universität innsbruck

Faculty of
Economics and Statistics

Experimenting with Financial Professionals

Christoph Huber, Christian König-Kersting, Matteo M. Marini

Working Papers in Economics and Statistics

2022-07 aktualisierte Fassung vom 25. Juni 2024

University of Innsbruck
https://www.uibk.ac.at/EconStat/

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
Faculty of Economics and Statistics
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 96136
E-mail: Dean-EconStat@uibk.ac.at

The most recent version of all working papers can be downloaded at
https://www.uibk.ac.at/fakultaeten/volkswirtschaft_und_statistik/forschung/wopec/

For a list of recent papers see the backpages of this paper.

Experimenting with Financial Professionals*

Christoph Huber[1], Christian König-Kersting[2], and Matteo M. Marini[3]

[1]Institute for Markets and Strategy, WU Vienna University of Economics and Business

[2]Department of Banking and Finance, University of Innsbruck

[3]Department of Public Economics & Masaryk University Experimental Economics Laboratory, Masaryk University

June 25, 2024

Abstract

As key players in financial markets and the broader industry, financial professionals are increasingly used as experimental research participants. We review over 50 studies comparing financial professionals to laypeople and conduct systematic meta-analyses of 24 eligible studies spanning from 1986 to 2023. Our findings reveal persistent and robust support for financial professionals being more risk- and uncertainty-loving, but little evidence of superior forecasting accuracy. Further analyses indicate that larger monetary payments result in greater behavioral differences between financial professionals and laypeople, suggesting an increased susceptibility to incentives among professionals. This systematic review not only synthesizes experimental results, contributing to recent discussions about external validity and generalizability, but also highlights critical methodological considerations when experimenting with financial professionals.

JEL: B40, B41, C83, C90, C93, G41
Keywords: experimental finance, experimental methodology, finance industry, generalizability, meta-analysis

*We thank Geert Bekaert (the editor), an associate editor, two anonymous reviewers, and conference participants at the Experimental Finance 2024 in Stavanger for helpful comments and suggestions, as well as Lucas Paul Unterweger for excellent research assistance. We also thank the authors of the included studies for generously providing data and additional clarifications, which greatly contributed to the comprehensiveness and accuracy of our analysis. Financial support from the Austrian Science Fund FWF (P29362-G27 J. Huber and SFB F63) is gratefully acknowledged.
Email addresses: christoph.huber@wu.ac.at (C. Huber), christian.koenig@uibk.ac.at (C. König-Kersting), matteom.marini@gmail.com (M.M. Marini).

1 Introduction

Whereas the early experimental literature in economics and finance primarily investigated student behavior, recent studies have increasingly focused on financial professionals, the agents making the most consequential decisions on financial markets. In this study, we provide a comprehensive review and apply meta-analyses to this literature that employs financial professionals as experimental participants and compares their behavior to other samples. So far, this line of research has yielded somewhat mixed results: several studies identify differences between professionals and students (e.g., Haigh & List, 2005; Alevy et al., 2007; Kaustia et al., 2008; Cohn et al., 2014; Kirchler et al., 2018), another set of studies reports little or no behavioral differences (e.g., Rahwan et al., 2019; Holzmeister et al., 2020). Some studies also argue that differences only occur when professionals operate in tasks that mimic their usual work decisions more closely and for which they have extensive experience (e.g., Mikhail et al., 1997; Weber et al., 2005; Groysberg et al., 2008; Huber & Huber, 2020). We thus aim to synthesize this literature and to provide quantitative evidence on the behavioral differences between financial professionals and other experimental participants, as well as on whether they are driven by how much the experimental setting resembles the usual environment of professionals. In addition, we investigate and discuss methodological aspects, such as payments and study environments, that may affect decision-making when experimenting with financial professionals. Indeed, higher monetary payments and a financial framing, for example, would move the setting closer to an environment professionals are familiar with.

Endorsing laboratory studies in economics and finance, in general, one of the earliest proponents of experimental economics, Charles Plott (1982), calls such laboratory processes "[...] real [...] in the sense that real people participate for real and substantial profits and follow real rules in doing so. It is precisely because they are real that they are interesting." (p. 1486). Nevertheless, external validity is undeniably a concern with all experimental studies (e.g., Guala, 1999; Schram, 2005; Levitt & List, 2007). A key component of this issue is whether the behavior of experimental participants is representative of the behavior of people in the non-experimental, "real world" situation being modeled. This concern is even aggravated by the use of convenience samples, typically students, which are in high supply and relatively inexpensive to compensate for their participation in research studies. Plott (1982), however, refutes this as not being a "criticism of experimental methods [, but] a hypothesis about behavior in different subject pools". As such, it is actually "a call for more experiments (with businessmen subjects)" (p. 1522).

In the following decades, researchers studying finance topics, in particular, have indeed started to conduct experiments employing subjects with relevant task experience, i.e., individuals working in the finance industry: financial professionals.[1] Some studies exclusively use financial professionals as participants and study their behavior in hand-picked decision situations. While these studies reveal actual behavioral patterns of trained professionals, this approach has its limits. Oftentimes it is not feasible to conduct studies exclusively with financial professionals because their supply is limited, the costs of paying them adequately are comparatively high, and access can be complicated by compliance, privacy, and scheduling conflicts. At the same time, only employing financial professionals as participants prevents researchers from understanding which findings generalize to a broader set of individuals.

The natural solution is to conduct studies involving both financial professionals and other participants, exposing both groups to the same experimental stimuli. Such studies, which have been conducted since the late 1980s, are at the heart of this article. In particular, the aim of our endeavor is to review and synthesize all studies based on lab, lab-in-the-field, or online experiments that include a group of financial professionals and at least one comparison sample of laypeople to shed light on two fundamental questions: Do financial professionals behave differently from non-professionals? Does the existing literature reveal any methodological or thematic aspects that predict whether professionals and non-professionals differ in their behavior?

Overall, we identified more than 50 studies published in a variety of economics, finance, accounting, psychology, and general science journals, as well as several recent working papers, that analyze a variety of different behavioral outcome variables in a wide array of topics relating to the finance industry: risk and uncertainty, (financial) forecasting, asset markets, but also (dis)honesty, business culture, and other individual characteristics. As a key feature of the present study, we augment our review of the literature with systematic meta-analyses. The method of meta-analysis, however, requires reasonably comparable outcome variables, which can be found for studies in the context of *risk and uncertainty* (i.e., a preference for uncertainty), *asset markets* (i.e., market efficiency), and *forecasting* (i.e., forecasting accuracy), but not for the other research topics we identified. Moreover, it requires a sufficiently large sample of studies for which the relevant data is available – a requirement that is not met for *asset market* studies in our sample. For *risk and uncertainty* and *forecasting* we therefore provide both narrative reviews and meta-analyses,

[1] Throughout this study, we use the terms "financial professionals", "finance professionals", "people in the finance industry", and "bankers" interchangeably. In all instances, we refer to all kinds of people associated with the finance industry – that is, employees and managers, self-employed traders, brokers, and other entrepreneurs in the realm of financial markets.

while for the remaining topics, we present only narrative reviews. Together, our study represents a rich body of evidence on the differences and commonalities between financial professionals and non-professionals, and allows us to provide the most exhaustive review of this literature to date.

Our meta-analysis yields strong evidence that finance professionals are more risk- and uncertainty-loving than comparison samples of laypeople across many different studies. In terms of their forecasting abilities, however, evidence for professionals being able to forecast more accurately than laypeople is scarce and considerably less compelling: the majority of individual effects is small and the overall meta-effect does not hold up in a meta-regression that accounts for correlated errors within studies. In both meta-regressions, we find larger differences between finance professionals and non-professionals when differences in payments between the two groups are larger; other potential moderators, such as a financially-framed decision environment or comparing stated preferences with non-incentivized survey measures, however, yield no significant differences.

The remaining part of this paper proceeds as follows. Section 2 outlines details about the methodology for the systematic review and meta-analyses. In sections 3 and 4 we then focus on *risk and uncertainty* and *forecasting*, respectively, summarizing the main findings for each research topic while highlighting differences between financial professionals and non-professionals and following up with quantitative results in the form of meta-analyses. In sections 5 and 6 we summarize the main findings from the study categories not feasible for meta-analyses, i.e., on *asset markets*, but also on further results such as professionals' individual characteristics, the finance industry's business culture, and option pricing. We go on to explore methodological aspects on experimental methods and procedures and examine how studies that reveal differences between professionals and non-professionals differ from an experimental perspective (Section 7). Lastly, we discuss the overall body of evidence on financial professionals' particularities in laboratory experiments, highlight potential future directions in this line of research, and conclude (Section 8).

2 Methodology

In this study we follow two general approaches to review and aggregate the experimental literature comparing professionals in the finance industry with other populations: narrative review and systematic meta-analysis. By the method of narrative review, we aim to synthesize a broad range of studies, highlighting key themes, methodologies, and findings from seminal works. Our aim is to provide a comprehensive analysis that includes an extensive overview of all published studies in this field (either as a peer-reviewed publication or as a working paper) – from the first study we identified, published in 1986, up until very recent contributions published in 2023. This includes studies in a wide array of topics relating to the finance industry: risk and uncertainty, (financial) forecasting, asset markets, but also further results on, for example, (dis)honesty, business culture, and other individual characteristics.

As such a survey of the literature comes with limitations, we also conduct meta-analyses whenever feasible to provide a quantitative aggregation of the results which is based on systematic and statistically robust evidence. While the narrative review might suffer from potentially nontransparent weighting of individual studies, one drawback of the meta-analyses is that not all eligible studies will have sufficient information to be included – this is particularly relevant for older studies, for which the data can no longer be recovered. By conducting both a qualitative survey and quantitative meta-analyses, we are thus able to provide a comprehensive picture of the evidence for behavioral differences and similarities between finance professionals and other populations.

While our survey considers studies comparing finance professionals and laypeople in various contexts, conducting a meta-analysis also requires an outcome variable that is comparable across studies. We thus focus on two main topics of financial decision-making, *risk and uncertainty* and *forecasting*. For *risk and uncertainty*, we define the outcome variable as a stated or revealed preference measure for risk, ambiguity, and related concepts of uncertainty. For *forecasting*, we define the outcome variable as forecasting inaccuracy, i.e., as the absolute difference between a point prediction and the respective true value.

We then apply standard procedures to get the respective effect sizes in Hedges‘s g units, i.e., by calculating Cohen's $d = (\overline{Y}_{PROF} - \overline{Y}_{LAY})/\sigma_{pooled}$ where $\overline{Y}_{PROF}$ is the mean outcome for the treatment employing finance professionals, $\overline{Y}_{LAY}$ is the mean outcome for the treatment employing laypeople (e.g., students), and σ_{pooled} is the pooled standard deviation,[2] and applying a correction factor to overcome potential a small-sample bias (Cohen, 1988; Hedges & Olkin, 1985). Note that, for comparability, our meta-analyses consider the difference between financial professionals and non-professionals in a given study as an effect, but cannot address potential variations in treatment differences between these groups.

[2]The pooled standard deviation is calculated as $\sigma_{pooled} = \sqrt{\frac{(n_{PROF}-1)\sigma^2_{PROF}+(n_{LAY}-1)\sigma^2_{LAY}}{n_{PROF}+n_{LAY}-2}}$, where n_{PROF} and n_{LAY} are the sample sizes for finance professionals and laypeople. Note that we divide the respective sample size by the number of effects to avoid double counting participants when the same subjects generate data for multiple effects (Borenstein et al., 2009).

For *risk and uncertainty*, we align the signs of the effect size measures across studies such that a positive (negative) effect corresponds to financial professionals exhibiting a stronger (weaker) preference for risk, ambiguity, or a related measure of uncertainty than laypeople. For *forecasting*, the signs are aligned such that a positive (negative) effect corresponds to finance professionals being less (more) accurate in their forecasts than laypeople.

The prototypical study that we aim to include is based on a lab, lab-in-the-field, or online experiment with at least two different groups of participating subjects, namely, a group of financial professionals and a comparison sample of laypeople. We thus apply the following general inclusion criteria:

1. The study involves a laboratory, lab-in-the-field, or online experiment.
2. The study employs financial professionals as participants in comparison to at least another participant group of laypeople (e.g., students, general population samples).
3. The experimental procedures for financial professionals and non-professionals are comparable in the sense that the only difference between treatments with professionals and non-professionals are the subject's profession and expertise.

Note that criterion 1 excludes conventional lab experiments with only student participants (but no comparison group) and natural field experiments.[3] Criterion 2 excludes studies which employ other, non-financial professionals; studies which exclusively use financial professionals; and audit studies etc. which might employ professionals but in which subjects do not know that they are participating in an experiment.[4,5] Criterion 3 makes sure that the experimental design (e.g., the tasks performed) and procedures (i.e., online or in a physical laboratory) are the same across professionals and non-professionals – that is, we only compare effects from professionals observed in online experiments with effects from non-professionals observed in online experiments and we

[3]While conventional lab experiments use a standard subject pool (students), an abstract framing, and an imposed set of rules, artefactual field experiments employ a nonstandard subject pool (such as financial professionals) and framed field experiments might apply the laboratory method to a field context. Natural field experiments, in contrast, would loosen experimental control and are conducted in a naturally occurring environment in which subjects are not aware of their participation in an experiment (see Harrison & List, 2004). Our systematic review thus includes only studies that involve an artefactual field experiment with finance professionals in addition to a comparison group of non-professionals.

[4]See Fréchette (2015, 2016), for selective reviews of experimental studies with professionals as subjects with relevant task experience, more generally.

[5]We apply a comparatively narrow definition of *financial* professionals and do not include studies with "businessmen" or other professionals (e.g., we do not include Burns (1985), which employs experienced "wool buyers" in an auction experiment). Füllbrunn et al. (2022) provide a recent methodological discussion of a selection of experimental studies with financial professionals including descriptive studies and studies without a comparison group.

only compare effects from professionals observed in in-person experiments with those from non-professionals observed in in-person experiments.[6]

For each of the two main topics under consideration, *risk and uncertainty* and *forecasting*, we aim to identify studies through appropriate keyword searches in Google Scholar, EconLit, and IDEAS as relevant academic databases (for the complete list of keywords and search queries see Table A.1 in Online Appendix A). This search resulted in a list from the three databases of 116, 21, and 52 studies for *risk and uncertainty*; and 38, 21, and 111 studies for *forecasting* – including several duplicates. From the resulting items, the eligibility criteria enable us to locate 22 unique eligible studies (15 for risk and uncertainty, 2 for forecasting, and 5 that are eligible for both research topics) up until January 2024.

In addition, we performed manual search queries on the relevant databases, followed an ancestry approach by screening the references of recent contributions and related studies, and asked the relevant scientific community to send us papers that fall under our inclusion criteria by posting a request on the ESA Announcement email list in November 2023. With this approach, we identified another 13 eligible studies. Hence, we identified a total of 35 studies eligible for the meta-analyses.

As a next step, we tried to locate the relevant data for our meta-analyses, either from the publisher of the article, from public repositories linked to in the study, or by requesting them from the original authors. Overall, we were able to code 183 effects from 20 studies for *risk and uncertainty*, and 76 effects from 4 studies for *forecasting*.[7]

In addition to calculating the overall effect size, we also conduct meta-regressions including moderators that might explain the heterogeneity in effects between studies. In particular, we include the difference in level of financial incentivization between professionals and non-professionals, whether the study was conducted online or in-person (i.e., in a physical laboratory or lab-in-the-field), whether the decision was framed in a financial context, and whether the outcome variable relies on stated or revealed preferences. Each moderator variable is described in detail in Online Appendix A.

[6]The comparison of effects from professionals in in-person experiments with those from non-professionals in in-person experiments is equivalent to comparing laboratory and artefactual field experiments (Harrison & List, 2004), i.e., laboratory experiments with a standard convenience sample and a comparable experiment with a nonstandard subject pool.

[7]This includes 21 out of the 35 unique eligible studies we identified as 3 studies are suitable for meta-analyses in both *risk and uncertainty* and *forecasting*. 14 eligible studies could not be included because the necessary data is not publicly available and the authors did not respond to our requests, stated that the data were no longer retrievable, or were not willing to share the data (e.g., because their manuscript is not yet in a peer-reviewed journal); 7 of those 14 papers (50%) were published before 2010.

3 Risk and Uncertainty

Attitudes towards risk and uncertainty are believed to be core determinants of financial decision making. In this section, we explore potential differences between finance professionals and non-professionals in that area of research. First, we provide a comprehensive overview by summarizing and reviewing all identified studies that compare financial professionals and non-professionals in the context of risk and uncertainty.[8] As a second step, we augment our discussion with quantitative results from a systematic meta-analysis.

3.1 Overview

As a notable recent contribution relating to risk and uncertainty, Holzmeister et al. (2020) study what individuals perceive as risk using large samples of financial professionals and laypeople. While they do not find that the two populations differ in their perception, they show that the skewness of the return distribution and the probability of suffering losses have the largest predictive power when it comes to investments in equal expected return prospects. Their results hold for different cultural backgrounds, different countries, and different job fields of professionals. In a different experimental setting focused on responses to experimentally-induced price and volatility shocks, Huber et al. (2022) find similar results with respect to risk perception among students; financial professionals' perceived risk, on the other hand, increases as long as volatility goes up, regardless of a price change.

Hanaki (2022) shows that professionals and non-professionals also differ in their susceptibility to misperceptions. His experiment, an incentivized test of Kunz et al. (2017), demonstrates that students perceive Barrier Reverse Convertibles – a common type of structured financial products – to become less risky when a comparatively safe asset is added to the basket of underlying assets when in fact it becomes more risky. Financial professionals do not make the same mistake. While the pattern of probability misperceptions among students is reminiscent of the "dieter's paradox" (Chernev, 2011), more financially sophisticated professionals do not appear to be similarly affected.

Moving from risk perception to risk preferences and risky decision-making (see, for example, Holzmeister et al., 2022), we first consider studies on the recent COVID-19 pandemic. On the

[8] Online Appendix B summarizes information on the decision environments, duration, incentive structures, payments, sample sizes, and other notable aspects of each experiment covered by our review.

one hand, Angrisani et al. (2020) conduct risk preference elicitations of professional traders and students using the Bomb Risk Elicitation Task (BRET). They find traders to be significantly less risk averse than students in both their Pre-COVID and the COVID treatments and conclude that in the short term, the pandemic did not affect risk preferences of either group significantly. On the other hand, Huber et al. (2021) conduct a similar study with an investment task constructed from historical stock index patterns. They compare investments between financial professionals and students before and during the pandemic. For both treatments they find that financial professionals invest more than students. Among students, the level of investments is hardly affected by the pandemic but financial professionals invest significantly less. While the effect of the pandemic on risk-taking seems to be inconclusive so far, both studies have reported financial professionals to take more risk (pre-pandemic) than students. In the related study mentioned above, Huber et al. (2022) also report more pronounced responses to experimentally-induced price and volatility shocks by financial professionals in comparison to a student sample. In particular, professionals decrease their investments in a risky asset after price surge and increase their investments after a price drop. Overall, professionals' investment levels are significantly lower than those of students.

