The effect of experts' and laypeople's forecasts on others' stock market forecasts

Christoph Huber^{*,†} Jürgen Huber^{*} Laura Hueber^{*}

This draft: August 19, 2019

Abstract

With a large-scale online experiment with 1593 participants from the U.S. and the U.K. we explore whether and how people working in the finance industry and laypeople from the general population are influenced by information on other people's forecasts when making forecasts on the future development of two indices and two stocks. We find that (i) laypeople's forecasts are strongly influenced by information they get on other subjects' forecasts, while financial professionals are much less influenced by information signals; (ii) signals by financial professionals influence all subject groups more than forecasts by laypeople; (iii) we observe a home bias in all subject groups, which can be mitigated by information signals; (iv) all subject groups expect lower forecast errors for financial professionals than for laypeople, hence we find evidence for trust in experts.

JEL: C91, D82, G10, G40

Keywords: stock market forecasts, expert advice, investment decision, home bias, overconfidence, social information

The accepted version of this manuscript is available here:

Huber, C., Huber, J., Hueber L. *The effect of experts' and laypeople's forecasts on others' stock market forecasts*. Journal of Banking & Finance. DOI: [10.1016/j.jbankfin.2019.105662](https://doi.org/10.1016/j.jbankfin.2019.105662).

^{*}University of Innsbruck, Department of Banking and Finance, Universitätsstrasse 15, 6020 Innsbruck, Austria.

[†]Corresponding author, E-mail: christoph.huber@uibk.ac.at.
Declarations of interest: none.

1 Introduction

While for some avid investors and financial professionals the stock market is like a natural habitat, for many laypeople it is a “scary” place, as their money is at stake, they lack control, and they think they know too little about the assets being traded. Many people thus prefer to rely on expert advice when investing in the stock market – either by consulting a financial advisor, or by handing their funds to a professional fund manager for investment.¹ However, clearly not all professionals will (can) always invest their clients’ funds successfully.²

Previous research on the role of financial advice and the importance of expert information provided to clients yields mixed results. For example, [Calcagno and Monticone \(2015\)](#) present survey evidence that a higher level of financial literacy among customers raises the probability of calling on expert advisors but lowers the probability of delegating investment decisions. In a similar direction, based on a large data set of a German brokerage firm [Bhattacharya et al. \(2012\)](#) find that people who are more financially literate are more likely to consult advice, but they do not necessarily act on the given advice. In this vein, investigations on the quality of available information are closely connected to the question whether people should act on the given information. Psychological research by [Önköl et al. \(2009\)](#) shows that the framing on the source of advice – calling it ‘human experts’ or a ‘statistical method’, in particular – matter for its influence on stock market predictions.³ However, the influence of information signals by actual people with different characteristics has not yet been investigated.

To provide clients with sufficiently informed financial advice, experts naturally have to form beliefs about future stock market developments. Thus, stock market forecasts play a central role. So far, several studies have observed overconfident forecasts (e.g. [Lichtenstein and Fischhoff, 1977](#)) and have looked into potential differences between laypeople and financial professionals, i.e. experts: [Törngren and Montgomery \(2004\)](#) find that professionals’ forecasts are generally not better than those of laypeople but both groups share the erroneous belief that financial professionals would make more accurate predictions. Similarly, [Menkhoff et al. \(2006\)](#) and [Deaves et al. \(2010\)](#) report miscalibration among financial professionals but find mixed evidence on the impact of market experience on the self-assessment of their estimation accuracy. [Ben-David et al. \(2013\)](#) provide additional evidence of overconfident professionals as they report considerable

¹As a consequence, the share of U.S. stocks held directly by households has steadily decreased over the past decades from over 90 percent in 1950 to 37 percent in 2017. This huge change can mostly be attributed to an increasing share of stocks held through mutual funds (24 percent of the total in 2017), pension funds (13 percent), ETFs (6 percent), and hedge funds (3 percent), all of which are managed by financial professionals. See: Federal Reserve Board and Goldman Sachs Global Investment Research, <https://www.cnbc.com/2017/06/26/goldman-markets-gains-etfs-mutual-funds.html>

²Among many others, [Jensen \(1968\)](#), [Gruber \(1996\)](#), and [Malkiel \(2003, 2004\)](#) have shown that over decades actively managed mutual funds have underperformed the market by between 1 and 2 percentage points per year, while [Lack \(2011\)](#) has provocatively stated, that hedge funds made their respective managers rich, but not the investors, as managers captured most of the profits through the fees they demanded.

³In experimental economics it is generally regarded as very important not to deceive subjects (e.g. [Friedman and Sunder, 1994](#)), hence we use data from real subjects to construct our information signals. In the study by [Önköl et al. \(2009\)](#), however, a ‘signal’ is framed to be from ‘human experts’, while it is actually the same as the one framed to be from a ‘statistical method.’

miscalibration among U.S. CFOs in conducting stock market predictions based on a 10-year panel. [Kilka and Weber \(2000\)](#) relate overconfidence among investors to preferences for domestic assets (“home bias”) as they observe subjects to report higher competence levels for return predictions of domestic stocks and to form more optimistic beliefs about future returns of stocks associated with higher self-assessed competence levels.

With this study we aim to directly relate financial advice to the level of and confidence in stock market forecasts by clients as well as other professionals. Thus, we examine experts’ and laypeople’s stock market forecasts and how their judgements are influenced by others’ opinions – i.e., whether they react similarly to information signals by laypeople or by experts. In a novel experiment with 617 and 607 subjects from the general population in the United States and the United Kingdom, respectively, as well as 269 and 100 subjects from the finance industry in these two countries, we explore how predictions of future stock market returns and their respective volatility are influenced by the presence or lack of information on other people’s estimates. In a between-subjects design with three treatments, participants receive either (i) no information on others’ predictions, (ii) the average prediction of a sample of the general population, or (iii) the average prediction of a sample from the finance industry. Each subject then makes stock market predictions on the one-month return of a U.S.-American stock index and a stock as well as on a British stock index and a stock. In addition, participants estimate both their own prediction errors as well as the prediction errors of each sub-group (laypeople and financial professionals).

This comprehensive experimental design allows us to contribute to the literature by analyzing how different groups of subjects are influenced by information signals from different sources: either from the general population or from financial experts. To the best of our knowledge, we are the first who examine the effect of information signals provided from pre-gathered data from the general population and financial professionals on forecasts. We are thus able to directly compare stock market forecasts across different treatments varying the presence and source of information in a rich, cross-country data set. We thus contribute to the ongoing discussion about the highly relevant expert-customer relationship in the financial advice industry. In addition, we investigate whether people’s stock market predictions are biased in the sense that people are overly optimistic regarding domestic assets and whether they are aware of their own and other people’s potentially biased beliefs.

Our key findings are that (i) laypeople’s forecasts are strongly influenced by information they get on other subjects’ forecasts, while financial professionals are much less influenced by such signals. (ii) Signals on forecasts by financial professionals influence all subject groups more than signals by laypeople. (iii) We observe a home bias in all subject groups, which can be mitigated by information signals. (iv) All subject groups expect lower forecasting errors for financial professionals than for other laypeople, hence we find evidence for trust in experts.

2 The experiment

2.1 Experimental tasks

The experiment consists of four main tasks. Tasks I and II are concerned with forecasting stock market returns; Task III asks for subject’s estimated accuracy of their own and others’ forecasts; and Task IV lets subjects allocate a given endowment into two different stock markets. In the following, we will describe each of the four tasks in turn. Instructions and screenshots of the tasks are provided in Appendices A and B, respectively.

In Task I (PREDICTION-task) subjects have to forecast the return of four distinct assets for the following 30 days. In particular, we ask participants to enter their point prediction as well as to estimate the respective 10% and 90% quantiles (Ben-David et al., 2013).⁴ We ask for forecasts for two stock indices as well as for two individual stocks from the same countries, the United States and the United Kingdom. The relevant stock indices are the S&P500 and the FTSE100. For the individual stocks we choose the largest oil and gas company in each of the two countries, ExxonMobil and BP, respectively. In order to increase the saliency of the assets’ nationality, we always add either ‘U.S.-American’ or ‘British’ before the respective asset’s name and depict the corresponding country flags behind. Additionally, subjects are presented with a table depicting the twelve previous monthly returns of the respective asset (Törngren and Montgomery, 2004). To keep the incentive structure simple, for each of the four estimates subjects earn 10 USD minus the number of percentage points they are off the actual return.

In Task II (CHOICE-task) we ask subjects to compare first the two stock indices’ return forecasts and then the two stock return forecasts. In particular, we ask which of the two stock indices or stocks, respectively, subjects expect to perform better in the following 30 days – that is, which of the two will yield a higher return. In addition, we ask participants to state the confidence that their forecast will be correct, i.e., that their chosen stock index or stock will actually perform better.

