

Bubbles and Financial Professionals

Utz Weitzel

Vrije Universiteit Amsterdam and Radboud University

Christoph Huber, Jürgen Huber and Michael Kirchler

University of Innsbruck

Florian Lindner

Max Planck Institute for Research on Collective Goods

Julia Rose

University of Innsbruck

The efficiency of financial markets and 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)

Received October 8, 2018; editorial decision May 27, 2019 by Editor Lauren Cohen. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

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, and Stefan Zeisberger; seminar participants at Chapman University, the Helsinki Center of Economic Research (HECER), and the Universities of Innsbruck and Tilburg; and 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.” Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Michael Kirchler, Department of Banking and Finance, University of Innsbruck, Universitätsstrasse 15, 6020 Innsbruck, Austria; telephone: +43 512 507 73014. E-mail: michael.kirchler@uibk.ac.at.

The Review of Financial Studies 33 (2020) 2659–2696

© The Authors 2019. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

For commercial re-use, please contact journals.permissions@oup.com

doi:10.1093/rfs/hhz093

Advance Access publication August 28, 2019

Financial bubbles and crashes have been recurring phenomena in economic history. Galbraith (1994), Kindleberger and Aliber (2011), Brunnermeier and Oehmke (2013), and Brunnermeier and Schnabel (2016) have shown that bubbles have been observed in different time periods (dating back to the 17th century), in economies at different stages of development (from developing economies to highly industrialized economies in the 21st century), and across a wide range of asset classes (e.g., real estate markets, asset markets, and derivative markets; Xiong and Yu 2011). Bubbles, crashes, and their underlying drivers are highly interesting to economists for at least two reasons. First, they represent periods of inefficient prices, that is, 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).¹

Bubbles and crashes are generally considered by many to be 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.² Despite considerable literature on mispricing in financial markets, empirical evidence remains elusive (see, e.g., Gürkaynak 2008; Greenwood, Shleifer, and You 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 simultaneously test an underlying equilibrium model that may be misspecified. One way to tackle this problem is to use experimental asset markets as test-beds, where fundamental values are 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,

¹ We 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), so 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, Huber, and Kirchler (2017), which we elaborate on in Section 2.2. Moreover, whenever we refer to other studies on this topic, as we do in this introduction, we follow the nomenclature of the original authors and refer to market inefficiencies as bubbles when they do.

² 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” (Cassidy 2010). 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” (Udland 2015).

financial professionals are found to be less prone to anchoring than students (Kaustia, Alho, and Puttonen 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, Haigh, and List 2007), to more accurately assess others, risk preferences (Roth and Voskort 2014), and to take more risk in competitive situations involving rankings (Kirchler, Lindner, and Weitzel 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.³ We readily acknowledge that there are several ways to balance internal and external validity.⁴ Moreover, each experiment, whether lab or field, represents a well-defined, controlled situation. Generalizations from experiments with financial professionals to behavior on real financial markets therefore require caution. However, given that the question to which degree markets are efficient is central to the academic and industry-wide debate, it is surprising that no large-scale experimental evidence on professionals' contribution to price efficiency in financial asset markets exists. We 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, Porzio, and Song (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, Porzio, and Song 2017).⁵

³ Smith, Suchanek, and Williams (1988) and King et al. (1993) provided the first anecdotal evidence of the behavior of business professionals. In Smith, Suchanek, and Williams (1988) one market was run with businessmen and businesswomen 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 over-the-counter (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.

⁴ 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).

⁵ The seminal framework of Smith, Suchanek, and Williams (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, Van Boening, and Wellford 2000; Noussair, Robin, and Ruffieux 2001; Kirchler, Huber, and Stöckl 2012; Huber and Kirchler 2012).

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), that is, a high initial level of the monetary supply (cash) relative to the asset value in the market (see Caginalp, Porter, and Smith 1998, 2001; Noussair and Tucker 2016). In another treatment we allowed capital inflows and thereby created an increasing CA ratio over time (Kirchler, Huber, and Stöckl 2012; Razen, Huber, and Kirchler 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). Galbraith (1994) and Kindleberger and Aliber (2011) document how the expansion phase of many historic bubbles was fueled by various forms of capital inflows. Similarly, Brunnermeier and Schnabel (2016) analyze twenty-three bubble episodes spanning the last 400 years and conclude that the emergence of bubbles is often preceded or accompanied by expansive monetary policy, high leverage of market participants, lending booms, and capital inflows. The two treatments we apply – high initial CA ratio and increasing CA ratio over time – capture these features of high or increasing capital supply in a simplified way.

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, Huber, and Kirchler 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.⁶

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 thirty-eight 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, Huber, and Kirchler (2017), we detect

⁶ Experimentally, in lab markets with student subjects, all four bubble drivers and bubble moderators have been shown to affect price efficiency. For example, Caginalp, Porter, and Smith (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, Huber, and Kirchler (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 even can become negative (Haruvy and Noussair 2006).

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, Porter, and Smith 1998; Allen and Gale 2000; Brunnermeier and Schnabel 2016). We also add to the emerging experimental literature analyzing behavior of financial professionals (e.g., Haigh and List 2005; Alevy et al. 2007; Cohn, Fehr, and Maréchal 2014, 2017; Kirchler, Lindner, and Weitzel 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, Lindqvist, and Moore 2005), gender (Eckel and Füllbrunn 2015), cognitive sophistication (Bosch-Rosa, Meissner, and Bosch-Domènech 2018), or speculative behavior (Janssen, Füllbrunn, and Weitzel 2018). The theory does not discriminate about 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, Suchanek, and Williams 1988; Lei, Noussair, and Plott 2001; Dufwenberg, Lindqvist, and Moore 2005; Kirchler, Huber, and Stöckl 2012; Sutter, Huber, and Kirchler 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 markedly contributed 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 accelerated their purchases even during the crash.⁷ 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 (that is, beliefs about the rationality of others; Cheung, Hedegaard, and Palan 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 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 1982; Corgnet, DeSantis, and Porter 2019). Translated to our setting, this indicates that professionals act as stabilizing device in markets with less experienced (nonprofessionals) 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

⁷ Cheng, Raina, and Xiong (2014), however, show that professionals might face difficulties in detecting bubbles. The authors focus on the bubble in the U.S. housing market from 2004 to 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.

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.

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, Porzio, and Song (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.⁸ 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, Porzio, and Song (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, that is,

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

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

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

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

⁸ See Online 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, Füllbrunn, and Weitzel 2018). A rich history of theoretical models in finance (e.g., Brunnermeier 2008; Scherbina and Schlusche 2014) shows 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, Lahav, and Noussair (2007), Kirchler et al. (2015), and Razen, Huber, and Kirchler (2017). Specifically, at the beginning of each period t , subjects are asked to predict average period prices for the three upcoming periods. $\hat{P}_{t,t+k}^i$ 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, Porzio, and Song (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.⁹ Because of these model characteristics put forward by Smith et al. (2014) and Holt, Porzio, and Song (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, Porzio, and Song 2017; Razen, Huber, and Kirchler 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).¹⁰ 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 B"). This account is not available for trading, but the holdings are added to final wealth and thus converted to euro at the end of the experiment. This procedure ensures the absence of capital inflow in the market with all other model features being identical to Treatment INC.

