



UvA-DARE (Digital Academic Repository)

Do investors trade too much?

A laboratory experiment

da Gama Batista, João; Massaro, D.; Bouchaud, J.P.; Challet, D.; Hommes, C.

DOI

[10.1016/j.jebo.2017.05.013](https://doi.org/10.1016/j.jebo.2017.05.013)

Publication date

2017

Document Version

Final published version

Published in

Journal of Economic Behavior & Organization

License

Article 25fa Dutch Copyright Act

[Link to publication](#)

Citation for published version (APA):

da Gama Batista, J., Massaro, D., Bouchaud, J. P., Challet, D., & Hommes, C. (2017). Do investors trade too much? A laboratory experiment. *Journal of Economic Behavior & Organization*, 140, 18-34. <https://doi.org/10.1016/j.jebo.2017.05.013>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Do investors trade too much? A laboratory experiment



João da Gama Batista^a, Domenico Massaro^{b,*}, Jean-Philippe Bouchaud^d,
Damien Challet^{a,e}, Cars Hommes^c

^a *Chaire de Finance Quantitative, Laboratoire de Mathématiques Appliquées aux Systèmes, CentraleSupélec, France*

^b *Università Cattolica del Sacro Cuore & Complexity Lab in Economics, Milan, Italy*

^c *CeNDEF, University of Amsterdam & Tinbergen Institute, The Netherlands*

^d *Capital Fund Management, Paris, France*

^e *Encelade Capital SA, Lausanne, Switzerland*

ARTICLE INFO

Article history:

Received 13 May 2016

Received in revised form 3 May 2017

Accepted 15 May 2017

Available online 22 May 2017

JEL classification:

C91

C92

D84

G11

G12

Keywords:

Experimental asset markets

Trading volumes

Crashes

Expectations

Risk attitude

ABSTRACT

We run an experiment to investigate the emergence of excess and synchronised trading activity leading to market crashes. Although the environment clearly favours a buy-and-hold strategy, we observe that subjects trade too much, which is detrimental to their wealth given the implemented market impact (known to them). We find that preference for risk leads to higher activity rates and that price expectations are fully consistent with subjects' actions. In particular, trading subjects try to make profits by playing a *buy low, sell high* strategy. Finally, we do not detect crashes driven by collective panic, but rather a weak but significant synchronisation of buy activity.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Financial bubbles and crises are potent reminders of how far investors' behaviour may deviate from perfect rationality. Many behavioural biases of individual investors are now well documented, such as the propensity for trend following or extrapolative expectations (Greenwood and Shleifer, 2014), herding behavior (Cipriani and Guarino, 2014), the disposition effect (Grinblatt and Keloharju, 2001), home bias (Solnik and Zuo, 2012), and over and underreaction to news (Barber and Odean, 2008); see Barberis and Thaler (2003) and Barber and Odean (2013) for comprehensive overviews.

A well established fact about individual trading behaviour which is in stark contrast with the predictions of rational models is the tendency of individual investors to trade too much (Odean, 1999). Many investors trade actively, speculatively, and to their detriment. Odean (1999), Barber and Odean (2000), Odean and Barber (2011) and Barber et al. (2009) among others show that the average return of individual investors is well below the return of standard benchmarks and that the

* Corresponding author at: Università Cattolica del Sacro Cuore, Largo Gemelli 1, 20123 Milano, Italy.
E-mail address: domenico.massaro@unicatt.it (D. Massaro).

more active traders usually perform worse on average. In other words, these investors would do a lot better if they traded less. Moreover, as noted in [Barber and Odean \(2013\)](#), individual investors make systematic, not random, buying and selling decisions.

Based on the empirical evidence mentioned above, the goal of this paper is to study the emergence of excess trading in an experimental financial market where trading is clearly detrimental for investors' wealth. One advantage of using a controlled laboratory environment is that it allows us to directly measure individual characteristics and relate them in a systematic way to trading activity. We measure individual risk attitudes and elicit asset price forecasts, and link them to observed trading activity. Moreover, we aim to gain a deeper understanding not only of why agents trade, but also of how correlated is their activity. In particular, we want to study whether synchronisation of trading activity leads to unstable market behaviour, such as crashes driven by panic, herding and cascade effects.

Market laboratory experiments now have a rather long history. The most influential paradigm for multi-period laboratory asset markets was developed in [Smith et al. \(1988\)](#). The asset traded in their experiment has a known finite life span and pays a stochastic dividend at the end of each period. The fundamental value of the asset falls deterministically over time, and lacking a terminal value the asset expires worthless. A salient result is that asset prices in the experimental markets follow a "bubble and crash" pattern which is similar to speculative bubbles observed in real world markets.¹ This seminal work has spawned a large number of replications and follow-ups, see [Palan \(2013\)](#) for an extensive overview.

The present study belongs to the above tradition of experimental markets but implements a different market mechanism. Instead of trading in a continuous double auction, subjects can submit buying and selling orders, executed by a market maker, for a fictitious asset that increases in value at a known average rate and has an indefinite horizon. One particularly interesting feature of our laboratory market is that we model and implement price impact, namely the fact that the very action of agents modifies the price trajectory. This is now believed to be a crucial aspect of real financial markets, which may lead to feedback loops and market instabilities ([Bouchaud et al., 2009](#); [Bouchaud, 2011](#); [Cont and Wagalath, 2014](#); [Caccioli et al., 2014](#)). What is of particular interest in our experiment is that excess trading significantly impacts the price trajectory and is strongly detrimental to the wealth of our economic agents. In other words, unwarranted individual decisions can lead to a substantial loss of collective welfare, when mediated by the mechanics of financial markets.

