

**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





# **Bubbles and Financial Professionals**

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

July 2018

revised March 2019

# Bubbles and Financial Professionals

Utz Weitzel<sup>\*a,b</sup>    Christoph Huber<sup>c</sup>    Jürgen Huber<sup>c</sup>    Michael Kirchler<sup>c,d</sup>  
Florian Lindner<sup>e</sup>    Julia Rose<sup>c</sup>

March 14, 2019

## Abstract

The efficiency of financial markets, but also their potential to produce bubbles are central topics in academic and professional debates. Yet, little is known about the contribution of financial professionals to price efficiency. We run 116 experimental markets with 412 professionals and 502 students. We find that professional markets with bubble-drivers – capital inflows or high initial capital supply – are susceptible to bubbles, although they are more efficient than student markets. In mixed markets with students, bubbles also occur, but professionals act as price stabilizers. We show that heterogeneous price beliefs drive overpricing, especially in bubble-prone market environments.

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, Austrian Working Group on Banking and Finance 2017, the Dutch Authority for Financial Markets (AFM), the EEG Inaugural Conference at the Max Planck Institute for Research on Collective Goods in Bonn 2018, the Behavioral Finance Working Group Conference 2018, Queen Mary University of London 2018, and the Innsbruck-Vienna Experimental Day 2018 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”. All materials for the experiments (e.g., source codes, data files) can be found public in the OSF repository <https://osf.io/265dp/>.

\*Corresponding author, Email: [u.weitzel@uu.nl](mailto:u.weitzel@uu.nl).

<sup>a</sup>Utrecht University, Utrecht School of Economics, Kriekenpitplein 21-22, 3584 EC Utrecht.

<sup>b</sup>Radboud University, Institute for Management Research, Thomas van Aquinostraat 5.1.26, 6525 Nijmegen.

<sup>c</sup>University of Innsbruck, Department of Banking & Finance, Universitätsstrasse 15, 6020 Innsbruck.

<sup>d</sup>University of Gothenburg, Department of Economics, Centre for Finance, Vasagatan 1, 40530 Gothenburg.

<sup>e</sup>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 17<sup>th</sup> century), in economies at different stages of development (from developing economies to highly industrialized economies in the 21<sup>st</sup> 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 drivers are of high interest to economists for at least two reasons. First, they represent periods of inefficient prices, i.e., of prices that strongly deviate from fundamentals. Second, bubbles and subsequent crashes have the power to severely affect the real economy through misallocation of resources and impaired balance sheets (Brunnermeier and Schnabel, 2016).<sup>1</sup>

For many people, 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.<sup>2</sup> Despite considerable literature on mispricing in financial markets, empirical evidence remains elusive (see, e.g., Gürkaynak, 2008; Greenwood et al., 2019, for an overview and an interesting approach). 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

---

<sup>1</sup>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 elaborate on 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.

<sup>2</sup>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>.

be misspecified. One way to tackle this problem is to use experimental asset markets as test-beds, where fundamental values are clearly 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 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.<sup>3</sup> We readily acknowledge that there are several ways to balance internal and external validity.<sup>4</sup> Moreover, each experiment, whether lab or field, represents a well-defined, controlled situation. Generalizations from experiments with financial professionals to behavior

---

<sup>3</sup>First anecdotal evidence of the behavior of business professionals was provided by Smith et al. (1988) and King et al. (1993). In Smith et al. (1988) one market was run with business men and women from the Tucson area, showing one of the largest deviations from fundamentals compared to the student markets they ran. In King et al. (1993) one market was set up with six OTC traders, exhibiting moderate levels of overpricing compared to the student markets in the paper. However, it is important to emphasize that drawing inferences and conclusions from two markets is impossible.

<sup>4</sup>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). Moreover, Hogan et al. (2004) introduce the concept of statistical arbitrage that also circumvents the joint hypothesis problem because its definition is independent of any equilibrium model. 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).

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 fill this gap by running lab-in-the-field trading experiments with financial professionals. As trading environment we use the design of Smith et al. (2014) and Holt et al. (2017) 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 common knowledge, the design has been shown to be able to consistently produce price bubbles (Smith et al., 2014; Holt et al., 2017).<sup>5</sup> The latter is important, because it provides room for our four base treatments – two bubble-driver and two bubble-moderator treatments – to take effect.

As bubble-driver (base) treatments, we administered two classical setups from the literature: in one treatment we implemented 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). In another treatment we allowed capital inflows and thereby created 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

---

<sup>5</sup>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 with decreasing fundamental values (Smith et al., 2000; Noussair et al., 2001; Kirchler et al., 2012; Huber and Kirchler, 2012).

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.

As bubble-moderating (base) treatments, we administered two designs that are known to reduce bubbles: in one treatment we allowed short sales (Ackert et al., 2006; Haruvy and Noussair, 2006), and in another treatment we provided 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.<sup>6</sup>

We recruited 294 financial professionals from high-skilled investment areas – such as trading, fund management, and portfolio management – to run the four base treatments in 38 financial asset markets. 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. In fact, following the definition of Razen et al. (2017) we detect bubbles in a quarter of all markets with financial professionals in the two bubble-driver base treatments. In the two bubble-moderator treatments, however, none of the markets populated by professionals exhibit bubble patterns.

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,

---

<sup>6</sup>Experimentally, in lab markets with student subjects, all four bubble-drivers and bubble-moderators have been shown to affect price efficiency. 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. Moreover, 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).

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' crucial role in allocating capital and the importance of financial markets for the functioning of a modern economy, it is important to learn more about their preferences and behavior in decision making. Here we add as our first main finding that even high-skilled financial professionals are not immune to bubble-drivers.

As a second contribution, we examine whether bubble phenomena are robust to subject pools. For comparison we administered the same four base treatments to 384 students in 48 lab markets. To keep the student population as comparable as possible to the sample of professionals, we 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 certain characteristics of student participants, 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 the direction and statistical significance of the treatment effects is comparable. In both groups, professionals and students, bubble-drivers increase overvaluation and bubble-moderators yield more efficient markets, though the effect sizes are significantly smaller for professionals. With these results our study also complements the experimental finance literature investigating long-lived laboratory asset markets with student subjects and thereby largely supports previous treatment effects of

student samples (e.g., Smith et al., 1988; Lei et al., 2001; Dufwenberg et al., 2005; Kirchler et al., 2012; Sutter et al., 2012). These findings also 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) show that, during the Tech Bubble, institutional investors started 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.<sup>7</sup> We contribute by providing controlled evidence on whether and how professionals promote price efficiency and that (relatively) inexperienced private investors play a pronounced role in bubble formation.

As a third contribution we explore potential drivers of bubbles and explanations for differences between subject pools, particularly the role of cognitive skills, beliefs about future prices, and beliefs about the behavior of others (i.e., beliefs about the rationality of others; Cheung et al., 2014). Regarding the latter, we investigate whether beliefs about the rationality of traders from the other subject pool can explain differences in price efficiency between professionals and students. For this, we administered three additional treatments to a newly recruited sample of 118 professionals and 118 students in 30 markets. In one treatment, four professionals and four students traded together in one of the bubble-driver base treatments and this was public knowledge. Hence, the only difference to the base treatment was the trader mix in the market. In two other treatments, professionals and students were seated in the same room, but traded in separate markets without knowledge of the trader composition in each market. Here, the only difference to the base treatment was the lack of knowledge about the homogeneous composition of the market. With this setup we can isolate the effect of beliefs about others from their actual influence on the market. We contribute by showing that beliefs about the rationality of others do

---

<sup>7</sup>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) increased their private housing exposure during the boom.

not affect price efficiency or the formation of bubbles. We observe, however, that mixed markets with public knowledge about the trader composition exhibit levels of price efficiency similar to markets solely populated by professionals (with public knowledge of trader composition). This finding is related to literature on insider trading, showing that even a small fraction of insiders can bring prices close to fundamentals in markets (e.g., Plott and Sunder, 1988). Translated to our setting, this indicates that professionals act as stabilizing device in markets with less experienced (non-professionals) traders. With an additional online survey on cognitive skills administered to market participants, we find that the more efficient markets of professionals cannot be explained by professionals' cognitive skills, but probably by their experience derived from real-world markets. Finally, our results show that heterogeneity in beliefs about future prices drives overvaluation and bubbles in the most bubble-prone treatments, which is in line with and contributes to the literature on heterogeneous beliefs and bubbles (e.g., Miller, 1977; Harrison and Kreps, 1978; Morris, 1996; Ofek and Richardson, 2003; Scheinkman and Xiong, 2003; Xiong, 2013). This finding is especially pronounced in markets populated by professionals.

The paper is structured as follows. In Section 1 we introduce the experimental design and theoretical background, followed by the results of the base treatments in Section 2. In Section 3 we investigate possible drivers for observed price (in)efficiencies and present the results of the additional market treatments, traders' beliefs and cognitive skills. Section 4 discusses and concludes.

## **1 The Base Experiment**

### **1.1 Market Setup and Theoretical Background**

The experimental market closely mimics the design of Smith et al. (2014) and Holt et al. (2017). Subjects buy and sell shares of a fictitious company for experimental currency (Taler) for a sequence of 20 periods of 120 seconds each. Share and Taler (cash) holdings are carried over from

one period to the next. We implement a typical continuous double-auction trading protocol with limit and market orders, which is standard in the literature.<sup>8</sup> 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.

Each subject is initially endowed with cash and shares. Similar to Smith et al. (2014) and Holt et al. (2017), shares pay dividends  $\tilde{d}$  of either 1.2 or 1.6 Taler with equal probability at the end of each period. Additionally, interest  $r$  of 5% is paid on cash holdings at the end of a period, but before dividends are added. The publicly known redemption value  $K$  for each share at the end of the final period  $T = 20$  is 28 Taler. The expected dividend return (expected dividend of 1.4 divided by 28) is equal to the interest rate on cash of 5% and therefore the share's risk-neutral fundamental value ( $FV_t$ ) is constant at 28 in all periods. In other words, the  $FV_t$  at the beginning of period  $t$  is calculated as the net present value of the sum of all remaining dividend payments plus the redemption value, i.e.,

$$FV_t = E(\tilde{d}) \left[ \sum_{\tau=1}^{T-t+1} (1+r)^{-\tau} \right] + K(1+r)^{-(T-t+1)}, \quad (1)$$

$$= E(\tilde{d})/r + (K - E(\tilde{d})/r)(1+r)^{-(T-t+1)} \quad \text{if } r \neq 0. \quad (2)$$

The second term of equation (2) drops out as the redemption value  $K$  equals  $E(\tilde{d})/r$ . Hence,

$$FV_t = E(\tilde{d})/r. \quad (3)$$

---

<sup>8</sup>See Appendix D for a screenshot and a detailed explanation of the trading screen.

Thus, the fundamental value is 28 ( $= 1.4/0.05$ ) and constant across all periods. Assuming risk-neutrality, market prices above 28 Taler indicate that other considerations than the fundamental value play a role. For example, investors can seek capital gains by holding an asset above its fundamental value, because they expect to sell it at an even higher price to another investor, often referred to as ‘greater fool’ (Kindleberger and Aliber, 2011). Experimental evidence shows that, even when traders are endowed with the same information about the underlying fundamentals, different expectations about resale prices and asset bubbles can still emerge because of heterogeneous beliefs about the future price predictions of others (Moinas and Pouget, 2013; Janssen et al., 2018). There is also a rich history of theoretical models in finance (e.g., Brunnermeier, 2008; Scherbina and Schlusche, 2014) showing that speculation on capital gains above the fundamental value, driven by heterogeneous beliefs, can be a rational strategy. A starting point is the study of Miller (1977) introducing traders with heterogeneous beliefs about the asset’s fundamental value. As short selling is prohibited in his model, all units of the asset are held by the investors with the most optimistic estimates of returns of the asset. Consequently, the market price equals the beliefs of the most optimistic traders. Harrison and Kreps (1978) and Ofek and Richardson (2003) show that, as soon as some traders adjust their beliefs upward, prices can rise even above the beliefs of the most optimistic traders. This resale-option theory builds on the joint effects of heterogeneous price beliefs and short-selling constraints (Morris, 1996; Scheinkman and Xiong, 2003; Palfrey and Wang, 2012), which can drive overvaluation and bubbles in our setting as well.

In the experiment, we therefore elicit beliefs about future market prices in each period, as in Haruvy et al. (2007), Kirchler et al. (2015), and Razen et al. (2017). Specifically, at the beginning of each period  $t$ , subjects are asked to predict average period prices for the three upcoming periods.  $\tilde{P}_{t,t+k}^i$  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  $\pm 5\%$  of the

average market price in the corresponding period, 50 Taler (175 Taler in Treatment HIGH, see below) are added to the cash holdings at the end of the experiment in all treatments. Subjects receive feedback on their forecast accuracy only after the final period.

## 1.2 Base Treatments

All four base treatments, two bubble-driver and two bubble-moderator treatments, are based on the market setup described above. In the first bubble-driver treatment INC (for increasing CA-ratio), each subject is initially endowed with 560 Taler in cash and 20 shares and, moreover, receives an income of 100 Taler from an exogenous source (the experimenter), which is paid to each subject at the beginning of each period.<sup>9</sup> 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 shares and borrowing money are not allowed. All this information is public knowledge. With these market characteristics we test capital inflow as one of the most prominent bubble-drivers (Brunnermeier and Schnabel, 2016; Holt et al., 2017; Razen et al., 2017).

The first bubble-moderator treatment SHORT is identical to Treatment INC except for the possibility to short up to 40 shares (i.e., share holdings can fall to a value of  $-40$ ).<sup>10</sup> With this bubble-moderator treatment we can analyze whether potential overpricing induced by the cash inflow in Treatment INC can be mitigated when allowing short-sales.

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

---

<sup>9</sup>In real life, most financial professionals invest on behalf of their clients and not own money. The latter, however, is fully incentive compatible, in the sense that it relates higher (lower) trading performance to correspondingly higher (lower) payoffs.

<sup>10</sup>Subjects with negative share holdings have to pay the respective dividends in each period and the buyback price of 28 Taler for each outstanding share at the end of the experiment. We do not impose additional cash reserve requirements (Haruvy and Noussair, 2006).

B”). This account is not available for trading but the holdings there 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 setting 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). See Table A1 in Appendix A for an overview of the treatment parameterization.

### 1.3 Implementation of the Experiment

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 markets with as many different treatments per session as possible, administering 2 to 4 treatments within a session simultaneously. The planned size of markets was 8 traders and in most cases we managed to keep this market size. However, when running experiments with professionals, some subjects participate or cancel on short notice, because of last-minute shifts in their schedules.<sup>11</sup> We thus also ran a few markets with 7 or 9 traders.

For Experiment PROF, we recruited 294 professionals from major financial institutions in several Central and Northern European countries, who are 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. Ex-

---

<sup>11</sup>Moreover, it was very difficult to deny access to the experimental market when a 9<sup>th</sup> 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.

periments with professionals were run in the Netherlands, Austria, Germany, and Slovakia. In total, we conducted 15 experimental sessions. Moreover, we signed non disclosure agreements (NDAs) for not disclosing the identity of the participating financial institutions. 89.5 percent of participating professionals 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 G). 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 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.<sup>12</sup>

---

<sup>12</sup>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.

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 shares 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 Experiment STUD we ran 48 markets with the two student subject pools from the University of Innsbruck and Radboud University Nijmegen. 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.

## 2 Results of the Base Treatments

### 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 A10 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 A11 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 by the FV of 28) and RAD (relative absolute deviation of prices to fundamentals, normalized by the FV of 28) as measures for overpricing and mispricing, respectively (Stöckl et al., 2010). Specifically,  $RAD = \sum_{t=1}^T \frac{|\bar{P}_t - FV_t|}{T}$  and  $RD = \sum_{t=1}^T \frac{\bar{P}_t - FV_t}{T}$ , with  $\bar{P}$  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  $t^*$ .  $RDMAX = \max_t \left\{ \frac{\bar{P}_t - FV_t}{FV_t} \right\} = \frac{\bar{P}_{t^*} - FV_{t^*}}{FV_{t^*}}$ , is calculated as RD of the peak average period price ( $\bar{P}$ ). 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,  $AMPLITUDE = \frac{\bar{P}_{t^*} - FV_{t^*}}{FV_{t^*}} - \min_{0 \leq k < t^*} \left\{ \frac{\bar{P}_{t^* - k} - FV_{t^* - k}}{FV_{t^* - k}} \right\}$ . 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,  $CRASH = \min_{0 \leq l \leq T - t^*} \left\{ \frac{\bar{P}_{t^* + l} - FV_{t^* + l}}{FV_{t^* + l}} \right\} - \frac{\bar{P}_{t^*} - FV_{t^*}}{FV_{t^*}}$ , 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).<sup>13</sup> To test for significant pairwise differences between subject pools or treatments, we first compute the market mean of all period

<sup>13</sup>See also Section 2.2 for more explanations and figures B2 to B12 in Appendix B for individual transaction price charts of each market.

values for the variable of interest. Then we employ a Mann-Whitney U-test (MW U-test) with the market as unit of observation.<sup>14</sup>

**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 A10 in Appendix A mispricing and overpricing in treatments SHORT and LOW are very small, with values below 7.0 percent of the fundamental value. In contrast, median overpricing in treatments INC and HIGH is substantially higher with median values reaching 13.3 and 57.7 percent, respectively. Similarly, amplitude and crashes are large, particularly in Treatment HIGH with median AMPLITUDE of 70.8 percent and median CRASH equaling  $-128.6$  percent (both as a percentage of the FV of 28). By running pairwise Mann-Whitney U-tests<sup>15</sup> in Table 2 for RD (for both experiments) and in Table A10 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 at 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.

---

<sup>14</sup>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.

<sup>15</sup>Throughout the paper we report significance levels of statistical tests at 5% and lower. Due to the elimination of a minor coding error some of the numbers reported on the base treatments (e.g., RD, RAD) differ slightly from the first working paper. However, statistical results and magnitudes of significance did not change.

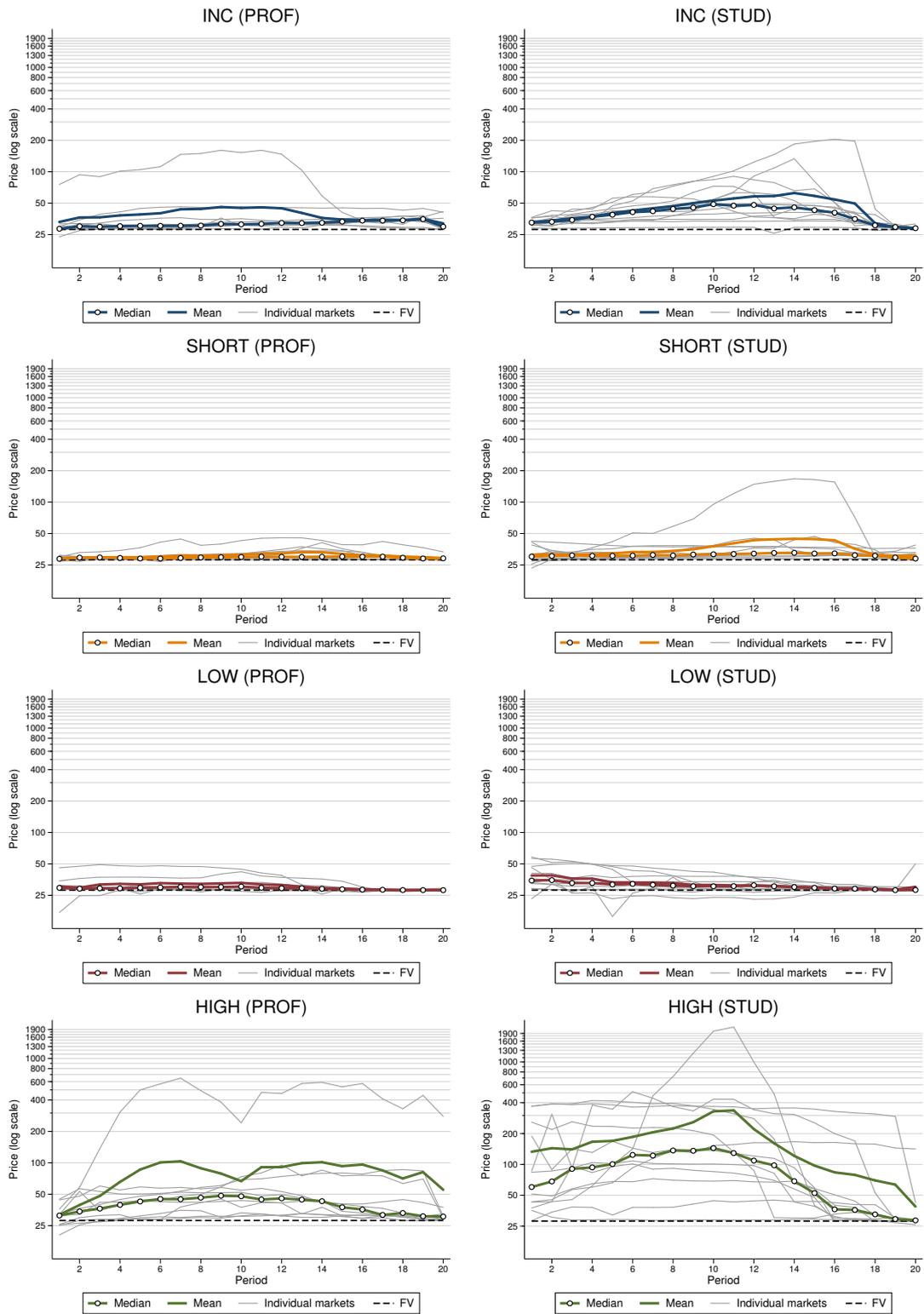


Figure 1: **Log-price developments across treatments in Experiment PROF (left column) and in Experiment STUD (right column):** This figure depicts median treatment prices (bold and colored lines with circles) and mean treatment prices (bold and colored lines) as a function of period for treatments INC (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.