Task III (ACCURACY-task) then is concerned with how good subjects can assess their own and other’s forecasting errors. We ask subjects to assess their own (absolute) forecast errors and also the forecast errors of others regarding each of the two stock indices (in percentage points). Apart from their own forecasts subjects have to assess the forecasting accuracy of four distinct groups: the U.S.-American general population, U.S.-American financial professionals, the British general population, and British financial professionals. In this task, subjects earn 4 USD minus the number of percentage points they are off the actual forecasting error of the respective group for each of the ten forecast error estimates.

In Task IV (ALLOCATION-task) subjects have to invest a given endowment of 40 USD, which they should think of as if it were the equivalent of their annual income, in the stock market. We ask subjects to allocate this amount between two stock indices, the U.S.-American S&P500 and the British FTSE100, and enter the respective percentages of the endowment they want to invest in each of the two. A subject’s earnings in this task then is the value of the invested amount after 30 days.

⁴We ask, for example, “Over the next 30 days, for the S&P500, I expect” (i) “... the return to be ___” (point prediction), (ii) “... there is a 1-in-10 chance the actual return will be less than ___” (10% quantile), and (iii) “... there is a 1-in-10 chance the actual return will be more than ___” (90% quantile) (see Ben-David et al., 2013).

2.2 Treatments

We implement a $2 \times 2 \times 3$ between-subjects design in which we vary the *information*, i.e. the information about other participants' answers in each of the four tasks, for samples from two different countries – the United States (US) or the United Kingdom (UK) – and for either a representative sample of the general population or for a sample working in the finance industry (indicated by subscripts GP or FI, respectively).

In treatment I_{No} , participants have no information on how other people answered the respective questions and therefore have to rely solely on their own judgment and intuition. In treatment I_{GP} , we present subjects with the average answers from a subsample of the respective domestic general population sample next to each input field.^{5,6} Similarly, in treatment I_{FI} , we present subjects with average answers from the finance sample of their respective country and tell them about this subsample's composition (i.e., we mention examples of their professions and state truthfully that they have on average 10 years of experience in this industry; see Appendix A).

To be able to provide a comprehensive picture of how information about financial forecasts affect a representative sample and people from the finance industry, we run each treatment with each of the four different subject groups mentioned above. This way, we are able to distinguish the effect of 'peer' information from a representative sample as well as information from 'experts', i.e. a sample of finance professionals, within two different countries.

2.3 Implementation of the experiment

In total we recruited 617 and 607 subjects from general population samples in the United States and the United Kingdom, respectively, as well as 269 and 100 subjects from the finance industry in the two respective countries.⁷ Overall, 39.5% of subjects were female, with similar ratios across all groups; subjects' mean age was 46 years.

The experiment was conducted online using oTree by [Chen et al. \(2016\)](#). The whole experiment took place between Friday, June 29th and Wednesday, July 4th 2018. We made sure to start the experiment only after stock markets in New York had closed for the weekend (21:00 UTC) such that subjects from all four subject groups had the same – i.e., the most recent – return information. More than 90% of all data was collected while markets in the two relevant countries were closed (before Monday, 8:00 UTC).

⁵We first collected 25 observations from each of the four subject groups in treatment I_{No} and then presented subjects in the respective treatments I_{GP} and I_{FI} with the average answers from one of these subsamples. We deliberately refrained from pooling signals across subject groups and countries to avoid ambiguity about the source of information and presented only averages to keep the provided information simple and easy to understand.

⁶The screenshots of the decision tasks in Appendix B show how the information was presented next to each input field. The instructions contain a paragraph describing the subsample from which the average answers are shown as a group of 25 people from the respective country and it explicitly states that the answers were collected on the same or the previous day – after markets had closed for the weekend (see Appendix A).

⁷The samples are maintained by a large, globally operating market research company, which is widely used in economics research (e.g. [Almås et al., 2016](#); [Kirchler et al., 2018](#); [De Quidt et al., 2018](#); [Haaland and Roth, 2018](#); [Roth and Wohlfart, 2018](#)).

On average, the experiment lasted 16 minutes, which included reading on-screen instructions for each part of the experiment. After the main experimental tasks, we conducted a questionnaire on attitudes toward risk (from the German Socio-Economic Panel (SOEP), [Dohmen et al., 2011](#)), financial literacy ([Lusardi and Mitchell, 2014](#)), as well as several demographic questions. The experimental instructions for each treatment as well as screenshots of the decision screens are provided in Appendices A and B.

Subjects are incentivized in the following way. After the actual 30-day returns of the respective assets have been realized, one out of five subjects is randomly chosen to be paid her earnings from one randomly chosen task (Task I, III, or IV). Subjects from the U.S. samples can earn up to 40 USD in each of tasks I and III, and even more in task IV (depending on index performance); for the British samples the respective maximum in tasks I and III is 30 GBP. Average payments per chosen subjects were 24.87 USD in the U.S. and 19.27 GBP in the U.K.⁸

3 Results

3.1 How signals affect the *level* of predicted returns

To give a first overview of our data we start by presenting subjects' point predictions of expected returns over the next 30 days from the PREDICTION-task (Task I). We first discuss forecasts in treatment I_{NO} as they are independent of the information provided in treatments I_{GP} and I_{FI} and, thus, serve as a benchmark.

Table 1 shows average point predictions of each of our four sub-groups for each of the four assets in treatment I_{NO} .⁹ We see that average U.S. subjects' forecasts are higher than those of Britons for each asset. Additionally, in both countries financial professionals submit higher estimates than those of the respective general population in all but one of the eight pairwise comparisons (the exception is the estimate of the S&P500 by U.K. samples; though, none of the differences is statistically significant except for ExxonMobil for Britons, where $p < 0.05$) and U.S. financial professionals predict the highest average returns compared to all other subjects for each asset. Furthermore, all four population samples are more optimistic regarding the U.S. index compared to the British one, while they all expect a higher return for the British than for the American stock (all pairwise comparisons are significantly different at least at the 5% level, except for British financial professionals regarding the index return prediction and U.S. financial professionals regarding the stock return prediction). Most of the results discussed so far, however, are time-specific and unlikely to be generalizable. Presenting them has the main purpose of giving the reader a first impression of our results.

⁸At the time of the experiment, the exchange rate was approximately $\text{GBP } 1.00 = \text{USD } 1.32$.

⁹As our data contains several extreme outliers, most likely caused by typos, inattentiveness, or misunderstandings by some subjects we remove the highest and lowest 5% of subjects' point predictions and forecast errors for our further analyses unless stated otherwise. However, all our main results also hold when taking the full dataset into account. All data and replication files are available upon request.

Table 1: **Point predictions in treatment I_{NO} .** This table shows average point predictions in percentage points within each subject group and for each asset in treatment I_{NO} , i.e. independent from other people’s forecasts. Numbers in parentheses denote the respective standard deviations.

Asset	Subject group			
	FI_{US}	GP_{US}	FI_{UK}	GP_{UK}
S&P500	3.36 (3.97)	3.04 (3.49)	2.40 (3.35)	2.70 (3.77)
FTSE100	2.56 (2.73)	2.21 (2.36)	1.88 (1.91)	1.75 (1.75)
ExxonMobil	3.24 (2.71)	2.76 (2.32)	2.73 (2.02)	2.19 (1.77)
BP	4.05 (2.77)	3.69 (2.58)	3.58 (2.49)	2.97 (2.19)

Having presented a brief descriptive overview of the forecasts in treatment I_{NO} , we now turn to the question whether the provided signals on other people’s forecasts in treatments I_{GP} and I_{FI} influence these estimates. Figure 1 shows for each of the four assets (top two panels: indices; bottom two panels: stocks) and each of the four subject groups (people working in the U.S. finance industry – FI_{US} ; U.S. general population – GP_{US} ; U.K. finance industry – FI_{UK} ; U.K. general population – GP_{UK}) the one-month ahead return forecast. For each subject group we show three capped spikes, each representing one information treatment (no information I_{NO} ; information from the general population I_{GP} ; information from financial professional I_{FI}).¹⁰ The yellow circle within each spike represents subjects’ mean point prediction, while the lower and upper capped spikes represent the means of subjects’ self-reported 10% and 90% quantile for their return estimates. The line between the capped spikes thus shows the average size of subjects’ 80% confidence interval for their forecasts. The green diamonds represent the corresponding information signals that were provided to subjects before they entered their estimates.¹¹ The leftmost of each set of three capped spikes represents subjects’ predictions in treatment I_{NO} .¹²

First, we focus on the yellow circles representing the corresponding average point prediction for each spike. On the middle of each group of three spikes (representing treatment I_{GP}) we see that the provided signal is always comparatively high (between 6 and 10 percent expected return for the next month), which is welcome for our analysis, as we can then clearly see whether subjects’ forecasts are influenced by these signals. By visual inspection it seems that this is indeed the case: subjects take the provided information signal into account when making their forecasts as the average forecasts in treatment I_{GP} are always higher, by up to two percentage points, than in the baseline treatment I_{NO} . A second main observation is that subjects only partly follow the provided signals, as their forecasts are higher, but the (comparatively high)

¹⁰Note that there is only one capped spike for U.K. financial professionals as we only had access to 100 financial professionals in the U.K. and thus only conducted treatment I_{NO} with this sub-group.