Our work relates to the literature on experimental asset markets that investigates how excess trading and mispricing is affected by the characteristics of market participants. In particular, we relate our results on trading activity to market experience, individual risk preferences and price forecasts.² Previous experimental studies show that repeated participation in identical markets plays an important role in eliminating bubbles and crashes ([Smith et al., 1988](#); [King, 1991](#); [Haruvy et al., 2007](#)). Moreover, it is often argued that non-zero trade volumes are observed in experimental asset markets due to the heterogeneity in risk preferences. Previous empirical studies show that individual risk attitude could have an impact on trading in asset markets. For example, [Robin et al. \(2012\)](#) and [Fellner and Maciejovsky \(2007\)](#) find that risk-aversion leads to smaller bubbles and less trade in asset markets. Moreover, [Keller and Siegrist \(2006\)](#) did a mail survey and found that financial risk tolerance is a predictor for the willingness to engage in asset markets. In the light of the aforementioned empirical evidence, we measure subjects' risk aversion using a standard [Holt and Laury \(2002\)](#) procedure and link it to individuals' trading activity. Finally, [Palan \(2013\)](#) highlights the importance of investigating the dynamics of expectations regarding future prices in multi-period asset markets. We also elicit individual price expectations in order to better understand what leads to investors' decision of engaging actively in trading activity as well as being inactive in the market. Since we collect data on both individual trading decisions and price expectations, we are able to give a consistent picture of activity and inactivity as a consequence of price return expectations. See also [Smith et al. \(1988\)](#), [Hommes et al. \(2005, 2008\)](#) and [Haruvy et al. \(2007\)](#) among others for studies on the role of expectations in generating bubbles and crashes in experimental asset markets. Another personality trait often linked to individual investment choices is overconfidence. Theoretical models predict that overconfident investors trade too much and to their detriment ([Gervais and Odean, 2001](#)). [Odean \(1999\)](#) links high turnover rates and poor performance of individual investors to overconfidence, while [Odean and Barber \(2011\)](#) show that men, who are more prone to overconfidence than women, trade more than women. Finally, [Michailova and Schmidt \(2016\)](#) present experimental evidence of the positive association between overconfidence and individual engagement in trading activity.

Our findings can be summarised as follows. Although the market environment clearly favours a buy-and-hold strategy (see below), we observe that our subjects engage in excessive trading activity, which is both individually and collectively detrimental, since the negative impact of sellers reduces the price of our artificial asset. When the experiment is immediately repeated with the same subjects, we see a significant improvement of the collective performance, which is however still substantially lower than the (optimal) buy-and-hold strategy.

¹ In fact, [Smith \(2010\)](#) blames a failure of backward induction for the existence of bubbles in these simple experimental markets and he suggests that "price bubbles were a consequence of . . . homegrown expectations of prices rising". [Oechssler \(2010\)](#) shares this view and argues that "backward induction is only useful when there is a finite number of periods which most asset markets do not have. Subjects are told that they trade assets on a market so they probably expect to see something similar to what they see on real markets: stochastic processes with increasing or at least constant trend in most cases." See [Noussair and Powell \(2010\)](#), [Giusti et al. \(2012\)](#), [Breaban and Noussair \(2014\)](#) and [Stöckl et al. \(2015\)](#) for previous experimental studies on markets with (partly) increasing fundamental values.

² See [Powell and Shestakova \(2016\)](#) for a recent overview of the literature on experimental asset markets connecting market outcomes to the structure of experimental markets, properties of the traded asset and trader characteristics.

Moreover, although our subjects are physically separated and cannot communicate, we have seen that a significant amount of synchronisation takes place in the decision process that can therefore only be mediated by the price trajectory itself. This resonates with what happens in real financial markets, where price changes themselves appear to be interpreted as news, leading to self-reflexivity and potentially unstable feedback loops (see for example [Bouchaud et al. \(2009\)](#); [Hommes \(2013\)](#)). In fact, our experimental setting was such that panic and crashes were possible but this did not happen. Although we observed a significant level of synchronisation, no cascades or “fire sales” effects could be detected for our particular choice of parameters.

Consistent with the empirical evidence mentioned above, we find that risk-loving attitudes lead to higher trading activity and this is detrimental to individual wealth.

Finally, subjects seem to have a desire to trade actively, motivated by a willingness to “beat the market”, as revealed by the analysis of individual price expectations. Most subjects engage in trading activity trying to *buy low and sell high*, and the decisions of being inactive and hold cash (shares) depend on whether they expect price returns lower (higher) than average.

The outline of this paper is as follows. Section 2 describes the experimental motivation and set-up. The precise instructions given to our subjects are reported in the Online Appendix. We then summarize our main results in Section 3, which includes a refined statistical analysis of agents’ trading activity in Section 3.1. In Section 3.2 we relate individual risk attitude to activity rate and final wealth, while in Section 3.3 we link price forecasts and trading behaviour. Section 4 concludes.