<b>Experiment PROF</b>				
Variable (Median) in percent	Treatment			
	INC	SHORT	LOW	HIGH
<b>RAD</b> (mispricing)	13.68	6.97	5.46	61.47
<b>RD</b> (overpricing)	13.32	6.58	2.81	57.67
<b>RDMAX</b> (max overpricing)	35.21	16.41	9.19	108.21
<b>AMPLITUDE</b> (price amplitude)	41.55	16.77	9.19	70.81
<b>CRASH</b> (price crash)	-28.89	-19.22	-44.91	-128.62
<i>N</i>	10	9	9	10

<b>Experiment STUD</b>				
Variable (Median) in percent	Treatment			
	INC	SHORT	LOW	HIGH
<b>RAD</b> (mispricing)	47.76	15.28	13.18	309.31
<b>RD</b> (overpricing)	47.76	13.73	10.49	309.27
<b>RDMAX</b> (max overpricing)	90.27	38.62	33.49	497.70
<b>AMPLITUDE</b> (price amplitude)	76.99	10.66	5.10	177.85
<b>CRASH</b> (price crash)	-87.75	-53.61	-66.41	-393.66
<i>N</i>	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). The numbers in parentheses show the Z-value of the MW U-test statistic. \*\* and \*\*\* represent the 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						
Treatment	PROF			STUD		
	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	6.74** (2.37)	10.50** (2.29)	-44.36 (-0.91)	34.03*** (2.71)	37.27*** (3.18)	-261.51*** (-2.89)
SHORT	.	3.76 (-0.75)	-51.10*** (2.61)	.	3.24 (-0.35)	-295.54*** (3.35)
LOW	.	.	-54.86*** (-2.78)	.	.	-298.78*** (-3.58)

**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 A11 in Appendix A overpricing (RD) of 13.7 and 10.5 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 47.8 and 309.3 percent, respectively. Moreover, amplitude and crashes are substantial, especially in Treatment HIGH with median AMPLITUDE of 177.9 percent and median CRASH of -393.66 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 A11 for all other variables.<sup>16</sup> We find no differences between SHORT

<sup>16</sup>As results are very similar in all other variables compared to RD, we only outline RD in Table 2 for exemplary purpose.

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 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). When 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 34.4 and 251.6 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 55.1 and 389.5 percentage points in both treatments, respectively, as well as in the median CRASH with differences of 58.9 (INC) and 265.0 (HIGH) percentage points.

## **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 introduction, there is no generally accepted bubble definition, also

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 5% and 1% significance levels of a double-sided test. Sample size  $N$  for each test is either 21 or 22.

Treatment	RAD			RD			RDMAX		
	PROF	STUD	Z	PROF	STUD	Z	PROF	STUD	Z
INC	13.68	47.76	(2.11)**	13.32	47.76	(2.11)**	35.21	90.27	(1.71)
SHORT	6.97	15.28	(1.63)	6.58	13.73	(1.49)	16.41	38.62	(1.35)
LOW	5.46	13.18	(1.42)	2.81	10.49	(0.92)	9.19	33.49	(1.92)
HIGH	61.47	309.31	(2.18)**	57.67	309.27	(2.18)**	108.21	497.70	(1.98)**

Treatment	AMPLITUDE			CRASH		
	PROF	STUD	Z	PROF	STUD	Z
INC	41.55	76.99	(1.32)	-28.89	-87.75	(-1.98)**
SHORT	16.77	10.66	(0.04)	-19.22	-53.61	(-1.42)
LOW	9.19	5.10	(-0.87)	-44.91	-66.41	(-1.21)
HIGH	70.81	177.85	(1.12)	-128.62	-393.66	(-1.78)

not in the experimental markets literature.<sup>17</sup> Some of the challenges in identifying bubbles in the lab concern the variables to measure a bubble with and the threshold values that 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 the 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 95<sup>th</sup> 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 implemented 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.<sup>18</sup>

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  $RDMAX_m > MAX\{0; \overline{RDMAX}^{LOW} + t(df)_{0.95} \cdot \sigma(RDMAX^{LOW})\}$ ,

<sup>17</sup>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.

<sup>18</sup>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.

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  $\text{RDMAX}_m$  is higher than the 95th percentile, its peak period price for a particular market is considered to be significantly higher than in the benchmark. Note that we also impose that  $\text{RDMAX}_m$  must exceed zero to rule out potential price paths that do not exceed the FV.

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

$\text{AMPLITUDE}_m > \overline{\text{AMPLITUDE}}^{\text{LOW}} + t(df)_{0.95} \cdot \sigma(\text{AMPLITUDE}^{\text{LOW}})$ , 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  $\text{CRASH}_m < \overline{\text{CRASH}}^{\text{LOW}} - t(df)_{0.95} \cdot \sigma(\text{CRASH}^{\text{LOW}})$ , 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.

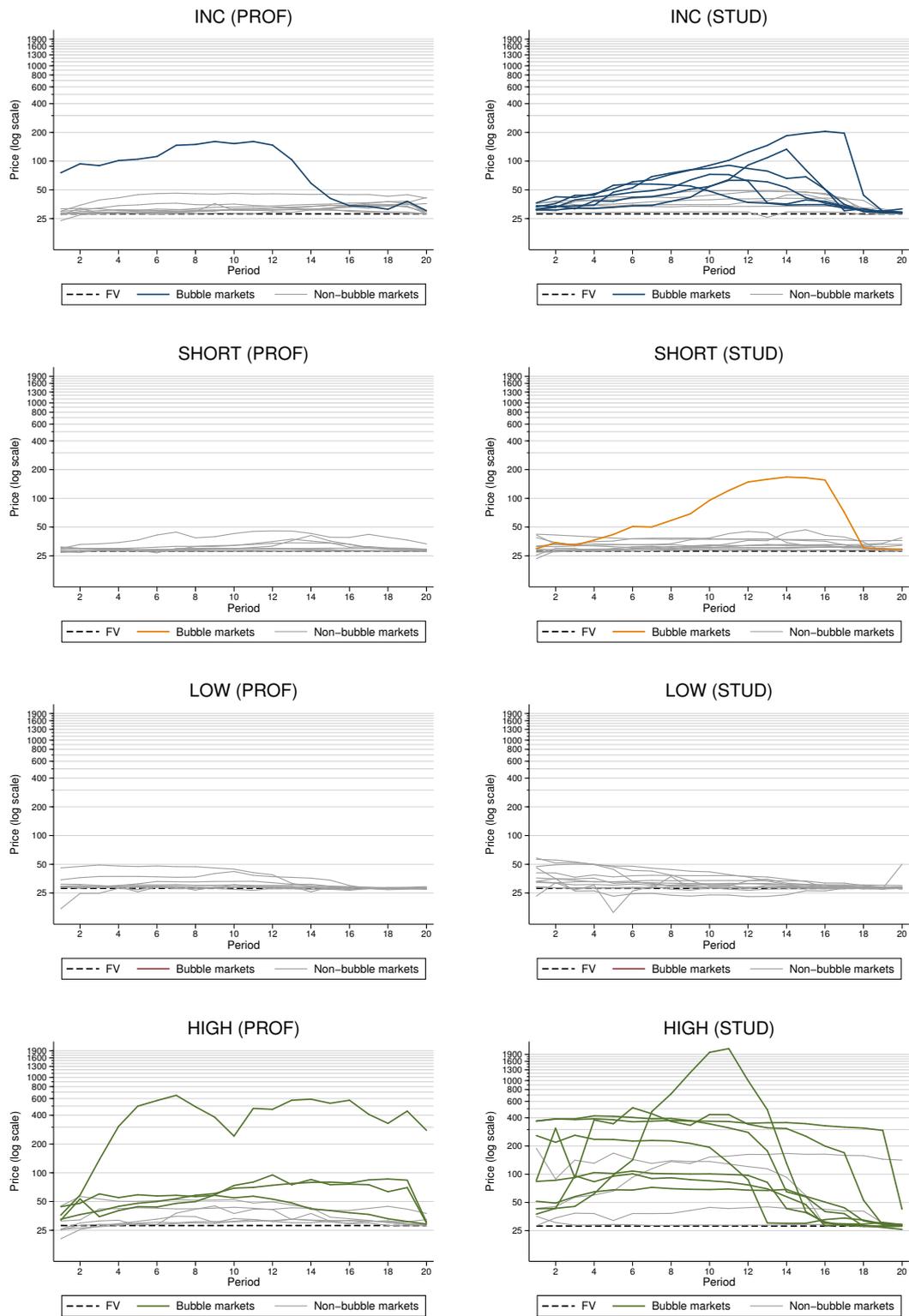


Figure 2: **Bubble identification across treatments in Experiment PROF (left column) and in Experiment STUD (right column):** Following the bubble definition in Section 2.2 this figure depicts volume-weighted mean prices for bubble markets (bold and colored lines) and non-bubble markets (grey lines) 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 4:** *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: 50 and 67 percent of the markets in treatments INC and HIGH, respectively, show bubble patterns (see tables A3 to A6 in Appendix A for details on all measures, separated for each market).

### 3 Possible Drivers of Price (In)Efficiency

In this section we explore possible drivers of price efficiency. First, we focus on possible differences between subject pools; specifically, on different beliefs about others and different cognitive skills. We administered three additional treatments of type HIGH (NPK<sup>PROF</sup>, NPK<sup>STUD</sup>, and PK<sup>MIXED</sup>) with 10 markets run for each treatment. With these additional treatments, we investigate whether beliefs about the rationality of others (i.e., other subject pools with perceived different experience levels) drive our main results. In treatments NPK<sup>PROF</sup> and NPK<sup>STUD</sup>, professionals and students were seated in the same room, traded in separate markets like in the other treatments, but did not know the trader composition in each market (the acronym NPK stands for “no public knowledge”). In Treatment PK<sup>MIXED</sup> four professionals and four students

traded in the same market which was “public knowledge” (hence, PK as treatment acronym). To test whether cognitive skills differ between professionals and students, which could possibly explain professionals’ lower inclination to bubble markets, we administered a survey with various cognitive skill tasks to the participants of the experiment and also to newly recruited subjects. Second, we analyze potential drivers of price inefficiency and bubbles in general and across all treatments. Here, we focus on the role of heterogeneous beliefs about future prices and on other variables, such as trading volume and market liquidity.

### 3.1 Additional Treatments $\text{NPK}^{\text{PROF}}$ , $\text{NPK}^{\text{STUD}}$ , and $\text{PK}^{\text{MIXED}}$

#### 3.1.1 Experimental Design and Results

Following laboratory asset market experiments with students by Lei et al. (2001), Cheung et al. (2014), Akiyama et al. (2017), and Bosch-Rosa et al. (2018), beliefs about the “rationality” of other market participants can impact price efficiency.<sup>19</sup> It is indeed possible that professionals, knowing that they are trading with other professionals, trade more efficiently, because they developed common knowledge of each other’s rationality. In contrast, students could believe that trading with other students offers plenty of irrationality in the market, leaving room for speculation.

In both treatments without public knowledge of trader composition in the market ( $\text{NPK}^{\text{PROF}}$  and  $\text{NPK}^{\text{STUD}}$ ), professionals and students were seated in the same room, traded in separate markets like in the other base treatments, and did not know the exact trader composition. They were informed about the demographics of the other group and that 8 subjects in the room constitute one market. Identical to Cheung et al. (2014) they were not told that professionals

---

<sup>19</sup>For instance, Lei et al. (2001) show that mispricing in markets following the design of Smith et al. (1988), arises from uncertainty about the behavior of others – i.e., that some subjects doubt the rationality of others. Moreover, Cheung et al. (2014) manipulate agents’ information regarding the rationality of others and distinguish the direct effect of training on the protocol (experience) from the indirect effect of public knowledge about this training. The authors find that there is a distinct and significant effect of public knowledge over and above the effect of training alone.

will only trade with professionals and students only with students (see the exact wording in the experimental instructions in Appendix E). By comparing both treatments to the respective base treatments HIGH, we can isolate beliefs about the perceived rationality of others. Hence, markets in treatments  $\text{NPK}^{\text{PROF}}$  and HIGH are set up identically except for subjects' beliefs about other traders in the market.

In the third additional treatment,  $\text{PK}^{\text{MIXED}}$ , we ran mixed markets with public knowledge about the exact trader composition in each market. Here, four professionals and four students constitute one market. Comparing Treatment  $\text{PK}^{\text{MIXED}}$  with the respective non-mixed treatments of type HIGH allows us to measure the impact of heterogeneous skills on price efficiency. Where possible we ran all three treatments within the same session to control for idiosyncratic session effects.

In total, we conducted 30 additional markets, equally split across all three treatments. We recruited 118 professionals from major financial institutions in Austria and the Netherlands, who were regularly confronted with investment and trading decisions in their daily work. 86.3% percent were male, the average age was 35.1 years, and they had been working in the finance industry for 9.0 years on average. We made sure none of the subjects had participated in any market of our base treatments. We applied the same recruitment and implementation strategy and the same software package as for the base treatments outlined above (market type HIGH). Professionals received an average payout of 71.3 euro with a standard deviation of 10.0. The average duration of the experiment was 75 minutes. Moreover, we recruited 118 students from the same subject pools as in the base treatments.<sup>20</sup> We recruited 80.0% percent male students, the average age was 22.9 years and 95.8 percent were students at management and economics departments. Students received an average payout of 17.5 euro with a standard deviation of 3.9.

---

<sup>20</sup>Of course, each subject participated in only one market in this paper and we made sure that subjects had not participated in earlier asset market experiments of similar design.

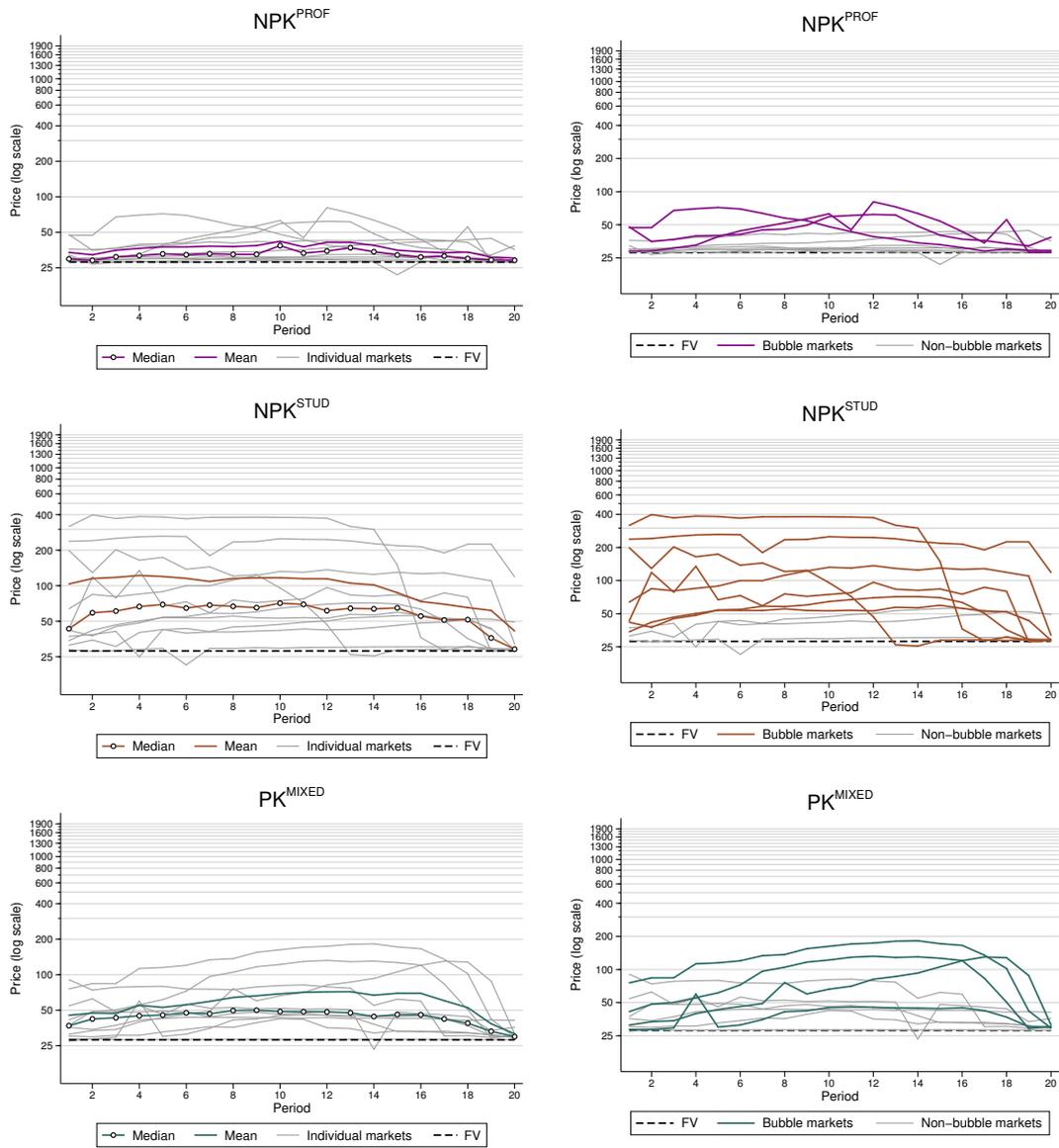


Figure 3: **Log-price developments (left column) and bubble identification (right column) across the treatments:** Left column: this figure depicts median treatment prices (bold and colored lines with circles) and mean treatment prices (bold and colored lines) as a function of period for treatments  $NPK^{\text{PROF}}$ / $NPK^{\text{STUD}}$  (Treatment HIGH with non public knowledge about composition of market) and  $PK^{\text{MIXED}}$  (Treatment HIGH with public knowledge about equal split of professionals/students in the market) in log-scale. The dashed lines represent the risk-neutral fundamental value of 28 and the grey lines show volume-weighted mean prices for individual markets. 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 the corresponding treatments in log-scale.

Table 4: **Treatment medians of mispricing (RAD), overpricing (RD), maximum overpricing (RDMAX), price run-ups (AMPLITUDE), and crash (CRASH) in the additional treatments in percent:** This table depicts median treatment values for treatments HIGH (high and constant CA-ratio of 10.2), NPK<sup>PROF</sup> (Treatment HIGH populated by professionals, but with non-public knowledge about trader composition of market), NPK<sup>STUD</sup> (Treatment HIGH populated by students, but with non-public knowledge about trader composition of market), and PK<sup>MIXED</sup> (Treatment HIGH with public knowledge about equal split of professionals/students in the market). 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.

Variable (Median) in percent	Treatment				
	HIGH(PROF)	NPK <sup>PROF</sup>	HIGH(STUD)	NPK <sup>STUD</sup>	PK <sup>MIXED</sup>
<b>RAD</b> (mispricing)	61.47	18.93	309.31	131.98	60.57
<b>RD</b> (overpricing)	57.67	18.93	309.27	131.94	59.73
<b>RDMAX</b> (max overpricing)	108.21	35.61	497.70	268.47	118.40
<b>AMPLITUDE</b> (price amplitude)	70.81	18.34	177.85	115.95	43.61
<b>CRASH</b> (price crash)	-128.62	-34.19	-393.66	-262.98	-123.39
<i>N</i>	10	10	12	10	10

***Result 5:** Professionals' markets without public knowledge of trader composition are, in line with Results 3 and 4, more efficient than students' markets without public knowledge. Mixed markets of professionals and students with public knowledge about the trader composition exhibit levels of price efficiency that are close to markets with professionals, indicating that professionals can act as a stabilizing force in our markets.*

Support: Figure 3 outlines average volume-weighted period prices of individual markets, treatment medians and means (left column) and bubble identification analyses (right column). Table 4 and Figure 4 provide measures for price efficiency including significance tests for treatment differences. Professionals' markets without public knowledge of market composition (NPK<sup>PROF</sup>) are much more efficient with fewer bubbles and significantly lower overpricing (RD) compared to student markets without public knowledge (NPK<sup>STUD</sup>), but also compared to student markets with public knowledge (HIGH). Moreover, pairwise Mann-Whitney U-tests for RD show no differences between PK<sup>MIXED</sup>, NPK<sup>PROF</sup>, and the Treatment HIGH with professionals at

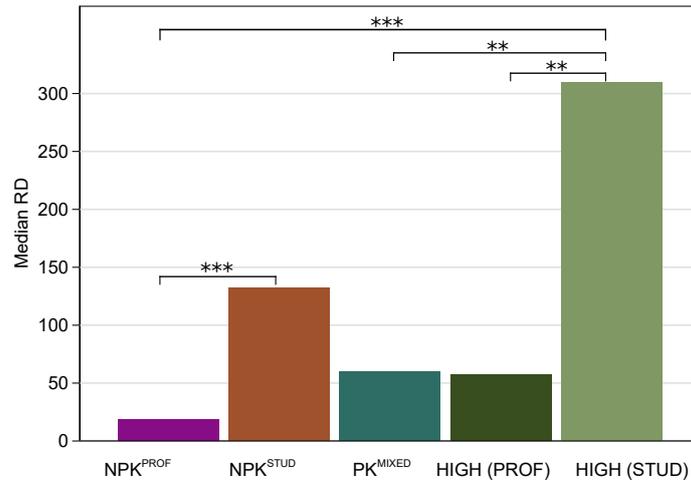


Figure 4: **Pairwise Mann-Whitney U-tests of overpricing (RD) in the additional treatments:** This figure shows pairwise treatment comparisons for treatments HIGH (high and constant CA-ratio of 10.2), NPK<sup>PROF</sup> (Treatment HIGH populated by professionals, but with non-public knowledge about trader composition of market), NPK<sup>STUD</sup> (Treatment HIGH populated by students, but with non-public knowledge about trader composition of market), and PK<sup>MIXED</sup> (Treatment HIGH with public knowledge about equal split of professionals/students in the market). \*\* and \*\*\* represent the 5% and 1% significance levels of a double-sided test.

the 5 percent level. In contrast, we do find significantly less overpricing in mixed markets with public knowledge (PK<sup>MIXED</sup>) compared to markets populated solely by students with public knowledge (HIGH). Jonckheere's trend tests show that overpricing decreases with the number of professionals in the market, independent of public or non-public knowledge. In fact, we find a statistically significant ordering of treatments (at  $p < 0.01$ ), both, for  $RD(HIGH (PROF)) < RD(PK^{MIXED}) < RD(HIGH (STUD))$ , and for  $RD(NPK^{PROF}) < RD(PK^{MIXED}) < RD(NPK^{STUD})$ . The picture is similar for the other variables RDMAX, AMPLITUDE, and CRASH as outlined in Table A12 in Appendix A.