Finally, the second bubble driver treatment HIGH is identical to Treatment LOW, except for the level of the CA ratio. Here, we implement a constant CA ratio of 10.2 (i.e., the final level in treatments INC and SHORT) by 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, that is, a high initial monetary base in the market (e.g., Caginalp, Porter, and Smith 1998, 2001; Noussair and Tucker 2016). See Table A1 in Online Appendix A for an overview of the treatment parameterization.

1.3 Experiment implementation

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

⁹ 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.

¹⁰ 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).

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.¹¹ 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, that is, professionals from private banking, trading, investment banking, portfolio management, fund management, and wealth management. Experiments with professionals were run in the Netherlands, Austria, Germany, and Slovakia. In total, we conducted 15 experimental sessions. Moreover, we signed nondisclosure agreements (NDAs) for not disclosing the identity of the participating financial institutions. Most participating professionals were male (89.5%); 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 used in Kirchler, Lindner, and Weitzel (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 Online 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, Lindner, and Weitzel 2018), as well as some demographic questions (see Online 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, Lindner, and Weitzel (2018), who report hourly payments of 96 U.S. dollars (equivalent to 73 euro at the time of their experiment) and 72 euro,

¹¹ Moreover, it was very difficult to deny access to the experimental market when a ninth subject arrived unexpectedly (often despite prior cancellation) or when one market participant did not show up in time, endangering the participation of the seven other market participants.

respectively.¹² Subjects' payout comprises 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 forty-eight 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% male participants. The average age was 22.2 years and 87.3% 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, Haigh, and List 2007; Cohn, Fehr, and Maréchal 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, Baetge, and Nicklisch (2014) in Innsbruck and ORSEE by Greiner (2004) in Nijmegen.

2. Base Treatments Results

2.1 Price efficiency

Figure 1 outlines average volume-weighted period prices of individual markets and treatment medians and means in Experiment PROF (left panel) and Experiment STUD (right panel). The upper part of Table 1 and Table A10 in Online 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 Online Appendix A show corresponding numbers for

¹² 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% 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%) of their salary.

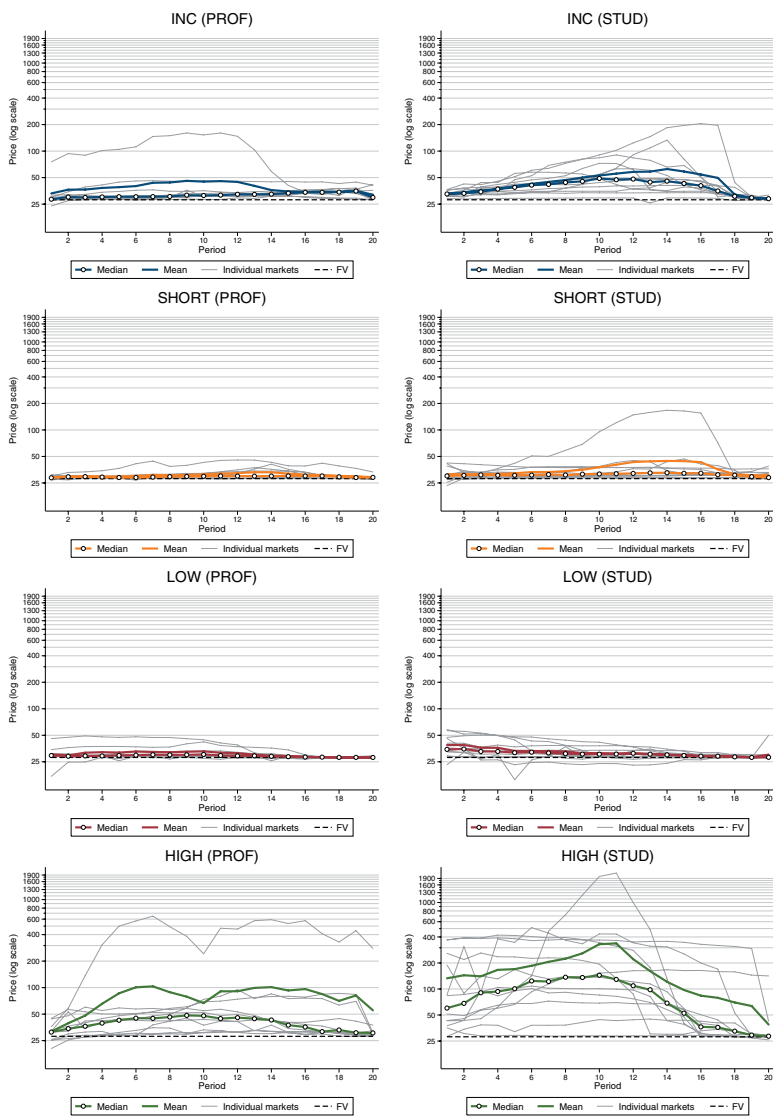


Figure 1
Log-price developments across treatments in Experiment PROF (left column) and in Experiment STUD (right column)

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 gray lines represent 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) as a percentage

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

Experiment STUD				
Variable (Median, %)	Treatment			
	INC	SHORT	LOW	HIGH
RAD (mispricing)	47.76	15.28	13.18	309.31
RD (overpricing)	47.76	13.73	10.49	309.27
RDMAX (max overpricing)	90.27	38.62	33.49	497.70
AMPLITUDE (price amplitude)	76.99	10.66	5.10	177.85
CRASH (price crash)	-87.75	-53.61	-66.41	-393.66
<i>N</i>	12	12	12	12

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

Experiment STUD. We follow Stöckl, Huber, and Kirchler (2010) and Razen, Huber, and Kirchler (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, Huber, and Kirchler 2010). Specifically,

$$RAD = \sum_{t=1}^T \frac{|\bar{P}_t - FV_t|}{FV_t} \text{ and } RD = \sum_{t=1}^T \frac{\bar{P}_t - FV_t}{FV_t}, \text{ with } \bar{P} \text{ being the average price in period } t \text{ and } T \text{ the total number of periods.}$$

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 prepeak 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, Huber, and Kirchler (2017).¹³ To test for significant pairwise differences between subject pools or treatments, we first compute the market mean of all period values for the variable of interest. Then we employ a Mann-Whitney U-test (MW U-test) with the market as unit of observation.¹⁴

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 Online Appendix A mispricing and overpricing in treatments SHORT and LOW are very small, with values below 7.0% 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%, respectively. Similarly, amplitude and crashes are large, particularly in Treatment HIGH with median AMPLITUDE of 70.8% and median CRASH equaling -128.6% (both as a percentage of the FV of 28). By running pairwise Mann-Whitney U-tests¹⁵ in Table 2 for RD (for both experiments) and in Table A10 in Online 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% 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.

Result 2: *In Experiment STUD, markets with bubble drivers capital inflow (INC) and high initial CA ratio (HIGH) exhibit significantly less efficient*

¹³ See also Section 2.2 for more explanations and Figures B1 to B12 in Online Appendix B for individual transaction price charts of each market.

¹⁴ 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.

¹⁵ Throughout the paper we report significance levels of statistical tests at 5% and lower. Because of 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.

