Cognitive Bubbles*

Ciril Bosch-Rosa[†]
Thomas Meissner[‡]
Antoni Bosch-Domènech[§]

November 10, 2015

Abstract

Smith et al. (1988) reported large bubbles and crashes in experimental asset markets, a result that has been replicated many times. Here we test whether the occurrence of bubbles depends on the experimental subjects' cognitive sophistication. In a two-part experiment, we first run a battery of tests to assess the subjects' cognitive sophistication and classify them into low or high levels. We then invite them separately to two asset market experiments populated only by subjects with either low or high cognitive sophistication. We observe classic bubble and crash patterns in the sessions populated by subjects with low levels of cognitive sophistication. Yet, no bubbles or crashes are observed with our sophisticated subjects, indicating that cognitive sophistication of the experimental market participants has a strong impact on price efficiency.

Keywords Asset Market Experiment · Bubbles · Cognitive Sophistication

JEL Classification C91 · D12 · D84 · G11

*The authors acknowledge the helpful comments of Michael Greenacre, Frank Heinemann, Joaquim Silvestre, Stefan Palan, Luba Petersen and participants of the Berlin Behavioral Economics colloquium. They also thank Ernan Haruvy, Yaron Lahav, and Charles Noussair for their help with the experimental software, and Pablo Lopez-Aguilar for helping with its implementation. The first two authors acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG) through CRC 649 "Economic Risk". The last author acknowledges the financial support from the Spanish *Ministerio de Educación, Cultura y Deporte* under research project ECO2011-25295. The first author would also like to thank the International Foundation for Research in Experimental Economics and their Small Grants Program for financial support.

[†]Berlin University of Technology and CUNEF Madrid

[‡]Berlin University of Technology and Grenoble Ecole de Management

[§]Universitat Pompeu Fabra and Barcelona GSE

1 Introduction

In 1988, Vernon Smith, Gerry Suchanek and Arlington Williams (SSW) (Smith et al., 1988) published a seminal paper reporting the results of experiments on the efficiency of asset markets. In their experiment, subjects are first endowed with assets and experimental currency, and then are allowed to trade assets for currency in a multi-period double auction market. At the end of each period, assets pay a stochastic dividend whose distribution is common knowledge. At the end of the experiment, assets have no buyback value and subjects are paid in cash according to the amount of experimental currency they have accumulated. The asset's fundamental value (FV) at any period can be calculated as the number of periods left times the expected dividend per period. The advantage of such experimental asset markets is that, contrary to real world financial markets, the asset's fundamental value is known to all participants of the market and also to any observer attempting to assess the efficiency of these markets.

SSW observed large positive price deviations from the FV (also called bubbles) followed by dramatic crashes towards the end of the experiment. To the surprise of most, these bubbles turned out to be extremely resilient to replications under different treatments.[1,2]

Thus, the results became canonical to the extent that seldom a paper in the economic experimental literature has spawned such a large industry of replications and follow-ups. Stefan Palan in a recent survey (Palan, 2013) documents the main findings based on the results from 41 published papers, 3 book chapters and 20 working papers, describing them under 32 observations. Palan concludes with an optimistic appraisal: "Hundreds of SSW markets have been run, yielding valuable insights into the behavior of economic actors and the factors governing bubbles"(p. 570).

[1]E.g.: Porter and Smith (1995), Caginalp et al. (1998), Caginalp et al. (2000), Smith et al. (2000), Dufwenberg et al. (2005), Noussair and Tucker (2006), Haruvy and Noussair (2006), Haruvy et al. (2007), Hussam et al. (2008), Williams (2008).

[2]In a recent interview Vernon Smith reminisced about his earlier experiments and declared that the design of his SSW experiment was transparent and, consequently, he could not understand why subjects would not trade at the fundamental value: "We then turned to asset markets in the 1980s, and we started with a *very transparent* market, an asset that could be re-traded but there was a yield, a dividend on it that was common information. And we thought that would be very simple. It would be transparent and people would trade at fundamental value. Well, wrong [...] *These markets are very subject to bubbles in the lab.* And people get caught up in self-reinforcing expectations of rising prices. We don't know where that comes from. It's incredible, but they do." (Emphasis added) http://www.econtalk.org/archives/2014/11/vernon_smith_on_2.html. November 17 2014.

We are not so sure about that. We show below that the bubbles and crashes observed in experimental asset markets disappear when the participants have a sufficient level of cognitive sophistication. This would suggest that bubbles and crashes are not necessarily intrinsic to SSW experimental asset markets, but rather that they could depend on the cognitive profile of the experimental subjects.[3]

The idea that the observed bubbles and crashes in the SSW-type experiments may be due to some lack of understanding by the participants of the experiments is not entirely new. Huber and Kirchler (2012) and Kirchler et al. (2012) have managed to reduce bubbles in their experiments by either offering a more thorough rendering of the market or describing the asset as a "stock from a depletable gold mine". According to them, an easier understanding of the market diminishes the bubbles. However, this interpretation has been challenged. Baghestanian and Walker (2015) argue that particular features of the experimental design by Kirchler et al. (2012) generate asset prices equal to the fundamental value through increased focalism or anchoring, and not because agents are less confused.[4]

In this paper, we test whether the occurrence of bubbles in SSW-type experiments depends on the subjects' cognitive sophistication. Building on previous evidence relating some degree of misunderstanding with asset price bubbles, we would expect markets populated only by high sophistication subjects to generate fewer bubbles compared to markets populated by less sophisticated ones. To test this hypothesis we design a two-part experiment: In the first part we invite subjects to participate in a battery of tasks that, we reasonably believe, allow us to approximate their "cognitive sophistication". In

[3]If SSW experimental asset markets are not per se prone to bubble, then some comments that rely on the external validity of this type of experiments become questionable. See, e.g. (Knott, 2012, p.86) who referencing SSW experiments writes: "In simulated economic markets played with student participants, the results show that price bubbles occur naturally. [...] These analyses of incentives and institutional relationships in the economy in the past decade help to explain in part the private market failure that led economic actors to engage in increasingly risky behavior. Experimental economics also shows why the dramatic economic changes and financial innovations during this period may have added to risk taking and the failure in the market. "

[4]In a paper on trust and reciprocity, McCabe and Smith (2000) present the result of one asset experiment with 22 subjects whose decisions track the fundamental value of the asset *from the very first period*. We value this as an inconsequential result since the participants were advanced graduate students (in the third or fourth year of their Ph.D) from all over the world who had traveled to Arizona to participate in a 5-days course on experimental economics *with Vernon Smith*. One should suspect that these grad students interested in experimental economics had prepared well for their expensive trip and had read or were already familiar with some of Prof. Smith's most prominent papers, among them his famous 1988 paper about the asset market experiment.

part two, which is scheduled for a later date, we invite subjects who score low (high) in our tasks of cognitive sophistication to participate in an asset market experiment populated only by low (high) sophistication subjects. The results of the experiment verify our expectations. Bubbles and crashes persist when the experimental subjects are selected because of their lower cognitive scores. Interestingly, bubbles vanish completely when we run the experiment with the more sophisticated subjects.

2 The Cognitive Tasks

In the first part of the experiment, subjects were asked to participate in a number of time-constrained tasks to evaluate, among other items, their cognitive abilities. A total of 352 subjects participated in these tasks. All subjects were recruited through ORSEE (Greiner, 2015). The invitation mails instructed subjects to only sign up if they were available on a second date in which a new round of experiments would take place. The second dates proposed in the email varied between one and five weeks after the initial session. Except for the dates, no further information was given about what was expected of them in the second part of the experiment. Subjects were mostly undergraduate students with a variety of backgrounds, ranging from anthropology to electrical engineering or even musicology. Sessions were run at the Experimental Economics Laboratory of the Berlin University of Technology. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

Subjects began this first part of the experiment with a "Cognitive Reflection Test" (CRT)(Frederick, 2005), followed by playing a "Guessing Game" (Nagel, 1995) against other subjects, then a "Guessing Game Against Oneself", and finally 12 rounds of "Race to 60". There was no feedback to the participants during or in-between tasks.[5] We selected the first three tasks in order to test subjects on three dimensions that we deem relevant for understanding SSW asset markets: cognitive reflection, strategic sophistication and backward induction ability. We also elicited risk preferences by using a Holt and Laury price list (Holt and Laury, 2002). However, this task was not time constrained, and not included in our measure of cognitive sophistication. In the following, we provide a brief

[5]Exceptionally, in the Race to 60 game there was some feedback as subjects learned at the end of each round whether they have won that round.

description of the cognitive tasks. For a more detailed description we refer the reader to the appendices.

The CRT is a three-item task of an algebraic nature, designed to measure the ability to override an intuitive response that is incorrect and to engage in further *reflection* that leads to the correct response. It has been shown that the test results are highly correlated with IQ level, with compliance to expected utility theory, as well as with lower discount rates (higher patience) for short horizons and lower levels of risk aversion (see e. g. Frederick (2005) and Oechssler et al. (2009)). With respect to experimental asset markets, Corgnet et al. (2014) and Noussair et al. (2014) find that CRT scores correlate positively with earnings. In the Guessing Game (against others), participants were asked to guess a number between 0 and 100 and were paid based on how close their choice was to 2/3 of the average of all the guesses within their session. The guess gives an indication of the participant's capacity to perform iterative reasoning in a strategic environment. A simpler way (because devoid of any strategic concerns) of testing the basic capacity for *iterative reasoning* is the Guessing Game Against Oneself, where a participant has to pick two numbers between 0 and 100, and each number is paid independently, according to how close it is to 2/3 of the average of the two chosen numbers.[6] Finally, participants played Race to 60, a variant of the race game (Gneezy et al. (2010), Levitt et al. (2011)), for 12 rounds against a computer. In this game, the participants and the computer sequentially choose numbers between 1 and 10, which are added up. Whoever is first to push the sum to or above 60 wins the game. The game is solvable by *backward induction*, and the first mover can always win. Subjects always move first and therefore, independently of the computer sophistication, they can always win the game by applying backward induction.

