Discussion Papers of the Max Planck Institute for Research on Collective Goods Bonn 2018/9

Bubbles and Financial Professionals

Utz Weitzel
Christoph Huber
Jürgen Huber
Michael Kirchler
Florian Lindner
Julia Rose

MAX PLANCK SOCIETY

Discussion Papers of the
Max Planck Institute
for Research on Collective Goods

Bonn 2018/9

Bubbles and Financial Professionals

Utz Weitzel / Christoph Huber / Jürgen Huber / Michael Kirchler / Florian Lindner / Julia Rose

July 2018

Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn
http://www.coll.mpg.de

Bubbles and Financial Professionals

Utz Weitzel[*a,b] Christoph Huber[c] Jürgen Huber[c] Michael Kirchler[c,d]
Florian Lindner[e] Julia Rose[c]

June 18, 2018

Abstract

The efficiency of financial markets and their potential to produce bubbles are central topics in academic and professional debates. Yet, surprisingly little is known about the contribution of financial professionals to price efficiency. To close this gap, we run 86 experimental markets with 294 professionals and 384 students. We report that professional markets with bubble-drivers—capital inflows or high initial capital supply—are susceptible to bubbles, but they are significantly more efficient than student markets. In a survey with 245 professionals and students we show that cognitive skills and risk attitudes do not explain subject pool differences in bubble formation.

JEL: C92, D84, G02, G14

Keywords: Experimental finance, financial professionals, price efficiency, financial bubbles.

We thank Peter Bossaerts, Markus Brunnermeier, Alain Cohn, Brice Corgnet, Christoph Engel, Sascha Füllbrunn, Hanjo Hamann, Daniel Kleinlercher, Jörg Oechssler, Stefan Palan, Charles Plott, David Porter, David Schindler, Christoph Siemroth, Wei Xiong, Stefan Zeisberger, seminar participants at Chapman University, the Helsinki Center of Economic Research (HECER), and the Universities of Innsbruck and Tilburg, as well as conference participants of ESA 2017 in Vienna, Experimental Finance 2017 in Nice, the EEG Inaugural Conference at the Max Planck Institute for Research on Collective Goods in Bonn 2018, the Behavioral Finance Working Group Conference, Queen Mary University of London 2018, and workshop participants of the Austrian Working Group on Banking and Finance 2017 and the Dutch Authority for Financial Markets (AFM) for very valuable comments. We are grateful to Parampreet Christopher Bindra, Laura Hueber, and Julia Scheunert for excellent research assistance. We particularly thank all financial institutions and participating professionals for the excellent collaboration and their enthusiasm. Financial support from the Austrian Science Fund FWF (START-grant Y617-G11 and SFB F63), Radboud University, and the Swedish Research Council (grant 2015-01713) is gratefully acknowledged. This study was ethically approved by the IRB of the University of Innsbruck. A prior version of this paper was presented with the title "Market Experience and Price Efficiency – Evidence from Experiments with Financial Professionals".

[*]Corresponding author, Email: u.weitzel@uu.nl.

[a]Utrecht University, Utrecht School of Economics, Kriekenpitplein 21-22, 3584 EC Utrecht.

[b]Radboud University, Institute for Management Research, Thomas van Aquinostraat 5.1.26, 6525 Nijmegen.

[c]University of Innsbruck, Department of Banking & and Finance, Universitätsstrasse 15, 6020 Innsbruck.

[d]University of Gothenburg, Department of Economics, Centre for Finance, Vasagatan 1, 40530 Gothenburg.

[e]Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, 53113 Bonn.

Financial bubbles and crashes have been recurring phenomena in economic history. Following Galbraith (1994), Kindleberger and Aliber (2011), Brunnermeier and Oehmke (2013), and Brunnermeier and Schnabel (2016), bubbles have been observed in different time periods (dating back to the 17th century), in economies at different stages of development (from developing economies to highly industrialized economies in the 21st century), and across a wide range of asset classes (e.g., real estate markets, asset markets, and derivative markets; Xiong and Yu, 2011). Bubbles, crashes, and their underlying mechanics are of high interest to economists for at least two major reasons. First, they represent periods of inefficient prices, i.e., of prices that strongly deviate from fundamentals. Second, bubbles and their subsequent crashes have the power to severely affect the real economy through misallocation of resources and impaired balance sheets.

For many, bubbles and crashes are typical manifestations of inefficiencies in financial markets. Yet, whether and to which degree financial markets are efficient is still one of the most controversially debated questions in economics.[1] Despite considerable literature on mispricing in financial markets, empirical evidence remains elusive (see, e.g., Gürkaynak, 2008, for an overview). As fundamental values are usually not observable in data from financial markets, the empirical identification of bubbles and price inefficiencies often suffers from the joint hypothesis problem (Fama, 1970): tests of market efficiency are simultaneously also tests of an underlying equilibrium model that may be misspecified. One way to tackle this problem is to use experimental asset markets as test-beds, where fundamental values are defined and price deviations can be measured in a controlled setting (e.g., Bloomfield and Anderson, 2010). However, the experimental literature almost exclusively investigates the behavior of students, not of the main protagonists in financial markets: financial professionals. This is a potentially important limitation as professionals' and students' behavior *can* differ substantially. For example, financial professionals are found to be less prone to anchoring than students (Kaustia et al., 2008), to exhibit a higher degree of myopic loss aversion (Haigh and List, 2005), to better discern the quality of public signals in information cascades (Alevy et al., 2007), to more accurately assess others' risk preferences (Roth and Voskort, 2014), and to take more risk in competitive situations involving rankings (Kirchler et al., 2018). Hence, research on mispricing and bubbles faces a fundamental dilemma between internal and external validity: studies with data from financial professionals' behavior in real markets are externally more valid, but have a limited internal

[1] The scientific oeuvre of two Nobel laureates from 2013, Eugene Fama and Robert Shiller, highlights the discrepancies on this topic in the scientific profession. When asked about bubbles Fama stated that "*[i]f a bubble is defined as an irrational strong price increase that implies a predictable strong decline, then there's not much evidence that such things exist.*" Shiller by contrast believes that bubbles exist and states "*I define a bubble as a social epidemic that involves extravagant expectations for the future. Today, there is certainly a social and psychological phenomenon of people observing past price increases and thinking that they might keep going.*" See `http://www.newyorker.com/news/john-cassidy/interview-with-eugene-fama` and `http://uk.businessinsider.com/robert-shiller-stock-market-bubble-2015-5?r=US&IR=T`.

validity; experimental asset markets provide more internal validity and causal inference, but mostly rely on student subjects, which limits their external validity.

As a first contribution, this paper tackles the above dilemma by combining the higher internal validity of controlled market experiments with the externally more valid behavior of financial professionals. We readily acknowledge that there are several ways to balance internal and external validity.[2] Moreover, each experiment, whether lab or field, represents a well-defined, controlled situation. Generalizations from experiments with financial professionals to behavior on real financial markets therefore require caution. However, given that the question to which degree markets are efficient is central to the academic and industry-wide debate, it is surprising that no large-scale experimental evidence on professionals' contribution to price efficiency in financial asset markets exists. We recruited 294 financial professionals from high-skilled investment areas—such as trading, fund management, asset liability management, and portfolio management—for lab-in-the-field experiments with 38 financial asset markets. To get a comprehensive picture of professionals' impact on price efficiency, we administered two bubble-driver treatments (capital inflows and high initial capital supply in the market) and two bubble-moderator treatments (short-selling and a low ratio of cash to asset-value in the market), which prior literature has shown to be effective. We then observed subjects' trading behavior and price deviations from the fundamental value. As one of the main results we find that professionals are susceptible to bubble-drivers, such as capital inflow or high initial capital supply in the absence of short-selling. Moreover, in the bubble-driver treatments, we detect bubbles in roughly a quarter of all markets with financial professionals.

With these results we do not only contribute to the ongoing debate on the degree of financial market efficiency, but also to the literature that identifies various forms of capital inflows in financial markets as important bubble-drivers (e.g., Caginalp et al., 1998; Allen and Gale, 2000; Brunnermeier and Schnabel, 2016). We also add to the emerging experimental literature analyzing behavior of financial professionals (e.g., Haigh and List, 2005; Alevy et al., 2007; Cohn et al., 2014, 2017; Kirchler et al., 2018). This research is still in its infancy, but because of financial professionals' prominent role in the economy and the importance of financial markets for the functioning of society, it is crucial to learn more about their preferences and behavior in decision making. Here we add the finding that that even high-skilled financial professionals are not immune to bubble-drivers.

As a second contribution, we examine whether the bubble phenomena is robust to subject pools. For comparison we administered the same treatments to 384 students in 48 lab markets. To keep the student population as comparable as possible to the sample of professionals, we

[2]For example, an econometric method that is less susceptible to the joint hypothesis problem is to test for explosive roots in stock prices, which are difficult to argue for if the underlying dividend or earnings process follows a linear unit-root process (e.g., a random walk). However, despite these and other tests (e.g., variance bounds tests), it is very difficult to econometrically detect asset price bubbles with a satisfactory degree of certainty (Gürkaynak, 2008; Brunnermeier, 2008; Scherbina and Schlusche, 2014).

mainly recruited male students from management and economics. By assigning subjects to markets based on specific characteristics (i.e. being a professional or a student) we follow earlier studies that study price efficiency and bubble formation by composing markets according to student characteristics such as prior market experience (Dufwenberg et al., 2005), gender (Eckel and Füllbrunn, 2015), cognitive sophistication (Bosch-Rosa et al., 2018), or speculative behavior (Janssen et al., 2018). The theory does not discriminate by who is participating. We show that there is a difference: markets with professionals show significantly less overpricing and also fewer and smaller bubbles than markets with students in bubble-driver treatments. In bubble-moderator treatments, however, professionals and students show similar levels of high price efficiency. The good news for experimenters with student subjects is that, qualitatively, bubble-drivers and bubble-moderators have similar effects on financial professionals and students. In others words, the direction and statistical significance of the treatment effects is comparable, though the effect sizes are significantly smaller for professionals. With these results our study also complements the experimental finance literature investigating bubble-drivers and -moderators in classical laboratory experiments with student subjects (e.g., Smith et al., 1988; Lei et al., 2001; Dufwenberg et al., 2005; Kirchler et al., 2012; Sutter et al., 2012).

As a third contribution we explore whether it is possible to identify more specific or alternative explanations for the observed differences between subject pools. For example, it is possible that people with higher cognitive skills preferably select into the finance industry. Alternatively, it is possible that higher cognitive skills are not selected but acquired in the industry. If this is the case, we would expect professionals to exhibit higher cognitive skills than students, which may explain higher price efficiency (Corgnet et al., 2018; Bosch-Rosa et al., 2018). We therefore measured—with an online survey—fluid intelligence, cognitive reflection, theory of mind, backward induction, and risk preferences of both professionals and students. To prevent confounding with the experiments, we administered the survey separately to newly recruited, but similar samples of 121 professionals and 124 students. The results show that deviations in price efficiency, observed in the experiments, cannot be explained by superior cognitive skills of professionals or elevated risk attitudes of students. This suggests that other factors are at play, such as real-world market experience, which may include a richer set of dimensions than specific cognitive skills or risk preferences.

Our study is set up as follows. For the asset markets in the lab-in-the-field experiments (with professionals) and lab experiments (with students) we use the design of Smith et al. (2014) and Holt et al. (2017). We chose this design for two reasons. First, this market design has a number of features (like dividend and interest payments) that, from the perspective of financial professionals, are comparatively close to their experience of real-world markets. Second, although the fundamental value is transparent, the design has been shown to be able to consistently

produce price bubbles (Smith et al., 2014; Holt et al., 2017), which provides room for different bubble-driver and -moderator treatments to take effect.[3]

We administered two classical bubble-driver treatments from the literature by either (i) implementing a high initial cash to asset-value ratio (CA-Ratio), i.e., a high initial level of the monetary supply (cash) relative to the asset value in the market (see Caginalp et al., 1998, 2001; Noussair and Tucker, 2016) or (ii) allowing capital inflows and thereby creating an increasing CA-Ratio over time (Kirchler et al., 2012; Razen et al., 2017). The CA-Ratio is calculated as the total amount of money in the market over the product of shares outstanding and the fundamental value (FV). Following Galbraith (1994) and Kindleberger and Aliber (2011) the expansion phase of many historic bubbles was fueled by various forms of capital inflows. Similarly, Brunnermeier and Schnabel (2016) analyze 23 bubble episodes spanning the last 400 years and conclude that the emergence of bubbles is often preceded or accompanied by expansive monetary policy, high leverage of market participants, lending booms, and capital inflows. The two treatments we apply—high initial CA-Ratio and increasing CA-Ratio over time—capture these features of high or increasing capital supply in a simplified way.

We also administered two treatments implementing classical bubble-moderating factors by either (iii) allowing short sales (Ackert et al., 2006; Haruvy and Noussair, 2006) or (iv) providing a low initial cash to asset-value ratio with no capital inflow over time (Kirchler et al., 2015; Razen et al., 2017). Following the theoretical literature, market frictions like short-sale constraints can lead to bubble formation even in finite horizon models with asymmetric information (see, e.g., Allen and Gorton, 1993; Brunnermeier, 2001, 2009). Short-sale constraints are also a necessary requirement for bubbles to form in heterogeneous beliefs models (Miller, 1977; Harrison and Kreps, 1978). Empirically, Ofek and Richardson (2003) relate the combination of heterogeneous beliefs and short-sale frictions to the formation of the dot-com bubble in the late 1990s. Experimentally, in lab markets with student subjects, all these four bubble-drivers and bubble-moderators have been shown to affect price efficiency.[4]

In summary, we find the following results: *across* subject pools, markets populated by professionals exhibit significantly more efficient prices, are less prone to bubbles, and these bubbles are smaller compared to student markets. This finding holds for the bubble-driver treatments but not for the bubble-moderator treatments which show similar levels of high efficiency across both subject pools. Following the definition of Razen et al. (2017) we find that 25 percent of pro-

[3]The seminal framework of Smith et al. (1988) could have been an alternative, but a number of studies have shown that inefficiencies arise due to the particular design of decreasing fundamental values (Smith et al., 2000; Noussair et al., 2001; Kirchler et al., 2012; Huber and Kirchler, 2012).

[4]For example, Caginalp et al. (1998, 2001), Haruvy and Noussair (2006), and Noussair and Tucker (2016) find that high initial CA-Ratios lead to strong overpricing in markets with declining and constant fundamental values. With respect to monetary inflow over time, Kirchler et al. (2015) show that the inflow of new traders with cash endowments triggers strong and consistent bubbles; and Razen et al. (2017) find that capital inflow to already active traders can fuel bubbles in case trading horizons are long. Concerning the role of short-selling for price efficiency, experimental evidence shows that overpricing is deflated (Ackert et al., 2006) and can even turn negative (Haruvy and Noussair, 2006).

fessional markets and 58 percent of student markets are classified as bubble markets in the two bubble-driver treatments. In the bubble-moderator treatments none of the markets populated by professionals and only four percent of all student markets exhibit bubble patterns.[5] Despite these differences, we also find qualitatively very similar patterns *within* each subject pool. Most importantly, bubble-drivers reduce price efficiency and increase the likelihood of mispricing and bubbles in both groups. In other words, bubble-drivers do not only affect students but also professionals. Bubble-moderators yield efficient markets in both groups—in specific market environments even inexperienced subjects price efficiently. Finally, we probe for potential drivers of the above results in our online survey and find that cognitive skills and risk attitudes cannot serve as explanations for differences in price efficiency between professionals and students. For both groups we find equal levels of cognitive skills and students self-report lower levels of financial risk-taking than professionals, which is at odds with more bubbles in student markets.[6]

The paper is structured as follows. In Section 1 we introduce the experimental design, followed by the results in Section 2. Section 3 discusses and concludes.

1 The Experiment

1.1 The Market

The experimental markets closely mimic the design of Smith et al. (2014) and Holt et al. (2017). Subjects buy and sell assets of a fictitious company for experimental currency (Taler) for a sequence of 20 periods of 120 seconds each. Asset and Taler holdings are carried over each period. We implement a continuous double auction protocol where each market was set up for 8 traders.

In our baseline treatment, INC (for increasing CA-Ratio), each subject is initially endowed with 560 Taler in cash and 20 shares. Similar to Smith et al. (2014) and Holt et al. (2017) dividends of either 1.2 or 1.6 Taler are paid with equal probability at the end of each period. Additionally, interest of 5% is paid on cash holdings at the end of a period but before dividends are added. The publicly known redemption value for each stock at the end of Period 20 is 28 Taler. The expected dividend return is equal to the interest rate on cash at 5% (1.4 divided by 28) and therefore the asset's risk-neutral fundamental value (FV) is constant at 28 in all

[5] We elicit subjects' price beliefs and report differences in professionals' and students' price forecast errors. We find that, although forecast errors are similar across pools in price upswings, professionals predict prices significantly better than students in price downswings in the bubble-driver treatments.

[6] We would like to stress the importance of nomenclature when referring to the term "bubble". The definition of bubbles is still controversially discussed (e.g., Brunnermeier, 2009; Engsted, 2016), which is why we also refer to overpricing and other terms/variables measuring price inefficiencies in a more precise way (e.g., price amplitude, maximum overpricing). When we refer to bubble markets in our study, we follow the definition of Razen et al. (2017), which we apply in Section 2.2. Moreover, whenever we refer to other studies on this topic, as in this introduction, we follow the nomenclature of the original authors and refer to market inefficiencies as bubbles when they do.

periods. Moreover, an income of 100 Taler from an exogenous source is paid to each subject at the beginning of each period. Because of these model characteristics put forward by Smith et al. (2014) and Holt et al. (2017) the CA-Ratio in the market (i.e., total cash divided by the product of numbers of shares outstanding and FV of 28) is increasing from 1.0 to 10.2 from the beginning of Period 1 to the end of Period 20. Shorting assets and borrowing money are not allowed. All this information is public knowledge. With these market characteristics we test the prominent bubble-driver capital inflow (Brunnermeier and Schnabel, 2016; Holt et al., 2017; Razen et al., 2017).

Treatment SHORT is identical to Treatment INC except for the possibility to short up to 40 shares (i.e., asset holdings can fall to a value of -40).[7] With this bubble-moderator treatment we can analyze whether potential overpricing induced by the cash inflow in Treatment INC can be mitigated when allowing for short-selling.

The second bubble-moderator treatment (LOW) is also identical to Treatment INC, except that we keep the CA-Ratio constant at 1.0 in all 20 periods. We transfer the exogenous period-income of 100 Taler as well as the dividend and interest payments to a separate account ("Account B"). This account is not available for trading but the holdings are added to final wealth and thus converted to euro at the end of the experiment. This procedure ensures the absence of capital inflow in the market with all other model features being identical to Treatment INC.

Finally, the second bubble-driver treatment HIGH is identical to Treatment LOW except for the level of the CA-Ratio. Here, we implement a constant CA-Ratio of 10.2 (i.e., the final level in treatments INC and SHORT) by increasing the initial cash endowment of each subject to 5,700 Taler. Dividends, interest, and income are again transferred to Account B which cannot be used for trading. With this treatment we can test the role of the most prominent bubble-driver in the experimental finance literature, i.e., a high initial monetary base in the market (e.g., Caginalp et al., 1998, 2001; Noussair and Tucker, 2016).

We apply a typical continuous double-auction trading protocol, which is standard in the literature (see the Appendix for a screenshot and a detailed explanation of the trading screen). All orders are executed according to price and then time priority in an open order-book framework. Market orders have priority over limit orders and are always executed instantaneously. When posting limit orders, traders specify the price and quantity they want to trade for. When posting market orders traders only specify the quantity they want to trade and the order is executed immediately at the price of the currently best limit order. Any order size, the partial execution of limit orders, and deleting already posted limit orders are possible.

As in Haruvy et al. (2007), Kirchler et al. (2015), and Razen et al. (2017), we elicit subjects' beliefs about future market prices in each period. Specifically, at the beginning of each period

[7] Subjects with outstanding assets have to pay the respective dividends in each period and the buyback price of 28 Taler for each outstanding unit at the end of the experiment. We do not impose additional cash reserve requirements (Haruvy and Noussair, 2006).

t, subjects are asked to predict average period prices for the three upcoming periods. $\widetilde{\mathrm{P}}^{i}_{t,t+k}$ indicates subject i's beliefs in period t of each average period price from t to $t+k$ with k indicating values in the range of $\{0, 1, 2\}$. Following Holt et al. (2017) payout depends on prediction accuracy. If a prediction lies within a range of +/− 5% of the average market price in the corresponding period, 50 Taler (175 Taler in Treatment HIGH) are added to the cash holdings at the end of the experiment in all treatments. Moreover, subjects receive feedback on their forecast accuracy only after the final period.[8]

1.2 Implementation of the Experiment

In total we conducted 38 markets with the professional sample in Experiment PROF. We ran sessions with 16 to 35 professionals, which resulted in 2 to 4 markets per session. We randomized subjects into as many treatments per session as possible, administering 2 to 4 treatments within a session simultaneously. The planned size of markets was 8 traders and in many cases we managed to keep this market size. However, when running experiments with professionals, some subjects participate or cancel on short notice, because of constant shifts in their schedules. Moreover, it was very difficult to deny access to the experimental market when a 9th subject arrived unexpectedly (often despite prior cancellation), or when one market participant did not show up in time, endangering the participation of the other 7 market participants. We thus also ran some markets with 7 or 9 traders.