2. Experimental design

The basic idea of our experiment is to propose to subjects a simple investment “game” where they can use the cash they are given at the beginning of the experiment to invest in a fictitious asset that will – they are told – increase in value at an average rate of $m = 2\%$ per period. The asset “lives” for an indefinite horizon. Specifically, the game may stop randomly at each time step with probability $p = 0.01$. The game is thus expected to last around 100 time steps. Subjects are also informed that random shocks impact the price of the asset in each period. In the absence of trade, price dynamics are described by

$$p_{t+1} = p_t \cdot \exp(m + s\eta_t), \quad (2.1)$$

where $m = 2\%$, η_t is a noise term drawn from a Student’s t -distribution with 3 degrees of freedom and unit variance,³ as commonly observed in financial markets ([Gopikrishnan et al., 1998](#); [Jondeau and Rockinger, 2003](#); [Bouchaud and Potters, 2003](#)), and s is a constant that sets the actual amplitude of the noisy contribution to the evolution of the price and is chosen to be $s = 10\%$. These numbers correspond roughly to the average return and the volatility of a stock index over a quarter. Therefore, in terms of returns and volatility, one time step in our experiment roughly corresponds to three months in a real market, and 100 steps to 25 years.

If an amount w_0 is invested in the asset at time $t = 0$, the wealth of the inactive investor will accrue to

$$w_T = w_0 \exp [mT + \xi s \sqrt{T}] \quad (2.2)$$

at time T , where the second term of the exponential implies random fluctuations of root mean square (RMS) s per period with ξ defined as a noise term with zero mean and unit variance. The numerical value of the term in the exponential is therefore equal to $2 \pm \xi$ for $T = 100$, leading to a substantial most probable profit of $e^2 - 1 \approx 640\%$. As we shall show below, the *fully rational* decision in the presence of risk-neutral agents is to buy and hold the asset until the game ends; for the students participating in the experiment, the most probable gain would represent roughly EUR 160, a very significant reward for spending two hours in the lab. In other words, the financial motivation to “do the right thing” is voluntarily strong.

In order to make the experiment more interesting, and trading even more unfavourable, the asset price trajectory is made to react to the subjects decisions, in a way that mimics *market impact* in real financial markets. The idea is that while a buying trade pushes the price up, a selling trade pushes the price down (for a short review on price impact, see [Bouchaud \(2010\)](#)). As an agent submits a (large) buying or selling order at time t , the price p_{t+1} at which the transaction is going to be fully executed is (a) not known to him at time t and (b) adversely impacted by the very order that is executed. It is made very clear to the subjects that their transaction orders will be executed at the impacted price, meaning that the impact will amount for them as a cost. This should therefore be a strong incentive not to trade.

³ Mathematically, the average of the exponential of a Student distribution is infinite, because of rare, but extreme values of η_t . In order not to have to deal with this spurious problem, we impose a cut-off beyond $|\eta| = 10$, with no material influence on the following discussion.

The price updating rule described in Eq. (2.1) is easily modified to include price impact and now reads⁴

$$p_{t+1} = p_t \cdot \exp(m + s\eta_t + I_t). \quad (2.3)$$

The term I_t is the price impact caused by all the orders submitted at time t , which we model as:

$$I_t = \frac{N_t B_t - S_t}{N B_t + S_t}, \quad (2.4)$$

where N_t is the number of subjects who submitted an order at time t and N is the total number of subjects in a given market session (in other words, the “depth” of our market). B_t and S_t , in currency units, are the total amount of buying orders and selling orders, respectively. Note that for a single buying (or selling) order, the impact is given by $1/N$, that is around 3% for a market with 30 participants (and less if the market involved more participants, as is reasonable). On the other hand, if all agents decide to buy (or sell) simultaneously, the ensuing impact factor would be 100%.

With the introduction of market impact, subjects have to guess if the observed price fluctuations are due to “natural” fluctuations, namely to the noise term they are warned about at the beginning of the game, or if they are due to the action of their fellow subjects. This represents an important difference with the design of Smith et al. (1988), where subjects can in principle disentangle the fundamental component and the trade component of the asset price, which was meant to provide a potentially destabilizing channel, where mild sell-offs could spiral into panic and crashes.

We are therefore interested in studying whether excess trading emerges in a market environment which clearly penalises trading, and whether synchronisation of trading activity, such as avalanches of selling orders triggered by panic, results in big market crashes.

2.1. Implementation

The experiment was programmed in Java and it was conducted at the CREED laboratory at the University of Amsterdam in May 2014. In the beginning of the experiment each subject is randomly assigned a computer in the laboratory and physical barriers guarantee that there is no communication between subjects during the experiment. The experiment was conducted with 201 subjects, divided in 9 experimental markets with a median number of participants per market of 24 subjects. The smallest market includes 15 subjects while the largest includes 29 subjects. The experiment lasted about two hours and before starting the experiment subjects had to answer a final quiz to make sure they understood the rules of the game (see Online Appendix). At the beginning of the experiment each subject is endowed with 100 francs. Each period lasts 20 s, during which subjects have to decide whether they want to hold cash or shares in the next period.⁵ If subjects did not make any decision within the limit of 20 s, their current market position would simply carry over to the next period. If they have cash at period t , they can decide either to use it all to buy shares or to stay out of the market at period $t + 1$. Conversely, if they have shares at period t , they have to choose between selling them all for cash or staying in the market at period $t + 1$. In order to simplify the market environment and compare experimental results with theoretical benchmarks, fractional orders are not allowed, therefore each player is either in the market or out of the market at all times.