We finally computed an index of cognitive sophistication, S_i, as a weighted average of the results obtained by each subject (i) in the tasks described above. This index has a value between zero and one, and we use it to rank our subjects. A subject is classified as having Low (High) cognitive sophistication if she is in the lower (upper) 30% of the

[6] To our knowledge, this is the first experiment in which a guessing game against oneself is played. Petersen and Winn (2014) have a similar setup in which subjects compete against themselves in a monopolistic competition environment.

distribution of S_i.[7,8] We counted 84 subjects with low sophistication and 83 with high sophistication.[9]

3 The Experiment

All sessions of the asset market experiment followed the design of Haruvy et al. (2007), except that our subjects participated in groups of seven (instead of nine), we did not allow for practice runs, and had three (instead of four) repetitions of the market. Subjects were endowed with a bundle composed of Talers (our experimental currency) and a number of assets. Three subjects received 1 asset and 472 Taler, one subject received 2 assets and 292 Taler, and three subjects received 3 assets and 112 Taler.[10] Each session consisted of three repetitions (that are called rounds) and each round lasted 15 periods. In each period subjects were able to trade units of the asset (called "shares" in the instructions) in a call market with other subjects.[11,12] At the end of every period, each share paid a stochastic

[7]See Appendix C for detailed results of each task, the construction of the Cognitive Sophistication measure S_i, as well as its final distribution.

[8]The index aggregates the results of all cognitive tasks. One may wonder how differently subjects would have been selected if one of the tests had not been used in constructing the index S_i and, ultimately, how different the results of our asset market experiment would have been. In Table 4 of Appendix C we show that the percentage of overlapping subjects when one test is dropped from the index is high for both High and Low sophistication groups, ranging from 72% to 86%.

[9]After the first batch of sessions, and in order to run three additional High sessions (see 4.2 below for an explanation), we invited more subjects to be tested at a later time. We classified these subjects as being of High Sophistication if they were above the boundaries of our first batch of tested subjects. In total we ended up inviting 92 subjects with high scores. Participants who were not classified as having either Low or High cognitive sophistication, i.e. the remaining 40%, were not subjects in the asset market experiment.

[10]Subjects knew about their private endowment and were told that participants could have different endowments.

[11]In order to trade, subjects post buy or sell orders, specifying the amount of shares they want to buy (sell) and the maximal (minimal) price they want to pay (get). The price at which trades happen is then set by the experimental software as the lowest price at which there is an equal number of shares offered for purchase and sale.

[12]The SSW-type of asset market experiment has been run in the literature with different institutional arrangements, basically either a continuous double auction or a call market. A call market, as in Haruvy et al. (2007), allows only one price per period, as opposed to the possibility of multiple prices in the continuous double auction, thus yielding a crisp description of the price dynamics. It also helps participants to better understand the price prediction process, and mitigates the possibility of subjects trying to manipulate prices to improve their prediction scores. Importantly, these advantages come at no cost, as call markets and continuous double auction markets do not differ in their results. See Palan (2013) (in particular his Observation 27: "A two-sided sealed-bid call auction does not significantly attenuate the bubble") for a detailed discussion on the matter and references to experiments comparing both

dividend of either 0, 4, 14 or 30 Taler with equal probability (expected dividend, 12 Taler). Shares had no buy back value at the end of the 15 periods. Hence, the fundamental value of the asset in period t is $12(16-t)$. At the end of the experiment, subjects were paid in cash according to the sum of Talers they have accumulated at the end of all three rounds. At every period and before any trade took place, subjects were asked to predict the price of the asset for all upcoming periods. So, in period 1 subjects were asked to predict 15 prices, in period 2 they were asked to predict 14, and so on. Subjects were incentivized to give accurate predictions: They were paid 5 extra Taler if a price prediction was within 10% of the actual price, 2 Taler if a prediction was within 25%, 1 Taler if a prediction was within 50% of the price, and nothing otherwise.[13]

At the end of each period, subjects were told the price at which the asset was traded, the dividend they collected, their profits, their share and cash holdings, and their accumulated profits from their price predictions. Each session (which, as mentioned above, is composed of three rounds) was programmed to last for two and a half hours, but a few sessions went somewhat beyond.[14]

Before turning to the results, recall that our experiment had two different treatments:

- Low Sophistication treatment: all subjects that took part in this treatment were from the lower 30% of the distribution of S_i
- High Sophistication treatment: all subjects that took part in this treatment were from the upper 30% of the distribution of S_i,

and that the main purpose of the experiment was to compare the asset price dynamics in the two treatments.

4 Results

4.1 Low Sophistication Treatment

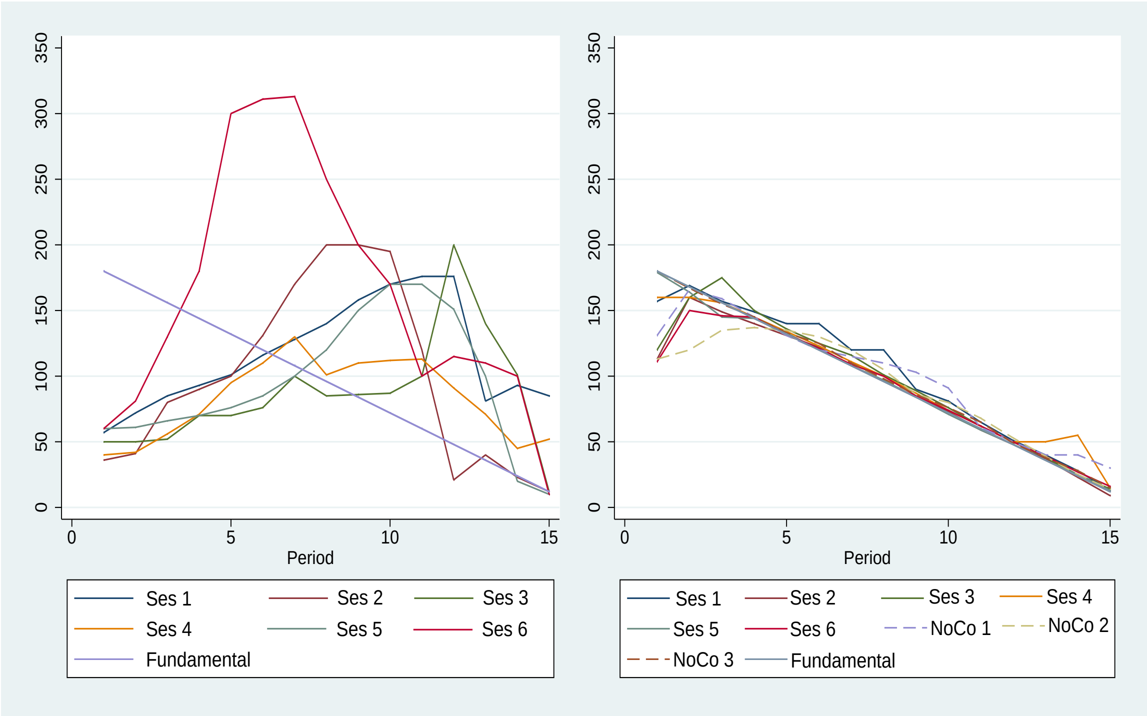
We ran six sessions of the experiment under the Low Sophistication treatment. The results in all six sessions are the usual ones reported in the literature. The diagram

institutions.

[13]Notice that subjects were paid independently for all predictions they made of the price for a certain period. For example, for the price in period 2 subjects were paid twice; once for the prediction they made in period 1, and once again for the prediction they made in period 2.

[14]The instructions for the experiment can be found in Appendix D.

Figure 1: Asset Price Dynamics



Notes: Asset prices in the first rounds of the two treatments: Six sessions in the Low Sophistication treatment (on the left) and nine sessions in the High Sophistication treatment (on the right). The diagonal line corresponds to the asset fundamental value.

on the left of Figure 1 shows the price dynamics for the first round of each of the six sessions. Prices begin below the fundamental value of the asset, climbing in the following periods well above and beyond it, to finally crash near the last period. In summary, when the experimental subjects belong to the lower end of the distribution of Cognitive Sophistication, we observe the classic price dynamics of bubbles and crashes. As in previous experiments reported in the literature, bubbles tend to diminish somewhat in the second and third rounds of a session with the same subjects and endowments.[15]

4.2 High Sophistication Treatment

Under the High Sophistication treatment we ran a total of nine sessions where all subjects were chosen from to the upper 30% of the distribution of S_i. In six of these sessions subjects were told that everyone in the session had "scored above average" in the cognitive tasks. The results for these six High Sophistication sessions are striking by how markedly they differ from the standard results of bubbles and crashes. In all six sessions, asset prices

[15]The second and third rounds of each session are not the focus of this paper. Therefore, they are not reported in its main body. See Appendix D for the results from rounds two and three.

track the fundamental value (almost) perfectly, as shown in the diagram on the right of Figure 1 with the labels Sessions 1 to 6. While in both treatments, Low and High, prices start below the fundamental value (as one would expect if subjects are risk averse and begin the experiment by testing the market), in the High Sophistication treatment prices reach the fundamental value sooner and hover close to it for the remaining periods. Because we were in doubt whether the disappearance of the bubbles was due to the high cognitive scores of the experimental subjects or to their shared knowledge of it, we ran three additional sessions.These sessions were populated by High Sophistication subjects who were *not told* that they had been selected because of their high scores (dashed lines in Figure 1 with labels NoCo1 to 3).[16] Again, we observe that prices approach the fundamental value of the asset from below and stay close to it for the remaining periods. In essence, as before, bubbles and crashes vanish.[17] Since we do not observe any differences whether subjects share or not a knowledge for their common sophistication, we pool the nine sessions together in the graph to the right of Figure 1, to facilitate the comparison with the Low Sophistication treatment on the left of it. In the second and third rounds of the High Sophistication treatment we observe basically the same price dynamics.[18]

4.3 Measurement of Mispricing

In order to formally compare the asset price dynamics in our two treatments, we make use of standard bubble measures (see e.g. Stöckl et al. (2010) and Porter and Smith (1995)). These measures are relative absolute deviation (RAD), relative deviation (RD), duration (DUR) and price amplitude (AMP). We also use a measure of our own, which we call

[16] Admittedly, one cannot rule out completely that subjects could independently come to the conclusion that they, and all other invited subjects, were of high cognitive ability, that all of them shared the same beliefs that everyone else was of high ability, that they believed that all other subjects believed what they believed, and so on ad infinitum. Yet, given the temporal spacing between the task sessions and the asset market sessions, the participants' different backgrounds that did not facilitate communication among them, and the fact that the tasks session involved more than cognitive tests, it seems highly unlikely that this potential "common-knowledge-of-sorts" would be driving our treatment effect.

