then enter a percentage number.

Total method: What do you think is the percentage change that the S&P 500 makes over the next 10 years in total? Please select a direction first, then enter a percentage number.

In Stage 3, we collect additional information, and with a further sample split for Group A. For half of the participants, we elicit the average expectations of each of the ten years separately. This allows us to measure whether they follow the geometric mean or the arithmetic mean when answering the expectation question of the average method (in cases where the two means differ). For the other half of the participants in Group A, we elicit probabilistic measures of their expectations about the average return per year of the S&P 500. Based on their probabilistic responses, we can construct an additional measure of average expectations and derive a measure of subjective uncertainty. Similarly, in Group T, we elicit probabilistic measures of their expectations about the total return (from all participants in the group).

In Stage 4, we swap the elicitation method between the two groups and collect their long-run expectations again: now with the total method for Group A and the average method for Group T. Accordingly, in Stage 5, we follow the strategy of Stage 3 to collect information about subjective uncertainty. Therefore, we obtain both measures of expectations (average and total) from every individual and the two groups differ only in the order of the elicitation methods. Since the answers for the second expectation question (Stage 4) might be influenced by the earlier questions (Stages 2-3), we pre-registered that for the main analysis, we discard the information from Stages 4 and 5. In a robustness check, we include all observations.

Finally, in Stage 6, we collect information on financial literacy, risk attitudes, demographic characteristics and personal experience with financial markets.

2.2 Data

We implemented the survey with Prolific, with a sample that is designed to be representative of the U.S. population by age, gender, and race. A total of 3,060 participants completed our questionnaire between March and April 2024. As pre-registered, we exclude individuals with extreme expectations at the top and the bottom one percent. This leaves us with 1,479 individuals in Group A and 1,459 individuals in Group T.

We perform a balancing test between the two groups using all of the observable characteristics of group members. The small differences that occur between the two groups are insignificant, except for a few variables. The joint null hypothesis of identical distributions cannot be rejected at the level of ten percent statistical significance (see Table A.1).

3 Results

3.1 The level of stated expectations

In Table 2, we compare the stated expectations about the stock market return between the two groups. To make the average and total expectations comparable, we convert the ten-year total expectations into geometrically annualized returns using Annualized $= \sqrt[10]{1 + \text{Total}} - 1$.

For the comparison of levels of expectations between the two methods, the table's first row shows a clear pattern: average expectations are far higher than the corresponding values from the total question. When asked about annual growth, the average answer is 5.315 percent (Column 1), whereas it is only 1.386 percent when asked about total returns (Column 2).[7] The difference (Column 3) amounts to 3.930 percentage points for the full sample of participants, with a p-value for rejecting equality that lies below 0.01 percent. For robustness, Column 4 shows the difference between annual and total expectations for the case that total expectations are annualized using the arithmetic mean. The conclusion remains the same (i.e., compounding of interest does not strongly affect the result): annual expectations are larger than annualized total expectations.

The same pattern holds across all subgroups that we pre-registered. There is, however, considerable heterogeneity in the levels of the expectations – men report higher expectations than women, financially literate people report higher expectations than less literate people, people with a graduate degree report higher expectations than those with less education, and people who have experience of equity investment report higher expectations than those without such experience. The corresponding differences between the annual and total expectations are fairly similar across subgroups, and they are strongly significant irrespective of the annualization approach (geometric or arithmetic).

Next, we compare the geometric and arithmetic means of year-by-year expectations with the reported average annual expectations. This confirms that the use of geometric versus arith-

[7] In Appendix Figure A.1 we examine the distribution of total and average annual expectations and confirm that this finding does not merely appear at the mean.

metic mean does not significantly affect the results. Two thirds of the respondents (964/1446) reported expectations that did not vary over time. Among the remaining respondents, while arithmetic means more frequently align exactly with the reported average annual expectations, both methods are equally accurate if we allow for rounding of responses.[8]

3.2 The relevance of stated expectations

Next, we turn to the paper's main question: Are annual or total expectation more predictive of investments? Under the standard view of the nature of expectations – the approach of de Finetti and Savage, stating that expectations are subjective probability distributions that serve as rationalizations of choice – the attempt to measure expectations is relevant only insofar as the measured expectations are related to behavior. We therefore quantify the predictive power of the two expectation measures for our portfolio choice. We standardize the total and average annual expectations (separately), to generate variables that have zero mean and unit standard deviation, and measure the predictive power in a simple regression analysis.