Turning to our bubble classification, 3 of the 10 markets in Treatment NPK<sup>PROF</sup>, but a relatively high number 7 of 10 markets in the corresponding Treatment NPK<sup>STUD</sup> with students exhibit bubbles. In Treatment PK<sup>MIXED</sup> we find an intermediate number of bubble markets (4 out of 10 markets). See also tables A7 to A9 in Appendix A for details on all measures, separated for each market of the additional treatments.

These results corroborate the finding that markets populated by professionals show more efficient prices and are less prone to bubbles compared to student markets. This applies, as the additional treatments show, even when the trader composition is unknown or half of the traders are (inexperienced) students. Hence, in our setting, professionals seem to act as a stabilizing force in markets, irrespective of their beliefs about the behavior of others.

### 3.1.2 Cognitive Skills and Risk-Taking

In this section we explore whether cognitive skills and/or risk preferences – both of which prior literature suggests as natural candidates – can explain why professionals act as price stabilizers. Professionals can differ from students in cognitive skills through selection into the industry or through learning on the job (or both). Empirical studies suggest that high IQ-investors exhibit 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). The experimental finance literature extends these findings by analyzing the impact of various types of cognitive skills: fluid intelligence, cognitive reflection, Theory of Mind (ToM), and backward induction ability.<sup>21</sup>

Professionals can also differ in risk attitudes. The pertinent experimental literature is not very extensive, but it does provide some indications for potential effects of risk aversion on

---

<sup>21</sup>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). However, DeMartino et al. (2013) show that ToM-skills can also be detrimental when trading on financial markets. In a study using fMRI techniques, the authors report a mechanism by which social signals affect value computations in ventromedial prefrontal cortex, thereby increasing subject's proneness to ride financial bubbles. Moreover, Corgnet et al. (2015a) and Bosch-Rosa et al. (2018) find a causal relationship between traders' cognitive sophistication and price efficiency.

trading behavior and price efficiency.<sup>22</sup> Hence, both, cognitive skills and risk attitudes, can drive behavior and performance on stock markets (e.g., Fellner and Maciejovsky, 2007; Grinblatt et al., 2011, 2012; Kleinlercher et al., 2014; Hefti et al., 2016; Corgnet et al., 2018) and explain differences in price efficiency (Corgnet et al., 2015a; Bosch-Rosa et al., 2018).

To investigate whether higher cognitive abilities and, possibly, differences in risk attitudes can explain higher price efficiency in markets with professionals, we administered a separate online survey  $\text{EXP}^{\text{COGRISK}}$  to all participants of treatments  $\text{NPK}^{\text{PROF}}$ ,  $\text{NPK}^{\text{STUD}}$ , and  $\text{PK}^{\text{MIXED}}$  after the experiment. This survey was run online to limit the length of the experiment and to separate both parts from each other to avoid confounding effects. From the pool of subjects of treatments  $\text{NPK}^{\text{PROF}}$ ,  $\text{NPK}^{\text{STUD}}$ , and  $\text{PK}^{\text{MIXED}}$ , 101 professionals and 106 students participated (out of 118 each).

Although we administered the survey  $\text{EXP}^{\text{COGRISK}}$  a few days after the experiment, we cannot fully exclude confounding effects by the preceding experiment. We therefore administered the same online survey to two newly recruited samples of 121 financial professionals and 124 students from the same countries as in the market experiment. We refer to this additional online survey as  $\text{ONLINE}^{\text{COGRISK}}$ . Professionals were employed in the same areas as the ones from the market experiments and they shared the same characteristics. 84.3 percent of the professionals were male, their 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 were male and average age was 23.8 years.

In contrast to  $\text{EXP}^{\text{COGRISK}}$ , the data from  $\text{ONLINE}^{\text{COGRISK}}$  cannot be merged with the experimental trading data. In our analyses we therefore focus on  $\text{EXP}^{\text{COGRISK}}$  and report  $\text{ONLINE}^{\text{COGRISK}}$  as a robustness check. Both surveys,  $\text{EXP}^{\text{COGRISK}}$  and  $\text{ONLINE}^{\text{COGRISK}}$ ,

---

<sup>22</sup>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. However, when faced with bonus incentives, Kleinlercher et al. (2014) show that even less risk averse subjects invest more in the risky asset, resulting in asset overvaluation.

were programmed and conducted with oTree (Chen et al., 2016) and professionals (students) received a flat payment of 40 (10) euro as 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).<sup>23</sup> For cognitive reflection skills, we used the extended cognitive reflection test (CRT) from Toplak et al. (2014) with seven items. Cognitive reflection adds to fluid intelligence, because it helps individuals to avoid commonly-observed heuristics and biases with effortful reasoning. The CRT rests on the dual-process theory framework (Kahneman, 2011).<sup>24</sup> ToM defines one's capacity to infer others' intentions. 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 describe the mental state of the person whose eyes are shown.<sup>25</sup> Details on the tests can be found in Appendix F. In addition, we administered a HIT15 test (Burks et al., 2009) analyzing individuals' backward induction abilities, which are important in finite horizon markets.<sup>26</sup> 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).<sup>27</sup>

---

<sup>23</sup>Fluid intelligence 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.

<sup>24</sup>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 is correct) which requires effortful System 2 processes.

<sup>25</sup>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 as possible without losing explanatory power. See Bilker et al. (2012) and Olderbak et al. (2015) for a discussion of the high validity of short versions of the Raven's advanced progressive matrices and the eye-gaze test.

<sup>26</sup>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 player to reach 15 points. The initial number of the game is randomly determined. The task was played for 6 rounds.

<sup>27</sup>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

**Result 6:** *Professionals do not differ from students in most cognitive skills. Although professionals' self-reported levels of financial risk attitudes are significantly higher than those of students, general risk attitudes are not. Regressions show that higher levels of price efficiency in professionals' markets cannot be explained by risk attitudes or cognitive skills.*

Support: Figure 5 and Tables A13 and A14 in Appendix A outline the results of both surveys on all cognitive tasks.<sup>28</sup> Encouragingly, we find almost identical patterns in both surveys. In particular, we report that professionals do not differ significantly from students in most of the cognitive tests in both surveys. The only significant difference we find in cognitive variables is CRT in EXP<sup>COGRISK</sup>, but the difference is rather small with 5.3 correct answers (out of 7) for professionals and 4.7 for students. Moreover, the difference is not robust when compared to ONLINE<sup>COGRISK</sup>, where we find no statistically significant difference for any of the four cognitive skills, including CRT.<sup>29</sup> These results indicate very good performance of both pools and differences are clearly much smaller compared to other studies comparing bubble formation of high and low CRT students (Bosch-Rosa et al., 2018). 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. Moreover, in contrast to Corgnet et al. (2018) we do not find that cognitive skills, elicited in Survey EXP<sup>COGRISK</sup>, explain trader performance in the three additional treatments (see Table A14 in Appendix A). This finding

---

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.

<sup>28</sup>A detailed overview of the distribution of scores for professionals and student subjects can be found in Figure A1 in the appendix for both EXP<sup>COGRISK</sup> and ONLINE<sup>COGRISK</sup> separately.

<sup>29</sup>All results for CRT also 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.

holds both for professionals and students separately and provides indirect evidence that the slight differences in cognitive skills are probably not responsible for driving price efficiency. The potential differences to Corgnet et al. (2018) could be based on our subject pools, as subjects from both samples show high average CRT scores, leaving potentially little room for variation in the cognitive measures.

Importantly, we test whether cognitive skills (plus self-reported risk attitudes and years in industry) explain overpricing in markets of the additional treatments which would provide direct evidence of a link between (low) cognitive skills and bubble formation. As outlined in Table A15 in Appendix A, we calculate for each market and variable the mean across all subjects. We then regress these variables against market's overvaluation RD for the pooled data of the three additional treatments. We find clear evidence that neither any cognitive measure, nor financial risk-taking, explain overvaluation in our markets. We report that all coefficients are not significant on any conventional significance levels.

Our findings in this section show that neither cognitive abilities nor risk preferences can explain higher levels of price efficiency in markets with professionals. Hence, we conjecture that "professional skills" that may be rooted in real-world market experience, possibly including a more intuitive understanding of markets that goes beyond specific cognitive skills, may affect trading behavior and lead to more efficient pricing.

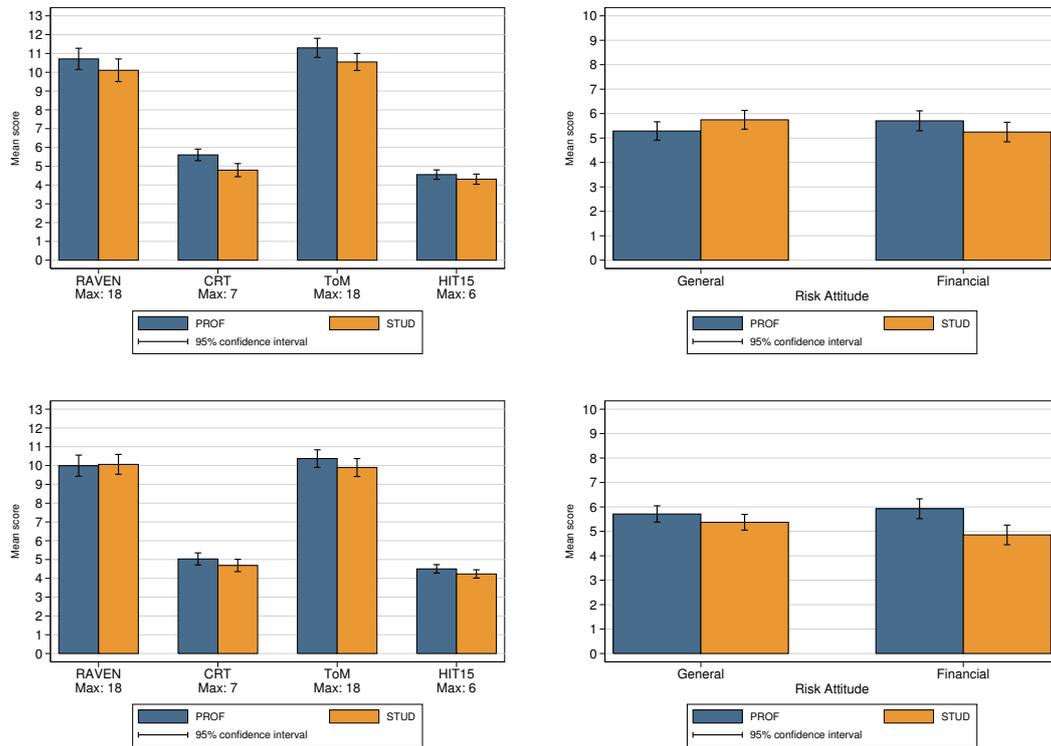


Figure 5: **Differences in cognitive skills (left panel) and risk attitudes (right panel) between professionals (PROF) and students (STUD) in surveys EXP<sup>COGRISK</sup> (top panel) and ONLINE<sup>COGRISK</sup> (bottom panel):** The left panel of this figure depicts mean test scores of fluid intelligence (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. 95% confidence intervals are displayed for each bar. The maximum score in the tests was 18 (RAVEN and ToM), 7 (CRT), and 6 (HIT15). EXP<sup>COGRISK</sup> was administered to participants of treatments NPK<sup>PROF</sup>, NPK<sup>STUD</sup>, and PK<sup>MIXED</sup> as second part of the experiment (to be taken online at another day). ONLINE<sup>COGRISK</sup> was run with newly recruited samples of 121 financial professionals and 124 students from the same countries as in the market experiment.

## 3.2 Beliefs and Trading Behavior

### 3.2.1 Beliefs

In this section we analyze whether heterogeneous beliefs about future prices drive bubbles in our markets. For this, we calculate the standard deviation of normalized price beliefs  $SD(\widetilde{NP}_{t,t+k})$  among all traders in a market in period  $t$  with  $k$  indicating values in the range  $\{0,1,2\}$ . In particular, normalized price beliefs  $\widetilde{NP}_{t,t+k}^i$  of subject  $i$  in period  $t$  are calculated as  $\frac{\widetilde{P}_{t,t+k}^i}{\bar{P}_{t-1}}$ . Here  $\widetilde{P}_{t,t+k}^i$  indicates subject  $i$ 's belief in period  $t$  of the mean market price in  $t+k$ , divided by the average price in the previous period ( $\bar{P}_{t-1}$ ). With this normalization on the last period's mean market price we control for the absolute price level and thereby take a conservative approach for measuring heterogeneous beliefs.

To investigate the impact of heterogeneous beliefs on price efficiency, we run OLS regressions reported in Table 5. Here, we take overpricing  $RD_t$  as dependent variable and include  $SD(\widetilde{NP}_{t,t+k})$ , treatment dummies for all treatments except for Treatment LOW (which is captured with the intercept), and interaction effects for all treatments except for Treatment LOW (e.g.,  $HIGH \times SD(\widetilde{NP}_{t,t+k})$ ), measuring the impact of heterogeneous beliefs in the various treatments on overpricing.

Table 5: **Heterogeneous beliefs and overvaluation.** This table outlines OLS-regressions measuring the impact of heterogeneous beliefs on overvaluation across treatments.  $RD_t$  serves as dependent variable and measures overvaluation in period  $t$ .  $SD(\widetilde{NP}_{t,t+k})$  stands for the standard deviation of normalized price beliefs of all traders in a market in period  $t$  with  $k$  indicating values in the range  $\{0,1,2\}$ . Normalized price beliefs  $\widetilde{NP}_{t,t+k}^i$  of subject  $i$  in period  $t$  are calculated as  $\widetilde{P}_{t,t+k}^i/\overline{P}_{t-1}$ . Moreover, treatment dummies control for level effects and are not reported: INC (increasing CA-ratio), SHORT (increasing CA-ratio, short-selling allowed), HIGH (high and constant CA-ratio of 10.2), LOW (low and constant CA-ratio of 1—captured with the intercept),  $NPK^{PROF}/NPK^{STUD}$  (Treatment HIGH with non public knowledge about composition of market), and  $PK^{MIXED}$  (Treatment HIGH with public knowledge about equal split of professionals/students in the market). Interaction effects of treatment and heterogeneous beliefs (e.g.,  $HIGH \times SD(\widetilde{NP}_{t,t+k})$ ) measure the impact of heterogeneous beliefs in the various treatments with respect to Treatment LOW. \*\* and \*\*\* represent the 5% and 1% significance levels of a double-sided test; for interaction terms we test whether the respective coefficient is different from zero. Clustered standard errors on a market level are provided in parentheses.

Dep. variable: $RD_t$	STUD						PROF					
	t,t		t,t+1		t,t+2		t,t		t,t+1		t,t+2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$SD(\widetilde{NP}_{t,t+k})$	2.575*** (0.669)	0.0154 (0.0354)	2.847*** (0.750)	0.0183 (0.0351)	3.242*** (1.029)	0.0120 (0.0407)	4.693 (2.498)	0.0727 (0.307)	5.469** (2.548)	-0.238*** (0.0602)	5.120 (2.608)	-0.247*** (0.0580)
$INC \times SD(\widetilde{NP}_{t,t+k})$		0.349 (0.753)		1.468** (0.575)		2.138*** (0.676)		0.308 (0.624)		1.074 (0.751)		1.278 (0.700)
$HIGH \times SD(\widetilde{NP}_{t,t+k})$		8.908 (4.995)		9.752** (4.800)		10.54 (6.959)		10.77*** (0.907)		11.54*** (0.802)		11.44*** (2.148)
$SHORT \times SD(\widetilde{NP}_{t,t+k})$		2.002** (0.868)		1.604*** (0.418)		0.788 (0.615)		2.226** (0.952)		2.419*** (0.726)		1.362 (0.701)
$NPK^{STUD} \times SD(\widetilde{NP}_{t,t+k})$		2.831*** (0.470)		3.203*** (0.522)		3.627*** (0.599)						
$NPK^{PROF} \times SD(\widetilde{NP}_{t,t+k})$								2.738** (1.224)		2.575 (1.302)		1.398 (0.753)
$PK^{MIXED} \times SD(\widetilde{NP}_{t,t+k})$		1.241 (1.530)		1.931 (1.133)		2.306*** (0.684)		1.184 (1.563)		2.187 (1.137)		2.565*** (0.687)
Constant	-0.218 (0.144)	0.130*** (0.0426)	-0.298 (0.166)	0.132*** (0.0457)	-0.385 (0.220)	0.140*** (0.0484)	-0.314 (0.240)	0.0904** (0.0377)	-0.410 (0.260)	0.104** (0.0438)	-0.446 (0.301)	0.113** (0.0473)
Treatment dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$N$	1,268	1,268	1,202	1,202	1,138	1,138	1,075	1,075	1,017	1,017	960	960
$R^2$	0.199	0.217	0.212	0.234	0.228	0.260	0.211	0.341	0.265	0.415	0.272	0.441

**Result 7:** *Heterogeneity in beliefs about future prices drives price inefficiency and bubbles.*

Support: Table 5 outlines the results. We find that, on aggregate, heterogeneous price beliefs predict overpricing when interaction terms are dropped. In the specifications with interaction terms, we observe that all coefficients of these terms are positive and most of them are significantly different from Treatment LOW and from zero (when focusing on the joint effects of the interaction terms and  $SD(\widetilde{NP}_{t,t+k})$ , significance levels stay identical to what we observe in Table 5). This particularly holds for treatments of type HIGH, indicating that heterogeneous price beliefs significantly contribute to overpricing. Interestingly, markets in Treatment SHORT show partly significant coefficients of the interaction terms as well, although the magnitude of the coefficients is clearly smaller compared to Treatment HIGH. This suggests that short-selling possibilities increase the diversity of beliefs.

As a robustness check we run a modified regression without treatment dummies, but with a dummy variable BUBBLE, indicating a bubble market following the definition in Section 2.2, and interaction terms with the measure of heterogeneous beliefs. The sample includes all treatments and markets. Table 6 reports strong evidence that heterogeneous beliefs in bubble markets (see coefficients of  $BUBBLE \times SD(BeP_{t,t+k})$ ) are significantly related to overvaluation. This finding applies to both students and professionals, and, in terms of effect size, is particularly relevant for markets populated by professionals. Section C in the Appendix outlines further details on beliefs such as forecast errors and belief time series data.

### 3.2.2 Trading Behavior

Market liquidity is considered to be crucial for the efficient functioning of financial markets, for reducing transaction costs, and for creating firm value (Holmström and Tirole, 1993; Fang et al., 2009). To get a more comprehensive picture how markets with professionals differ from those with students, we analyze several variables that measure market liquidity across subject pools. Table 7 provides treatment medians of various measures including significance tests between

Table 6: **Heterogeneous beliefs, bubble formation and overvaluation.** This table reports OLS regressions measuring the impact of heterogeneous beliefs on overvaluation across all 11 treatments.  $RD_t$  serves as dependent variable and measures overvaluation in period  $t$ .  $SD(\widetilde{NP}_{t,t+k})$  stands for the standard deviation of normalized price beliefs of all traders in a market in period  $t$  with  $k$  indicating values in the range  $\{0,1,2\}$ . Normalized price beliefs  $\widetilde{NP}_{t,t+k}^i$  of subject  $i$  in period  $t$  are calculated as  $\widetilde{P}_{t,t+k}^i/\overline{P}_{t-1}$ . Moreover, BUBBLE is a dummy for bubble markets defined in Section 2.2. Interaction effects of whether a market was a bubble market and heterogeneous beliefs (e.g.,  $BUBBLE \times SD(\widetilde{NP}_{t,t+k})$ ) measure the impact of heterogeneous beliefs in the bubble markets across all treatments. \*\* and \*\*\* represent the 5% and 1% significance levels of a double-sided test. Clustered standard errors at market level are in parentheses.