We recruited 294 professionals from major financial institutions in several OECD-countries who were regularly confronted with investment and trading decisions in their daily work—i.e., professionals from private banking, trading, investment banking, portfolio management, fund management, and wealth management—in Experiment PROF.[9] 89.5 percent were male, the average age was 35.6 years, and they had been working in the finance industry for 10.4 years on average. We applied the same recruitment and implementation strategy of the lab-in-the-field experiments as in Kirchler et al. (2018). For each session of Experiment PROF we booked a conference room on location, set up our mobile laboratory and invited professionals to participate. Our mobile laboratory is identical to the Innsbruck EconLab at the University of Innsbruck and the NSM Decision lab at the Radboud University in Nijmegen, where the corresponding student markets were administered. It consists of notebooks and partition walls on all sides for each participant, ensuring conditions as in regular experimental laboratories (see pictures in Appendix F). In total, we ran 10, 9, 9, and 10 markets in treatments INC, SHORT, LOW, and HIGH, respectively, with corresponding numbers of participants of 78, 71, 68, and 77, respectively. After the conclusion of the market experiment we administered a questionnaire on attitudes toward risk (from the German Socio-Economic Panel (SOEP; Dohmen

[8] See Table A1 in Appendix A for an overview of the treatment parameterization.

[9] We signed non disclosure agreements (NDAs) for not disclosing the identity of the participating financial institutions.

et al., 2011)), social status, financial success, relative performance, and competitiveness (as in Kirchler et al., 2018), as well as some demographic questions (see Appendix D for the questions and the instructions of the experiment). We programmed and conducted the experiment using z-Tree 3.6.7 (Fischbacher, 2007) and GIMS 7.2.4 (Palan, 2015).

Professionals received an average payout of 76.5 euro with a standard deviation of 12.7. The average duration of the experiment was 70 minutes. This is in line with prior studies of, for instance, Haigh and List (2005) and Kirchler et al. (2018) who report hourly payments of 96 US dollars (equivalent to 73 euro at the time of their experiment) and 72 euro, respectively.[10] Subjects' payout is composed of earnings from the asset market including the belief elicitation tasks. For the market experiment, the buyback price of 28 was multiplied by a subject's units of the asset held at the end of the experiment and added to the end holdings in Taler (including the holdings in Account B in treatments LOW and HIGH). Finally, the amount in Taler was exchanged to euro at a conversion rate of 350:1 in Treatment HIGH and 100:1 in the three other treatments with professional subjects to account for the different cash endowments across treatments. The procedure included 10 minutes to study the written instructions, two trial periods, the market experiment and the survey questions as outlined above. Finally, we administered the payout privately by handing out sealed envelopes containing the payout from the experiment.

In total we ran 48 markets with the two student subject pools from the University of Innsbruck and Radboud University Nijmegen in Experiment STUD. As in Experiment PROF each subject participated in only one market and we made sure that subjects had not participated in earlier asset market experiments of similar design. Students represent the most prevalent "classic" lab participants in experimental studies. In an attempt to resemble the gender ratio of the professionals, we recruited 75.8 percent male participants. The average age was 22.2 years and 87.3 percent were students at management and economics departments. The market setup, handouts, and the experimental protocol were identical to Experiment PROF except for the stake size. Similar to other studies (List and Haigh, 2005; Alevy et al., 2007; Cohn et al., 2014), we scaled down student stakes to 25% of the professionals' payoffs (i.e., conversion rate of Taler to euro of 1,400:1 vs. 350:1 in Treatment HIGH for students and professionals, respectively, and 400:1 vs. 100:1, respectively, in all other treatments). Students received an average payout of 18.6 euro with a standard deviation of 4.5. Student subjects were recruited using hroot by Bock et al. (2014) in Innsbruck and ORSEE by Greiner (2004) in Nijmegen.

[10] Professionals reported an average annual gross salary of 121,701 euro in the questionnaire. Accordingly, the average (maximum) hourly payoff from the experiment amounted to roughly 1.9 times (2.7 times) the average professional's hourly wage after taxes. For this calculation, we assumed a working time of 45 hours/week for 47 weeks/year and 40 percent taxes to calculate an hourly net wage (34.5 euro). In our experiment, subjects' average (maximum) hourly payment was 65.6 (93.1) euro ($76.5 \times 60/70$ and $108.6 \times 60/70$), resulting in 190 (270) percent of their salary.

2 Results

2.1 Price Efficiency

Figure 1 outlines average volume-weighted period prices of individual markets and treatment medians and means in Experiment PROF (left panel) and Experiment STUD (right panel). The upper part of Table 1 and Table A6 in Appendix A provide measures for price efficiency including significance tests for treatment differences in Experiment PROF. The lower part of Table 1 and Table A7 in Appendix A show corresponding numbers for Experiment STUD. We follow Stöckl et al. (2010) and Razen et al. (2017) in identifying mispricing, overvaluation and potential bubbles. We use RD (relative deviation of prices to fundamentals, normalized at the FV of 28) and RAD (relative absolute deviation of prices to fundamentals, normalized at the FV of 28) as measures for overpricing and mispricing, respectively (Stöckl et al., 2010). Specifically, RAD $= \sum_{t=1}^{T} \frac{|\frac{\overline{P}_t - FV_t}{FV_t}|}{T}$ and RD $= \sum_{t=1}^{T} \frac{\frac{\overline{P}_t - FV_t}{FV_t}}{T}$, with $\overline{P}$ being the average price in period t and T the total number of periods. We further use RDMAX to measure overpricing at the peak period price, denoting the corresponding period by t^*. RDMAX $= \max_t \{\frac{\overline{P}_t - FV_t}{FV_t}\} = \frac{\overline{P}_{t^*} - FV_{t^*}}{FV_{t^*}}$, is calculated as RD of the peak average period price ($\overline{P}$). Additionally, we measure the difference from the pre-peak minimum to the maximum period price as a percentage of FV with the variable AMPLITUDE. We compare the minimum average period price at period $t^* - k$ and the maximum average period price at t^*, normalized at the FV, AMPLITUDE $= \frac{\overline{P}_{t^*} - FV_{t^*}}{FV_{t^*}} - \min_{0 \leq k < t^*} \{\frac{\overline{P}_{t^*-k} - FV_{t^*-k}}{FV_{t^*-k}}\}$. Finally, we calculate the difference between the minimum price after the peak in period $t^* + l$ and the peak average price at t^*, normalized at the FV, CRASH $= \min_{0 \leq l \leq T - t^*} \{\frac{\overline{P}_{t^*+l} - FV_{t^*+l}}{FV_{t^*+l}}\} - \frac{\overline{P}_{t^*} - FV_{t^*}}{FV_{t^*}}$, to learn about the severity of crashes. Note that k and l indicate the lead and lag in periods with respect to the average peak price. RDMAX, AMPLITUDE, and CRASH are taken from Razen et al. (2017).[11] To test for significant pairwise differences between subject pools or treatments, we first compute the market mean of all period values for the variable of interest. Then we employ a Mann-Whitney U-test (MW U-test) with the market as unit of observation.[12]

Result 1: *In Experiment* PROF*, the bubble-driver treatments* INC *and* HIGH *exhibit significantly less efficient market prices compared to markets in which bubble-moderators are implemented (treatments* SHORT *and* LOW*). The latter markets show efficient prices.*

Support: As shown in the upper part of Table 1 and in Table A6 in Appendix A mispricing and overpricing in treatments SHORT and LOW are very small, with values below 7.0 percent. In contrast, median overpricing in treatments INC and HIGH is substantially higher with

[11] See also Section 2.2 for more explanations, figures B2 to B9 in Appendix B for individual transaction price charts of each market.

[12] We employ the user-written command `ranksumex` by Harris et al. (2013) in Stata to calculate exact p-values as the built-in command only provides asymptotic results by assuming normality, which is inappropriate for a small sample size of $N < 25$. Therefore, results from the MW U-tests in this paper are rather conservative.

median values reaching 12.8 and 63.1 percent, respectively. In a related way, amplitude and crashes are large, particularly in Treatment HIGH with median AMPLITUDE of 57.2 percent and median CRASH equaling –146.2 percent (both as a percentage of the FV of 28). By running pairwise Mann-Whitney U-tests in Table 2 for RD (for both experiments) and in Table A6 in Appendix A, we find no differences between SHORT and LOW in any of the five variables, hinting at similar and efficient prices. However, we observe significant differences between Treatment HIGH and treatments SHORT and LOW in all variables (most differences are significant on the 1 percent level). With the exception of CRASH, Treatment INC exhibits significantly higher values compared to both bubble-moderator treatments in all other variables as well. Moreover, we report no differences in any of the five variables between both bubble-driver treatments INC and HIGH.

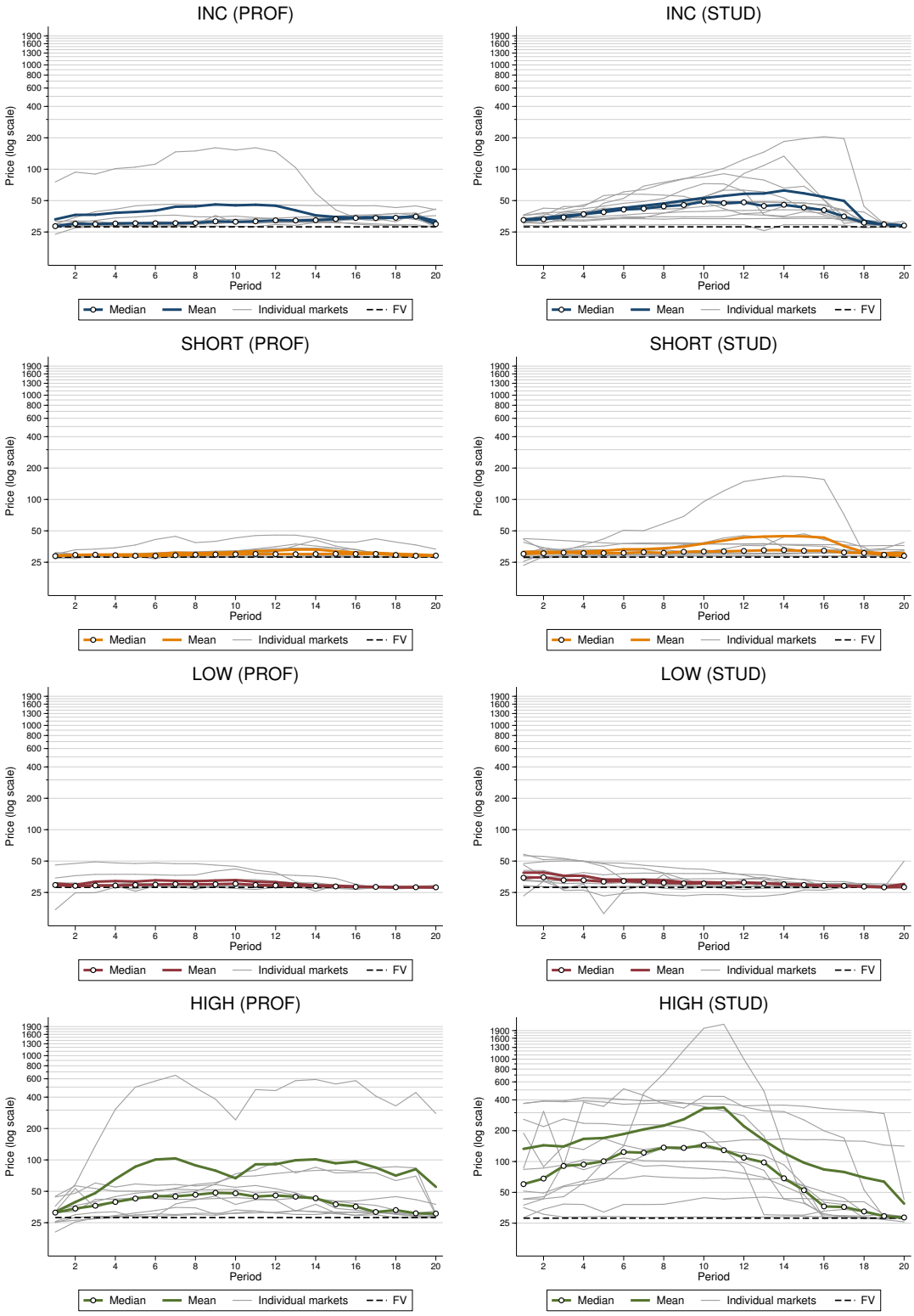


Figure 1: **Log-price developments across treatments in Experiment PROF (left column) and in Experiment STUD (right column)**: This figure depicts median treatment prices (bold and colored lines with circles) and mean treatment prices (bold and colored lines) as a function of period for treatments INC (increasing CA-Ratio), SHORT (increasing CA-Ratio, short-selling allowed), LOW (low and constant CA-Ratio of 1), and HIGH (high and constant CA-Ratio of 10.2) in log-scale. Treatments of Experiment PROF are displayed in the left column and the corresponding treatments in Experiment STUD are shown in the right column. The dashed lines represent the risk-neutral fundamental value of 28 and the grey lines show volume-weighted mean prices for individual markets.

Table 1: **Treatment medians of mispricing (RAD), overpricing (RD), maximum overpricing (RDMAX), price run-ups (AMPLITUDE), and crash (CRASH) in experiments PROF (top) and STUD (bottom) in percent**: This table depicts median treatment values for treatments INC (increasing CA-Ratio), SHORT (increasing CA-Ratio, short-selling allowed), LOW (low and constant CA-Ratio of 1), and HIGH (high and constant CA-Ratio of 10.2) for both experiments. RAD measures mispricing and is calculated as the absolute difference of mean period prices and FVs averaged across all periods of a market and RD measures overpricing by using the raw difference of mean period prices to FVs. RDMAX denotes overpricing at the peak (maximum mean period price) and AMPLITUDE measures price run-ups (amplitude) before the peak price by comparing the minimum average period price and the following maximum average period price, normalized at the FV of 28. Finally, CRASH measures the severity of a crash by taking the difference between the minimum average price after the peak and the peak average price, normalized at the FV.

Experiment PROF

	Treatment			
Variable (Median) in percent	INC	SHORT	LOW	HIGH
RAD (mispricing)	13.43	6.97	5.46	66.26
RD (overpricing)	12.77	6.58	2.81	63.11
RDMAX (max overpricing)	34.76	16.41	9.19	114.70
AMPLITUDE (price amplitude)	33.31	16.77	9.19	57.19
CRASH (price crash)	−28.67	−19.22	−44.91	−146.16
N	10	9	9	10

Experiment STUD

	Treatment			
Variable (Median) in percent	INC	SHORT	LOW	HIGH
RAD (mispricing)	47.32	13.58	13.74	342.27
RD (overpricing)	47.32	12.61	11.97	342.27
RDMAX (max overpricing)	74.93	38.49	28.57	499.89
AMPLITUDE (price amplitude)	66.62	15.72	3.36	147.59
CRASH (price crash)	−74.33	−47.56	−53.99	−500.46
N	12	12	12	12

Table 2: **Pairwise Mann-Whitney U-tests of overpricing (RD) in experiments PROF and STUD**: This table shows pairwise treatment comparisons for treatments INC (increasing CA-Ratio), SHORT (increasing CA-Ratio, short-selling allowed), LOW (low and constant CA-Ratio of 1), and HIGH (high and constant CA-Ratio of 10.2) in Experiment STUD. The numbers identify the difference in the treatment medians in percentage points, i.e., the value of the "row" treatment minus the value of the "column" treatment (a positive value implies that, for instance, INC is larger than SHORT). *, ** and *** represent the 10%, 5%, and 1% significance levels of a double-sided test. Sample size N for each test is between 18 and 20 for PROF and 24 for STUD.

	RD (overpricing), treatment differences in percentage points					
	PROF			**STUD**		
Treatment	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	6.19**	9.95**	−50.35	34.71***	35.35***	−294.96***
SHORT	.	3.76	−56.54***	.	0.64	−329.66***
LOW	.	.	−60.30***	.	.	−330.30***

Result 2: *In Experiment* STUD*, markets with bubble-drivers capital inflow (*INC*) and high initial CA-Ratio (*HIGH*) exhibit significantly less efficient prices compared to the bubble-moderator treatments* SHORT *and* LOW*. Again, markets of the bubble-moderator treatments show efficient prices.*

Support: As outlined in the lower part of Table 1 and in Table A7 in Appendix A overpricing (RD) of 12.6 and 12.0 percent in treatments SHORT and LOW is comparatively low. The numbers for mispricing are very similar. Again, the bubble-driver treatments INC and HIGH show substantially higher values with median overpricing reaching 47.3 and 342.3 percent, respectively. Moreover, amplitude and crashes are substantial, especially in Treatment HIGH with median AMPLITUDE of 147.6 percent and median CRASH of −500.5 percent (both as a percentage of the FV of 28). We run pairwise Mann-Whitney U-tests in Table 2 for RD (for both experiments) and in Table A7 for all other variables.[13] We find no differences between SHORT and LOW in any of the five outlined variables, hinting at very similar and in general rather efficient prices. However, we find significant differences between Treatment HIGH and all other treatments in all variables (with the only exception being Treatment INC for AMPLITUDE). Most of these differences are significant at the 1 percent level. The other bubble-driver treatment INC also exhibits significantly higher levels of price inefficiency compared to treatments SHORT and LOW, which is evident in almost all variables.

Result 3: *Markets populated by professionals are significantly more efficient compared to student markets. This result holds for the bubble-driver treatments* INC *and* HIGH*, but not for*

[13] As results are very similar in all other variables compared to RD, we only outline RD in Table 2 for exemplary purpose.

the bubble-moderator treatments SHORT *and* LOW*, which have similarly high levels of price efficiency.*

Support: As already outlined with the previous results, mispricing and overpricing are low for both subject pools in treatments SHORT and LOW. In particular, for bubble-moderator treatments, differences in RAD and RD between students and professionals are below 10 percentage points (see Table 3). By running pairwise Mann-Whitney U-tests in Table 3 we find no statistical differences between SHORT and LOW in all variables but RDMAX, hinting at similar and very efficient prices in both subject pools. In contrast, with differences in median overpricing (RD) of 34.6 and 279.2 percentage points in treatments INC and HIGH, respectively, student markets are significantly more inefficient compared to markets populated by professionals. In addition, subject pool differences in the median RDMAX are substantial with values of 40.2 and 385.2 percentage points in both treatments, respectively, as well as in the median CRASH with differences of 45.7 (INC) and 354.3 (HIGH) percentage points. In particular, we find significant differences in all variables except price amplitude between student and professional markets in both treatments testing bubble-drivers, INC and HIGH.

Table 3: **Pairwise Mann-Whitney U-tests of mispricing (RAD), overpricing (RD), maximum overpricing (RDMAX), price run-ups (AMPLITUDE), and crash (CRASH) between experiments PROF and STUD**: This table shows pairwise subject pool comparisons for each treatment: INC (increasing CA-Ratio), SHORT (increasing CA-Ratio, short-selling allowed), LOW (low and constant CA-Ratio of 1), and HIGH (high and constant CA-Ratio of 10.2). The table outlines median treatment values of the respective variables in percent and the numbers in parentheses show the Z-values of the MW U-test statistic. *, ** and *** represent the 10%, 5%, and 1% significance levels of a double-sided test. Sample size N for each test is either 21 or 22.

	RAD			**RD**			**RDMAX**		
Treatment	PROF	STUD	Z	PROF	STUD	Z	PROF	STUD	Z
INC	13.43	47.32	(2.11)**	12.77	47.32	(2.11)**	34.76	74.93	(1.71)*
SHORT	6.97	13.58	(1.63)	6.58	12.61	(1.49)	16.41	38.49	(1.35)
LOW	5.46	13.74	(1.42)	2.81	11.97	(0.92)	9.19	28.57	(1.92)*
HIGH	66.26	342.27	(2.18)**	63.11	342.27	(2.18)**	114.70	499.89	(1.98)**
	AMPLITUDE			**CRASH**					
Treatment	PROF	STUD	Z	PROF	STUD	Z			
INC	33.31	66.62	(1.32)	−28.67	−74.33	(−1.98)**			
SHORT	16.77	15.72	(0.04)	−19.22	−47.56	(−1.42)			
LOW	9.19	3.36	(−0.87)	−44.91	−53.99	(−1.21)			
HIGH	57.19	147.59	(1.12)	−146.16	−500.46	(−1.78)*			

Finally, we turn to price beliefs and investigate whether forecast accuracy of prices differs between professionals and students and between bubble-driver and bubble-moderator treatments. For this, we calculate the forecast error of price beliefs, $\mathrm{FE}^i_{t,t+k}$, of subject i in period t for period $t+k$ as follows: $\mathrm{FE}^i_{t,t+k} = ln(\frac{\overline{P}_{t+k}}{\widetilde{\mathrm{P}}^i_{t,t+k}})$. Here, $\overline{P}_{t+k}$ stands for the mean period price in period $t+k$ with k indicating values in the range {0,1,2} and $\widetilde{\mathrm{P}}^i_{t,t+k}$ indicates subject i's beliefs in period t of the mean market price in $t+k$. Hence, with $\mathrm{FE}^i_{t,t+k}$ we measure the percentage difference of future market prices in $t+k$ and subject's price beliefs for $t+k$, elicited in t.