The estimates are shown in Table 3. Column (1) displays how the respective outcome is related to the standardized annual expectation, and analogously in Column (2) for the standardized total expectation. In Column (3), we pool the two groups and estimate the baseline and interaction effects of asking the annual question. Here, the differences in estimates between the two expectation measures are identified by the interaction terms.

In Panels I-III, the outcome variable is the proportion of "stocks" investment in the incentivized portfolio choice of Stage 1. For the main analysis (Panel I), we exploit the between-individual design and only use the first stated expectation, i.e., answers given in Stage 2 of the experiment. The results are very clear: Column 1 shows that a one-standard-deviation increase in the annual expectation is associated with an increase of 4.3 percentage points in the share of stocks in the portfolio. In contrast, a one-standard-deviation increase in the stated total expectation relates to a 7.5 percentage points higher investment share in stocks (Column 2). Both estimates are significant at conventional levels, and Column 3 shows the differences between these two estimates, which is also highly significant: the total method performs far better. An increase of one standard deviation in expectations elicited by the total method shifts investment behavior by 3.2 percentage points more than an increase of the same size in the annual expectations.

[8]More generally, we find that the arithmetic mean is closer to the reported annual expectation than the geometric mean (321 vs. 161) for the majority of respondents. However, this difference, too, is largely due to rounding errors. Once both means are rounded to match the level of precision used by each respondent in reporting their expectation (i.e., integer, one or two decimal places), they become identical for 86% of respondents (413/482).

Table 2: Comparison of annualized expectations

	(1) Average method		(2) Total method		(3) Diff	(4) Diff arithmetic
	Mean	Std. dev.	Mean	Std. dev.	Mean (S.E.)	Mean (S.E.)
Full sample	5.315	5.067	1.386	1.690	3.930*** (0.140)	3.681*** (0.146)
Male	5.614	5.027	1.793	2.009	3.821*** (0.202)	3.438*** (0.215)
Female	5.017	5.092	0.996	1.190	4.021*** (0.192)	3.902*** (0.194)
Age: 30 or below	5.284	5.610	1.252	1.592	4.032*** (0.310)	3.822*** (0.318)
Age: 31-44	4.913	5.065	1.466	1.986	3.446*** (0.287)	3.120*** (0.305)
Age: 45-59	5.526	4.912	1.483	1.659	4.043*** (0.270)	3.787*** (0.281)
Age: 60 or above	5.593	4.610	1.342	1.498	4.251*** (0.249)	4.045*** (0.258)
Financial literacy score < 5	4.905	5.863	0.981	1.156	3.924*** (0.216)	3.810*** (0.218)
Financial literacy $= 5$	5.808	3.850	1.834	2.040	3.974*** (0.166)	3.576*** (0.184)
Education: no college	4.850	5.091	1.091	1.476	3.759*** (0.254)	3.586*** (0.261)
Education: 2-year college	4.840	5.114	1.052	1.141	3.788*** (0.341)	3.669*** (0.346)
Education: 4-year college	5.493	4.773	1.543	1.809	3.950*** (0.223)	3.653*** (0.236)
Education: graduate degree	6.151	5.477	1.857	2.011	4.294*** (0.358)	3.902*** (0.377)
No equity inv. experience	4.511	5.724	0.865	1.087	3.645*** (0.248)	3.551*** (0.250)
Having equity inv. experience	5.781	4.584	1.703	1.901	4.078*** (0.164)	3.735*** (0.175)

Notes: Only responses to the first expectation question are included. In Column (2), expectations about total price change over the next ten years are annualized in an exponential way, i.e. an increase of 100% over the next ten years is converted into an average increase of 7.177% per year. Column (3) presents the difference between the average annual expectations in Column (1) and the annualized total expectations in Column (2). In Column (4), total expectations are annualized in a linear way, i.e. an increase of 100% over the next ten years is converted into an average increase of 10% per year, and then the difference from the average annual expectations in Column (1) are taken. The size of each subgroup can be inferred from Table A.1.