Dep. variable: $RD_t$	STUD			PROF		
	$t, t$	$t, t+1$	$t, t+2$	$t, t$	$t, t+1$	$t, t+2$
$SD(BeP_{t,t+k})$	0.012 (0.166)	0.070 (0.209)	0.163 (0.235)	0.023 (0.146)	0.093 (0.182)	0.097 (0.162)
BUBBLE	3.559*** (1.018)	3.659*** (1.072)	3.494*** (1.043)	0.627 (0.736)	0.237 (0.594)	-0.0477 (0.375)
$BUBBLE \times SD(BeP_{t,t+k})$	2.198*** (0.809)	2.332*** (0.814)	3.059** (1.331)	9.389*** (1.263)	9.688*** (1.528)	9.201*** (2.827)
Constant	0.362*** (0.100)	0.366*** (0.098)	0.364*** (0.098)	0.147*** (0.026)	0.146*** (0.026)	0.148*** (0.027)
Observations	1,078	1,022	968	885	837	790
$R^2$	0.209	0.222	0.240	0.359	0.414	0.419

experiments PROF and STUD in each treatment. In particular, SPREAD measures the bid-ask spread at the end of a period by using the absolute difference between the best bid and the best ask normalized by the FV.<sup>30</sup> VOLA measures price volatility by using log-returns of all market prices within a period. TURNOVER stands for total trading volume normalized by the total number of shares outstanding (TSO) in a period. Table A2 in Appendix A outlines details on the variables used.

**Table 7: Pairwise Mann-Whitney U-tests of bid-ask spread, intra-period price volatility, and turnover:** 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), NPK<sup>PROF</sup> (Treatment HIGH populated by professionals, but with non-public knowledge about trader composition of market), NPK<sup>STUD</sup> (Treatment HIGH populated by students, but with non-public knowledge about trader composition of market), and PK<sup>MIXED</sup> (Treatment HIGH with public knowledge about equal split of professionals/students in the market). 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. SPREAD is measured as the absolute difference between the best bid and the best ask at the end of a period normalized by the FV. VOLA measures price volatility by using all log-returns of all market prices within a period. TURNOVER stands for total trading volume normalized by the total number of shares outstanding (TSO) in a period. \*\* and \*\*\* represent the 5% and 1% significance levels of a double-sided test. Sample size  $N$  for each test is between 20 and 22.

Treatment	SPREAD			VOLA			TURNOVER		
	PROF	STUD	Z	PROF	STUD	Z	PROF	STUD	Z
INC	5.36	14.77	(1.65)	4.03	8.56	(1.78)	19.38	16.69	(-0.46)
SHORT	1.96	4.48	(1.85)	4.36	7.01	(1.35)	28.11	28.44	(-0.43)
LOW	2.93	7.87	(1.49)	4.38	7.18	(0.50)	10.97	14.83	(0.00)
HIGH	18.75	43.49	(2.37)**	12.93	19.10	(0.66)	22.83	17.00	(-1.65)
NPK	6.28	38.36	(2.72)***	5.17	9.46	(1.89)	26.08	23.03	(-0.08)
PK <sup>MIXED</sup>	17.20			12.18			20.16		

**Result 8:** *On aggregate, markets populated by professionals show higher levels of market liquidity than student markets.*

Support: Based on the results in Table 7 we find a tendency of smaller spreads in markets populated by professionals compared to those of students, indicating more liquid markets when professionals trade. In particular, the median bid-ask spreads at the end of a period in the

<sup>30</sup>As in all other Mann-Whitney U-tests we calculate the market mean across all period means and run statistical tests with the market as unit of observation.

professional markets vary between 0.5 (2.0 percent of the FV) in SHORT and 5.3 (18.8 percent of the FV) in HIGH, whereas treatment medians are larger with 1.3 (4.5 percent) and 12.2 (43.5 percent) in the student markets, respectively. This qualitative pattern holds for each treatment separately with significant subject pool differences in HIGH and NPK<sup>PROF</sup>. Within both subject pools patterns are similar, as the treatment with the highest frequency of bubble markets, HIGH, shows significantly higher spreads compared to all other base treatments (see tables A16, A17, A18 in Appendix A for details). The price volatility measure VOLA goes in a qualitatively similar direction within and across subject pools (and is naturally correlated with SPREAD by construction), but the differences between professionals and students are not statistically significant at a 5 percent level. Although effect sizes on VOLA appear to be large, there is, in addition to low sample sizes in the tests, strong heterogeneity across markets within treatments. Overall, these results indicate that professionals provide more liquidity to the market compared to students. This leads to lower bid-ask spreads and to higher order book depth, reducing transaction costs and the possibility for individuals to idiosyncratically drive prices for speculative purposes. In Table A19 in Appendix A we run tests on additional variables such as the submission rate SR (the number of limit orders posted divided by the sum of limit and market orders posted in a period), LIQUIDITY (the quantity of all open bids and asks at the end of a period normalized by total shares outstanding), and DEPTH (the average percentage log-difference among the best five bids and best five asks, respectively, in the order book). We again find that markets populated by professionals are more liquid than the student counterparts.

With respect to turnover we find no significant differences between the two subject pools, and no clear patterns. However, as several theoretical and empirical papers predict higher trading volume in the run-up of a bubble than after its crash (e.g., Lee and Swaminathan, 2000; Scheinkman and Xiong, 2003; Baker and Stein, 2004; Hong and Stein, 2007; Barberis, 2018), we explore this in an additional analysis on the development of trading volumes in our markets. We

find a moderate pattern of higher trading volume before than after the price peak – in line with the above-mentioned theoretical and empirical literature. In Table A20 in Appendix A we show that the effects are clearly more pronounced for students than for professionals. However, we would like to emphasize that conclusions on trading volume in our markets should be taken with caution, because of the pronounced idiosyncratic characteristics of individual markets, resulting in large level differences of trading volume across markets. In combination with the limited number of markets, this leaves us with a substantial fraction of insignificant results although effect sizes appear to be large.

## 4 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. In total, we ran 10 mixed markets, 48 asset markets with financial professionals from high-skilled investment areas, and 58 markets with student subjects without professional market experience. We set up two classical bubble-driver treatments by either implementing a high initial level of the monetary supply relative to the asset value or by administering capital inflows over time. We also implemented two classical bubble-moderator treatments by either allowing short-selling or by keeping the level of capital inflow constant and low over time. In addition to these four base treatments, we ran three treatments to analyze whether beliefs about the rationality of others or superior skills of professionals drive differences in price efficiency between subject pools. Finally, we administered an extensive survey to measure several cognitive skills (fluid intelligence, cognitive reflection, theory of mind, backward induction).

We found that professionals are not immune to bubbles in experimental asset markets. In fact, across the two bubble-driver (base) 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 showing that professionals generate market inefficiencies and bubbles, even in relatively simple and controlled market environments with only one trade-able 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 comparing professionals' behavior with that of students (*across* subject pools), we found that markets populated by professionals generated less overpricing, fewer bubbles, and smaller bubbles 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.

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 likelihood for 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 running long-lived laboratory markets with student subjects: even though the treatment effect sizes are smaller for professionals and bubbles are less likely, qualitatively, the effects are comparable in direction and statistical significance.

Finally, we probed for potential drivers of the results with three additional treatments and an online survey. We found that beliefs about the rationality of others do not influence price

efficiency significantly, but that professionals – similar to insiders in markets with asymmetric information (e.g., Plott and Sunder, 1988) – can act as a stabilizing force. We found that mixed markets with an equal split of subjects from both pools and public knowledge about the trader composition show levels of price efficiency similar to markets solely populated by professionals (also with public knowledge about the trader composition). This finding is remarkable as we did not find pronounced differences in cognitive skills between professionals and students. Moreover, none of these skills were able to explain trader performance or overvaluation in the markets. We did find, however, that heterogeneity in future price beliefs affected efficiency in our laboratory markets; a finding that applies to both students and professionals, but which was particularly pronounced for the professional sample and for the most bubble-prone treatments of type HIGH.

In our setting, despite some overpricing in bubble-prone market environments, professionals act as price stabilizers, even in mixed markets with students, where they constitute only a fraction of all traders. Our results suggest that professionals' higher level of price efficiency is not due to superior cognitive skills and, hence, also unlikely to be due to a better cognitive understanding of the experimental task. We readily acknowledge that the lack of evidence for differences in cognitive skills cannot serve as proof, but it provides indications that professionals' superior price efficiency could be due to a more intuitive understanding of market developments and to professional expertise that goes beyond specific cognitive skills. This expertise arguably originates from real-world market experience, including, for instance, experience with price dynamics, and with financial investments and trading in general, all of which may lead to a more intuitive understanding of what “drives” a market. Although we cannot pin down a singular determinant (if that exists at all), our results suggest the exclusion of a number of candidate explanations such as risk attitudes, beliefs about the rationality of others, and specific cognitive skills (measured with classic tasks). Moreover, we can identify heterogeneous beliefs about future prices as a strong predictor of price inefficiencies, particular for professionals in bubble-prone

market environments. Future research may want to focus on this as a fruitful avenue for further insights.

## 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.
- Akiyama, Eizo, Nobuyuki Hanaki, Ryuichiro Ishikawa. 2017. It is not just confusion! strategic uncertainty in an experimental asset market. *Economic Journal* **127** F563–F580.
- 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.
- Baker, Malcom, Jeremy Stein. 2004. Market liquidity as a sentiment indicator. *Journal of Financial Markets* **3**(1) 271–299.
- Barberis, Nicholas. 2018. Psychology-based models of asset prices and trading volume. *NBER Working Paper No. 24723* 1–99.
- 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.
- Bilker, Warren B, John A Hansen, Colleen M Brensinger, Jan Richard, Raquel E Gur, Ruben C Gur. 2012. Development of abbreviated nine-item forms of the raven’s standard progressive matrices test. *Assessment* **19**(3) 354–369.
- 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 Dictionary 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, Øyvind 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.
- Corngnet, 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.
- DeMartino, Benedetto, John O’Doherty, Debajyoti Ray, Peter Bossaerts, Colin Camerer. 2013. In the mind of the market: Theory of mind biases value computation during financial bubbles. *Neuron* **79** 1222–1231.
- 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.
- Fang, Vivian W., Thomas H. Noe, Sheri Tice. 2009. Stock market liquidity and firm value. *Journal of Financial Economics* **94**(1) 150–169.

- 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.
- Greenwood, Robin, Andrei Shleifer, Yang You. 2019. Bubble for fama. *Journal of Financial Economics* **131** 20–43.
- 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.
- Hefti, Andreas, Steve Heinke, Frederic Schneider. 2016. Mental capabilities, trading styles, and asset market bubbles: Theory and experiment. *Working Paper* .
- Hogan, Steve, Robert Jarrow, Melvyn Teo, Mitch Warachka. 2004. Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial Economics* **73**(3) 525–565.
- Holmström, Bengt, Jean Tirole. 1993. Market liquidity and performance monitoring. *Journal of Political Economy* **101**(4) 678–709.
- Holt, Charles A., Megan Porzio, Michelle Yingze Song. 2017. Price bubbles, gender, and expectations in experimental asset markets. *European Economic Review* **100** 72–94.
- Hong, Harrison, Jeremy Stein. 2007. Disagreement and the stock market. *Journal of Economic Perspectives* **21**(2) 109–128.
- 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.
- Lee, Charles, Bhaskaran Swaminathan. 2000. Price momentum and trading volume. *Journal of Finance* **55**(5) 2017–2069.

- 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.
- Moinas, S., S. Pouget. 2013. The bubble game: an experimental study of speculation. *Econometrica* **81**(4) 1507–1539.
- Morris, Stephen. 1996. Speculative investor behavior and learning. *Quarterly Journal of Economics* **111** 1111–1133.
- 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.
- Olderbak, Sally, Oliver Wilhelm, Gabriel Olaru, Mattis Geiger, Meghan W Brenneman, Richard D Roberts. 2015. A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. *Frontiers in Psychology* **6** 1503.
- Palan, Stefan. 2015. GIMS – Software for asset market experiments. *Journal of Behavioral and Experimental Finance* **5** 1–14.
- Palfrey, Thomas R., Stephanie W. Wang. 2012. Speculative overpricing in asset markets with information flows. *Econometrica* **80**(5) 1937–1976.
- Plott, Charles R., Shyam Sunder. 1988. Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica* **56**(5) 1085–1118.
- 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 Stráznická, 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.
- Scheinkman, Jose A., Wei Xiong. 2003. Overconfidence and speculative bubbles. *Journal of Political Economy* **111** 1183–1219.

- 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. 2013. Bubbles, crises, and heterogeneous beliefs. *NBER Working Paper No. 18905*.
- 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 the base and additional treatments:** 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	NPK <sup>PROF</sup>	NPK <sup>STUD</sup>	PK <sup>MIXED</sup>
Number of periods	20	20	20	20	20	20	20
Number of traders per market	8	8	8	8	8	8	8
Buyback price	28 Taler	28	28	28	28	28	28
Interest rate	5%	5%	5%	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	1.20; 1.60	1.20; 1.60	1.20; 1.60
Probabilities of the dividends	50%; 50%	50%; 50%	50%; 50%	50%; 50%	50%; 50%	50%; 50%	50%; 50%
Initial endowment (subject)	20 shares / 560 Taler	20 / 560	20 / 560	20 / 5,700	20 / 5,700	20 / 5,700	20 / 5,700
Additional cash per period (subject)	100 Taler	100	0	0	0	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	10.2/10.2/10.2	10.2/10.2/10.2	10.2/10.2/10.2
Short-selling capacity (subject)	0 shares	-40	0	0	0	0	0

The planned size of markets with professionals in the base treatments 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 9<sup>th</sup> 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 a few markets with 7 or 9 traders.

Table A2: **Variables measuring price efficiency:** RAD measures mispricing and is calculated as the absolute difference of mean period prices and FVs averaged across all periods of a market. 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 maximum average period price with the preceding minimum average period price, normalized by the FV. CRASH measures the severity of a crash by taking the difference between the minimum price after the peak and the peak average price, normalized at the FV. SPREAD calculates the absolute difference between the best bid and the best ask at the end of a period normalized by the FV. VOLA measures log-returns of all market prices within a period. TURNOVER stands for total trading volume normalized by the total number of shares outstanding (TSO) in a period. SR is defined as the number of limit orders posted divided by the sum of limit and market orders posted in a period. LIQUIDITY defines the quantity of all open bids and asks at the end of a period normalized by TSO. DEPTH measures the average percentage difference (log) among the best five bids and best five asks in the order book at the end of a period. All variables below are calculated on the market level.

Relative absolute deviation	$\text{RAD} = \sum_{t=1}^T \frac{ \bar{P}_t - FV_t }{T}$
Relative deviation	$\text{RD} = \sum_{t=1}^T \frac{\bar{P}_t - FV_t}{T}$
Peak price	$\text{RDMAX} = \max_t \left\{ \frac{\bar{P}_t - FV_t}{FV_t} \right\} = \frac{\bar{P}_{t^*} - FV_{t^*}}{FV_{t^*}}$
Amplitude	$\text{AMPLITUDE} = \frac{\bar{P}_{t^*} - FV_{t^*}}{FV_{t^*}} - \min_{0 \leq k < t^*} \left\{ \frac{\bar{P}_{t^* - k} - FV_{t^* - k}}{FV_{t^* - k}} \right\}$
Crash	$\text{CRASH} = \min_{0 \leq l \leq T - t^*} \left\{ \frac{\bar{P}_{t^* + l} - FV_{t^* + l}}{FV_{t^* + l}} \right\} - \frac{\bar{P}_{t^*} - FV_{t^*}}{FV_{t^*}}$
Bid-ask spread	$\text{SPREAD} = \sum_{t=1}^T \frac{1}{FV_t} \frac{1}{T} \left[ \min_{j \in N_t} \{S_{t,j}\} - \max_{j \in N_t} \{B_{t,j}\} \right]$
Intra-period price volatility	$\text{VOLA} = \sum_{t=1}^T \frac{1}{T} \sqrt{\frac{1}{N_t} \sum_{j=1}^{N_t} (R_{t,j} - \bar{R}_t)^2}$
Turnover	$\text{TURNOVER} = \sum_{t=1}^T \frac{1}{T} \frac{VOL_t}{TSO}$
Submission rate	$\text{SR} = \sum_{t=1}^T \sum_{j=1}^{N_t} \frac{1}{T} \frac{LO_{j,t}}{LO_{j,t} + MO_{j,t}}$
Order book liquidity	$\text{LIQUIDITY} = \frac{1}{TSO} \sum_{t=1}^T \sum_{o=1}^{O_t} \frac{1}{T} O_o^j$
Order book depth	$\text{DEPTH} = \sum_{t=1}^T \frac{1}{T} \frac{1}{8} \left( \ln\left(\frac{B_t^1}{B_t^2}\right) + \dots + \ln\left(\frac{B_t^4}{B_t^5}\right) + \left  \ln\left(\frac{S_t^1}{S_t^2}\right) \right  + \dots + \left  \ln\left(\frac{S_t^4}{S_t^5}\right) \right  \right)$

$\bar{P}_t$  is the volume-weighted mean price in period  $t$ ;  $FV_t$  is the fundamental value in period  $t$ ;  $t^*$  denotes the period with the highest volume-weighted mean price. Limit orders  $LO_t$  are the number of shares offered to trade in period  $t$ . Market order  $MO_t$  are the number of shares traded based on accepted  $LO$ s posted by other subjects in period  $t$ . log-return of a trade:  $R_{t,j} = \ln(P_{t,j}/P_{t,j-1})$ ; total number of trades in period  $t$ :  $N_t$ ; average log return in period  $t$ :  $\bar{R}_t$ ; price of sell order  $j$  at the end of period  $t$ :  $S_{t,j}$ ; price of buy order  $j$  at the end of period  $t$ :  $B_{t,j}$ . number of open (buy and sell) orders at the end of period  $t$ :  $O_t$ ; Quantity offered in (buy or sell) order  $o$ :  $O_o^j$ ;  $n$ -th best buy order at the end of period  $t$ :  $B_t^n$ ;  $n$ -th best sell order at the end of period  $t$ :  $S_t^n$ .

Table A3: **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
<i>Threshold for bubble classification</i>			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
<b>M8**</b>	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
Experiment STUD					
<b>M1**</b>	229.3	229.3	632.3*	601.8*	-625.5*
<b>M2**</b>	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
<b>M6**</b>	48.2	48.2	105.6*	87.4*	-101.2*
M7	3.8	2.8	4.8	2.6	-12.9
<b>M8**</b>	51.1	51.1	124.9*	114.6*	-123.0*
M9	31.5	31.5	46.3	38.8	-44.8
<b>M10**</b>	80.7	80.7	376.5*	366.4*	-374.0*
M11	17.0	17.0	21.0	0.0	-20.4
<b>M12**</b>	96.9	96.9	223.2*	203.5*	-221.1*

Table A4: **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	AMPLITUDE	CRASH
<i>Threshold for bubble classification</i>			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
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
<b>M12**</b>	181.0	181.0	497.5*	491.2*	-492.5*

Table A5: **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	AMPLITUDE	CRASH
<i>Threshold for bubble classification</i>			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
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

Table A6: **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	AMPLITUDE	CRASH
<i>Threshold for bubble classification</i>			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
<b>M3**</b>	1336.7	1336.7	2215.3*	2186.8*	-1450.3*
<b>M4**</b>	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
<b>M8**</b>	120.7	120.7	207.5*	191.1*	-198.3*
M9	30.5	30.4	55.4	43.8	-55.8
<b>M10**</b>	71.0	71.0	114.7*	57.2*	-111.1*
Experiment STUD					
M1	37.7	37.7	60.4	59.4*	-60.5
<b>M2**</b>	869.8	869.8	1732.1*	1524.9*	-1729.2*
<b>M3**</b>	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
<b>M7**</b>	1478.2	1478.2	7727.4*	7674.3*	-7723.8*
<b>M8**</b>	1118.9	1118.9	1296.2*	75.5*	-1244.6*
<b>M9**</b>	106.2	106.2	158.9*	82.8*	-158.9*
<b>M10**</b>	394.2	394.2	831.3*	147.6*	-827.3*
<b>M11**</b>	138.9	138.0	261.1*	227.1*	-269.0*
<b>M12**</b>	165.5	165.3	286.2*	86.9*	-286.9*

Table A7: **Bubble measures for all markets in Treatment NPK<sup>PROF</sup>**: 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). RD\_MAX 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. RD\_MAX, 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	AMPLITUDE	CRASH
			94.5	44.8	-98.8
<b>M1**</b>	65.9	65.9	188.9*	185.8*	-188.7*
<b>M2**</b>	59.0	59.0	121.6*	95.9*	-106.8*
M3	9.8	9.8	16.1	8.9	-12.5
M4	41.0	41.0	55.1	27.6	-55.1
<b>M5**</b>	66.4	66.4	156.4*	88.2*	-153.3*
M6	4.2	4.2	6.8	4.6	-6.8
M7	7.9	7.9	10.6	9.0	-7.4
M8	28.1	28.1	58.6	51.2*	-31.3
M9	2.6	-0.2	15.1	0.0	-37.1
M10	3.0	3.0	7.1	6.3	-7.1

Table A8: **Bubble measures for all markets in Treatment NPK<sup>STUD</sup>**: 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). RD\_MAX 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. RD\_MAX, 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	AMPLITUDE	CRASH
<i>Threshold for bubble classification</i>			94.5	44.8	-98.8
<b>M1**</b>	92.8	92.8	155.3*	121.0*	-152.3*
<b>M2**</b>	715.8	715.8	838.0*	88.7*	-513.5*
M3	52.3	52.3	87.2	78.0*	-11.4
<b>M4**</b>	279.0	279.0	387.9*	260.3*	-373.7*
<b>M5**</b>	867.4	867.4	1320.4*	287.3*	-1319.6*
<b>M6**</b>	229.4	227.8	623.4*	262.6*	-632.5*
M7	65.0	63.8	99.9*	110.9*	-47.5
<b>M8**</b>	171.1	171.0	381.7*	327.3*	-382.6*
M9	6.4	3.9	8.2	32.1	-7.5
<b>M10**</b>	80.2	80.2	113.1*	90.7*	-110.6*

Table A9: **Bubble measures for all markets in Treatment PK<sup>MIXED</sup>**: 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). RD\_MAX 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. RD\_MAX, 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	AMPLITUDE	CRASH
<i>Threshold for bubble classification</i>			94.5	44.8	-98.8
M1	48.0	48.0	75.2	42.0	-70.5
<b>M2**</b>	39.3	39.3	113.7*	111.9*	-106.8*
M3	1.0	0.7	1.9	4.4	-1.7
M4	135.8	135.8	223.6*	0.0	-219.0*
M5	48.1	48.1	55.3	33.0	-9.0
<b>M6**</b>	209.0	209.0	371.6*	323.7*	-367.8*
<b>M7**</b>	352.2	352.2	551.4*	381.2*	-546.3*
<b>M8**</b>	152.6	152.6	364.3*	352.6*	-351.9*
M9	21.0	21.0	51.3	45.2*	-46.8
M10	73.0	71.4	123.1*	29.0	-140.0*

Table A10: **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 5% and 1% significance levels of a double-sided test. Sample size  $N$  for each test is between 18 and 20.