Result 4: *Differences in professionals' and students' price forecast errors are statistically insignificant in all treatments before the price peak. However, professionals predict prices significantly better than students in both bubble-driver treatments after the price peak.*

Support: All results are reported in detail in Appendix C. For the bubble-moderator treatments, SHORT and LOW, forecasts are very accurate both for professionals and students, which is not surprising given the high level of price efficiency in tracking the FV (Table C8). For the bubble driver treatments, INC and HIGH, we find that both groups, professionals and students, find it similarly difficult to predict prices before they peak (upswings). Professionals underestimate real prices in upswings by 5.0 to 14.5 percent across all forecasting periods (t to $t+2$). Students' corresponding underestimation is statistically not different from professionals and lies between 3.5 and 13.8 percent (see upper panel of Table C8). In downswings (after price peaks) in bubble driver treatments, professionals' price predictions are very accurate with prices being very close to price beliefs in the range of 0.3 to –2.1 percent for all forecasting periods. Students, in contrast, significantly underestimate price downswings in all bubble driver treatments: prices fall below beliefs ranging from –4.8 to –26.7 percent (see lower panel of Table C8). In part, the difference between subject pools in bubble driver downswings may be due to the fact that markets with professionals are generally closer to fundamentals, particularly in INC and HIGH (see Table 3), thus making forecasts easier for professionals.

2.2 Bubble Identification

In this section we attempt to identify bubble markets and to separate them from non-bubble markets. As outlined in the introductory section, there is heterogeneity regarding bubble definitions, particularly in the experimental markets literature.[14] Some of the challenges in identifying bubbles in the lab concern the variables to measure a bubble with and the threshold values that

[14]For instance, according to the survey of Brunnermeier (2009) "*[b]ubbles refer to asset prices that exceed an asset's fundamental value because current owners believe that they can resell the asset at an even higher price in the future.*" In the experimental literature, King et al. (1993) speak of a bubble when "*...traders invariably trade in high volume at prices that are considerably at variance from intrinsic value...*". Noussair et al. (2001) follow this definition and quantify a bubble according to two criteria; i) the median transaction price in five consecutive periods is at least 50 units of experimental currency (about 13.9 percent) greater than the fundamental value and ii) the average price is at least two standard deviations (of transaction prices) greater than the fundamental value for five consecutive periods.

are used to separate bubble markets from non-bubble markets. We follow the approach of Razen et al. (2017), who developed an endogenous bubble definition by using a "benchmark treatment" in absence of any bubble-driver, assuming that it represents the expected price characteristics of the particular asset market without treatment intervention. They consider price developments to constitute a bubble if the deviations in *all three* of their measures, RDMAX, AMPLITUDE, and CRASH exceed the 95th percentile of the corresponding measure in the benchmark treatment (distribution) and therefore can be considered "significant" deviations from the benchmark. For our purpose we take all 21 markets of Treatment LOW as a benchmark because there are no bubble-drivers in this treatment. We pool student and professional markets of this treatment as we find no treatment differences between students and professionals in any of the variables.[15]

In particular, Razen et al. (2017) define a bubble episode to be characterized by the time interval between the periods with the lowest average market prices before and after the price peak (relative to the fundamental value). With this definition a bubble requires a subsequent crash to separate it from other forms of overpricing such as information mirages (Camerer and Weigelt, 1991). Moreover, we follow their approach and define three criteria (C1–C3) that have to be jointly fulfilled to term a market a bubble market.

C1: Price bubbles are characterized by an extraordinarily high peak average period price. A market m fulfills criterion C1 iff

$$\begin{gathered}\text{RDMAX}_m > \\ MAX\{0; \overline{\text{RDMAX}}^{\text{LOW}} + t(df)_{0.95} \cdot \sigma(\text{RDMAX}^{\text{LOW}})\},\end{gathered} \tag{1}$$

with $\overline{\text{RDMAX}}^{\text{LOW}}$ indicating the mean of the maximum period peak prices of the 21 benchmark markets of both treatments LOW. $t(df)_{0.95}$ stands for the 95 percent quantile of a student t-distribution with $N-1$ degrees of freedom (df) and N is the number of markets in the benchmark. $\sigma(\text{RDMAX}^{\text{LOW}})$ stands for the standard deviation of RDMAX in the N baseline markets. If RDMAX_m is higher than the 95 percent quantile, its peak period price for a particular market is considered to be significantly higher than in the benchmark.[16]

C2: Price bubbles are characterized by exhibiting extraordinary price rallies toward the peak price. A market m fulfills criterion C2 iff

$$\begin{gathered}\text{AMPLITUDE}_m > \\ \overline{\text{AMPLITUDE}}^{\text{LOW}} + t(df)_{0.95} \cdot \sigma(\text{AMPLITUDE}^{\text{LOW}}),\end{gathered} \tag{2}$$

[15]Results of bubble identification remain identical if we only take the 12 (9) markets of the students (professionals) as benchmark for all other student (professional) markets of the other treatments.

[16]Note that we also impose that RDMAX_m must exceed zero to rule out potential price paths that do not exceed the FV.

with $\overline{\text{AMPLITUDE}}^{\text{LOW}}$ and $\sigma(\text{AMPLITUDE}^{\text{LOW}})$ indicating the mean and the standard deviation of AMPLITUDE of the benchmark markets.

C3: Price bubbles are characterized by exhibiting extraordinary crashes. A market m fulfills criterion C3 iff

$$\text{CRASH}_m < \overline{\text{CRASH}}^{\text{LOW}} - t(df)_{0.95} \cdot \sigma(\text{CRASH}^{\text{LOW}}), \tag{3}$$

with $\overline{\text{CRASH}}^{\text{LOW}}$ and $\sigma(\text{CRASH}^{\text{LOW}})$ defining the mean and the standard deviation of CRASH of the benchmark markets.

We consider price developments to constitute a bubble if the deviations in *all three* measures, i.e., C1 the maximum period price (RDMAX), C2 the price run-up (AMPLITUDE), and C3 the CRASH are above the 95 percent quantile (below the 5 percent quantile for C3) of the corresponding measure in the benchmark distribution. If only one of the measures fails to exceed this defined range, we do not classify the market as a bubble market.

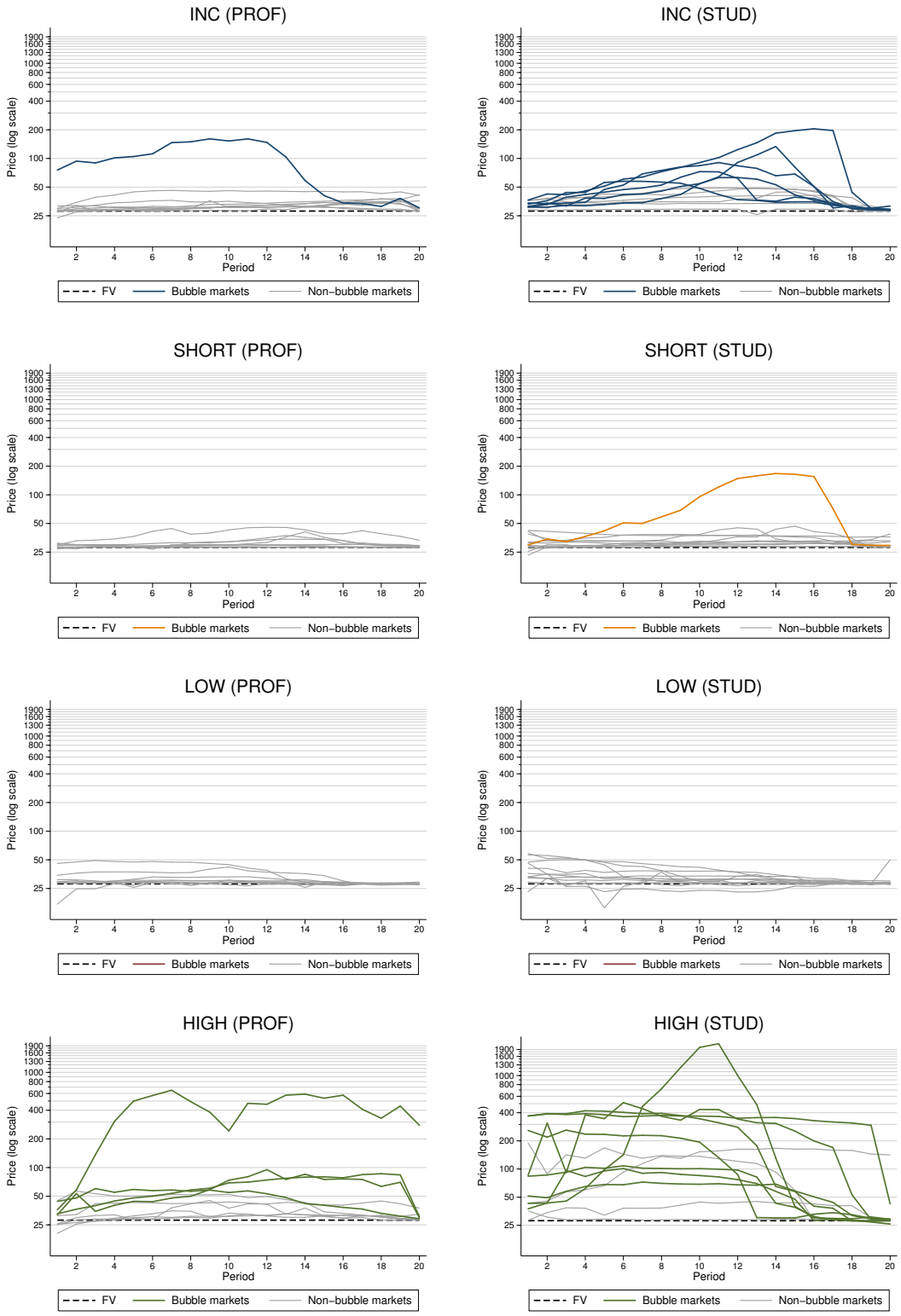


Figure 2: **Bubble identification across treatments in Experiment PROF (left column) and in Experiment STUD (right column)**: Following the bubble definition in Section 2.2 this figure depicts volume-weighted mean prices for bubble markets (bold and colored lines) and non-bubble markets (grey lines) as a function of period for treatments INC (increasing CA-Ratio), SHORT (increasing CA-Ratio, short-selling allowed), LOW (low and constant CA-Ratio of 1), and HIGH (high and constant CA-Ratio of 10.2) in log-scale. Treatments of Experiment PROF are displayed in the left column and the corresponding treatments in Experiment STUD are shown in the right column. The dashed lines represent the risk-neutral fundamental value of 28.

Result 5: *Following the bubble classification, 25 percent of markets with professionals are defined as bubble markets in the two bubble-driver treatments* INC *and* HIGH*. In markets with students, bubble markets are more frequent and account for 58 percent of all markets in treatments* INC *and* HIGH*.*

Support: Figure 2 depicts the individual markets of all treatments, separated into bubble markets (bold and colored lines) and non-bubble markets (grey lines). Following our classification none of the 18 markets in the bubble-moderator treatments SHORT and LOW exhibits bubbles in the professional sample. In the student sample it is only 1 out of 24 markets with an identified bubble pattern, indicating that both treatments show consistent and non-bubble price patterns across both subject pools. In contrast, 10 and 40 percent of markets with professionals are defined as bubble markets in the bubble-driver treatments INC and HIGH, respectively. Bubble markets are even more frequent in the student sample as 50 and 67 percent of the markets in treatments INC and HIGH, respectively, show bubble patterns (see tables A2 to A5 in Appendix A for details on all measures, separated for each market).

2.3 Potential Explanations of Differences in Price Efficiency between Professionals and Students

In this section we test potential explanations for the observed differences in price efficiency between professionals and student subjects. In doing so we focus on two dimensions which prior literature suggests as natural candidates and along which professionals can potentially differ from students: first, professionals can differ from students in cognitive skills, through selection into the industry or through learning on the job (or both). Second, professionals can differ in risk attitudes. Both, cognitive skills and risk attitudes, have been shown to drive behavior and performance on stock markets (e.g., Fellner and Maciejovsky, 2007; Grinblatt et al., 2011, 2012; Kleinlercher et al., 2014; Corgnet et al., 2018).

Empirical studies outline that high IQ-investors show higher levels of stock market participation (Christelis et al., 2010), earn higher Sharpe ratios (Grinblatt et al., 2011), are less prone to the disposition effect, exhibit superior market timing, and stock-picking skills, which results in outperformance compared to low-IQ investors (Grinblatt et al., 2012). Experimental finance literature extends these findings by analyzing the impact of various cognitive skills: Fluid intelligence measures the capacity to reason and solve novel problems and is necessary for logical problem solving (Mackintosh, 2011). Cognitive reflection adds to fluid intelligence because it helps individuals to avoid commonly-observed heuristics and biases and measures the ability to engage in effortful reasoning (e.g., Oechssler et al., 2009; Toplak et al., 2011, 2014). Theory of Mind (ToM) defines one's capacity to infer others' intentions, which is considered important in detecting the informational content of trading by inferring others' intentions from order books and prices (Bruguier et al., 2010). Experimental evidence suggests that various forms of cognitive abilities are conducive to trader performance: high cognitive reflection scores predict

subjects' earnings in asset markets with student subjects (Noussair et al., 2014; Corgnet et al., 2015b), ToM correlates with subjects' skills in predicting price changes (Bruguier et al., 2010), and all three concepts are joint predictors of trader performance (Corgnet et al., 2018). Moreover, Corgnet et al. (2015a) and Bosch-Rosa et al. (2018) find a causal relationship between traders' cognitive sophistication and price efficiency.

The experimental literature on the potential effects of risk aversion on trading behavior and price efficiency is, according to our knowledge, surprisingly thin and only provides correlational evidence. Kleinlercher et al. (2014) show that less risk averse subjects invest more in the risky asset when faced with bonus incentives, resulting in asset overvaluation. Fellner and Maciejovsky (2007) report that the higher the degree of risk aversion among subjects in the market, the lower the observed market activity. Similarly, Robin et al. (2012) find that both mispricing and asset turnover are lower when the pool of traders exhibits a higher level of risk-aversion.

To investigate whether higher cognitive abilities and, possibly, differences in risk attitudes can explain higher price efficiency in markets with professionals, we administered an online survey to newly recruited samples of 121 financial professionals from various OECD countries and 124 students. We deliberately separated this survey from the experiments for two reasons. First, we wanted to prevent confounding effects, such as, for example, within-subject order effects from the trading activity on cognitive measurements. Second, the limited time financial professionals had for the experiment did not allow to add the extensive cognitive tasks to the main experiment within a session. Therefore, we took great care to recruit a sample from the same subject pool of financial professionals as in the main experiment. Professionals are employed in the same areas as the ones from the market experiments and they share the same characteristics. Among all professionals, 84.3 percent were male, average age was 37.0 years and they have been working in the industry for 12.3 years. Student subjects were selected from the same subject pool as in the student market experiments. Here, 87.1 percent of all students were male and average age was 23.8 years. The survey was programmed and conducted with oTree (Chen et al., 2016) and professionals (students) received a flat payment of 40 (10) euro for compensation.

To test fluid intelligence, we administered a test similar to Corgnet et al. (2018), i.e., 18 of the Raven's advanced progressive matrices (Raven, 2000). For cognitive reflection skills, we used the extended cognitive reflection test (CRT) from Toplak et al. (2014) with seven items. The CRT rests on the dual-process theory framework (Kahneman, 2011).[17] To measure ToM-skills, we administered 18 pictures of the eye-gaze test from Baron-Cohen et al. (2001). In this test, participants look at images of people's eyes and choose one of four feelings that best

[17]The questions of cognitive reflection tests are constructed in a way that they have an intuitive, but on reflection incorrect, response put forward by System 1. The correct response requires the effortful activation of System 2. For instance, "A bat and a ball cost $1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost", (Frederick, 2005). The (incorrect) intuitive answer (10 cents) can be "overruled" upon reflection (5 cents) which requires effortful System 2 processes.

describe the mental state of the person whose eyes are shown.[18] Details on the tests can be found in Appendix E. In addition, we administered a HIT15 test (Burks et al., 2009) analyzing individuals' backward induction abilities, which are important in finite horizon markets.[19] The order of the four tasks was randomized across all subjects. To measure risk attitudes we took the survey question concerning general risk taking from the German Socio-Economic Panel (SOEP; Dohmen et al., 2011).[20]

Result 6: *Professionals' and students' cognitive skills are statistically indifferent from each other. Moreover, professionals self-report significantly higher levels of financial risk-taking than students. Higher levels of price efficiency in Experiment* PROF*, however, cannot be explained by the reported equal levels of cognitive skills or higher levels of risk taking by professionals.*

Support: Table 4 and Figure A1 in Appendix A outline the results. Although we find that professionals are better than students in 3 out of 4 of the cognitive tests, differences are very small and not significant at the 5% level for any test.[21] Thus, cognitive abilities cannot explain higher levels of price efficiency in markets with professionals. When turning to subjects' self-perception of risk attitudes, we find no statistical difference for general risk taking, but significantly higher levels for professionals concerning financial risk taking. This indicates that differences in risk attitudes cannot serve as an explanation either, because professionals report even higher levels of risk taking in financial matters. This would work in the opposite direction, because with less risk

[18]For the Raven's and eye-gaze test part of the survey, we used a shortened version. The original tasks comprise 36 questions each, out of which we took every second question, starting with the first one of the original task. This was done to keep the overall time needed to complete the survey as short a possible without losing explanatory power.

[19]The HIT15 is a game between the subject and the computer. The computer and the subject take turns in adding points (from 1 to 3) to a basket. The goal of the game is to be the first player to reach 15 points. The initial number of the game is randomly determined. The task was played for 6 rounds.

[20]Subjects answered the question on general risk taking: "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid risks?" The answers were provided on a Likert scale from 0 (not at all willing to take risks) to 10 (very willing to take risks). This question was also administered to professionals and students in the study of Kirchler et al. (2018). Dohmen et al. (2011) find that the self-reported SOEP measure can represent a valid substitute for incentivized lottery schemes and that it performs reasonably well in predicting risk taking behavior of individuals. Crosetto and Filippin (2013) report that the single-item SOEP measure is highly and significantly correlated with the Domain-Specific Risk-Taking Scale (DOSPERT), which is a validated measure of risk attitudes across domains and contexts (Blais and Weber, 2006). As risk attitudes can differ between contexts (Blais and Weber, 2006), we also administered the SOEP questions about risk-taking in the financial domain. Specifically, we asked: "People can behave differently in different situations. How would you rate your willingness to take risks in financial matters?" We used the same coding as for the general SOEP question.

[21]For the CRT, results still hold if we compare alternative measures to just counting the number of correct answers. We additionally calculated ECRT1 and ECRT2 measures as proposed by Noussair et al. (2016), where answers are given a weight according to whether they are (a) correct, (b) wrong, but correspond to the intuitive answer, (c) all other answers. ECRT1 punishes type (b) answers more severely, whereas ECRT2 punishes type (c) answers more severely. We also use the measure developed by Jimenez et al. (2018), classifying subjects into *reflective subjects* (at least 5 out of 7 answers are correct), *impulsive subjects* (at least 5 out of 7 answers correspond to the intuitive answer), and *other* for all other combinations of answers. Results are available upon request.

Table 4: **Pairwise Mann-Whitney U-tests of task scores of fluid intelligence (RAVEN: Raven's advanced progressive matrices), cognitive reflection (CRT), theory of mind (ToM: eye-gaze test), and backward induction skills (HIT15), respectively, between STUD and PROF:** This table shows pairwise subject pool comparisons for each task. The table outlines summary statistics and the numbers in parentheses show the Z-values of the MW U-test statistic. The maximum score in the tests was 18 (RAVEN and ToM), 7 (CRT), and 6 (HIT15). *, ** and *** represent the 10%, 5%, and 1% significance levels of a double-sided test. Sample size N for each test is 245 (121 professionals and 124 students).