Similar patterns can also be observed in Haigh & List (2005), who test whether students and professional traders exhibit Myopic Loss Aversion (MLA) to a similar degree using the investment task of Gneezy & Potters (1997). They find that MLA is significantly more pronounced among traders than among students. While not at the core of their study, their results nevertheless also reveals that in the control condition, in which participants face the same investment decision over the course of nine rounds and receive frequent feedback, students invest significantly less in the risky asset than professionals.

Kirchler et al. (2018) study the effects of rank and tournament incentives on financial professionals and students in an investment task over multiple rounds. Supporting the findings of the previously mentioned studies, on average students are found to invest significantly less in a risky asset than financial professionals. With regard to this study's main research question, the authors report that financial professionals are susceptible to relative performance and that tournament incentives increase risk taking, but do not affect the rank-dependent investment behavior. Students, on the other hand, only react to ranking incentives if they come with monetary consequences.

Gajewski et al. (2020) directly study whether risk preferences of wealth advisors differ from those of students in laboratory and online experiments. They use the method of Tanaka et al. (2010) to estimate risk aversion, probability weighting, and loss aversion from three choice lists. While they do not find significant differences in risk aversion, the participants' choices reveal a gender-dependence in loss aversion. Female wealth advisors are found to be less loss averse than their student counterparts, while no statistically significant differences appear for males when controlling for demographic characteristics. A major caveat for this result is the low number of only eleven female wealth advisors included in the study.

In a recent large-scale study, Stefan et al. (2022) include self-reported measures of risk tolerance for a sample of the Swedish general population and financial professionals. The Likert-type measures reveal less risk aversion among the professionals than the general population sample for both aspects: risk-taking in general and risk-taking in financial matters. In addition, the authors use the investment task of Banks et al. (2019) to measure decision-making quality. Their data reveals that risk averse financial professionals do not make better portfolio choices than risk averse members of the general population. However, among the more risk tolerant, financial professionals exhibit higher decision-making quality in constructing their portfolios.

Up to now, we have presented studies that directly speak to the differences in risk preferences between financial professionals and non-professionals. Yet, the literature has studied a much broader set of issues in the context of risk and uncertainty. There are two studies, in particular, that put the focus on risk-tolerance assessments. Roszkowski & Grable (2005) study whether financial advisors and their clients differ in their ability to correctly estimate their own risk tolerance. Based on responses to a developmental version of the Survey of Financial Risk Tolerance (SOFRT), the authors conclude that clients are statistically significantly better at assessing their own risk preferences than financial advisors. In addition, they report that financial advisors show a greater risk tolerance than their clients, which is in line with much of the previously presented literature.

Similarly, Roth & Voskort (2014) study how financial agents gauge the risk preferences of their clients. Students as well as junior and senior financial professionals are asked to predict two risk preference measures (a multiple price list and a survey question) from a list of demographic characteristics and a self-reported risk preference of their clients. Senior professionals exhibit a statistically significantly stronger false consensus effect than junior professionals and non-professionals. That is, their own risk preferences correlate more strongly with their predictions for their clients than those of junior professionals and non-professionals. Junior and senior professionals are both found to be more accurate in predicting risk preferences than students.

A further branch of the literature is concerned with different biases and behavioral phenomena that affect decision making of professionals and non-professionals in decisions involving uncertainty. As a first entry in this category, List et al. (2005) let CBOT traders and undergraduate students make choices in the classic Allais paradox situation to test expected utility theory. The authors report both students and traders to exhibit choice patterns that are in line with the Allais paradox. While not formally tested, the patterns also suggest that traders are somewhat less likely to make choices in line with the paradox. While the students do not seem to reduce compound lotteries to simple lotteries, the authors cannot reject the hypothesis that traders do.

Second, List & Haigh (2010) pit the options model against the neoclassical investment model. They find that the decisions of both CBOT traders as well as undergraduate students are more in line with the options model than the classical model and that both groups seem to follow the "bad news principle" (Bernanke, 1983), i.e., taking only the expected severity of future bad news into account in deciding whether to invest in an asset today. The authors highlight that traders seem to be less responsive to payoff changes than students.

Stuying whether risk preferences are malleable, Gilad & Kliger (2008) conduct an experiment with investment advisors and undergraduates studying economics. They prime their participants with stories that are supposed to either induce risk-seeking or risk-averse behavior and elicit certainty equivalents for binomial lotteries based on stock returns. They find that both financial professionals and students are affected by the priming manipulation. Participants primed with the risk seeking story behave less risk averse than those primed with the risk averse story. Notably, professionals are reported to react stronger to the priming than students.

Most studies so far have focussed exclusively on financial decisions incolving risk and uncertainty. Razen et al. (2020) take this a step further and ask whether behavior in non-financial and financial decision contexts is the same for both financial professionals and non-professionals. They run lab-in-the-field experiments with financial professionals and participants from the general population targeted at measuring domain-dependent risk-taking. For non-financial decision contexts, they find that both professionals and non-professionals are affected by the outcome domain, i.e. the framing of outcomes as gains or losses. Both samples show a higher tendency to take the risky choice option in the loss domain than in the gain domain. For explicitly financial decision contexts, their professionals behave differently from their non-professional participants. For professionals they find behavior to be in line with the disposition effect (they are less likely to hold on to a winning stock than a losing stock), but they do not find this effect for non-professionals. Both samples are found to be similarly affected by the narrow framing bias.

3.2 Meta-analysis

For a quantitative analysis of differences in preferences towards risk and uncertainty among finance professionals and other samples we conduct a meta-analysis with 20 studies (17 published in peer-reviewed journals and 3 unpublished), containing 183 tests with 88,609 data points from at least 11 different countries.[9] While our narrative review above discusses all studies we identified in the broader context of risk and uncertainty that compare financial professionals to non-professionals, this meta-analysis focuses on one particular aspect: i.e., differences in risk and uncertainty preferences between the two participant groups. Out of the 183 included tests, 137 (75%) show a positive effect – indicating that finance professionals are more risk- or uncertainty-loving than the respective comparison sample – and only 46 (25%) show a negative effect. With 93 out of the 183, more than half of all effects are small in absolute values (i.e., Hedges's $g \leq 0.2$). A random-effects meta-analysis corroborates that first indication and shows a small but highly statistically significant effect. The overall mean effect size g is 0.195 (95% confidence interval (CI): $[0.154, 0.236]$; $p < 0.001$) with an intermediate level heterogeneity across studies ($I^2 = 29.71$; $\tau^2 = 0.022$).

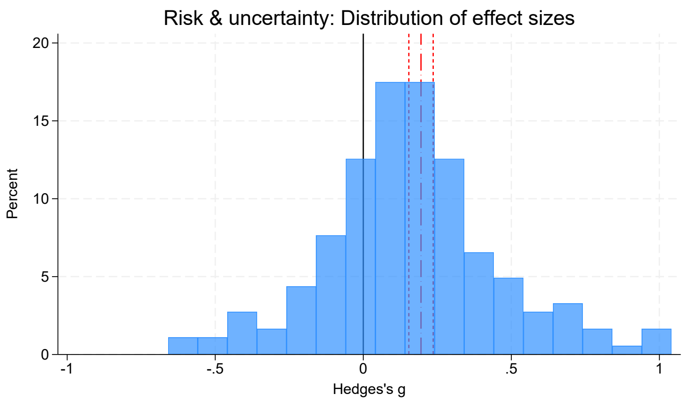

Figure 1: Risk and uncertainty, distribution of effect sizes. This figure shows the distribution of the 183 effect sizes among the 20 studies included in the meta-analysis for *risk and uncertainty* in Hedges's g units. The dashed red line corresponds to the mean meta-analytical effect $g = 0.195$ with the dotted red lines representing the 95% confidence interval.

Next, we conduct meta-regressions with moderator variables to explain the heterogeneity in effects; see Table 1. Including only the difference in financial incentives between professionals and non-professionals as a moderator yields a positive coefficient for the *Incentives diff.* variable, sug-

[9]We were able to code the country in which the data was collected for 14 of the 20 included studies. Among those studies, 23% of participants originate from the United States, 13% from Germany, 10% from the United Kingdom, and the remainder from different countries in Africa, Asia, and Europe.

gesting that the larger the gap in incentives, the greater the risk and uncertainty taken by professionals in comparison to non-professionals ($p < 0.001$, see Column (1)). The bubble plot in Fig. 2 visualizes this result. This effect is also robust to including more moderators, while the heterogeneity is reduced to $I^2 = 28.26$; see Column (2). In addition, our results suggest that the observed effect is larger in online compared to in-person experiments ($p < 0.001$), whereas framing the task in a financial context has no significant effect on the difference between financial professionals and laypeople ($p = 0.567$). To account for potentially dependent effects coming from the same study, we apply a weighted least squares meta-regression with clustered standard errors at the study level which confirms our initial results (see Column (3)). Including the average level of incentivization instead of the difference between professionals and non-professionals does not change our results: higher incentives and online experiments yield larger effects, while using a financial framing cannot explain the resulting effect size (see Table A.2 in Online Appendix A).

Table 1: Risk and uncertainty, random-effects and weighted least squares meta-regressions. This table shows the estimated coefficients from random-effects (Columns 1-2) and weighted least squares (Column 3) meta-regressions. The dependent variable is the effect size g; *Incentives diff.* denotes the difference in incentives between professionals and non-professionals; *Online* is a binary variable taking the value 1 if the study is conducted online and 0 otherwise; *Stated* is a binary variable taking the value 1 if the study measures stated (in contrast to revealed) preferences and 0 otherwise (see Table A in the Online Appendix). Standard errors are in parentheses; the WLS estimation in Column 3 uses clustered standard errors at the study level. *** p<0.01, ** p<0.05, * p<0.1.

	Dependent variable: *Effect size* (g)		
	(1) RE-MR	(2) RE-MR	(3) WLS
Incentives diff.	0.003***	0.004***	0.004**
	(0.001)	(0.001)	(0.002)
Online		0.185***	0.187
		(0.063)	(0.119)
Financial		−0.028	−0.044
		(0.048)	(0.089)
Stated		−0.048	−0.016
		(0.050)	(0.077)
Constant	0.162***	0.052	0.066
	(0.024)	(0.070)	(0.098)
Observations	183	183	183
R-squared	0.028	0.061	0.083
τ^2	0.021	0.021	—
I^2	29.12	28.26	—

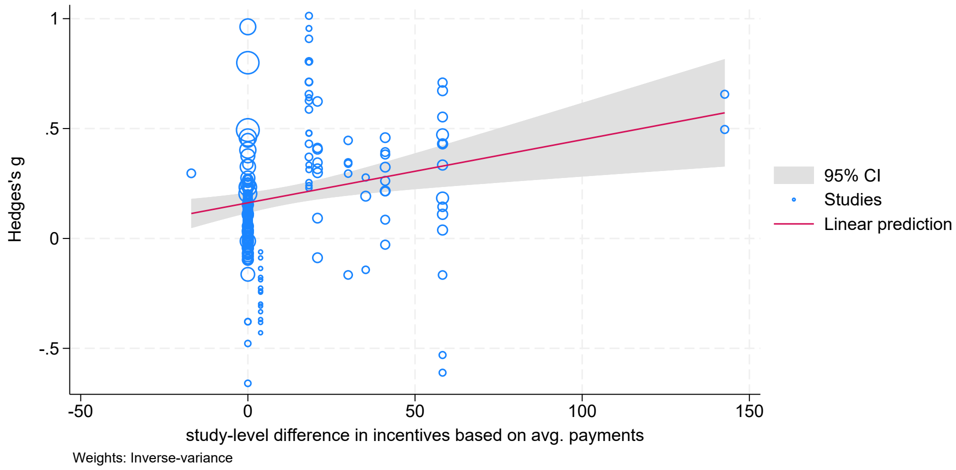

Figure 2: Risk and uncertainty, bubble plot. This figure shows a bubble plot from a meta-regression with the difference in incentives between professionals and non-professionals based on average payments as a moderator.

An Egger test suggests a small-study effect ($p < 0.001$) – i.e., results of larger studies differ from results of smaller studies. When imputing missing studies in a contour-enhanced funnel plot by means of the trim-and-fill method, however, in case of publication bias we would expect mainly non-significant studies to be missing, which is not the case (see Fig. A.3 in Online Appendix A). Also, after imputing missing studies the overall effect is even larger than observed, increasing from 0.195 to 0.258.

3.3 Summary

Looking at the large body of literature comparing risk and uncertainty preferences among financial professionals and non-professionals, a persistent finding that emerges is that financial professionals show less risk and uncertainty aversion than non-professionals.[10] This result is further corroborated by the quantitative analysis shown above, in which a systematic meta-analysis reveals a robust difference between the two participant groups in the reported direction: financial professionals are more risk- and uncertainty-loving than respective comparison groups. As an important aspect revealed by the meta-analyses, the larger the between-group gap in incentives in a given study, the greater is the observed difference in risk-taking between professionals and non-professionals. At the same time, considering study characteristics not included in the meta-

[10] If we assume that financial professionals typically have higher wealth levels than non-professionals and the wealth levels are incorporated when making the decisions in the experiments, this result is in line with the theoretical predictions of the Expected Utility framework with commonly used utility function specifications.

analysis, the two groups seem to differ in their susceptibility to psychological phenomena such as context-dependent framing, priming, and differing perceptions of outcome domains. However, in this regard, the evidence is less conclusive because the individual pieces of evidence largely stem from single studies that do not explicitly or implicitly replicate previous findings, which would allow for an accumulation of results over time.

4 Forecasting

Besides risk and uncertainty preferences, forecasts – beliefs about future asset prices – are another key aspect in all financial markets as they relate to trading behavior (e.g., Hong & Stein, 2007; Carlé et al., 2019). Heterogeneous beliefs among students and professionals have, for example, been shown to foster market inefficiencies in experimental asset markets; likewise, professionals' and students' beliefs in those markets similarly relate to the respective group's trading behavior (Füllbrunn et al., 2024). However, when it comes to forecasting naturally occurring asset prices, the first question one might ask is whether professionals – with their experiences and exposure to financial markets – are actually better forecasters than students and other laypeople. Hence, in this section, we first provide a comprehensive review of all studies that compare the forecasting abilities of financial professionals with those of non-professionals. As a second step, we again augment our discussion with quantitative results from a systematic meta-analysis.

4.1 Overview

The first study we identified in this area, Muradoğlu & Önkal (1994), elicits probabilistic stock price forecasts from portfolio managers working for a bank-affiliated brokerage house ("experts") and from what they call "semi-experts," i.e., internal auditors and managers who completed a training program on portfolio management. They find the experts' calibration to be significantly better than the semi-experts' one across all performance measures in short-term forecasts (one-week horizon). For a longer horizon (four weeks), however, semi-experts tend to be better calibrated. For the most part, this "inverse expertise effect" has, however, not been found in a follow-up study by Önkal & Muradoğlu (1996), in which they similarly compare probabilistic stock price forecasts from "experts," "semi-experts," and student subjects as "'novices" across two different task formats.

These early studies on forecasting abilities suggest that finance professionals are indeed better forecasters in some contexts, but can be even more biased than some control group in other contexts. A more recent study by Bao et al. (2022) corroborates these results by comparing financial professionals' and students' forecasting performance across four incentivized lab and field tasks. In the most abstract forecasting task, they find no performance differences. Counterintuitively, however, in more realistic lab and field tasks, they find differences but professionals do not necessarily outperform students. In forecasting a historical time series, the S&P500 stock index, without information on the stock's/index's name and the selected time period, students actually outperformed professionals. In forecasting the Nikkei index, however, – a field task in which expertise and better access to information might give professionals an advantage – financial professionals indeed have the upper hand. Barron et al. (2021) also reports results on professional investors from various financial institutions not necessarily performing better in forecasting than non-professional investors from non-financial industries; however, at the group-level, their mean forecasts seem to provide a better estimate as professionals' individual errors tend to be less correlated than those from non-professionals.

Building on her earlier work, Muradoğlu (2002) also raises the important question to what extent financial professionals' forecast errors are systematic, predictable, by experimentally comparing their stock market forecasts to those from business students. Overall, she finds prevalent optimism in real-time stock market forecasting when the stock's name is known; however, finance professionals in her sample are generally even more optimistic than the student novices. Looking into price forecasts and investor satisfaction in a sample of 150 finance professionals and 576 students, Schwaiger et al. (2020) find that professionals and students show very similar patterns across different price paths, for which they compare positive and negative final returns and vary how they are achieved (i.e., an upswing followed by a downswing and vice versa). The authors report professionals' expectations to be less prone to framing effects than students' ones and do not find professionals to be more optimistic than a non-financial control group.

As a related concept, several studies have shown that finance professionals are not just overoptimistic about potential stock returns, but also tend to be overconfident with regard to their own forecasting ability. In two studies with 43 stock market professionals and 63 students, Törngren & Montgomery (2004) find that professionals' errors in forecasting are similarly-sized than those by laypeople, but professionals are worse calibrated – i.e., they erroneously expect their own forecasts to be more accurate; thus, they are more overconfident than laypeople. In a similarly-sized sample of professionals from a large German bank and finance students, Glaser et al. (2007)

find professionals to be more overconfident than students in trend prediction tasks abstracted from specific stock markets. Comparing financial analysts' and laypeople's financial forecasts during the financial crisis of 2009/2010, Zaleskiewicz (2011) find that experts are only slightly more accurate in their stock forecasts but not in exchange rate forecasts, whereas they are more confident about their forecasts in both markets. Corroborating these earlier results with a large sample of 369 and 1224 U.K and U.S. participants from the finance industry and the general population, respectively, Huber et al. (2019) report widespread miscalibration and overconfidence among all subject groups across several stock market forecasting tasks: they vastly underestimate stock market volatility, set the respective confidence intervals for their point predictions too narrowly, and wrongly expect smaller forecast errors for their own (i.e., professionals') forecasts. In addition, Huber et al. (2019) find that finance professionals are less influenced in their forecasting by "social information", i.e., by being presented with other people's forecasts.