Treatment	RAD			RD			RDMAX		
	SHORT	LOW	HIGH	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	6.72** (2.37)	8.23** (2.29)	-47.79 (-1.36)	6.74** (2.37)	10.50** (2.29)	-44.36 (-0.91)	18.80** (1.96)	26.02** (2.12)	-73.00 (-1.21)
SHORT	.	1.51 (0.22)	-54.51*** (2.86)	.	3.76 (-0.75)	-51.10*** (2.61)	.	7.22 (0.13)	-91.80** (2.53)
LOW	.	.	-56.01*** (-2.86)	.	.	-54.86*** (-2.78)	.	.	-99.02** (-2.53)
Treatment	AMPLITUDE			CRASH					
	SHORT	LOW	HIGH	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	24.78 (1.88)	32.36** (2.04)	-29.26 (-1.44)	-9.68 (-1.55)	16.02 (-0.33)	99.73 (1.21)			
SHORT	.	7.58 (-0.66)	-54.04*** (2.78)	.	25.69 (-1.19)	109.40** (-2.45)			
LOW	.	.	-61.62*** (-2.86)	.	.	83.71 (1.71)			

Table A11: **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 5% and 1% significance levels of a double-sided test. Sample size  $N$  for each test is 24.

Treatment	RAD			RD			RDMAX		
	SHORT	LOW	HIGH	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	32.48*** (2.66)	34.58*** (3.06)	-261.55*** (-2.89)	34.03*** (2.71)	37.27*** (3.18)	-261.51*** (-2.89)	51.66** (2.31)	56.78 (1.91)	-407.43** (-2.54)
SHORT	.	2.09 (-0.12)	-294.03*** (3.35)	.	3.24 (-0.35)	-295.54*** (3.35)	.	5.12 (0.46)	-459.08** (3.35)
LOW	.	.	-296.13*** (-3.46)	.	.	-298.78*** (-3.58)	.	.	-464.21*** (-3.41)
Treatment	AMPLITUDE			CRASH					
	SHORT	LOW	HIGH	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	66.33 (1.71)	71.88*** (3.00)	-100.86 (-1.53)	-34.13 (-1.85)	-21.34 (-1.33)	305.91** (2.25)			
SHORT	.	5.55 (-1.21)	-167.19*** (2.81)	.	12.80 (-1.15)	340.05*** (-3.23)			
LOW	.	.	-172.74*** (-3.58)	.	.	327.25*** (3.12)			

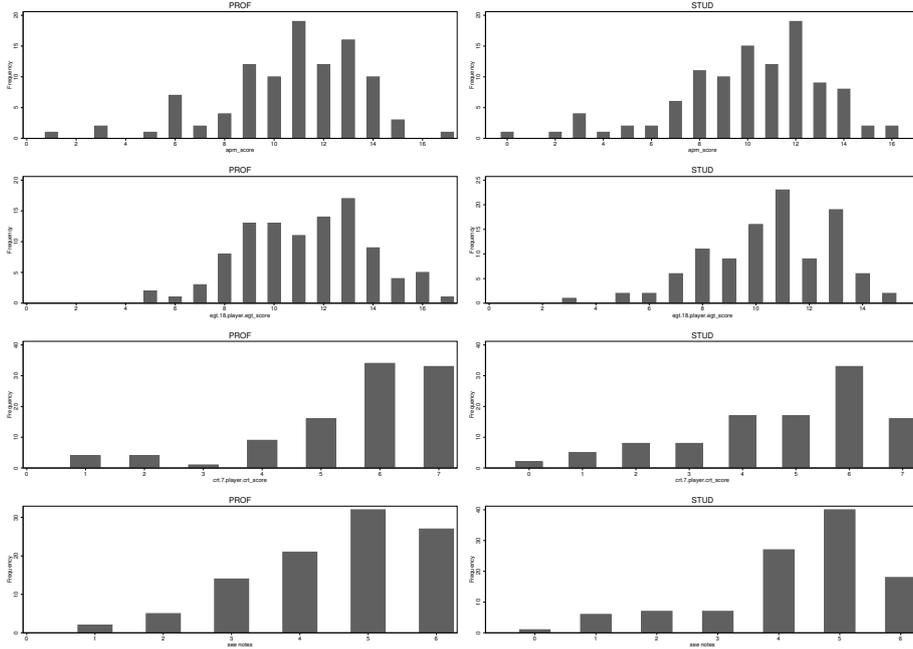
Table A12: **Pairwise Mann-Whitney U-tests of mispricing (RAD), overpricing (RD), maximum overpricing (RDMAX), price run-ups (AMPLITUDE), and crash (CRASH) between the additional treatments NPK<sup>PROF</sup>, NPK<sup>STUD</sup>, and PK<sup>MIXED</sup>**: This table shows pairwise treatment comparisons for treatments HIGH (high and constant CA-ratio of 10.2) of both samples, NPK<sup>PROF</sup> (Treatment HIGH populated by professionals, but with non-public knowledge about trader composition of market), NPK<sup>STUD</sup> (Treatment HIGH populated by students, but with non-public knowledge about trader composition of market), and PK<sup>MIXED</sup> (Treatment HIGH with public knowledge about equal split of professionals/students in the market). The table outlines median treatment differences in percentage points and the numbers in parentheses show the Z-values of the MW U-test statistic. \*\* and \*\*\* represent the 5% and 1% significance levels of a double-sided test. Sample size  $N$  for each test is between 19 and 22.

Pairwise comparisons	Variable				
	<b>RAD</b>	<b>RD</b>	<b>RDMAX</b>	<b>AMPLITUDE</b>	<b>CRASH</b>
NPK <sup>PROF</sup> vs. NPK <sup>STUD</sup>	-113.05***	-113.01***	-232.88**	-97.61***	228.79**
NPK <sup>PROF</sup> vs. PK <sup>MIXED</sup>	-41.64	-40.80	-82.79	-25.27	89.20
NPK <sup>PROF</sup> vs. HIGH (PROF)	-42.54	-38.75	-72.59	-52.47**	94.43
NPK <sup>STUD</sup> vs. PK <sup>MIXED</sup>	71.41	72.21	150.09	72.34	-139.59
NPK <sup>STUD</sup> vs. HIGH (STUD)	-177.32	-177.33	-229.20	-61.90	130.68
PK <sup>MIXED</sup> vs. HIGH (PROF)	-0.90	2.05	10.20	-27.19	5.23
PK <sup>MIXED</sup> vs. HIGH (STUD)	-248.74**	-249.54**	-379.29**	-134.24	270.27

Table A13: **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 professionals and students in surveys EXP<sup>COGRISK</sup> and ONLINE<sup>COGRISK</sup>**: 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 5% and 1% significance levels of a double-sided test. Sample size  $N$  for the tests in the data EXP<sup>COGRISK</sup> is 207 (101 professionals and 106 students) and for the tests in the data ONLINE<sup>COGRISK</sup> it is 245 (121 professionals and 124 students).

EXP <sup>COGRISK</sup>						
Cognitive Skill Task		Mean	SD	Min.	Max.	Z
RAVEN	PROF	10.71	2.85	1.00	17.00	(-1.465)
	STUD	10.10	3.11	0.00	16.00	
CRT	PROF	5.60	1.56	1.00	7.00	(-3.628)***
	STUD	4.79	1.81	0.00	7.00	
ToM	PROF	11.30	2.57	5.00	17.00	(-1.916)
	STUD	10.55	2.35	3.00	15.00	
HIT15	PROF	4.55	1.27	1.00	6.00	(-1.134)
	STUD	4.31	1.40	0.00	6.00	
<b>Risk Attitudes</b>						
General Risk	PROF	5.29	1.92	2.00	9.00	(1.505)
	STUD	5.74	2.00	1.00	10.00	
Financial Risk	PROF	5.70	2.07	1.00	10.00	(-1.723)**
	STUD	5.25	2.08	0.00	10.00	
ONLINE <sup>COGRISK</sup>						
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	
<b>Risk Attitudes</b>						
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	

## EXP<sup>COGRISK</sup>



## ONLINE<sup>COGRISK</sup>

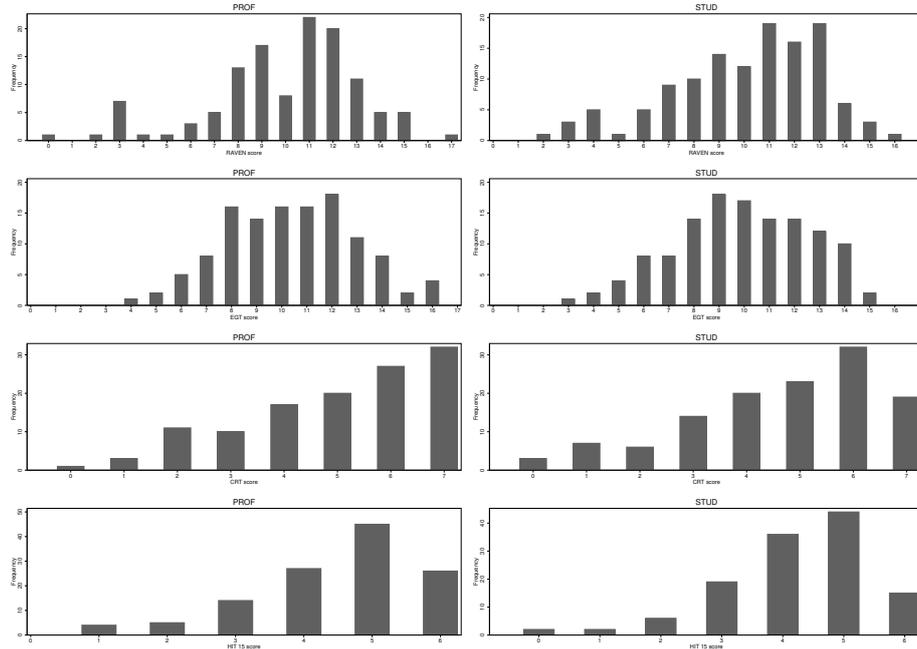


Figure A1: **Histograms of distributions of scores in the different tasks for the data of surveys EXP<sup>COGRISK</sup> (top) and ONLINE<sup>COGRISK</sup> (bottom):** The respective left panels depict the distribution of scores in the different tasks (from top to bottom: RAVEN, ToM, CRT, HIT15) for professionals; the right panels show the respective distributions for students. The data consists of 101 professional subjects and 106 student subjects for EXP<sup>COGRISK</sup> and 121 professionals and 124 students for ONLINE<sup>COGRISK</sup>.

Table A14: **Cognitive abilities and individual performance in the market:** This table outlines OLS-regressions measuring the impact of cognitive skills as well as financial risk taking and years in industry on individual market earnings in EUR for professionals, students, and the pooled data of treatments NPK<sup>PROF</sup>, NPK<sup>STUD</sup>, and PK<sup>MIXED</sup>. EARNINGS\_MARKET is the dependent variable and denotes the final earnings from the market experiment (incentives for beliefs are not included). RAVEN is a measure of fluid intelligence (Raven's advanced progressive matrices), CRT measures cognitive reflection, ToM (eye gaze test) is an indicator of theory of mind, whereas backward induction skills are measured by the HIT15 task. Moreover, FIN\_RISK represents answers on the survey question concerning financial risk taking taken from the German Socio-Economic Panel and YEARS\_INDUSTRY stands for years in industry. The matched market and survey data stems from EXP<sup>COGRISK</sup> data. In contrast to the EXP<sup>COGRISK</sup> data, we report results of 165 observations because of matching issues. \*\* and \*\*\* represent the 5% and 1% significance levels of a double-sided test. Clustered standard errors on a market level are provided in parentheses.

Dep. variable:					
EARNINGS_MARKET	PROF			STUD	
	(1)	(2)	(3)	(1)	(2)
RAVEN	0.000 (.028)	0.000 (0.030)	-0.002 (0.030)	-0.015 (0.052)	-0.014 (0.054)
CRT	0.058 (0.056)	0.061 (0.058)	0.064 (0.059)	0.179 (0.142)	0.164 (0.181)
ToM	0.009 (0.027)	0.013 (0.028)	0.010 (0.029)	0.048 (0.093)	0.044 (0.107)
HIT15	-0.046 (0.068)	-0.042 (0.072)	-0.043 (0.075)	0.017 (0.117)	0.024 (0.134)
FIN_RISK		-0.032 (0.045)	-0.028 (0.046)		-0.039 (0.130)
YEARS_INDUSTRY			-0.003 (0.007)		
Constant	4.248*** (0.622)	4.352*** (0.571)	4.392*** (0.553)	3.330*** (1.009)	3.600** (1.838)
Observations	84	84	84	81	81
R <sup>2</sup>	0.016	0.022	0.024	0.036	0.038

Table A15: **Average cognitive abilities of traders in markets and overvaluation RD:** This table outlines OLS-regressions measuring the impact of cognitive skills, financial risk taking and years in industry of all subjects in a market on overvaluation RD of the pooled data set of treatments NPK<sup>PROF</sup>, NPK<sup>STUD</sup>, and PK<sup>MIXED</sup>. RD is the dependent variable and measures overvaluation. For all cognitive variables, means across all traders in the particular market are calculated: RAVEN is a measure of fluid intelligence (Raven’s advanced progressive matrices), CRT measures cognitive reflection, ToM (eye gaze test) is an indicator of theory of mind, whereas backward induction skills are measured by the HIT15 task. Moreover, FIN\_RISK represents answers on the survey question concerning financial risk taking taken from the German Socio-Economic Panel and YEARS\_INDUSTY stands for years in industry, averaged across all professionals in a market. We report data of 25 markets because of matching issues. \*\* and \*\*\* represent the 5% and 1% significance levels of a double-sided test. Clustered standard errors on a market level are provided in parentheses.

Dep. variable: RD	(1)	(2)
RAVEN	13.768 (34.227)	26.320 (37.078)
CRT	-15.308 (33.408)	17.203 (38.267)
ToM	6.249 (29.125)	-14.582 (33.694)
HIT15	-66.780 (86.528)	-30.706 (90.387)
FIN_RISK		-49.193 (60.025)
YEARS_INDUSTY		-12.209 (6.209)
Constant	276.943 (314.421)	382.508 (268.047)
Observations	25	25
$R^2$	0.055	0.261

Table A16: **Pairwise Mann-Whitney U-tests of bid-ask spread, intra-period price volatility, and turnover 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 5% and 1% significance levels of a double-sided test. Sample size  $N$  for each test is between 18 and 20.

Treatment	SPREAD			VOLA			TURNOVER		
	SHORT	LOW	HIGH	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	3.41 (1.55)	2.43 (1.63)	-13.38** (-2.04)	-0.33 (0.57)	-0.36 (-0.16)	-8.90** (-2.57)	-8.73 (-1.55)	8.41 (1.71)	-3.45 (-0.83)
SHORT	.	-0.98 (0.49)	-16.79*** (2.86)	.	-0.02 (0.66)	-8.57*** (3.02)	.	17.14*** (-2.69)	5.28 (-1.14)
LOW	.	.	-15.81*** (-3.02)	.	.	-8.55 (-1.55)	.	.	-11.86** (-2.37)

Table A17: **Pairwise Mann-Whitney U-tests of bid-ask spread, intra-period price volatility, and turnover 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 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 5% and 1% significance levels of a double-sided test. Sample size  $N$  for each test is 24.

Treatment	SPREAD			VOLA			TURNOVER		
	SHORT	LOW	HIGH	SHORT	LOW	HIGH	SHORT	LOW	HIGH
INC	10.29 (1.73)	6.90 (1.73)	-28.72 (-2.83)	1.56 (1.10)	1.39 (0.00)	-10.54** (-2.48)	-11.75** (-2.19)	1.86 (1.30)	-0.31 (0.35)
SHORT	.	-3.39 (0.17)	-39.01*** (3.18)	.	-0.17 (0.64)	-12.10*** (2.83)	.	13.61*** (-2.92)	11.44** (-2.19)
LOW	.	.	-35.62*** (-3.58)	.	.	-11.92** (-2.42)	.	.	-2.17 (-0.92)

Table A18: **Pairwise Mann-Whitney U-tests of bid-ask spread, intra-period price volatility, and turnover in the additional treatments NPK<sup>PROF</sup>, NPK<sup>STUD</sup>, and PK<sup>MIXED</sup>**: This table shows pairwise treatment comparisons for treatments HIGH (high and constant CA-ratio of 10.2) of both samples, NPK<sup>PROF</sup> (Treatment HIGH populated by professionals, but with non-public knowledge about trader composition of market), NPK<sup>STUD</sup> (Treatment HIGH populated by students, but with non-public knowledge about trader composition of market), and PK<sup>MIXED</sup> (Treatment HIGH with public knowledge about equal split of professionals/students in the market). The table outlines median treatment differences in percentage points. \*\* and \*\*\* represent 5% and 1% significance levels of a double-sided test. Sample size  $N$  for each test is between 19 and 22.

Pairwise comparisons:	Variable		
	SPREAD	VOLA	TURNOVER
NPK <sup>PROF</sup> vs. NPK <sup>STUD</sup>	-30.05**	-4.29	2.98
NPK <sup>PROF</sup> vs. PK <sup>MIXED</sup>	-11.07	-7.01	6.49
NPK <sup>PROF</sup> vs. HIGH (PROF)	-12.61	-7.76	3.82
NPK <sup>STUD</sup> vs. PK <sup>MIXED</sup>	18.98	-2.72	3.51
NPK <sup>STUD</sup> vs. HIGH (STUD)	-7.49	-9.64	6.24
PK <sup>MIXED</sup> vs. HIGH (PROF)	-1.54	0.75	-2.67
PK <sup>MIXED</sup> vs. HIGH (STUD)	-26.47	-6.92	2.73

Table A19: **Pairwise Mann-Whitney U-tests of additional variables measuring market liquidity 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), NPK<sup>PROF</sup> (Treatment HIGH populated by professionals, but with non-public knowledge about trader composition of market), NPK<sup>STUD</sup> (Treatment HIGH populated by students, but with non-public knowledge about trader composition of market), and PK<sup>MIXED</sup> (Treatment HIGH with public knowledge about equal split of professionals/students in the market). 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. SR is defined as the number of limit orders posted divided by the sum of limit and market orders posted in a period. LIQUIDITY defines the quantity of all open bids and asks at the end of a period normalized by TSO and DEPTH measures the average percentage log-difference among the best five bids and best five asks, respectively, in the order book at the end of a period. \*\* and \*\*\* represent the 5% and 1% significance levels of a double-sided test. Sample size  $N$  for each test is between 20 and 22.

Treatment	SR			LIQUIDITY			DEPTH		
	PROF	STUD	Z	PROF	STUD	Z	PROF	STUD	Z
INC	78.63	73.97	(1.65)	22.64	18.97	(0.40)	4.87	8.19	(-1.19)
SHORT	76.24	70.11	(1.21)	18.56	16.71	(0.85)	2.67	7.92	(-1.85)
LOW	82.20	75.43	(2.20)**	12.00	7.84	(2.70)**	5.61	6.88	(-1.28)
HIGH	80.33	74.38	(1.85)	40.28	11.93	(3.23)***	7.11	11.35	(-2.24)**
NPK	81.65	73.55	(2.12)**	47.85	11.80	(2.57)***	5.36	5.36	(-0.15)
PK <sup>MIXED</sup>	78.21			21.17			7.63		

Table A20: **Pairwise Mann-Whitney U-tests of turnover for the pre-peak and the post-peak phases in all markets, separated for professionals (top) and students (bottom):** This table outlines average turnover (trading volume divided by total shares outstanding) per period before and after the price peak across all markets and separated for all, bubble, and non-bubble markets. The numbers in parentheses show the Z-values of the MW U-test statistic. \*\* and \*\*\* represent the 5% and 1% significance levels of a double-sided test.