Cognitive Skill Task		Mean	SD	Min.	Max.	Z
RAVEN	PROF	9.99	3.14	0.00	17.00	(0.167)
	STUD	10.07	2.98	2.00	16.00	
CRT	PROF	5.03	1.79	0.00	7.00	(−1.594)
	STUD	4.69	1.83	0.00	7.00	
ToM	PROF	10.37	2.60	4.00	16.00	(−1.159)
	STUD	9.90	2.68	3.00	15.00	
HIT15	PROF	4.50	1.25	1.00	6.00	(−1.930)*
	STUD	4.23	1.24	0.00	6.00	
Risk Attitudes						
General Risk	PROF	5.71	1.86	1.00	10.00	(−1.252)
	STUD	5.37	1.81	1.00	9.00	
Financial Risk	PROF	5.93	2.26	1.00	10.00	(−3.764)***
	STUD	4.85	2.25	0.00	9.00	

aversion we would expect, particularly in the bubble-driver treatments, more price inefficiencies and bubbles in markets populated by professionals.

As subject pool differences cannot be explained by specific cognitive skills and risk attitudes, the observed differences in price efficiency must be due to other, unobserved characteristics. As we compare the two subject pools in a correlational study, this unobserved characteristic could in principle be any difference (or combination of differences) between subject pools, ranging from age to wealth to even competitiveness (Kirchler et al., 2018). However, a natural and intuitive candidate is the years- and sometimes decades-long real-world market experience of financial professionals. This includes a whole package of differences, for instance, experience with price developments on real financial markets, experience with real-world trading (e.g., with the formation of order books and inferring information from them), and also experience and expectations about what other market participants (i.e., professionals in PROF) might do.

3 Conclusion

In this study we investigated the impact of financial professionals' behavior on price efficiency and bubble formation in a large-scale lab-in-the-field experiment and, for comparison, of students in a lab experiment. We ran 38 asset markets with financial professionals from high-skilled investment areas and contrasted their behavior to 48 markets with student subjects without any professional market experience. We administered two classical bubble-driver treatments by either implementing a high initial level of the monetary supply relative to the asset value or by imposing capital inflows over time. We also ran two classical treatments featuring bubble-moderators by either allowing short-selling or by keeping the level of capital inflow constant and low over time. Moreover, we administered an extensive survey to 121 professionals and 124 students to measure several cognitive skills (fluid intelligence, cognitive reflection, theory of mind, backward induction) and risk attitudes as potential drivers of the observed differences in price efficiency between professionals and students.

We found that professionals are not immune to bubbles in experimental asset markets. In fact, in both bubble-driver treatments 25 percent of all markets with professionals generated bubbles (following the definition of Razen et al., 2017). Moreover, we found significant overpricing by professionals in both bubble-driver treatments. With this finding we contribute to the ongoing debate on the degree of price efficiency on financial markets by presenting a large-scale experiment with financial professionals from high-skilled investment areas. We add the finding that professionals generate market inefficiencies and bubbles, even in relatively simple and controlled market environments with only one tradeable asset. We also add to the emerging experimental literature analyzing behavior of financial professionals (e.g., Haigh and List, 2005; Alevy et al., 2007; Cohn et al., 2014, 2017; Kirchler et al., 2018).

When we compared professionals' behavior with that of students (*across* subject pools), markets populated by professionals generated less overpricing, fewer bubbles and the bubbles were smaller than in student markets. These findings apply to both bubble-driver treatments. In the bubble-moderator treatments we found a high level of price efficiency, very close to the fundamental value, for both subject pools. With this we contribute to the experimental finance literature investigating bubble-drivers and moderators in classical laboratory experiments with student subjects (e.g., Smith et al., 1988; Lei et al., 2001; Dufwenberg et al., 2005; Kirchler et al., 2012; Sutter et al., 2012). The theory does not discriminate by who is participating. We show that there is a significant difference: faced with bubble-drivers, markets populated with professionals are significantly more efficient than student markets.

We also observed differences in price forecasting quality between professional and student markets. Both groups, professionals and students, underestimated prices in upswings (i.e., before the price peak) in bubble-driver treatments to a similar extent. In downswings, however,

professionals predicted prices remarkably well in the bubble-driver treatments, and significantly better than students, who significantly overestimated prices after the price peak.

Despite all these differences, we also found qualitatively very similar patterns *within* each subject pool. We showed that bubble-drivers reduced price efficiency and increased the proneness to bubbles, whereas bubble-moderators yielded efficient markets. In other words, bubble-drivers did not only affect students but also professionals, and in specific market environments with bubble-moderators even inexperienced subjects priced efficiently. This is good news for experimenters with student subjects, because, even though the treatment effect sizes are smaller for professionals and bubbles are less likely, the effects are comparable in direction and statistical significance.

Finally, we probed for potential drivers of the results and found that cognitive skills and risk attitudes cannot serve as explanations for differences in price efficiency between professionals and students. For both subject pools we found equal levels of cognitive skills and students self-reported lower levels of financial risk taking than professionals, which is at odds with more bubbles in student markets. The results suggest that the higher levels of price efficiency in markets with professionals may be explained by real-world market experience, including a set of characteristics that go beyond specific cognitive skills. These characteristics and skills are difficult to observe, but can include, for instance, experience with price developments in real-world markets, experience with trading in general, and experience and expectations about what other market participants might do. The latter refers to beliefs about the level of "rationality" of other market participants. It is indeed possible that experienced traders, knowing that they are trading with other professionals, trade more efficiently, because they developed a common knowledge of each other's rationality (in contrast to markets with students).[22] However, it is not clear how market experience (and which aspect of it) affects price efficiency. On the one hand, professionals with year-long experience in financial decision-making may have learned about themselves (e.g., how to avoid costly mistakes) and about the behavior of others (as explained above). This can be conducive to price efficiency, as suggested by evidence from experiments with students. Dufwenberg et al. (2005) and Sutter et al. (2012) find that re-running market experiments with intact cohorts, thus conditioning subjects to the market environment and their fellow traders, leads to more efficient markets. On the other hand, financial professionals' daily exposure to real-world markets may also aggravate bubbles, for example, if market experience increases the readiness to speculate and take risks. A rich history of theoretical models shows that it can be profitable for rational agents to fuel bubbles (Brunnermeier, 2008; Brunnermeier and Oehmke, 2013). Informational frictions, for example, allow more experienced or sophisticated traders to "ride a bubble", where "*investors can rationally expect an asset price to move in one*

[22]The importance of beliefs about others and the common rationality assumption could be tested with hybrid sessions of professionals and students like in Cheung et al. (2014). For operational reasons we were not able to do this, but we consider this to be a promising avenue for future research.

direction in the short run and in the opposite direction in the long run" (DeLong et al., 1990, p. 394).

Ultimately, the answer to the question whether market experience fuels or mitigates bubbles is an empirical one. Our results corroborate the notion that professionals' market experience supports price efficiency and that (relatively) inexperienced private investors play an important role in bubble formation. Following this interpretation, our results shed light on the role of different investor groups (inexperienced investors vs financial professionals) for price efficiency and speculative bubbles. According to the narratives of Kindleberger and Aliber (2011) and the analyses of Brunnermeier and Schnabel (2016), private investors contributed significantly to speculative bubbles in history. Moreover, Griffin et al. (2011) investigate the behavior of different investor groups during the Tech Bubble and show that institutional investors start pulling capital out of the market at the peak in mid-March 2000, while various individual investor groups accelerate their purchases even during the crash. Cheng et al. (2014), however, show that professionals might face difficulties in detecting bubbles. The authors focus on the bubble in the US housing market from 2004-2006 and find that securitization investors and issuers (i.e., mid-level employees in the mortgage securitization business) increase their private housing exposure during the boom. We contribute to this emerging strand of literature by providing controlled evidence on how professionals behave in bubble and non-bubble environments. However, more research is needed to establish the role of market experience in bubble formation.

References

Ackert, Lucy F., Narat Charupat, Bryan K. Church, Richard Deaves. 2006. Margin, short sell, and lotteries in experimental asset markets. *Southern Economic Journal* **73**(2) 419–436.

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.

Allen, Franklin, Douglas Gale. 2000. Bubbles and crises. *The Economic Journal* **110**(1) 236–255.

Allen, Franklin, Gary Gorton. 1993. Churning bubbles. *The Review of Economic Studies* **60**(4) 813–836.

Baron-Cohen, Simon, Sally Wheelwright, Jacqueline Hill, Yogini Raste, Ian Plumb. 2001. The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines* **42**(2) 241–251.

Blais, Ann-Renée, Elke U Weber. 2006. A Domain-Specific Risk-Taking (DOSPERT) scale for adult populations. *Judgment and Decision Making* **1**(1) 33–47.

Bloomfield, R., A. Anderson. 2010. *Behavioral Finance: Investors, Corporations, and Markets*. John Wiley & Sons.

Bock, Olaf, Ingmar Baetge, Andreas Nicklisch. 2014. hroot: Hamburg registration and organization online tool. *European Economic Review* **71** 117–120.

Bosch-Rosa, Ciril, Thomas Meissner, Antoni Bosch-Domènech. 2018. Cognitive bubbles. *Experimental Economics* **21** 132–153.

Bruguier, Antoine J., Steven R. Quartz, Peter Bossaerts. 2010. Exploring the nature of "trader intuition". *The Journal of Finance* **65**(5) 1703–1723.

Brunnermeier, Markus K. 2001. *Asset pricing under asymmetric information: Bubbles, Crashes, Technical Analysis, and Herding*. Oxford University Press.

Brunnermeier, Markus K. 2008. Bubbles. Steven N. Durlauf, Lawrence E. Blume, eds., *The New Palgrave Dictonary of Economics*. Elsevier, 1221–1288.

Brunnermeier, Markus K. 2009. Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic Perspectives* **23**(1) 77–100.

Brunnermeier, Markus K., Martin Oehmke. 2013. Bubbles, financial crises, and systemic risk. George Constantinides, Milton Harris, Rene Stulz, eds., *Handbook of the Economics of Finance*, vol. 2. Elsevier, 1221–1288.

Brunnermeier, Markus K., Isabel Schnabel. 2016. Bubbles and central banks: Historical perspectives. Michael D. Bordo, Eitrheim, Marc Flandreau, Jan F. Qvigstad, eds., *Central Banks at a Crossroads: What Can We Learn from History?*. Cambridge University Press.

Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, Aldo Rustichini. 2009. Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Science* **106**(19) 7745–7750.

Caginalp, Gunduz, David Porter, Vernon L. Smith. 1998. Initial cash/asset ratio and asset prices: An experimental study. *Proceedings of the National Academy of Sciences* **95**(2) 756–761.

Caginalp, Gunduz, David Porter, Vernon L. Smith. 2001. Financial bubbles: Excess cash, momentum and incomplete information. *The Journal of Psychology and Financial Markets* **2**(2) 80–99.

Camerer, Colin, Keith Weigelt. 1991. Information mirages in experimental asset markets. *The Journal of Business* **64**(4) 463–493.

Chen, Daniel L., Martin Schonger, Chris Wickens. 2016. oTree - an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* **9** 88–97.

Cheng, Ing-Haw, Sahil Raina, Wei Xiong. 2014. Wall street and the housing bubble. *American Economic Review* **104**(9) 2797–2829.

Cheung, Stephen L., Morten Hedegaard, Stefan Palan. 2014. To see is to believe - common expectations in experimental asset markets. *European Economic Review* **66** 84–96.

Christelis, Dimitris, Tullio Jappelli, Mario Padula. 2010. Cognitive abilities and portfolio choice. *European Economic Review* **54**(1) 18–38.

Cohn, Alain, Ernst Fehr, Michel André Maréchal. 2014. Business culture and dishonesty in the banking industry. *Nature* **516** 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.

Corgnet, Brice, Mark DeSantis, David Porter. 2015a. Revisiting information aggregation in asset markets: Reflective learning and market efficiency. *Working Paper* .

Corgnet, Brice, Mark DeSantis, David Porter. 2018. What makes a good trader? On the role of quant skills, behavioral biases and intuition on trader performance. *The Journal of Finance* **forthcoming**.

Corgnet, Brice, Roberto Hernán-González, Praveen Kujal, David Porter. 2015b. The effect of earned versus house money on price bubble formation in experimental asset markets. *Review of Finance* **19**(4) 1455–1488.

Crosetto, Paolo, Antonio Filippin. 2013. The "bomb" risk elicitation task. *Journal of Risk and Uncertainty* **47**(1) 31–65.

DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, Robert J. Waldmann. 1990. Noise trader risk in financial markets. *Journal of Political Economy* **98**(4) 703–738.

Dohmen, Thomas J., Armin Falk, David Huffman, Juergen Schupp, Uwe Sunde, Gert Wagner. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* **9**(3) 522–550.

Dufwenberg, Martin, Tobias Lindqvist, Evan Moore. 2005. Bubbles and experience: An experiment. *American Economic Review* **95**(5) 1731–1737.

Eckel, Catherine C., Sascha Füllbrunn. 2015. Thar she blows? gender, competition, and bubbles in experimental asset markets. *American Economic Review* **105**(2) 906–920.

Engsted, Tom. 2016. Fama on bubbles. *Journal of Economic Surveys* **30**(2) 370–376.

Fama, Eugene F. 1970. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance* **25**(2) 383–417.

Fellner, Gerlinde, Boris Maciejovsky. 2007. Risk attitude and market behavior: Evidence from experimental asset markets. *Journal of Economic Psychology* **28**(3) 338–350.

Fischbacher, Urs. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* **10**(2) 171–178.

Frederick, Shane. 2005. Cognitive reflection and decision making. *Journal of Economic Perspectives* **19**(4) 25–42.

Galbraith, John Kenneth. 1994. *A short history of financial euphoria*. Penguin.

Greiner, Ben. 2004. An Online Recruitment System for Economic Experiments. Kurt Kremer, Volker Macho, eds., *Forschung und wissenschaftliches Rechnen. GWDG Bericht 63*. Gesellschaft fuer Wissenschaftliche Datenverarbeitung, 79–93.

Griffin, John, Jeffrey Harris, Tao Shu, Selim Topaloglu. 2011. Who drove and burst the tech bubble? *The Journal of Finance* **66**(4) 1251–1290.

Grinblatt, Mark, Matti Keloharju, Juhani Linnainmaa. 2011. IQ and stock market participation. *The Journal of Finance* **66**(6) 2121–2164.

Grinblatt, Mark, Matti Keloharju, Juhani T. Linnainmaa. 2012. IQ, trading behavior, and performance. *Journal of Financial Economics* **104**(2) 339–362.

Gürkaynak, Refet S. 2008. Econometric tests of asset price bubbles: taking stock. *Journal of Economic Surveys* **22**(1) 166–186.

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.

Harris, Tammy, James W. Hardin, et al. 2013. Exact Wilcoxon signed-rank and Wilcoxon Mann–Whitney ranksum tests. *Stata Journal* **13**(2) 337–343.

Harrison, Michael, David Kreps. 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics* **92**(2) 323–336.

Haruvy, Ernan, Yaron Lahav, Charles N. Noussair. 2007. Traders' expectations in asset markets: experimental evidence. *American Economic Review* **97**(5) 1901–1920.

Haruvy, Ernan, Charles N. Noussair. 2006. The effect of short selling on bubbles and crashes in experimental spot asset markets. *The Journal of Finance* **61**(3) 1119–1157.

Holt, Charles A., Megan Porzio, Michelle Yingze Song. 2017. Price bubbles, gender, and expectations in experimental asset markets. *European Economic Review* **100** 72–94.

Huber, Jürgen, Michael Kirchler. 2012. The impact of instructions and procedure on reducing confusion and bubbles in experimental asset markets. *Experimental Economics* **15**(1) 89–105.

Janssen, Dirk-Jan, Sascha Füllbrunn, Utz Weitzel. 2018. Individual speculative behavior and overpricing in experimental asset markets. *Experimental Economics* **forthcoming**.

Jimenez, Natalia, Ismael Rodriguez-Lara, Jean-Robert Tyran, Erik Wengström. 2018. Thinking fast, thinking badly. *Economics Letters* **162** 41–44.

Kahneman, Daniel. 2011. *Thinking, fast and slow*. Farrar, Straus and Giroux.

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.

Kindleberger, Charles P., Robert Z. Aliber. 2011. *Manias, Panics, and Crashes: A History of Financial Crises*. 6th ed. Palgrave Macmillan.

King, Ronald, Vernon L. Smith, Arlington W. Williams, Mark Van Boening. 1993. The robustness of bubbles and crashes in experimental stock markets. Richard Day, Ping Chen, eds., *Nonlinear Dynamics and Evolutionary Economics*. Oxford University Press, 183–200.

Kirchler, Michael, Caroline Bonn, Jürgen Huber, Michael Razen. 2015. The "inflow-effect" – Trader inflow and price efficiency. *European Economic Review* **77** 1–19.

Kirchler, Michael, Jürgen Huber, Thomas Stöckl. 2012. Thar she bursts: Reducing confusion reduces bubbles. *American Economic Review* **102**(2) 865–883.

Kirchler, Michael, Florian Lindner, Utz Weitzel. 2018. Rankings and risk-taking in the finance industry. *The Journal of Finance* **forthcoming**.

Kleinlercher, Daniel, Jürgen Huber, Michael Kirchler. 2014. The impact of different incentive schemes on asset prices. *European Economic Review* **68** 137–150.

Lei, Vivian, Charles N. Noussair, Charles R. Plott. 2001. Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. actual irrationality. *Econometrica* **69**(4) 831–859.

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

Mackintosh, N.J. 2011. History of theories and measurement of intelligence. Robert J. Sternberg, Scott Barry Kaufman, eds., *The Cambridge Handbook of Intelligence*. Cambridge University Press, 3–19.

Miller, Edward M. 1977. Risk, uncertainty, and divergence of opinion. *The Journal of Finance* **32**(4) 1151–1168.

Noussair, Charles N., Stephane Robin, Bernard Ruffieux. 2001. Price bubbles in laboratory asset markets with constant fundamental values. *Experimental Economics* **4**(1) 87–105.

Noussair, Charles N., Steven Tucker. 2016. Cash inflows and bubbles in asset markets with constant fundamental values. *Economic Inquiry* **54**(3) 1596–1606.

Noussair, Charles N., Steven Tucker, Yilong Xu. 2014. A futures market reduces bubbles but allows greater profit for more sophisticated traders. Working Paper.

Noussair, Charles N, Steven Tucker, Yilong Xu. 2016. Futures markets, cognitive ability, and mispricing in experimental asset markets. *Journal of Economic Behavior & Organization* **130** 166–179.

Oechssler, Jörg, Andreas Roider, Patrick W Schmitz. 2009. Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization* **72**(1) 147–152.

Ofek, Eli, Matthew Richardson. 2003. DotCom Mania: The rise and fall of internet stock prices. *The Journal of Finance* **58**(3) 1113–1138.

Palan, Stefan. 2015. GIMS – Software for asset market experiments. *Journal of Behavioral and Experimental Finance* **5** 1–14.

Raven, John. 2000. The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology* **41**(1) 1–48.

Razen, Michael, Jürgen Huber, Michael Kirchler. 2017. Cash inflow and trading horizon in asset markets. *European Economic Review* **92** 359–384.

Robin, Stéphane, Katerina Strázniczká, Marie C. Villeval. 2012. Bubbles and incentives: An experiment on asset markets. Working paper.

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

Scherbina, Anna, Bernd Schlusche. 2014. Asset price bubbles: a survey. *Quantitative Finance* **14**(4) 589–604.

Smith, Alec, Terry Lohrenz, Justin King, Read Montague, Colin Camerer. 2014. Irrational exuberance and neural crash warning signals during endogenous experimental market bubbles. *Proceedings of the National Academy of Sciences* **111**(29) 10503–10508.

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.

Smith, Vernon L., Mark Van Boening, Charissa P. Wellford. 2000. Dividend timing and behavior in laboratory asset markets. *Economic Theory* **16**(3) 567–583.

Stöckl, Thomas, Jürgen Huber, Michael Kirchler. 2010. Bubble measures in experimental asset markets. *Experimental Economics* **13**(3) 284–298.

Sutter, Matthias, Jürgen Huber, Michael Kirchler. 2012. Bubbles and information: An experiment. *Management Science* **58**(2) 384–393.

Toplak, Maggie E., Richard F. West, Keith E. Stanovich. 2011. The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition* **39**(7) 1275–1289.

Toplak, Maggie E., Richard F. West, Keith E. Stanovich. 2014. Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning* **20**(2) 147–168.

Xiong, Wei, Jialin Yu. 2011. The Chinese Warrants Bubble. *American Economic Review* **101**(6) 2723–2753.