Table 2
Pairwise Mann-Whitney U-tests of overpricing (RD) in experiments PROF and STUD

Treatment	RD (overpricing), treatment differences in percentage points					
	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)

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, that is, 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. Sample size *N* for each test is between 18 and 20 for PROF and 24 for STUD. ** $p < .05$; *** $p < .01$ (double-sided test).

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 Online Appendix A overpricing (RD) of 13.7% and 10.5% 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%, respectively. Moreover, amplitude and crashes are substantial, especially in Treatment HIGH with median AMPLITUDE of 177.9% and median CRASH of -393.66% (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.¹⁶ We find no differences between SHORT and LOW in any of the five outlined variables, hinting at very similar and in general rather efficient prices. However, we find significant differences between Treatment HIGH and all other treatments in all variables (with the only exception being Treatment INC for AMPLITUDE). Most of these differences are significant at the 1% 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.

¹⁶ Results are very similar in all other variables compared to RD, so we only outline RD in Table 2 for exemplary purposes.

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

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)

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 as a percentage and the numbers in parentheses show the Z-values of the MW U-test statistic. Sample size N for each test is either 21 or 22. ** $p < .05$; *** $p < .01$ (double-sided test).

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 nonbubble markets. As outlined in the introduction, there is no generally accepted bubble definition, also not in the experimental markets literature.¹⁷ 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

¹⁷ 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...” Nousseir, Robin, and Ruffieux (2001) follow this definition and quantify a bubble according to two criteria: (1) the median transaction price in five consecutive periods is at least 50 units of experimental currency (about 13.9%) greater than the fundamental value and (2) the average price is at least two standard deviations (of transaction prices) greater than the fundamental value for five consecutive periods.

bubble markets from nonbubble markets. We follow the approach of Razen, Huber, and Kirchler (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 95th percentile of the corresponding measure in the benchmark treatment (distribution) and therefore can be considered “significant” deviations from the benchmark. For our purpose we take all twenty-one markets of Treatment LOW as a benchmark, because no bubble drivers are 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.¹⁸

In particular, Razen, Huber, and Kirchler (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 $RDMAX_m > MAX\{0; \overline{RDMAX}^{LOW} + t(df)_{0.95} \cdot \sigma(RDMAX^{LOW})\}$, with \overline{RDMAX}^{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% quantile of a student t-distribution with $N - 1$ degrees of freedom (df) and N is the number of markets in the benchmark. $\sigma(RDMAX^{LOW})$ stands for the standard deviation of RDMAX in the N baseline markets. If $RDMAX_m$ 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 $RDMAX_m$ 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 $AMPLITUDE_m > \overline{AMPLITUDE}^{LOW} + t(df)_{0.95} \cdot \sigma(AMPLITUDE^{LOW})$, with $\overline{AMPLITUDE}^{LOW}$ and $\sigma(AMPLITUDE^{LOW})$ 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 $CRASH_m < \overline{CRASH}^{LOW} - t(df)_{0.95} \cdot$

¹⁸ Results of bubble identification remain identical if we only use the 12 (9) markets of the students (professionals) as a benchmark for all other student (professional) markets of the other treatments.

$\sigma(\text{CRASH}^{\text{LOW}})$, 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; that is, C1 the maximum period price (RDMAX), C2 the price run-up (AMPLITUDE), and C3 the CRASH are above the 95% quantile (below the 5% quantile for C3) of the corresponding measure in the benchmark distribution.

Result 4: *Following the bubble classification, 25% 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% 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 (gray 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 only 1 out of 24 markets have an identified bubble pattern, indicating that both treatments show consistent and nonbubble price patterns across both subject pools. In contrast, 10% and 40% 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% of the markets in treatments INC and HIGH, respectively, show bubble patterns (see Tables A3 to A6 in Online 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^{PROF}, NPK^{STUD}, and PK^{MIXED}) with ten 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^{PROF} and NPK^{STUD}, 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^{MIXED} 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

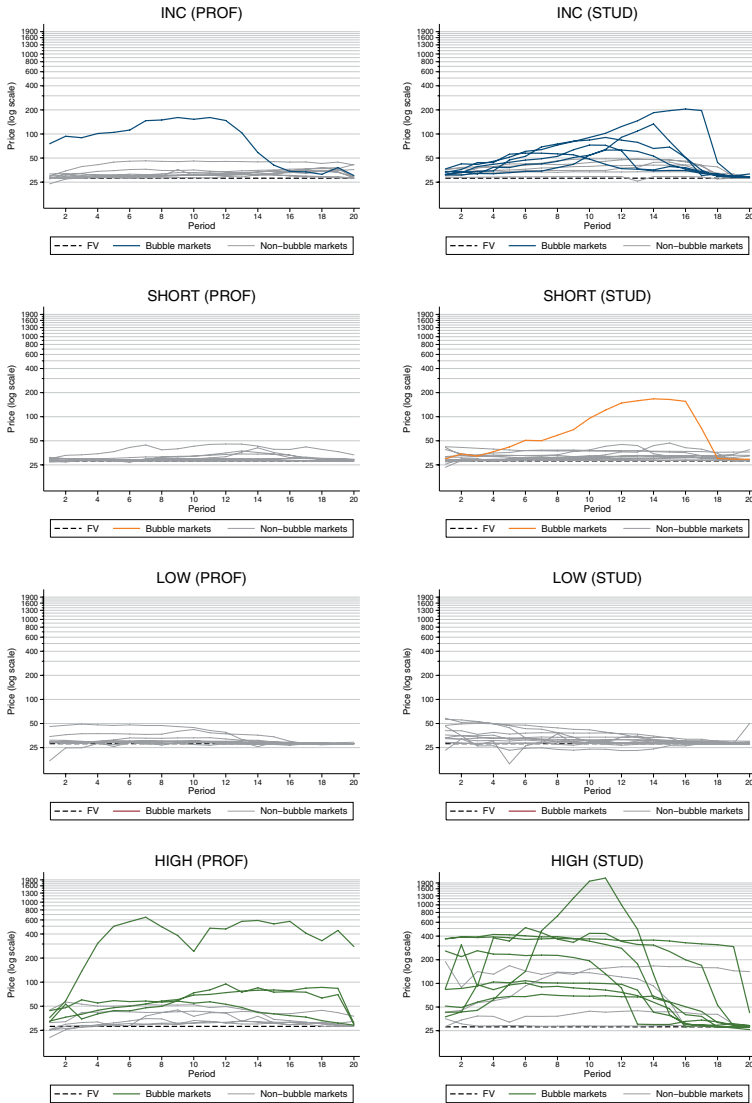


Figure 2
Bubble identification across treatments in Experiment PROF (left column) and in Experiment STUD (right column)

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 (gray 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.

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 NPK^{PROF} , NPK^{STUD} , and PK^{MIXED}

3.1.1 Experimental design and results. Following laboratory asset market experiments with students by Lei, Noussair, and Plott (2001), Cheung, Hedegaard, and Palan (2014), Akiyama, Hanaki, and Ishikawa (2017), and Bosch-Rosa, Meissner, and Bosch-Domènech (2018), beliefs about the “rationality” of other market participants can affect price efficiency.¹⁹ Professionals, knowing that they are trading with other professionals, may possibly trade more efficiently, because they have 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 (NPK^{PROF} and NPK^{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 eight subjects in the room constitute one market. Identical to the subjects in Cheung, Hedegaard, and Palan (2014), they were not told that professionals will only trade with professionals and students only with students (see the exact wording in the experimental instructions in Online 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 NPK^{PROF} and HIGH are set up identically except for subjects’ beliefs about other traders in the market.