The "social information" shown to participants in Huber et al. (2019) essentially operates as an anchor; an initial benchmark or starting value, often irrelevant, which has been shown to alter numerical estimates (Tversky & Kahneman, 1974). In a series of experiments, Kaustia et al. (2008) specifically examine the responsiveness of finance professionals and a control group of students to different anchors in stock market forecasting. Overall, they find professionals' long-term stock return expectations to be influenced by anchors to a smaller degree than students' ones. Yet, finance professionals are not immune to and still affected by such anchors.

4.2 Meta-analysis

For a quantitative analysis of differences in forecasting performance among finance professionals and other samples we conduct a meta-analysis with 4 studies (all published in peer-reviewed journals), containing 76 tests with 25,622 data points from at least 3 different countries.[11] Out of the 76 tests, 45 (59%) show a negative effect – indicating that the forecasting error among finance professionals is smaller, i.e., finance professionals perform better – and 31 (41%) show a positive effect. Note, however, that the majority of effects is small in absolute values (Hedges's $g \leq 0.2$ for 62 (82%) out of 76 effects). Conducting a random-effects meta-analysis seems to confirm these indications and yields a small negative mean effect $g = -0.110$ (CI: $[-0.194, -0.027]$; $p = 0.010$)

[11] We were able to code the country in which the data was collected for 2 of the 4 included studies. Among those studies, 42% of participants originate from the United States, 42% from the United Kingdom, and 15% from Poland.

with little heterogeneity between studies ($I^2 = 0.00$; $\tau^2 = 0.000$).[12] The negative overall effect size suggest that professionals are on average more accurate in their forecasting. Nevertheless, this result does not hold up in a weighted least squares meta-regression, taking clustered standard errors at the study level into account ($g = -0.110$ with $p = 0.295$, CI: $[-0.386, 0.166]$).

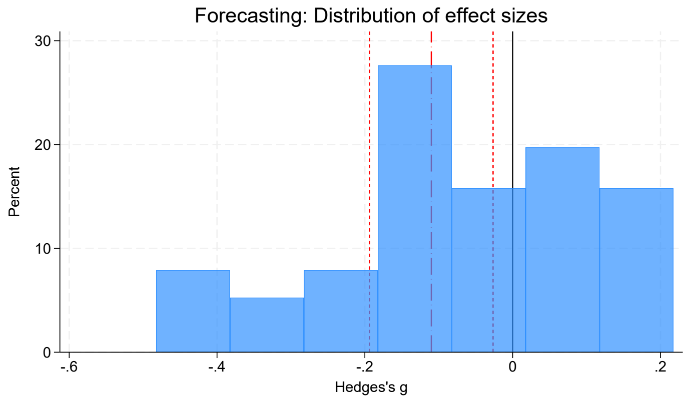

Figure 3: Forecasting: Distribution of effect sizes. This figure shows the distribution of the 76 effect sizes among the 4 studies included in the meta-analysis for *forecasting* in Hedges's g units. The dashed red line corresponds to the mean meta-analytical effect $g = -0.110$ with the dotted red lines representing the 95% confidence interval.

Next, we explore the extent to which the difference in financial incentives affects the difference in forecasting accuracy between finance professionals and laypeople. We thus conduct meta-regressions including the *Incentives diff.* variable as a moderator; see Table 2.[13] The negative coefficient suggests that the larger the gap in incentives, the better the forecasts provided by finance professionals in comparison with non-professionals ($p = 0.016$, see Column (1) in Table 2; also see Fig. 4). This result is robust to taking clustered standard errors at the study level into account in a weighted least squares estimation ($p = 0.010$, see Column (2) in Table 2). Including the average level of incentivization instead of the difference between professionals and non-professionals yields a positive coefficient with a negative constant, suggesting that forecasting accuracy among professionals and non-professionals converges with higher average payments (see Table A.3 in Online Appendix A).

[12]Note that the minimal heterogeneity here is due to the fact that 74% of effects are derived from a single study and on average 19 out of 76 effects are from the same paper.

[13]Considering other moderator variables such as *Online* and *Financial* (see Online Appendix A) in this analysis is neither feasible nor informative. All included studies are framed in a financial context and for only 3 out of the 76 tests the respective study has been conducted in person.

Table 2: Forecasting, random-effects and weighted least squares meta-regressions. This table shows the estimated coefficients from random-effects (Column 1) and weighted least squares (Column 2) meta-regressions. The dependent variable is the effect size g; *Incentives diff.* denotes the difference in incentives between professionals and non-professionals (see Table A in the Online Appendix). Standard errors in parentheses; the WLS estimation in Column 2 uses clustered standard errors at the study level. *** p<0.01, ** p<0.05, * p<0.1.

	Dependent variable: *Effect size* (g)	
	(1) RE-MR	(2) WLS
Incentives diff.	-0.012^{**}	-0.012^{***}
	(0.005)	(0.002)
Constant	-0.029	-0.029
	(0.054)	(0.036)
Observations	76	76
R-squared	0.00	0.32
τ^2	0.00	—
I^2	0.00	—

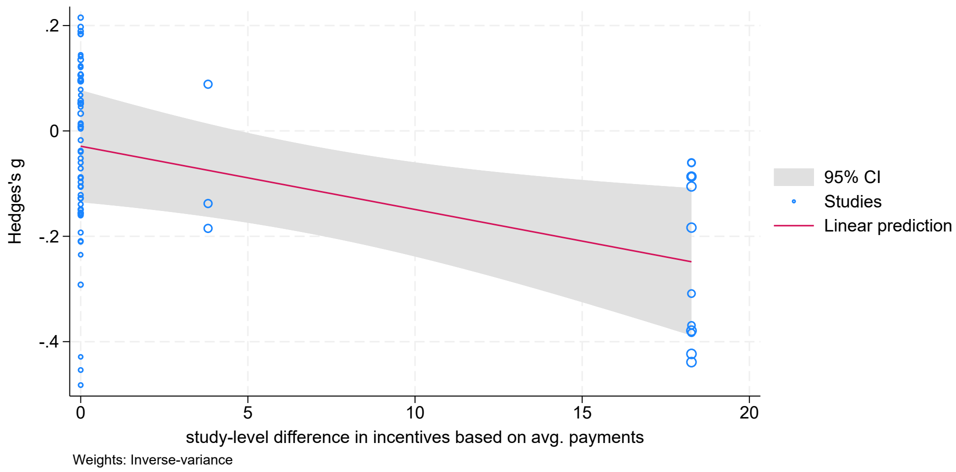

Figure 4: Forecasting, bubble plot. This figure shows a bubble plot from a meta-regression with the difference in incentives between professionals and non-professionals based on average payments as a moderator.

An Egger test suggests a small-study effect ($p = 0.037$) that may potentially be due to publication bias as smaller studies will tend to show larger effects when publication bias is present. Imputing missing studies to the dataset by means of the trim-and-fill method and recalculating the overall effect size further shows that we are indeed missing non-significant studies (see the contour-enhanced funnel plot presented in Figure A.4 in Online Appendix A). Hence, publication bias might be a reasonable explanation for the observed small-study effect. As the overall effect size increases in absolute values from -0.110 to -0.189 after imputing 26 missing studies, one can infer that these non-significant studies would have large effect sizes but would be based on small sample sizes.

4.3 Summary

So far, it seems that in most contexts – even in those relating to financial markets – financial professionals are, overall, neither better nor worse forecasters than students or laypeople. While some earlier studies have found professionals to outperform others in forecasting stock market prices, our review suggests that these results seem to be sensitive to the particular asset class (and potentially different familiarity thereof), time horizon, or context, and could not be reinforced in later studies. Also note that the earlier studies have vastly smaller sample sizes and several other differences in their experimental design: they were mostly take-home surveys conducted over several days, while later ones were conducted either online or in person within only a few minutes; only the forecasting studies since Kaustia et al. (2008) were incentivized, that is, more accurate forecasts resulted in higher monetary payouts. Nevertheless, the recent study by Bao et al. (2022) reinforces the view that forecasting performance might be context-dependent as professionals outperform students in a field task. Our meta-analysis with 76 effects – albeit from only 4 included studies – largely corroborates these results. In particular, one can identify a small difference between financial professionals and non-professionals suggesting that professionals are somewhat better forecasters, but this effect does not hold up when considering clustered standard errors at the study level. With regard to incentives, however, meta-regressions reveal higher accuracy on the part of financial professionals compared with non-professionals, the larger the between-group gap in incentives. On top of that, one fairly robust finding across most studies in our narrative review is that financial professionals tend to be more optimistic and overconfident in their probabilistic forecasts than other subject groups.

5 Asset Markets

One of the most prominent lines of research within the field of experimental finance is the work on experimental asset markets, originating in early studies by Smith (1962), Forsythe et al. (1982), Friedman et al. (1984), and Plott & Sunder (1982, 1988), among others – all looking into different aspects of asset pricing by applying the laboratory method with student participants. One particular study, Smith, Suchanek, & Williams (1988), proved pioneering in examining the foundations of bubbles and crashes in experimental asset markets, and their so-called "SSW" design became the leading paradigm in this line of research (see Palan, 2013, for a comprehensive review). With student participants, they report that price bubbles and crashes tend to form in long-lived markets, i.e. when an asset lives for multiple consecutive trading periods, where each asset pays a risky dividend at the end of each period: in a vast majority of sessions, inexperienced subjects trade assets at prices considerably above their fundamental value. To counter the argument that their results might be "an artifact of student subjects, and that businessmen who 'run the real world' would quickly learn to have rational expectations", they run one experimental session employing "professional and business people from the Tucson community" (p. 1130). While they indeed find no more rational behavior (i.e., no more efficient prices) and even larger deviations from fundamentals than in the students sessions, this early result can only be regarded as anecdotal evidence for it only comprises one independent observation and it is not clear whether the sample consists of finance professionals, in particular. Below, we summarize and qualitatively evaluate key findings from additional attempts of comparing financial professionals and non-professionals in this line of literature.

In a series of experiments, King et al. (1993) extend Smith et al. (1988) and test the robustness of their results against several modifications. Besides introducing "experienced" student subjects to the experiment (i.e., subjects participated in the same experiment once or twice more), one of these modification is the inclusion of "experienced business persons," in contrast to inexperienced students as experimental participants. They conducted one session exclusively with corporate executives from different industries, as well as one session with six over-the-counter traders and three experimenters as "insiders." While King et al. (1993) reports somewhat smaller or no bubbles with once- and twice-experienced student subjects, they still find considerable overpricing with professionals, i.e., with corporate executives or traders who are first-time participants in the laboratory experiment. Hence, they conclude that professionals show indeed similar, general patterns to inexperienced students – that is, bubbles do not disappear.

DeJong et al. (1988) run one sealed offer laboratory market experiment each with student subjects as well as with "businessmen subjects," who include accounting firm partners and corporate financial officers, to examine the price and quality choices in a principle-agent framework. Students were incentivized by monetary payouts, whereas professionals had the possibility to win a university souvenir if they manage to outperform their student counterpart in the experiment (i.e., the corresponding student subject in the same role and with the same endowment). They observe very similar results for businessmen and for students along three different performance measures (average prices, sellers' expected profits, and market efficiencies) and find no statistically significant differences between the two groups of participants.

In examining whether individual ambiguity aversion persists with trading in experimental markets, Sarin & Weber (1993) conduct two out of 14 experimental sessions with bank executives described as "bond or currency traders or advisors" with "a minimum of two years of work experience" (p. 604). Albeit only considering two market sessions, the authors report no differences in behavior compared to markets populated by students: with both subject groups, an ambiguous asset tends to yield lower prices than an unambiguous (risky) asset.

Anderson & Sunder (1995) compare students' and professionals' market outcomes and behavior in double oral auction experiments. More particularly, they analyze how well market outcomes approximate equilibrium predictions and whether experience is conducive to alleviating the level of bias which market participants exhibit in the experiment. Overall, they find that participants' prior market experience matters for price and allocation outcomes as students' behavior tends to be best predicted by a representativeness model, while prices in professionals' markets can be better approximated by a Bayesian model. Moreover, experienced professionals exhibit a considerably reduced price bias, which tends to decrease over time. Nevertheless, Anderson & Sunder (1995) conclude that the exposure to market forces which professionals clearly experienced, "does not appear to be sufficient ... to eliminate bias." (p. 196).

A similar conclusion, albeit in a different experimental set-up, is provided in the study by Weitzel et al. (2020). Weitzel et al. (2020) run a series of lab and lab-in-the-field experiments comparing market efficiency and the emergence of bubbles across several treatments. Incorporating previous results on student samples, they conduct two treatments with market characteristics previously shown to be conducive to mis- and overpricing, as well as two treatments which tend to produce comparatively efficient prices. Overall, markets with professionals exhibited less overpricing as well as fewer and smaller bubbles – prices were, on average, more efficient. Yet, looking into treatment differences within each group of subjects, Weitzel et al. (2020) report qualitatively similar patterns for students and for finance professionals: bubbles did arise even in markets populated by professionals, and the treatment differences – that is, significantly more efficient prices with a constrained cash-to-asset ratio or with short selling, and significantly less efficient prices with a comparatively high supply of cash – held for both subject pools. In a series of additional tasks, they find hardly any significant differences between students and professionals with regard to their cognitive skills. Moreover, professionals reported a higher willingness to take financial risk than students, but showed no differences in their general risk attitudes. Weitzel et al. (2020) suggest that the higher level of price efficiency with professionals could be a result of their real-world market experience and their experience with price dynamics, financial investments, and trading, more generally.

In a closely related study, Cipriani et al. (2020) contrast students and professionals traders in three experiments relating to financial markets: an SSW-type market experiment, a guessing game (Nagel, 1995), and an individual-decision variation of the guessing game. Their results confirm that finance professionals and traders, in particular, trade at prices close to fundamentals and thus foster market efficiency. Nevertheless, a classic bubble-crash pattern did emerge in one out of seven professionals markets, demonstrating that markets can be inefficient and overpriced even with professionals traders. Similarly, the guessing game reveals that professionals behave more in line with the Nash Equilibrium than students. Corroborating the results by Weitzel et al. (2020), conducting a number of side tasks, Cipriani et al. (2020) observe that the differences between professional traders' and students' behavior in the market experiment and the guessing game do not arise from the former's superior cognitive abilities, a higher level of overconfidence, or a difference in risk attitudes.

While the early studies of professionals in experimental asset markets are subject to rather vague definitions of "financial professionals" and small sample sizes, whereby they might be underpowered, by now the literature paints a more convincing picture: financial professionals and traders, in particular, tend to produce more efficient prices than student subjects.[14] Note that without any exception, all studies looking into this question also find that bubbles and market inefficiencies can and do arise even with an experienced subject pool such as financial profes-

[14]In principle, one can also conduct a meta-analysis on the 5 effects on mispricing from Weitzel et al. (2020), which indeed yields an overall effect size $g = -0.559$ (CI: $[-.936, -.182]$, $p = 0.004$), indicating a lower level of mispricing among finance professionals. With the same criteria and procedures described in Section 2 we identified another 4 eligible studies, but we could not retrieve the required data on the respective means, standard deviations, and sample sizes to include them in a meta-analysis.

sionals. Being an experienced professional in the finance industry surely helps, but alone, it is not sufficient to eliminate being susceptible to biases and other commonly observed treatment effects, such as overpricing in a high-liquidity environment.

6 Further Results

Besides potential differences between finance professionals and laypeople with regard to the core themes in finance discussed above – decisions under risk and uncertainty, asset markets and pricing, and financial forecasting – a more recent development is that researchers are increasingly interested in other aspects constituting the financial industry profession such as, for example, finance professionals' individual characteristics, their personality traits, as well as the identification and potential effects of a prevalent "business culture". meta-analyses are not feasible here as the respective outcome variables are not comparable and as most topics are only being examined by a single study. The remainder of this section thus aims to provide a qualitative review of the relevant literature, summarizing further results from all identified studies which compare finance professionals and non-professionals and which are not discussed in the sections above.

In a prominent study, Cohn et al. (2014) experimentally examine the role of a prevailing business culture within the finance industry on (dis)honest behavior using a coin tossing task, in which participants anonymously report the outcome of ten coin tosses and are compensated depending on the outcomes of the coin tosses – leaving the possibility to misreport the coin toss for one's monetary benefit. Bank employees from a large, international bank, half of whom work as private bankers, asset managers, traders or investment managers, participated in this study. As a control group, the authors employ workers from outside the banking industry as well as university students. In the treatment condition, bankers were primed by being asked several questions about their professional background to render their professional identity salient, whereas in the control condition they were asked questions unrelated to their profession. With 58.2% reportedly successful coin flips, participants from the finance industry who were primed with their professional identity behaved significantly less honestly than bankers in the control condition (51.6% successful coin flips reported). For non-banking employees and students, however, the treatment variation had no significant effect on (dis)honest behavior, whereas students were not significantly more honest than bankers.

In a large-scale replication attempt, Rahwan et al. (2019) follows up on these initial results and conduct a series of experiments with bankers and non-bankers from five different populations across three continents, all applying the same task as in Cohn et al. (2014). Overall, they do find dishonesty among finance professionals but cannot replicate the original result of a significant effect of priming bankers' professional identity on subsequent dishonesty – calling into question its generalizability beyond the originally sampled population. In a closely related study, Huber & Huber (2020) examine finance professionals' (dis)honest behavior from a different perspective by varying the situational context of a controlled, experimental cheating task. They find finance professionals to be more – not less – honest than students in two out of three treatments and a financial context framing leads to significantly more honesty compared to neutral and abstract situations among professionals, while students do not react to changes in the framing. Moreover, they suggest social norms and reputational concerns to be driving the observed behavioral differences.

Developing this idea of a prevailing business culture particular to the finance industry, which comes with social norms and informal rules on top of its legal and institutional framework, further, Cohn et al. (2017) analyze whether priming bankers on their professional identity affects their risk attitude in an experimental investment task. They apply the same priming method as in Cohn et al. (2014) with a sample of employees of a large international bank and non-banking employees. They find bankers to take significantly less risk in the priming condition and are able to replicate their initial results with bankers from several other, smaller and larger banks, but do not find the effect among non-bankers.

Extending the earlier work discussed above, Lindner et al. (2021) examine how social motives such as reputational concerns and intrinsic (self-image) motivations affect risk-taking in decision-situations involving relative performance comparisons by running lab and lab-in-the-field experiments with students and finance professionals. Their results show that professionals' behavior is to a large extent driven by intrinsic motives, with reputation playing only a minor role, while for students, in contrast, social image and reputational motives tend to be the keys determinants in their risk-taking behavior.