Online Appendix

A Additional Figures and Tables

Table A1: **Treatment parameterization in both experiments**: This table outlines model parameters across the different treatments. Similar to Smith et al. (2014) and Holt et al. (2017), assets pay dividends of either 1.2 or 1.6 Taler with equal probability at the end of each period. Additionally, interest of 5% is paid on cash holdings at the end of a period but before dividends are added. The expected dividend return is equal to the interest rate on cash at 5% (1.4 divided by 28) and therefore the asset's risk-neutral fundamental value FV is constant at 28 in all periods. An additional income of 100 Taler from an exogenous source is paid to each subject at the beginning of each period in two treatments. CA-Ratio stands for the cash to asset-value ratio in the respective periods 1/10/20 (i.e., total cash divided by the product of the number of shares outstanding and the FV of 28).

Treatment	INC	SHORT	LOW	HIGH
Number of periods	20	20	20	20
Number of traders per market	8	8	8	8
Buyback price	28 Taler	28	28	28
Interest rate	5%	5%	5%	5%
Possible dividends per period	1.20 Taler; 1.60 Taler	1.20; 1.60	1.20; 1.60	1.20; 1.60
Probabilities of the dividends	50%; 50%	50%; 50%	50%; 50%	50%; 50%
Initial endowment (subject)	20 shares / 560 Taler	20 / 560	20 / 560	20 / 5,700
Additional cash per period (subject)	100 Taler	100	0	0
CA-Ratio (Periods 1/10/20)	1.0/4.1/10.2	1.0/4.1/10.2	1.0/1.0/1.0	10.2/10.2/10.2
Short-selling capacity (subject)	0 shares	–40	0	0

Table A2: **Bubble measures for all markets in Treatment INC**: This table outlines bubbles measures, such as RD (relative deviation of prices to fundamentals, normalized at the FV of 28) and RAD (relative absolute deviation of prices to fundamentals, normalized at the FV of 28), measuring overpricing and mispricing, respectively (Stöckl et al., 2010). RDMAX depicts overpricing at the peak period price, AMPLITUDE measures the difference from the pre-peak minimum to the maximum period price as a percentage of FV and CRASH calculates the difference between the minimum period price after the peak and the peak average price normalized at the FV to learn about the severity of crashes. RDMAX, AMPLITUDE, and CRASH are taken from Razen et al. (2017). Threshold values for bubble measures qualifying a bubble according to the criteria outlined in Section 2.2 in the main text are displayed in the first row. Values marked with * exceed the corresponding threshold. Markets in **bold** and marked with ** are classified as bubble markets.

Market	RAD	RD	RD_MAX	AMLITUDE	CRASH
Threshold for bubble classification			94.5	44.8	−98.8
Experiment PROF					
M1	9.2	9.2	14.1	1.0	−14.2
M2	13.4	11.7	34.8	49.8*	−28.7
M3	53.7	53.7	64.9	58.3*	−18.6
M4	9.6	9.6	28.6	28.6	−28.6
M5	12.8	12.8	28.1	28.1	−28.1
M6	10.4	10.1	48.5	51.0*	−51.0
M7	10.3	10.2	23.4	24.2	−17.9
M8**	233.0	233.0	473.5*	303.9*	−465.5*
M9	16.7	16.7	35.7	33.3	−35.2
M10	15.3	15.3	29.6	23.1	−29.1
Median	13.4	12.8	34.8	33.3	−28.7
Experiment STUD					
M1**	229.3	229.3	632.3*	601.8*	−625.5*
M2**	59.1	59.1	159.5*	147.0*	−156.0*
M3	44.9	44.9	72.9	44.3	−69.6
M4	47.7	47.7	74.9	66.6*	−74.3
M5	28.4	28.4	59.4	38.9	−55.4
M6**	48.2	48.2	105.6*	87.4*	−101.2*
M7	3.8	2.8	4.8	2.6	−12.9
M8**	51.1	51.1	124.9*	114.6*	−123.0*
M9	31.5	31.5	46.3	38.8	−44.8
M10**	80.7	80.7	376.5*	366.4*	−374.0*
M11	17.0	17.0	21.0	0.0	−20.4
M12**	96.9	96.9	223.2*	203.5*	−221.1*
Median	47.3	47.3	74.9	66.6	−74.3

Table A3: **Bubble measures for all markets in Treatment SHORT**: This table outlines bubbles measures, such as RD (relative deviation of prices to fundamentals, normalized at the FV of 28) and RAD (relative absolute deviation of prices to fundamentals, normalized at the FV of 28), measuring overpricing and mispricing, respectively (Stöckl et al., 2010). RDMAX depicts overpricing at the peak period price, AMPLITUDE measures the difference from the pre-peak minimum to the maximum period price as a percentage of FV and CRASH calculates the difference between the minimum period price after the peak and the peak average price normalized at the FV to learn about the severity of crashes. RDMAX, AMPLITUDE, and CRASH are taken from Razen et al. (2017). Threshold values for bubble measures qualifying a bubble according to the criteria outlined in Section 2.2 in the main text are displayed in the first row. Values marked with * exceed the corresponding threshold. Markets in **bold** and marked with ** are classified as bubble markets.

Market	RAD	RD	RD_MAX	AMLITUDE	CRASH
Threshold for bubble classification			94.5	44.8	−98.8
Experiment PROF					
M1	0.0	0.0	0.2	0.0	−0.2
M2	9.6	9.1	45.9	49.1*	−42.3
M3	3.5	3.2	7.2	10.5	−5.9
M4	11.9	11.9	22.2	16.8	−19.2
M5	0.6	0.5	1.9	1.3	−2.8
M6	6.5	6.1	9.3	0.0	−13.2
M7	12.4	12.4	33.5	30.7	−29.8
M8	39.6	39.6	62.7	56.2*	−43.6
M9	3.1	3.1	6.9	3.1	−6.9
Median	6.5	6.1	9.3	10.5	−13.2
Experiment STUD					
M1	13.6	12.6	38.5	47.6*	−47.6
M2	1.7	1.3	3.1	5.6	−1.4
M3	17.0	17.0	38.7	0.0	−27.5
M4	2.4	2.2	4.7	5.6	−2.9
M5	19.1	17.3	67.4	84.1*	−68.8
M6	8.3	8.3	17.3	15.7	−16.1
M7	10.0	10.0	17.4	4.2	−14.2
M8	27.4	27.4	47.4	0.0	−45.7
M9	21.7	21.7	60.9	49.6*	−60.1
M10	7.2	7.2	14.9	11.4	−11.4
M11	35.1	35.1	50.5	0.0	−22.2
M12**	181.0	181.0	497.5*	491.2*	−492.5*
Median	13.6	12.6	38.5	11.4	−22.2

Table A4: **Bubble measures for all markets in Treatment LOW**: This table outlines bubbles measures, such as RD (relative deviation of prices to fundamentals, normalized at the FV of 28) and RAD (relative absolute deviation of prices to fundamentals, normalized at the FV of 28), measuring overpricing and mispricing, respectively (Stöckl et al., 2010). RDMAX depicts overpricing at the peak period price, AMPLITUDE measures the difference from the pre-peak minimum to the maximum period price as a percentage of FV and CRASH calculates the difference between the minimum period price after the peak and the peak average price normalized at the FV to learn about the severity of crashes. RDMAX, AMPLITUDE, and CRASH are taken from Razen et al. (2017). Threshold values for bubble measures qualifying a bubble according to the criteria outlined in Section 2.2 in the main text are displayed in the first row. Values marked with * exceed the corresponding threshold. Markets in **bold** and marked with ** are classified as bubble markets.

Market	RAD	RD	RD_MAX	AMLITUDE	CRASH
Threshold for bubble classification			94.5	44.8	−98.8
Experiment PROF					
M1	9.5	9.5	18.6	16.4	−17.8
M2	2.3	1.8	4.0	0.0	−8.0
M3	2.8	2.8	7.3	2.3	−7.3
M4	4.5	−4.1	4.1	43.1	−8.8
M5	5.3	5.3	9.2	9.2	−9.2
M6	25.0	24.5	50.2	27.4	−53.4
M7	4.2	3.4	7.6	6.1	−15.7
M8	3.3	1.6	10.5	0.0	−18.7
M9	37.7	37.5	75.6	11.8	−77.5
Median	4.5	3.4	9.2	9.2	−15.7
Experiment STUD					
M1	25.9	25.8	45.4	0.0	−46.3
M2	6.0	−3.2	14.4	31.7	−58.5
M3	13.7	−4.1	64.4	0.0	−82.5
M4	2.4	2.3	4.5	3.4	−6.0
M5	38.4	38.2	79.5	10.5	−80.3
M6	9.1	9.0	27.8	9.2	−28.1
M7	26.1	25.9	101.3*	0.0	−103.3
M8	12.2	12.0	28.6	0.0	−29.9
M9	8.4	8.3	78.6	78.8*	−78.8
M10	16.5	16.5	21.1	6.8	−21.1
M11	34.8	34.8	107.9*	0.0	−104.8*
M12	9.8	9.8	15.6	0.0	−8.4
Median	13.7	12.0	28.6	0.0	−46.3

Table A5: **Bubble measures for all markets in Treatment HIGH**: This table outlines bubbles measures, such as RD (relative deviation of prices to fundamentals, normalized at the FV of 28) and RAD (relative absolute deviation of prices to fundamentals, normalized at the FV of 28), measuring overpricing and mispricing, respectively (Stöckl et al., 2010). RDMAX depicts overpricing at the peak period price, AMPLITUDE measures the difference from the pre-peak minimum to the maximum period price as a percentage of FV and CRASH calculates the difference between the minimum period price after the peak and the peak average price normalized at the FV to learn about the severity of crashes. RDMAX, AMPLITUDE, and CRASH are taken from Razen et al. (2017). Threshold values for bubble measures qualifying a bubble according to the criteria outlined in Section 2.2 in the main text are displayed in the first row. Values marked with * exceed the corresponding threshold. Markets in **bold** and marked with ** are classified as bubble markets.

Market	RAD	RD	RD_MAX	AMLITUDE	CRASH
Threshold for bubble classification			94.5	44.8	−98.8
Experiment PROF					
M1	58.4	58.4	101.7*	41.8	−99.1*
M2	13.1	9.3	25.2	52.5*	−19.7
M3**	1336.7	1336.7	2215.3*	2186.8*	−1450.3*
M4**	114.6	114.6	239.3*	225.3*	−233.1*
M5	8.6	7.7	15.3	23.6	−13.2
M6	8.6	7.6	16.7	27.1	−15.3
M7	33.3	31.5	60.7	69.9*	−27.0
M8**	120.7	120.7	207.5*	191.1*	−198.3*
M9	30.5	30.4	55.4	43.8	−55.8
M10**	71.0	71.0	114.7*	57.2*	−111.1*
Median	58.4	58.4	101.7	54.9	−99.1
Experiment STUD					
M1	37.7	37.7	60.4	59.4*	−60.5
M2**	869.8	869.8	1732.1*	1524.9*	−1729.2*
M3**	782.9	782.7	1400.0*	182.1*	−1401.1*
M4	4.1	4.1	27.0	0.0	−26.5
M5	282.7	282.6	576.1*	0.0	−576.7*
M6	342.3	342.3	495.5*	441.4*	−89.9
M7**	1478.2	1478.2	7727.4*	7674.3*	−7723.8*
M8**	1118.9	1118.9	1296.2*	75.5*	−1244.6*
M9**	106.2	106.2	158.9*	82.8*	−158.9*
M10**	394.2	394.2	831.3*	147.6*	−827.3*
M11**	138.9	138.0	261.1*	227.1*	−269.0*
M12**	165.5	165.3	286.2*	86.9*	−286.9*
Median	342.3	342.3	576.1	86.9	−576.7

Table A6: **Pairwise Mann-Whitney U-tests of mispricing (RAD), overpricing (RD), maximum overpricing (RDMAX), price run-ups (AMPLITUDE), and crash (CRASH) in Experiment PROF**: This table shows pairwise treatment comparisons for treatments INC (increasing CA-Ratio), SHORT (increasing CA-Ratio, short-selling allowed), LOW (low and constant CA-Ratio of 1), and HIGH (high and constant CA-Ratio of 10.2) in Experiment PROF. The numbers identify the difference in the treatment medians in percentage points, i.e., the value of the "row" treatment minus the value of the "column" treatment (a positive value implies that, for instance, INC is larger than SHORT). The numbers in parentheses show the Z-value of the MW U-test statistic. *, ** and *** represent the 10%, 5%, and 1% significance levels of a double-sided test. Sample size N for each test is between 18 and 20.

	RAD			**RD**			**RDMAX**		
Treatment	SHORT	LOW	HIGH	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	6.46**	7.97**	−52.84	6.19**	9.95**	−50.35	18.36**	25.58**	−79.93
	(2.37)	(2.29)	(−1.36)	(2.37)	(2.29)	(−0.91)	(1.96)	(2.12)	(−1.21)
SHORT	.	1.51	−59.30***	.	3.76	−56.54***	.	7.22	−98.29**
	.	(0.22)	(2.86)	.	(−0.75)	(2.61)	.	(0.13)	(2.53)
LOW	.	.	−60.81***	.	.	−60.30***	.	.	−105.51**
	.	.	(−2.86)	.	.	(−2.78)	.	.	(−2.53)

	AMPLITUDE			**CRASH**		
Treatment	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	16.54*	24.12**	−23.88	−9.45	16.24	117.49
	(1.88)	(2.04)	(−1.44)	(−1.55)	(−0.33)	(1.21)
SHORT	.	7.58	−40.42***	.	25.69	126.94**
	.	(−0.66)	(2.78)	.	(−1.19)	(−2.45)
LOW	.	.	−48.00***	.	.	101.25*
	.	.	(−2.86)	.	.	(1.71)

Table A7: **Pairwise Mann-Whitney U-tests of mispricing (RAD), overpricing (RD), maximum overpricing (RDMAX), price run-ups (AMPLITUDE), and crash (CRASH) in Experiment STUD**: This table shows pairwise treatment comparisons for treatments INC (increasing CA-Ratio), SHORT (increasing CA-Ratio, short-selling allowed), LOW (low and constant CA-Ratio of 1), and HIGH (high and constant CA-Ratio of 10.2) in Experiment STUD. The numbers identify the difference in the treatment medians in percentage points, i.e., the value of the "row" treatment minus the value of the "column" treatment (a positive value implies that, for instance, INC is larger than SHORT). The numbers in parentheses show the Z-value of the MW U-test statistic. *, ** and *** represent the 10%, 5%, and 1% significance levels of a double-sided test. Sample size N for each test is 24.

	RAD			**RD**			**RDMAX**		
Treatment	SHORT	LOW	HIGH	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	33.74***	33.58***	−294.96***	34.71***	35.35***	−294.96***	36.43**	46.36*	−424.96**
	(2.66)	(3.06)	(−2.89)	(2.71)	(3.18)	(−2.89)	(2.31)	(1.91)	(−2.54)
SHORT	.	−0.16	−328.69***	.	0.64	−329.66***	.	9.92	−461.39**
	.	(−0.12)	(3.35)	.	(−0.35)	(3.35)	.	(0.46)	(3.35)
LOW	.	.	−328.54***	.	.	−330.30***	.	.	−471.32***
	.	.	(−3.46)	.	.	(−3.58)	.	.	(−3.41)

	AMPLITUDE			**CRASH**		
Treatment	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	50.89*	63.26***	−80.97	−26.76*	−20.34	426.13**
	(1.71)	(3.00)	(−1.53)	(−1.85)	(−1.33)	(2.25)
SHORT	.	12.36	−131.86***	.	6.42	452.90***
	.	(−1.21)	(2.81)	.	(−1.15)	(−3.23)
LOW	.	.	−144.23***	.	.	446.47***
	.	.	(−3.58)	.	.	(3.12)

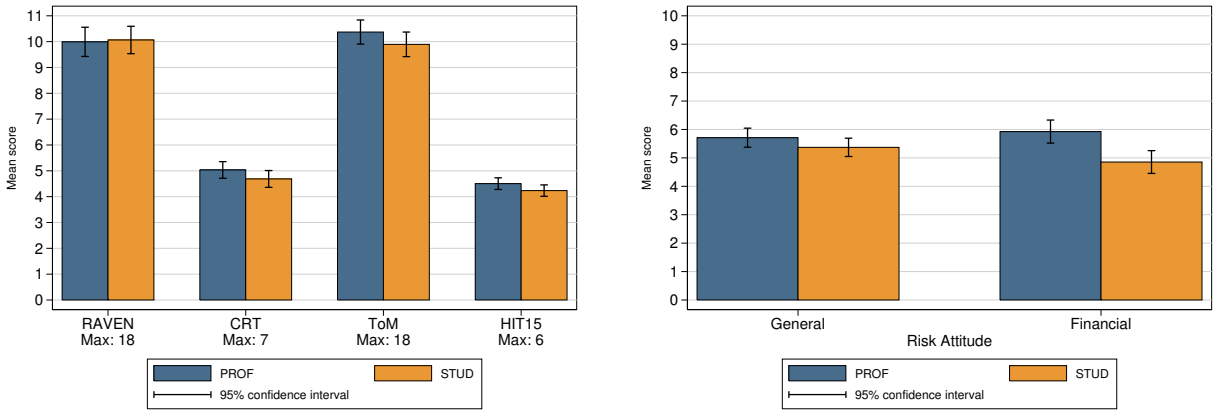


Figure A1: **Differences in cognitive skills (left panel) and risk attitudes (right panel) between professionals (PROF) and students (STUD):** The left panel of this figure depicts mean test scores of fluid intelligence (RAVEN: Raven's advanced progressive matrices), cognitive reflection (CRT), theory of mind (ToM: eye-gaze test), and backward induction skills (HIT15). The right panel shows risk attitudes taken from two survey questions concerning general risk taking (General) and financial risk taking (Financial) from the German Socio-Economic Panel SOEP. Fluid intelligence (RAVEN) measures the capacity to reason and solve novel problems and is necessary for logical problem solving. It is a nonverbal test typically used as an IQ-test. Cognitive reflection (CRT) adds to fluid intelligence because it helps individuals avoid commonly-observed heuristics and biases and measures the ability to engage in effortful reasoning. Questions of CRT have an obvious (intuitive but incorrect) response. The correct response requires effortful reasoning. ToM defines one's capacity to infer others' intentions. In the associated eye gaze test subjects look at images of people's eyes and chose one of four feelings that best describe the mental state of the person whose eyes are shown best. Abilities to engage in backward induction are measured by the The HIT15 test. The computer and the subject take turns in adding points (1, 2, or 3) to a basket with the goal to be the first player to reach 15 points. 95% confidence intervals are displayed for each bar. The maximum score in the tests was 18 (RAVEN and ToM), 7 (CRT), and 6 (HIT15).

B Individual Price Graphs

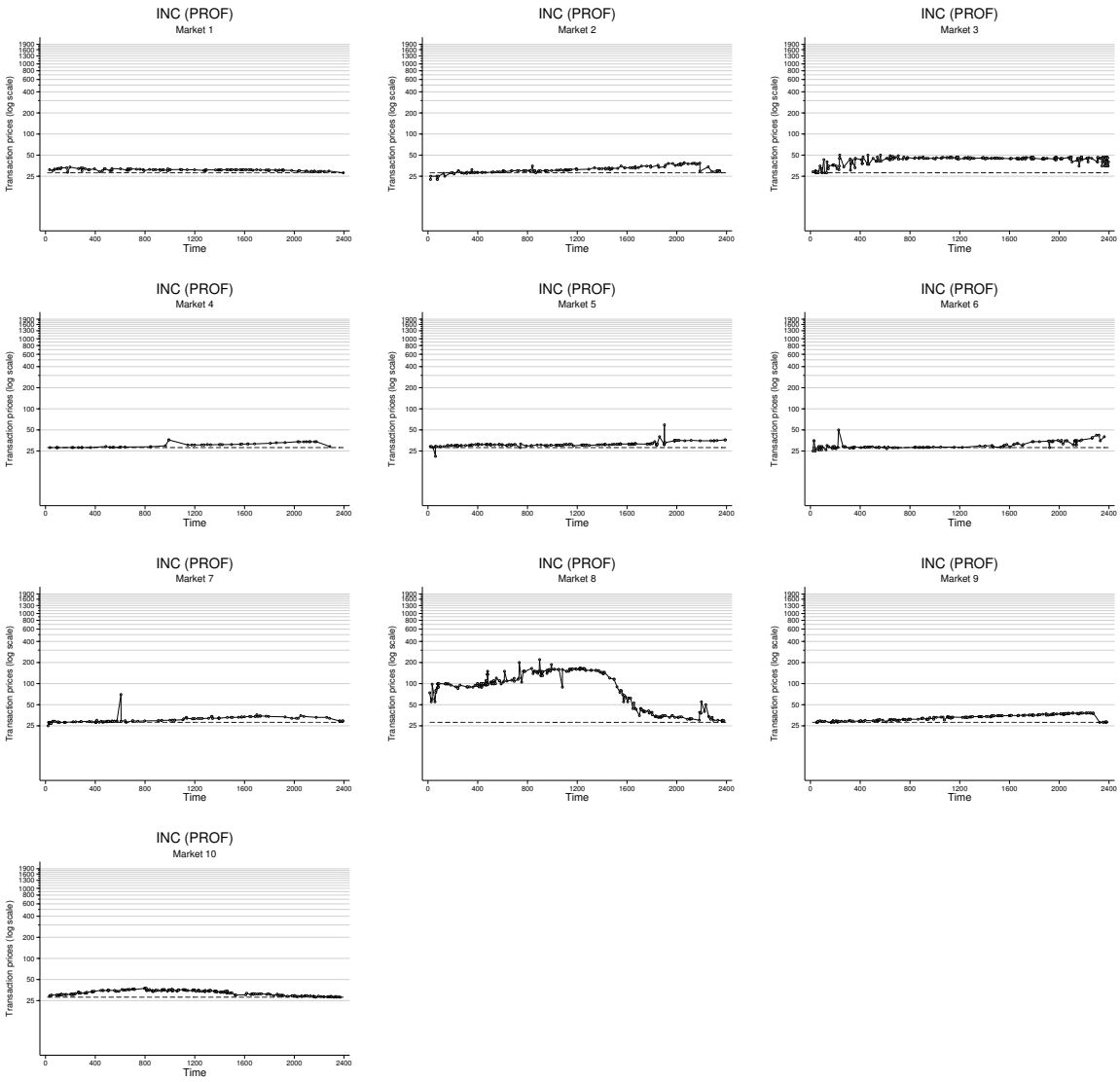


Figure B2: **Individual transaction prices for each market of Treatment INC in Experiment PROF**: The dashed line represents the risk-neutral fundamental value of 28. In Market 6, one trade with price < 5 was dropped for presentation purposes.