¹¹I.e., in treatment I_{GP} this is the mean point prediction of the first 25 subjects from the general population and in I_{FI} the mean point prediction of the first 25 subjects from the finance industry in the respective country completing treatment I_{NO} of the experiment.

¹²Note that the respective means represented by yellow circles are equivalent to the means presented in Table 1.

Figure 1: **Overview of forecasts across treatments, assets, and subject groups.** This figure shows the average forecasts for each of the four stock indices and stocks, respectively, of each subject group and for each of the three treatments. Each panel shows forecasts for one asset and each group of three dots and three capped spikes represent the average point predictions (yellow circles), 10% and 90% quantiles (bottom and top caps), as well as the respective provided information signals (green diamonds) in percentage points for the three treatments within a group of subjects.

signal in treatment I_{GP} is always above the 90% quantile of forecasts – this holds for all twelve sub-groups in this treatment. As a third observation we see that also subjects from the finance industry tend to follow the signal from the general population, even though to a slightly smaller degree than subjects from the general population, as on all four pairwise comparisons the mean estimate of the general population is closer to the provided signal than the mean of people from the finance industry.

Turning to treatment I_{FI} , in which we provide forecasting signals from the finance industry (green diamonds in Figure 1), we see that the signals are on average markedly lower than the (relatively high) ones in treatment I_{GP} and usually very close to the average forecasts elicited in treatment I_{NO} . Thus, we cannot as

unequivocally judge whether subjects follow the signal and find only minor differences between the level of return predictions in treatments I_{NO} and I_{FI} .¹³

Table 2, showing differences in point predictions between treatments, corroborates these observations on treatment differences between I_{GP} and I_{FI} and the control treatment I_{NO} , respectively, using Wilcoxon rank-sum tests.

Table 2: Treatment differences in the level of predicted returns. The table shows differences in the level of predicted returns between the benchmark treatment, I_{NO} , and both information treatments, I_{GP} and I_{FI} . The numbers represent treatment differences of average point predictions for each asset and each subject group in percentage points. *, **, and *** represent the 5%, 1%, and 0.1% significance levels from Wilcoxon rank-sum tests.

Treatment	Asset	Difference in point predictions		
		FI_{US}	GP_{US}	GP_{UK}
I_{NO} vs. I_{GP}	S&P500	0.25	0.78***	0.94***
I_{NO} vs. I_{GP}	FTSE100	1.06	1.70***	1.53***
I_{NO} vs. I_{GP}	ExxonMobil	1.19	2.31***	1.22***
I_{NO} vs. I_{GP}	BP	1.62**	2.19***	1.39***
I_{NO} vs. I_{FI}	S&P500	-0.21	0.06***	-0.15***
I_{NO} vs. I_{FI}	FTSE100	0.06*	0.77***	0.41***
I_{NO} vs. I_{FI}	ExxonMobil	-0.57	0.07**	0.74***
I_{NO} vs. I_{FI}	BP	-0.60	-0.09	0.32**

While laypeople’s predictions are strongly influenced by the provided signals of other people’s forecasts ($p < 0.01$ in 15 out of 16 tests), subjects working in the finance industry react much less to these signals ($p < 0.01$ in just one out of eight tests). Hence, we conclude that while subjects from the general population in both the U.K and the U.S are strongly influenced by signals they receive on other people’s forecasts, U.S. financial professionals, by contrast, are usually not much influenced by this information. In addition, we examine whether subjects’ individual characteristics influence their forecasts’ distance from the provided information signals using ordinary least squares (OLS) regressions. We find that for general population samples in treatment I_{GP} a higher level of education and higher financial literacy are related to being further away from the signal – i.e., people rely more on their own intuition and judgement – but this does not hold for other subject groups or treatments.¹⁴

¹³Providing subjects with artificially high or low signals of our choice might have delivered clearer results, but it would have violated our pledge never to deceive subjects – as we stated that we would provide them with the average forecasts of 25 financial professionals, that is what we did.

¹⁴See Table C.4 in Appendix C. Analogously, tables C.5 and C.6 show OLS estimates relating subjects’ individual characteristics to the size of the self-reported confidence intervals (see section 3.2) and allocation decisions (see section 3.4), respectively. Yet, none of the individual explanatory variables shows a consistent and statistically significant pattern of influence regarding these two regressands.

The influence of information signals on subjects' forecasts can be further analyzed using the binary measure of predicted performance elicited in the CHOICE-task (Task II), where people have to forecast the better performing of two assets and how sure they are about this forecasts (certainty level 50-100%). Without information signals (treatment I_{NO} , top panel of Table 3), on average, 62.57% of subjects expect their domestic index/stock to perform better than the respective foreign index/stock. This holds for people from the finance industry as well as for laypeople and for the index as well as the stock. (The only exception is Britons where a slightly larger proportion of people expect the U.S. index to perform better than the FTSE100.)

In treatments I_{GP} and I_{FI} we find that these forecasts are influenced by the information signal subjects receive: the proportion of subjects who predict the better-performing asset to be in line with the provided signal increases by on average 5.75 percentage points compared to the respective proportions in treatment I_{NO} ($p < 0.01$, two-sided Wald test). In addition, information by financial professionals is given more weight than information from laypeople, as the shift is on average 8.30 percentage points in treatment I_{FI} , compared to 3.02 percentage points in treatment I_{GP} . Furthermore, we again find that forecasts by subjects from the general population sample are significantly influenced by the signals from financial professionals as the proportion of forecasts in line with the expert signal (treatment I_{FI}) increases by 12.32 percentage points ($p < 0.01$); forecasts in line with the signal from laypeople (treatment I_{GP}) only shift by less than a third of that amount, by 3.80 percentage points ($p = 0.1159$). Forecasts by subjects from the finance industry, however, are again not markedly influenced by the provided signals, neither from laypeople nor from experts: on average, the share of forecasts in line with the signal decreases slightly but statistically insignificantly by 3.86 percentage points ($p = 0.1899$). Table 3 thus shows that all subgroups are influenced by the signals they receive about other people's forecasts, as in all but two pairwise comparisons the independent judgements in treatment I_{NO} are farther from the signal provided than choices in treatment I_{GP} and I_{FI} .¹⁵ Again, this effect is larger for the expert signals in treatment I_{FI} and 3 out of 6 comparisons with I_{NO} are statistically significant at the 5% level, while this is the case for only 1 out of 6 comparisons between treatments I_{GP} and I_{NO} . Hence, subjects also tend to adjust their binary performance forecasts towards experts' forecasts but not or not as much, respectively, towards laypeople one's.

¹⁵In the remaining two comparisons, both for U.S. financial professionals the difference is marginal at 41.84% (I_{NO}) vs. 41.79% (I_{GP}) and 40.20% (I_{FI}), respectively.

Table 3: **Forecasts about the better-performing asset.** This table depicts the percentage amount of people in each subject group forecasting that the S&P500 and/or ExxonMobil will perform better over the next month and the corresponding information signals provided. The remainder then expects the FTSE100 and/or BP, respectively to outperform.

	Subject group			
	FI _{US}	GP _{US}	FI _{UK}	GP _{UK}
I _{NO}				
% amount choosing S&P500	80.61	78.33	52.00	53.23
% amount choosing ExxonMobil	58.16	68.97	38.00	42.29
I _{GP}				
Signal (better performing index)	FTSE100		S&P500	
% amount choosing S&P500	67.16	74.88	-	57.71
Signal (better performing stock)	BP		ExxonMobil	
% amount choosing ExxonMobil	58.21	60.87	-	41.79
I _{FI}				
Signal (better performing index)	FTSE100		S&P500	
% amount choosing S&P500	78.64	64.08	-	58.91
Signal (better performing stock)	BP		ExxonMobil	
% amount choosing ExxonMobil	59.80	51.71	-	54.46

3.2 How signals affect subjects' *confidence* in their forecasts

After having examined the influence of information signals on experts' and laypeople's forecasts on the *level* of predicted returns, we now turn to the question of how others' forecasts shape people's *confidence* in their predictions.

First, we look more closely at the subjective 80% confidence intervals elicited in the PREDICTION-task (Task I) and represented by capped spikes for each asset, each sub-group, and each information treatment (see Figure 1). Overall, these average reported confidence bounds are between 1.5 to 4.0 percentage points wide – as we will see below, this is too small with regard to historical data and points to overconfidence in the sense of miscalibration of subjects.¹⁶ As mentioned above, the middle spike in each set of three spikes shows the values for treatment I_{GP} and the rightmost spike gives results from treatment I_{FI}. We can now compare the size of the subjective confidence intervals between treatments to examine the information signals' influence on subjects' confidence in their forecasts.