	Before	After	
	Price Peak ( $t \leq t^*$ )	Price Peak ( $t > t^*$ )	Z
Professionals			
All markets	26.98	24.07	(1.47)
Bubble markets	25.70	22.27	(1.39)
Non-bubble markets	27.28	24.46	(0.48)
Students			
All markets	26.38	20.91	(2.67)***
Bubble markets	24.57	17.68	(2.07)**
Non-bubble markets	27.19	22.48	(1.81)

## B Individual Price Graphs

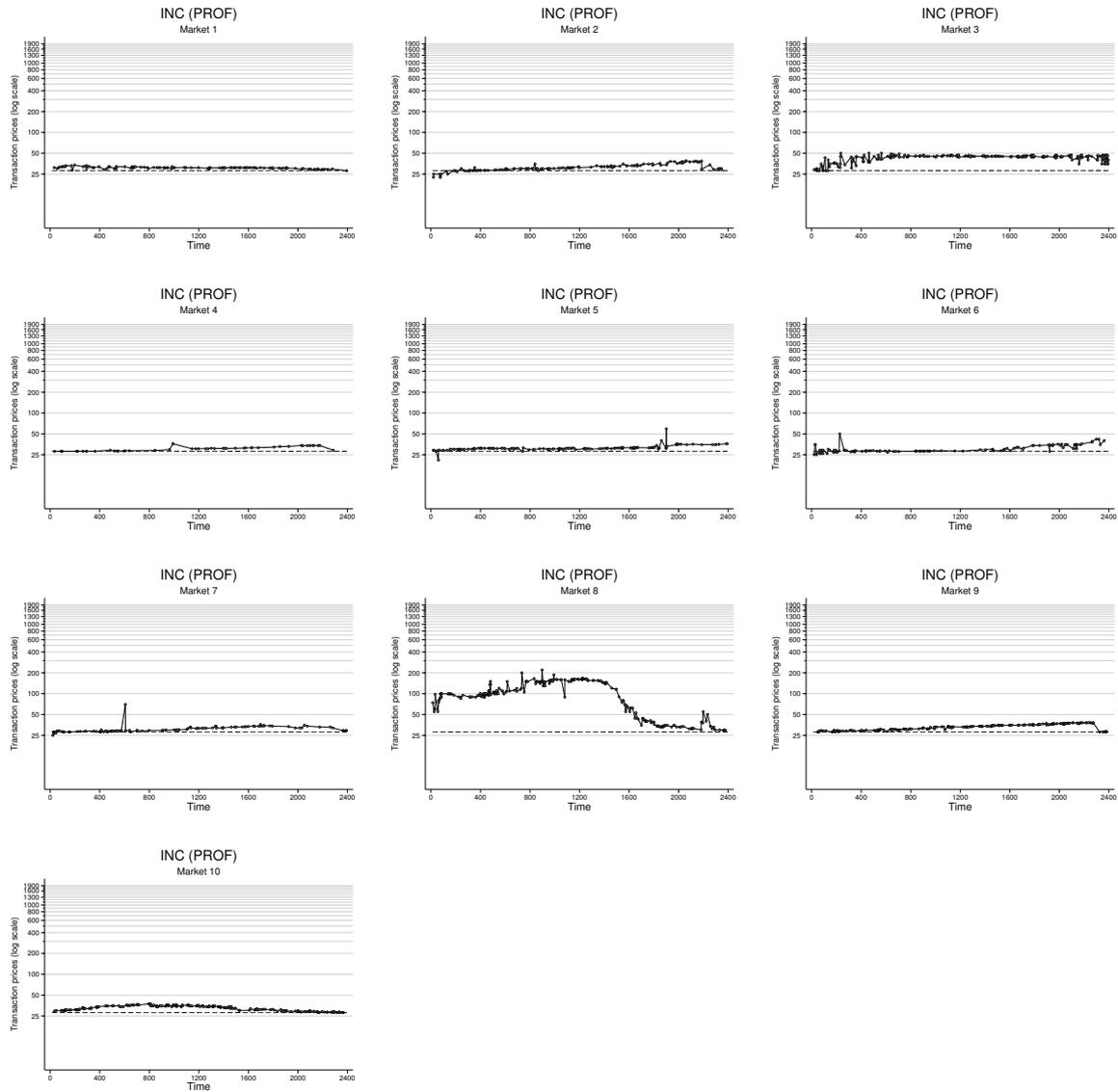


Figure B2: **Individual transaction prices for each market of Treatment INC in Experiment PROF**: The dashed line represents the risk-neutral fundamental value of 28. In Market 6, one trade with price  $< 5$  was dropped for presentation purposes.

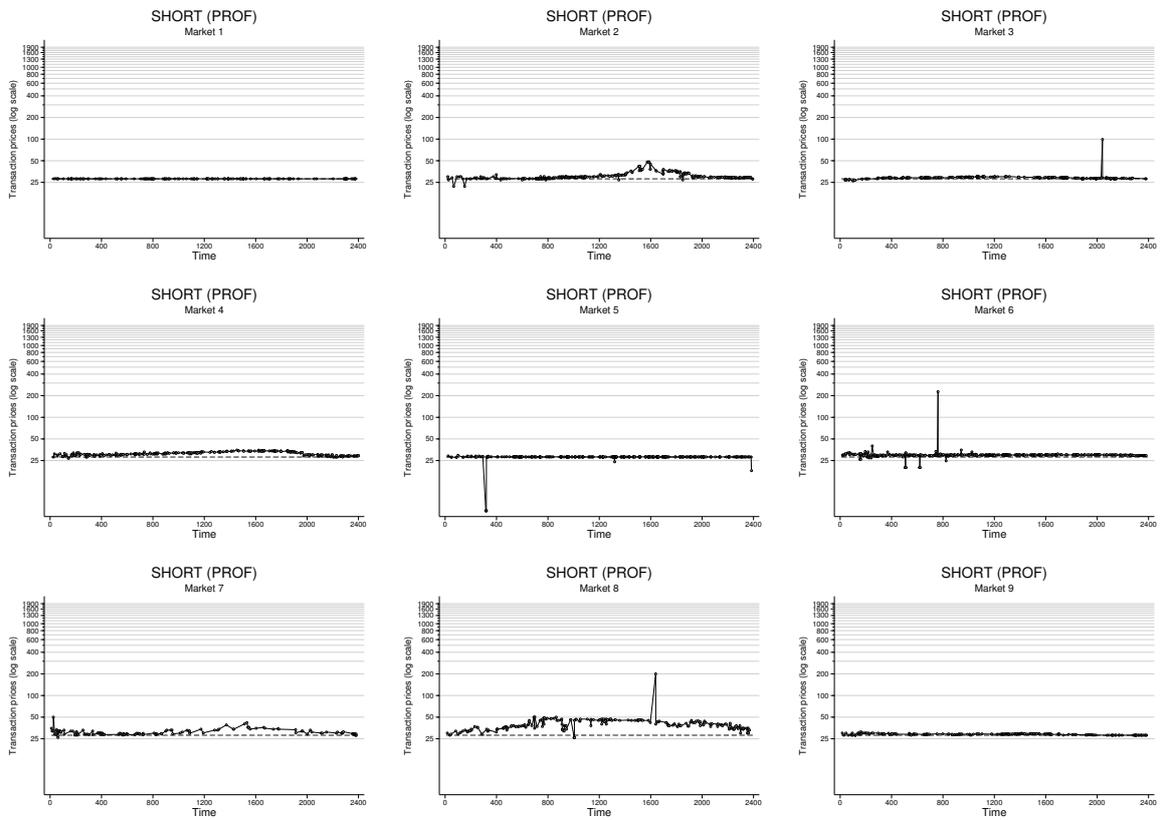


Figure B3: **Individual transaction prices for each market of Treatment SHORT in Experiment PROF**: The dashed line represents the risk-neutral fundamental value of 28.

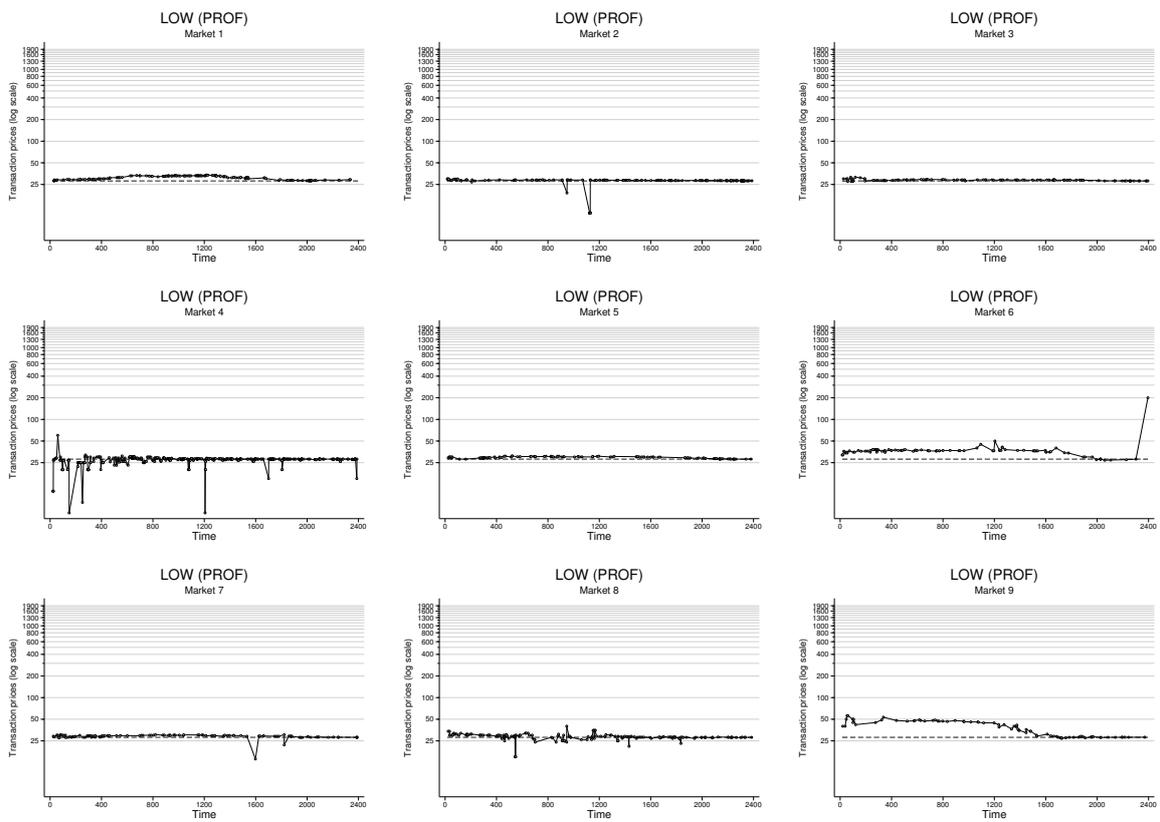


Figure B4: **Individual transaction prices for each market of Treatment LOW in Experiment PROF:** The dashed line represents the risk-neutral fundamental value of 28. In Markets 4, 7, and 8, ten, one, and seven trades, respectively, with prices  $< 5$  were dropped for presentation purposes.

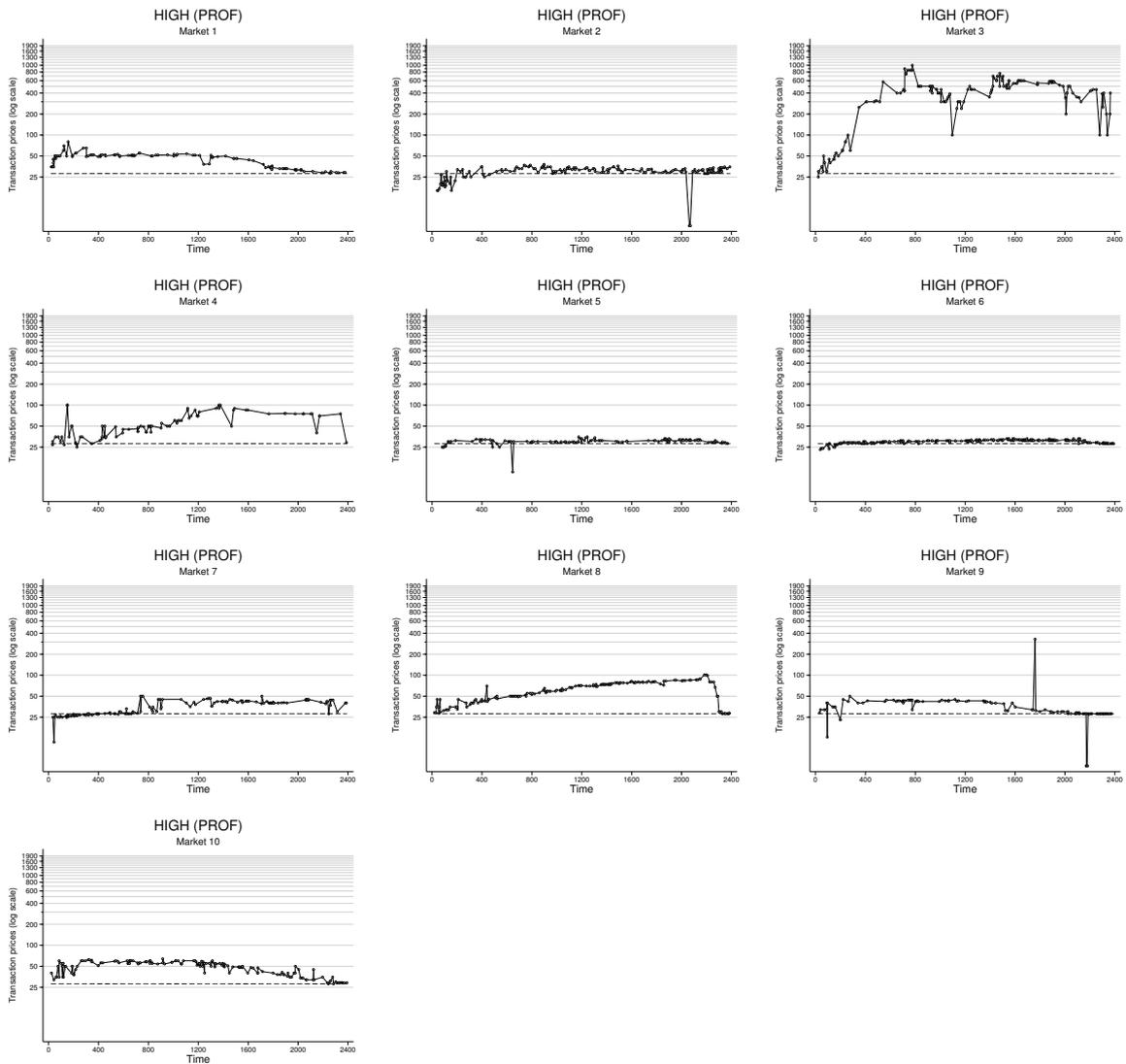


Figure B5: **Individual transaction prices for each market of Treatment HIGH in Experiment PROF**: The dashed line represents the risk-neutral fundamental value of 28. In Markets 1 and 4, four and one trades, respectively, with prices  $< 5$  were dropped for presentation purposes.

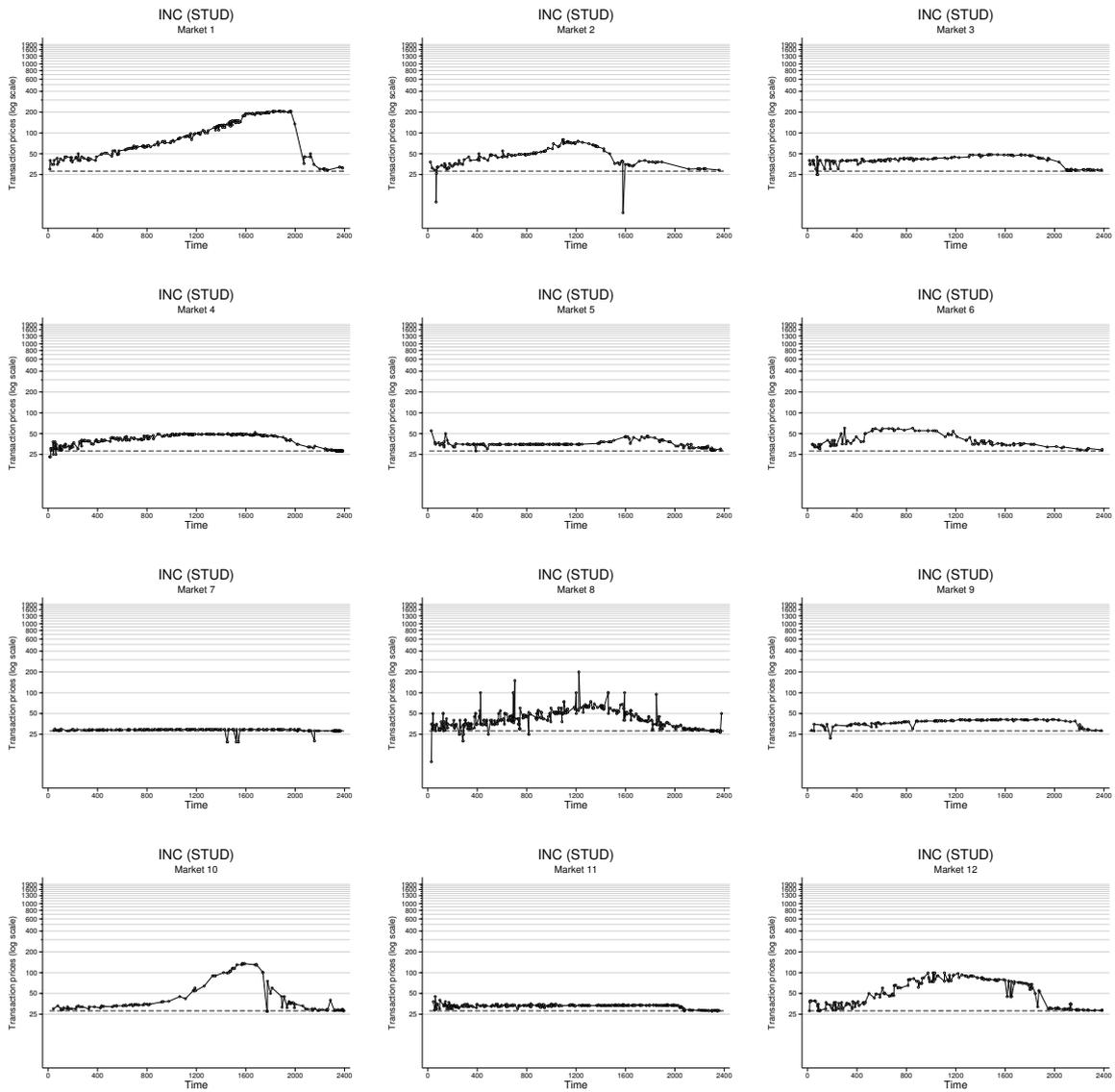


Figure B6: **Individual transaction prices for each market of Treatment INC in Experiment STUD:** The dashed line represents the risk-neutral fundamental value of 28. In Market 4, one trade with price  $< 5$  was dropped for presentation purposes.

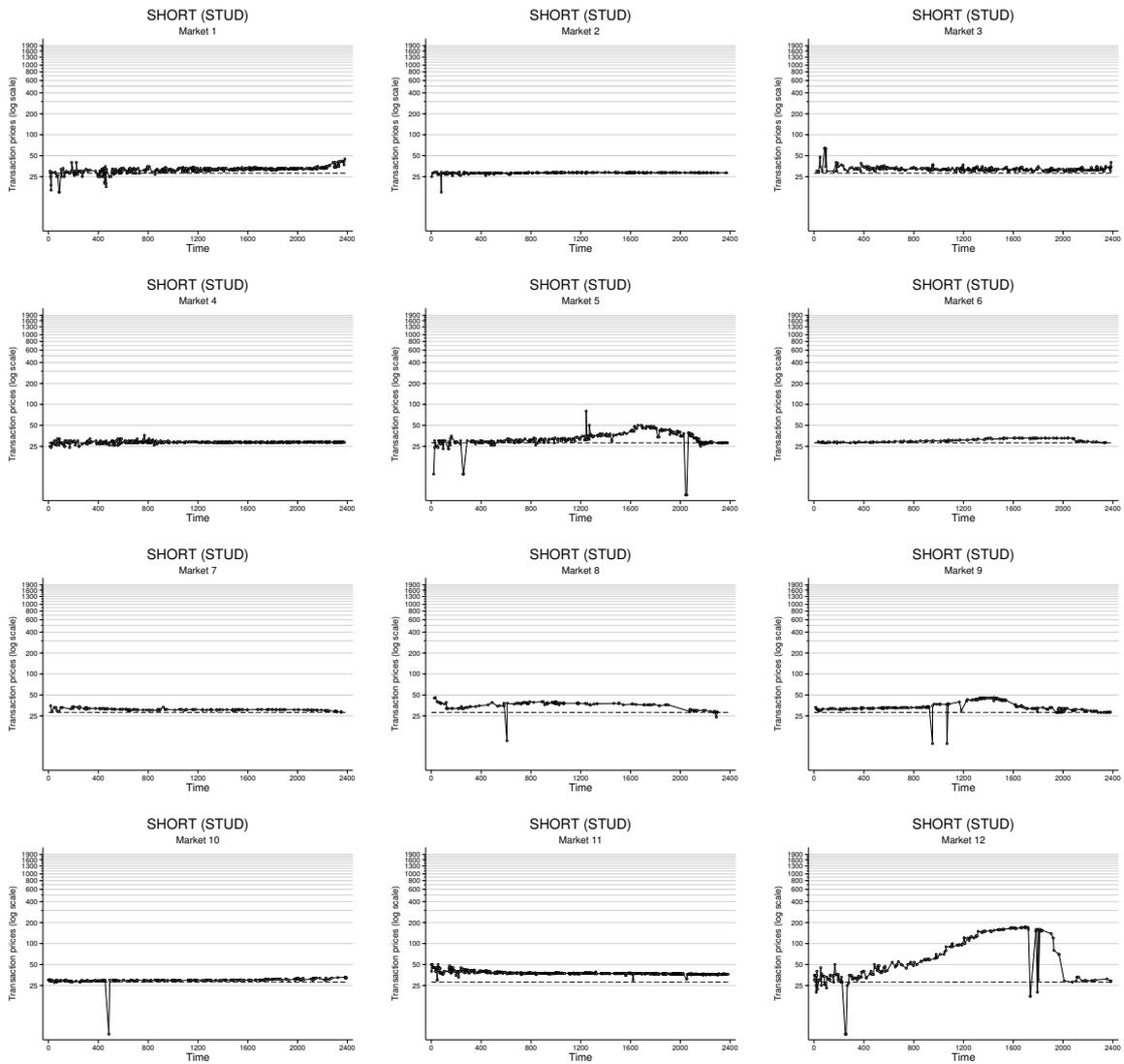


Figure B7: **Individual transaction prices for each market of Treatment SHORT in Experiment STUD:** The dashed line represents the risk-neutral fundamental value of 28. In Market 6, four trades with prices  $< 5$  were dropped for presentation purposes.