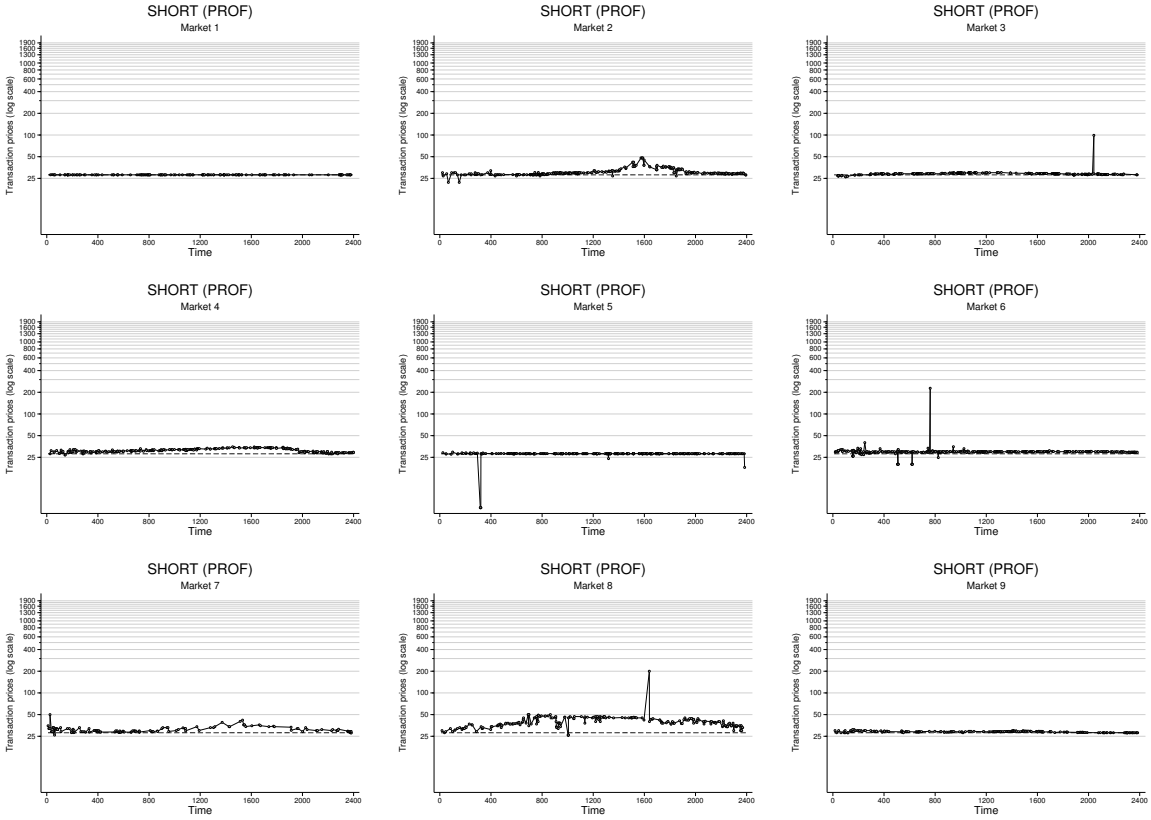


Figure B3: **Individual transaction prices for each market of Treatment SHORT in Experiment PROF**: The dashed line represents the risk-neutral fundamental value of 28.

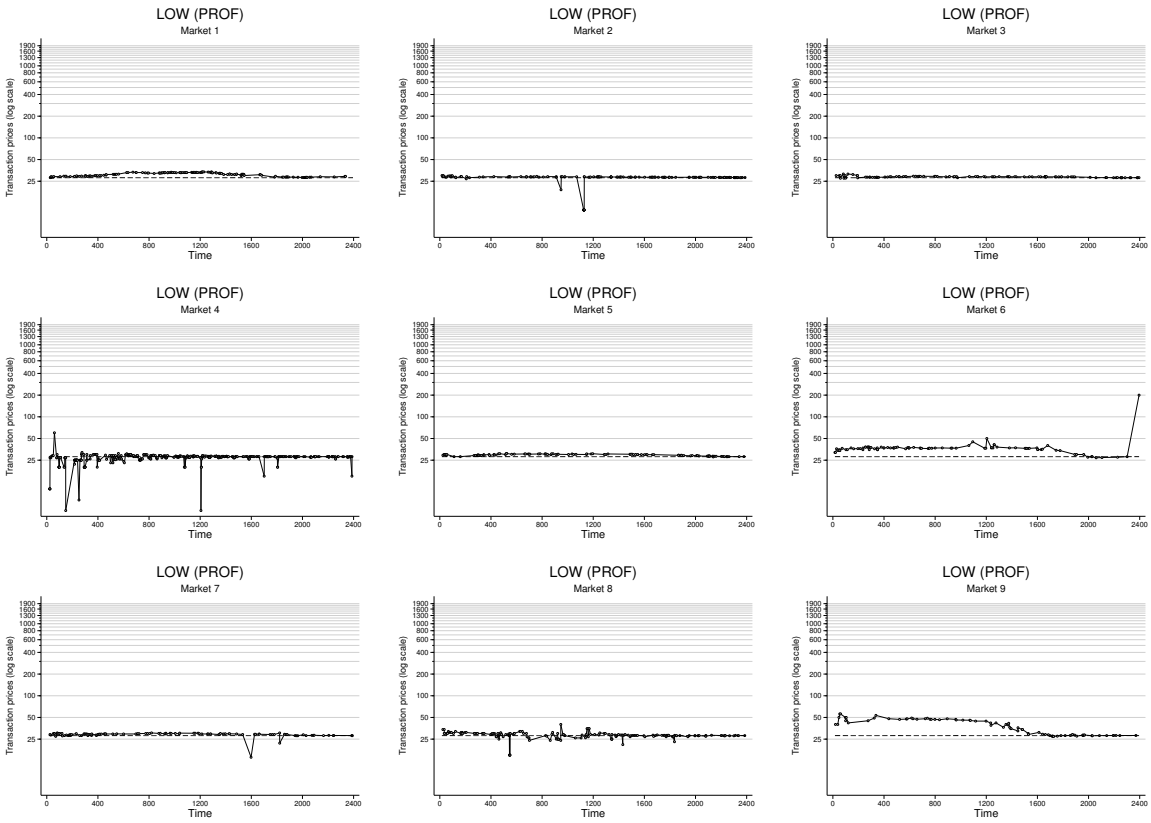


Figure B4: **Individual transaction prices for each market of Treatment LOW in Experiment PROF**: The dashed line represents the risk-neutral fundamental value of 28. In Markets 4, 7, and 8, ten, one, and seven trades, respectively, with prices < 5 were dropped for presentation purposes.

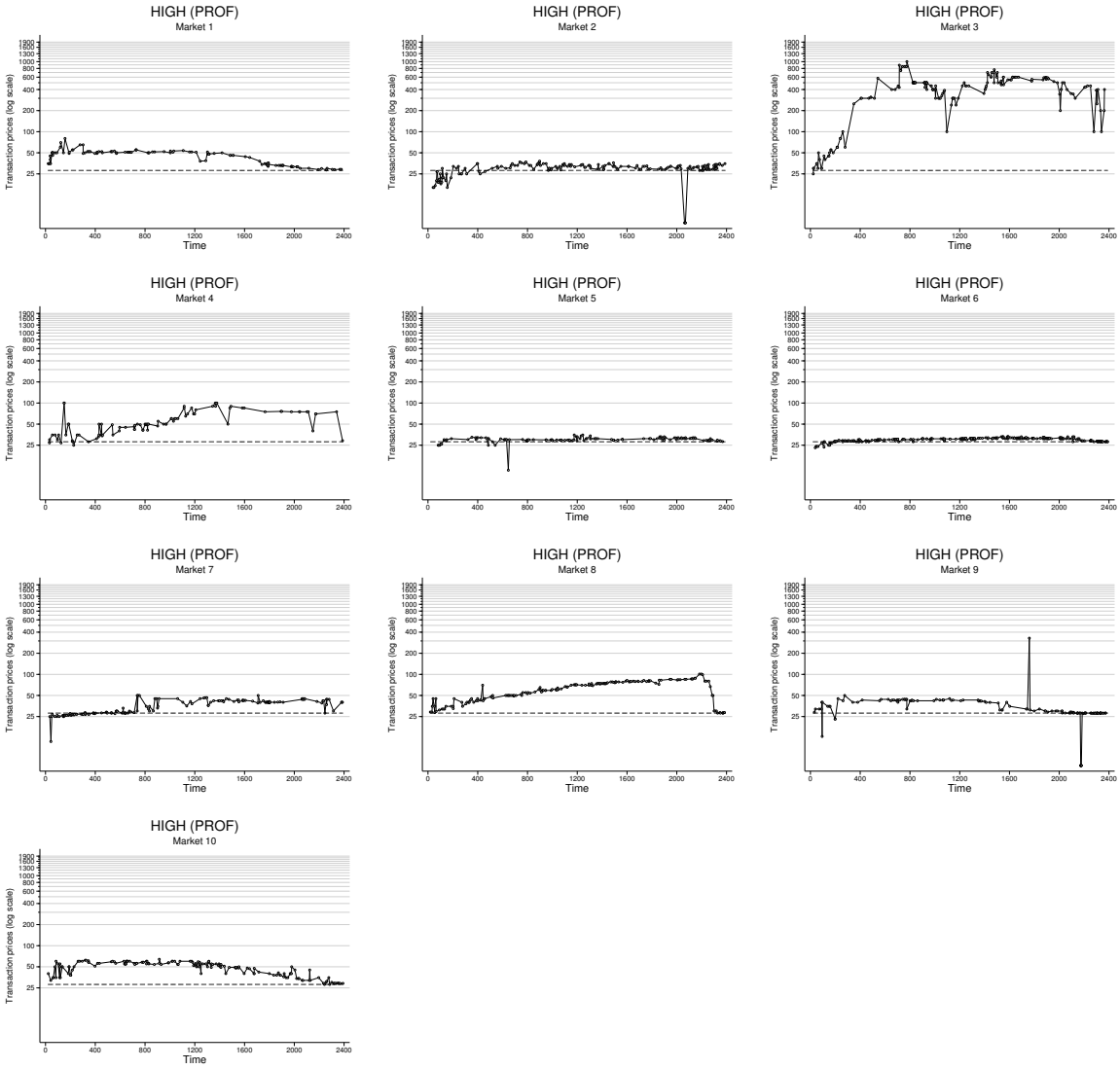


Figure B5: **Individual transaction prices for each market of Treatment HIGH in Experiment PROF**: The dashed line represents the risk-neutral fundamental value of 28. In Markets 1 and 4, four and one trades, respectively, with prices < 5 were dropped for presentation purposes.

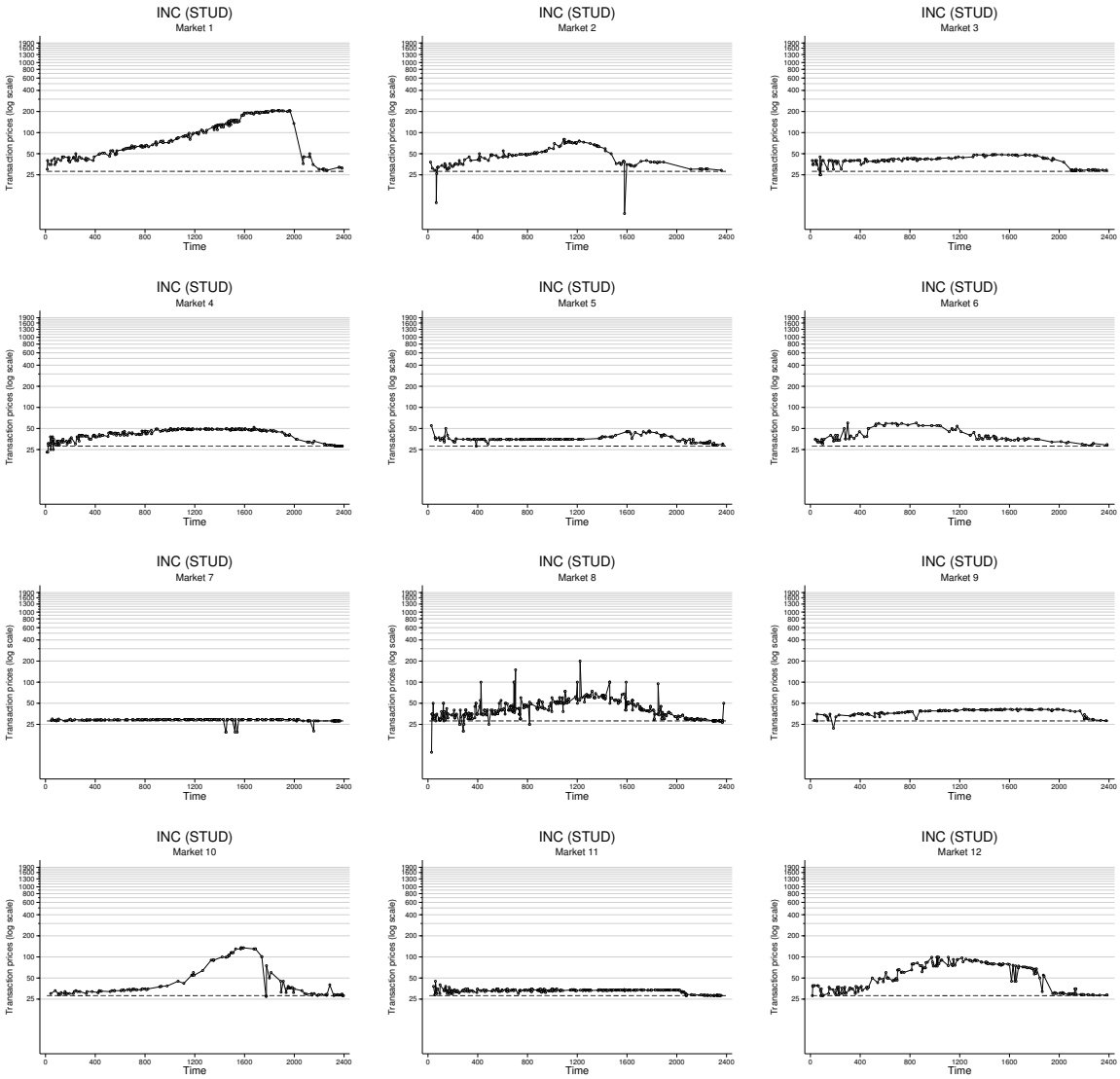


Figure B6: **Individual transaction prices for each market of Treatment INC in Experiment STUD**: The dashed line represents the risk-neutral fundamental value of 28. In Market 4, one trade with price < 5 was dropped for presentation purposes.

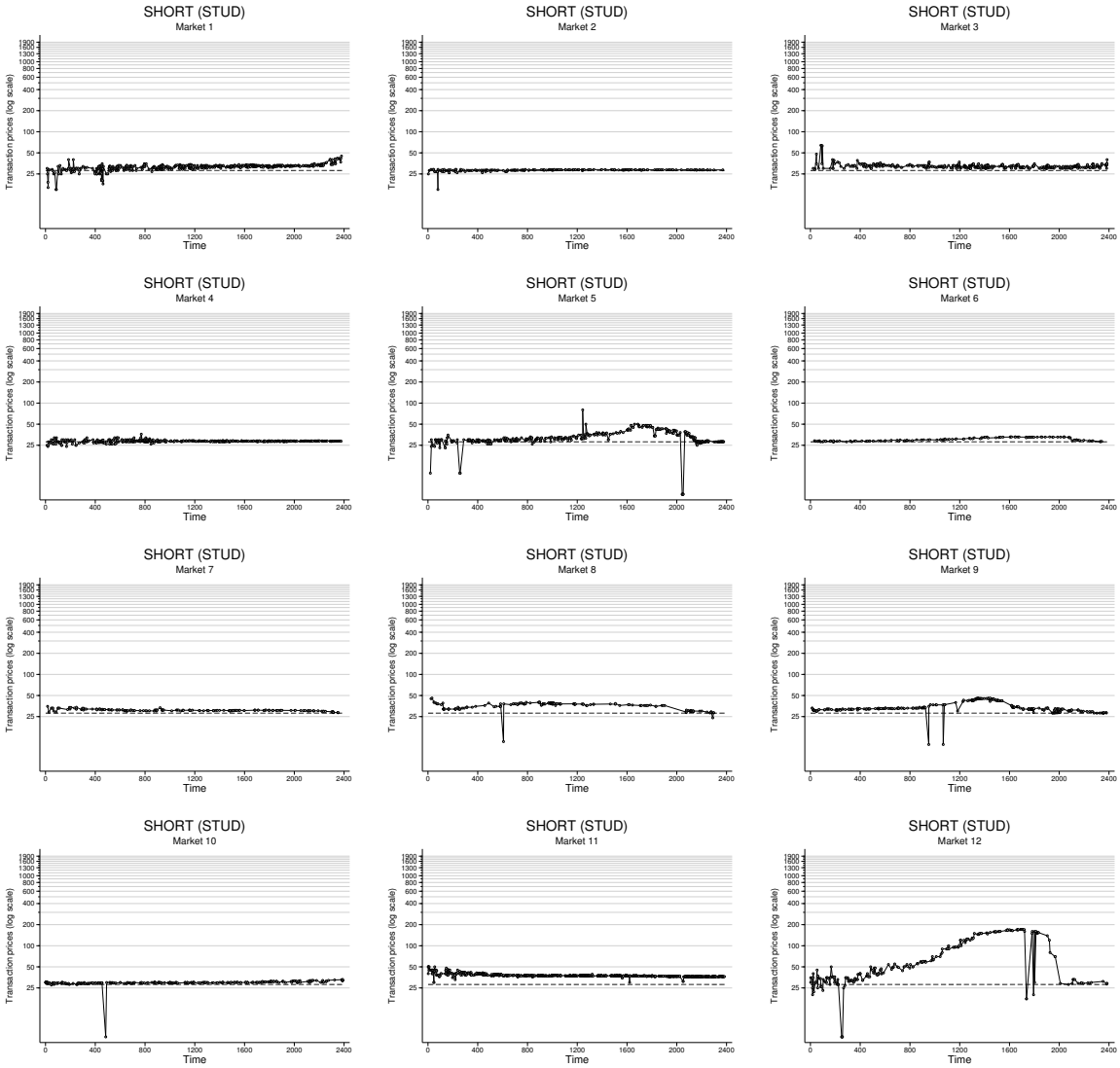


Figure B7: **Individual transaction prices for each market of Treatment SHORT in Experiment STUD**: The dashed line represents the risk-neutral fundamental value of 28. In Market 6, four trades with prices < 5 were dropped for presentation purposes.