For the general population samples, we see that the reported 80% confidence intervals are larger in the information treatment I_{GP} than in the benchmark treatment without signals, I_{NO} ($p < 0.01$). Hence, while the signal provided influences point predictions, they do not lead to increased confidence in subjects'

¹⁶Note that the term "overconfidence" is generally thought of as subsuming three distinct characteristics: overestimation (i.e., thinking that you are better than you are), overplacement (i.e., the exaggerated belief that you are better than others), and miscalibration (i.e., an excessive faith that you know the truth; see Moore and Schatz, 2017, for example, for a discussion of this terminology).

forecasts in the sense of narrower intervals. Rather, as the signal provided is comparatively high, many subjects may even have become more insecure and may thus report wider confidence bounds.

Comparing these estimates with those in treatment I_{FI} , one clear observation is that the self-reported confidence bounds are also not narrower (as would be the case if subjects clustered around the provided signal), but become mostly even larger with information signals by professionals (see Table 4).

Turning to subjects from the finance industry, the elicited confidence bounds are mostly very similar between treatment I_{NO} and either of the information treatments, while for the general population samples the size of the reported confidence intervals is significantly larger compared to either of the treatments in which we provide information signals ($p < 0.01$ in 14 out of 16 comparisons, see Table 4). Hence, we observe that while the general population samples' calibration in both the U.K and the U.S is strongly influenced by signals they receive on other people's estimates, U.S. financial professionals, by contrast, are hardly influenced by these signals.

Table 4: **Treatment differences in the size of self-reported confidence intervals.** The table shows differences in the sizes of self-reported confidence intervals between the benchmark treatment, I_{NO} , and both information treatments, I_{GP} and I_{FI} . The numbers represent treatment differences of average self-reported sizes of confidence intervals for each asset and each subject group in percentage points. *, **, and *** represent the 5%, 1%, and 0.1% significance levels from Wilcoxon rank-sum tests.

Treatment	Asset	Difference in the size of confidence intervals		
		FI_{US}	GP_{US}	GP_{UK}
I_{NO} vs. I_{GP}	S&P500	0.09	1.10***	0.39**
I_{NO} vs. I_{GP}	FTSE100	0.25	1.51***	0.73***
I_{NO} vs. I_{GP}	ExxonMobil	0.29	1.03***	0.63***
I_{NO} vs. I_{GP}	BP	0.25	0.97***	1.76***
I_{NO} vs. I_{FI}	S&P500	0.86**	1.21***	1.07***
I_{NO} vs. I_{FI}	FTSE100	0.86**	0.67***	1.56***
I_{NO} vs. I_{FI}	ExxonMobil	0.44	0.24***	1.56***
I_{NO} vs. I_{FI}	BP	0.11	0.17*	2.22***

To further analyze how well subjects' forecasts are calibrated with respect to historical stock market fluctuations, we follow Ben-David et al. (2013) in comparing the individual imputed volatility of subjects' self-reported 10% and 90% quantiles of expected returns with the distributional characteristics of historical returns of the same asset (also see Davidson et al., 1976).¹⁷ Thus, we transform the individual estimates of

¹⁷For the historical return distribution presented in Figure 2 we use data from September 1987 to September 2018, i.e. from the latest initiation among our four assets (BP) until the beginning of the experiment. In Figure C.1 in Appendix C we also use shorter windows (2008 to 2018, 2013 to 2018 and the financial crisis

the upper and lower bounds of the self-reported 80% confidence intervals into volatility estimates resulting in the imputed individual volatility as $\hat{\sigma}_i = \frac{x(0.90) - x(0.10)}{z}$. $x(0.10)$ and $x(0.90)$ denote the 10% and 90% quantile of the subjects' individual estimates regarding the certainty of their return predictions and z represents the number of standard deviations within the confidence interval. We apply $z = 2.65$ as the appropriately approximated value for a confidence interval of 80% (Davidson et al., 1976; Keefer and Bodily, 1983).

Figure 2 depicts the distribution of historical volatilities (yellow line) and the distribution of the imputed volatilities of our subjects (blue line) for each index (top two panels) and each stock (bottom two panels). In each panel the distribution for the individual imputed volatility is centered markedly more to the left than the distribution of historical returns. Specifically, the individual imputed volatility peaks from 0 to 2 (percent monthly volatility) for each asset, while the historical return distribution is centered at 5 in each case except for the S&P500 where it is centered at 4. Wilcoxon signed-rank tests confirm this observation that the individual imputed volatility is significantly different from the historical distribution ($p < 0.001$ for each asset). Hence, subjects are miscalibrated regarding their forecasting accuracy as they clearly underestimate the volatility of the return predictions compared to historical returns.

To corroborate that subjects are overconfident, rather than just miscalibrated, we calculate in which fraction of all observations the actually realized return is inside of the 80% confidence bound submitted by our subjects. While this number should be close to 80% if subjects were well-calibrated, the actual numbers are between 58.82% and 0.87% with an overall average of 19.33%, and thus only a quarter of what it should be, as is shown in Table 5. We also note that experts are only very marginally better calibrated than laypeople, with on average 19.90% vs. 18.77% of estimates inside the confidence bounds, respectively, for the two groups.

Table 5: **Percentage of correctly calibrated forecasts.** This table shows the percentage of forecasts for which the actual realized returns are inside the self-reported 80% confidence interval.

Asset	Subject group			
	FI _{US}	GP _{US}	FI _{UK}	GP _{UK}
S&P500	32.13	35.01	17.65	32.52
FTSE100	38.50	26.23	58.82	50.76
ExxonMobil	1.75	1.50	4.75	1.50
BP	0.87	1.51	4.71	1.11

Similarly, in the CHOICE-task (Task II), subjects assess the probability that their choice of the better-performing asset turns out to be correct between 59% and 62% for the stock and the index, respectively.

period 2007-2009) which represent historical periods of comparatively high and low volatility, respectively; results remain unchanged.

Figure 2: **Volatility of returns.** The graph shows the kernel density of the individual imputed volatility (blue line) and the historical return volatility (yellow line) for each stock and index. The individual imputed volatility was calculated from the experimental data of the return predictions. For the historical volatility, we include data from September 1987 until September 2018 and compute one-year rolling windows of monthly returns.

3.3 Beliefs about experts' and laypeople's forecasting accuracy

While we have seen above that the forecasting accuracy of all sub-groups was low compared to subjects' self-reported confidence, with less than 20% of realized returns lying inside of the subjective 80% confidence bounds, and with subjects from the financial sector being only marginally more accurate than laypeople, we now explore the beliefs members of each group hold about the forecasting accuracy of others. This allows us to assess subjects' individual confidence concerning other people's forecasts and whether and how much they trust in "experts" when it comes to stock market forecasts.

Figure 3 depicts for each subject group and for each asset the difference between estimated forecasting errors and experts' average own estimated forecasting error in the respective country.¹⁸

¹⁸That is, we subtract the average own estimated forecast error of FI_{US} from the estimated errors of FI_{US} and GP_{US} and the estimated average own forecast error of FI_{UK} from the estimated errors of FI_{UK} and GP_{UK} , respectively, to display differences to experts' own estimated forecasting accuracy.

Figure 3: **Estimated forecast errors.** This table shows the differences between subjects' estimated forecast errors for each subject group/asset and the average error subjects from the finance industry in the respective country estimate for their own forecasts. Forecast errors are standardized by subtracting the estimated average own forecast error of FI_{US} from the estimated errors for FI_{US} and GP_{US} and the estimated average own forecasting error of FI_{UK} from the estimated errors for FI_{UK} and GP_{UK} , respectively. In particular, the estimated average own forecasting error regarding the S&P500 is taken for U.S. subjects and the estimated average own forecasting error regarding the FTSE100 for U.K. subjects. The blue bars represent the errors estimated for people working in the finance industry and the yellow bars show forecasting errors estimated for the general population. The labeling on the horizontal axis depicts the subject group for whom the respective error was estimated. For example, the second bar in the first panel is the average forecasting error estimated by FI_{US} for GP_{US} regarding the S&P500. *, **, and *** represent the 5%, 1%, and 0.1% significance levels from Wilcoxon rank-sum tests.

For each asset and for each subject group, the estimated forecasting errors for the general population (yellow bars) are consistently higher than the estimated forecasting errors for financial professionals (blue bars) as the blue bar is lower than the yellow one in each panel. This difference is statistically significant at the 5% level in 13 out of 16 comparisons. As a key observation, we thus report that all subject groups expect more accurate forecasts by financial professionals than by people from the general population. This confidence, however, is unjustified in our study, as we find no significant difference between the actual forecasting errors of financial professionals and laypeople. Hence, we can draw two conclusions: First, financial professionals are too optimistic about their own forecasting accuracy compared to the predictions of the general population. And second, the latter are also convinced of the superior performance of financial experts compared to their own forecasting abilities. This result is consistent with – and might be a reason for – our findings above that especially laypeople act on the information signals provided by financial professionals, supposedly as they trust these signals more than their own judgement.