In one of the most ambitious contributions so far, Holmen et al. (2023) provide a comprehensive analysis of finance professionals' economic preferences and personality traits in comparison to a general population sample. In an online study with professionals working as financial analysts, financial advisers, traders, fund managers, and financial brokers, and with people from the general population, – both samples from the Swedish population – they conduct a series of experimental tasks eliciting their attitudes towards risk, losses, and skewness; their distributional (social) preferences; their trust and trustworthiness; their (dis)honesty behavior; as well as their personality traits. A key aspect of this study is that the experimental data has been merged with registry data on socio-economic characteristics provided by Statistics Sweden, allowing the authors to estimate the difference between finance professionals and the general population sample controlled for the variation in these variables. The authors report financial professionals to be more risk tolerant, more selfish, less trustworthy[15], and that they show higher levels of narcissism, psychopathy, and Machiavellianism. After adjusting for the available socio-demographic background variables, however, many of the reported effects disappear or are considerably deflated. Nevertheless, Holmen et al. (2023) observe professionals to be less risk averse, less trustworthy, more competitive, and slightly more psychopathic then a general population sample, even after controlling for their socio-economic background.

With regard to finance professionals' psychological profile, an earlier contribution by Noll et al. (2012) compares the behavior of professional traders with the behavior of psychopaths (inpatients from two German high-security psychiatric hospitals) and people from the general population. They find finance professionals' psychological profiles to be closer to laypeople than to psychopaths. In a prisoner's dilemma game, they observe professionals making more uncooperative decisions than both psychopaths and people from the general population and maximizing the difference between their own and their respective partner's profit without necessarily optimizing their own total profit.

In a lab-in-the field experiment on the impact of environmental externalities on portfolio decisions with financial professionals and students, a recent study by Duchêne et al. (2022) find professionals to act more pro-environmental than students. Nevertheless, unlike for students, the professionals' pro-environmental (as well as pro-social) preferences cannot explain their portfolio decisions.

Behavioral differences between financial professionals and non-professionals have also been studied in a variety of other contexts, including auditing, arbitrage exploitation, and information processing. Frederick & Libby (1986), for example, analyze how experienced auditors and students make predictions about how weaknesses in companies' internal control processes translate into

[15]This is also supported by the results of Gill et al. (2023), who find that university students aspiring to work in the financial industry are less trustworthy than students aiming for non-finance careers. In addition, students who actually enter the finance industry are less trustworthy than students entering other industries.

errors in financial statements. In line with their predictions, they authors find that experienced auditors have acquired knowledge that sets them apart from students when assessing the probabilities of errors occurring jointly rather than separately.

In the context of option pricing, Abbink & Rockenbach (2006) experimentally investigate professional traders and students building upon Cox et al.'s (1979) option-pricing model. They find that economics students with training in mathematical methods estimate the separating price based on the probability of the underlying stock moving in price, while professionals do not exhibit this pattern. Their behavior, in contrast, is more in line with the theoretical prediction of the option-pricing model.

Finally, relating to how professionals and non-professionals process information, Alevy et al. (2007) conduct a field experiment on information cascades with financial market professionals (CBOT) and students. They find professionals to rely more heavily on their private information and on the quality of the publicly available signal than students. Therefore, students, despite being more in line with Bayesian reasoning, do not outperform professionals market professionals in earnings. While students appear to be differently affected by gains and losses, no such domain-dependence is evident from the professionals' behavior.

7 Discussion: Experimental Methods and Procedural Aspects

After having provided a comprehensive review of the experimental results on differences and similarities between financial professionals and non-professionals, in the following section we explore the methodological aspects on experimental methods and procedures when experimenting with financial professionals.[16] This discussion includes the experiments' sample definition and characteristics (Section 7.1), recruitment and selection (Section 7.2), different decision environments (Section 7.3), and incentives (Section 7.4). Apart from our treatment of decision environments and incentives, which also incorporates quantitative results from meta-regressions, we have to rely on a qualitative reading of the related literature.[17]

[16]For more general treatments of experimental methods and procedures with standard participants, we refer to Friedman & Sunder (1994), Friedman & Cassar (2004), and Schram & Ule (2019), for example.

[17]Note that apart from incentives, the methodological aspects discussed in this section are not readily suited for inclusion in meta-analytical assessment in the context of this study, due to either a limited number of relevant studies available or the complexity involved in coding a variable to accurately represent the respective aspect. In addition, we aim to present a more comprehensive view of differences among finance professionals and non-professionals that also accommodate studies not part of the meta-analyses reported above.

7.1 Sample definition and characteristics

Several studies restrict their recruitment of professionals only to a limited extent and employ a relatively broad definition of financial professionals of multiple career stages and specializations (e.g., Törngren & Montgomery, 2004; Glaser et al., 2007; Huber et al., 2019; Holzmeister et al., 2020; Rahwan et al., 2019). In these kinds of studies, "financial professionals" appears to be used mainly as an umbrella term to describe members of the general working population that are employed in the finance industry. While the general idea of conducting finance experiments with finance professionals as participants is to examine the behavior of actual protagonists in financial markets, this broad definition not only covers a variety of different types of financial institutions (e.g., small, locally-operating commercial banks and large, internationally-operating investment banks), but crucially also a multitude of job descriptions and business divisions. A common concern is that bank tellers, loan officers, fund managers, and executives, for example, are too different from each other to be treated as a homogeneous sample. Moreover, these different groups of finance professionals might also, naturally, exhibit differential expertise necessary for particular tasks relating to the experimental setup.

Catering to concerns about external validity, experimental participants should be "representative" of the relevant decision-makers in naturally occurring situations. As such, it depends on the particular research question and experimental set-up, what type of finance professionals are appropriate participants that can generate results that generalize. A number of studies take this approach and more strongly focus on "high-skilled" employees from core finance units as the relevant agents to address their research question and to account for the complexity of the decision task, utilizing their greater experience in financial markets (e.g., Alevy et al., 2007; Cohn et al., 2014, 2017; Kaustia et al., 2008; Kirchler et al., 2018; Holmen et al., 2023; Weitzel et al., 2020). In fact, there are only very few examples of studies, which exclusively employ one particular type of financial professionals (List & Haigh, 2010, for example, specifically recruit commodity and options traders).

7.2 Recruitment and selection

A question closely connected to the definition of financial professionals is the issue of recruitment and selection as getting financial professionals to participate in studies is not an easy task. Besides the obvious challenge of getting access to a pool of potential participants in the first place, company policies, compliance considerations, and data protection laws might increase the barrier to this kind of research. Researchers have met these challenges in different ways: Some have recruited their participants at seminars, workshops, conferences and trade fairs attended by finance professionals (e.g., Kaustia et al., 2008). Some have fostered connections to financial institutions to recruit their employees as participants and have also built proprietary databases of participants (e.g., Weitzel et al., 2020). Another approach has been to recruit professionals via market research companies who maintain large international samples (e.g., Huber et al., 2019; Holzmeister et al., 2020; Kirchler et al., 2020a), or via a government agency with access to people's employment information (Holmen et al., 2023; Holzmeister et al., 2023; Stefan et al., 2022). Recently, online labor markets such as Amazon MTurk or Prolific have added options to filter potential participants by profession and job description, giving a much larger group of researchers access to self-declared financial professionals as participants for their studies (e.g., Angrisani et al., 2020; Huber & Huber, 2020).

The way of recruitment largely determines the particular group of financial professionals researchers are able to target (see Section 7.1), but also comes with potential selection issues. Close connections to financial institutions, for example, make it easier to recruit selected sub-samples of professionals that fit the study at hand. Yet, researchers lack control over whether participants are strictly participating voluntarily (one could imagine cases where invitations are circulated from their respective higher-ups). Proprietary participant pools may appear like black boxes, requiring the reader to trust that the pool contains the professionals that it claims to. At the same time, the possibility to contact professionals directly avoids having to go through and disrupt the business operations of financial institutions for future experiments. It might also make it easier to have professionals from multiple institutions partake in the same experiments, reducing concerns about institution-specific effects and selection bias. Turning to online labor markets has the advantage of gaining access to potentially much larger sample sizes than would be possible through other means. Of course, this comes at the cost of control, as researchers and readers alike face the issue of not knowing exactly who the self-reported financial professionals on the online platforms really are.

Overall, we have seen a development from small experiments with only single digit numbers of professional participants from single institutions, to more recent studies involving hundreds, if not thousands of financial professionals spanning multiple institutions and different geographic regions (e.g., Holzmeister et al., 2020; Rahwan et al., 2019). It stands to reason that these more comprehensive studies, some of which also attempt to replicate their own (and others') findings, allow us to gain a better understanding about which observations are robust and apply universally.

Common to all forms of recruitment is the issue of (self-)selection. The financial professionals who are interested in research and are willing to take part in experiments (repeatedly) may not be a random sample of all financial professionals. When participants know ex-ante that they will receive a monetary compensation for their participation, this issue might be aggravated. Employees with comparatively lower salaries might be more inclined to take part than a company's top-earners. This raises the question whether results from experiments with volunteering financial professionals, possibly even recruited from a single institution and across very different business divisions, generalize to a truly random sample of financial professionals.

7.3 Decision environments

Entwined with the issue of recruitment is the challenge of actually conducting the study. Clearly, professionals (and their respective superiors) prefer as little interruption of their usual work day as possible. At the same time, researchers are interested in having close control over the decision environment, the communication, and the interaction between participants. In the early days of experimenting with financial professionals, experiments would be conducted by recruiting professionals directly at their workplace and asking them to participate in a study. Typically, study materials were pen-and-paper-based and the sessions were conducted in conference rooms on-site at financial institutions (e.g., List & Haigh, 2005; Haigh & List, 2005). While the level of control of the decision situation can be described as rather high in these settings, the personal approach and individual recruitment have implications for the perceived (lack of) anonymity between experimenter and participants. Participants may feel identifiable and potentially perceive an obligation towards the experimenter, which may affect their decisions in the experiments. Whether this is a concern depends on the experimental task and the topic being studied.

Some studies were conducted by providing participants with the study materials to take home over the weekend and return a couple of days later (e.g., Muradoğlu & Önkal, 1994; Önkal & Muradoğlu, 1996; Muradoğlu, 2002). In these cases, some control over the decision situation, participant's focus on the task, as well as the order of and the time between individual tasks, is given up in exchange for greater flexibility for participants. Compared to individual interviews and small group experiments on-site, take-home experiments also reduce the time that institutions and participants need to set aside from their usual working hours. As such, they are a fairly unobtrusive option that may be favored my many institutions.

While very few studies have brought professional participants to traditional experimental laboratories at universities and research facilities (e.g., Roth & Voskort, 2014), the laboratory has been brought to the professionals instead. Teams that have set up temporary computerized laboratories at financial institutions and were able to largely replicate the tightly-controlled decision-environment on-site (e.g., Kirchler et al., 2018; Weitzel et al., 2020; Lindner et al., 2021). Naturally, the trade-off for institutions lies in the rather large disruption of the work day with relatively large groups of employees simultaneously taking part in an experimental session, potentially over the course of several days. For researchers, this setting comes with the added challenge of acquiring, transporting, preparing, and managing a mobile laboratory setup. Yet, in terms of the decision environment, privacy, and procedures, experiments conducted in mobile laboratories are probably closest to traditional laboratory experiments with student participants.

With fast access to the internet becoming ever more prevalent, experiments have also moved online. Online studies trade off control over the decision environment for substantial reductions in time and cost for experimenters and participants alike. As for any online studies, researchers have to prepare for participants being distracted, interrupted, or generally less attentive than in a dedicated laboratory environment. Some studies, especially those involving a large number of decisions or groups proceeding through the experiment simultaneously, might simply not be suitable for the online setting.

Our meta-analysis in the context of *risk and uncertainty* also allows us to compare the differences in effects among online and in-person experiments (both on-site and in a laboratory). While in-person experiments might offer more control, the respective meta-regressions reported above provide no indication that giving up control by conducting an experiment online affects the results in a negative way; in fact, they show an even larger gap between financial professionals and non-professionals in online experiments.

Another critical aspect of the decision environment and choice architecture in experiments comparing professionals and non-professionals is its contextual alignment with the specific task presented to participants (Harrison & List, 2004), which often comes with extensive experience in the task at hand. For experiments with financial professionals, such a contextual alignment would relate to a financial framing of the experimental task, in particular. In comparing the performance of managers and students in a strategic game, for example, Cooper et al. (1999) finds that having a relevant domain context in the experiment facilitated the development of strategic play among managers but not among students. Similarly, Huber & Huber (2020), for example, directly compare honesty levels among professionals and students in various situational context framings and find finance professionals to be more honest in a financial context (for exhibits in the context of investment decisions and forecasting, see Weber et al., 2005 and Groysberg et al., 2008, for example). However, a qualitative review of studies not included in the meta-analysis is not able to identify systematic patterns regarding the contextual alignment of experiments with finance professionals as participants. With the meta-regressions in the context of *risk and uncertainty*, in contrast, we directly compare the effects between studies set in a financial and those in a non-financial context. The respective coefficient is not statistically significant, however, suggesting that a framing in a financial domain does not affect the observed differences in risk and uncertainty preferences between financial professionals and other participant groups.

7.4 Incentives

For many economists, the issue of incentives is a sanctuary in experiments. At the very least, experimental participants should be compensated adequately for the time they spend participating in the experiment. Better yet, experiments should link the compensation to participants' performance, such that incentives exist to exert cognitive effort and make choices in line with true preferences (see Smith, 1976, for example). As a consequence, most experiments compensate participants with a combination of a fixed payment for participation and a performance-based component for their choices in the experiment. With financial professionals participating in experiments, however, deviating from these practices might be inevitable. When compliance guidelines outright forbid monetary payments for participation, compensation and incentivization have to fall back on other reward media. For example, extensive debriefing information including the research question(s), background information on the experimental methodology, and the results can be provided to participants after data collection has concluded. If advertised, this may act as an incentive to participate. When it comes to incentivizing performance in the experiment, results by Kirchler et al. (2018) suggest that for finance professionals public rankings could be used as a reward medium in lieu of monetary incentives. Others have argued that (monetarily) incentivizing decisions in experiments might not be necessary at all (see Camerer & Hogarth, 1999; Hackethal et al., 2023, for example).

If monetary payments to professionals are feasible nevertheless, the next question is on the appropriate stake size in order to sufficiently motivate participants and therefore induce meaningful behavior. Naturally, the compensation should be adjusted to participants' opportunity costs, i.e., to their foregone income from participation. While student samples have comparatively homogeneous earnings, commonly used samples of financial professionals can be considerably more heterogeneous with respect to their salaries (e.g., support staff, clerks, and c-level executives). It is thus not clear how stake sizes should be determined. For any given amount, it is likely that it would be too low for some participants of the sample and simultaneously too high for others. Assuming experimenters are indeed able to strike a suitable balance for studies involving financial professionals, the issue becomes even more apparent when the same study comprises additional samples, such as students. The most common approach to tackling this concern is compensating professionals by a multiple of the student's compensation for the same number of experimental currency units (mostly between two to four times the students' compensation; e.g., Haigh & List, 2005; Alevy et al., 2007; Cohn et al., 2014; Kirchler et al., 2018; Weitzel et al., 2020).[18]

The meta-analyses on differences among finance professionals and non-professionals in the contexts of *risk and uncertainty* and *forecasting* reported above also allow for a quantitative analysis of incentives. In particular, meta-regressions including the average difference in payments between professionals and non-professionals at the study level as a moderator show that both the gap in risk-taking and the gap in forecasting accuracy between professionals and non-professionals is larger with a higher difference in financial incentives. When considering the average level of incentivization, we find a larger difference in risk and uncertainty preferences. These results suggest that professionals are actually more susceptible to monetary incentives, and non-professionals are comparatively more reluctant to take risks as the setting gets closer to the usual high-stakes environment of professionals. Also, differences in behavior are more likely to manifest with higher incentives for professionals in comparison to those for non-professionals.

[18]Paying different groups of participants different amounts for the completion of identical tasks is not without controversy. From an ethics perspective, for example, one might reasonably question why equivalent work should result in divergent pay. It should also be noted that setting payments to be competitive with financial professionals' outside options can be prohibitively expensive for many researchers. If studies with large samples of highly selected financial professionals become the norm (and de facto requirement for publication), an undesirable compartmentalization of experimental and behavioral finance research can occur.

8 Conclusion

In this study, we surveyed more than 50 studies in the time period 1986–2023 which compare experimental results from financial professionals with those from students and other laypeople – covering a number of different topics relevant to financial economics, such as risk and uncertainty, asset markets, and (financial) forecasting. In addition to a narrative review, where we synthesize and discuss the existing literature to provide a comprehensive understanding of the disparities between financial professionals and non-professionals, we complement our analysis with a quantitative assessment in the form of systematic meta-analyses. This method is feasible and thus applied to 20 of the identified studies on *risk and uncertainty* and 4 of the identified studies on *forecasting*.

On risk and uncertainty, the meta-analysis yields strong evidence that finance professionals are more risk- and uncertainty-loving than comparison samples of laypeople across many different studies. In addition, meta-regressions reveal that incentives matter in the sense that larger monetary payments for professionals yield an overall bigger gap between professionals and non-professionals. On the relevance of how much the experimental setting resembles professional's usual environment, our meta-analysis yields no evidence for a financial framing to produce larger differences between professionals and non-professionals. While Charness et al. (2013) report risk preferences being dependent on the domains in which they are elicited, we find no significant interaction with whether the domain matches the respective participant group as hypothesized by Hanoch et al. (2006), for example. Taking other study characteristics such as priming, and differing perceptions of outcome domains into account is less conclusive, however, and each of those results only relies on a single study and is thus not part of our quantitative analysis.

On forecasting, in contrast, evidence for professionals being able to forecast more accurately than laypeople is scarce. The majority of individual effects included in our meta-analysis is small and the meta-analytical effect is not robust to taking clustered standard errors at the study level into account. Nevertheless, we find a significantly negative effect of the difference in the level of incentives between professionals and non-professionals, indicating that with a larger gap in incentives, professionals tend to outperform laypeople in forecasting tasks. Treading carefully, this results might suggest that higher incentives bring results closer to the performance standard (e.g., Hertwig & Ortmann, 2001).

In addition to the meta-analytical results, the qualitative evidence on whether financial professionals behave differently to non-professionals or whether experimental results from convenience samples of laypeople generalize to professionals is mixed. Recent large-scale experiments show finance professionals to produce fewer and smaller price bubbles in experimental asset markets. Common treatment effects which have been found among student subjects, however, also hold among professional participants – despite smaller effect sizes. Assuming that there is indeed a comparatively small but non-zero treatment effect among finance professionals, it is not surprising that early studies with rather small sample sizes did not detect statistically significant differences. Similarly, earlier reports of an inherent banking culture of dishonesty among finance professionals could not be replicated in later studies, and several other differences in their individual characteristics subside after controlling for socio-economic characteristics.