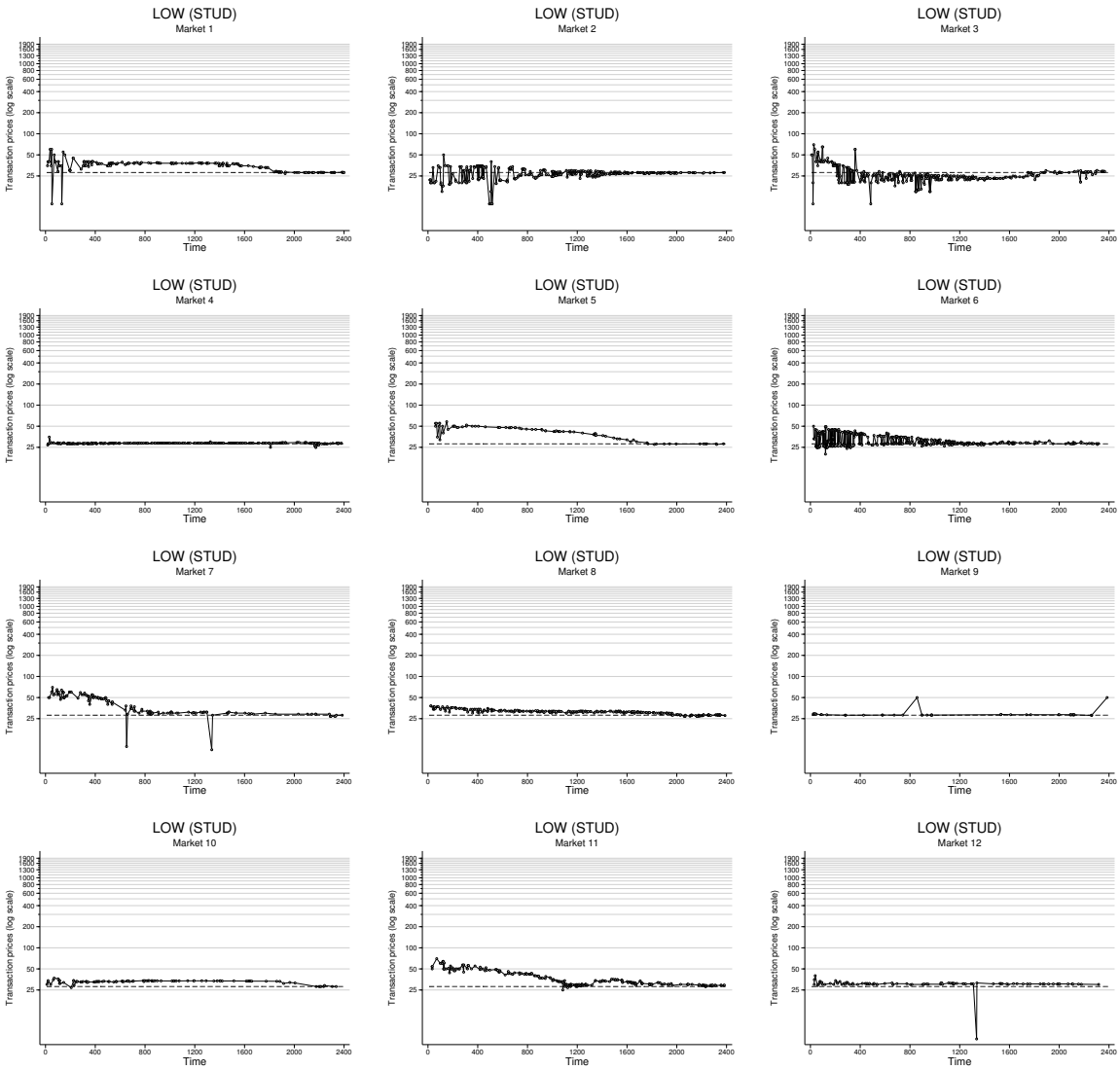


Figure B8: **Individual transaction prices for each market of Treatment LOW in Experiment STUD**: The dashed line represents the risk-neutral fundamental value of 28.

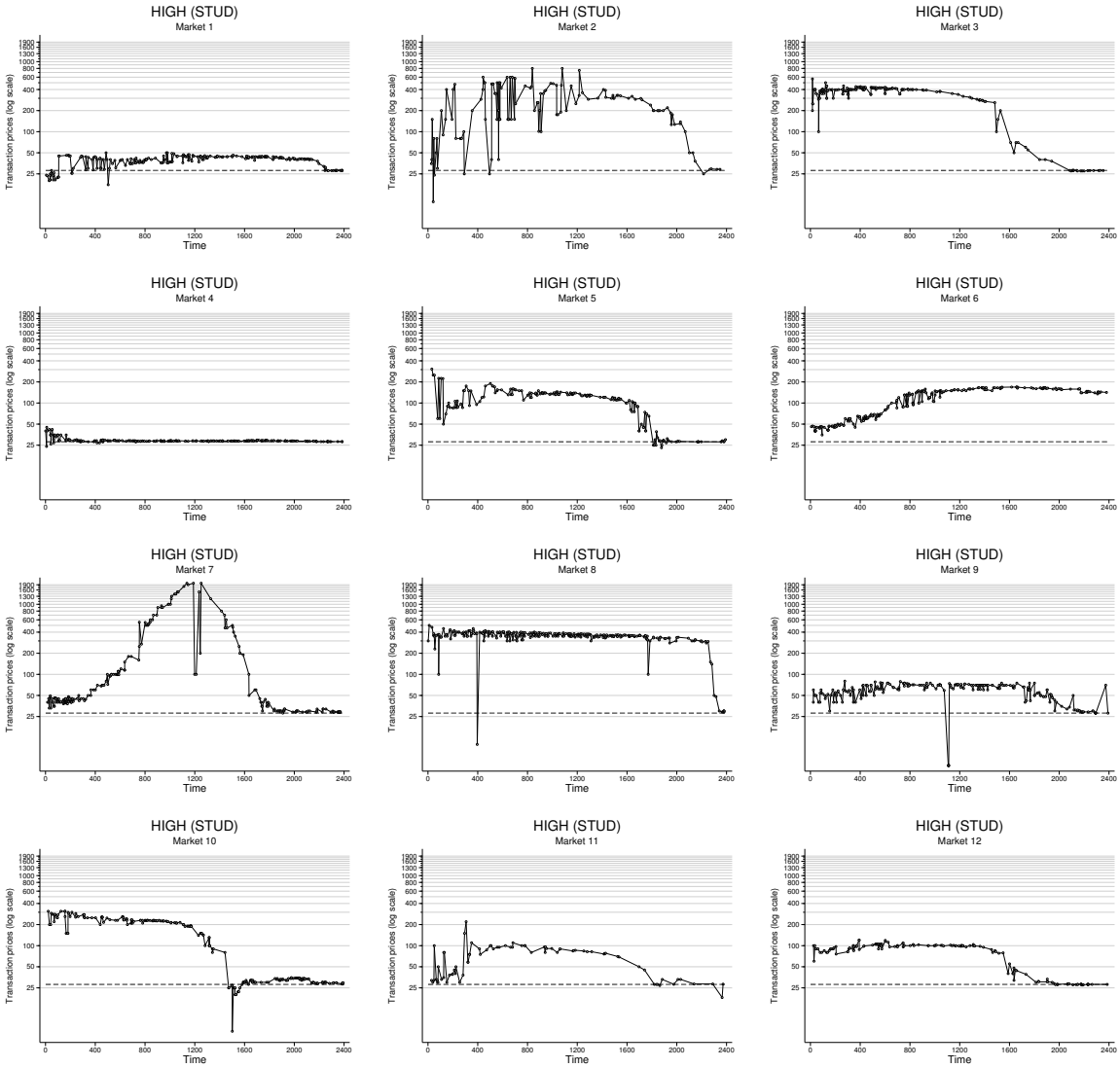


Figure B9: **Individual transaction prices for each market of Treatment HIGH in Experiment STUD**: The dashed line represents the risk-neutral fundamental value of 28. In Markets 7, 9, and 10, six, one, and three trades, respectively, with prices < 5 or > 2000 were dropped for presentation purposes.

C Price Beliefs

Figure C10 provides a descriptive overview of median treatment prices and median price forecasts for periods up to $t + 2$, elicited in t. Table C8 provides statistical tests between professionals and students, separated for treatments and periods before/after price peaks. For the statistical tests we pool across subjects and periods, separately for the periods before and after the price peak to arrive at two values for each market.

The upper part of Table C8, referring to forecasts before price peaks (upswings), shows no systematic patterns in forecast accuracy between professionals and students as most pairwise comparisons turn out to be insignificant. In particular, forecasts are very accurate in the bubble-moderator treatments SHORT and LOW which is not surprising given their high level of efficiency. In contrast, real market prices exceed price beliefs by 5.0 (INC) to 14.7 (HIGH) percent in the bubble-driver treatments in the professional sample (across all forecasting periods t, t to $t, t + k$). The patterns are similar in the student sample with respective forecasts underestimating actual prices by 3.3 (INC) to 13.8 (HIGH) percent. The z-statistics show that professionals and students underestimate price upswings in both treatments to a similar extent.

In contrast, professionals predict price downswings very accurately in all treatments with price realizations that are remarkably close to price beliefs (within the range of 0.3 to –2.1 percent for all beliefs up to $t + 2$). Student markets in the bubble-driver treatments INC and HIGH, however, show significantly higher forecast errors as market prices drop below price beliefs between –4.8 and –26.7 percent across all prediction intervals.[23]

[23] Pairwise Mann-Whitney U-tests of forecast errors for periods $t + k$ with $k = 0, 1, 2$ within each subject pool show similar patterns across treatments. For professionals and students we find that price forecasts are significantly more inaccurate in the bubble-driver treatments compared to the bubble-moderator treatments before the price peak. Results can be provided upon request.

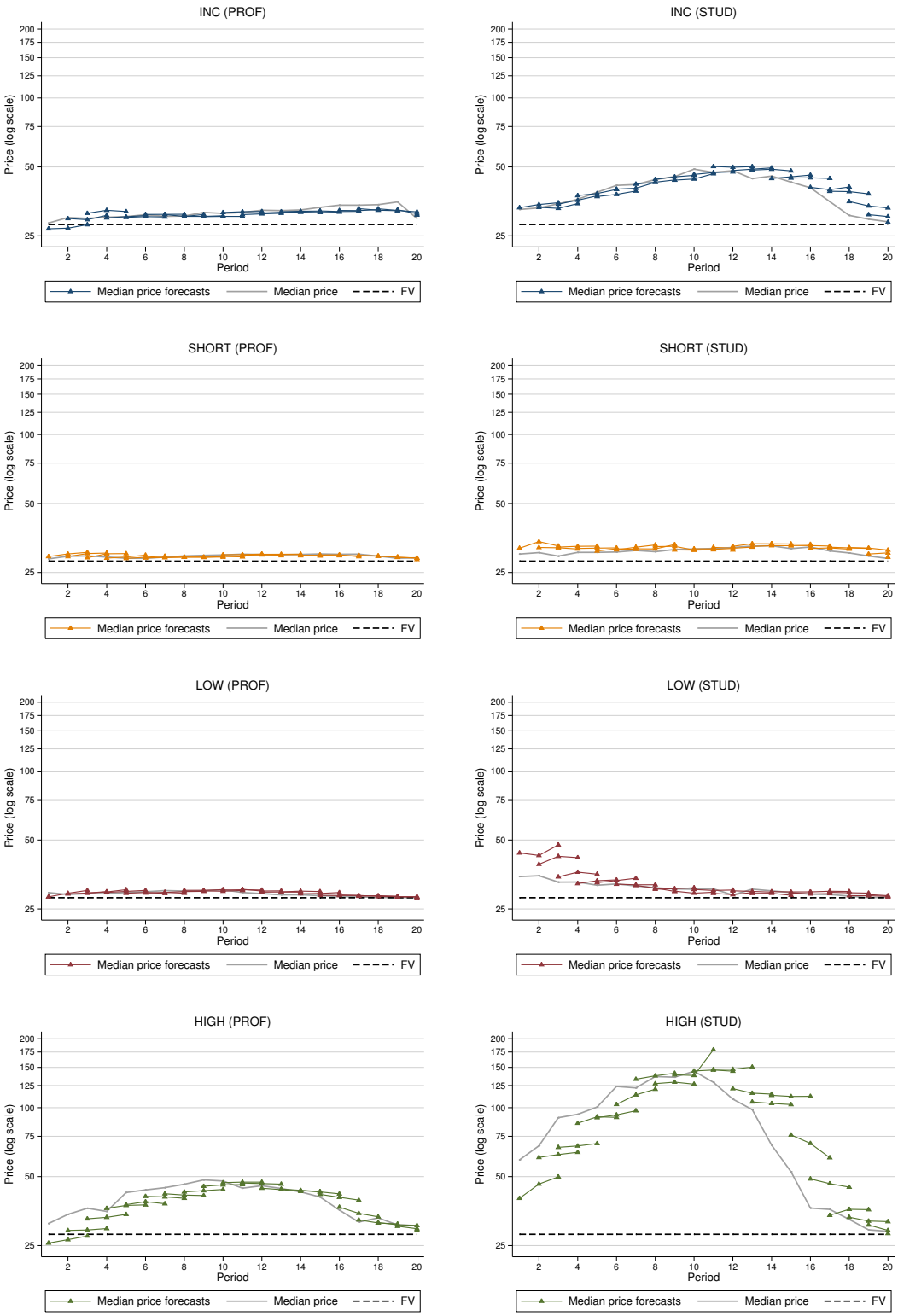


Figure C10: **Log-Price developments and subjects' median price forecasts across treatments in Experiment PROF (left column) and in Experiment STUD (right column)**: This figure depicts median treatment prices (bold grey lines) and median price forecasts for three upcoming periods (colored lines with triangles) as a function of period for treatments INC (increasing CA-Ratio), SHORT (increasing CA-Ratio, short-selling allowed), LOW (low and constant CA-Ratio of 1), and HIGH (high and constant CA-Ratio of 10.2) in log-scale. Treatments of the professional sample of Experiment PROF are displayed in the left column and the corresponding treatments in the student sample of Experiment STUD are shown in the right column. The dashed lines represent the risk-neutral fundamental value of 28.

Table C8: **Pairwise Mann-Whitney U-tests of forecast errors for periods $t+k$ with $k \in \{0,1,2\}$ between experiments PROF and STUD**: This table shows pairwise subject pool comparisons for each treatment: INC (increasing CA-Ratio), SHORT (increasing CA-Ratio, short-selling allowed), LOW (low and constant CA-Ratio of 1), and HIGH (high and constant CA-Ratio of 10.2). The table outlines median treatment values of forecast errors in percent and the numbers in parentheses show the Z-values of the MW U-test statistic. Here, forecast errors ($\text{FE}_{t,t+k}$) measure the percentage difference of market prices in $t+k$ and subject's price beliefs for $t+k$, elicited in t. The data are divided into forecast errors before (top panel) and after (bottom panel) the price peak in each market. *, ** and *** represent the 10%, 5%, and 1% significance levels of a double-sided test. Sample size N for each test is either 21 or 22.

	Before Price Peak ($t \leq t^*$)								
	t,t			$t,t+1$			$t,t+2$		
Treatment	PROF	STUD	Z	PROF	STUD	Z	PROF	STUD	Z
INC	4.95	3.45	(−1.19)	5.29	5.19	(0.07)	5.41	3.32	(−0.26)
SHORT	0.82	−2.00	(−1.71)*	0.31	−2.56	(−0.71)	0.34	−3.64	(−1.35)
LOW	0.88	0.04	(−1.28)	0.44	−15.24	(−2.35)**	−1.41	−18.65	(−1.42)
HIGH	12.65	8.49	(−0.73)	14.66	11.07	(−1.32)	14.51	13.78	(−0.33)
	After Price Peak ($t > t^*$)								
	t,t			$t,t+1$			$t,t+2$		
Treatment	PROF	STUD	Z	PROF	STUD	Z	PROF	STUD	Z
INC	−0.70	−4.83	(−2.08)**	−1.02	−11.06	(−2.03)**	0.33	−20.87	(−2.25)**
SHORT	0.01	−3.58	(−2.04)**	−0.37	−2.81	(−1.39)	−0.46	−3.32	(−1.39)
LOW	−1.33	−1.99	(−0.91)	−2.12	−4.05	(−1.32)	−2.73	−6.05	(−0.91)
HIGH	0.15	−7.58	(−2.04)**	−0.94	−16.07	(−2.24)**	−2.10	−26.72	(−2.91)***

D Instructions of the Experiments[24]

Background of the experiment

This experiment replicates an asset market in which 8 traders can trade shares of a fictitious company over 20 periods, where each period lasts for 120 seconds. You receive an initial endowment of 20 shares and 560 (**5700**) Taler (experimental currency, converted to Euro at the end of the experiment). Your asset and Taler holdings carry over from one period to the next. Your asset and Taler holdings cannot drop below zero. (`Your Taler holdings cannot drop below zero.`)

To familiarize you with the software and the trading mechanism there will be 2 trial periods which are not relevant for your final payment.

Information on the market architecture and your tasks as a trader

1) Trading

Participating in the market as a trader you can sell and buy assets. Trade is accomplished in form of a continuous double auction. That is, every trader can buy as well as sell assets. You can submit as many buy and sell orders (with at most 2 decimal places) as you like. You have to specify the number of stocks you want to trade for every order.

If you buy assets, your Taler holdings will be decreased by the respective expenditures (price x quantity) and the number of assets will be increased by the quantity of newly bought assets. Inversely, if you sell assets, your Taler holdings will be increased by the respective revenues (price x quantity) and the number of assets will be decreased by the quantity of newly sold assets. Please note that you can only buy (sell) as many assets as are covered by your Taler (asset) holdings - this includes also your active offers in the market. `Negative asset holdings (short-selling) are possible for up to -40 assets.`

Each share held at the of a trading period will pay a dividend of either 1.20 Taler or 1.60 Taler per asset with equal probability. `(For each asset shorted you have to pay the dividend.)` The randomly selected dividend is the same for each share and is newly determined each period. Additionally, you receive interest payments of 5% on your current Taler holdings. Dividend and interest payments will be added directly to your Taler holdings. (*Dividend and interest payments will be paid to a separate Account B and are not available for trading in later periods.*) *Account B will pay the same interest of 5% on all holdings and its total value will be added to your Taler holdings at the end of the experiment.*

[24]The following instructions are from the Experiment with the professional sample PROF for Treatment INC. Additional text for Treatment LOW is in *italic*, for Treatment SHORT is in `teletype` and for Treatment HIGH is written in **bold** font. Note that instructions for all four treatments in Experiment STUD were identical except for the stake size (see Section 1 in the main text for further details on the different exchange rates from Taler to euro). Of course, original instructions of each treatment can be provided upon request.

At the end of the experiment the assets you hold are bought back by the experimenter at a buyback price of 28 Taler per share `(for each asset shorted you have to pay 28 Taler)`.

Prior to the beginning of each new period you receive an income of 100 Taler, which will be added directly to your Taler holdings. (*Prior to the beginning of each new period you receive an income of 100 Taler, which will be transferred to Account B.*)

Example for the calculation of the dividend and your asset and Taler holdings: Suppose you begin the experiment with 560 Taler in cash and 20 shares. If you make no purchases or sales, then your interest earnings will be 28 Taler, that is $560 \times 0.05 = 28$ Taler. If the randomly determined dividend turns out to be 1.20 Taler, then the total dividend income will be $20 \times 1.20 = 24$ Taler. These $28 + 24 = 52$ Taler, as well as your income of 100 Taler, will be added to your Taler holdings at the end of the period. Hence, your initial endowment at the beginning of the next period will be 20 shares and 712 Taler ($560 + 52 + 100$).

(*Example for the calculation of the dividend and your asset and Taler holdings: Suppose you begin the experiment with 560 (**5700**) Taler in cash and 20 shares. If you make no purchases or sales, then your interest earnings will be 28 Taler, that is* $560 \times 0.05 = 28$ *(*$5700 \times 0.05 = 285$*) Taler. If the randomly determined dividend turns out to be 1.20 Taler, then the total dividend income will be* $20 \times 1.20 = 24$ *Taler. These* $28 + 24 = 52$ *(*$285 + 24 = 309$*) Taler, as well as your income of 100 Taler, will be transferred to Account B. The Taler holdings on Account B will be added to your Taler holdings at the end of the experiment. Hence, your initial endowment at the beginning of the next period will be 560 (**5700**) Taler and 20 shares again.*)

2) Market predictions

Additionally to your trading activity you will be asked to predict the development of market prices over the three subsequent periods. Exceptions are the penultimate period with two predictions and the last period with one prediction.
If your prediction is within $+-$ 5% of the average market price in the corresponding period, you earn 50 (**175**) Taler. That is, per period, you can earn a maximum of 150 (**525**) Taler for your three predictions. These earnings will be added to your Taler holdings at the end of the last period.
Note that you have just 30 seconds to enter your predictions in each period.