In order to avoid possible “demand effects”, namely trading volumes in the markets related to the fact that participation in the asset market is the only activity available for subjects (Lei et al., 2001), we introduce a forecasting game in which subjects are asked to forecast the price of the asset and rewarded according to the precision of their forecast.⁶ In order to minimize the possibility of subjects trying to strategically influence prices in the direction of their prediction, we implement a forecast prize of 0.10 EUR per period, which is a monetary incentive several order of magnitudes lower than potential market earnings.⁷

The information visualised on subjects’ screens include a chart displaying the asset price evolution together with a table reporting time series of prices, returns, individual holdings of cash and shares and price forecasts. Figure A2 in the Online Appendix shows a screenshot of the experiment.

⁴ We interpret I_t as the permanent component of the price impact, which we assume to be non-zero, meaning that even random trades do affect the long term trajectory of the price. This is clearly at odds with the efficient market theory, within which the price impact of uninformed trades should be zero. We tend to believe that the anchoring to the “fundamental price” in real markets is very weak and is only relevant on very long time scales (Bondt and Thaler, 1985), so that the assumption made here is of relevance for understanding financial markets. See the discussion in (Bouchaud et al., 2009; Donier and Bouchaud, 2015a,b). We also remark that there is clear evidence that random trades impact real markets, though this impact decays slowly. It is however hard to empirically assess whether the long term value of such impact is permanent or transient. We believe that implementing an impact that slowly decays over time in our setting would not change much, as it would still lead to considerable trading costs for active players, and it could generate feedback loops in the same conditions as with permanent impact. The only difference would be that excess trading would be potentially less detrimental to virtuous traders.

⁵ In a pilot session we implemented a duration of 40 s per period. Decision times were, however, well below the threshold of 20 s, so we reduced the duration of each time step.

⁶ Moreover, we remark that practically speaking, the choice of being active or inactive in the market involved the same type of action, namely a click on the correspondent radio button (see Online Appendix for details).

⁷ Given the high number of market participants, the individual impact on price is limited. Huck et al. (2004) show that a number of subjects $n \geq 4$ are enough to eliminate this sort of strategic reasoning and ensure convergence to either Cournot or Walrasian equilibrium.

In order to mitigate behaviour bias towards the end of the experiment, we implement an indefinite horizon (see [Crockett and Duffy \(2013\)](#) for an example in artificial asset markets). In each period of the experiment there is a known constant continuation probability $1 - p = 0.99$. For practical reasons we selected in advance exponential-distributed end-times to ensure that sessions would not stop too early or last too long. Moreover, each subject participates in two consecutive market sessions, after a short initial practice session to familiarise with the software, so that learning mechanisms can be studied. In order to minimise changes in the market environments and facilitate subjects' learning we use the same noise realisations in the first and the second market sessions.

When a market session is terminated, subjects' final wealth is defined as follows: if they are holding cash, their amounts of francs will determine their wealth; if they are holding shares, their wealth is defined as the market value of their shares, given by their amounts of shares times the market price at the end of the sequence. In fact, at the end of the market sequence shares are liquidated at the market price without any price impact. At the end of the experiment subjects' net market earnings are computed as their end-of-sequence balance minus their initial endowment. If net profits are negative at the end of a sequence, earnings from market activity for that sequence are set to zero.⁸ At the end of the experiment, each subject rolls a dice to determine which of the two market sessions will be used to calculate his take-home profits (which also include the earnings from the forecasting game). The exchange rate is 100 francs to EUR 25.

Finally, we asked subjects to participate in a second experiment involving the [Holt and Laury \(2002\)](#) paired lottery choice instrument. This second experiment occurred after the asset market experiment had concluded and was not announced in advance, to minimize any influence on decisions in the asset market experiment.

The complete experimental instructions can be found in the Online Appendix.

2.2. Theoretical benchmark

We now devote attention to the theoretical rational benchmark for this experiment and derive the equilibrium of the game under the assumption of full rationality and common knowledge of rationality. Let the wealth of the agent be w_t at time t . If we assume that the session ends at $t = t_F$, fully rational agents have two possible strategies. The first one is to stay out of the market for $t > 1$ and hold cash until $t = t_F$ which yields an expected final wealth $\mathbb{E}(w_{t_F}) = w_0$.

The second strategy consists in being fully invested in the market at $t = 1$ and hold the shares until $t = t_F$. In fact, in every period holding shares is an equilibrium of the stage game since the strategy "Hold" is the best response to any number of players n that an agent might expect to sell.⁹ This strategy yields $\mathbb{E}(w_{t_F}) \approx w_0 \cdot (1 + m + s^2/2)^{t_F}$. The average outcome of the first strategy is zero profit (remember that subjects only keep net profits), while the second strategy provides large expected profits when the experiment session lasts for a long time.

Therefore the rational benchmark, with a risk neutral population, is to get in the market at $t = 1$ and hold the shares until the end of the experiment.¹⁰ The longer the duration of the experiment, the larger the expected profits. In the Online Appendix we consider the case of risk aversion and study a few cases of bounded rationality with myopic optimization, which may lead traders to sell the asset in the presence of high wealth levels and large price fluctuations.