Table 3: Correlations between standardized expectations and investment

	Avearge method	Total method	Difference
	(1)	(2)	(3)
Outcome: Proportion of investment in stocks in the experiment			
Panel I: Using first expectation only			
Std. expectation	0.043***	0.075***	-0.032***
	(0.008)	(0.007)	(0.011)
Panel II: Using expectation inferred from reported tail probabilities			
Std. expectation	0.038***	0.071***	-0.033***
	(0.011)	(0.007)	(0.013)
Panel III: Using both reported expectations (pooled sample)			
Std. expectation	0.026***	0.078***	-0.052***
	(0.005)	(0.005)	(0.007)
Outcome: Past equity investment experience			
Panel IV: Using first expectation only			
Std. expectation	0.058***	0.111***	-0.053***
	(0.014)	(0.009)	(0.017)

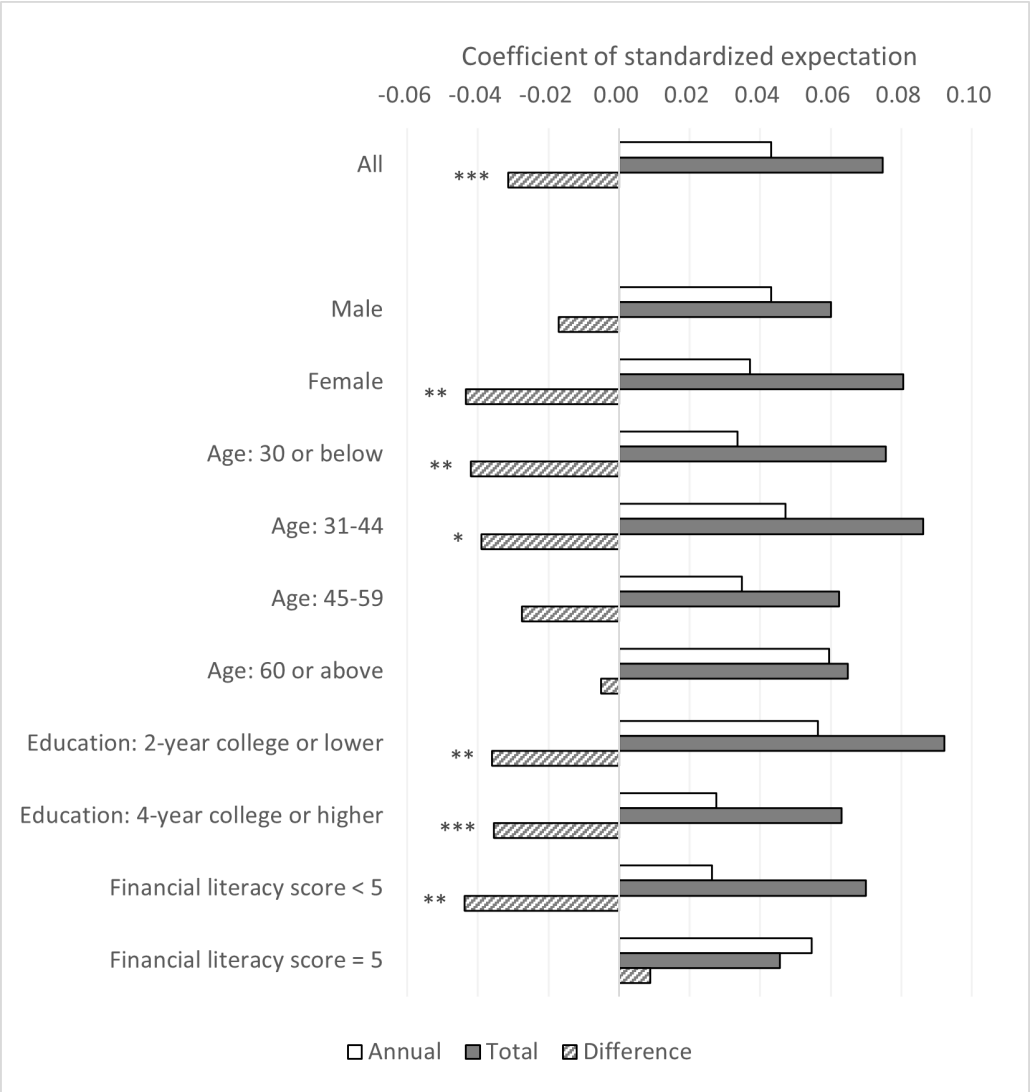
Notes: Linear regression of portfolio choice (0-1) in Panel I-III and equity investment experience in the past (0 or 1) in Panel IV on standardized expectation. Column 3 uses expectations from both methods and identifies the differences in coefficients by adding a dummy for the average method its interaction term with the standardized expectation. Panel I uses only the first expectation of each respondent. Panel II uses expectations inferred from reported tail probabilities. Panel III pools responses from Stage 2 and 4 of the experiment. Number of observations for Columns 1-3: (Panel I & IV) 1,479; 1,459; 2,938. (Panel II) 730; 1,410; 2,140. (Panel III) 2,938; 2,938; 5,876. All regressions include a constant term. Robust standard errors in parentheses. *** indicate significance at 1% level.