Calculation of your payment

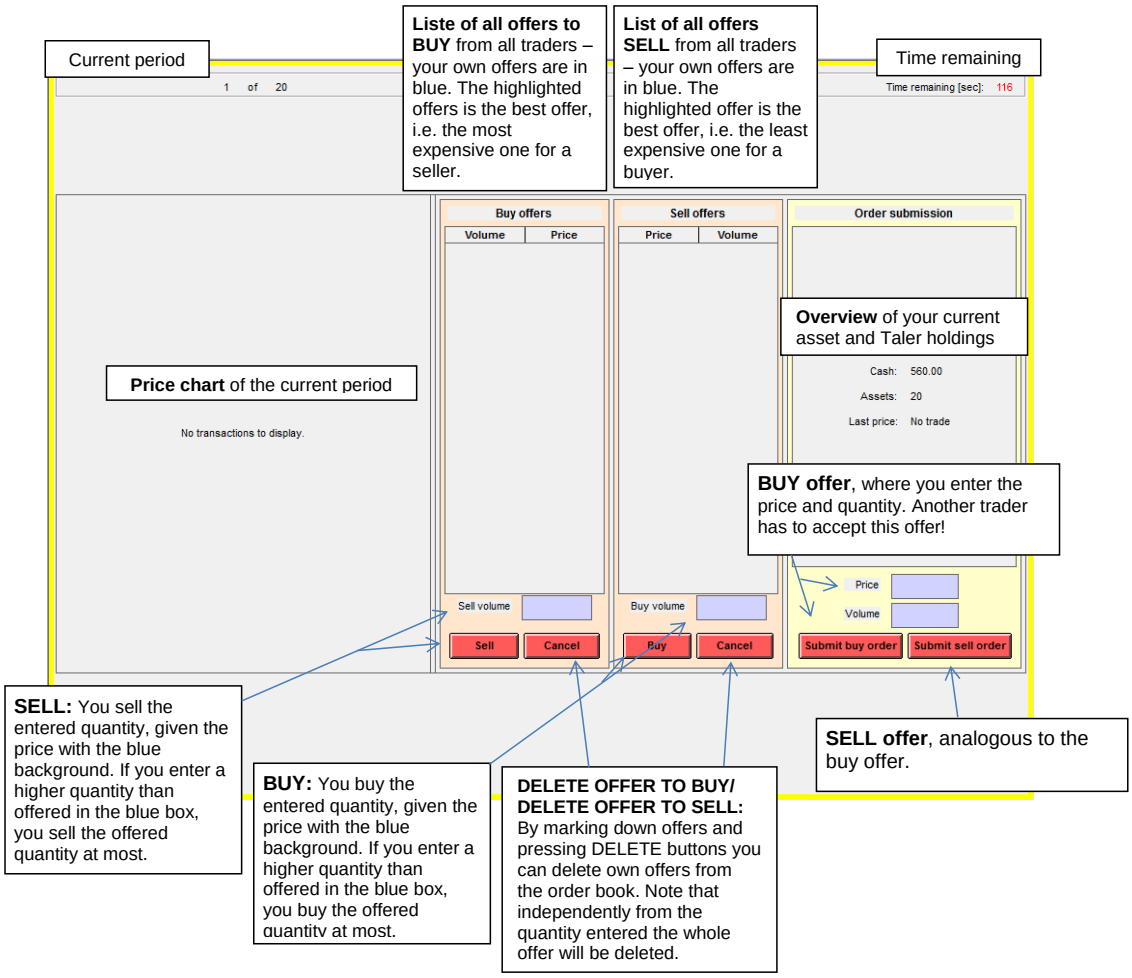
At the end of the experiment, your payment as a trader is calculated as follows:
The number of assets you hold are bought back by the experimenter at the end of the experiment (after Period 20). You will receive 28 Taler for each asset you hold. `In case your asset holdings are negative, your final wealth will be reduced by 28 Taler per asset.`
The total amount is added to your final cash (Taler) holdings. Additionally, your earnings from all your predictions will be added to your Taler holdings.

$$\text{Final Wealth in Taler} = \text{asset holdings} \times 28 \text{ Taler} + \text{Taler holdings } (+ \textit{Account B}) + \text{income from market predictions}$$

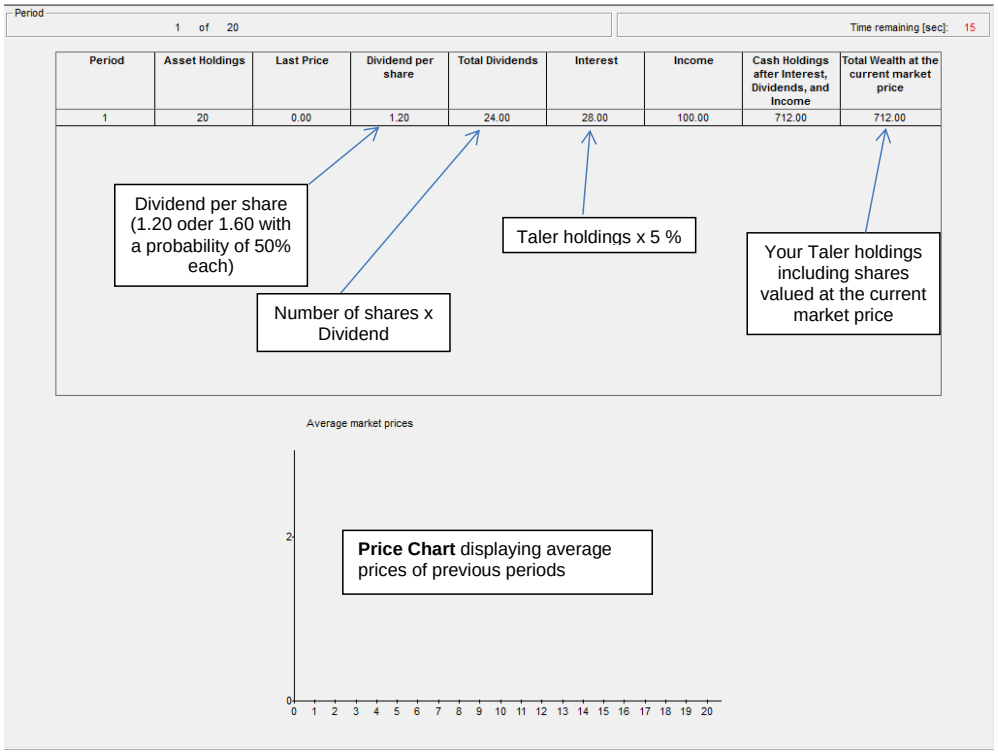
Your earnings from this experiment will then be converted to Euro using a conversion rate of 1 Euro for 100 (**350**) Taler.

$$\text{Final Wealth in Euro} = \text{Final Wealth in Taler} \ / \ 100 \ (\mathbf{350})$$

Trading Screen



History Screen



Questionnaire

Please answer the following questions or statements honestly. The analysis will be anonymous.

How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?
(0 = "unwilling to take risks"; 10 = "fully prepared to take risk")

How do you see yourself: Are you a person who is fully prepared to take risks in investment decisions or do you try to avoid taking risks?
(0 = "unwilling to take risks"; 10 = "fully prepared to take risk")

How would you rate your knowledge about financial markets compared to an average person?
(1 = "far below average"; 7 = "far below average")

Social status is primarily defined by financial success.
(1 = "completely disagree"; 7= "fully agree")

How important is it to you what others think about you?
(1 = "not important"; 7 = "very important")

How important is it to you to be the best at what you do?
(1 = "not important"; 7 = "very important")

——————— (new page) ———————

Please answer the following questions or statements honestly. The analysis will be anonymous.

I enjoy working in situations involving competition with others.
(1 = "disagree strongly"; 5 = "agree strongly")

It is important to me to perform better than others on a task.
(1 = "disagree strongly"; 5 = "agree strongly")

I feel that winning is important in both work and games.
(1 = "disagree strongly"; 5 = "agree strongly")

It annoys me when other people perform better than I do.
(1 = "disagree strongly"; 5 = "agree strongly")

I try harder when I'm in competition with other people.
(1 = "disagree strongly"; 5 = "agree strongly")

——————— (new page) ———————

Please answer the following questions or statements honestly. The analysis will be anonymous.

Age:
Gender:
What is your highest level of education?
What is your profession? Please be as specific as possible (E.g..: Risk manager in a bank, mechanical engineer)
How long have you been working in this industry (in years)?

——————— (new page) ———————

Please answer the following questions or statements honestly. The analysis will be anonymous.

What is your profession?
Which asset class(es) are you primarily involved in?
What is your annual gross salary (in euro)?

E Instructions of the Online Survey for Professionals and Students [25]

Welcome[26]

Dear participant,

Thank you very much for accepting our invitation to take part in this survey. We are researchers from several universities conducting a short study which is intended to take about 25 minutes. With your participation, you will make an important contribution to research and you can earn money: you will receive EUR 40 for participating in this survey. You will receive this amount via bank transfer. At the end of the survey we ask you to provide your e-mail address to be able to contact you regarding your bank details.

All data will be anonymous and no individual results will be disclosed publicly or to other participants of the experiment. The data will only be used for scientific purposes. This online study adheres to the principles of economic experiments: participants are not deceived and earnings are paid out in real. We guarantee at each stage of the data analyses that we will not trace back experimental decisions to participants' identities. Moreover, we will never mention the participating institutions in any paper and presentation.

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

The link to this study will be active until October 15.

Thank you very much for participating!

Prof. DDr. Jürgen Huber (University of Innsbruck)
Prof. Dr. Michael Kirchler (University of Innsbruck, Gothenburg University)
Prof. Dr. Utz Weitzel (Utrecht University, Radboud University)
Florian Lindner, PhD (University of Innsbruck)
Christoph Huber, MSc (University of Innsbruck)
Julia Rose, MSc (University of Innsbruck)

[25] The following instructions and screenshots are from the online survey analyzing differences in cognitive skills and economic preferences between professionals and students. The instructions are identical for both subject pools except for the payout (40 euro for the professionals and 10 euro for the students).

[26] For the Raven's and eye-gaze test part of the survey, we used a shortened version. The original tasks comprise 36 questions each, out of which we took every second question, starting with the first one of the original task. This was done to keep the overall time needed to complete the survey as short a possible without losing explanatory power. The order of the four tasks was randomized across all subjects.

(new page)

Overview

This survey consists of **4 different tasks**. Each task of the survey (including introductory instructions) will be presented on a separate screen. When you have completed a task, the study will continue directly with the next task (i.e., there is no immediate feedback). You will be informed about your results in the respective tasks at the very end of the experiment. Additionally, we ask you to answer a few short questions at the end of the survey.

Intro PROF

(new page)

Survey

In which industry sector do you work?
For how many years have you been working in the stated industry sector?

(new page)

Survey

In which specific field do you work?
For how many years have you been working in the stated field?

Intro STUD

(new page)

Survey

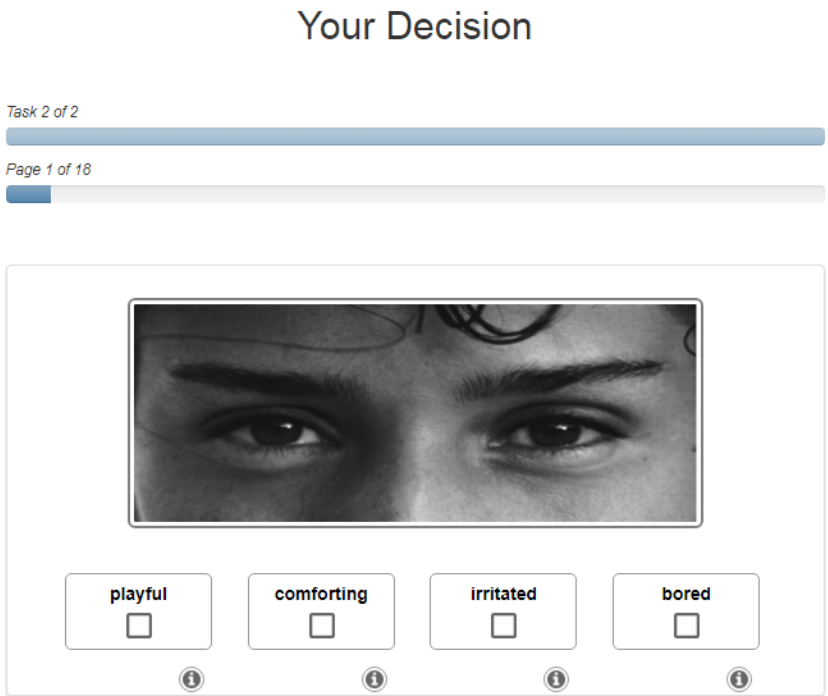
What is your field of study/your major?
Which semester are you in?
What is your country of origin?
Gender:

ToM

In the following, you will be shown 18 pictures showing just the eyes part of people's faces with four emotion labels below it. You are asked to select which one of the four emotion words best describes the emotion that the eyes are showing. Please provide your best guess for each item.

For each of the emotion words, synonyms and an example sentence are available via the small info sign in the bottom right corner.

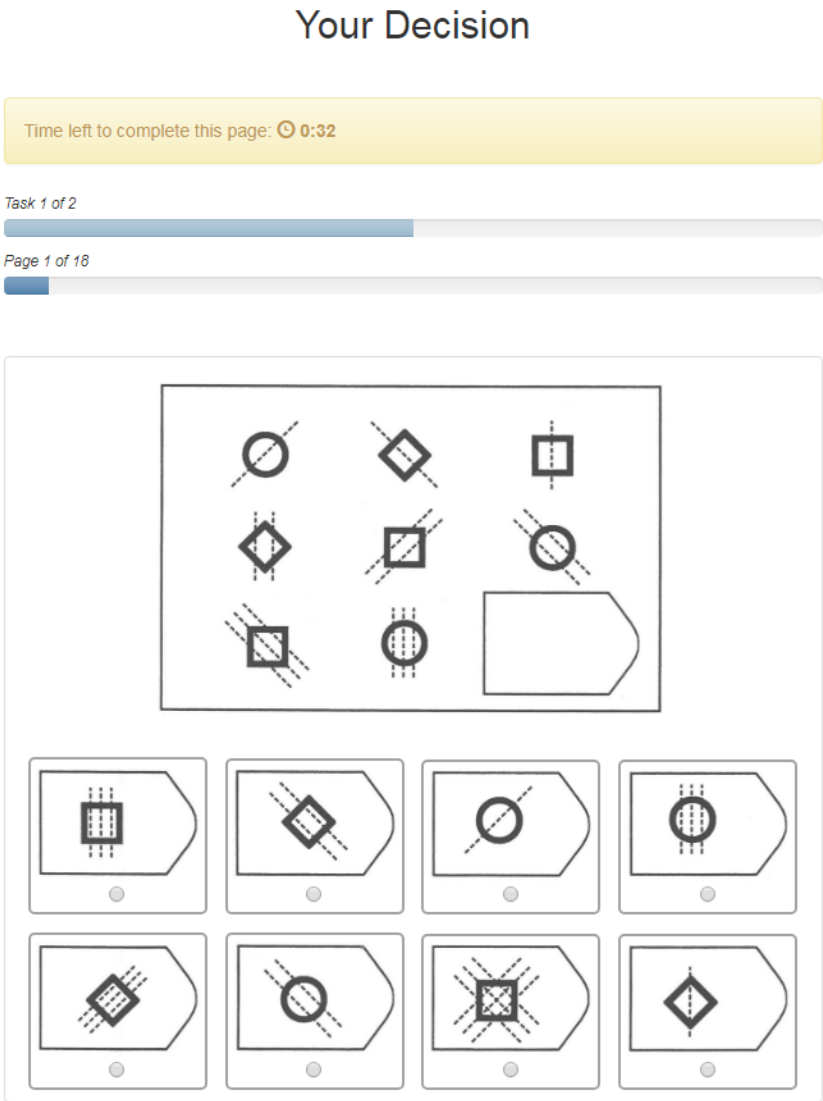
Screenshot:



Raven

In the following, you will be shown 18 test items. Each item comprises a pattern of diagrammatic puzzles with one piece missing. You are asked to choose the correct missing piece from a series of possible answers. The patterns in each item are presented in the form of a 3x3 matrix with the missing piece in the bottom right corner. You have 10 minutes in total for this task.

Screenshot:

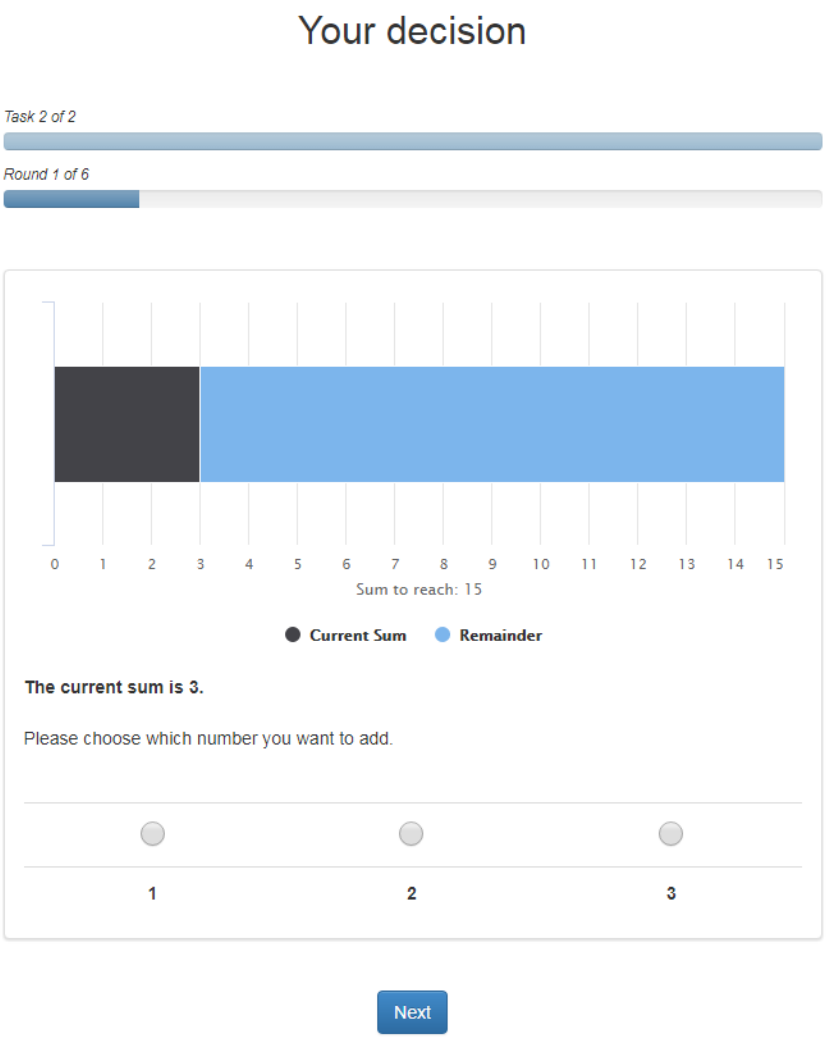


HIT15

In the following task you will play a short game against the computer for several subsequent rounds. At the beginning of the game, the computer draws a random initial value between 0 and 14. Then you are asked to add an integer between 1 and 3 to the initial value. Afterwards, the computer adds a number between 1 and 3, respectively. You and the computer then take turns. The goal of the game is to reach a total sum of 15.

If you are the one reaching 15 by adding your number, you win the game.
If the computer reaches 15 at its turn, then the computer wins the game.

Screenshot:



CRT7

Please answer the following seven questions.

Each question will be shown on a separate screen.

Screenshot:

CRT7 - Set of questions

1. A bat and a ball cost 110 cents in total. The bat costs a dollar more than the ball. How much does the ball cost?
   (intuitive answer: 10; correct answer: 5)

2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?
   (intuitive answer: 100; correct answer: 5)

3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?
   (intuitive answer: 24; correct answer: 47)

4. If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together?
   (intuitive answer: 9; correct answer: 4)

5. Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?
   (intuitive answer: 30; correct answer: 29)

6. A man buys a pig for $60, sells it for $70, buys it back for $80, and sells it finally for $90. How much has he made?
   (intuitive answer: 10; correct answer: 20)

7. Simon decided to invest $8,000 in the stock market one day early in 2008. Six months after he invested, on July 17, the stocks he had purchased were down 50%. Fortunately for Simon, from July 17 to October 17, the stocks he had purchased went up 75%. At this point, Simon:
   (a) has broken even in the stock market
   (b) is ahead of where he began
   (c) has lost money
   (intuitive answer: b; correct answer: c)

Demographics and Risk Attitudes

PROF

(new page)

Please answer the following questions:

Gender:
Year of birth:
What is your country of residence?

(new page)

Please answer the following questions:

How do you see yourself:
Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

People can behave differently in different situations.
How would you rate your willingness to take risks in financial matters?

STUD

(new page)

Please answer the following questions:

How do you see yourself:
Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

People can behave differently in different situations.
How would you rate your willingness to take risks in financial matters?

(new page)

Please answer the following questions:

Year of birth:

Please enter your matriculation number (this is needed for your payment):

F Pictures of the Experimental Laboratories

Figure F11: **Experimental Laboratories**: This figure shows one example of a mobile laboratory in the conference room of a financial institution (top) and the laboratory at Innsbruck EconLab (bottom).