3.4 Preferences for domestic assets

Both the CHOICE-task (Task II) and the ALLOCATION-task (Task IV) allow us to analyze whether people’s judgements and preferences are biased towards domestic assets – i.e., whether there is evidence of a “home bias” (French and Poterba, 1991; Tesar and Werner, 1995).

Table 6 shows the average percentages subjects in each subject group invest in the S&P500; the remainder is then the proportion allocated to the FTSE100. In treatment I_{No} (top panel) we do observe for U.S.-Americans that both laypeople and participants from the finance industry tend to allocate more of their wealth into their domestic stock index (S&P500, $p < 0.001$). This tendency is stronger for laypeople than for subjects from the finance industry, with 4 to 5 percentage points more invested into the American index ($p = 0.098$). For subjects from the U.K. we observe a preference to invest in the British index (FTSE100), but the allocation is much more equally split as on average 46.73% of all U.K. subjects from the finance industry and 48.72% of all U.K. laypeople allocate their wealth into the foreign S&P500 in treatment I_{No} ($p < 0.01$ for a Wilcoxon signed-rank test against an allocation of 50%).¹⁹

¹⁹We want to tread carefully in interpreting our results as evidence for a “home bias” as it might be perfectly reasonable to hold a higher share of one’s wealth in the more diversified U.S.-American stock market.

Table 6: **Allocations to domestic and foreign assets.** This table shows the average percentage allocations for each subject group and for each treatment to the U.S.-American S&P500 stock index and the corresponding information signals provided. The remaining share of subjects' wealth is then allocated to the British FTSE100.

	Subject group			
	FI _{US}	GP _{US}	FI _{UK}	GP _{UK}
I _{NO}				
Allocation S&P500 in %	57.78	62.15	46.73	48.72
I _{GP}				
Signal: allocation S&P500 in %	64.00		48.84	
Allocation S&P500 in %	57.93	62.65	-	48.50
I _{FI}				
Signal: allocation S&P500 in %	58.16		54.36	
Allocation S&P500 in %	58.01	59.36	-	53.26

We now again consider how these results are influenced by signals on other people's allocations. The signal provided in treatment I_{GP} (middle panel of Table 6) is very close to the allocation subjects from the general population also choose in treatment I_{NO}, hence it is no surprise that the allocations of the general population samples in treatment I_{GP} are almost identical to the ones in treatment I_{NO}. However, the one observation to be made is that subjects from the finance industry are not influenced at all by this signal, as their allocations in treatment I_{GP} are almost identical to the ones in I_{NO}, while the provided signal is more than 6 percentage points higher than their average allocation to the S&P500 in treatment I_{NO} ($p = 0.95$).

When subjects receive the average allocation of 25 subjects from the finance industry in their country as the information signal (treatment I_{FI}; see bottom panel of Table 6), this signal clearly influences their allocation choices, as all allocations are shifted into the direction of the signal, usually very close to the allocation from the signal. While the signal reduces the difference in the allocation between I_{NO} and the signal by 70 to 80 percent and, hence, seems to strongly influence subjects choices, this treatment difference is only statistically significant for GP_{UK} ($p < 0.01$).²⁰ To conclude: In the ALLOCATION-task changes in allocation were always in the direction of the signal provided, especially when the signal came from financial professionals, and the results are thus consistent with what we reported for Task II, in which 62.57 percent of subjects expected their domestic index/stock to perform better than the respective foreign index/stock.

4 Discussion and conclusion

To examine the forecasting behavior and the impact of information signals on subjects' forecasts we conducted a large-scale online experiment, recruiting 1593 U.S. and U.K. subjects from the general population and the finance industry. We collected forecasts including the 10% and 90% quantiles of the corresponding

²⁰E.g. for the U.S. general population the difference is reduced from 3.99 ($= 62.15 - 58.16$) to 1.20 ($= 59.36 - 58.16$) percentage points and analogously for the British general population.

estimates and subjects' self-reported confidence in their estimates. Additionally, we were able to conduct a cross-country comparison between the U.S. and U.K. estimates, identifying potential preferences for domestic assets ("home bias"). In a baseline treatment subjects entered their judgemental forecasts independent of others' estimations, while in two additional treatments we provided pre-collected averages of others' estimates – either from a general population sample or from financial professionals – as information signals. This allowed us to analyze the influence of others' forecasts on subjects' own forecasting behavior in connection with their self-reported confidence about their own and other subject groups' forecasting accuracies. We thereby connected two important topics which were so far studied in isolation: whether subjects' forecasts are biased when they predict future asset returns, and whether their forecasts are influenced by information signals on others' forecasts, namely by either financial experts or by a general population sample.

Overall, we found that return forecasts by laypeople were substantially influenced by information about others' forecasts. On average, we observe subjects return forecasts to move in the direction of the signal and we found substantial differences between return forecasts provided in either information treatments compared to the no-information treatment. Financial professionals, in contrast, were much less influenced by the provided information signals. Our results are therefore in line with previous findings – albeit in a different setting – by [Önkal et al. \(2009\)](#). Furthermore we found that all subject groups were more influenced by signals from financial experts than by laypeople.

Regarding people's forecasting accuracy, we reported miscalibration and even overconfidence among all subject groups. We compared the distribution of the individual imputed volatility based on subjects' self-reported 10% and 90% quantile and the distribution of the historical volatility. Subjects imputed estimated volatility was significantly lower than the realized historical volatility for each asset and subject group. This result is similar to [Ben-David et al. \(2013\)](#), who examined managerial forecasting behavior. In addition, we find that more than 80% of all confidence intervals the subjects in the present study stated were too narrow and the realized return was located outside the estimated interval. With these results, we add to existing literature documenting miscalibration and overconfidence among both financial professionals and laypeople (e.g. [Menkhoff et al., 2006](#); [Deaves et al., 2010](#); [Kaustia and Perttula, 2012](#)).

Forecasts on whether an American or British index would perform better and which share of a given amount of money subjects would invest in each of the two markets revealed a pronounced "home bias", i.e. each subject group tended to invest more in their home country than the respective foreign subject group would. This effect was stronger for laypeople than people from the finance industry, and it could be reduced by the information signals subjects received on others' decisions.

Our results further suggest that people expect lower forecasting errors for financial professionals compared to the general population. In particular, people overestimated the forecasting performance of financial professionals, as their forecasting error was not significantly different from that of the general population. This is in line with earlier papers who report higher overconfidence of financial professionals than laypeople (e.g. [Törngren and Montgomery, 2004](#); [Glaser et al., 2007](#); [Zaleskiewicz, 2011](#)). Similar to previous research, we were also able to show that financial professionals were not substantially better in forecasting future asset returns in comparison to laypeople (e.g. [Törngren and Montgomery, 2004](#); [Zaleskiewicz, 2011](#)). The observation that financial professionals were considered as more competent in forecasting asset returns

most likely explains why information signals from financial professionals were more strongly taken into account than signals from the general population. However, financial professionals were also not immune to being influenced by the provided information signals.

With this high-powered study with both financial professionals and laypeople from two different countries we are able to provide a comprehensive analysis of forecasting behavior, miscalibration, and asset allocation. With a novel design, introducing information signals about others' forecasts, in particular, this study is the first to examine forecasting behavior in the presence of peer and expert information, which in turn might be important when considering the client-expert relationship as it is the case in the financial advice industry. As laypeople's estimates and decisions are influenced by financial professionals' estimates and decisions – given that this information is available –, we find that people do trust experts in their judgements regarding stock market forecasts and follow their implicit recommendations at least partially. Our results thereby also relate to the growing literature on social finance and peer effects, in particular, which shows that people's investment decisions tend to be related to those of their coworkers (Hvide and Östberg, 2015), neighbors (Pool et al., 2015), and housemates (Zhang et al., 2018).

While expert advice has long been readily available to consumers e.g. by financial data providers and financial media, with the emergence of social networks and social trading platforms, in particular, *social* information in the sense of information about others' (other laypeople's) judgements and behavior has also become an increasingly important aspect in financial decision-making. In this regard, our results demonstrate that subjects' forecasts are indeed influenced by other laypeople's estimates – however, we find that people give more weight still to expert information: professionals' judgements have a larger influence on subjects forecasts and are expected to be more accurate than laypeople one's.

Acknowledgements

We thank Jun Honda and Michael Kirchler, as well as workshop participants of the Austrian Working Group on Banking and Finance 2018 in Salzburg and conference participants at Experimental Finance 2019 in Copenhagen for helpful comments and suggestions. Financial support from the Austrian Science Fund FWF (P29362-G27 and SFB F63) is gratefully acknowledged. This study was ethically approved by the IRB of the University of Innsbruck.