[17] Cheung et al. (2014) show that public knowledge of training on the experimental environment reduces bubbles. They also show that even well-trained subjects can create mispricing when they think that others in the market are non-trained. In our NoCo sessions, subjects were not only not told the cognitive sophistication of the subjects they traded with but, as explained in the previous footnote, had few reasons to believe that the participants' cognitive sophistication was a crucial variable in the asset market.

[18] In two of the nine sessions in this treatment, prices tend to rise somewhat towards the end of the third round. We do not attribute any significance to this pattern, which might well be due to simple boredom from the previous uneventful rounds. See Appendix E for additional comments.

Table 1: Definition of Bubble Measures

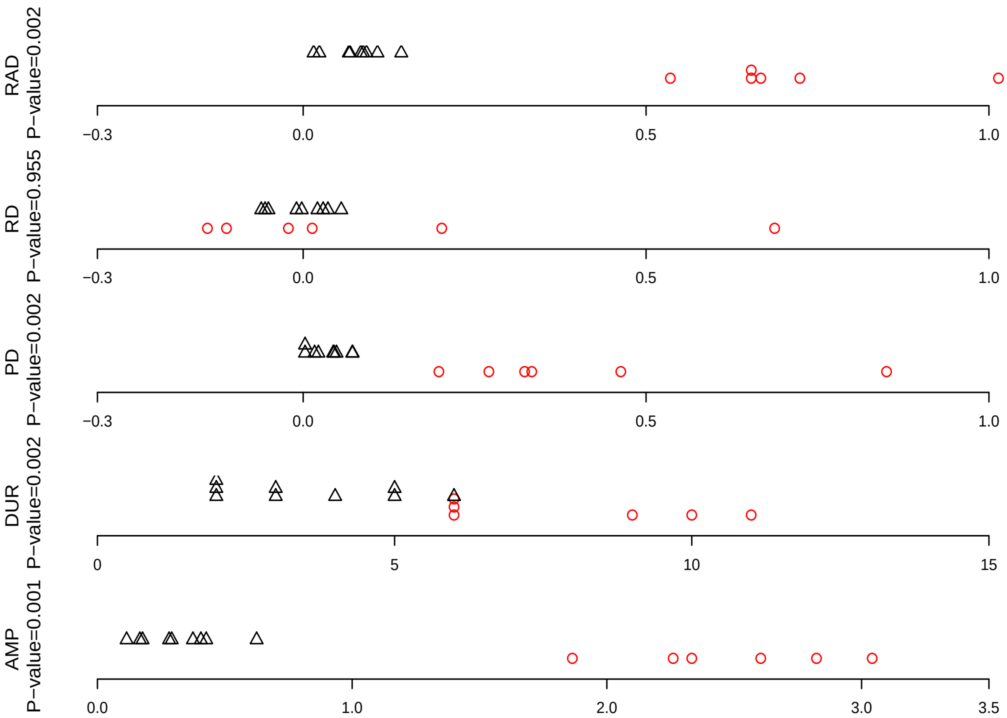
Measure	Formula
RAD	$\frac{1}{N}\sum_{t=1}^{N}\lvert P_t - FV_t\rvert/\overline{FV}$
RD	$\frac{1}{N}\sum_{t=1}^{N}(P_t - FV_t)/\overline{FV}$
PD	$\frac{1}{N_+}\sum_{t=1}^{N}\max\{0,(P_t - FV_t)/\overline{FV}\}$
DUR	$\max\{m : P_t - FV_t < P_{t+1} - FV_{t+1} < ... < P_{t+m} - FV_{t+m}\}$
AMP	$\max\left\{\frac{P_t - FV_t}{FV_1} : t = 1, ..., 15\right\} - \min\left\{\frac{P_t - FV_t}{FV_1} : t = 1, ..., 15\right\}$

positive deviation (PD). This last measure is analogous to RD except that it only takes into account positive deviations from the fundamental value, i.e. overvaluations of the asset. The measures are described in Table 1.

P_t and FV_t denote the observed price and the fundamental value in period t respectively. The number of total periods is $N = 15$, and N_+ denotes the number of rounds in which the deviations from the fundamental have a positive sign. In Figure 2 we show the values of all five measures (RAD, RD, PD, DUR and AMP) for the first round across all sessions: dots correspond to the six Low Sophistication sessions and triangles to the nine High Sophistication sessions. In the High Sophistication treatment, the measures of deviation from the fundamental value (RAD, RD and PD) have a low variance and are grouped together close to zero (the means are 0.077, -0.004, and 0.036, for RAD, RD, and PD, respectively), confirming that in this treatment asset prices stay close to the fundamental value. In contrast, the values of these measures for the Low Sophistication sessions are dispersed with means above zero (means are 0.70, 0.10, and 0.33 for RAD, RD and PD, respectively)[19]. A Mann-Whitney U-test (p-values for RAD, RD and PD are 0.002, 0.955 and 0.002, respectively), indicates that we can safely reject the hypothesis that RAD and PD values come from the same distribution in the two treatments. The measures DUR and AMP paint a similar picture: both the values for the bubble duration and its amplitude are significantly higher in the Low Sophistication treatment compared to the High Sophistication treatment (p-values from a Mann-Whitney U-test are 0.002

[19]While RAD aggregates the absolute distances of the prices to the fundamental values, and therefore, the larger the deviations above and below, the larger is the value it takes, RD can give a result close to zero even if the distances from below and from above are large, provided they are similar in absolute value.

Figure 2: Bubble Measures



Notes: Bubble measures RAD, RD, PD, DUR and AMP in the first rounds for all sessions. Dots and triangles represent Low Sophistication and High Sophistication respectively. P-values were calculated by use of Mann-Whitney U-tests.

and 0.001, for DUR and AMP respectively).

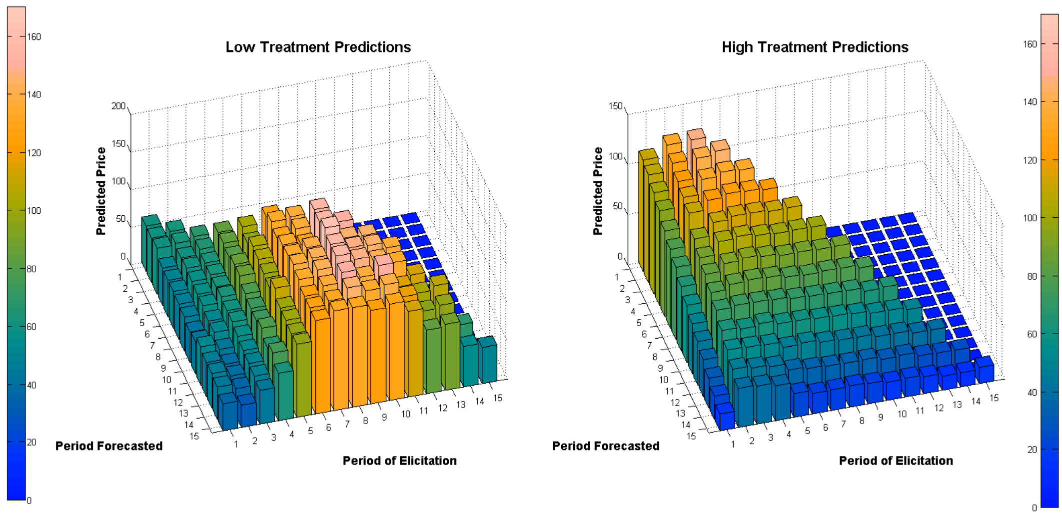
4.4 Predictions

As mentioned above, in every period subjects were asked to predict asset prices for the actual and the remaining periods of the round before posting a bid or ask. These predictions were incentivized to nudge subjects to give their best guess of present and future prices. Figure 3 shows the average predictions (over the first round of all sessions) for the Low (left) and High (right) treatments respectively. One axis indicates the period in which the prediction was elicited (t), while the other shows the predictions for all remaining periods ($16 - t$). The coloring of the bars indicates their height, with lighter colors representing higher price predictions and darker colors representing lower price predictions.[20]

In the Low Treatment, we observe the color pattern running perpendicular to the x-axis as subjects, in each period of elicitation, do not anticipate the price changes across the remaining periods. In contrast, in the High Treatment, bar colors remain unaltered along

[20]While we included the numerical values on the z-axis, it is easier to read the levels of the price predictions from their color coding, as the perspective distorts the vertical view.

Figure 3: Price Predictions



Notes: Average price predictions in the Low Treatment (left) and High Treatment (right). "Period of Elicitation" indicates the period in which the price predictions are made. "Period Forecasted" indicates the periods for which the predictions are made. The colors of the bars code for the average prices predicted, from beige for high prices to dark blue for low ones.

the x-axis, indicating that subjects on average predicted the same price for each period independently of the period in which prices were elicited. In other words, they anticipated from the beginning of the experiment what was bound to happen and, therefore, did not have to change their predictions as the experiment proceeded.

In a nutshell, Figure 3 shows that in the Low Treatment subjects keep adjusting their predictions to the current price, so that there is little difference between the current price and their next-price prediction. Instead, in the High Treatment, the price predictions stay very close to the fundamental values. That subjects in Low passively predict the last price observed for *all remaining periods* indicates a degree of backward induction incompetence. Low Sophistication subjects are apparently lost, unable to anticipate what is coming next. High Sophistication subjects, on the contrary, appear to understand what the experiment is all about. They predict well, and bubbles basically vanish in their sessions.