Panel II repeats the main analysis but uses inferred expectations rather than stated ones. We infer the expectation from the tail probabilities reported in Stage 3, assuming normality for each respondent's subjective distributions of returns. The conclusion that the stated total expectations correlates with investment more than the average annual expectations does not change, not even in terms of its quantity.

In Panel III, we present estimates for the pooled sample, disregarding whether respondents were asked the total or average annual question first. In spite of an order effect that we find on the total expectation,[9] the coefficient for the total expectation remains largely unchanged and the main result is robust to the inclusion of all data.

While the portfolio choices that our participants make as part of the survey experiment are incentive compatible, simple, and have been successfully implemented in previous studies,

[9]The stated total expectation is larger if it is elicited after asking about the annual expectation (see Table A.2).



Notes: */**/*** indicate significance difference at 10%, 5%, and 1% level.

Figure 1: Correlation between portfolio choice and standardized expectations

concerns about their external validity may remain. Especially, one may worry that the payments that we can offer in such a survey experiment are relatively small. Thus, Panel IV, reports the analogous results for an alternative outcome variable: a self-reported indicator of whether one has experience in equity investment.[10] The results support our main findings: responses under the total method have significantly more predictive value than responses under the annual method. Here, too, one-standard-deviation increase in the total expectation measure has a coefficient that is about twice as large as that of the annual expectation.

3.3 Heterogeneity in the relevance of stated expectations

As the last step of analysis, we examine heterogeneity in the relation between stated expectations and investments, expressed as sizes of the relevant regression coefficients. For comparison, the coefficients for the full population (Panel I of Table 3) are reported in the top of Figure 1. The bars represent the coefficients of the standardized annual expectation (hollow), the standardized total expectation (solid), and the difference between the effects of the two measures (dashed), as in Column (3) of Table 3. Stars at the end of the bars for the difference represent the significance level of the t-test for the interaction coefficient.

The following group-specific results stand out: First, for all subgroups, the coefficient of the expectation elicited with the total method is larger, except for individuals with the highest score in financial literacy (bottom end), for whom this difference matters little. Second, the subgroups for which the difference is not statistically significant at conventional levels include men, respondents above the age of 45, as well as the highly financially literate. This pattern confirms our pre-registered hypothesis that for individuals resembling the *homo economicus*, the elicitation method makes less of a difference. However, it is important to stress that these participants do not necessarily 'get it right': for example, people with the highest financial literacy score show some of the lowest estimates (for both questions) among all groups considered.

To better understand how the documented heterogeneity interacts with our main result, we include the relevant variables into our main analysis, in Appendix Table A.3. First, we add controls that are readily available in most surveys – age, education, and household composition. Adding these standard controls does not affect the results, but they do change when we add financial control variables, such as household net income, risk attitude, frequency of tracking S&P 500, and the financial literacy score: here, the correlations for both elicitation methods shrink. Their difference becomes smaller and insignificant once all financial variables are accounted for.[11] However, when repeating the analysis using the expectations from the pooled sample, which has the highest power and most precise estimates, the difference in coefficients remains negative and significant, with the total method again yielding a correlation twice the size of that from the average method.

[10] This part of the analysis was not pre-registered.

[11] We note that one may regard it as not surprising that the choice of a single-question elicitation method may become less important when numerous other financial variables, which are correlated with expectations, are also included in the regression. For the decision of what elicitation method to use in future work, it may or may not be relevant to control for such variables.

4 Conclusion

We aim to make a methodological contribution by comparing two ways of eliciting long-run expectations: asking about the total price change over a multiple-year period or the average change per year over the same period. All subgroups of the general population report expectations of stock market returns over the 10-year period that are pessimistic relative to the historical average, including those who get the full score on the financial literacy test and those with experience in equity investment. However, what is most important is the predictive power for economic choices (Brunnermeier et al., 2021). Contrary to the standard practice, the total method performs far better in predicting investment behaviors in the related domain than the average method, even though the obtained forecasts are less accurate than under the average method. Higher accuracy of expectations does not necessarily translate into economic relevance.