In the third additional treatment, PK^{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 PK^{MIXED} with the respective nonmixed treatments of type HIGH allows us to measure the impact of heterogeneous skills on price efficiency.

In total, we conducted thirty 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. Most participating professionals were male (86.3%); 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

¹⁹ For instance, Lei, Noussair, and Plott (2001) show that mispricing in markets following the design of Smith, Suchanek, and Williams (1988), arises from uncertainty about the behavior of others, that is, that some subjects doubt the rationality of others. Moreover, Cheung, Hedegaard, and Palan (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 a distinct and significant effect of public knowledge over and above the effect of training alone.

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. We recruited 80.0% male students, the average age was 22.9 years and 95.8% were students at management and economics departments. Students received an average payout of 17.5 euro with a standard deviation of 3.9. Where possible we ran all three treatments within the same session to control for idiosyncratic session effects.²⁰

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^{PROF}) are much more efficient with fewer bubbles and significantly lower overpricing (RD) compared to student markets without public knowledge (NPK^{STUD}), but also compared to student markets with public knowledge (HIGH). Moreover, pairwise Mann-Whitney U-tests for RD show no differences between PK^{MIXED} , NPK^{PROF} , and the Treatment HIGH with professionals at the 5% level. In contrast, we do find significantly less overpricing in mixed markets with public knowledge (PK^{MIXED}) compared to markets populated solely by students with public knowledge (HIGH). Following related literature on insider trading (e.g., Plott and Sunder 1982; Corgnet et al. 2019), we postulate that a higher number of professionals in the market improves price efficiency. Jonckheere's trend tests show that overpricing indeed decreases with the number of professionals in the market, independent of public or nonpublic 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 Online Appendix A.

Turning to our bubble classification, 3 of the 10 markets in Treatment NPK^{PROF} , but a relatively high number 7 of 10 markets in the corresponding Treatment NPK^{STUD} with students exhibit bubbles. In Treatment PK^{MIXED} we

²⁰ Session sizes were 15, 2x16, 23, 24, 30, 32, and 2x40 subjects, and we randomized subjects into treatments.

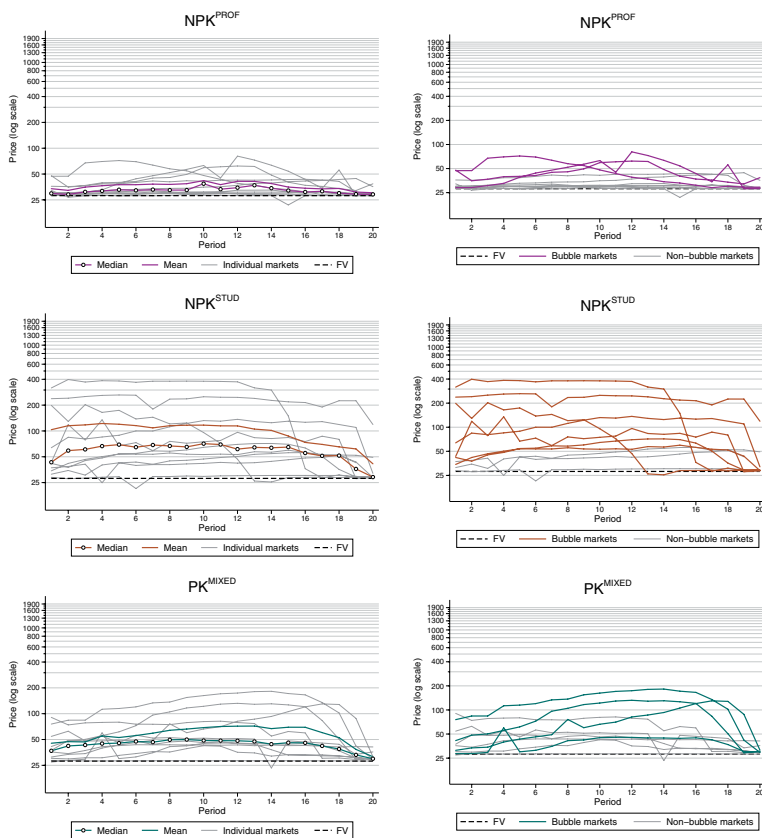


Figure 3
Log-price developments (left column) and bubble identification (right column) across the treatments
 Left column: this figure depicts median treatment prices (bold and colored lines with circles) and mean treatment prices (bold and colored lines) as a function of period for treatments NPK^{PROF}/NPK^{STUD} (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) in log-scale. The dashed lines represent the risk-neutral fundamental value of 28, and the gray lines represent 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 nonbubble markets (gray lines) as a function of period for the corresponding treatments in log-scale.

find an intermediate number of bubble markets (4 of 10 markets). See also Tables A7 to A9 in Online 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, regardless of their beliefs about the behavior of others. Moreover, these results

Table 4
Treatment medians of mispricing (RAD), overpricing (RD), maximum overpricing (RDMAX), price run-ups (AMPLITUDE), and crash (CRASH) in the additional treatments as a percentage

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

This table depicts median treatment values for treatments HIGH (high and constant CA ratio of 10.2), NPK^{PROF} (Treatment HIGH populated by professionals, but with nonpublic knowledge about trader composition of market), NPK^{STUD} (Treatment HIGH populated by students, but with nonpublic knowledge about trader composition of market), and PK^{MIXED} (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.

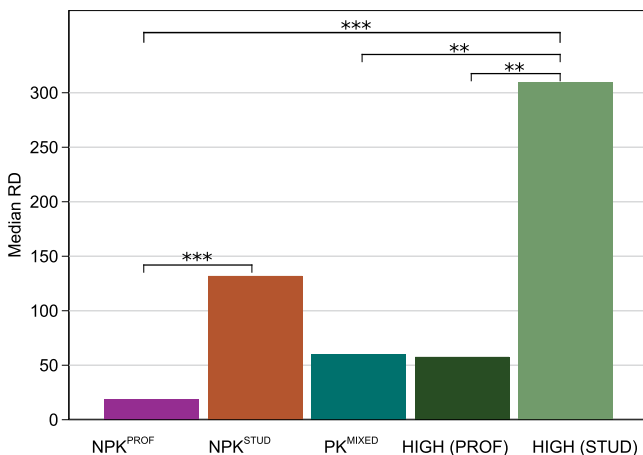


Figure 4
Pairwise Mann-Whitney U-tests of overpricing (RD) in the additional treatments

This figure shows pairwise treatment comparisons for treatments HIGH (high and constant CA ratio of 10.2), NPK^{PROF} (Treatment HIGH populated by professionals, but with nonpublic knowledge about trader composition of market), NPK^{STUD} (Treatment HIGH populated by students, but with nonpublic knowledge about trader composition of market), and PK^{MIXED} (Treatment HIGH with public knowledge about equal split of professionals/students in the market). ** $p < .05$; *** $p < .01$ (double-sided test).

also show that the presence of “arbitrageurs” (i.e., professionals) does not fully eliminate mispricing, indicating the fundamental importance of limits to arbitrage.