4.5 Risk Preferences

Whether risk aversion is correlated with cognitive ability is a question that has received some attention in the literature recently. Dohmen et al. (2010) find that risk aversion is correlated negatively with cognitive ability, i.e. cognitively more able people tend to be less risk averse. Andersson et al. (2015) however, argue that this result is spurious, an artifact caused by the method used to elicit risk aversion. With respect to experimental asset markets, Eckel and Füllbrunn (2015) attribute some of the differences they find in bubble formation between their male and female subjects to differences in risk aversion. In the first part of our experiment, we asked subjects to complete a Holt and Laury price list (Holt and Laury, 2002) before playing the cognitive tasks described in Section 2, in order to elicit their risk preferences.[21] We find no significant differences in risk aversion between our High and Low sophistication groups (Mann-Whitney U-test, p-value $= 0.88$). Moreover, none of the individual components of our measure of cognitive sophistication are significantly correlated with our measure of risk aversion. These findings imply that, at least for our sample and the measures used, risk aversion and cognitive sophistication are unrelated. Furthermore, differences in risk aversion cannot explain the differences in bubble formation between our High and Low sophistication treatments.

5 Conclusion

Our goal was to test the hypothesis that bubbles and crashes observed in SSW-type experimental asset markets are due to the subjects' lack of cognitive sophistication. We use a battery of cognitive tests to separate our pool of subjects into two groups (High and Low Sophistication) and run separate asset market experiments with each group. The results are striking. While the asset markets populated by Low Sophistication subjects show the usual pattern of bubbles and crashes, these vanish when the experimental subjects belong to the High Sophistication group. These results support the hypothesis that the bubbles and crashes observed in SSW-type experimental asset markets are not intrinsic to such markets, but are dependent on the cognitive sophistication of the experimental subjects.[22]

Further explorations can be performed on our data, including individual-level analysis of the subjects' decisions and beliefs. Moreover, additional treatments can be run with mixed-sophistication sessions in order to explore, for instance, whether there is learning on the part of the less sophisticated subjects from the behavior of more sophisticated

[21]see Appendix B.4 for a description of this price list

[22]While it is the case that some experimental asset markets have shown a high degree of price efficiency without cognitively sophisticated subjects (e.g. Smith et al. (2000), Noussair et al. (2001), Kirchler et al. (2012), Stöckl et al. (2014), Kirchler et al. (2015)), the design of these particular markets tend to be even simpler than in SSW, having constant FVs and, in some cases, no dividends.

ones. But these explorations belong to future papers. In the present one we decided to remain focused on the result that high cognitive sophistication eliminates the mispricing in SSW-type asset markets. Whether the effect of the subjects' cognitive sophistication can have an impact in other equally simple or, especially, in more complex experimental markets becomes now a plausible hypothesis.

References

Andersson, O., J.-R. Tyran, E. Wengström, and H. Holm (2015): "Risk Aversion Relates to Cognitive Ability: Preferences or Noise?," *Journal of the European Economic Association*. Cited on page 13.

Baghestanian, S., and T. B. Walker (2015): "Anchoring in experimental asset markets," *Journal of Economic Behavior & Organization*, 116, 15–25. Cited on page 3.

Caginalp, G., D. Porter, and V. Smith (1998): "Initial cash/asset ratio and asset prices: An experimental study," *Proceedings of the National Academy of Sciences*, 95(2), 756–761. Cited on page 2.

Caginalp, G., D. Porter, and V. L. Smith (2000): "Overreactions, Momentum, Liquidity, and Price Bubbles in Laboratory and Field Asset Markets," *Journal of Psychology and Financial Markets*, 1(1), 24–48. Cited on page 2.

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. Cited on page 9.

Corgnet, B., R. Hernán-González, P. Kujal, and D. Porter (2014): "The Effect of Earned Versus House Money on Price Bubble Formation in Experimental Asset Markets," *Review of Finance*, rfu031. Cited on page 5.

Costa-Gomes, M. A., and V. P. Crawford (2006): "Cognition and behavior in two-person guessing games: An experimental study," *The American economic review*, 1737–1768. Cited on page 24.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2010): "Are Risk Aversion and Impatience Related to Cognitive Ability?," *The American Economic Review*, 100(3), 1238–1260. Cited on page 12.

Dufwenberg, M., T. Lindqvist, and E. Moore (2005): "Bubbles and Experience: An Experiment," *American Economic Review*, 95(5), 1731–1737. Cited on page 2.

Eckel, C. C., and S. Füllbrunn (2015): "Thar SHE Blows? Gender, Competition, and Bubbles in Experimental Asset Markets," *American Economic Review*, 105(2), 906–20. Cited on page 13.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental economics*, 10(2), 171–178. Cited on page 4.

Frederick, S. (2005): "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19(4), 25–42. Cited on pages 4 and 5.

Gneezy, U., A. Rustichini, and A. Vostroknutov (2010): "Experience and insight in the race game," *Journal of Economic Behavior & Organization*, 75(2), 144–155. Cited on page 5.

Greiner, B. (2015): "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1, 114–125. Cited on page 4.

Güth, W., M. Kocher, and M. Sutter (2002): "Experimental ‚Äòbeauty contests‚Äô with homogeneous and heterogeneous players and with interior and boundary equilibria," *Economics Letters*, 74(2), 219–228. Cited on page 24.

Haruvy, E., Y. Lahav, and C. N. Noussair (2007): "Traders' Expectations in Asset Markets: Experimental Evidence," *American Economic Review*, 97(5), 1901–1920. Cited on pages 2, 6, and 37.

Haruvy, E., and C. N. Noussair (2006): "The Effect of Short Selling on Bubbles and Crashes in Experimental Spot Asset Markets," *The Journal of Finance*, 61(3), 1119–1157. Cited on page 2.

Holt, C. A., and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5), 1644–1655. Cited on pages 4, 13, and 25.

Huber, J., and M. Kirchler (2012): "The impact of instructions and procedure on reducing confusion and bubbles in experimental asset markets," *Experimental Economics*, 15(1), 89–105. Cited on page 3.

Hussam, R. N., D. Porter, and V. L. Smith (2008): "Thar She Blows: Can Bubbles Be Rekindled with Experienced Subjects?," *American Economic Review*, 98(3), 924–37. Cited on page 2.

Kirchler, M., C. Bonn, J. Huber, and M. Razen (2015): "The "inflow-effect"–Trader inflow and price efficiency," *European Economic Review*, 77, 1–19. Cited on page 13.

Kirchler, M., J. Huber, and T. Stöckl (2012): "Thar She Bursts: Reducing Confusion Reduces Bubbles," *American Economic Review*, 102(2), 865–83. Cited on pages 3 and 13.

Knott, J. H. (2012): "The President, Congress, and the Financial Crisis: Ideology and Moral Hazard in Economic Governance," *Presidential Studies Quarterly*, 42(1), 81–100. Cited on page 3.

Kocher, M. G., and M. Sutter (2006): "Time is money – Time pressure, incentives, and the quality of decision-making," *Journal of Economic Behavior & Organization*, 61(3), 375–392. Cited on page 24.

Levitt, S. D., J. A. List, and S. E. Sadoff (2011): "Checkmate: Exploring Backward Induction among Chess Players," *American Economic Review*, 101(2), 975–90. Cited on page 5.

McCabe, K. A., and V. L. Smith (2000): "A comparison of naive and sophisticated subject behavior with game theoretic predictions," *Proceedings of the National Academy of Sciences*, 97(7), 3777–3781. Cited on page 3.

Nagel, R. (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85(5), 1313–26. Cited on page 4.

Noussair, C., S. Robin, and B. Ruffieux (2001): "Price bubbles in laboratory asset markets with constant fundamental values," *Experimental Economics*, 4(1), 87–105. Cited on page 13.

Noussair, C., and S. Tucker (2006): "Futures Markets and Bubble Formation in Experimental Asset Markets*," *Pacific Economic Review*, 11(2), 167–184. Cited on page 2.

Noussair, C. N., S. J. Tucker, and Y. Xu (2014): "A Futures Market Reduces Bubbles But Allows Greater Profit for More Sophisticated Traders," SSRN Scholarly Paper ID 2490326, Social Science Research Network, Rochester, NY. Cited on page 5.

Oechssler, J., A. Roider, and P. W. Schmitz (2009): "Cognitive abilities and behavioral biases," *Journal of Economic Behavior & Organization*, 72(1), 147–152. Cited on page 5.

Palan, S. (2013): "A Review of Bubbles and Crashes in Experimental Asset Markets," *Journal of Economic Surveys*, 27(3), 570–588. Cited on pages 2 and 6.

Petersen, L., and A. Winn (2014): "Does Money Illusion Matter? Comment," *American Economic Review*, 104(3), 1047–62. Cited on page 5.

Porter, D. P., and V. L. Smith (1995): "Futures Contracting and Dividend Uncertainty in Experimental Asset Markets," *Journal of Business*, 509–541. Cited on pages 2 and 9.

Smith, V. L., M. v. Boening, and C. P. Wellford (2000): "Dividend timing and behavior in laboratory asset markets," *Economic Theory*, 16(3), 567–583. Cited on pages 2 and 13.

Smith, V. L., G. L. Suchanek, and A. W. Williams (1988): "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," *Econometrica*, 56(5), 1119. Cited on pages 1 and 2.

Stöckl, T., J. Huber, and M. Kirchler (2010): "Bubble measures in experimental asset markets," *Experimental Economics*, 13(3), 284–298. Cited on page 9.

——— (2014): "Multi-period experimental asset markets with distinct fundamental value regimes," *Experimental Economics*, 18(2), 314–334. Cited on page 13.

Williams, A. W. (2008): "Chapter 29 Price Bubbles in Large Financial Asset Markets," in *Handbook of Experimental Economics Results*, ed. by C. R. P. a. V. L. Smith, vol. 1, 242–246. Elsevier. Cited on page 2.

Appendix

For Online Publication

A Instructions

The instructions below are translated from the original German instructions. The instructions were read aloud to the participants.