3. Experimental results

In the following we report the experimental results. In Section 3.1 we document the emergence of excess trading in our experimental markets. Consistently with the goals set forth in the introduction of the paper, we then link our results on market activity to individual risk preferences and price forecasts. Section 3.2 relates trading behaviour to individual risk attitudes elicited in the experiment, while Section 3.3 depicts individual decisions of being active or inactive in the market conditional to expected returns.

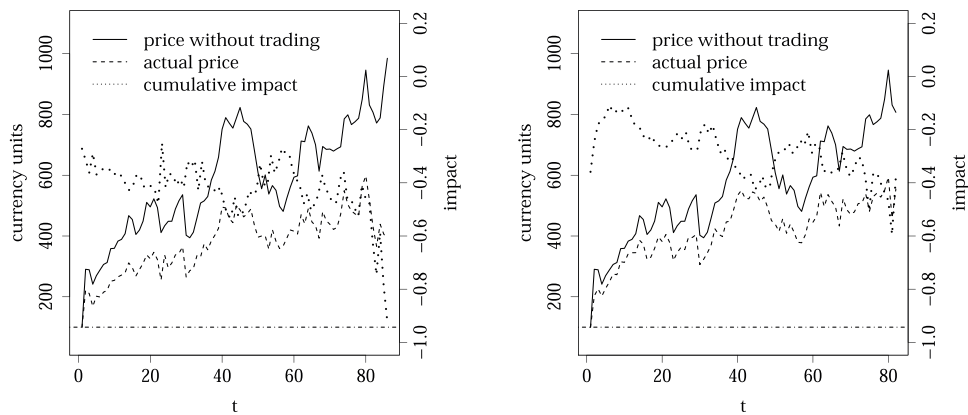
3.1. Trading activity

In this section we report on the trading activity observed in the experiment. Section 3.1.1 describes results on the volume of trading activity in relation to subjects' wealth, while Section 3.1.2 studies synchronization in subjects' trading activity.

⁸ As standard, we truncated the payoff function in order to avoid negative earnings. Negative net wealth was observed in first market runs but rarely in second market runs. One could argue that a truncated payoff function might lead single traders to act in a lazy way or to start experimenting. On the other hand, unsatisfactory performances in first runs could force participants to think harder about the game and could be one of the driver of subjects' learning that lead to the much higher earnings observed in second runs.

⁹ In fact, given the set of actions $\mathcal{A}_t = \{\text{Hold, Sell}\}$ available to the agent holding shares, the action "Hold" vs. "Sell" would result in a price $p_{t+1}^H = p_t \cdot \exp(m - (n-1)/N + s\eta_t)$ vs. $p_{t+1}^S = p_t \cdot \exp(m - n/N + s\eta_t)$, with $p_{t+1}^H > p_{t+1}^S$.

¹⁰ We remark however that, due to the indefinite horizon nature of the game, we can not rule out the presence of other equilibria on the basis of the Folk theorem. Nevertheless, coordination on such equilibria would require an extremely demanding coordination device for subjects. Given the laboratory environment in which the game is played (not to mention the presence of the noise in the price process), coordination on these equilibria can be ruled out as a concrete possibility.



(a) First sessions: quantities averaged over 198 subjects (18 for the last point). The correlation between realized price log returns and price log returns in the absence of trading is 0.85.

(b) Second sessions; quantities averaged over 201 subjects (23 for the last point). The correlation between realized price log returns and price log returns in the absence of trading is 0.89.

Fig. 1. Average price time series for first and second market sessions.

3.1.1. Excess trading and wealth

It is clear that if all subjects adopted the strategy of buying and holding the shares until the end of the game, the price impact would be $I(t) \equiv 0$ for $t > 1$ and the realized price at the end of the market sequence would lead to a most probable 640% increase in wealth. What we observe instead is that subjects keep on trading in and out of the market. The trading activity is in fact so high in the first market sessions that they barely break even, earning a meagre 10.7% on average. Remarkably, all groups learn to some extent and trade much less in the second market sessions, leading to a much higher average earning of 105%.¹¹ Given that the durations of the sessions were all different due to the indefinite time horizon, when comparing final wealth we refer to wealth at a stage that was reached in all cases ($t=63$). Our results are somewhat in line with findings in the literature on experimental asset markets showing that sufficient experience with an asset in certain environments eliminates mispricing and the emergence of bubbles (Smith et al., 1988; King, 1991; Van Boening et al., 1993; Dufwenberg et al., 2005; Haruvy et al., 2007; Lei and Vesely, 2009).¹² In fact we observe, as a result of learning, consistently different behaviour of subjects in the second market sequences when compared to their behaviour in the first sequences. The Welch two-sample t -test applied to the final wealth and average activity rate of each subject in first and second runs statistically confirms this difference (p -value $\approx 2^{-16}$). Therefore in the following analysis we treat separately first and second market sessions. Aggregating data for first and second market sessions leads to the price time series illustrated in Fig. 1.

Due to the different durations of the sessions, the number of data points used in the averaging process is not the same for each time t but a decreasing step function of t (this is clearly visualised in Fig. 2 below). The last data point includes 18 subjects for first sessions and 23 subjects for second sessions. Moreover, during the first experimental market sessions we experienced network problems with three computers. Hence the number of subjects included in the pool for the first runs is 198 instead of 201.