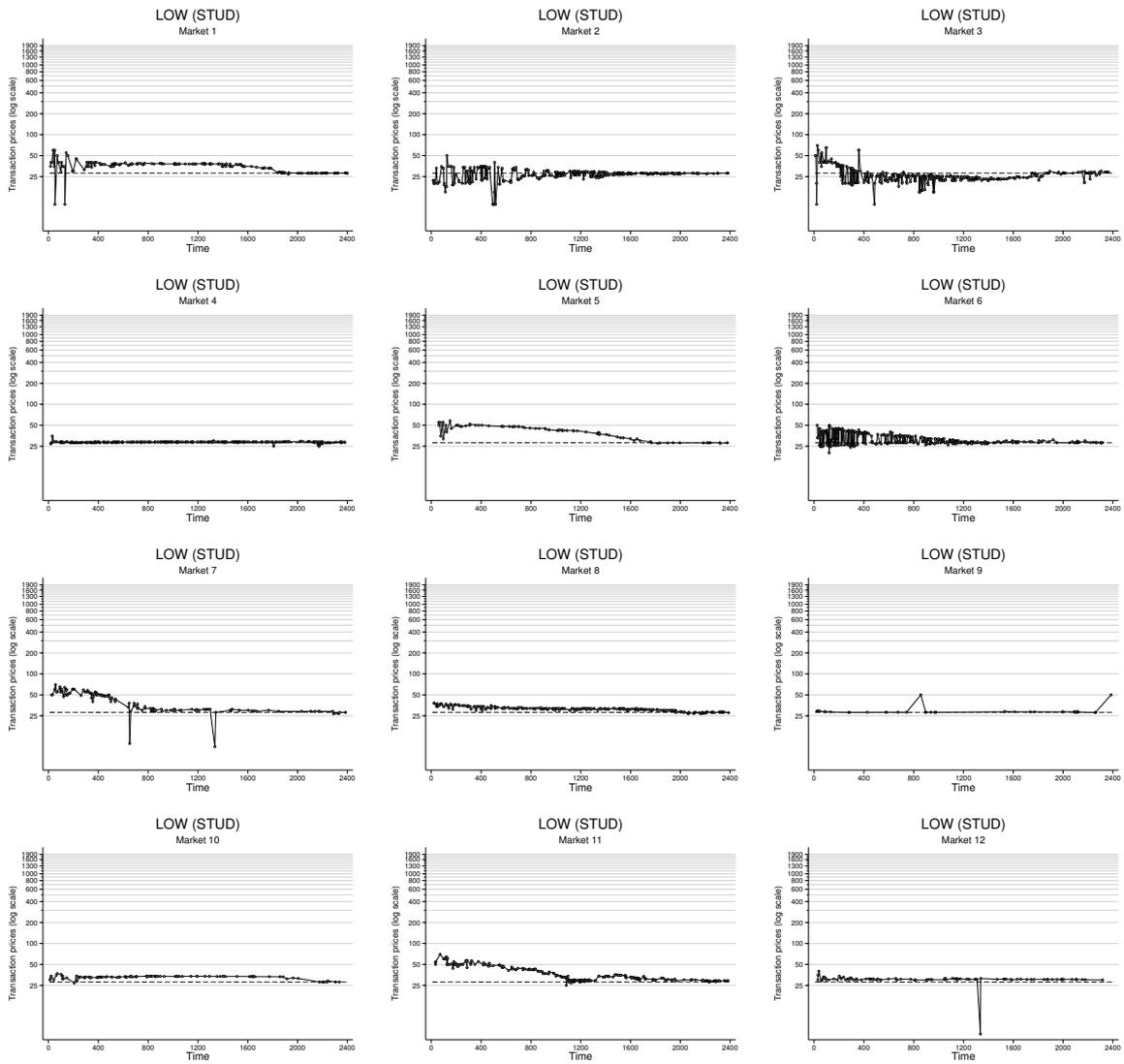


Figure B8: Individual transaction prices for each market of Treatment LOW in Experiment STUD: The dashed line represents the risk-neutral fundamental value of 28.

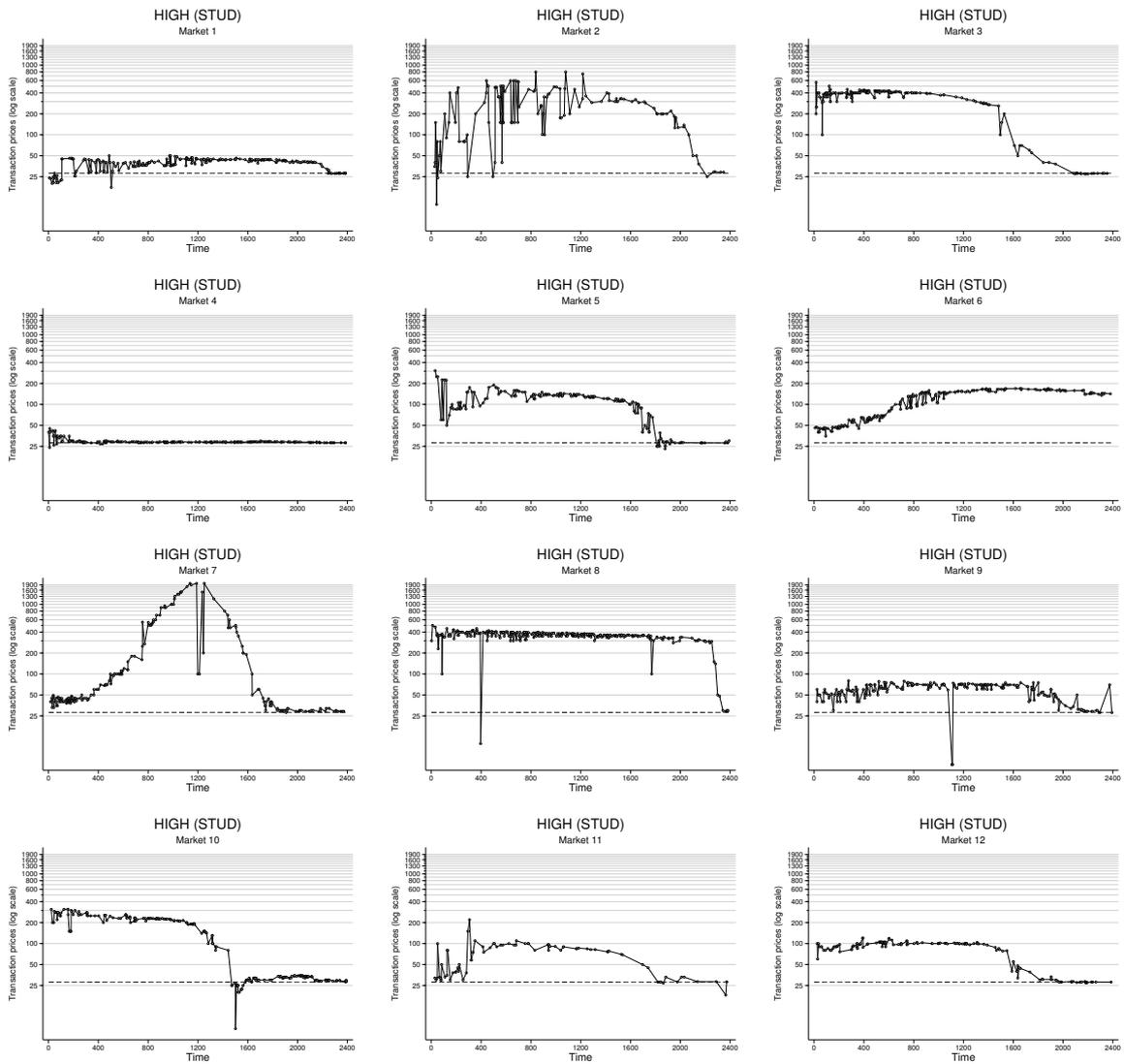


Figure B9: Individual transaction prices for each market of Treatment HIGH in Experiment STUD: The dashed line represents the risk-neutral fundamental value of 28. In Markets 7, 9, and 10, six, one, and three trades, respectively, with prices  $< 5$  or  $> 2000$  were dropped for presentation purposes.

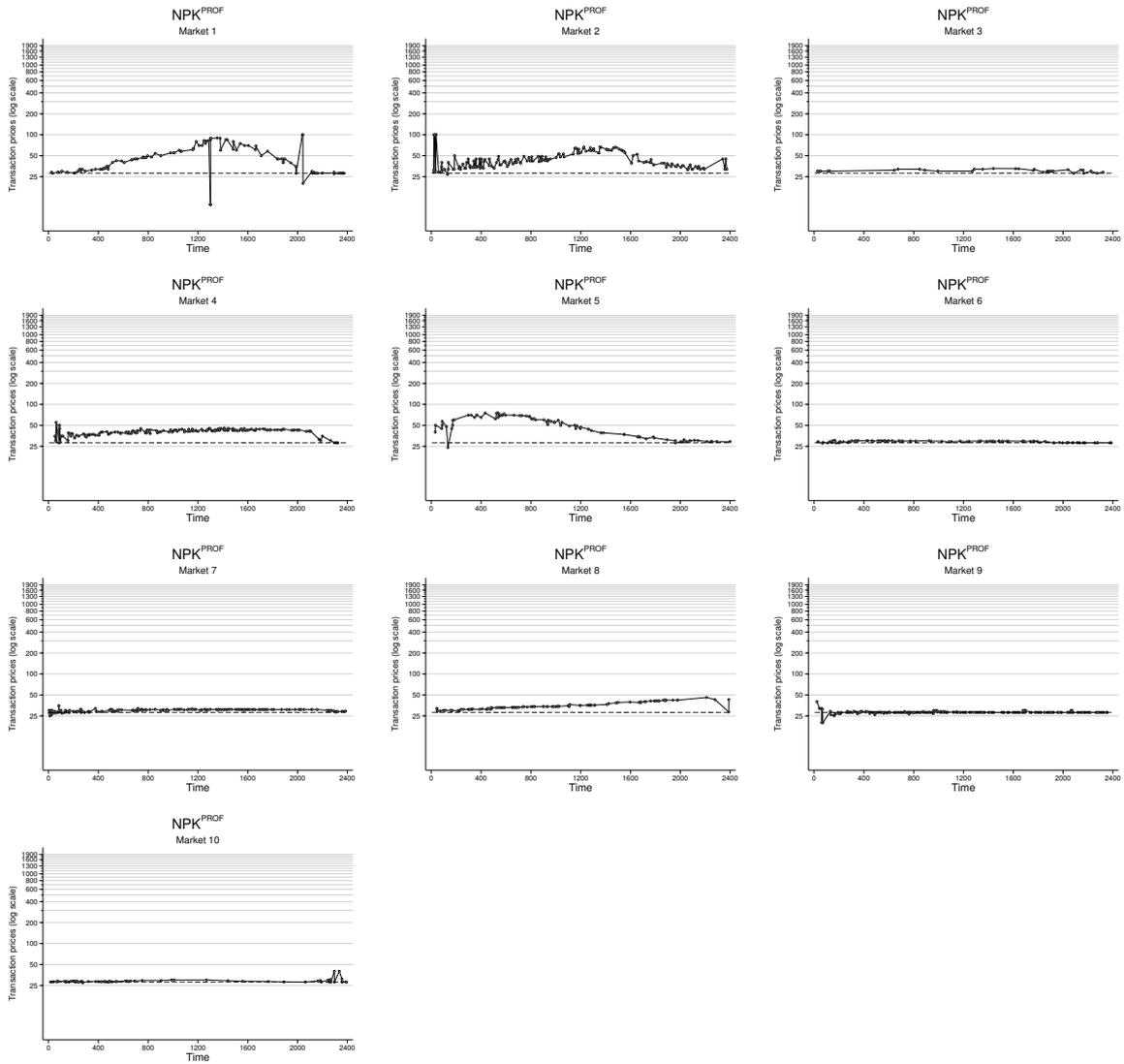


Figure B10: **Individual transaction prices for each market of Treatment  $NPK^{\text{PROF}}$** : The dashed line represents the risk-neutral fundamental value of 28. In Market 9, nine trades with prices  $< 5$  or  $> 2000$  were dropped for presentation purposes.

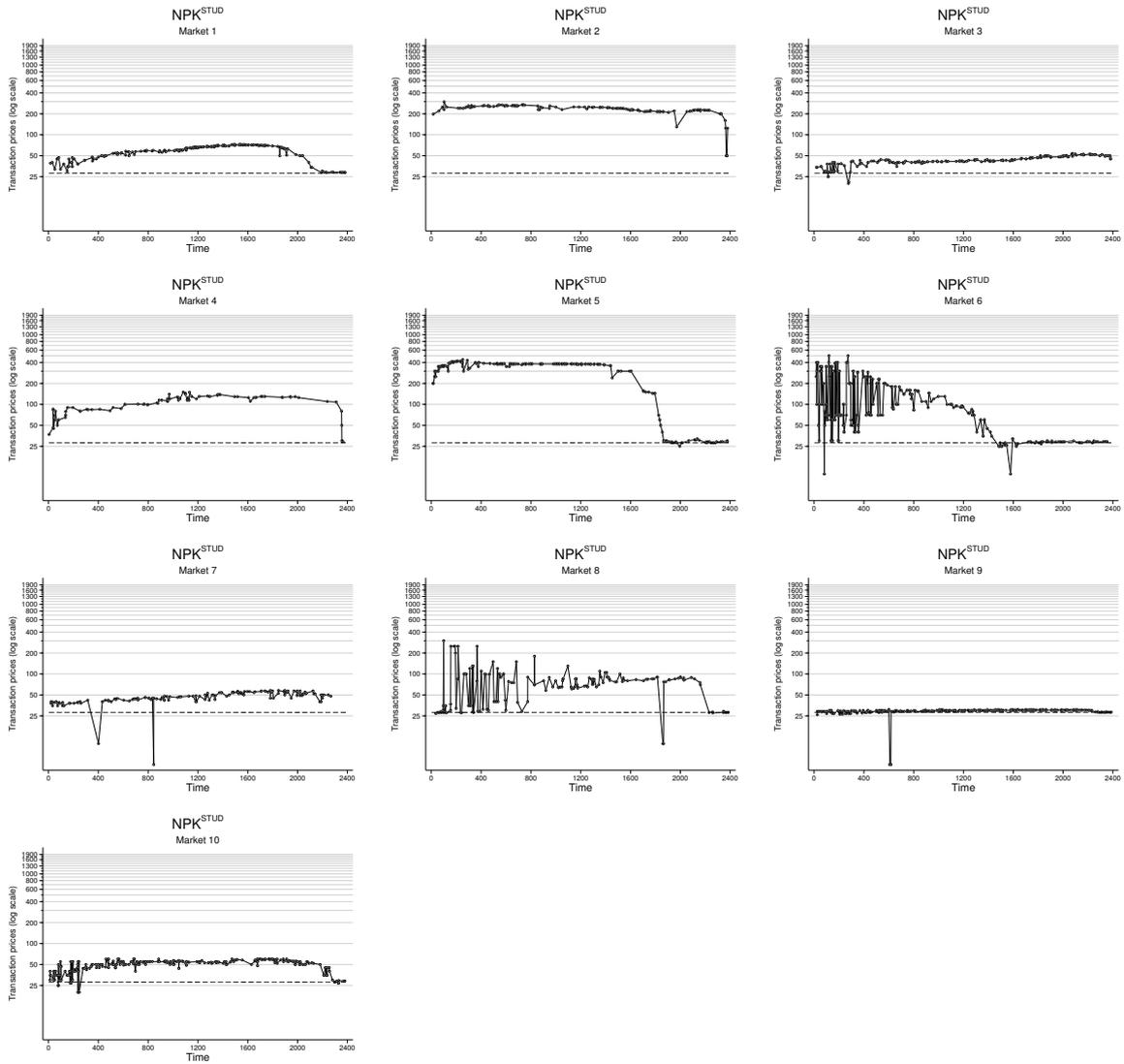


Figure B11: **Individual transaction prices for each market of Treatment  $NPK^{STUD}$** : The dashed line represents the risk-neutral fundamental value of 28. In Market 2, two trades with prices  $< 5$  or  $> 2000$  were dropped for presentation purposes.

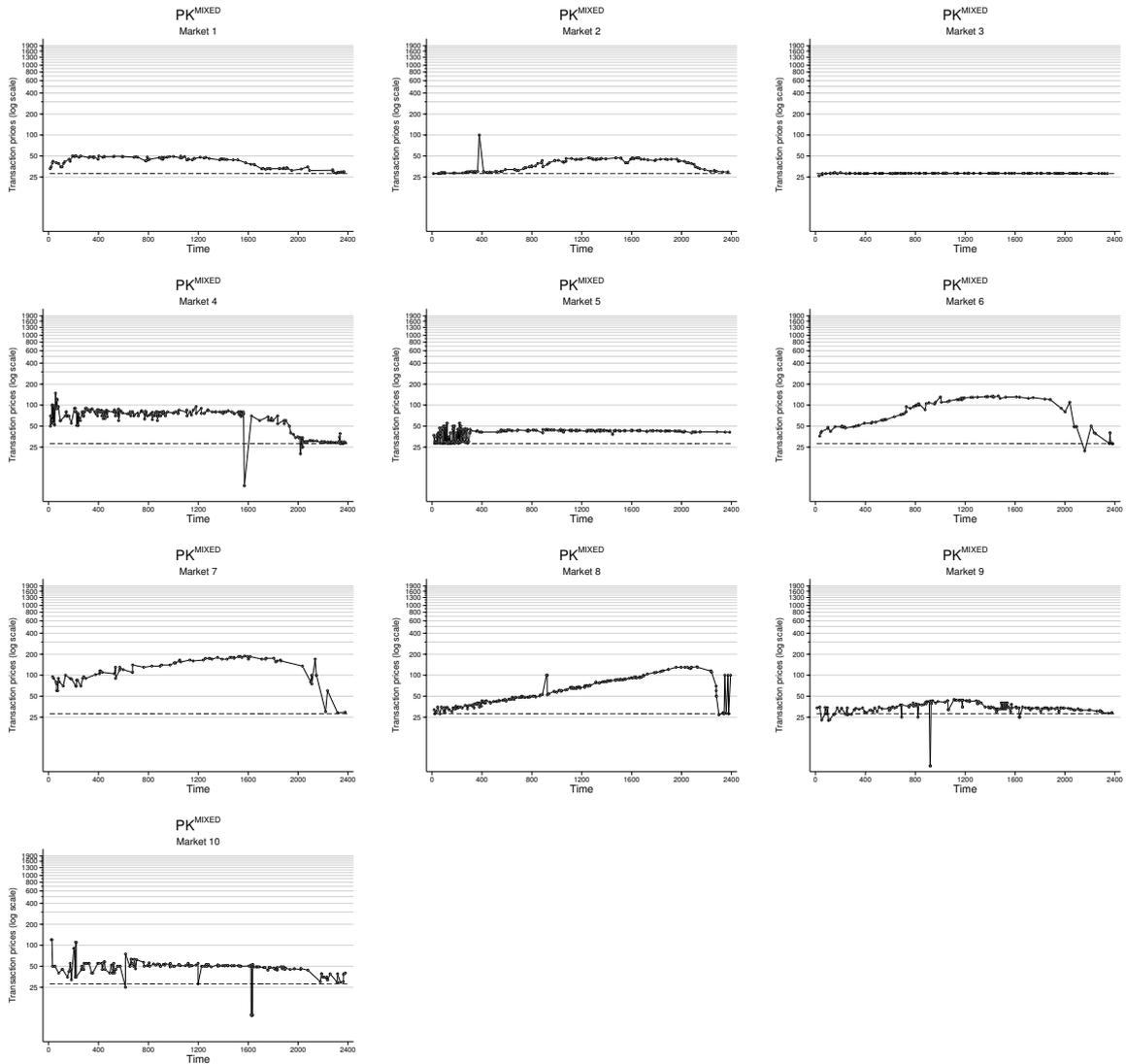


Figure B12: Individual transaction prices for each market of Treatment  $PK^{\text{MIXED}}$ : The dashed line represents the risk-neutral fundamental value of 28. In Market 8, one trade with a price  $< 5$  or  $> 2000$  was dropped for presentation purposes.

## C Price Beliefs

Figures C13 and C14 provide a descriptive overview of median treatment prices and median price forecasts for periods up to  $t + 2$ , elicited in  $t$ . Moreover, we investigate whether forecast accuracy of prices differs between professionals and students and between bubble-driver and bubble-moderator treatments. Here, we calculate the forecast error of price beliefs,  $FE_{t,t+k}^i$ , of subject  $i$  in period  $t$  for period  $t + k$  as follows:  $FE_{t,t+k}^i = \ln\left(\frac{\bar{P}_{t+k}}{\tilde{P}_{t,t+k}^i}\right)$ .  $\bar{P}_{t+k}$  stands for the mean period price in period  $t + k$  with  $k$  indicating values in the range  $\{0,1,2\}$  and  $\tilde{P}_{t,t+k}^i$  indicates subject  $i$ 's beliefs in period  $t$  of the mean market price in  $t + k$ . Hence, with  $FE_{t,t+k}^i$  we measure the percentage difference of future market prices in  $t + k$  and subject's price beliefs for  $t + k$ , elicited in  $t$ . Table C21 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. For Treatment PK<sup>MIXED</sup> we separate price beliefs of professionals and students.