From a methodological perspective, experimenting with financial professionals comes with a number of challenging questions on the experimental design and procedures, for which no gold standard has emerged yet. Each benefit of conducting the experiments in one fashion comes with its own set of limitations and researchers must carefully consider these individual trade-offs in the context of their research agenda. The level of incentivization represents one particularly aspect of every economic experiment, for which our meta-analyses provides important results, showing that larger incentives yield a bigger behavioral differences between professionals and non-professionals. Despite the multitude of researcher degrees of freedom (see Simmons et al., 2011) – with substantially more flexibility in data collection than in standard laboratory experiments –, however, we identify no other systematic patterns in design choices predicting differences between financial professionals and other participant groups.

While the meta-analytical conclusions from this study do not seem to be the result of publication bias, one might still wonder whether our review suffers from published studies (in both peer-reviewed journals and working paper series) being biased towards statistically significant effects as studies showing non-significant differences might end up "in the file-drawer" and not be published (see, for example, Brodeur et al., 2016, 2020). For studies examining differences between financial professionals and students, however, it seems somewhat more complex and the expected direction of a potential bias is not intuitive. Many early studies in this particular area are mainly concerned with the question of whether experimental results with student subjects generalize to financial professionals (see, for example, Füllbrunn et al., 2022), aiming to demonstrate the experimental method's relevance and (external) validity. With this intention in mind, one would expect published studies to be biased towards showing *no* differences between subject groups. And indeed, most of the early studies we identified yield no significant differences between professionals and other participants – albeit with very limited sample sizes –, while more recent studies testing the same hypotheses with larger sample sizes do reveal significant differences. Nevertheless, there are also cases when a primary study reports differences between financial professionals and non professionals (Cohn et al., 2014), while a more recent study is not able to replicate this result with a larger, more diverse sample (Rahwan et al., 2019). Several potential limitations arise from this example. As mentioned above, experimental results might differ between different groups of financial professionals. Moreover, seemingly insignificant design choices such as disclosing the purpose of the study to participants might also affect results. Lastly, as a related issue, there might be potential (self-)selection: in Rahwan et al. (2019), for example, only 2 out of 27 approached financial institutions agreed to participate (Cohn et al., 2019) – information which is generally not revealed in other studies but might bear important implications for experimenting with (financial) professionals.

Since the first studies involving financial professionals as participants in a controlled experiment in the 1980s, experimental finance has come a long way in examining their behavior in financial decision contexts. This literature already spans more than 50 studies and is growing rapidly, with roughly half of them published since 2016 alone. Each individual study, however, portrays one particular experimental design and one particular series of analyses, while many more "forking paths" leading to potentially different outcomes would be available (e.g., Simmons et al., 2011; Gelman & Loken, 2013). With limited sample sizes in early studies and analytical (Botvinik-Nezer et al., 2020; Menkveld et al., 2024) and design heterogeneity (Landy et al., 2020; Huber et al., 2023) limiting the generalizability of individual study results, we believe the future of experimenting with financial professionals lies in direct and conceptual replication attempts, extensions of previous results, and a stronger focus on studies leveraging the financial professionals' unique experience and expertise in financial decision-making situations.

References

Abbink, Klaus, Bettina Rockenbach (2006). Option pricing by students and professional traders: a behavioural investigation. *Managerial and Decision Economics*, *27*(6), 497–510.

Alevy, Jonathan E., Michael S. Haigh, John A. List (2007). Information cascades: Evidence from a field experiment with financial market professionals. *The Journal of Finance*, *62*(1), 151–180.

Anderson, Matthew J., Shyam Sunder (1995). Professional Traders as Intuitive Bayesians. *Organizational Behavior and Human Decision Processes*, *64*(2), 185–202.

Angrisani, Marco, Marco Cipriani, Antonio Guarino, Ryan Kendall, Julen Ortiz de Zarate (2020). Risk Preferences at the Time of COVID-19: An Experiment with Professional Traders and Students. Working Paper.

Arnold, Vicky, Jean C Bedard, Jillian R Phillips, Steve G Sutton (2011). Do section 404 disclosures affect investors' perceptions of information systems reliability and stock price predictions? *International Journal of Accounting Information Systems*, *12*(4), 243–258.

Ba, Bocar A, Roman Rivera, Alexander Whitefield (2023). Forecasting the impact of racial uprisings, market versus stakeholders' expectations. Working Paper.

Banks, James, Leandro S. Carvalho, Francisco Perez-Arce (2019). Education, Decision Making, and Economic Rationality. *The Review of Economics and Statistics*, *101*(3), 428–441.

Bao, Te, Brice Corgnet, Nobuyuki Hanaki, Katsuhiko Okada, Yohanes E. Riyanto, Jiahua Zhu (2022). Financial forecasting in the lab and the field: Qualified professionals vs. smart students. Working Paper.

Barron, Orie E, Charles R Enis, Hong Qu (2021). Do financial professionals process information better as a group than non-professionals? *Journal of Risk and Financial Management*, *14*(5), 230.

Bernanke, Ben S (1983). Irreversibility, uncertainty, and cyclical investment. *The Quarterly Journal of Economics*, *98*(1), 85–106.

Borenstein, Michael, Harris Cooper, L Hedges, J Valentine (2009). Effect sizes for continuous data. *The handbook of research synthesis and meta-analysis*, *2*, 221–235.

Botvinik-Nezer, Rotem, Felix Holzmeister, Colin F Camerer, et al. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*(7810), 84–88.

Brodeur, Abel, Nikolai Cook, Anthony Heyes (2020). Methods matter: P-hacking and publication bias in causal analysis in economics. *American Economic Review*, *110*(11), 3634–60.

Brodeur, Abel, Mathias Lé, Marc Sangnier, Yanos Zylberberg (2016). Star wars: The empirics strike back. *American Economic Journal: Applied Economics*, *8*(1), 1–32.

Burns, Penny (1985). Experience and decision making: A comparison of students and businessmen in a simulated progressive auction. In V. L. Smith (Ed.) *Research in Experimental Economics*, vol. 3, (pp. 139–157). JAI Press.

Camerer, Colin F, Robin M Hogarth (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty*, *19*(1), 7–42.

Carlé, Tim A, Yaron Lahav, Tibor Neugebauer, Charles N Noussair (2019). Heterogeneity of beliefs and trade in experimental asset markets. *Journal of Financial and Quantitative Analysis*, *54*(1), 215–245.

Charness, Gary, Uri Gneezy, Alex Imas (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior and Organization*, *87*, 43–51.

Chernev, Alexander (2011). The dieter's paradox. *Journal of Consumer Psychology*, *21*(2), 178–183.

Cipriani, Marco, Roberta De Filippis, Antonio Guarino, Ryan Kendall (2020). Trading by Professional Traders: An Experiment. Working Paper.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates.

Cohn, Alain, Ernst Fehr, Michel André Maréchal (2019). Selective participation may undermine replication attempts. *Nature*, *575*(7782), E1–E2.

Cohn, Alain, Ernst Fehr, Michel André Maréchal (2014). Business culture and dishonesty in the banking industry. *Nature*, *516*(7529), 86–89.

Cohn, Alain, Ernst Fehr, Michel André Maréchal (2017). Do Professional Norms in the Banking Industry Favor Risk-taking? *The Review of Financial Studies*, *30*(11), 3801–3823.

Cooper, David J, John H Kagel, Wei Lo, Qing Liang Gu (1999). Gaming against managers in incentive systems: Experimental results with chinese students and chinese managers. *American Economic Review*, *89*(4), 781–804.

Cox, John C, Stephen A Ross, Mark Rubinstein (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, *7*(3), 229–263.

DeJong, Douglas V., Robert Forsythe, Wilfred C. Uecker (1988). A note on the use of businessmen as subjects in sealed offer markets. *Journal of Economic Behavior & Organization*, *9*(1), 87–100.

Duchêne, Sébastien, Adrien Nguyen-Huu, Dimitri Dubois, Marc Willinger (2022). Risk-return trade-offs in the context of environmental impact: An experiment with finance professionals and students. Working Paper.

Forsythe, Robert, Thomas R. Palfrey, Charles R. Plott (1982). Asset Valuation in an Experimental Market. *Econometrica*, *50*(3), 537–567.

Fréchette, Guillaume R. (2015). Laboratory experiments: Professionals versus students. In G. R. Fréchette, & A. Schotter (Eds.) *The Methods of Modern Experimental Economics*, (p. 360–390). Oxford University Press.

Fréchette, Guillaume R. (2016). Experimental economics across subject populations. In J. H. Kagel, & A. E. Roth (Eds.) *The Handbook of Experimental Economics*, vol. 2, (p. 435). Princeton University Press.

Frederick, David M., Robert Libby (1986). Expertise and Auditors' Judgments of conjunctive Events. *Journal of Accounting Research*, *24*(2), 270–290.

Friedman, Daniel, Alessandra Cassar (2004). *Economics Lab. An intensive course in experimental economics*. Routledge.

Friedman, Daniel, Glenn W Harrison, Jon W Salmon (1984). The Informational Efficiency of Experimental Asset Markets. *Journal of Political Economy*, *92*(3), 349.

Friedman, Daniel, Shyam Sunder (1994). *Experimental Methods: A Primer for Economists*. Cambridge University Press.

Füllbrunn, Sascha, Christoph Huber, Christian König-Kersting (2022). Experimental finance and financial professionals. In S. Füllbrunn, & E. Haruvy (Eds.) *Handbook of Experimental Finance*, (p. 64–72). Edward Elgar Publishing.

Füllbrunn, Sascha, Christoph Huber, Catherine C. Eckel, Utz Weitzel (2024). Heterogeneity of Beliefs and Trading Behavior: A Reexamination. *Journal of Financial and Quantitative Analysis*, *59*(3), 1337–1361.

Gajewski, Jean-Francois, Luc Meunier, et al. (2020). Risk preferences: are students a reasonable sample to make inferences about the decision-making of finance professionals? *Economics Bulletin*, *40*(4), 3000–3009.

Gelman, Andrew, Eric Loken (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. Working Paper.

Gilad, Dalia, Doron Kliger (2008). Priming the risk attitudes of professionals in financial decision making. *Review of Finance*, *12*(3), 567–586.

Gill, Andrej, Matthias Heinz, Heiner Schumacher, Matthias Sutter (2023). Social preferences of young professionals and the financial industry. *Management Science*, *69*(7), 3759–4361.

Glaser, Markus, Thomas Langer, Martin Weber (2007). On the Trend Recognition and Forecasting Ability of Professional Traders. *Decision Analysis*, *4*(4), 176–193.

Gneezy, U., J. Potters (1997). An Experiment on Risk Taking and Evaluation Periods. *The Quarterly Journal of Economics*, *112*(2), 631–645.

Groysberg, Boris, Paul Healy, Craig Chapman (2008). Buy-side vs. sell-side analysts’ earnings forecasts. *Financial Analysts Journal*, *64*(4), 25–39.

Guala, Francesco (1999). The problem of external validity (or “parallelism”) in experimental economics. *Social science information*, *38*(4), 555–573.

Hackethal, Andreas, Michael Kirchler, Christine Laudenbach, Michael Razen, Annika Weber (2023). On the role of monetary incentives in risk preference elicitation experiments. *Journal of Risk and Uncertainty*, *66*(2), 189–213.

Haigh, Michael S., John A. List (2005). Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis. *The Journal of Finance*, *60*(1), 523–534.

Hanaki, Nobuyuki (2022). Risk misperceptions of structured financial products with worst-of payout characteristics revisited. *Journal of Behavioral and Experimental Finance*, *33*, 100604.

Hanoch, Yaniv, Joseph G Johnson, Andreas Wilke (2006). Domain specificity in experimental measures and participant recruitment: An application to risk-taking behavior. *Psychological science*, *17*(4), 300–304.

Harrison, Glenn W, John A List (2004). Field Experiments. *Journal of Economic Literature*, *42*(4), 1009–1055.

Hedges, Larry V, Ingram Olkin (1985). *Statistical methods for meta-analysis*. Academic Ppress.

Hertwig, Ralph, Andreas Ortmann (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and brain sciences*, *24*(3), 383–403.

Holmen, Martin, Felix Holzmeister, Michael Kirchler, Matthias Stefan, Erik Wengström (2023). Economic preferences and personality traits among finance professionals and the general population. *Economic Journal*, *133*(656), 2949–2977.

Holzmeister, Felix, Martin Holmén, Michael Kirchler, Matthias Stefan, Erik Wengström (2023). Delegation decisions in finance. *Management Science*, *69*(8), 4363–4971.

Holzmeister, Felix, Christoph Huber, Stefan Palan (2022). A critical perspective on the conceptualization of risk in behavioral and experimental finance. *Handbook of Experimental Finance*, (pp. 408–413).

Holzmeister, Felix, Jürgen Huber, Michael Kirchler, Florian Lindner, Utz Weitzel, Stefan Zeisberger (2020). What Drives Risk Perception? A Global Survey with Financial Professionals and Laypeople. *Management Science*.

Hong, Harrison, Jeremy C Stein (2007). Disagreement and the stock market. *Journal of Economic Perspectives*, *21*(2), 109–128.

Hopfensitz, Astrid, Tanja Wranik (2009). How to adapt to changing markets: experience and personality in a repeated investment game. Working Paper.

Huber, Christoph, Anna Dreber, Jürgen Huber, et al. (2023). Competition and moral behavior: A meta-analysis of forty-five crowd-sourced experimental designs. *Proceedings of the National Academy of Sciences*, *120*(23), e2215572120.

Huber, Christoph, Jürgen Huber (2020). Bad bankers no more? Truth-telling and (dis) honesty in the finance industry. *Journal of Economic Behavior & Organization*, *180*, 472–493.

Huber, Christoph, Jürgen Huber, Laura Hueber (2019). The effect of experts' and laypeople's forecasts on others' stock market forecasts. *Journal of Banking & Finance*, *109*, 105662.

Huber, Christoph, Jürgen Huber, Michael Kirchler (2021). Market shocks and professionals' investment behavior–evidence from the covid-19 crash. *Journal of Banking & Finance*, *133*, 106247.

Huber, Christoph, Jürgen Huber, Michael Kirchler (2022). Volatility shocks and investment behavior. *Journal of Economic Behavior & Organization*, *194*, 56–70.

Kaustia, Markku, Eeva Alho, Vesa Puttonen (2008). How Much Does Expertise Reduce Behavioral Biases? The Case of Anchoring Effects in Stock Return Estimates. *Financial Management*, *37*(3), 391–412.

King, R. R., V. L. Smith, A. W. Williams, M. V. van Boening (1993). The robustness of bubbles and crashes in experimental markets. RH Day, Chen, P., ed. *Nonlinear Dynamics and Evolutionary Economics*.

Kirchler, Michael, Florian Lindner, Utz Weitzel (2018). Rankings and risk-taking in the finance industry. *The Journal of Finance*, *73*(5), 2271–2302.

Kirchler, Michael, Florian Lindner, Utz Weitzel (2020a). Delegated investment decisions and rankings. *Journal of Banking & Finance*, *120*, 105952.

Kirchler, Michael, Florian Lindner, Utz Weitzel (2020b). Delegated investment decisions and rankings. *Journal of Banking & Finance*, *120*, 105952.

Kunz, Alexis H, Claude Messner, Martin Wallmeier (2017). Investors' risk perceptions of structured financial products with worst-of payout characteristics. *Journal of Behavioral and Experimental Finance*, *15*, 66–73.

Lambert, Jérôme, Véronique Bessière, Gilles N'Goala (2012). Does expertise influence the impact of overconfidence on judgment, valuation and investment decision? *Journal of Economic Psychology*, *33*(6), 1115–1128.

Landy, Justin F, Miaolei Liam Jia, Isabel L Ding, et al. (2020). Crowdsourcing hypothesis tests: Making transparent how design choices shape research results. *Psychological Bulletin*, *146*(5), 451.

Leuermann, Andrea, Benjamin Roth (2012). Does good advice come cheap?—-on the assessment of risk preferences in the lab and the field. Working Paper.

Levitt, Steven D, John A List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, *21*(2), 153–174.

Lindner, Florian, Michael Kirchler, Stephanie Rosenkranz, Utz Weitzel (2021). Social motives and risk-taking in investment decisions. *Journal of Economic Dynamics and Control*, *127*, 104116.

List, John A., Michael S. Haigh (2005). A simple test of expected utility theory using professional traders. *Proceedings of the National Academy of Sciences*, *102*(3), 945–948.

List, John A., Michael S. Haigh (2010). Investment Under Uncertainty: Testing the Options Model with Professional Traders. *The Review of Economics and Statistics*, *92*(4), 974–984.

List, John A., Michael S. Haigh, Marc Nerlove (2005). A simple test of expected utility theory using professional traders. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(3), 945–948.

Menkveld, Albert J., Anna Dreber, Felix Holzmeister, et al. (2024). Nonstandard errors. *The Journal of Finance*, *79*(3), 2339–2390.

Mikhail, Michael B, Beverly R Walther, Richard H Willis (1997). Do security analysts improve their performance with experience? *Journal of Accounting Research*, *35*, 131–157.

Muradoğlu, Gülnur (2002). Portfolio managers' and novices' forecasts of risk and return: are there predictable forecast errors? *Journal of Forecasting*, *21*(6), 395–416. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/for.839.

Muradoğlu, Gülnur, Dilek Önkal (1994). An exploratory analysis of portfolio managers' probabilistic forecasts of stock prices. *Journal of Forecasting*, *13*(7), 565–578. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/for.3980130702.

Nagel, Rosemarie (1995). Unraveling in Guessing Games: An Experimental Study. *American Economic Review*, *85*(5), 1313–1326.

Noll, Thomas, Jérôme Endrass, Pascal Scherrer, Astrid Rossegger, Frank Urbaniok, Andreas Mokros (2012). A Comparison of Professional Traders and Psychopaths in a Simulated Non-Zero Sum Game. *Catalyst: A Social Justice Forum*, *2*(2).

Palan, Stefan (2013). A Review of bubbles and crashes in experimental asset markets. *Journal of Economic Surveys*, *27*(3), 570–588.

Plott, Charles R. (1982). Industrial Organization Theory and Experimental Economics. *Journal of Economic Literature*, *20*(4), 1485–1527.

Plott, Charles R., Shyam Sunder (1982). Efficiency of Experimental Security Markets with Insider Information: An Application of Rational-Expectations Models. *Journal of Political Economy*, *90*(4), 663–698.

Plott, Charles R., Shyam Sunder (1988). Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica*, (pp. 1085–1118).

Rahwan, Zoe, Erez Yoeli, Barbara Fasolo (2019). Heterogeneity in banker culture and its influence on dishonesty. *Nature*, *575*(7782), 345–349.