The two methods' predictive power is less different for subpopulations who are more *homo economicus*, suggesting that the advantage of the total expectation mainly comes from its informativeness about financial sophistication. But finding an appropriate way of controlling for sophistication is hard, and we conclude that in the absence of reliable measures for sophistication, it is better to ask the respondents for their total expectations, rather than average annual expectations.

References

Aiolfi, M., C. CapistrÃ¡n, and A. Timmermann (2010): "Forecast Combinations," CREATES Research Papers 2010-21, Department of Economics and Business Economics, Aarhus University.

Ameriks, J., G. Kézdi, M. Lee, and M. D. Shapiro (2020): "Heterogeneity in Expectations, Risk Tolerance, and Household Stock Shares: The Attenuation Puzzle," *Journal of Business & Economic Statistics*, 38, 633–646.

Amromin, G. and S. A. Sharpe (2014): "From the Horse's Mouth: Economic Conditions and Investor Expectations of Risk and Return," *Management Science*, 60, 845–866.

Ang, A., G. Bekaert, and M. Wei (2007): "Do macro variables, asset markets, or surveys forecast inflation better?" *Journal of Monetary Economics*, 54, 1163–1212.

Arrondel, L., H. Calvo-Pardoz, and D. Tasx (2014): "Subjective Return Expectations, Information and Stock Market Participation: Evidence from France," Discussion Papers in Economics and Econometrics 1415, University of Southampton.

Bachmann, R., G. Topa, and W. van der Klaauw (2023): *Handbook of Economic Expectations*, Academic Press.

Bailey, M., R. Cao, T. Kuchler, and J. Stroebel (2018): "The Economic Effects of Social Networks: Evidence from the Housing Market," *Journal of Political Economy*, 126, 2224–2276.

Berge, T. (2018): "Understanding survey-based inflation expectations," *International Journal of Forecasting*, 34, 788–801.

Breunig, C., I. Grabova, P. Haan, F. Weinhardt, and G. Weizsäcker (2021a): "Long-run expectations of households," *Journal of Behavioral and Experimental Finance*, 31, 100535.

Breunig, C., S. Huck, T. Schmidt, and G. Weizsäcker (2021b): "The Standard Portfolio Choice Problem in Germany," *The Economic Journal*, 131, 2413–2446.

Brunnermeier, M., E. Farhi, R. S. J. Koijen, A. Krishnamurthy, S. C. Ludvigson, H. Lustig, S. Nagel, and M. Piazzesi (2021): "Perspectives on the Future of Asset Pricing," *Review of Financial Studies*, 34, 2126–2160.

Clements, M. P. (2015): "Are Professional Macroeconomic Forecasters Able To Do Better Than Forecasting Trends?" *Journal of Money, Credit and Banking*, 47, 349–382.

Diercks, A. M. and H. Jendoubi (2023): "Expectations of Financial Market Participants," in *Handbook of Economic Expectations*, ed. by R. Bachmann, G. Topa, and W. van der Klaauw, Elsevier.

Dominitz, J. and C. Manski (2011): "Measuring and Interpreting Expectations of Equity Returns," *Journal of Applied Econometrics*, 26, 352–370.

Faust, J. and J. Wright (2009): "Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset," *Journal of Business & Economic Statistics*, 27, 468–479.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus (2021): "Five Facts about Beliefs and Portfolios," *American Economic Review*, 111, 1481–1522.

Grothe, M. and A. Meyler (2018): "Inflation Forecasts: Are Market-Based and Survey-Based Measures Informative?" *International Journal of Financial Research*, 9, 171–188.

Hurd, M., M. Van Rooij, and J. Winter (2011): "Stock market expectations of Dutch households," *Journal of Applied Econometrics*, 26, 416–436.

Knüpfer, S., E. Rantapuska, and M. Sarvimäki (2017): "Formative Experiences and Portfolio Choice: Evidence from the Finnish Great Depression," *The Journal of Finance*, 72, 133–166.

Knüppel, M. and A. L. Vladu (2016): "Approximating fixed-horizon forecasts using fixed-event forecasts," Discussion Papers 28/2016, Deutsche Bundesbank.

Laudenbach, C., A. Weber, R. Weber, and J. Wohlfart (2023): "Beliefs About the Stock Market and Investment Choices: Evidence from a Survey and a Field Experiment," Tech. rep., SSRN Paper Series.