Overview This is the first part of a two-part experiment. The second part will take place this coming Friday, November 7th, 2014. Depending on your decisions in this experiment you may be invited to the second part of the experiment. However, not all participants of this experiment will be invited to the second part. The experiment today is made up of several games and questionnaires. After each game/questionnaire, you will receive new instructions for the next game/questionnaire. In total, the experiment will take approx. one hour. For your participation you will receive a minimum payment of 5 Euro. Depending on your actions during the experiment you can earn more than that. After all questionnaires and games are done, your payoff will be shown on your monitor. You will then be handed a receipt in which you enter your earned payoff as well as your name and address. Please go then to the adjoining room to receive your payment.

Quiz In this quiz, we ask you to answer three questions of differing difficulty. Please try to answer as many of them as possible. You have 5 minutes of time, and you will receive one Euro for each question answered correctly.

1. A bat and a ball cost \$1.10. The bat costs \$1.00 more than the ball. How much does the ball cost?

2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

Questionnaire On the screen before you, you see 10 decision situations. In each of these situations, you have the choice between two options, A or B. Both options contain a lottery with two possible amounts of money you can win, and their respective probabilities.

Example: In the first decision situation (the first row on your screen), Option A pays 2 € (with a probability of 10%) or 1.60 € (with a probability of 90%). Option B on the other hand pays 3.95 € (with a probability of 10%) or 0.10 € (with a probability of 90%).

The following 9 decision situations are very similar, and only the probabilities with which you can win the prizes change. Please choose between Option A and B by moving the scroll bar either to the left or to the right. Also note that you are restricted in the following way; after the first line in which you choose Option B over A, you have to choose Option B in all following lines. Your earnings from this lottery will be paid in cash after the end of the experiment. Which of the 10 decision situations will be paid is determined randomly by the computer. Depending on whether you chose Option A or B in this randomly chosen situation, either Lottery A or B will be played. Then a random number generator determines the amount that you win (of course with the stated probabilities).

Game 1 In this game you choose a number between 0 and 100 (both included). The other participants also choose a number between 0 and 100. Your payoff depends on how far away your number is from 2/3 of the average of all chosen numbers (yours included). The closer your number to 2/3 of the average, the higher your payoff. Your payoff is calculated as follows:

$$\text{Payoff (in Euro)} = 1 - 0.05 * |\text{your number} - 2/3 * \text{average}|$$

In words: your payoff (in Euro) is calculated as 1 minus 0.05 times the absolute difference between your number and two thirds of the average of all chosen numbers. Since the absolute difference (as indicated by the absolute value bars "|") is used, it does not matter whether your number is above or below two thirds of the average. Only the absolute distance is used to calculate your payoff. The smaller the difference, i.e. the distance of your number to two thirds of the average of the chosen numbers, the higher your payoff. Please note that your payoff cannot be negative. If your payoff, as calculated with the above formula, turns out to be negative, then you will receive 0 Euro. Since the payoff for the other participants is calculated in the same way, they too have an incentive to choose a number that is as close as possible to 2/3 of the average. You are playing this game with all other participants that are presently in the room. You have 90 seconds to enter your number.

Game 2 This game is very similar to the game played before. Again, it is your goal to choose numbers that are as close as possible to 2/3 of the average. This time, however, you will be playing against yourself. You are playing the same game as before, only this time the only player with whom you play, is yourself. This time you will be asked to enter two numbers between 0 and 100 (both included), and your payoff will depend on how close your numbers are to two thirds of the average of the two numbers that you chose. Since you play against yourself, the average number equals your first chosen number plus your second chosen number, divided by two. This time you will be paid twice, once for each number you choose. The payoff for your first chosen number is calculated as:

$$\text{Payoff (in Euro)} = 0.5 - 0.05|\text{Number1} - 2/3 * [((\text{Number1} + \text{Number2}))/2]|,$$

where Number1 is the first chosen number, and Number2 is the second chosen number. Your payoff for your second chosen number is calculated as:

$$\text{Payoff (in Euro)} = 0.5 - 0.05|\text{Number2} - 2/3 * [((\text{Number1} + \text{Number2}))/2]|,$$

You have 90 seconds to enter both numbers.

Game 3 (Race to 60) In this game, you play several repetitions of the game "Race to 60". Your goal is to win this game as often as possible against the computer. In this game you and the computer alternately choose numbers between 1 and 10. The numbers are added up, and whoever chooses the number that pushes the sum of numbers to or above 60 wins the game. In detail, the game works as follows: You start the game against the computer, by choosing a number between 1 and 10 (both included). Then the game follows these steps: The computer enters a number between 1 and 10. This number is added to your number. The sum of all chosen numbers so far is shown on the screen. If the sum is smaller than 60, you again enter a number between 1 and 10, which in turn will be added to all numbers chosen so far by you and the computer. This sequence is repeated until the sum of all numbers is above or equal to 60. Whoever (i.e. you or the computer) chooses the number that adds up to a sum equal or above 60 wins the game. You will be playing this game 12 times against the computer. For each of these games you have 90 seconds of time. For each game won, you receive 0,5 Euro.

B Description of Cognitive Tasks & Risk Preference Elicitation

B.1 CRT

The CRT tests the ability to overrule an initial intuitive response that is incorrect, and to engage in further reflection to find the correct answer. The test consists of three algebraic questions, which are:

1. A bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost?
2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?
3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

B.2 Guessing Game

Subjects play a total of two guessing games, one against the other subjects in the room and one against themselves. In these games, subjects are asked to state a number between 0 and 100, both inclusive. Subjects are paid according to the absolute distance of their guess to two thirds of the average guess. In the first case, this average is calculated as the average of all guesses of the subjects in the room. In the second case, where subjects play against themselves, we ask them to state two numbers. The average guess in this game is calculated as the average of these two numbers. In both cases, iterative deletion of dominated strategies leads to "0" as the equilibrium choice. The payoff function for the guessing game against others is given by:

$$\pi_{OS} = 1 - 0.05 \left| x - \frac{2}{3}\bar{x} \right|,$$

where x is the stated number, and $\bar{x}$ is the average number stated by all other subjects. In the guessing game against oneself, each player plays the game for two "selves". Hence she has two payoff functions:

$$\pi_{S1} = 0.5 - 0.05 \left| y - \frac{2}{3} \frac{y+z}{2} \right|,$$

$$\pi_{S2} = 0.5 - 0.05 \left| z - \frac{2}{3} \frac{y+z}{2} \right|,$$

where y and z denote the first and second number stated by subjects respectively. We decided to make payoffs based on absolute distance because this rule invokes the same equilibrium as the standard winner takes all scheme, while allowing to pay every subject for their choice. Note that this kind of payment scheme is common in the guessing game literature (e.g., Costa-Gomes and Crawford (2006), Güth et al. (2002). Moreover, Kocher and Sutter (2006) argue that continuous payoff schemes "resemble financial decision-making much more than the basic winner takes-all scheme with a boundary equilibrium").

B.3 Race to 60

In the Race to 60, the participants play a game against the computer in which both sequentially pick values between 1 and 10, which are added up into a "common pool". The goal of the game is to to be the one to push this common pool to or above 60. By picking numbers such that the common pool adds up to the sequence : $[5, 16, 27, 38, 49, 60]$ the first mover can always win this game. So, to always win the game the first mover should start by picking 5, then, after the computer has made its choice, pick whichever number pushes the common pool to 16, then to 27, then to 38, 49, and finally to 60 (or above). This game is used to measure the levels of backward induction subjects can make, by observing when they enter (and stay on) the optimal path.

Subjects played this game 12 times against a computer whose backward induction ability increased every two rounds. So, subjects started playing two rounds against a computer able to do only one backward induction step (i.e. the computer is able to pick the correct number to add up to 60 if the sum is above 49, otherwise the computer plays a random number). Then subjects played the following two rounds against a computer able to do two steps of backward induction (i.e. adding the numbers to 49 if the current sum is between 39 and 48, and to 60 if the sum is above 49), and so on. Subjects were not aware of this increase in ability of the computer.

We chose this procedure to be able to detect low levels of backward induction, since if the computer had played its best response all the time, we would have never been

Table 2: Price List

	Lottery A				Lottery B			
Line	p	Euro	p	Euro	p	Euro	p	Euro
1	0.1	2.00	0.9	1.60	0.1	3.85	0.9	0.10
2	0.2	2.00	0.8	1.60	0.2	3.85	0.8	0.10
3	0.3	2.00	0.7	1.60	0.3	3.85	0.7	0.10
4	0.4	2.00	0.6	1.60	0.4	3.85	0.6	0.10
5	0.5	2.00	0.5	1.60	0.5	3.85	0.5	0.10
6	0.6	2.00	0.4	1.60	0.6	3.85	0.4	0.10
7	0.7	2.00	0.3	1.60	0.7	3.85	0.3	0.10
8	0.8	2.00	0.2	1.60	0.8	3.85	0.2	0.10
9	0.9	2.00	0.1	1.60	0.9	3.85	0.1	0.10
10	1.0	2.00	0.0	1.60	1.0	3.85	0.0	0.10

able to identify backward induction levels below 6. For example, imagine a subject with less than 6 steps of backward induction ability; this subject will not (most likely!) start out on the optimal path (5) and would be instantly “kicked” off the optimal path by a perfectly backward inducting computer, not allowing us to observe her true level of backward induction.