Razen, Michael, Michael Kirchler, Utz Weitzel (2020). Domain-specific risk-taking among finance professionals. *Journal of Behavioral and Experimental Finance*, (p. 100331).

Roszkowski, Michael J., John E. Grable (2005). Estimating Risk Tolerance: The Degree of Accuracy and the Paramorphic Representations of the Estimate. SSRN Scholarly Paper ID 2252636, Social Science Research Network, Rochester, NY. Retrieved from https://papers.ssrn.com/abstract=2252636

Roth, Benjamin, Andrea Voskort (2014). Stereotypes and false consensus: How financial professionals predict risk preferences. *Journal of Economic Behavior & Organization*, *107*, 553–565.

Sarin, R. K., M. Weber (1993). Effects of Ambiguity in Market Experiments. *Management Science*, *39*(5), 602–615.

Schram, Arthur (2005). Artificiality: The tension between internal and external validity in economic experiments. *Journal of Economic Methodology*, *12*(2), 225–237.

Schram, Arthur, Aljaž Ule (2019). *Handbook of Research Methods and Applications in Experimental Economics*. Edward Elgar Publishing.

Schwaiger, Rene, Michael Kirchler, Florian Lindner, Utz Weitzel (2020). Determinants of investor expectations and satisfaction. A study with financial professionals. *Journal of Economic Dynamics and Control*, *110*, 103675.

Simmons, Joseph P, Leif D Nelson, Uri Simonsohn (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366.

Smith, Vernon L. (1962). An Experimental Study of Competitive Market Behavior. *Journal of Political Economy*, *70*(2), 111–137.

Smith, Vernon L. (1976). Experimental Economics: Induced Value Theory. *The American Economic Review*, *66*(2), 274–279.

Smith, Vernon L., Gerry L. Suchanek, Arlington W. Williams (1988). Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets. *Econometrica*, *56*(5), 1119–1151.

Stefan, Matthias, Felix Holzmeister, Martin Holmén, Michael Kirchler, Erik Wengström (2022). You can't always get what you want: An experiment on finance professionals' decisions for others. Working Paper.

Tanaka, Tomomi, Colin F Camerer, Quang Nguyen (2010). Risk and time preferences: Linking experimental and household survey data from vietnam. *American Economic Review*, *100*(1), 557–71.

Tatarnikova, Olga, Sebastien Duchene, Patrick Sentis, Marc Willinger (2023). Portfolio instability and socially responsible investment: experiments with financial professionals and students. *Journal of Economic Dynamics and Control*, *153*, 104702.

Tversky, Amos, Daniel Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Törngren, Gustaf, Henry Montgomery (2004). Worse Than Chance? Performance and Confidence Among Professionals and Laypeople in the Stock Market. *Journal of Behavioral Finance*, *5*(3), 148–153.

Weber, Elke U, Niklas Siebenmorgen, Martin Weber (2005). Communicating asset risk: How name recognition and the format of historic volatility information affect risk perception and investment decisions. *Risk Analysis: An International Journal*, *25*(3), 597–609.

Weitzel, Utz, Christoph Huber, Jürgen Huber, Michael Kirchler, Florian Lindner, Julia Rose (2020). Bubbles and financial professionals. *The Review of Financial Studies*, *33*(6), 2659–2696.

Zaleskiewicz, Tomasz (2011). Financial forecasts during the crisis: Were experts more accurate than laypeople? *Journal of Economic Psychology*, *32*(3), 384–390.

Önkal, Dilek, Gülnur Muradoğlu (1996). Effects of task format on probabilistic forecasting of stock prices. *International Journal of Forecasting*, *12*(1), 9–24.

Online Appendix

A Supplementary information on the meta-analyses

A.1 Literature search

Table A.1: Keywords and search queries. This table shows the the keywords and search queries used to identify relevant studies comparing finance professionals and non-professionals in *risk & uncertainty* and *forecasting*, which fulfill our inclusion criteria (see Section 2). For the search queries (bottom panel), we use several different combinations of keywords for identifying experiments, studies with financial professionals, and studies in the domain of our two main topics under investigation. The relevant databases are Google Scholar (https://scholar.google.com/), EconLit (https://www.aeaweb.org/econlit/, and IDEAS (https://ideas.repec.org/).

Experiment	Financial professionals	Risk and uncertainty	Forecasting
experiment	finance professionals	risk, risky	forecast
experiments	financial professionals	risk preferences	forecasts
experimental	practitioners	risk aversion	forecasting
laboratory	bankers	risk taking	expectations
field experiment	financial advisors	ambiguity	
field experiments	financial advisers	uncertainty	

Search queries

Risk and uncertainty	
Google Scholar	intext:"experiment" intitle:"finance OR financial" intitle:"professionals OR practitioners" intext:"risk OR risks OR risky"
EconLit	("finance professionals" OR "financial professionals" OR "bankers" OR "financial advisors" OR "financial advisers") AND ("risk preferences" OR "risk aversion" OR "risk taking" OR "ambiguity" OR "uncertainty") AND ("experiment" OR "experiments" OR "experimental" OR "laboratory" OR "field experiment" OR "field experiments")
IDEAS	("finance professionals" | "financial professionals" | "bankers" | "financial advisors" | "financial advisers") + ("risk preferences" | "risk aversion" | "risk taking" | "ambiguity" | "uncertainty") + ("experiment" | "laboratory" | "field experiment")
Forecasting	
Google Scholar	intext:"experiment" intitle:"finance OR financial" intitle:"professionals OR practitioners" intext:"forecast OR forecasts OR forecasting"
EconLit	("finance professionals" OR "financial professionals" OR "bankers" OR "financial advisors" OR "financial advisers") AND ("Forecasts" OR "Forecasting" OR "Expectations") AND ("experiment" OR "laboratory" OR "field experiment")
IDEAS	("finance professionals" | "financial professionals" | "bankers" | "financial advisors" | "financial advisers") + ("Forecasts" | "Forecasting" | "Expectations") + ("experiment" | "laboratory" | "field experiment")

A.2 Moderator variables

- *Incentives diff.*: The continuous variable *Incentives diff.* measures the absolute difference in the level of incentivization between professionals and non-professionals. We take the average payment of professionals and non-professionals treatments at the study level, and then use a PPP conversion factor by the World Bank to standardize them. The PPP conversion factor is (1) year-based: we use the publication (or WP) year given that in most cases the text does not report the date of the experiment. In Hackethal et al. (2023) and Holmen et al. (2023) the PPP conversion factors for 2022 are used, because those for 2023 are not yet available at the time of this writing; (2) country-based: we use "Europe" or "OECD" in case the experiment has been run in multiple (European or OECD) countries, respectively. We set the variable to 0 whenever professionals and non-professionals are by design equally rewarded and separate averages for the two groups are not available.

- *Incentives level*: The continuous variable *Incentives level* measures the level of incentivization among professionals and non-professionals. We take the average payment of professionals and non-professionals treatments at the study level, and then use a PPP conversion factor by the World Bank to standardize them.

- *Online*: The binary variable *Online* takes the value 1 if the study was conducted online, and 0 otherwise (i.e., for lab and lab-in-the-field studies).

- *Financial*: The binary variable *Financial* takes the value 1 if the study environment was framed in a financial context, and 0 otherwise.

- *Stated*: The binary variable *Stated* takes the value 1 if the outcome variable relies on stated preferences (i.e., non-incentivized survey measures), and 0 otherwise (i.e., when the outcome variable relies on revealed preferences).

Table A.2: Risk and uncertainty, random-effects and weighted least squares meta-regressions with the level of incentives. This table shows the estimated coefficients from random-effects (Columns 1-2) and weighted least squares (Column 3) meta-regressions. The dependent variable is the effect size g; *Incentives level* denotes the average level of incentives for professionals and non-professionals; *Online* is a binary variable taking the value 1 if the study is conducted online and 0 otherwise; *Stated* is a binary variable taking the value 1 if the study measures stated (in contrast to revealed) preferences and 0 otherwise (see Table A). Standard errors are in parentheses; the WLS estimation in Column 3 uses clustered standard errors at the study level. *** p<0.01, ** p<0.05, * p<0.1.

	Dependent variable: *Effect size* (g)		
	(1) RE-MR	(2) RE-MR	(3) WLS
Incentive level	0.005***	0.007***	0.008**
	(0.001)	(0.001)	(0.003)
Online		0.209***	0.227**
		(0.060)	(0.108)
Financial		0.046	0.052
		(0.050)	(0.061)
Stated		−0.022	0.003
		(0.049)	(0.071)
Constant	0.132***	−0.089	−0.121
	(0.026)	(0.080)	(0.131)
Observations	183	183	183
R-squared	0.214	0.316	0.191
τ^2	0.017	0.015	—
I^2	24.91	22.23	—

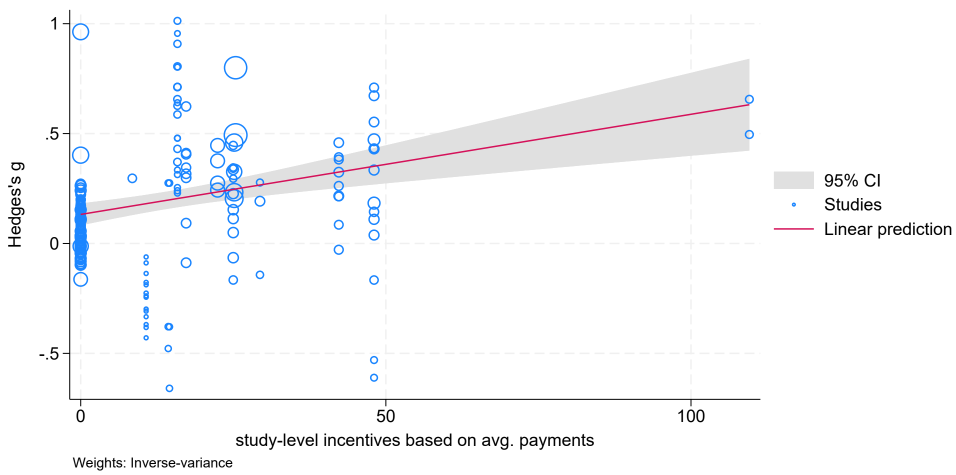

Figure A.1: Risk and uncertainty, bubble plot with the level of incentives. This figure shows a bubble plot from a meta-regression with the level of incentives among professionals and non-professionals based on average payments as a moderator.

Table A.3: Forecasting, random-effects and weighted least squares meta-regressions with the level of incentives. This table shows the estimated coefficients from random-effects (Column 1) and weighted least squares (Column 2) meta-regressions. The dependent variable is the effect size g; *Incentives level* denotes the average level of incentives for professionals and non-professionals (see Table A). Standard errors are in parentheses; the WLS estimation in Column 2 uses clustered standard errors at the study level. *** p<0.01, ** p<0.05, * p<0.1.

	Dependent variable: *Effect size* (g)	
	(1) RE-MR	(2) WLS
Incentives level	0.016**	0.016**
	(0.007)	(0.005)
Constant	−0.418***	−0.418**
	(0.135)	(0.125)
Observations	76	76
R-squared	0.00	0.32
τ^2	0.00	—
I^2	0.00	—

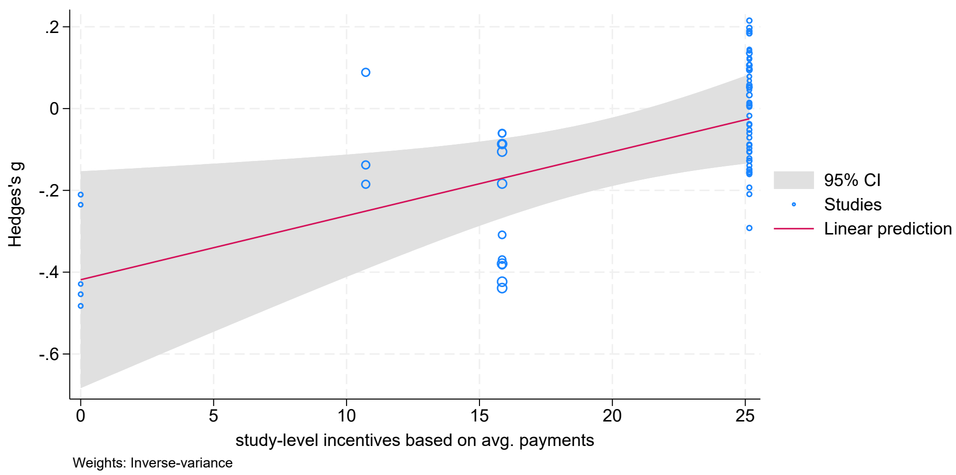

Figure A.2: Forecasting, bubble plot with the level of incentives. This figure shows a bubble plot from a meta-regression with the level of incentives among professionals and non-professionals based on average payments as a moderator.

A.3 Publication bias

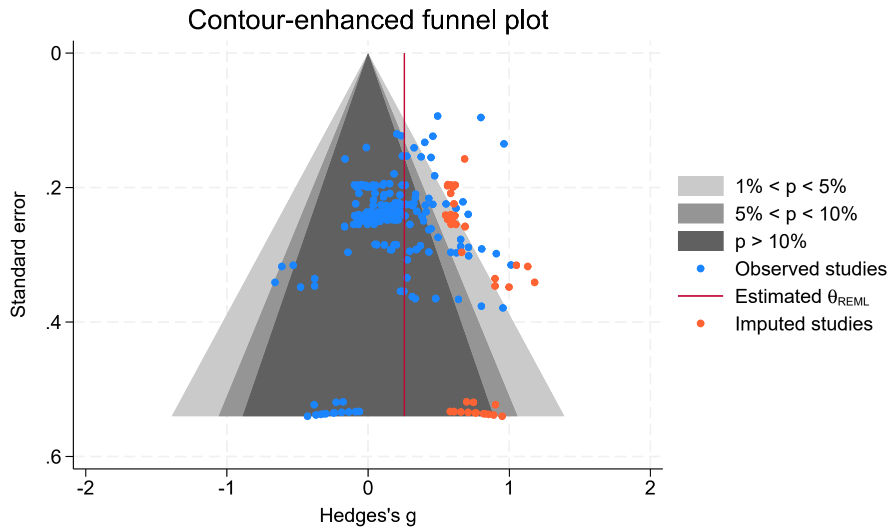

Figure A.3: Risk and uncertainty, contour-enhanced funnel plot.

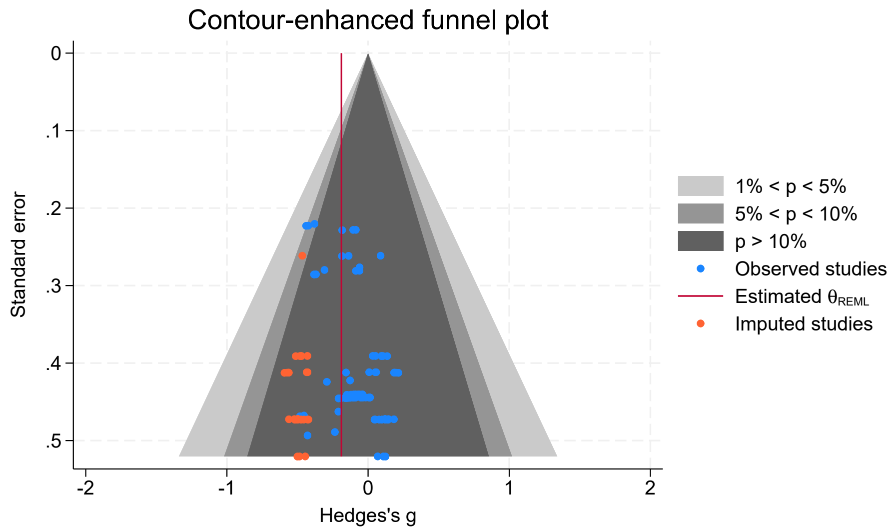

Figure A.4: Forecasting, contour-enhanced funnel plot.

B Study characteristics

B.1 Risk and Uncertainty

Haigh & List (2005). The professionals were 54 "locals, brokers, clerks and exchange employees (e.g., floor managers or and market reporters) who worked in the open outcry environment" (p. 527) with multiple years of experience from the Chicago Board of Trade (USA). No differences between different participant types among the professionals were found. The 64 undergraduate students were recruited at the University of Maryland. Student sessions were conducted in a laboratory-like setting on campus. Professionals took part in a dedicated room at the CBOT. Students earned USD 0.01 per unit while professionals received USD 0.04 per unit.

List & Haigh (2005). The professionals were 54 "locals, brokers, clerks and exchange employees (e.g., floor managers or and market reporters) who worked in the open outcry environment" (p. 946, footnote k) with multiple years of experience from the Chicago Board of Trade (USA). No differences between different participant types among the professionals were found. It is not explicitly stated whether these are the exact same professionals as in Haigh & List (2005). Undergraduate students were recruited at the University of Maryland (College Park). Student sessions were conducted in a laboratory-like setting on campus while the professionals took part at CBOT. Students received USD 0.01 per unit, professionals got USD 0.04 per unit.

Roszkowski & Grable (2005). The professionals were 386 financial advisors from all parts of the United States of America who had graduated The American College's Master's in Financial Services (MSFS) Program. The majority worked in the life and health insurance sector (64%) with the next biggest group working in financial planning (17%). Each participating advisor was asked to select some of their clients, resulting in a sample of 458 laypeople from all regions in the US. 45% of these participants worked in the private sector and 42% reported to be self-employed. No control was exercised over the environment while filling-in the SOFRT questionnaires.

Gilad & Kliger (2008). The professional participants were 44 investment advisors working in large commercial banks and accountants from CPA firms. The student sample consisted of 52 undergraduate students of economics. Although not explicitly stated, it is reasonable to assume that all participants were from Israel, as payments were made in Israeli New Shekel (NIS). The experiments took place in a controlled laboratory setting.

List & Haigh (2010). The professionals were 55 commodity (futures) and option traders from the Chicago Board of Trade (USA). The student sample consisted of 75 undergraduate students from the University of Maryland. Students earned USD 0.01 per unit while professionals received USD 0.04 per unit.