References

- Almås, Ingvild, Alexander W Cappelen, Bertil Tungodden. 2016. Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than scandinavians? Working Paper.
- Ben-David, Itzhak, John R Graham, Campbell R Harvey. 2013. Managerial miscalibration. *The Quarterly Journal of Economics* **128**(4) 1547–1584.
- Bhattacharya, Utpal, Andreas Hackethal, Simon Kaesler, Benjamin Loos, Steffen Meyer. 2012. Is unbiased financial advice to retail investors sufficient? Answers from a large field study. *The Review of Financial Studies* **25**(4) 975–1032.
- Calcagno, Riccardo, Chiara Monticone. 2015. Financial literacy and the demand for financial advice. *Journal of Banking & Finance* **50** 363–380.
- Chen, Daniel L, Martin Schonger, Chris Wickens. 2016. oTree – An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* **9** 88–97.
- Davidson, LB, DO Cooper, et al. 1976. A simple way of developing a probability distribution of present value. *Journal of Petroleum Technology* **28**(09) 1–069.
- De Quidt, Jonathan, Johannes Haushofer, Christopher Roth. 2018. Measuring and bounding experimenter demand. *American Economic Review* **108**(11) 3266–3302.
- Deaves, Richard, Erik Lüders, Michael Schröder. 2010. The dynamics of overconfidence: Evidence from stock market forecasters. *Journal of Economic Behavior & Organization* **75**(3) 402–412.
- Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, Gert G Wagner. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* **9**(3) 522–550.
- French, Kenneth R, James Michael Poterba. 1991. Investor diversification and international equity markets. *American Economic Review* **81**(2) 222–226.
- Friedman, Daniel, Shyam Sunder. 1994. *Experimental methods: A primer for economists*. Cambridge University Press.
- Glaser, Markus, Thomas Langer, Martin Weber. 2007. On the trend recognition and forecasting ability of professional traders. *Decision Analysis* **4**(4) 176–193.
- Gruber, Martin J. 1996. Another puzzle: The growth in actively managed mutual funds. *The Journal of Finance* **51**(3) 783–810.
- Haaland, Ingar, Christopher Roth. 2018. Labor market concerns and support for immigration. Working Paper.

- Hvide, Hans K, Per Östberg. 2015. Social interaction at work. *Journal of Financial Economics* **117**(3) 628–652.
- Jensen, Michael C. 1968. The performance of mutual funds in the period 1945–1964. *The Journal of Finance* **23**(2) 389–416.
- Kaustia, Markku, Milla Perttula. 2012. Overconfidence and debiasing in the financial industry. *Review of Behavioural Finance* **4**(1) 46–62.
- Keefer, Donald L, Samuel E Bodily. 1983. Three-point approximations for continuous random variables. *Management Science* **29**(5) 595–609.
- Kilka, Michael, Martin Weber. 2000. Home bias in international stock return expectations. *The Journal of Psychology and Financial Markets* **1**(3-4) 176–192.
- Kirchler, Michael, Florian Lindner, Utz Weitzel. 2018. Delegated decision making and social competition in the finance industry. Working Paper.
- Lack, Simon A. 2011. *The Hedge Fund Mirage: The Illusion of Big Money and Why It's Too Good to Be True*. Wiley.
- Lichtenstein, Sarah, Baruch Fischhoff. 1977. Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance* **20**(2) 159–183.
- Lusardi, Annamaria, Olivia S Mitchell. 2014. The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature* **52**(1) 5–44.
- Malkiel, Burton G. 2003. The efficient market hypothesis and its critics. *Journal of Economic Perspectives* **17** 59–83.
- Malkiel, Burton G. 2004. Models of stock market predictability. *The Journal of Financial Research* **27** 449–459.
- Menkhoff, Lukas, Ulrich Schmidt, Torsten Brozynski. 2006. The impact of experience on risk taking, overconfidence, and herding of fund managers: Complementary survey evidence. *European Economic Review* **50**(7) 1753–1766.
- Moore, Don A., Derek Schatz. 2017. The three faces of overconfidence. *Social and Personality Psychology Compass* **11**(8) e12331.
- Önköl, Dilek, Paul Goodwin, Mary Thomson, Sinan Gönöl, Andrew Pollock. 2009. The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making* **22**(4) 390–409.
- Pool, Veronika K, Noah Stoffman, Scott E Yonker. 2015. The people in your neighborhood: Social interactions and mutual fund portfolios. *The Journal of Finance* **70**(6) 2679–2732.

- Roth, Christopher, Johannes Wohlfart. 2018. How do expectations about the macroeconomy affect personal expectations and behavior? Working Paper.
- Tesar, Linda L, Ingrid M Werner. 1995. Home bias and high turnover. *Journal of International Money and Finance* **14**(4) 467–492.
- Törngren, Gustaf, Henry Montgomery. 2004. Worse Than Chance? Performance and Confidence Among Professionals and Laypeople in the Stock Market. *The Journal of Behavioral Finance* **5**(3) 148–153.
- Zaleskiewicz, Tomasz. 2011. Financial forecasts during the crisis: Were experts more accurate than laypeople? *Journal of Economic Psychology* **32**(3) 384–390.
- Zhang, Annie C, Jiali Fang, Ben Jacobsen, Ben R Marshall. 2018. Peer effects, personal characteristics and asset allocation. *Journal of Banking & Finance* **90** 76–95.

Appendix to

‘The effect of experts’ and laypeople’s forecasts on others’ stock market forecasts’

A Experimental instructions

The following instructions are from the experiment with the U.S.-American general population sample. Text which only applies for treatment I_{GP} is written in *italics*, text which only applies for treatment I_{FI} is written in **teletype**.²¹

General Instructions

Welcome to the Experiment!

Dear participant,

Thank you very much for accepting our invitation to take part in this experiment. We are researchers from [blinded] conducting a short study which is intended to take about 20 minutes. With your participation, you will make an important contribution to research and you can earn money throughout the experiment.

All data will be anonymous and no individual results will be disclosed publicly or to other participants of the experiment. The data will only be used for scientific purposes. This online study adheres to the principles of economic experiments: participants are not deceived and earnings are paid out in real. We guarantee at each stage of the data analyses that we will not trace back experimental decisions to participants’ identities.

Note that you will not be able to go back to previous pages throughout the whole study.

Thank you very much for participating!

————— (new page) —————

This experiment consists of **4 tasks**.

²¹Note that instructions for samples from the United States and Britain were identical except for the currency. Original instructions of each treatment are available upon request.

One month after the experiment, one in five participants will be randomly drawn for an additional payment. For the additional payment one task out of Tasks 1, 3, and 4 will be randomly drawn for payment. This task is then relevant for your payment and you will receive your earnings from this task if you are drawn for the additional payment. You receive this amount in points which you can exchange for cash via PayPal or other methods.

————— (new page) —————

In this experiment we ask you to **estimate future price changes** of U.S.-American and British stock indices and stocks and to assess your estimation error and the estimation errors of others concerning each forecast.

In each task you see the average answers from a sample of the general population in the United States. These answers come from a group of 25 people from the United States. These answers were collected on June 29th and June 30th, 2018.

In each task you see the average answers from financial professionals in the United States. These answers come from 25 financial professionals from the United States from the fields of trading, fund management, private equity management, etc. who have on average worked 10 years in this industry. These answers were collected on June 29th and June 30th, 2018.

Task 1

In Task 1 we ask you to estimate future price changes of U.S.-American and British stock indices and stocks.

In particular, you have to enter the **return you expect for the next 30 days** (the percentage price change in the month July). In addition, we ask you to predict the range of potential returns such that you estimate a 1-in-10 chance that the actual return will be less than the lower bound and such that you estimate a 1-in-10 chance that the actual return will be more than the upper bound.

In total we will ask you for four estimates. For each stock and for each stock index, we provide the past twelve monthly returns.

Payment: For each of the four estimates you have the possibility to earn up to \$10.0 depending

on how good your estimate is. For each estimate you receive \$10.0 minus 1.0 times the number of percentage points you are off the actual return after 30 days. For example: If you estimate a return of 6% and the actual return is 9%, then you earn $\$10.0 - 1.0 \times 3 = \7.0 .

Your total earnings from this task are the sum of your earnings from each of the four estimates. Hence, you can earn up to \$40.

Task 2

In Task 2 you have to estimate **which of two stock indices or stocks, respectively, performs better over the next 30 days**.

In addition, we ask you to enter the probability with which you think the stock index or stock you are choosing actually performs better in the next 30 days.

Note that this probability should be at least 50%. Otherwise, you expect the other stock index (or stock) to perform better and should choose the other one.