The solid lines represent the “bare” price time series, corresponding to the price evolution that would have occurred in the market if all agents had played the buy-and-hold strategy. A comparison with the dashed lines, depicting the realized prices, immediately reveals that trading significantly weakens the upwards price trend via the impact factor (dotted lines): the average slope (price trend) is divided by a factor of approximately 2 in the first sessions and 1.5 in the second sessions. The realized prices in Fig. 1 are proxies for the maximum earnings a subject would achieve if he used the buy-and-hold strategy within his group. However, the realized price log-returns remain highly correlated with the “bare” price time series: the correlation coefficient is 0.85 in first market sessions and 0.89 in second market sessions. The higher values of the slope (price trend) and of the correlation coefficient for the second sessions are due to a lower trading activity.

We used two simple ways to check whether the subjects arbitrated the fact that the shocks were identical in the two sessions. First, we tested if the subjects tried to avoid negative shocks larger than 10% in absolute value which occurred during the first run when trading in the second run but could not find any evidence of abnormal selling before such price drops: the p -value of the Welch test of the average impacts of the first and second runs one time step before the occurrence of large price drops is 0.53. Second, looking at the average price of the first run in Fig. 1, one clearly sees a fairly large price

¹¹ In practice the average payouts were higher since participants could not incur losses, therefore negative contributions did not play a role in actual average payouts.

¹² On the contrary, Hussam et al. (2008) and Xie and Zhang (2012) find that bubbles can be rekindled or sustained when the market experiences a shock from increased liquidity, dividend uncertainty and reshuffling, or from the admission of new subjects.

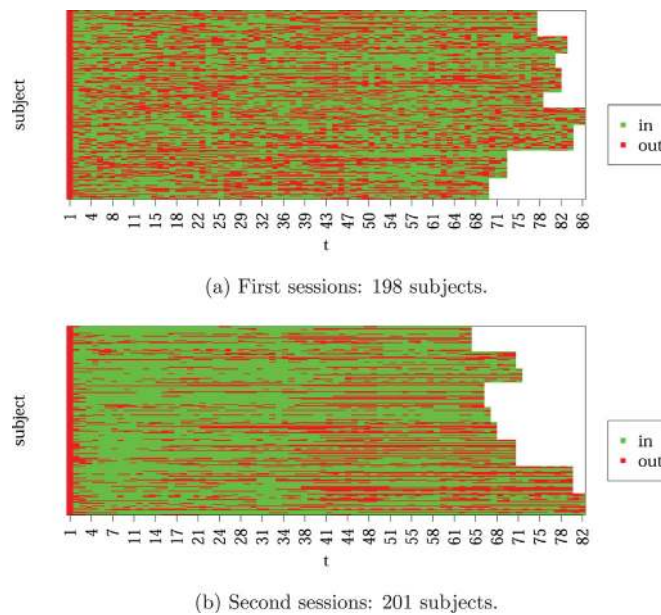


Fig. 2. Traders' positions – in the market (green) or out of the market (red) – in first sessions (above) and in second sessions (below). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

increase between times 31 and 40, followed by a large drop between times 45 and 56. If the agents tried to exploit this knowledge, one would have seen a spike of buying around time 31, and a massive sell-off between times 40 and 45. This is not the case: the agents buy just after a negative price return and sell just after a price increase in a quite systematic way in the second run, thus keep selling when during the price increase of times 31 to 40 and keep buying during the price decrease of times 45 to 56.

Fig. 2 displays the positions of the traders – in the market (green) or out of the market (red) – throughout the experimental market sequences. Notice that not all the sessions lasted the same number of periods, hence the white space in both Fig. 2a and b for large t . In fact, for each case, only one session – the longest – lasted until the maximum time t displayed, $t=86$ for first sessions (Fig. 2a) and $t=82$ for second sessions (Fig. 2b). The minimum session duration was $t=63$, time at which we therefore compare the average wealth between the runs.

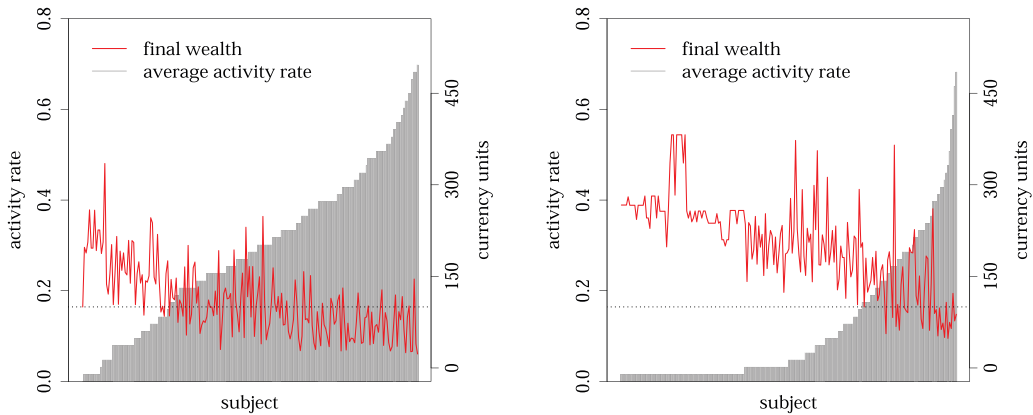
Subjects' positions – in (holding shares) or out of the market (holding cash) – are mostly intermittent, which implies excessive trading. However, comparing Fig. 2a and b, we observe that when the same subjects play for a second time, some of them actually learn the optimal strategy, which translates into “green corridors” in Fig. 2b.