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 has suggested 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, Jappelli, and Padula 2010), earn higher Sharpe ratios (Grinblatt, Keloharju, and Linnainmaa 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, Keloharju, and Linnainmaa 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.²¹

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 trading behavior and price efficiency.²² Hence, both cognitive skills and risk attitudes, can drive behavior and performance on stock markets (e.g., Fellner and Maciejovsky 2007; Grinblatt, Keloharju, and Linnainmaa 2011, 2012; Kleinlercher, Huber, and Kirchler 2014; Hefti, Heinke, and Schneider 2016; Corgnet, DeSantis, and Porter 2018 and explain differences in price efficiency (Corgnet, DeSantis, and Porter 2015; Bosch-Rosa, Meissner, and Bosch-Domènech 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 $EXP^{COGRISK}$ to all participants of treatments NPK^{PROF} , NPK^{STUD} , and PK^{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

²¹ 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 one's ability to engage in effortful reasoning (e.g., Oechssler, Roeder, and Schmitz 2009; Toplak, West, and Stanovich 2011, 2014). 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, Quartz, and Bossaerts 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, Tucker, and Xu 2014; Corgnet et al. 2015); ToM correlates with subjects' skills in predicting price changes (Bruguier, Quartz, and Bossaerts 2010); and all three concepts are joint predictors of trader performance (Corgnet, DeSantis, and Porter 2018). However, DeMartino et al. (2013) show that ToM skills also can be detrimental when trading on financial markets. In a study using functional magnetic resonance imaging (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, DeSantis, and Porter (2015) and Bosch-Rosa, Meissner, and Bosch-Domènech (2018) find a causal relationship between traders' cognitive sophistication and price efficiency.

²² 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, Stráznická, and Villevall (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, Huber, and Kirchler (2014) show that even less risk-averse subjects invest more in the risky asset, resulting in asset overvaluation.

of treatments NPK^{PROF} , NPK^{STUD} , and PK^{MIXED} , 101 professionals and 106 students participated (out of 118 each).

Although we administered the survey $EXP^{COGRISK}$ 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 used in the market experiment. We refer to this additional online survey as $ONLINE^{COGRISK}$. Professionals were employed in the same areas as the ones from the market experiments and they shared the same characteristics. Most participating professionals were male (84.3%); 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% were male, and the average age was 23.8 years.

In contrast to $EXP^{COGRISK}$, the data from $ONLINE^{COGRISK}$ cannot be merged with the experimental trading data. In our analyses we therefore focus on $EXP^{COGRISK}$ and report $ONLINE^{COGRISK}$ as a robustness check. Both surveys, $EXP^{COGRISK}$ and $ONLINE^{COGRISK}$, were programmed and conducted with oTree (Chen, Schonger, and Wickens 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, DeSantis, and Porter (2018), that is, eighteen of the Raven's advanced progressive matrices (Raven 2000). For cognitive reflection skills, we used the extended cognitive reflection test (CRT) from Toplak, West, and Stanovich (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).²³ ToM defines one's capacity to infer others' intentions. To measure ToM-skills, we administered eighteen pictures of the eye-gaze test from Baron-Cohen et al. (2001). In this test, participants looked 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.²⁴ Online Appendix F provides details on the tests. In addition, we administered a HIT15 test (Burks et al. 2009) analyzing individuals' backward induction abilities, which are important in finite horizon

²³ 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" on reflection (5 cents is correct), which requires effortful system 2 processes.

²⁴ For the Raven's and eye-gaze test part of the survey, we used a shortened version. The original tasks comprise thirty-six questions, 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.

markets.²⁵ 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).²⁶

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 Online Appendix A outline the results of both surveys on all cognitive tasks.²⁷ 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^{COGRISK}, but the difference is rather small with 5.3 correct answers (of 7) for professionals and 4.7 for students. Moreover, the difference is not robust when compared to ONLINE^{COGRISK}, where we find no statistically significant difference for any of the four cognitive skills, including CRT.²⁸ 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, Meissner, and Bosch-Domènech 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 Corngnet, DeSantis, and Porter (2018) we do not find that cognitive skills, elicited

²⁵ 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.

²⁶ 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, Lindner, and Weitzel (2018). Dohmen et al. (2011) find that the self-reported SOEP measure can represent a valid substitute for incentivized lottery schemes and that it performs reasonably well in predicting risk-taking behavior of individuals. Crosetto and Filippin (2013) report that the single-item SOEP measure is highly and significantly correlated with the Domain-Specific Risk-Taking Scale (DOSPRT), 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 used for the general SOEP question.

²⁷ A detailed overview of the distribution of scores for professionals and student subjects can be found in Figure A1 in the Online Appendix for both EXP^{COGRISK} and ONLINE^{COGRISK} separately.

²⁸ 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, Tucker, and Xu (2016), where answers are given a weight according to whether they are (a) correct, (b) wrong, but correspond to the intuitive answer, and (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 of 7 answers are correct), *impulsive subjects* (at least 5 of 7 answers correspond to the intuitive answer), and *other* for all other combinations of answers. Results are available on request.

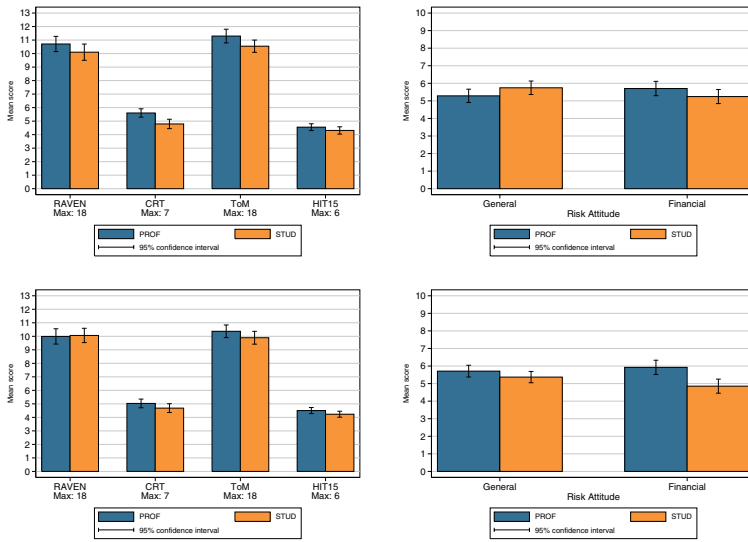


Figure 5 Differences in cognitive skills (left panel) and risk attitudes (right panel) between professionals (PROF) and students (STUD) in surveys EXP^{COGRISK} (top panel) and ONLINE^{COGRISK} (bottom panel). The left panel of this figure depicts mean test scores of fluid intelligence (RAVEN: Raven's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

in Survey EXP^{COGRISK}, explain trader performance in the three additional treatments (see Table A14 in Online Appendix A). This finding 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. Potential differences from Corngnet, DeSantis, and Porter (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 Online 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.

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 ordinary least squares (OLS) regressions, which are 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.

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.