Task 3

In Task 3 you have to assess your own **forecasting accuracy**, the forecasting accuracy of people from the U.S.-American and British general population, and the forecasting accuracy of U.S.-American and British financial professionals.

On each page of this task we ask you to enter your prediction of the respective **forecasting error** (in percentage points) - that is, the average of the absolute differences between the respective forecast and the actual return after 30 days. These values should represent how far off you think your own and the other groups' forecasts are.

For example: Assume you entered a forecast of 10% for the U.S.-American S&P 500 stock index. If this index actually realizes a return of 14% in the next 30 days, your forecasting error is 4 percentage points. Analogously, if the index actually realizes a return of 6%, your forecasting error is also 4 percentage points.

Payment: For each of the following ten estimates you have the possibility to earn up to \$4.0 depending on how good your estimate is. For each estimate you receive \$4.0 minus 1.0 times the

number of percentage points you are off the actual forecasting error of the respective group. For example: If you estimate a forecasting error of 6 percentage points and the actual forecasting error is 8 percentage points, then you earn $\$4.0 - 1.0 \times 2 = \2.0 .

Your total earnings from this task are the sum of your earnings from each of the ten estimates. Hence, you can earn up to \$40.

Task 4

In Task 4 we ask you to **allocate** money between two stock markets.

In particular, assume you inherit the equivalent of your annual income and want to invest this amount in the stock market. How would you distribute the money among the U.S.-American and British stock markets?

Please indicate the percentage of this amount you want to invest in the following stock indices: the U.S.-American S&P 500 and the British FTSE 100.

Note that the percentages must add up to 100.

Payment: To determine your earnings from this task, the amount of \$40 is allocated to the two stock markets according to the allocation you enter. Your earnings from this task are then the value of the invested amount after 30 days. For example: If your allocation generates a return of 10%, then you earn $\$40 \times (1 + 0.10) = \44.0 .

B Screenshots of the decision tasks

Forecast for the US-American S&P 500

Task 1 of 4

Page 1 of 9

Fenster ausschneiden

Previous monthly returns of the US-American S&P 500 stock index:

2017 Jul	2017 Aug	2017 Sep	2017 Oct	2017 Nov	2017 Dec	2018 Jan	2018 Feb	2018 Mar	2018 Apr	2018 May	2018 Jun
1.9%	0.1%	1.9%	2.2%	2.8%	1.0%	5.6%	-3.9%	-2.7%	0.3%	2.2%	0.4%

Please enter your forecast.

Over the next 30 days, for the US-American S&P 500 stock index , I expect

General population's
average forecasts

... the return to be %
(my point prediction)

7,44 %

... there is a 1-in-10 chance the actual return will be less
than %

6,68 %

... there is a 1-in-10 chance the actual return will be more
than %

8,15 %

Instructions

Next

Task I: PREDICTION-task (I_{GP})

Task 2 of 4

Page 5 of 9

Task II: CHOICE-task (I_{GP})

Assess your own and others' forecasting accuracy

Task 3 of 4

Page 7 of 9

Assess your own and others' forecasting accuracy.

What do you think is the [forecasting error](#) (in percentage points) of **your** return prediction regarding the **US-American S&P 500** stock index?

percentage points

What do you think is the [forecasting error](#) (in percentage points) of the **US-American general population's** return prediction regarding the **US-American S&P 500** stock index?

percentage points

What do you think is the [forecasting error](#) (in percentage points) of the **US-American financial professionals'** return prediction regarding the **US-American S&P 500** stock index?

percentage points

What do you think is the [forecasting error](#) (in percentage points) of the **British general population's** return prediction regarding the **US-American S&P 500** stock index?

percentage points

What do you think is the [forecasting error](#) (in percentage points) of the **British financial professionals'** return prediction regarding the **US-American S&P 500** stock index?

percentage points

Instructions

Next

Task III: ACCURACY-task (all treatments)

Task 4 of 4

Page 9 of 9

Instructions

Next

Task IV: ALLOCATION-task (I_{GP})

C Additional figures and tables

Table C.1: **Expected returns in treatment I_{NO} .** This table shows average point predictions within each sub-group and for each asset in treatment I_{NO} . The first number in each bracket denotes the standard deviations while the second number denotes the median point prediction.

Asset	Sub-group			
	FI_{US}	GP_{US}	FI_{UK}	GP_{UK}
S&P500	3.36 (3.97, 2.00)	3.04 (3.49, 2.00)	2.40 (3.35, 1.60)	2.70 (3.77, 1.80)
FTSE100	2.56 (2.73, 1.50)	2.21 (2.36, 1.00)	1.88 (1.91, 1.30)	1.75 (1.75, 1.00)
ExxonMobil	3.24 (2.71, 2.40)	2.76 (2.32, 2.00)	2.73 (2.02, 2.00)	2.19 (1.77, 1.95)
BP	4.05 (2.77, 3.45)	3.69 (2.58, 3.15)	3.58 (2.49, 3.00)	2.97 (2.19, 2.10)

Figure C.2: **Distribution of point predictions.** This figure shows histograms of point predictions for each stock index and stock including median (solid line) and mean (dashed line) estimations.

Figure C.1: **Volatility of returns.** For each asset the kernel density of the individual imputed volatility (blue line) and the historical return volatility with data from initiation (yellow line), from the last 10 years (green line), from the last 5 years (red line) and during the financial crisis (2007–2009; light green line) are shown.

Figure C.3: **Distribution of individual forecast errors.** The figure shows histograms of estimated individual forecast errors for both stock indices. Additionally, the median (solid line) and mean (dashed line) of the estimated individual forecast errors is depicted.

Figure C.4: **Distribution of other people's forecast errors.** Each graph depicts the distribution of estimated forecast errors of each subject group and both indices. For example, the first panel on the left shows the distribution of the sample's estimated forecast error of U.S. professionals' regarding the S&P500. Again, the median (solid line) and mean (dashed line) of the estimated forecast errors is plotted.

Table C.2: **Estimated forecast errors.** The table shows estimated average forecast errors and actual forecasts for both indices and each subject group.

	Sub-group			
	FI _{US}	GP _{US}	FI _{UK}	GP _{UK}
Estimated error own prediction S&P500	16.51 (10.00)	14.76 (6.39)	10.91 (6.00)	13.84 (8.00)
Estimated error of U.S. professionals S&P500	16.47 (10.00)	14.76 (6.00)	11.38 (5.00)	12.96 (7.50)
Estimated error of U.S. general population S&P500	17.45 (10.00)	16.26 (8.00)	12.21 (7.00)	15.50 (10.00)
Estimated error of U.S. professionals FTSE100	15.46 (10.00)	14.65 (6.00)	11.64 (5.00)	14.12 (8.00)
Estimated error of U.S. general population FTSE100	17.05 (10.00)	16.22 (7.00)	13.33 (7.00)	14.97 (9.00)
Estimated error own prediction FTSE100	16.62 (10.00)	15.16 (6.03)	12.69 (5.00)	13.81 (8.00)
Estimated error of U.K. professionals FTSE100	15.57 (10.00)	14.24 (6.00)	12.24 (5.00)	13.30 (7.00)
Estimated error of U.K. general population FTSE100	16.57 (10.00)	15.46 (7.00)	13.71 (6.50)	14.48 (9.00)
Estimated error of U.K. professionals S&P500	16.39 (10.00)	14.29 (6.20)	11.93 (5.00)	13.44 (7.00)
Estimated error of U.K. general population S&P500	17.65 (10.00)	15.71 (7.65)	12.31 (6.00)	14.65 (8.70)
Absolute actual error S&P500	2.31 (1.70)	2.04 (1.60)	2.55 (2.20)	2.37 (1.85)
Absolute actual error FTSE100	2.00 (1.36)	2.14 (1.53)	1.20 (0.57)	1.50 (0.73)

Table C.3: **Differences between actual and estimated prediction errors.** The table shows differences between the estimated median prediction error $\tilde{X}$ and the absolute actual median prediction error $\bar{X}$ for both indices and each sub-group of subjects. *, **, and *** represent the 5%, 1%, and 0.1% significance levels in Wilcoxon rank-sum tests.

	Sub-group			
	FI _{US}	GP _{US}	FI _{UK}	GP _{UK}
$\tilde{X}_{S\&P500} - \bar{X}_{S\&P500}$	8.40***	5.97***	4.00***	6.40***
$\tilde{X}_{FTSE100} - \bar{X}_{FTSE100}$	10.53***	7.06***	4.83***	8.43***

Significance levels: * 5%, ** 1%, *** 0.1%

Table C.4: **Individual characteristics and the distance of one's forecast from the information signal.** This table shows estimates from ordinary least squares (OLS) regressions with the absolute distance of one's forecast from the provided information signal in percentage points as the dependent variable for different treatments and subject groups. FEMALE is a dummy variable taking the value 0 for male subjects and 1 for female subjects. INCOME denotes their self-reported annual gross income. EDUCATION is an ordinal variable which takes on values between 0 and 4 and denotes different education levels, i.e. compulsory school, apprenticeship, technical college, high school and university, respectively. FINLIT is an ordinal variable between 0 and 3 where more financial literacy represents higher values. RISK_SEEKING is an ordinal variable between 1 and 7 where higher values denote more risk seeking. Clustered standard errors on the subject level are in parentheses. *, **, and *** represent the 5%, 1%, and 0.1% significance level, respectively.