Fig. 3 shows that the distribution of average trading activity changes when the same set of people play the game for the second time, and relates it to agents' final wealth. The number of people who keep trading to a minimum increases significantly in second sessions, where only a few outliers keep trading activity above 40%, meaning that they changed their market positions in more than 40% of the periods. In both cases, the final wealth of the agents is strongly anti-correlated with average trading activity, which is expected since trading is costly. In fact, if a trader decides to buy shares at period t and to sell them at period $t+1$, he will, on average, end up with less cash than he started because of his own contribution to price impact.

Fig. 4 show the average components of wealth over time. There are differences between first sessions (Fig. 4a) and second sessions (Fig. 4b). In first sessions, where overall trading activity is high (Fig. 3a), the average wealth does not follow the upward trend one would expect in a set-up with “guaranteed” average growth of 2% per period. In fact, players trade so much that they keep eroding their wealth when they sell (due to the negative impact on price) and affording fewer and fewer number of shares when they buy (due to the positive impact on price). This accounts for a negative impact bias on average and results in very low earnings at the end of the session.

On the other hand, in second sessions, where overall trading activity is much lower (Fig. 3b), the average wealth does increase with time. Nevertheless the number of shares owned eventually decreases, which is due not only to excessive trading but also to the fact that some subjects cash in their earnings before the end of the experiment and stay out of the market from that point onwards.

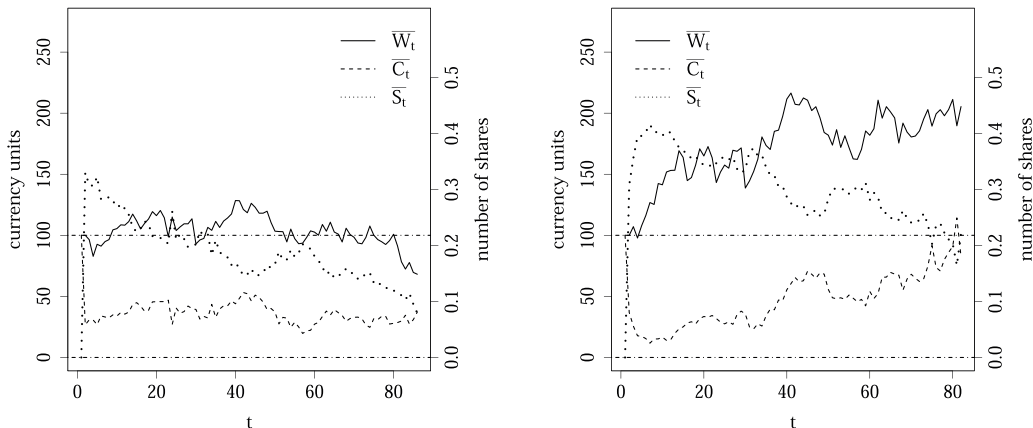
This is particularly visible in Fig. 4b, when a surge in price triggers selling orders over several periods which result in a higher average amount of cash and, naturally, in a lower average number of shares. Although this is also visible in the middle of the time series in first sessions (Fig. 4a), the difference between the two cases is that most of the resulting cash is eventually reinvested in the first sessions, while in the second sessions this does not happen: cash holdings consistently increase after the initial investment phase (Fig. 4b).



(a) First sessions: 198 subjects and 69-85 periods. The average final wealth at time $t = 63$ was 110.72 units of currency, the average activity rate 29% and the correlation between the two -65% .

(b) Second sessions: 201 subjects and 63-81 periods. The average final wealth at time $t = 63$ was 205.34 units of currency, the average activity rate 12% and the correlation between the two -73% .

Fig. 3. Activity rate and final wealth in first sessions (left) and in second sessions (right). Note the strong anti-correlation between activity and wealth.



(a) First sessions, 198-18 subjects.

(b) Second sessions, 201-23 subjects.

Fig. 4. Average wealth \bar{W}_t (solid), average cash \bar{C}_t (dashed) and average number of shares \bar{S}_t (dotted) in first sessions (left) and in second sessions (right).

As we showed in Section 2.2 the optimal strategy in our experimental set-up would be to collectively buy-and-hold and reap the benefits from the baseline average return of 2% per period in the absence of trading. We see that the behaviour of the agents is very far from this benchmark, even in the second sessions, where we observed a significant decrease in trading activity. The average performance in second sessions is indeed still far from what it would have been if all subjects used the optimal buy-and-hold strategy. Despite the learning, there is still excess trading activity which translates into detriment of collective welfare – since even the “virtuous” agents are adversely impacted by the trading activity of excessively “active” agents.