Table 5
Heterogeneous beliefs and overvaluation

Dep. variable: RD _{<i>t</i>}	STUD				PROF							
	t,t	t,t+1	t,t+2	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($\tilde{NP}_{t,t+k}$)	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($\tilde{NP}_{t,t+k}$)	0.349 (0.753)	0.349 (0.753)	1.468** (0.753)	1.468** (0.753)	2.138*** (0.676)	2.138*** (0.676)	0.308 (0.751)	0.308 (0.624)	1.074 (0.700)	1.074 (0.700)	1.278 (0.700)	1.278 (0.700)
HIGH \times SD($\tilde{NP}_{t,t+k}$)	8.908 (4.995)	8.908 (4.995)	9.752** (4.800)	9.752** (4.800)	10.54 (6.959)	10.54 (6.959)	10.77*** (6.907)	10.77*** (6.907)	11.54*** (8.02)	11.54*** (8.02)	11.44*** (8.02)	11.44*** (8.02)
SHORT \times SD($\tilde{NP}_{t,t+k}$)	2.002** (0.868)	2.002** (0.868)	1.604*** (0.418)	1.604*** (0.418)	2.226** (0.615)	2.226** (0.615)	2.226** (0.952)	2.226** (0.952)	2.419*** (0.726)	2.419*** (0.726)	2.419*** (0.726)	2.419*** (0.726)
NPK ^{STUD} \times SD($\tilde{NP}_{t,t+k}$)	2.831*** (0.470)	2.831*** (0.470)	3.203*** (0.522)	3.203*** (0.522)	3.627*** (0.599)	3.627*** (0.599)						
NPK ^{PROF} \times SD($\tilde{NP}_{t,t+k}$)							2.738** (1.224)	2.738** (1.224)	2.575 (1.302)	2.575 (1.302)	2.575 (1.302)	2.575 (1.302)
PK ^{MIXED} \times SD($\tilde{NP}_{t,t+k}$)	1.241 (1.530)	1.241 (1.530)	1.931 (1.133)	1.931 (1.133)	2.306*** (0.684)	2.306*** (0.684)	1.184 (1.563)	1.184 (1.563)	2.187 (1.137)	2.187 (1.137)	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
<i>N</i>	1,268	1,268	1,202	1,202	1,138	1,138	1,075	1,075	1,017	1,017	960	960
<i>R</i> ²	.199	.217	.212	.234	.228	.260	.211	.341	.265	.415	.272	.441

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($\tilde{NP}_{t,t+k}$) 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 $\tilde{NP}_{t,t+k}^i$ of subject *i* in period *t* are calculated as $\tilde{P}_{t,t+k}^i / \bar{P}_{t,t+k} - 1$. 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($\tilde{NP}_{t,t+k}$)) measure the impact of heterogeneous beliefs in the various treatments with respect to Treatment LOW. For interaction terms, we test whether the respective coefficient is different from zero. Clustered standard errors on a market level are provided in parentheses. ** *p* < .05; *** *p* < .01 (double-sided test).

Table 6
Heterogeneous beliefs, bubble formation, and overvaluation

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	.209	.222	.240	.359	.414	.419

This table reports OLS regressions measuring the impact of heterogeneous beliefs on overvaluation across all eleven treatments. RD_t serves as dependent variable and measures overvaluation in period t . $SD(\widehat{NP}_{t,t+k})$ 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 $\widehat{NP}_{t,t+k}^i$ of subject i in period t are calculated as $\widehat{P}_{t,t+k}^i / \widehat{P}_{t-1}$. 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., $BUBBLE \times SD(\widehat{NP}_{t,t+k})$) measure the impact of heterogeneous beliefs in the bubble markets across all treatments. Clustered standard errors at market level are in parentheses. ** $p < .05$; *** $p < .01$ (double-sided test).

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 Online Appendix outlines further details about 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, Noe, and Tice 2009). To obtain a more comprehensive picture of 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 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.²⁹ VOLA measures price volatility by using log-returns of all market prices within a period. TURNOVER stands for total

²⁹ Like 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.

Table 7
Pairwise Mann-Whitney U-tests of bid-ask spread, intraperiod price volatility, and turnover

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 ^{MIXED}		17.20			12.18			20.16	

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^{PROF} (Treatment HIGH populated by professionals, but with nonpublic knowledge about trader composition of market), NPK^{STUD} (Treatment HIGH populated by students, but with nonpublic knowledge about trader composition of market), and PK^{MIXED} (Treatment HIGH with public knowledge about equal split of professionals/students in the market). The table outlines median treatment values of the respective variables as a percentage, 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. Sample size N for each test is between 20 and 22. ** $p < .05$; *** $p < .01$ (double-sided test).

trading volume normalized by the total number of shares outstanding (TSO) in a period. Table A2 in Online Appendix A details the variables used.

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 professional markets vary between 0.5 (2.0% of the FV) in SHORT and 5.3 (18.8% of the FV) in HIGH, whereas treatment medians are larger with 1.3 (4.5%) and 12.2 (43.5%) in the student markets, respectively. This qualitative pattern holds for each treatment separately with significant subject pool differences in HIGH and NPK^{PROF}. 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 Online 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% 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 Online 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 Online Appendix A we show that the effects are clearly more pronounced for students than for professionals. However, we 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. We set up two classical bubble driver treatments and two classical bubble moderator treatments. In addition, 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.

We found that professionals are not immune to bubbles in experimental asset markets. In fact, we found significant overpricing by professionals in both bubble driver treatments. 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. In the bubble driver treatments, 25% of all professional markets, but 58% of all student markets generated bubbles.

Despite all these differences, we also found qualitatively very similar patterns *within* each subject pool. We showed that 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 1982; Corgnet, DeSantis, and Porter 2019) – 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 (with public knowledge about the trader composition).

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, L. F., N. Charupat, B. K. Church, and R. Deaves. 2006. Margin, short sell, and lotteries in experimental asset markets. *Southern Economic Journal* 73:419–36.
- Akiyama, E., N. Hanaki, and R. Ishikawa. 2017. It is not just confusion! strategic uncertainty in an experimental asset market. *Economic Journal* 127:F563–F580.
- Alevy, J. E., M. S. Haigh, and J. A. List. 2007. Information cascades: Evidence from a field experiment with financial market professionals. *Journal of Finance* 62:151–80.
- Allen, F., and D. Gale. 2000. Bubbles and crises. *Economic Journal* 110:236–55.
- Allen, F., and G. Gorton. 1993. Churning bubbles. *Review of Economic Studies* 60:813–36.
- Baker, M., and J. Stein. 2004. Market liquidity as a sentiment indicator. *Journal of Financial Markets* 3:271–99.