Dependent variable:						
Distance to inf. signal	I_{GP}			I_{FI}		
	GP_{US}	GP_{UK}	FI_{US}	GP_{US}	GP_{UK}	FI_{US}
AGE	0.004 (0.016)	0.031* (0.014)	0.012 (0.033)	−0.003 (0.007)	−0.020* (0.009)	−0.002 (0.009)
FEMALE	−0.182 (0.386)	0.330 (0.298)	0.047 (0.662)	−0.077 (0.160)	0.001 (0.225)	−0.483 (0.270)
INCOME	−0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	−0.000 (0.000)	−0.000 (0.000)	−0.000 (0.000)
EDUCATION	0.615*** (0.177)	0.215* (0.105)	−0.309 (0.402)	0.093 (0.083)	−0.017 (0.058)	−0.024 (0.161)
FINLIT	0.336* (0.132)	0.093 (0.120)	0.505 (0.256)	0.064 (0.053)	−0.037 (0.088)	−0.009 (0.089)
RISK_SEEKING	−0.160 (0.116)	−0.024 (0.091)	−0.175 (0.230)	−0.031 (0.044)	−0.052 (0.052)	−0.033 (0.085)
Constant	1.415 (1.151)	0.934 (0.784)	3.598 (2.264)	1.286* (0.535)	2.522*** (0.591)	2.623** (0.880)
Observations	697	731	219	748	724	343
R^2	0.091	0.052	0.137	0.011	0.026	0.024
F	5.278	2.846	2.962	0.686	1.341	0.933

Table C.5: **Individual characteristics and the size of the self-reported confidence interval.** This table shows estimates from ordinary least squares (OLS) regressions with the size of the self-reported confidence interval as the dependent variable for different treatments and subject groups. FEMALE is a dummy variable taking the value 0 for male subjects and 1 for female subjects. INCOME denotes their self-reported annual gross income. EDUCATION is an ordinal variable which takes on values between 0 and 4 and denotes different education levels, i.e. compulsory school, apprenticeship, technical college, high school and university, respectively. FINLIT is an ordinal variable between 0 and 3 where more financial literacy represents higher values. RISK_SEEKING is an ordinal variable between 1 and 7 where higher values denote more risk seeking. Clustered standard errors on the subject level are in parentheses. *, **, and *** represent the 5%, 1%, and 0.1% significance level, respectively.

Dependent variable: Size of CI															
	I _{NO}				I _{GP}				I _{FI}						
	GP _{US}	GP _{UK}	FI _{US}	FI _{UK}	GP _{US}	GP _{UK}	FI _{US}	GP _{US}	GP _{UK}	FI _{US}	GP _{US}	GP _{UK}	FI _{US}		
AGE	-0.018 (0.047)	-0.042 (0.050)	-0.057 (0.035)	-0.009 (0.059)	-0.026 (0.020)	-0.028 (0.031)	0.098 (0.071)	-0.023 (0.014)	-0.004 (0.041)	-0.000 (0.048)	-0.023 (0.014)	-0.004 (0.041)	-0.000 (0.048)		
FEMALE	-0.193 (1.069)	-0.362 (0.885)	-0.520 (1.329)	-0.774 (1.018)	-0.650 (0.712)	2.886 (2.870)	2.149 (1.715)	-0.324 (0.237)	-1.016 (0.903)	-2.018 (1.635)	-0.324 (0.237)	-1.016 (0.903)	-2.018 (1.635)		
INCOME	0.000*** (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000* (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)		
EDUCATION	-0.988 (1.081)	0.403 (0.232)	0.159 (0.788)	-0.397 (0.873)	-0.087 (0.332)	0.082 (0.171)	0.539 (0.792)	0.131 (0.141)	0.337 (0.308)	0.991 (1.064)	0.131 (0.141)	0.337 (0.308)	0.991 (1.064)		
FINLIT	0.041 (0.262)	0.182 (0.328)	-0.636 (0.584)	-0.316 (0.509)	0.148 (0.194)	-0.140 (0.467)	-0.094 (0.782)	0.412*** (0.092)	-0.673 (0.534)	-0.020 (0.587)	0.412*** (0.092)	-0.673 (0.534)	-0.020 (0.587)		
RISK_SEEKING	-0.283 (0.319)	0.397 (0.312)	-0.037 (0.266)	0.577 (0.571)	-0.451 (0.288)	1.205 (1.066)	0.043 (0.298)	-0.029 (0.065)	0.076 (0.164)	0.467 (0.655)	-0.029 (0.065)	0.076 (0.164)	0.467 (0.655)		
Constant	8.574 (4.810)	2.358 (2.962)	7.300** (2.218)	3.222 (3.184)	6.256** (2.297)	0.085 (3.030)	-3.956 (4.463)	2.542** (0.815)	5.049 (2.866)	0.232 (2.925)	2.542** (0.815)	5.049 (2.866)	0.232 (2.925)		
Observations	768	736	340	376	768	788	248	788	772	372	788	772	372		
R ²	0.039	0.022	0.020	0.053	0.040	0.023	0.051	0.053	0.024	0.008	0.053	0.024	0.008		
F	143.326	1.109	1.078	3.107	1.347	2.911	3.885	3.762	0.640	1.235	3.762	0.640	1.235		

Table C.6: **Individual characteristics and allocations to the U.S.-American S&P 500 index.** This table shows estimates from ordinary least squares (OLS) regressions with the percentage allocation to the U.S.-American S&P 500 index as the dependent variable for different treatments and subject groups. FEMALE is a dummy variable taking the value 0 for male subjects and 1 for female subjects. INCOME denotes their self-reported annual gross income. EDUCATION is an ordinal variable which takes on values between 0 and 4 and denotes different education levels, i.e. compulsory school, apprenticeship, technical college, high school and university, respectively. FINLIT is an ordinal variable between 0 and 3 where more financial literacy represents higher values. RISK_SEEKING is an ordinal variable between 1 and 7 where higher values denote more risk seeking. Clustered standard errors on the subject level are in parentheses. *, **, and *** represent the 5%, 1%, and 0.1% significance level, respectively.

Dependent variable:											
Perc. allocation to S&P500				I _{GP}			I _{FI}				
	GP _{US}	GP _{UK}	FI _{US}	FI _{UK}	GP _{US}	GP _{UK}	FI _{US}	GP _{US}	GP _{UK}	FI _{US}	FI _{UK}
AGE	0.365** (0.111)	-0.197 (0.122)	0.315* (0.152)	0.022 (0.146)	0.253* (0.125)	-0.202 (0.121)	0.301 (0.182)	-0.015 (0.096)	0.058 (0.157)	0.121 (0.140)	
FEMALE	-5.874* (2.535)	-1.804 (2.603)	-8.013* (3.874)	-5.781 (3.617)	-3.901 (2.938)	-1.813 (2.399)	-4.545 (5.447)	-0.993 (2.389)	3.698 (2.269)	-1.374 (2.702)	
INCOME	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	
EDUCATION	-0.162 (1.341)	-0.852 (0.974)	0.066 (2.333)	0.078 (1.541)	-1.833 (1.703)	-0.872 (0.929)	-3.320 (2.731)	0.424 (1.369)	0.146 (0.982)	-0.633 (1.658)	
FINLIT	0.708 (0.864)	0.549 (0.870)	-0.321 (1.596)	0.393 (1.415)	1.088 (0.982)	1.528 (1.074)	1.724 (1.205)	2.366* (0.987)	0.268 (0.936)	1.300 (1.012)	
RISK_SEEKING	-0.099 (0.696)	0.068 (0.828)	0.401 (1.199)	0.838 (1.374)	-1.070 (0.742)	0.576 (0.658)	0.552 (1.629)	-1.107 (0.641)	-0.629 (0.747)	-0.030 (1.048)	
Constant	46.568*** (8.554)	59.003*** (8.705)	48.576*** (11.167)	40.890*** (10.183)	60.008*** (9.432)	55.191*** (6.249)	53.072*** (15.859)	54.846*** (6.572)	48.920*** (9.643)	51.959*** (12.081)	
Observations	768	736	340	376	768	788	248	788	772	372	
R ²	0.123	0.023	0.093	0.067	0.063	0.038	0.131	0.072	0.029	0.054	
F	5.707	0.815	2.713	1.122	3.285	2.725	1.602	2.524	1.324	1.111	