Table C.1: Articles on risk and uncertainty

Article	Environment	Duration	Incentives	Average payments
Haigh & List (2005)	controlled	25 min	proper	Students: USD 10 Professionals: USD 40
List & Haigh (2005)	controlled	-	proper	not reported
Roszkowski & Grable (2005)	uncontrolled	30-60 min	none	-
Gilad & Kliger (2008)	controlled	-	Students: fixed Professionals: none	Students: NIS 45 (USD 10)
List & Haigh (2010)	controlled	30 min	proper	Students: USD 11.75 Professionals: USD 47
Roth & Voskort (2014)	controlled	50 min	proper	EUR 11.92
Kirchler et al. (2018)	controlled	45 min	proper	Professionals: EUR 52 Students: EUR 18
Angrisani et al. (2020)	controlled + online	-	proper	Students: GBP 25 Professionals: GBP 250
Gajewski et al. (2020)	Students: controlled + online Professionals: online	-	Students: proper Professionals: -	Students: EUR 5.70
Holzmeister et al. (2020)	online	-	-	-
Huber et al. (2021)	online	20 min	proper	Students: EUR 5.45 Professionals: EUR 20.27
Razen et al. (2020)	online	11 min	fixed	USD 25 (with 20% chance)
Hanaki (2022)	online	25 min	proper	not reported
Stefan et al. (2022)	online	45 min	proper	SEK 238.9 (USD 30)
Arnold et al. (2011)	online	-	none	-
Duchêne et al. (2022)	controlled	45 min	proper	Students: EUR 13.45 Professionals: EUR 216.81 (1 of 10 professionals paid)
Hackethal et al. (2023)	online	13.22 min	FLAT: fixed INCENTIVES: proper	FLAT: Students: EUR 3 Professionals: EUR 12 INCENTIVES: Students: EUR 7.67 Professionals: EUR 31.43
Holmen et al. (2023)	online	15 min	proper	SEK 211.13 (USD 23.50)

Article	Environment	Duration	Incentives	Average payments
Hopfensitz & Wranik (2009)	controlled	2x 60 min	proper	Students: CHF 31.30 (USD 27) Professionals: CHF 58.90 (USD 52)
Kirchler et al. (2020b)	online	10 min	proper	USD 62.90 (20% paid)
Lambert et al. (2012)	controlled	60 min	Students: fixed Professionals: none	Students: EUR 15 Professionals: none
Leuermann & Roth (2012)	controlled	50 min	proper	EUR 11.92
Tatarnikova et al. (2023)	controlled + online		proper	Students: EUR 21 Professionals: EUR 92 (1 of 10 paid)

Roth & Voskort (2014). There are three different samples in this study. The first sample of professionals were 38 senior professionals from large financial advisory agencies and local banks in Germany. The second sample consisted of 52 junior professionals from a banking specific advanced training institution (applied university) in Germany. The third sample included 77 students from Heidelberg University (Germany). All sessions took place in controlled environments either in the laboratory at Heidelberg University (all sessions with non-professionals and three sessions with professionals) or on-site at the institutions (four sessions with professionals).

Kirchler et al. (2018). We focus on the main treatments for the relevant comparison of financial professionals and students. A total of 252 professionals and 432 students participated in lab-in-the-field experiments. Professionals were recruited from “major financial institutions in several OECD countries” and worked in “private banking, trading, investment banking, portfolio management, fund management, and ealth management” (p. 2278). Professionals took part in a mobile laboratory which was set up in conference rooms at participating financial institution. To create some degree of anonymity, sessions were generally populated with professionals from different institutions. Students from multiple disciplines and programs of study were recruited at the University of Innsbruck (Austria) and took part in the local experimental laboratory. One fifth of the professionals participants was randomly selected for payment, with professionals receiving EUR 52 on average (maximum EUR 600) for 45 minutes. Average payments were approximately 2.7 times the professionals after tax hourly wage. Students’ incentives were “scaled down to one-third of the professionals’ payoffs” (p. 2283), resulting in average payments of EUR 18 (maximum EUR 323).

Angrisani et al. (2020). The study was conducted in two waves, about 13 months apart. The first wave was conducted in an experimental laboratory, while the second wave of data collection took place online. The same participants that took part in the first wave were invited to take part in the second. The professionals were traders, proprietary traders, sales-traders, portfolio managers, and others, with the majority being traders of some kind. They were described as working "in a variety of financial markets, such as equity, equity derivatives, FX, fixed income, and commodities" (p. 5). Students were undergraduates from various disciplines. Notably, 80% of the student sample is male, which is close to the male gender ration in the professional sample of 86% . The data analysis is based on 48 financial professionals and 60 students who took part in both waves. In the first wave, professionals (students) earned GBP 3.70 (4.90), while in the second wave professionals (students) earned GBP 4.10 (4.90) for the main task the article reports on. In the first wave, the experiment had multiple other parts resulting in average earnings of professionals (students) of GBP 250 (25). In the second wave, participants received an additional fixed fee of GBP 25 in addition to their earnings from the task.

Gajewski et al. (2020). The article reports on three samples. Professionals were 57 French wealth advisers recruited via an e-mail to the French professional association. The professionals took part online but the article does not mention any monetary compensation for participation. A sample of 102 French business school students participated in the laboratory. They faced proper incentives and earned EUR 5.70 on average. A second sample of 448 students from the same institution took part online. No monetary compensation is mentioned for this sample.

Holzmeister et al. (2020). The 2213 finance professionals in this study are split 86%/14% between the finance and the insurance industry. They work in accounting and controlling, advisory services, analysis and research, fund and portfolio management, administration, investment banking, private banking, risk management, sales, general management, trading, and brokerage. The laypeople sample consists of 4559 members of the general population (not working in finance or the insurance industry) from Brazil, China, Germany, India, Japan, Russia, United Kingdom, United States of America, and South Africa. The experiment was conducted online and no performance based payments were made. There is no mention of fixed payments in the article either.

Huber et al. (2021). Two waves of data collection are reported in this article. The first wave was conducted in December 2019. 202 financial professionals and 282 students participated. The second wave of data collection followed in the first month of the COVID-19 pandemic (March 2020) with an additional 113 professionals and 216 students. Notably, different participants took part in the two waves to ensure that wave two participants could not recall their previous experience in the experiment. All data was collected online. The professionals were recruited from the before.world participant pool and included job functions such as investment and portfolio management, trading, and financial advice. The student sample consisted of economics and business students from the Innsbruck EconLab subject pool at the University of Innsbruck, Austria. Decisions in the experiment were monetarily incentivized for both students and professionals. Students received an endowment of EUR 5 while professionals started with EUR 20. The experiment took about 20 minutes to complete and average total payments were EUR 20.27 for financial professionals and EUR 5.45 for students.

Razen et al. (2020). The professional sample comprised 202 US financial professionals working as advisors, in sales, as portfolio and risk managers, or in support functions. The non-financial professional sample included 408 participants from the US general working population. This sample included mostly people working in services, education, and manufacturing and construction. All data was collected online in May 2018. The participants were recruited on before.world and via an international market research company. One out of five participants were randomly selected to be paid for their participation and received a flat fee of USD 25. The experiment took on average 11 minutes of their time.

Hanaki (2022). Eighty-four Certified Financial Accountants (CFA) and 87 students from the University of Osaka, Japan, were recruited for the online experiment. The professionals had previously participated in finance experiments and indicated to be willing to participate again. After removing the 10% fastest and 10% slowest participants as well as enforcing monotonicity in responses, the analysis is based on the decisions of 64 professionals and 63 students. Professionals (students) received a fixed payment of JPY 1000 (500) and JPY 10 (5) for each experimental currency unit earned in the experiment. Ten percent of the participants were selected for real payments administered via Amazon gift certificates (emailed to participants). The experiment took 25 minutes to complete.

Stefan et al. (2022) and Holzmeister et al. (2023). The two articles report on on different elements of fundamentally the same experiment. As such, they share the same sample characteristics and experimental details. The financial professionals were 408 Swedish financial analysts, investment advisors, traders, fund managers, financial brokers, among others. The sample of laypeople consisted of 550 non-financial professionals. Invitations were sent to a representative sample of the Swedish working population. All observations were collected online.

Arnold et al. (2011). Participants in the experiment were 67 investment professionals from large and small financial firms recruited through a survey company. In addition, 100 non-professional investors took part. All participants stated that they regularly assess companies' financial data. Participants were given information about an anonymized high-tech manufacturing company based on 10-K forms and were asked to assess reliability and company risk as well as predict future stock prices. The study took place online. No information about duration of the study is given. The authors thank participants for their time but do not mention and monetary compensation.

Duchêne et al. (2022). The samples comprise 279 students and 190 finance professionals. Professionals took part in a mobile laboratory in Casablanca in 2019. Students participated in 2020 in a standard university laboratory. One out of every 10 professionals was paid, averaging EUR 216.81. All students were paid. The experiment took about 45 minutes to complete.

Hackethal et al. (2023). The study involved three samples: professionals, private investors, and students. The professional sample comprises 244 fund managers, portfolio managers, analysts, and risk managers, mainly recruited from before.world. The private investors make up 821 participants and were recruited from a panel of clients of a German bank, maintained by Goethe University Frankfurt. In addition, 415 students recruited at the University of Innsbruck took part.

The experiment was conducted online in 2020 and took 13.22 minutes (median) to complete. In the INCENTIVES treatments, professionals received an average final payment of EUR 31.43, private investors received EUR 31.72, and students received EUR 7.76. In the FLAT treatments they received EUR 12, EUR 12, and EUR 3 respectively.

Hopfensitz & Wranik (2009). The sessions were run in the laboratory of the University of Geneva (Switzerland) in 2009 and consisted of two sessions of 1 hour length each. One session involved personality questionnaires while the other was the experiment itself. Professionals completed both sessions on the same day, while there was a break between sessions for students. Thirty-one Professionals were recruited by their HR Manager at a small private bank in Switzerland by email and earned CHF 58.90 on average. In addition, 46 students took part and earned CHF 31.30 on average.

Kirchler et al. (2020b) and Holmen et al. (2023). Finance professionals were 298 analysits, advisors, traders, fund managers, and financial brokers from Sweden. They received hardcopy invitations to participate in an online study set up in cooperation with Statistics Sweden. In addition, a random population sample of 395 people from Sweden (excluding financial professionals) took part online. The study took 15 minutes to complete and participants receveid SEK 211.13 on average.

Lambert et al. (2012). The study involved 20 loan officers from major French banks who were contacted individually and filled in the questionnaires on laptop computers during appointments at their respective offices. In addition, 64 business students took part in four sessions at the University of Montpellier. While the professionals did not receive any compensation, students received a fixed amount of EUR 15 each. The experiment took 60 minutes to complete.

Leuermann & Roth (2012). The study consists of two parts, an online survey and a laboratory experiment. The online survey was conducted at the end of 2010 and involved a lottery incentive of EUR 50 for participation. The laboratory experiment was conducted between April 2011 and January 2012. Student participants were recruited at Heidelberg University, while professionals worked at local banks and a financial advisory agency. Some sessions with professionals took place in the laboratory, some at their offices in a controlled, laboratory-like environment. The experiment took about 50 min to complete and participants received EUR 11.92 on average.

Tatarnikova et al. (2023). Professionals were recruited for a lab in the field study in Paris, Marseille, and Montpellier. About half worked in bank branches while the other half works in asset management. Students were recruited at the University of Montpellier and took part online. The experiment contained multiple parts. The investment task relevant for this article took about 20 minutes to complete. Professionals received an average of EUR 94 while student participants received an average of EUR 15 for the investment task.

B.2 Forecasting

Table C.2: Articles on Forecasting

Article	Environment	Duration	Incentives	Average payments
Muradoğlu & Önkal (1994)	take-home	2.5 days	none	-
Önkal & Muradoğlu (1996)	take-home	2.5 days	none	-
Muradoğlu (2002)	take-home	2.5 days	none	-
Törngren & Montgomery (2004)	take-home	30 days	-	-
Glaser et al. (2007)	online	60 min	Students: fixed Professionals: none	-
Kaustia et al. (2008)	controlled	15-20 min	none	-
Zaleskiewicz (2011)	online	$\leq$ 1 day	none	-
Huber et al. (2019)	online	16 min	proper	USA: USD 24.87 UK: GBP 19.27 1 out of 4 tasks paid
Schwaiger et al. (2020)	controlled	10 min	fixed	Students: EUR 6 Professionals: EUR 18
Barron et al. (2021)	online / take-home	87 min	none	-
Bao et al. (2022)	online	3 × 15-30 min, 1 × 3-4 days	proper	Students: JPY 915 Professionals: JPY 4,877 (Tasks 1-3 out of 4)
Ba et al. (2023)	online	3 + 8 min	fixed	USD 1.56

Muradoğlu & Önkal (1994). The professionals sample consists of 7 licensed brokers and portfolio managers from Istanbul, Turkey, who are managing investment funds and give financial advice to clients. The second sample can be described as a sample of semi-professionals. These are 10 bank employees who were recently trained in portfolio management in Ankara, Turkey. Participants could take the study materials home and were asked to return them within 2.5 days. Participants were not paid for their participation.

Önkal & Muradoğlu (1996). This article uses a setting that is very similar to Muradoğlu & Önkal (1994). The professionals were 13 licensed brokers and portfolio managers from Istanbul, Turkey. The second sample consists of 9 bank employees that were recently trained in portfolio management. A third sample consisted of 64 university students from the Faculty of Business Administration of Bilkent University, Turkey. Participants could take the study materials home and complete them within 2.5 days. They did not receive any payments.

Muradoğlu (2002). This is a third paper using the familiar setting of Önkal & Muradoğlu (1996) and Muradoğlu & Önkal (1994). Professionals are 35 brokers, fund managers, analysts, and financial advisors from Istanbul, Turkey. The participants had between 8 months and 6 years of work experience and were participating in a 20 hour training program on portfolio management and financial forecasting. The student sample comprises 45 undergraduate and graduate students from the Faculty of Business Administration of Bilkent University, Turkey. The students had at least one finance course and were exposed to concepts like the efficient market hypothesis and methods of financial forecasting. Once again, participants could take the study materials home and were expected to return them after 2.5 days. No payments were made.

Törngren & Montgomery (2004). Financial professionals are described as stock market professionals such as portfolio managers, analysts, brokers, and investment counselors. The professionals had on average 12 years of experience. The student sample was recruited from undergraduate students in psychology at Stockholm University, Sweden. The article reports on two studies and highlights that a large overlap in the professional participants between the two studies is likely. There were 33 financial professionals and 29 students in study 1. In study 2, there were 21 financial professionals and 34 students. Participants received the study materials and had to return them after 30 days. No monetary compensation is reported.

Glaser et al. (2007). The professionals are 31 employees from a large bank in Germany. They had 5 years of experience on average and primarily worked in fields such as derivatives, proprietary trading, and market making. The student sample comprised 64 advanced students specializing in banking and finance at Mannheim University, Germany. The experiment was conducted online and took about 60 minutes to complete. Professionals did not receive any payments, while students received fixed payments.

Kaustia et al. (2008). Professionals are 300 financial advisers, institutional investors, asset managers, analysts, investment experts, brokers, wealth managers, stock specialists and administrative staff from Finland and Sweden. They were recruited at field seminars on financial markets and professional education sessions. 213 undergraduate finance students from Helsinki School of Economics, Finland, serve as the control group. The experiments were conducted in controlled, laboratory-like environments and took about 20 minutes to complete. No compensation was paid.

Zaleskiewicz (2011). Professional participants were 38 financial analysts from Poland, who worked for banks and mutual funds and had a mean work experience of 7 years. As part of their job, they were forecasting changes in the economics system. The comparison group are 43 members of the Polish general population without any specific knowledge or experience in the stock market. Participants were contacted personally or by email on the day of the study and asked to submit their forecasts. No information is given on the study materials, the duration of the forecasting task, or any monetary compensation for participation.

Huber et al. (2019). The experiment was conducted in the United Kingdom as well as the United States of America. For each country, a separate sample of financial professionals and separate sample from the general population was recruited. In the UK, 100 financial professionals and 607 members of the general population participated. In the USA, the experiments were conducted with 269 financial professionals and 617 laypeople. Recruitment was done by a large globally operating market research company. No further information is given about the job descriptions of the financial professionals. The experiments were conducted online and took about 16 minutes to complete on average. Participants received performance incentives with 20% being selected for actual payments. The average payment in the USA (UK) was USD 24.87 (GBP 19.27 / USD 25.44).

Schwaiger et al. (2020). Professionals are 150 individuals mainly working in financial advice, fund management, as well as investment and portfolio management. They were recruited from various financial institutions in northern and central Europe. The professionals had on average 13.2 years of experience. The student sample consists of 576 students of various disciplines from the University of Innsbruck, Austria, and was approximately gender matched to the professional sample (77% male). Payments were fixed at EUR 18 for professionals and EUR 6 for students. The experiment took about 10 minutes to complete. While students participated at the campus laboratory, professionals took part in a controlled lab-in-the-field environment.

Barron et al. (2021). The professional sample included 69 professional investors from various financial institutions. The sample includes financial analysts, brokers, investment advisors, fund managers, and portfolio managers among others. They were recruited via personal contacts, referrals, and on the professional social network LinkedIn. The comparison group are 121 non-professional investors who are members of the American Association of Individual Investors. Similar to earlier studies with take-home materials, participants were emailed the study documents and asked to return them later. It took 87 minutes on average to complete the tasks. While it is not explicitly stated in the article, it seems that the participants volunteered and did not receive any payments for their participation.

Bao et al. (2022). Professionals are 212 CFAs (93.4% male) who are certified members of the Securities Analysts Association of Japan (SAAJ) and were recruited via SAAJ. The comparison sample includes 228 students (53.5% male) from Osaka University. Participants were recruited by email. Tasks 1-3 each took between 15 and 30 minutes. For Task 4, participants had 3 or 4 days to submit their forecasts. Payoffs depended on forecasting accuracy and were paid as Amazon gift cards, whereby professionals received five times students' incentives. Average payments for professionals (students) were JPY 1,362 (JPY 316), JPY 1,675 (JPY 284), and JPY 1,840 (JPY 315), in Tasks 1, 2, and 3. In Task 4 the most accurate professional (student) forecaster received JPY 5,000 (JPY 1,000).

Ba et al. (2023). The study contains a survey and an online experiment. The experimental evidence is based on the responses of 467 finance professionals (20% female) and 2346 non-experts (44% female). Professionals were recruited by emailing 300,000 employees from the finance sector while non-experts were recruited on Prolific. Most participants were born in the USA. For non-experts the study was split into to parts of 3 and 8 minutes. Professionals completed both parts back to back. Participants received USD 1.56 for a median completion time of 10 minutes and 26 seconds.