- Barberis, N. 2018. Psychology-based models of asset prices and trading volume. Working Paper.
- Baron-Cohen, S., S. Wheelwright, J. Hill, Y. Raste, and I. 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. *Journal of Child Psychology and Psychiatry and Allied Disciplines* **42**:241–51.
- Bilker, W. B., J. A. Hansen, C. M. Brensinger, J. Richard, R. E. Gur, and R. C. Gur. 2012. Development of abbreviated nine-item forms of the raven's standard progressive matrices test. *Assessment* **19**:354–69.
- Blais, A.-R., and E. U. Weber. 2006. A Domain-Specific Risk-Taking (DOSPRT) scale for adult populations. *Judgment and Decision Making* **1**:33–47.
- Bloomfield, R., and A. Anderson. 2010. *Behavioral finance: Investors, corporations, and markets*. Hoboken, NJ: John Wiley & Sons.
- Bock, O., I. Baetge, and A. Nicklisch. 2014. hroot: Hamburg registration and organization online tool. *European Economic Review* **71**:117–20.
- Bosch-Rosa, C., T. Meissner, and A. Bosch-Domènech. 2018. Cognitive bubbles. *Experimental Economics* **21**:132–53.
- Bruguier, A. J., S. R. Quartz, and P. Bossaerts. 2010. Exploring the nature of "trader intuition," *Journal of Finance* **65**:1703–23.
- Brunnermeier, M. K. 2001. *Asset pricing under asymmetric information: Bubbles, crashes, technical analysis, and herding*. Oxford: Oxford University Press.
- . 2008. Bubbles. Steven N. Durlauf, Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*. Elsevier, 1221–288.
- . 2009. Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic Perspectives* **23**:77–100.
- Brunnermeier, M. K., and M. Oehmke. 2013. Bubbles, financial crises, and systemic risk. *Handbook of the economics of finance*, vol. 2. Eds. G. Constantinides, M. Harris, and R. Stulz. Amsterdam, the Netherlands: Elsevier.
- Brunnermeier, M. K., and I. Schnabel. 2016. Bubbles and central banks: Historical perspectives. In *Central banks at a crossroads: What can we learn from history?* Eds. M. D. Bordo, Ø. Eitheim, M. Flandreau, and J. F. Qvigstad. Cambridge, UK: Cambridge University Press.
- Burks, S. V., J. P. Carpenter, L. Goette, and A. Rustichini. 2009. Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Science* **106**:7745–50.
- Caginalp, G., D. Porter, and V. L. Smith. 1998. Initial cash/asset ratio and asset prices: An experimental study. *Proceedings of the National Academy of Sciences* **95**:756–61.
- . 2001. Financial bubbles: Excess cash, momentum and incomplete information. *Journal of Psychology and Financial Markets* **2**:80–99.
- Camerer, C., and K. Weigelt. 1991. Information mirages in experimental asset markets. *Journal of Business* **64**:463–93.
- Cassidy, J. 2010. Interview with Eugene Fama. *New Yorker*, January 13. <https://www.newyorker.com/news/john-cassidy/interview-with-eugene-fama>.
- Chen, D. L., M. Schonger, and C. Wickens. 2016. oTree - an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* **9**:88–97.
- Cheng, I.-H., S. Raina, and W. Xiong. 2014. Wall street and the housing bubble. *American Economic Review* **104**:2797–829.
- Cheung, S. L., M. Hedegaard, and S. Palan. 2014. To see is to believe - common expectations in experimental asset markets. *European Economic Review* **66**:84–96.

- Christelis, D., T. Jappelli, and M. Padula. 2010. Cognitive abilities and portfolio choice. *European Economic Review* **54**:18–38.
- Cohn, A., E. Fehr, and M. A. Maréchal. 2014. Business culture and dishonesty in the banking industry. *Nature* **516**:86–89.
- . 2017. Do professional norms in the banking industry favor risk-taking? *Review of Financial Studies* **30**:3801–23.
- Corgnet, B., M. DeSantis, and D. Porter. 2015. Revisiting information aggregation in asset markets: Reflective learning and market efficiency. Working Paper.
- . 2018. What makes a good trader? on the role of intuition and reflection on trader performance. *Journal of Finance* **73**:1113–37.
- . 2019. The distribution of information and the price efficiency of markets. *Journal of Economic Dynamics and Control*. Advance Access published February 15, 2019, 10.1016/j.jedc.2019.02.006.
- Corgnet, B., R. Hernán-González, P. Kujal, and D. Porter. 2015. The effect of earned versus house money on price bubble formation in experimental asset markets. *Review of Finance* **19**:1455–88.
- Crosetto, P., and A. Filippin. 2013. The “bomb” risk elicitation task. *Journal of Risk and Uncertainty* **47**:31–65.
- DeMartino, B., J. O’Doherty, D. Ray, P. Bossaerts, and C. Camerer. 2013. In the mind of the market: Theory of mind biases value computation during financial bubbles. *Neuron* **79**:1222–31.
- Dohmen, T. J., A. Falk, D. Huffman, J. Schupp, U. Sunde, and G. Wagner. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* **9**:522–50.
- Dufwenberg, M., T. Lindqvist, and E. Moore. 2005. Bubbles and experience: An experiment. *American Economic Review* **95**:1731–37.
- Eckel, C. C., and S. Füllbrunn. 2015. Thar she blows? gender, competition, and bubbles in experimental asset markets. *American Economic Review* **105**:906–20.
- Engsted, T. 2016. Fama on bubbles. *Journal of Economic Surveys* **30**:370–76.
- Fama, E. F. 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* **25**:383–17.
- Fang, V. W., T. H. Noe, and S. Tice. 2009. Stock market liquidity and firm value. *Journal of Financial Economics* **94**:150–69.
- Fellner, G., and B. Maciejovsky. 2007. Risk attitude and market behavior: Evidence from experimental asset markets. *Journal of Economic Psychology* **28**:338–50.
- Fischbacher, U. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* **10**:171–8.
- Frederick, S. 2005. Cognitive reflection and decision making. *Journal of Economic Perspectives* **19**:25–42.
- Galbraith, J. K. 1994. *A short history of financial euphoria*. New York: Penguin.
- Greenwood, R., A. Shleifer, and Y. You. 2019. Bubble for fama. *Journal of Financial Economics* **131**:20–43.
- Greiner, B. 2004. An online recruitment system for economic experiments. In *Forschung und wissenschaftliches Rechnen*. Eds. K. Kremer and V. Macho, 79–93. GWDG Bericht 63. Gesellschaft fuer Wissenschaftliche Datenverarbeitung.
- Griffin, J., J. Harris, T. Shu, and S. Topaloglu. 2011. Who drove and burst the tech bubble? *Journal of Finance* **66**:1251–90.
- Grinblatt, M., M. Keloharju, and J. Linnainmaa. 2011. IQ and stock market participation. *Journal of Finance* **66**:2121–64.
- . 2012. IQ, trading behavior, and performance. *Journal of Financial Economics* **104**:339–62.