As major result we find that differences in professionals' and students' price forecast errors are statistically insignificant in almost all treatments before the price peak. However, professionals predict prices significantly better than students in both bubble-driver treatments INC and HIGH and in Treatment NPK<sup>PROF</sup> after the price peak.

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. 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 14.7 percent across all forecasting periods ( $t$  to  $t + 2$ ). Students' corresponding underestimation is statistically not different from professionals and lies between 3.3 and 13.8 percent (see upper panel of Table C21).

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 C21). Importantly, these qualitative patterns hold for the additional treatments  $\text{NPK}^{\text{PROF}}$  and  $\text{NPK}^{\text{STUD}}$ , as professionals are again significantly better in predicting price downswings compared to students.

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, HIGH (see Table 3), and  $\text{NPK}^{\text{PROF}}$ , thus making forecasts easier for professionals.<sup>31</sup>

---

<sup>31</sup>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.

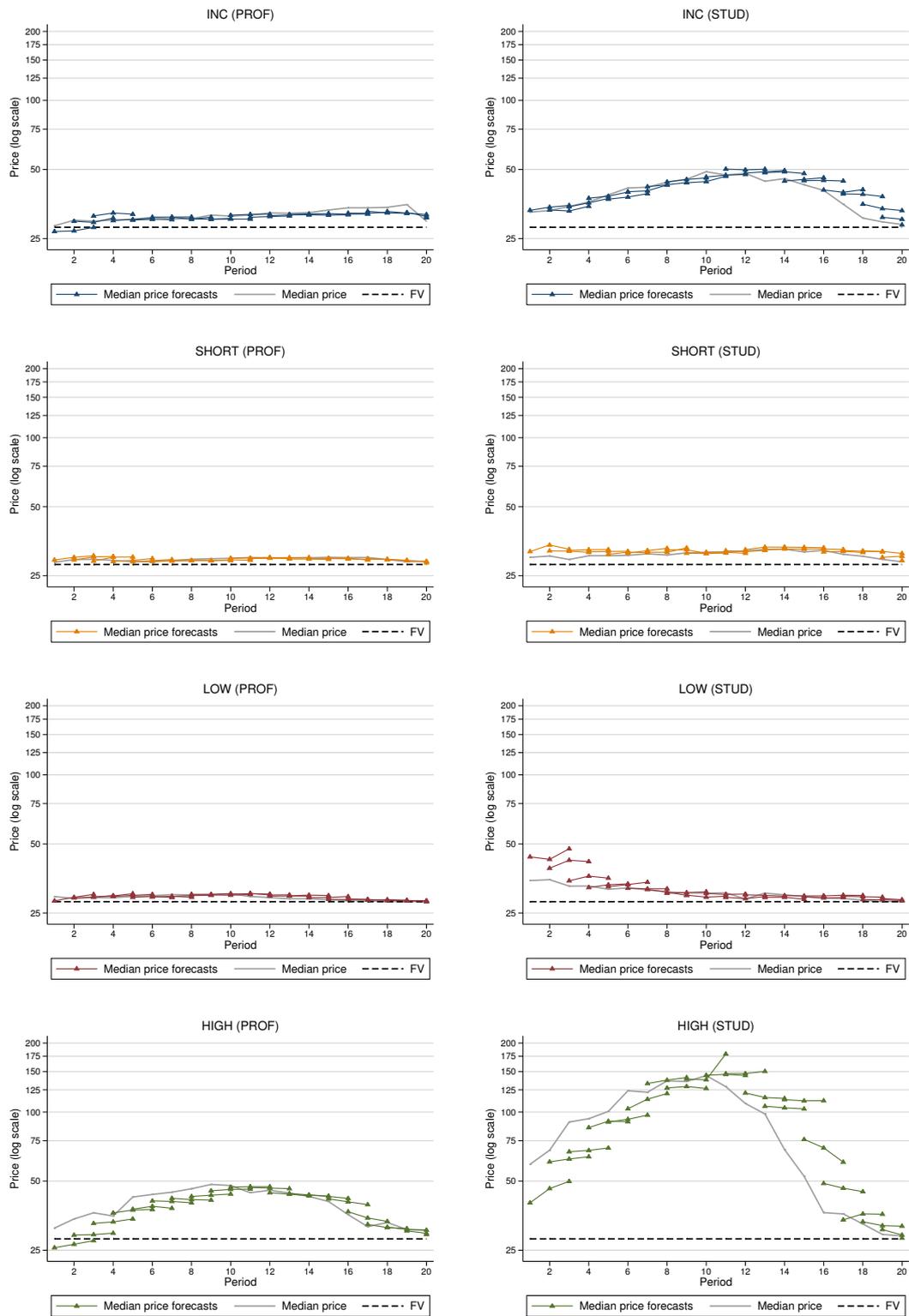


Figure C13: **Log-Price developments and subjects' median price forecasts across treatments in Experiment PROF (left column) and in Experiment STUD (right column):** This figure depicts median treatment prices (bold grey lines) and median price forecasts for three upcoming periods (colored lines with 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.

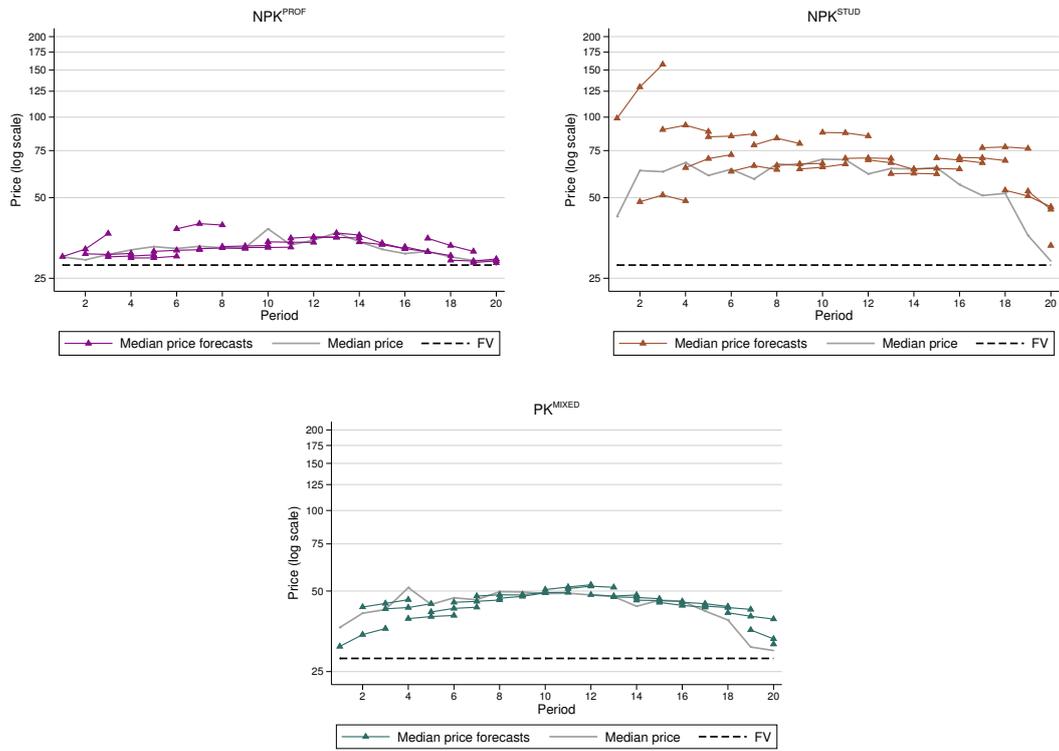


Figure C14: **Log-Price developments and subjects' median price forecasts across treatments in  $NPK^{PROF}$ ,  $NPK^{STUD}$ , and  $PK^{MIXED}$** : This figure depicts median treatment prices (bold grey lines) and median price forecasts for three upcoming periods (colored lines with triangles) as a function of period for treatments  $NPK^{PROF}$  (Treatment HIGH populated by professionals, but with non-public knowledge about trader composition of market),  $NPK^{STUD}$  (Treatment HIGH populated by students, but with non-public knowledge about trader composition of market), and  $PK^{MIXED}$  (Treatment HIGH with public knowledge about equal split of professionals/students in the market) in log-scale. The dashed lines represent the risk-neutral fundamental value of 28.

Table C21: **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), HIGH (high and constant CA-ratio of 10.2), NPK<sup>PROF</sup> (Treatment HIGH populated by professionals, but with non-public knowledge about trader composition of market), NPK<sup>STUD</sup> (Treatment HIGH populated by students, but with non-public knowledge about trader composition of market), and PK<sup>MIXED</sup> (Treatment HIGH with public knowledge about equal split of professionals/students in the market). 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 ( $FE_{t,t+k}$ ) 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 5% and 1% significance levels of a double-sided test.

<b>Before Price Peak (<math>t \leq t^*</math>)</b>									
Treatment	$t, t$			$t, t + 1$			$t, t + 2$		
	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)
NPK	1.57	3.34	(0.61)	0.77	4.81	(0.68)	0.48	7.81	(0.68)
PK <sup>MIXED</sup>	8.91	4.01	(-1.36)	11.08	3.99	(-1.13)	3.21	2.20	(-0.45)
<b>After Price Peak (<math>t &gt; t^*</math>)</b>									
Treatment	$t, t$			$t, t + 1$			$t, t + 2$		
	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)***
NPK	-0.25	-5.74	(2.50)**	-1.88	-9.17	(2.16)**	-2.08	-17.38	(1.25)
PK <sup>MIXED</sup>	-1.11	-1.74	(-1.13)	-1.87	-5.35	(-1.21)	-4.41	-6.87	(-1.13)

## D Instructions of the Base Experiment (Treatments INC, LOW, SHORT, and HIGH)<sup>32</sup>

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

---

<sup>32</sup>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.

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.*

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  $\pm 5\%$  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.

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

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

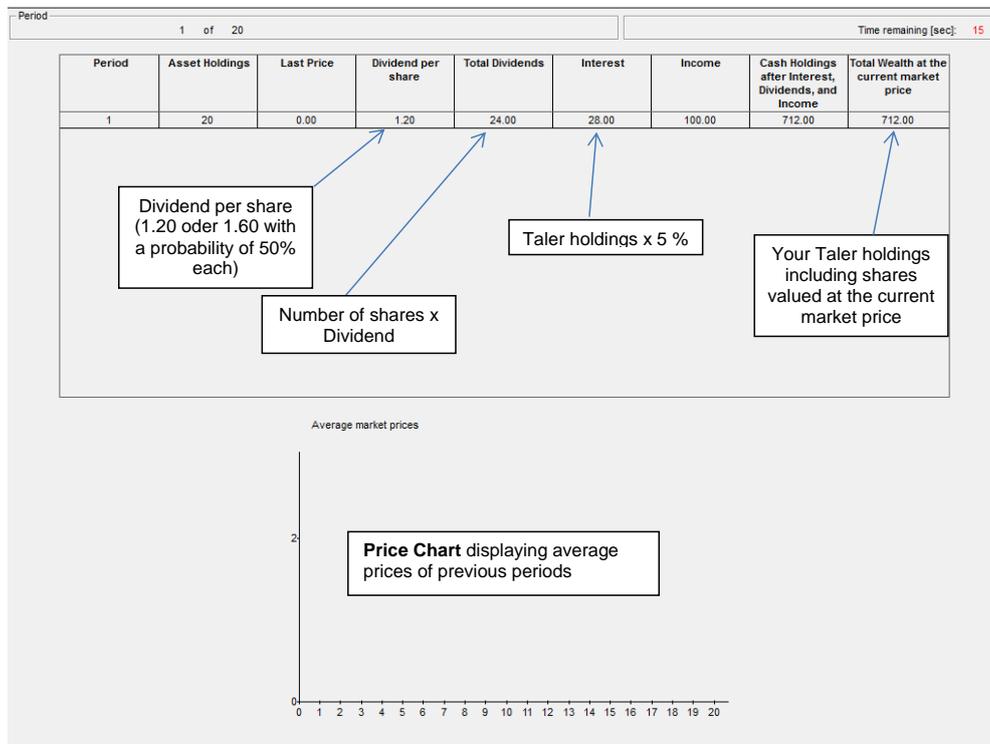
$$\text{Final Wealth in Euro} = \text{Final Wealth in Taler} / 100 \text{ (350)}$$

## Trading Screen

The screenshot shows a trading interface with several key components:

- Price chart of the current period:** A line graph on the left showing price over time (0 to 120 seconds).
- Liste of all orders to BUY from all traders:** A table in the top center showing buy orders from all participants, with the best offer highlighted in blue.
- List of all orders to SELL from all traders:** A table in the top right showing sell orders from all participants, with the best offer highlighted in blue.
- Order submission:** A panel on the right for entering and submitting orders, including fields for price and volume, and buttons for "Submit buy order" and "Submit sell order".
- Overview of your current holdings in asset and cash (Taler):** A panel in the middle right showing the trader's current position, including the last price (28.00).
- SELL:** A callout explaining that you sell the entered quantity at the price of the offer with the blue background.
- BUY:** A callout explaining that you buy the entered quantity at the price of the offer with the blue background.
- Submit BUY order:** A callout explaining that you must enter the offered price and quantity, and trading only occurs if another participant accepts the offer.
- Submit SELL order:** A callout explaining that it is analogous to "Submit BUY order".

# 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 above 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 Additional Treatments $\text{NPK}^{\text{PROF}}$ , $\text{NPK}^{\text{STUD}}$ , and $\text{PK}^{\text{MIXED}}$ ,<sup>33</sup>

————— Treatment  $\text{NPK}^{\text{PROF}}$  —————

### 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 and each of the other 7 traders receive an initial endowment of 20 shares and 5700 Taler each (experimental currency, converted to EUR 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.

Half of the traders in the room are financial professionals. The other half of the traders are students. These students were selected from a sample with the following characteristics: average age of 22 years; 75% male students; bachelor and master students from all fields with a focus on students of economics and management (fraction of 85%). You are matched with 7 other traders in the room to one market at the beginning of the experiment and, of course, this composition stays the same for all periods (i.e., in each period there are the same 8 traders in the market).

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

---

<sup>33</sup>Note that instructions for treatments  $\text{NPK}^{\text{PROF}}$ ,  $\text{NPK}^{\text{STUD}}$ , and  $\text{PK}^{\text{MIXED}}$  were identical to Treatment HIGH for the respective subject pool. The only difference was the explanation of the other subject pool in the second paragraph. Hence, we only report the first three paragraphs of the instructions, as the remainder was identical. Of course, original instructions of each treatment can be provided upon request.

### 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 and each of the other 7 traders receive an initial endowment of 20 shares and 5700 Taler each (experimental currency, converted to EUR 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.

Half of the traders in the room are students. The other half of the traders are financial professionals. These financial professionals were selected from a sample with the following characteristics: average age of 35 years; average of 10 years in industry; 90% male professionals; predominantly occupied as Private Banker, Trader, Investment Banker, Portfolio Manager, Fund Manager. You are matched with 7 other traders in the room to one market at the beginning of the experiment and, of course, this composition stays the same for all periods (i.e., in each period there are the same 8 traders in the market).

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

## 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 and each of the other 7 traders receive an initial endowment of 20 shares and 5700 Taler each (experimental currency, converted to EUR 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.

*Half of the traders in the market are financial professionals. The other half of the traders are students. These 4 students were selected from a sample with the following characteristics: average age of 22 years; 75% male students; bachelor and master students from all fields with a focus on students of economics and management (fraction of 85%). You are randomly matched with 3 other financial professionals and 4 students in the room to one market at the beginning of the experiment and, of course, this composition stays the same for all periods (i.e., in each period there are the same 8 traders in the market).*

Half of the traders in the market are students. The other half of the traders are financial professionals. These 4 financial professionals were selected from a sample with the following characteristics: average age of 35 years; average of 10 years in industry; 90% male professionals; predominantly occupied as Private Banker, Trader, Investment Banker, Portfolio Manager, Fund Manager. You are randomly matched with 3 other students and 4 financial professionals in the room at the beginning of the experiment and, of course, this composition stays the same for all periods (i.e., in each period there are the same 8 traders in the market).

---

<sup>34</sup>The following instructions are from Treatment PK<sup>MIXED</sup>. Text parts in the first three paragraphs of the instructions only applicable to the professional sample are in *italic* font, text parts only applicable to the students sample are in teletype.

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

## F Instructions of the Online Survey for Professionals and Students<sup>35</sup>

### Welcome<sup>36</sup>

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)

---

<sup>35</sup>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). The instructions were identical for both online surveys EXP<sup>COGRISK</sup> and ONLINE<sup>COGRISK</sup>.

<sup>36</sup>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 as possible without losing explanatory power. The order of the four tasks was randomized across all subjects.

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)

---

(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:*

### Your Decision

Task 2 of 2

---

Page 1 of 18

---



**playful**

**comforting**

**irritated**

**bored**

ⓘⓘⓘⓘ

## 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:*

### Your Decision

Time left to complete this page: ⌚ 0:32

Task 1 of 2

Page 1 of 18

The puzzle consists of a 3x3 grid of shapes. The bottom-right cell is empty. The shapes in the grid are:

Circle with a diagonal line from top-left to bottom-right.	Diamond with a diagonal line from top-left to bottom-right.	Square with a vertical dashed line.
Diamond with a vertical dashed line.	Square with a diagonal line from top-left to bottom-right.	Circle with a diagonal line from top-left to bottom-right.
Square with a diagonal line from top-left to bottom-right.	Circle with a vertical dashed line.	Empty cell with a tab on the right side.

Below the grid are eight possible answer options, each in a box with a tab on the right and a radio button below it:

- Option 1: Square with a vertical dashed line.
- Option 2: Diamond with a diagonal line from top-left to bottom-right.
- Option 3: Circle with a diagonal line from top-left to bottom-right.
- Option 4: Circle with a vertical dashed line.
- Option 5: Square with a diagonal line from top-left to bottom-right.
- Option 6: Circle with a diagonal line from top-left to bottom-right.
- Option 7: Square with a diagonal line from top-left to bottom-right.
- Option 8: Diamond with a vertical dashed line.

## **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:*

# Your decision

Task 2 of 2



Round 1 of 6



Sum to reach: 15

● Current Sum ● Remainder

The current sum is 3.

Please choose which number you want to add.

1  2  3

Next

## CRT7

Please answer the following seven questions.

Each question will be shown on a separate screen.

*Screenshot:*

# Survey

*Task 1 of 2*

*Page 1 of 7*

**Please answer the following question:**

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?

<input style="width: 90%;" type="text"/>	cents
--	-------

[Next](#)

### 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):

## G Pictures of the Experimental Laboratories

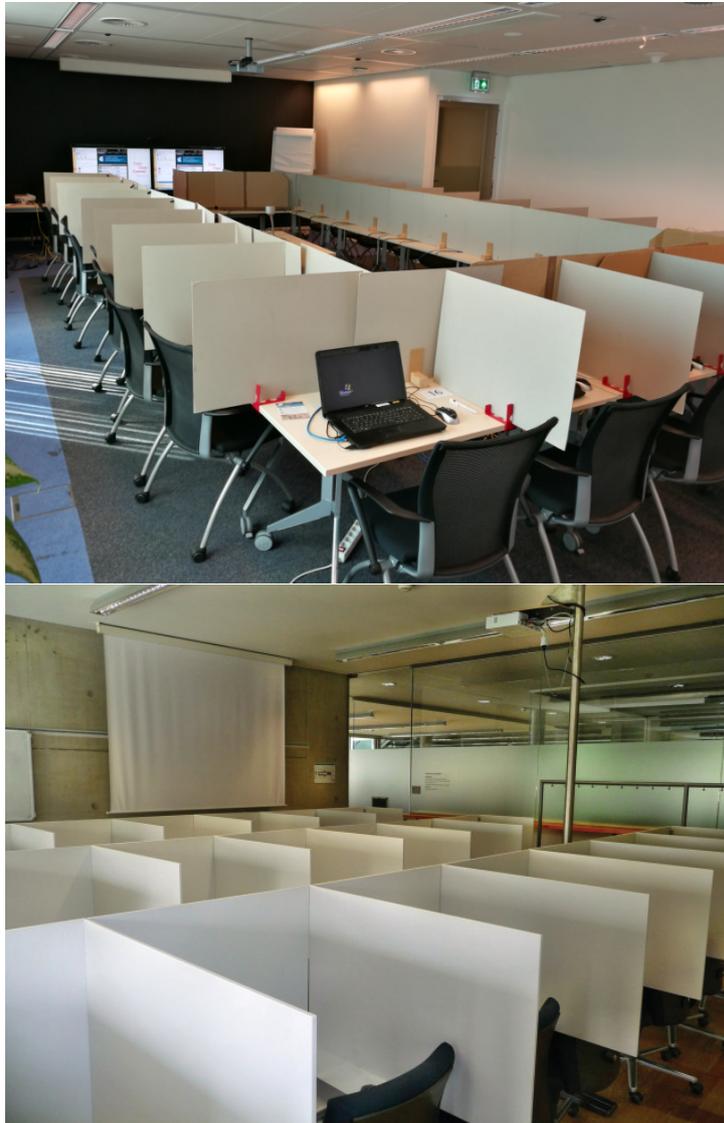


Figure G15: **Experimental Laboratories:** This figure shows one example of a mobile laboratory in the conference room of a financial institution (top) and the laboratory at Innsbruck EconLab (bottom).