Levy, M. and J. Tasoff (2016): "Exponential-growth bias and lifecycle consumption," *Journal of the European Economic Association*, 14, 545–583.

Meeuwis, M., J. A. Parker, A. Schoar, and D. Simester (2022): "Belief Disagreement and Portfolio Choice," *Journal of Finance*, 77, 3191–3247.

Merkle, C. and M. Weber (2014): "Do investors put their money where their mouth is? Stock market expectations and investing behavior," *Journal of Banking & Finance*, 46, 372–386.

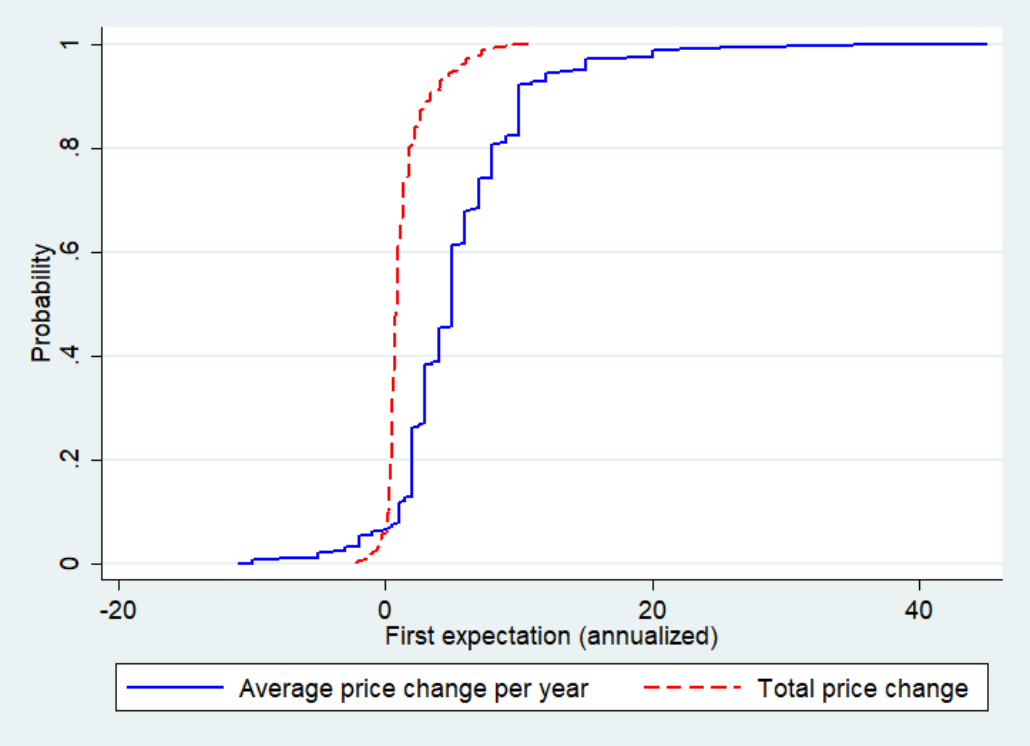
Shiller, R. J. and A. K. Thompson (2022): "What Have They Been Thinking? Home Buyer Behavior in Hot and Cold Markets: A Ten-Year Retrospect," *Brookings Papers on Economic Activity*, Spring, 307–344.

Snowberg, E., J. Wolfers, and E. Zitzewitz (2013): "Prediction Markets for Economic Forecasting," in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. Granger, and A. Timmermann, Elsevier, vol. 2 of *Handbook of Economic Forecasting*, chap. 0, 657–687.

Vissing-Jorgensen, A. (2004): "Perspectives on behavioral finance: does 'irrationality' disappear with wealth? Evidence from expectations and actions," in *NBER Macroeconomics Annual 2003*, ed. by M. Gertler and K. Rogoff, NBER.

Zimpelmann, C. (2021): "Stock Market Beliefs and Portfolio Choice in the General Population," Tech. Rep. 258, Discussion Paper Series - CRC TR 224.

Appendix



Notes: Expectations about total price change over the next ten years are annualized in an exponential way, i.e. an increase of 100% over the next ten years is converted into an average increase of 7.177% per year.