B.3 Asset Markets

Table C.3: Articles on asset markets

Article	Environment	Duration	Incentives	Average payments
DeJong et al. (1988)	controlled	Students: 180 min Professionals: 120 min	proper	Students: USD 10-25 Professionals: Prize or nothing
King et al. (1993)	controlled	90 - 120 min	proper	Students: USD 13 Professionals: ca. USD 21 (+60%)
Sarin & Weber (1993)	controlled	120 min	proper	Students: DEM 11-38 Professionals: DEM 46-64
Anderson & Sunder (1995)	controlled	180 min	proper	USD 6-65
Cipriani et al. (2020)	controlled	120 min	proper	Students: GBP 23.35 Professionals: GBP 234.93
Weitzel et al. (2020)	controlled	70-75 min	proper	Students: EUR 17-19 Professionals: EUR 70-75

DeJong et al. (1988). The professional sample consisted of 5 partners in public accounting and auditing firms as well as 2 corporate financial officers. All professionals had at least 15 years of experience. Student participants were recruited from the College of Business at the University of Iowa, USA. Students received between USD 10 and USD 25 for their participation. Professionals received a university souvenir if they earned more on average per round than a matched student participant. They did not receive anything if they did not earn more. According to the authors, paying professionals in cash would have been prohibitively expensive and receiving tangible evidence of having beaten the student (the souvenir) was believed to be a suitable alternative.

King et al. (1993). "Six over-the-counter traders familiar with computerized stock quotation systems" participated in "Experiment 293; 6, 3i" (p. 196). The most comparable experiment had 6 student participants with one round of experience and 3 informed student participants. Student participants were recruited at the University of Arizona in Tucson (AZ), Indiana University in Bloomington (IN), and Washington University in St. Louis (MO), USA. Students received between USD 3 and USD 34 (average USD 13), while professionals received about USD 21. Decisions were properly incentivized.

Sarin & Weber (1993). "[W]e created markets using eight executives of J. P. Morgan in Frankfurt, who were bond or currency traders or advisors and had a minimum of two years of work experience." (Experiments 9 and 10; p. 604). In addition, there where twelve markets with eight student subjects each, recruited at Aachen University or Cologne University (Experiments 1-8 and 11-14). All experimental sessions lasted around two hours; students earned between DEM 11 and DEM 38, professionals between DEM 46 and DEM 64 (at the time of the experiments the exchange rate was approximately USD 1 to DEM 2). Decisions were properly incentivized.

Anderson & Sunder (1995). The 21 professionals had about 5 years of experience working at stock and bond underwriting houses and the Minneapolis Commodity Exchange, USA. They took part in two markets with 12 and 9 traders, respectively. The student sample comprised MBA students trained in finance, statistical methods, and risk analysis from two state universities. They took part in 3 markets with 12, 11, and 8 traders, respectively. Experiments took about 180 minutes to conduct in a controlled, laboratory-like setting. Payments ranged from USD 6 to USD 65.

Cipriani et al. (2020). A total of 56 traders and portfolio managers from London (UK) who were working in a variety of different markets (equity, equity derivatives, foreign exchange, fixed income, commodities, etc.) and had an average tenure of 9.25 years took part in the experiment. The comparison sample of 56 undergraduate students was recruited at Central London University, UK. The student sample had approximately the same gender composition (79% male) as the professional sample (86% male). Experimental sessions took about 120 minutes to conduct and participants received performance-based pay. Professionals received GBP 2.50 per 100 experimental currency units, while students received GBP 0.25 per 100 units. Average task earnings were GBP 234.93 (USD 306) for professionals and GBP 23.35 (USD 30.45) for students. The experiment was conducted in the laboratory.

Weitzel et al. (2020). The paper reports on two sets of treatments. For the first set, the professional sample consisted of 294 financial professionals from central and northern European countries working in private banking, trading, investment banking, portfolio management, fund management, and wealth management. For the second set, it consisted of 118 professionals (avg. 9 year tenure) from major financial institutions in Austria and the Netherlands. The student samples both consisted of students from the University of Innsbruck (Austria) and Radboud University Nijmegen (the Netherlands). A total of 384 students participated in the first set and 118 additional students participated in the second set of treatments. The main sessions took place in laboratory-like settings. Sessions in the first set took about 70 minutes to complete and paid on average EUR 76.5 (EUR 18.6) to professionals (students). The second set was slightly longer at approximately 75 minutes. Payments were 71.3 EUR (EUR 17.5) on average for professionals (students).

B.4 Further Results on Financial Professionals' Behavior

Table C.4: Articles on individual characteristics, culture, and context

Article	Environment	Duration	Incentives	Average payments
Frederick & Libby (1986)	controlled	5-10 min	none	-
Abbink & Rockenbach (2006)	controlled	Students: 60-120 min Professionals: 60 min	proper	-
Alevy et al. (2007)	controlled	30 min	proper	see below.
Noll et al. (2012)	controlled, individual meetings	-	none	-
Cohn et al. (2014)	online	15 min	proper	Students: USD 50 Professionals: USD 200 Laypeople: USD 200
Cohn et al. (2017)	online	26 min	proper	up to USD 500
Lindner et al. (2021)	controlled	45 min	proper	Students: EUR 17 Professionals: EUR 48
Rahwan et al. (2019)	online	10 min	proper	Asia pacific: USD 14/coin toss max: USD 140
Huber & Huber (2020)	online	9 min	proper	Students: EUR 4.66 Professionals: EUR 8.16
Holmen et al. (2023)	online	15 min	proper	SEK 211.13
Duchêne et al. (2022)	controlled	45 min	proper	Students: EUR 13.45 Professionals: EUR 216.81

Frederick & Libby (1986). Five experiments are reported in the article. Experiments 1 and 2 were conducted with professionals. Experiments 3 to 5 with students. The professionals were auditors from one of the largest CPA firms with 2.5-3.5 years of experience, who were attending a two-week training program. Students were undergraduates taking auditing classes and MBA students of advanced accounting. Experiments 1 and 2 were conducted with 33 and 31 professionals, respectively. Experiments 3, 4, and 5 were conducted with 49, 40, and 24 student participants, respectively. Participants did not receive any payments.

Abbink & Rockenbach (2006). There were three samples of participants in this study. The professionals were 24 bank employees from Frankfurt, Germany, who mainly worked in foreign exchange, security, futures, bonds, and money trade. All reported to be decision-makers in their fields. The first student sample consisted of 108 students from Bonn University, Germany. They were mostly studying economics and law. The authors emphasize that their education is highly technical with many theory-oriented courses and a strong focus on mathematics. While option pricing is part of the curriculum, they did not have any prior experience with financial market experiments. The third sample were students with a mainly non-technical, social-science majors from the University of Erfurt, Germany. This group did not receive any formal training in option pricing as part of their curriculum. All participants took part in controlled, laboratory-like environments. Due to time limitations, sessions with professionals were shortened from 50 to 30 decision rounds. The exchange rate for experimental currency units was adjusted to yield payments comparable to the student treatments.

Alevy et al. (2007). Financial professionals are 55 "market professionals" from the Chicago Board of Trade (CBOT), USA. The students were undergraduates from the University of Maryland (College Park), USA. The experiments were conducted in controlled, lab-like environments at the CBOT (professionals) and on campus (students). Students started the experiment with an endowment of USD 6.25 while professionals received an endowment of USD 25. As losses could be incurred in the experiment, additional games were played in each session to ensure positive balances of all subjects at the end of the experiment. Average earnings are given separately by sample, urn type (asymmetric or symmetric), and gain / loss domain in their Table II. Payments to professionals were approximately 4x those of students.

Noll et al. (2012). The first sample consists of 28 professional bank traders (equities, commodities, etc.). One half worked for large international banks, the other half worked for medium-sized banks. No location is given. The second sample are 24 individuals diagnosed with moderate to severe levels of psychopathic personality disorder. These were recruited in German high security psychiatric hospitals. The third sample are 24 non-academic men from the general population from Regensburg, Germany. Notably, all participants took part in individual sessions and played against a computer opponent programmed to play a tit-for-two-tats strategy. Session lengths are not given in the article. Participants did not receive any payments.

Cohn et al. (2014). There are three samples in this study. The financial professionals are 128 employees from a large international bank with about 11.5 years of experience on average. About half of the financial professionals worked in "core business units, i.e., as private bankers, asset managers, traders, or investment managers" (supplementary material, p. 2), while the other half worked in supporting roles in risk management and human resources management. All participants are described as bank. A location is not revealed. The second sample are 222 students from an undisclosed university. The third sample are 133 members of the working population with 14.8 years of experience in their respective fields on average. These people were employed in the middle or upper management of manufacturing, pharmaceutics, telecommunications, and information technology companies. Participants took part in an online experiment that took approximately 15 minutes to complete. Financial professionals and members of the general working population received gift cards of up to USD 200 in value and 20% of the professional participants were paid. Students received up to USD 50 ("reduced the stake size by a factor of four.", supplementary material, p. 7).

Cohn et al. (2017). This paper apparently uses the same sample of financial professionals and the same sample of members of the general working population as Cohn et al. (2014), which becomes apparent from the identical sample sizes and summary statistics. In addition, it includes a sample of 142 banking employees from many smaller and larger banks. These financial professionals predominantly worked in asset management, private banking, and trading and investment banking. Work experience was relatively high with 25 years on average. All sessions were conducted online and took about 26 minutes to complete. Participants were endowed with USD 200 and could earn up to USD 500 in the tasks. About every fifth participant was paid.

Lindner et al. (2021). The professional sample consists of 330 employees from major financial institutions from several OECD countries. On average, the professionals reported 12.6 years of experience and worked in private banking, trading, investment banking, portfolio management, fund management, and wealth management. The student sample was recruited at the University of Innsbruck, Austria, and consisted of 864 bachelor and master students. The sessions took place in controlled, lab-like environments and took about 45 minutes to complete. Participants received performance-based payments with professionals (students) earning on average EUR 48 (EUR 17). Stakes for the professionals were three times the stakes of students.

Rahwan et al. (2019). The article reports on several samples and multiple studies. We focus on the samples most relevant to the comparison of financial professionals an non-professionals. First, there are 620 bankers from a "large bank in the Asia Pacific region" (p. 346). From the same region, they also collect a sample of 242 non-banking employees, aiming to be "nationally representative for gender and age" (p. 346). Then, there are 148 bankers from the a "medium-sized bank in the Middle East", as well as 67 "regulators of financial services" (p. 346). Participants could earn USD 14 in local currency for each of 10 coin tosses, resulting in a maximum pay of approximately USD 140. A lottery mechanism was used to pay about 10% of the participants. Participants received shopping vouchers. The non-banking participants did not receive monetary payments, but charitable donations were made instead.

Huber & Huber (2020). A total of 223 financial professionals participated in the experiment. Of these, 115 were recruited on before.world, while the remaining 108 participants were recruited on Prolific. Participants from the before.world pool worked mainly as portfolio managers, fund managers, investment managers, traders, analysts, consultants, and financial advisors. Selection on Prolific was based on participants reporting to work in the finance and insurance industry. The student sample consisted of 166 students from the University of Innsbruck, Austria. All participants took part online and completed the experiment in 9 minutes on average. Professionals from before.world earned 8.16 EUR on average, while professionals on Prolific received EUR 3.64 (paid in GBP) on average. Students earned EUR 4.66 on average. Payments for participants on Prolific as well as students were reduced by half compared to the financial professionals in the before.world sample.

Holmen et al. (2023). The first sample are 298 financial analysts, advisors, traders, fund managers, and financial brokers from Sweden. The second sample are 395 members of the Swedish general working population (excluding financial professionals). The experiment was conducted online and could be completed in 15 minutes. Participants received a participation fee of 100 SEK (EUR 10) and earned on average SEK 211.13 (EUR 23.50).

Duchêne et al. (2022). The professional sample consisted of 190 financial professionals from major financial institutions (investment banks and asset management companies) of Morocco. The experiment was conducted on-site in Casablanca using a mobile laboratory setup. The professional sample mainly included, among others, proprietary traders, sales traders, asset managers, trading room managers, quantitative engineers, structurers, and financial analysts. In addition, 279 students from the University of Montpellier, France, took part in a standard laboratory setting. Professionals received 1 EUR per experimental currency unit, while students received 0.04 EUR per currency unit. While only 10% of professionals were paid, all students received a payment. On average over those selected, professionals earned EUR 216.81. Students were paid EUR 13.45 (EUR 8.10 excluding show-up) on average.

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

https://www.uibk.ac.at/eeecon/wopec/

2022-07 **Christoph Huber, Christian König-Kersting, Matteo M. Marini:** Experimenting with Financial Professionals

2022-06 **Martin Gächter, Martin Geiger, Elias Hasler:** On the structural determinants of growth-at-risk

2022-05 **Katharina Momsen, Sebastian O. Schneider:** Motivated Reasoning, Information Avoidance, and Default Bias

2022-04 **Silvia Angerer, Daniela Glätzle-Rützler, Philipp Lergetporer, Thomas Rittmannsberger:** How does the vaccine approval procedure affect COVID-19 vaccination intentions?

2022-03 **Robert Böhm, Cornelia Betsch, Yana Litovsky, Philipp Sprengholz, Noel Brewer, Gretchen Chapman, Julie Leask, George Loewenstein, Martha Scherzer, Cass R. Sunstein, Michael Kirchler:** Crowdsourcing interventions to promote uptake of COVID-19 booster vaccines

2022-02 **Matthias Stefan, Martin Holmén, Felix Holzmeister, Michael Kirchler, Erik Wengström:** You can't always get what you want–An experiment on finance professionals' decisions for others

2022-01 **Toman Barsbai, Andreas Steinmayr, Christoph Winter:** Immigrating into a Recession: Evidence from Family Migrants to the U.S.

2021-32 **Fanny Dellinger:** Housing Support Policies and Refugees' Labor Market Integration in Austria

2021-31 **Albert J. Menkveld, Anna Dreber, Felix Holzmeister, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Sebastian Neusüss, Michael Razen, Utz Weitzel and et al:** Non-Standard Errors

2021-30 **Toman Barsbai, Victoria Licuanan, Andreas Steinmayr, Erwin Tiongson, Dean Yang:** Information and Immigrant Settlement

2021-29 **Natalie Struwe, Esther Blanco, James M. Walker:** Competition Among Public Good Providers for Donor Rewards

2021-28 **Stjepan Srhoj, Melko Dragojević:** Public procurement and supplier job creation: Insights from auctions

2021-27 **Rudolf Kerschbamer, Regine Oexl:** The effect of random shocks on reciprocal behavior in dynamic principal-agent settings

2021-26 **Glenn E. Dutcher, Regine Oexl, Dmitry Ryvkin, Tim Salmon:** Competitive versus cooperative incentives in team production with heterogeneous agents

2021-25 **Anita Gantner, Regine Oexl:** Respecting Entitlements in Legislative Bargaining - A Matter of Preference or Necessity?

2021-24 **Silvia Angerer, E. Glenn Dutcher, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** The formation of risk preferences throughsmall-scale events

2021-23 **Stjepan Srhoj, Dejan Kovač, Jacob N. Shapiro, Randall K. Filer:** The Impact of Delay: Evidence from Formal Out-of-Court Restructuring

2021-22 **Octavio Fernández-Amador, Joseph F. Francois, Doris A. Oberdabernig, Patrick Tomberger:** Energy footprints and the international trade network: A new dataset. Is the European Union doing it better?

2021-21 **Felix Holzmeister, Jürgen Huber, Michael Kirchler, Rene Schwaiger:** Nudging Debtors to Pay Their Debt: Two Randomized Controlled Trials

2021-20 **Daniel Müller, Elisabeth Gsottbauer:** Why Do People Demand Rent Control?

2021-19 **Alexandra Baier, Loukas Balafoutas, Tarek Jaber-Lopez:** Ostracism and Theft in Heterogeneous Groups

2021-18 **Zvonimir Bašić, Parampreet C. Bindra, Daniela Glätzle-Rützler, Angelo Romano, Matthias Sutter, Claudia Zoller:** The roots of cooperation

2021-17 **Silvia Angerer, Jana Bolvashenkova, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Children's patience and school-track choices several years later: Linking experimental and field data

2021-16 **Daniel Gründler, Eric Mayer, Johann Scharler:** Monetary Policy Announcements, Information Schocks, and Exchange Rate Dynamics

2021-15 **Sebastian Bachler, Felix Holzmeister, Michael Razen, Matthias Stefan:** The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

2021-14 **Jeppe Christoffersen, Felix Holzmeister, Thomas Plenborg:** What is Risk to Managers?

2021-13 **Silvia Angerer, Daniela Glätzle-Rützler, Christian Waibel:** Trust in health care credence goods: Experimental evidence on framing andsubject pool effects

2021-12 **Rene Schwaiger, Laura Hueber:** Do MTurkers Exhibit Myopic Loss Aversion?

2021-11 **Felix Holzmeister, Christoph Huber, Stefan Palan:** A Critical Perspective on the Conceptualization of Risk in Behavioral and Experimental Finance

2021-10 **Michael Razen, Alexander Kupfer:** Can increased tax transparency curb corporate tax avoidance?

2021-09 **Changxia Ke, Florian Morath, Anthony Newell, Lionel Page:** Too big to prevail: The paradox of power in coalition formation

2021-08 **Marco Haan, Pim Heijnen, Martin Obradovits:** Competition with List Prices

2021-07 **Martin Dufwenberg, Olof Johansson-Stenman, Michael Kirchler, Florian Lindner, Rene Schwaiger:** Mean Markets or Kind Commerce?

2021-06 **Christoph Huber, Jürgen Huber, and Michael Kirchler:** Volatility Shocks and Investment Behavior

2021-05 **Max Breitenlechner, Georgios Georgiadis, Ben Schumann:** What goes around comes around: How large are spillbacks from US monetary policy?

2021-04 **Utz Weitzel, Michael Kirchler:** The Banker's Oath And Financial Advice

2021-03 **Martin Holmen, Felix Holzmeister, Michael Kirchler, Matthias Stefan, Erik Wengström:** Economic Preferences and Personality Traits Among Finance Professionals and the General Population

2021-02 **Christian König-Kersting:** On the Robustness of Social Norm Elicitation

2021-01 **Laura Hueber, Rene Schwaiger:** Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

University of Innsbruck

Working Papers in Economics and Statistics

2022-07

Christoph Huber, Christian König-Kersting, Matteo M. Marini

Experimenting with Financial Professionals

Abstract
As key players in financial markets and the broader industry, financial professionals are increasingly used as experimental research participants. We review over 50 studies comparing financial professionals to laypeople and conduct systematic meta-analyses of 24 eligible studies spanning from 1986 to 2023. Our findings reveal persistent and robust support for financial professionals being more risk- and uncertainty-loving, but little evidence of superior forecasting accuracy. Further analyses indicate that larger monetary payments result in greater behavioral differences between financial professionals and laypeople, suggesting an increased susceptibility to incentives among professionals. This systematic review not only synthesizes experimental results, contributing to recent discussions about external validity and generalizability, but also highlights critical methodological considerations when experimenting with financial professionals.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)