B.4 Risk Preferences

To elicit risk preferences, we use a standard Holt and Laury price list (Holt and Laury, 2002). Subjects repeatedly choose between two lotteries (A and B), one involving relatively low risk, and one involving relatively high risk (i.e. a higher variance in potential payoffs). Table 2 describes all choices subjects face. In the Table, p represents the probability to win the “Euro” amount right of it in the table. The point at which subjects first prefer Option B over Option A can be used to assess their risk preferences.[23]

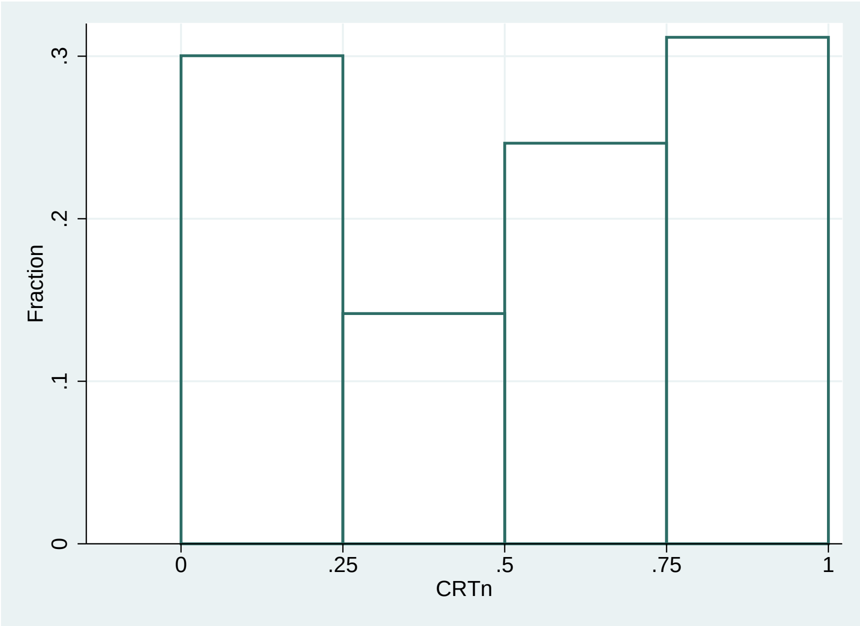
C Index of Cognitive Sophistication

The index S_i used to rank participants is constructed according to the following weighted average:

$$S_i = 1/3 * \text{CRTn}_i + 1/3 * \text{Guessingn}_i + 1/3 * \text{Racen}_i$$

[23]The software allowed to switch only once from Option A to B. See Appendix A for more details.

Figure 4: CRTn Distribution



C.1 CRTn

CRTn is the normalized result of the number of correct answers for the CRT questions (if all three answers are correct, CRTn=1, if only two are correct, CRTn=2/3, if only one, CRTn=1/3, and CRTn=0 if no correct answers.

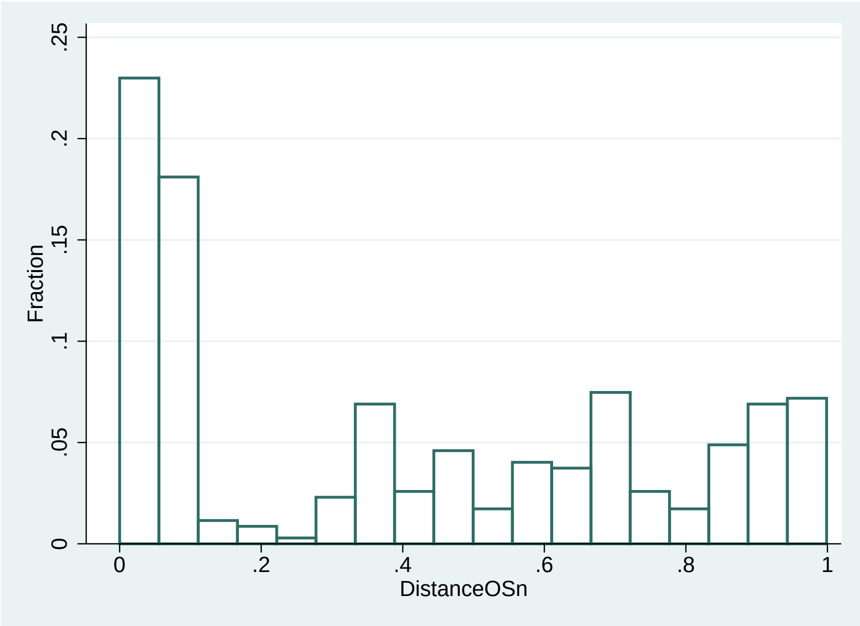
C.2 Guessingn

The measure Guessingn combines the outcomes of the Guessing Game and Guessing Game Against Oneself and is defined as $Guessingn = 0.5 * \text{DistanceOSn}_i + 0.5 * \text{Selfn}_i$, where:

DistanceOSn The variable DistanceOSn is our measure of how well a subject performed in the guessing game. To construct it we take the following steps. First, we separate the choices of all subjects (ChoiceOS_i) into two groups: those that played a dominated strategy (i.e. ChoiceOS $>$ 66) and the rest. Those in the former group are assigned a score of zero for their DistanceOSn. We then define our “objective” value, which is 2/3 of the average of choices all chosen numbers in the guessing game across all sessions , which is 25.587. With this, we create a measure called Distance_i as follows:

$$\text{Distance}_i = |(25.587 - \text{ChoiceOS}_i)/(66 - 25.587)|,$$

Figure 5: DistanceOSn Distribution

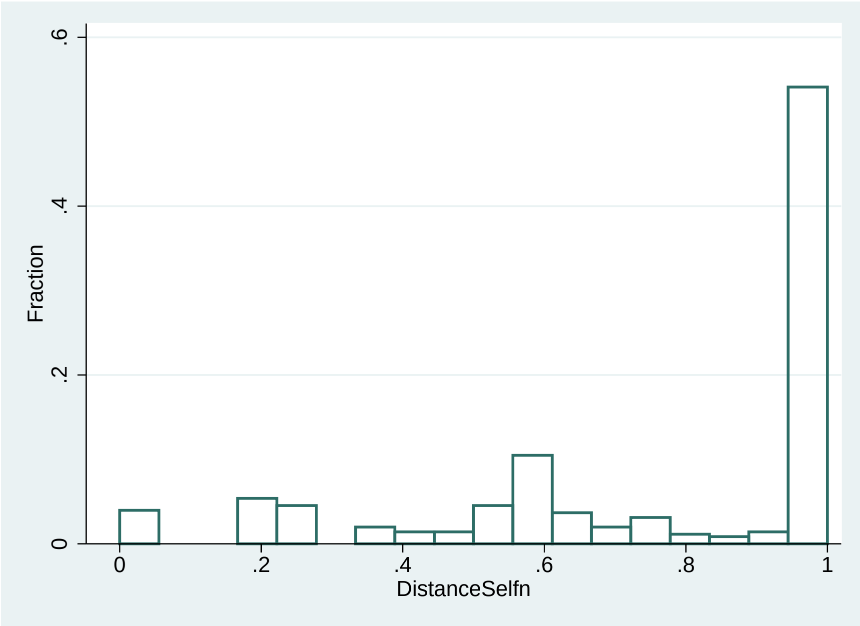


if $\text{ChoiceOS}_i \leq 66$. This allows us to rank all subjects in a range between zero and one, with zero being assigned to those players that played exactly the objective value and one to those subjects that played above 66. In addition, we posit that choosing a number below the objective value indicates a better understanding of the game than choosing a number above it. Accordingly, in our measure of cognitive sophistication for the guessing game we add an extra 50% to the "distance" for any choice above the objective value. This translates into the following equation:

$$\text{DistanceOSn}_i = \max\{0, \begin{cases} 1 - \text{Distance}_i * 1.5 & \text{if ChoiceOS}_i > 25.587 \\ 1 - \text{Distance}_i & \text{if ChoiceOS}_i < 25.587 \end{cases}\}$$

Selfn The measure Selfn, for cognitive sophistication in the "playing against self" game, is again a two-step procedure. We posit that the game has two dimensions of understanding: the first dimension is realizing that the numbers picked should always be close together (in fact they should be the same); the second dimension is realizing that there is a unique correct answer (zero for both choices). In order to evaluate both dimensions we first measure the distance of each choice (ProximitySelf^1 and ProximitySelf^2) to 2/3 of the average (AvgSelf) of both:

Figure 6: Selfn Distribution



$$\text{ProximitySelf}_i^1 = |\text{Self}_i^1 - 2/3\text{AvgSelf}|$$

$$\text{ProximitySelf}_i^2 = |\text{Self}_i^2 - 2/3\text{AvgSelf}|$$

,

where Self_i^1 is the first number chosen and Self_i^2 is the second number chosen by subject i. We then create a normalized measure for the proximity of both values:

$$\text{NormalizedSelf}_i^a = 1 - (\text{ProximitySelf}_i^1 + \text{ProximitySelf}_i 2)/100$$

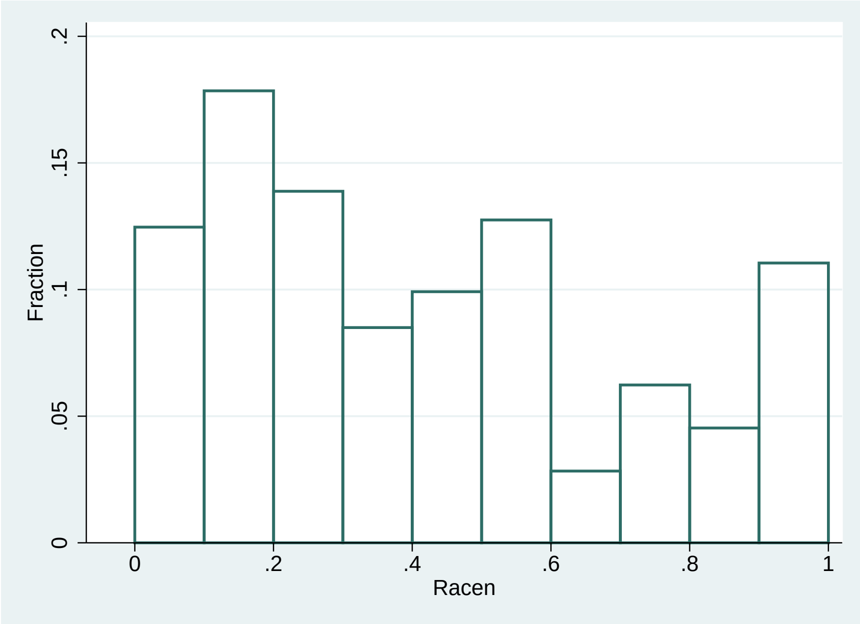
Next we compute the second measure:

$$\text{Normalizedself}_i^b = 1 - (\text{Self}_i^1 + \text{Self}_i^2)/200,$$

which penalizes subjects for picking numbers away from the solution of the game. Using both NormalizedSelf^a and NormalizedSelf^b we create the final measure:

$$\text{Selfn}_i = (\text{NormalizedSelf}_i^a + \text{NormalizedSelf}_i^b)/2$$

Figure 7: Wonn Distribution



C.3 Racen

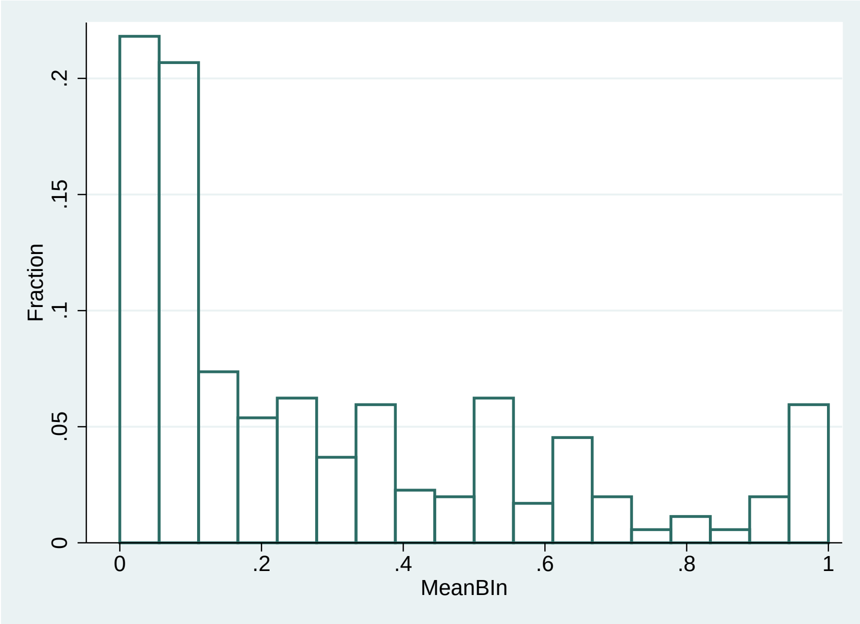
Racen is composed by two measures extracted from the Race to 60 game and is defined as $Racen = 0.5 * \text{Wonn}_i + 0.5 * \text{MeanBIn}_i$, where:

Wonn: This measure is the normalization of the number of rounds won by each subject in the Race to 60 game (Won_i):

$$\text{Wonn}_i = \text{Won}_i/12$$

MeanBIn This measure is the average number of backward induction steps (MeanBIn) that a subject made during the 12 Rounds of Race to 60. Race to 60 has a correct path [5, 16, 27, 38, 49, 60] that allows the first mover to always win the game. The number of backward induction steps is dependent on when a subject enters this optimal path and stays on it. If a subject starts out with a 5, and then stays on the correct path, we say that she has 6 backward induction steps. In this case she has solved the game completely. Consequently, if a subject enters the correct path at, say, 38 she thinks three steps ahead. We then create the measure $MeanBIn$ which is the normalized mean of backward induction steps that a subject has taken across all 12 rounds.

Figure 8: MeanBIn Distribution



$$\text{MeanBI}_i = \sum_{r=1}^{12} \frac{\text{BIsteps}_{ir}}{12}$$

C.4 Distribution of S_i

Finally we present the distribution of S_i in Figure 9. Any subject with a score $S_i > 0.678$ ($S_i < 0.28$) was considered to be of High (Low) Sophistication.

The symmetric weighting of S_i was picked because *a priori* any choice of weights is arbitrary. We felt comfortable with this solution as our measures are highly correlated (see Table 3), pointing towards an S_i that is robust to changes in its weights. In order to confirm this we sort our subjects into High and Low following the "No CRT", "No Guessing", and "No Race" criteria. In each of these cases one of the three main measures was dropped and equal weights were given to the remaining ones. The percentage of subjects that overlapped with our original symmetric measure and the robustness modifications are reported in Table 4. As is clear from the results our index S_i is robust to changes in its weights.

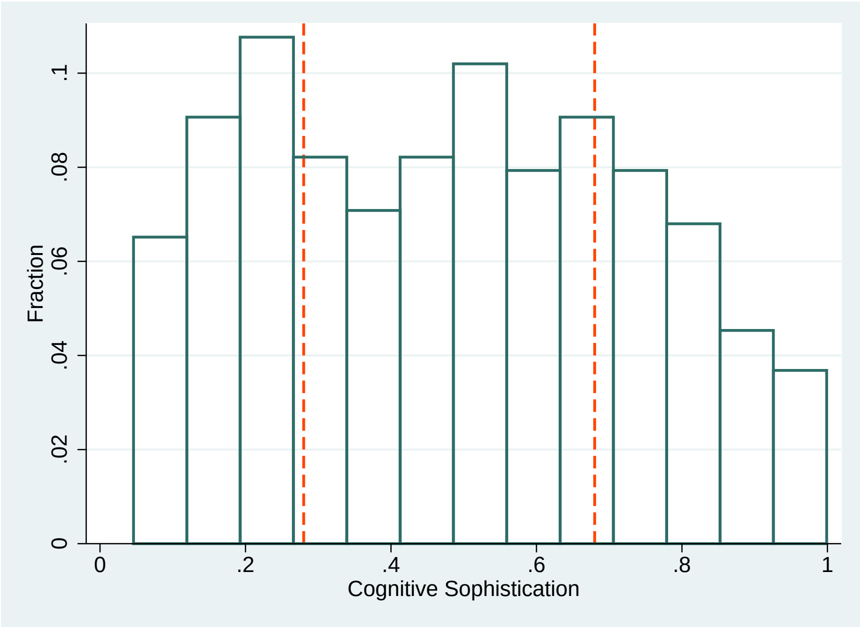


Figure 9: Cognitive Sophistication Measure (S_i) Distribution.

Notes: The red dashed lines mark the separation for Low and High Sophistication subjects.

Table 3: Correlation between measures.

	CRT	Guessing	Race
CRTn	1	-	-
Guessingn	0.422	1	-
Racen	0.477	0.396	1

Table 4: Overlap

	High	Low
No CRT	0.828	0.771
No Guessing	0.828	0.809
No Race	0.716	0.857

Notes: Percentage of overlapping subjects in the High and Low groups. No CRT is constructed by giving half of the weight to Guessing and the other half to Race, No Guessing is gives half the weight to CRTn and half to Race, while No Race gives half the weight to CRTn and half to Guessing.

D Asset Market Experiment Instructions

This is the second part of the experiment. [24]

Overview This is an economic experiment on decisions in markets. In this experiment we generate a market, in which you trade units of a fictitious asset with the other participants of the experiment. The instructions are not complicated, and if you follow them closely and make appropriate decisions, you can earn a considerable amount of money. The money that you earn during the experiment will be paid in cash at the end of the experiment. The experiment consists of 3 rounds. Each round consists of 15 periods (in the following also named trading periods) in which you have the opportunity to trade in the market, i.e. to buy and sell. The currency in which you trade is called "Taler". All transactions in the market will be denoted in this currency. The payoff that you receive will be paid in Euro. You will receive one Euro for every 90 Taler.

Experiment Software and Market You will be trading in one of two markets, each of which consists of 7 participants. Both markets are identical in their functionality and are independent of each other. Your assignment to one of these markets is random, and you will stay in this market for the duration of the experiment. You can make your decisions in the market through the experiment software. A screenshot of this software can be found on the next page. In every trading period you can buy and sell units of an asset (called "share" from now on). In the top left corner of the screen you can see how many Taler and shares you have at every moment (see screenshot). In case you want to buy shares, you can issue a buy order. A buy order contains the number of shares that you want to buy and the highest price that you are willing to pay per share. In case you want to sell shares, you can issue a sell order. Similar to the buy order, a sell order contains the number of shares that you want to sell as well as the lowest price that you are willing to accept for each share. The price at which you want to buy shares has to be lower than the price at which you want to sell shares. All prices refer to prices of a single share.

[24] In the instructions for the "shared-knowledge" High Sophistication treatments the following sentences were added at this point: "Based on your answers to the questionnaires and your actions in the games of the first part of the experiment, we have calculated a "performance score" that reflects the quality of your decisions. You have been invited to this experiment today because your score was above average."

The experiment software combines the buy and sell orders of all participants and determines the trading price, at which shares are bought and sold. This price is determined so that the number of shares with sell order prices at or below this price is equal to the number of shares with buy order prices at or above this price. All participants who submit buy orders above the trading price will buy shares, and those that have sell orders below the trading price will sell shares. Example of how the market works: Suppose there are four traders in the market and:

- Trader 1 submits a buy order for one share at the price of 60 Taler.
- Trader 2 submits a buy order for one share at the price of 20 Taler.
- Trader 3 submits a sell order for one share at the price of 10 Taler.
- Trader 4 submits a sell order for one share at the price of 40 Taler.

At any price above 40, there are more units offered for sale than units for purchase. At any price below 20, there are more units offered for purchase than for sale. At any price between 21 and 39, there is an equal number of units offered for purchase and sale. The trading price is the lowest price at which there is an equal number of units offered for purchase and sale. In this case, the trading price is 21 Taler. Trader 1 buys one share from Trader 3 at the price of 21 Taler. Trader 2 buys no shares, because her buy order price is below the trading price. Trader 4 does not sell any shares, because her sell order price is above the trading price.

Specific Instructions for this Experiment This experiment consist of 3 independent rounds, each consisting of 15 trading periods. In every period you can trade in the market, according to the rules stated above. At the start of each round, you receive an endowment of Taler and shares. This endowment does not have to be the same for every participant. As mentioned, you can see the amount of shares and Taler that you own on the top left corner of your screen. Shares have a life of 15 periods. The shares that you have purchased in one period are at your disposal at the next period. If you happen to own 5 shares at the end of period 1, you own the same 5 shares at the beginning of period 2. For every share you own, you receive a dividend at the end of each of the 15 periods. At the end of each period, including period 15, each share pays a dividend of either 0, 4, 14, or

Accuracy	Your Earnings
Within 10% of actual price	5 Taler
Within 25% of actual price	2 Taler
Within 50% of actual price	1 Taler

30 Taler, with equal probability. This means that the average dividend is 12 Taler. The dividend is added automatically to your Taler account at the end of each period. After the dividend of period 15 has been paid, the market closes and you will not receive any further dividends for the shares that you own. After this round is finished, a new round of 15 period starts, in which you can buy and sell shares. Since all rounds are independent, shares and Taler from the previous period are not at your disposal anymore. Instead, you receive the same endowment of shares and Taler that you had at the beginning of round one. The experiment consists of 3 rounds with 15 periods each.