Figure A.1: Cumulative distribution of annual and total expectations

Table A.1: Balance tests of variables between treatment groups

	Treatment A		Treatment T		Diff	
	Mean	Std. dev.	Mean	Std. dev.	Mean (S.E.)	p-value
Share of investment in stocks	0.605	0.275	0.614	0.263	-0.010 (0.010)	0.3288
Having equity investment experience	0.634	-	0.621	-	0.013 (0.018)	0.4815
Female	0.500	-	0.511	-	-0.010 (0.018)	0.5773
Age: 31-44	0.274	-	0.241	-	0.033 (0.016)	0.0436
Age: 45-59	0.262	-	0.252	-	0.011 (0.016)	0.5032
Age: 60 or above	0.225	-	0.264	-	-0.039 (0.016)	0.0146
Education: 2-year college	0.147	-	0.162	-	0.016 (0.013)	0.2387
Education: 4-year college	0.373	-	0.356	-	0.017 (0.018)	0.3439
Education: graduate degree	0.174	-	0.183	-	0.009 (0.014)	0.5450
Household net income	70,865	42,080	72,084	44,286	-1,219 (1,612)	0.4497
Log-household net income	10.941	0.743	10.933	0.788	0.008 (0.029)	0.7707
Number of adults	2.086	0.970	2.066	0.923	0.020 (0.035)	0.5781
Number of children	0.542	0.951	0.523	0.942	0.018 (0.035)	0.6024
Willingness to take risk (0-10)	5.041	2.480	5.017	2.494	0.024 (0.092)	0.7928
Frequently tracking S&P 500	0.108	-	0.134	-	-0.026 (0.012)	0.0295
Sometimes tracking S&P 500	0.385	-	0.363	-	0.022 (0.018)	0.2150
Financial literacy score (0-5)	4.080	1.076	4.089	1.094	-0.009 (0.040)	0.8291

Notes: Number of observations: Treatment A: 1,448 for household net income, 1,449 for number of adults, 1,462 for number of children, and 1,479 for all other variables; Treatment T: 1,423 for household net income, 1,440 for number of adults, 1,450 for number of children, and 1,459 for all other variables. A joint test does not reject the null hypothesis of equality at the 10% level with $F(16, 2812) = 1.27$ and p-value $= 0.2089$.

Table A.2: Order effect

	First		Second		Diff	
	Obs.	Mean (S.E.)	Obs.	Mean (S.E.)	Mean (S.E.)	p-value
Annual expectation	1,479	5.315 (0.132)	1,459	5.483 (0.184)	-0.167 (0.226)	0.4592
Total expectation	1,459	16.342 (0.603)	1,479	20.933 (0.748)	-4.591 (0.962)	0.0000

Table A.3: Correlations between standardized expectations and investment, with controls

Method:	Average	Total	Difference	Average	Total	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Panel I: Using first expectation only						
Std. expectation	0.039***	0.062***	-0.023**	0.030***	0.038***	-0.008
	(0.008)	(0.007)	(0.011)	(0.007)	(0.007)	(0.010)
Controls A	✓	✓	✓	✓	✓	✓
Controls B				✓	✓	✓
Observations	1,446	1,439	2,885	1,422	1,407	2,829
Panel II: Using expectation inferred from reported tail probabilities						
Std. expectation	0.036***	0.058***	-0.023	0.025***	0.036***	-0.010
	(0.012)	(0.007)	(0.014)	(0.009)	(0.007)	(0.012)
Controls A	✓	✓	✓	✓	✓	✓
Controls B				✓	✓	✓
Observations	711	1,390	2,101	701	1,359	2,060
Panel III: Using both reported expectations (pooled sample)						
Std. expectation	0.025***	0.066***	-0.042***	0.021***	0.040***	-0.020***
	(0.005)	(0.005)	(0.007)	(0.005)	(0.005)	(0.006)
Controls A	✓	✓	✓	✓	✓	✓
Controls B				✓	✓	✓
Observations	2,885	2,885	5,770	2,829	2,829	5,658

Notes: Linear regression of portfolio choice (0-1) in stated expectation. Column 3 uses expectations from both methods and identifies the differences in coefficients by adding a dummy for the average method its interaction term with the standardized expectation. Panel I uses each respondent's first expectation only. Panel II uses expectations inferred from reported tail probabilities. Panel III pools responses from Stage 2 and Stage 4 of the experiment. Controls A consist of gender, age, education, and household composition. Controls B consist of household net income, risk attitude, frequency to track S&P 500, and financial literacy score. Interaction between the treatment dummy and all the controls are also included. All regressions include a constant term. Robust standard errors in parentheses. */**/*** indicate significance at 10%, 5%, and 1% level.