- Gürkaynak, R. S. 2008. Econometric tests of asset price bubbles: taking stock. *Journal of Economic Surveys* **22**:166–86.
- Haigh, M. S., and J. A. List. 2005. Do professional traders exhibit myopic loss aversion? An experimental analysis. *Journal of Finance* **60**:523–34.
- Harris, T., and J. W. Hardin. 2013. Exact Wilcoxon signed-rank and Wilcoxon Mann–Whitney ranksum tests. *Stata Journal* **13**:337–43.
- Harrison, M., and D. Kreps. 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics* **92**:323–36.
- Haruvy, E., Y. Lahav, and C. N. Noussair. 2007. Traders' expectations in asset markets: experimental evidence. *American Economic Review* **97**:1901–20.
- Haruvy, E., and C. N. Noussair. 2006. The effect of short selling on bubbles and crashes in experimental spot asset markets. *Journal of Finance* **61**:1119–57.
- Hefti, A., S. Heinke, and F. Schneider. 2016. Mental capabilities, trading styles, and asset market bubbles: Theory and experiment. Working Paper.
- Hogan, S., R. Jarrow, M. Teo, and M. Warachka. 2004. Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial Economics* **73**:525–65.
- Holmström, B., and J. Tirole. 1993. Market liquidity and performance monitoring. *Journal of Political Economy* **101**:678–709.
- Holt, C. A., M. Porzio, and M. Y. Song. 2017. Price bubbles, gender, and expectations in experimental asset markets. *European Economic Review* **100**:72–94.
- Hong, H., and J. Stein. 2007. Disagreement and the stock market. *Journal of Economic Perspectives* **21**:109–28.
- Huber, J., and M. Kirchler. 2012. The impact of instructions and procedure on reducing confusion and bubbles in experimental asset markets. *Experimental Economics* **15**:89–105.
- Janssen, D.-J., S. Füllbrunn, and U. Weitzel. 2018. Individual speculative behavior and overpricing in experimental asset markets. *Experimental Economics*. Advance Access published March 15, 2018, 10.1007/s10683-018-9565-4.
- Jimenez, N., I. Rodriguez-Lara, J.-R. Tyran, and E. Wengström. 2018. Thinking fast, thinking badly. *Economics Letters* **162**:41–44.
- Kahneman, D. 2011. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- Kaustia, M., E. Alho, and V. Puttonen. 2008. How much does expertise reduce behavioral biases? The case of anchoring effects in stock return estimates. *Financial Management* **37**:391–412.
- Kindleberger, C. P., and R. Z. Aliber. 2011. *Manias, panics, and crashes: A history of financial crises*, 6th ed. New York: Palgrave Macmillan.
- King, R., V. L. Smith, A. W. Williams, and M. V. Boening. 1993. The robustness of bubbles and crashes in experimental stock markets. In *Nonlinear dynamics and evolutionary economics*, eds. R. Day and P. Chen, 183–200. Oxford, UK: Oxford University Press.
- Kirchler, M., C. Bonn, J. Huber, and M. Razen. 2015. The “inflow-effect” – Trader inflow and price efficiency. *European Economic Review* **77**:1–19.
- Kirchler, M., J. Huber, and T. Stöckl. 2012. Thar she bursts: Reducing confusion reduces bubbles. *American Economic Review* **102**:865–83.
- Kirchler, M., F. Lindner, and U. Weitzel. 2018. Rankings and risk-taking in the finance industry. *Journal of Finance* **73**:2271–302.
- Kleinlercher, D., J. Huber, and M. Kirchler. 2014. The impact of different incentive schemes on asset prices. *European Economic Review* **68**:137–50.

- Lee, C., and B. Swaminathan. 2000. Price momentum and trading volume. *Journal of Finance* **55**:2017–69.
- Lei, V., C. N. Noussair, and C. R. Plott. 2001. Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. actual irrationality. *Econometrica* **69**:831–59.
- List, J. A., and M. S. Haigh. 2005. A simple test of expected utility theory using professional traders. *Proceedings of the National Academy of Science* **102**:945–48.
- Mackintosh, N. J. 2011. History of theories and measurement of intelligence. In *The Cambridge handbook of intelligence*, R. J. Sternberg and S. B. Kaufman, eds., 3–19. Cambridge, UK: Cambridge University Press.
- Miller, E. M. 1977. Risk, uncertainty, and divergence of opinion. *Journal of Finance* **32**:1151–68.
- Moinas, S., and S. Pouget. 2013. The bubble game: an experimental study of speculation. *Econometrica* **81**:1507–39.
- Morris, S. 1996. Speculative investor behavior and learning. *Quarterly Journal of Economics* **111**:1111–33.
- Noussair, C. N., S. Robin, and B. Ruffieux. 2001. Price bubbles in laboratory asset markets with constant fundamental values. *Experimental Economics* **4**:87–105.
- Noussair, C. N., and S. Tucker. 2016. Cash inflows and bubbles in asset markets with constant fundamental values. *Economic Inquiry* **54**:1596–606.
- Noussair, C. N., S. Tucker, and Y. Xu. 2014. A futures market reduces bubbles but allows greater profit for more sophisticated traders. Working Paper.
- . 2016. Futures markets, cognitive ability, and mispricing in experimental asset markets. *Journal of Economic Behavior & Organization* **130**:166–79.
- Oechssler, J., A. Roider, and P. W. Schmitz. 2009. Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization* **72**:147–52.
- Ofek, E., and M. Richardson. 2003. DotCom mania: The rise and fall of internet stock prices. *Journal of Finance* **58**:1113–38.
- Olderbak, S., O. Wilhelm, G. Olaru, M. Geiger, M. W. Brennehan, and R. 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, S. 2015. GIMS – Software for asset market experiments. *Journal of Behavioral and Experimental Finance* **5**:1–14.
- Palfrey, T. R., and S. W. Wang. 2012. Speculative overpricing in asset markets with information flows. *Econometrica* **80**:1937–76.
- Plott, C., and S. Sunder. 1982. Efficiency of experimental security markets with insider information: An application of rational-expectations models. *Journal of Political Economy* **90**:663–98.
- Raven, J. 2000. The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology* **41**:1–48.
- Razen, M., J. Huber, and M. Kirchler. 2017. Cash inflow and trading horizon in asset markets. *European Economic Review* **92**:359–84.
- Robin, S., K. Stráznická, and M. C. Villeval. 2012. Bubbles and incentives: An experiment on asset markets. Working Paper.
- Roth, B., and A. Voskort. 2014. Stereotypes and false consensus: How financial professionals predict risk preferences. *Journal of Economic Behavior & Organization* **107**:553–65.
- Scheinkman, J. A., and W. Xiong. 2003. Overconfidence and speculative bubbles. *Journal of Political Economy* **111**:1183–219.
- Scherbina, A., and B. Schlusche. 2014. Asset price bubbles: A survey. *Quantitative Finance* **14**:589–604.

- Smith, A., T. Lohrenz, J. King, R. Montague, and C. Camerer. 2014. Irrational exuberance and neural crash warning signals during endogenous experimental market bubbles. *Proceedings of the National Academy of Sciences* **111**:10503–8.
- Smith, V. L., G. L. Suchanek, and A. W. Williams. 1988. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica* **56**:1119–51.
- Smith, V. L., M. Van Boening, and C. P. Wellford. 2000. Dividend timing and behavior in laboratory asset markets. *Economic Theory* **16**:567–83.
- Stöckl, T., J. Huber, and M. Kirchler. 2010. Bubble measures in experimental asset markets. *Experimental Economics* **13**:284–98.
- Sutter, M., J. Huber, and M. Kirchler. 2012. Bubbles and information: An experiment. *Management Science* **58**:384–93.
- Toplak, M. E., R. F. West, and K. E. Stanovich. 2011. The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition* **39**:1275–89.
- . 2014. Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning* **20**:147–68.
- Udland, M. 2015. Robert Shiller: 'There is a bubble element to what we're seeing.' *Business Insider*, May 31. www.businessinsider.com/robert-shiller-stock-market-bubble-2015-5?r=UK.
- Xiong, W. 2013. Bubbles, crises, and heterogeneous beliefs. Working Paper.
- Xiong, W., and J. Yu. 2011. The Chinese warrants bubble. *American Economic Review* **101**:2723–53.