Average Holding Value The table "Average Holding Value", which is attached to these instructions, is meant to facilitate your choices. The table shows how much dividend a share pays on average, if you hold it from the current period until the last period, i.e. period 15 of this round. The first column indicates the current period. The second column gives the average earnings of a share if it is held from this period until the end of the round. These earnings are calculated as the average dividend, 12, multiplied by the number of remaining periods, including the current period.

Predictions In addition to the money you earn by trading shares, you can earn additional money by predicting the trading prices. In every period, before you can trade shares, you will be asked to predict the trading prices in all future periods. You will indicate your forecasts in a screen that looks exactly like the screen in front of you. The cells correspond to the periods for which you have to make a forecast. Each cell is labeled with the period for which you are asked to make a forecast. The amount of Taler you can earn with your forecasts is calculated as follows.

You can earn money on each and every forecast. The accuracy is calculated separately for each forecast. For example, in period 2, your forecast from period 1 and your forecast from period 2 are evaluated separately. If both forecast are within 10% of the actual price, you earn 2*5=10 Taler. If one is within 10% of the actual price and one is within 25% of the actual price, but not within 10%, you earn 5 Taler + 2 Taler = 7 Taler.

Your Payoff For your participation you receive a fixed payment of 5 Euro and a payment that depends on your actions. The latter part of the payment is calculated for each round, as the amount of Taler that you have at the end of period 15, after the last dividend has been paid, plus the amount of Taler you receive for your forecasts. Your payoff for each round is calculated as:

The amount of Taler you have at the beginning of period 1
+ the dividends you receive
+ Taler that you receive from selling shares
– Taler that you spend on shares
+ Taler that you earn with your forecasts.

The total payment that you receive in Euro consists of the sum of Taler you earn in all three rounds, multiplied by 1/90, plus the fixed payment of 5 Euro.

Period	Average Holding Value
1	180
2	168
3	156
4	144
5	132
6	120
7	108
8	96
9	84
10	72
11	60
12	48
13	36
14	24
15	12

Table 5: Bubble Measures

Measure	Treatment	Round 1	Round 2	Round 3	Total
mean RAD	high	0.077	0.074	0.101	0.084
mean RAD	low	0.708	0.308	0.277	0.431
pvalue		0.002	0	0.018	< 0.001
mean RD	high	-0.004	0.065	0.095	0.052
mean RD	low	0.105	0.092	0.031	0.148
pvalue		0.955	0.272	0.388	0.529
mean PD	high	0.036	0.069	0.098	0.068
mean PD	low	0.406	0.2	0.154	0.253
pvalue		0	0.066	0.955	0.008
mean DUR	high	3.556	3.889	3	
mean DUR	low	8	5.333	4.833	
pvalue		0.003	0.08	0.112	
mean AMP	high	0.299	0.17	0.162	
mean AMP	low	1.433	1.327	1.199	
pvalue		0.002	< 0.001	< 0.001	

Notes: P-values are calculated using the Mann-Whitney U-test, the null hypothesis being that the distributions of the measures in the treatments high and low are identical.

E Second and Third Round Results

In this section we report the results for the second and third round of our market sessions, both for High and Low treatments.

In Figure 10 we present the evolution of prices in the Low Sophistication sessions. As usually found in the literature, prices appear to converge (slowly) to the fundamental value. We present standard bubble measures for all rounds in Table 5. Indeed, all bubble measures appear to decrease over rounds, indicating convergence to the fundamental value. The price dynamics for the High Sophistication treatment are presented in Figure 11. There seems to be a slight increase in deviations from the fundamental value in the last round, according to the bubble measures in Table 1 This deviation appears to be concentrated in the late periods of two of the nine sessions, which might very well indicate, as we mentioned in footnote 18, that some subjects were becoming bored from the third repetition of an uneventful market.

In this appendix, we also document the price predictions of our two treatments in rounds 2 and 3. (Figure 12 and 13). Two things are noteworthy in the Low Sophistication treatment; first, subjects in the second round predict a bubble and crash pattern, which is akin to what Haruvy et al. (2007) observe in their experiment. Second, in the third round subjects seem to have improved their understanding of the asset price dynamics and

Figure 10: All Rounds Low Sophistication Markets

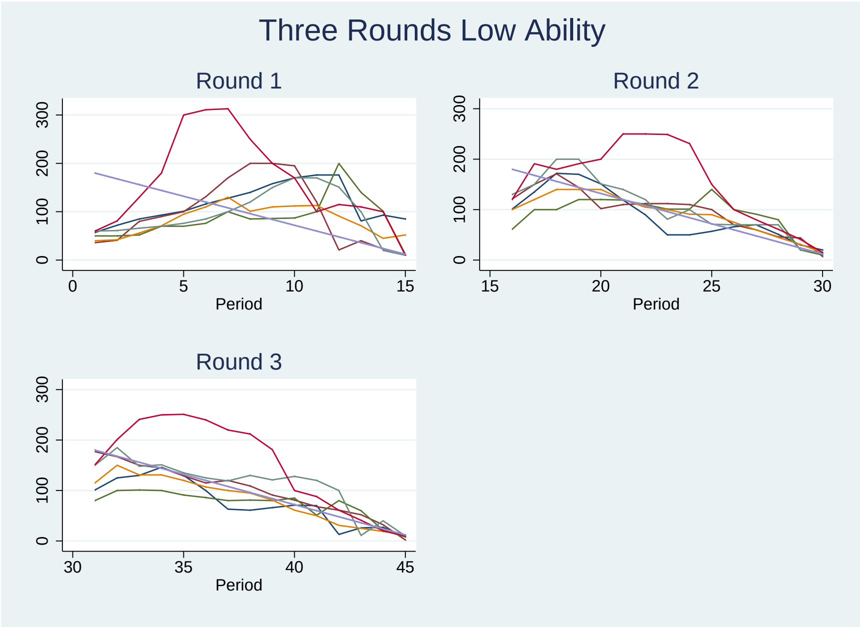
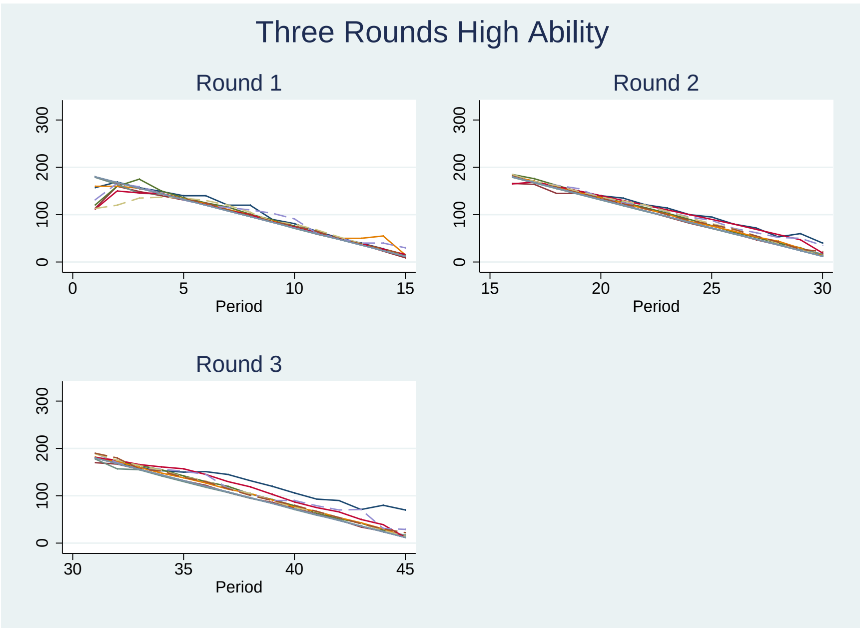


Figure 11: All Rounds High Sophistication Markets



predict a falling trajectory of prices instead of the perennial inverted-U shape of round 1 and 2. As in round 1, the price predictions in the High Sophistication treatment track the fundamental value almost perfectly.

Figure 12: Price Predictions for Round 2 and 3 of the Low Treatment

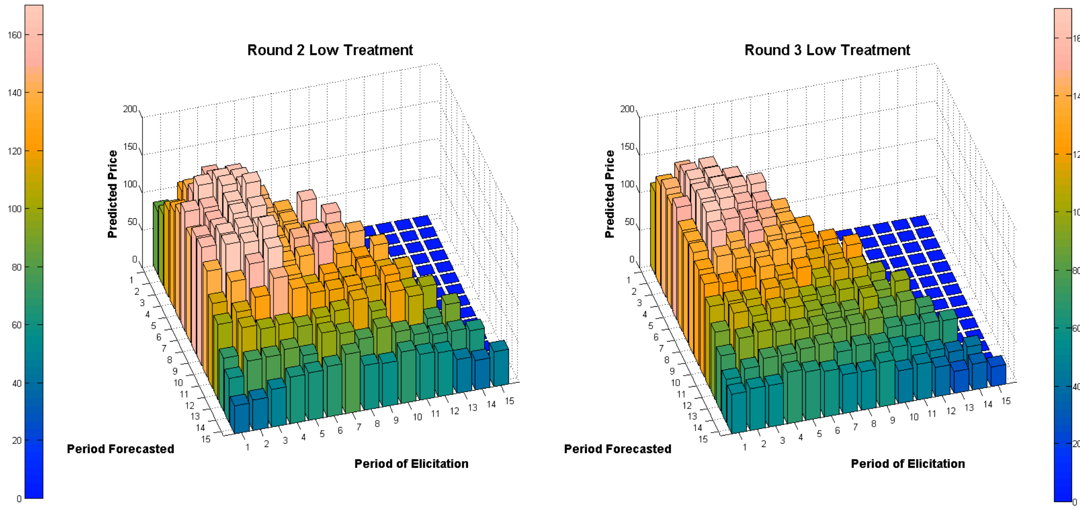


Figure 13: Price Predictions for Round 2 and 3 of the High Treatment