3.1.2. Collective trading modes – activity correlations, panic & euphoria

Our subjects trade too much, but can we describe in more detail how correlated their activity is? In fact, our initial intuition – that turned out to be quite far from what actually happened – was that subjects would initially buy the asset and then they would not trade in the initial phase of the game, letting the price rise from its initial value of 100 to quite high values, say 400 (EUR 100), before starting to worry that others might start selling, pushing the price down and potentially inducing a panic chain reaction. Given the results derived in Section 2.2, we remark that this conjecture was based on “irrational” behavior of traders in the market. We hypothesized that, due to the high amount of money at stake and the presence of random shocks (combined with the relatively short decision time), some traders could have panicked, overreacting to a negative shock, deciding therefore to sell the asset in order to secure their profits. The negative price impact of the asset sales, added to the initial negative shock, could have triggered herding behavior in the market with subjects selling the asset in order to cash in their profits, fearing a further price decrease due to the negative impact of sales. This would have translated into either a

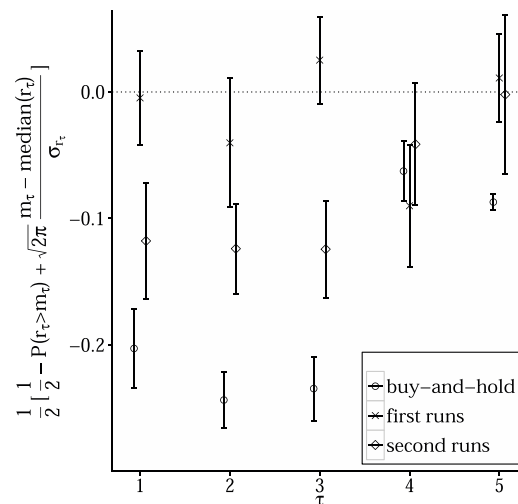


Fig. 5. Average skewness of price log returns as a function of aggregation length τ in first sessions (cross) and second sessions (diamond), together with the skewness of log returns in the buy-and-hold strategy (circle).

major crash, or perhaps smaller downward corrections, but in any case a significant *skewness* in the distribution of returns – absent in principle from the bare price series which is constructed to be perfectly symmetrical, since the noise term in Eq. (2.3) is symmetric. In fact, as we will see below, the empirical skewness of the particular realization of the noise turns out to be negative, so the reference point that we shall be comparing to must be shifted downwards.

We have therefore measured the relative skewness of the distribution of price changes, upon aggregation over time intervals of increasing length, from $\tau = 1$ round to $\tau = 5$ rounds. The idea is that a panic spiral would lead to a negative skewness that becomes larger and larger (in absolute value) when measured on larger time intervals, before going back down to zero after the typical correlation time of the domino effect. This is called the “leverage effect” in financial markets, and is observed in particular on stock indices where the negative skewness indeed grows as the time scale increases, before decreasing again, albeit very slowly (Bouchaud and Potters, 2003).

In order to reduce the measurement noise, it is convenient to measure the skewness using two low-moment quantities. One is $1/2 - P(r_\tau > m_\tau)$. If this quantity is negative, it means that large negative returns are more probable than large positive returns, as to compensate the excess number of positive returns larger than the mean. Another often used quantity is the mean m_t of the returns minus the median, normalised by the RMS of the returns on the same time scale. Again, if the median exceeds the mean, the distribution is negatively skewed (see Reigner et al. (2011) for further details about these estimators of skewness). Both quantities were found to give the same qualitative results, thus we chose to average these two definitions of skewness and plot them as a function of τ , averaged over all first and second sessions.

The result is shown in Fig. 5. The circles correspond exactly to the time series of bare prices because there is only one (collective) trade in the buy-and-hold strategy, right at the first period, which we discard from the computation. Although the bare returns were constructed using a Student’s t-distribution with 3 degrees of freedom, which by definition is not skewed, we see in Fig. 5 that the bare prices do not have zero skewness. This illustrates the role of the noise, which gives way to different values of skewness for bare prices depending on the number of periods of the session. We observe in Fig. 5 that the realized skewness of returns impacted by trade is typically larger (less negative) than bare returns, but without any significant time dependence. This suggests that buying orders tend to be more synchronized than selling orders, specially in the first sessions, but that neither buying nor selling orders induce further buy/sell orders. In short, there is no destabilising feedback loop in the present setting, which explains why we never observed any large crash in our experiments.

In order to detect more precisely the synchronisation of our agents, we compute the activity correlation matrix A associated to each session, defined as follows:

$$A_{ij} = \frac{1}{T-1} \sum_{t=2}^T \theta_i(t) \theta_j(t) - \frac{1}{T-1} \sum_{t=2}^T \theta_i(t) \times \frac{1}{T-1} \sum_{t=2}^T \theta_j(t), \quad (3.1)$$

where $\theta_i(t)$ is the activity of agent i at time t , $\theta_i(t) = 0$ if s/he is inactive, $\theta_i(t) = \pm 1$ if s/he buys or sells. The sums start at time $t = 2$ in order to avoid the expected synchronization of the first time step.

Global synchronization occurs when a substantial fraction of agents tend to act in exactly the same way across the experiment. Such phenomenon may be uncovered with Principal Component Analysis: the largest eigenvalues of A correspond to most important principal components of the subjects’ synchronization. More precisely, the “synchronized mode” of the subjects corresponds the eigenvalue whose associated eigenvector \vec{v} has the most components with the same sign. Mathematically, this eigenvector maximizes the absolute value of the dot product $\vec{v} \cdot \vec{e}$ where \vec{e} is the uniform vector $(1, 1, \dots, 1)/\sqrt{N}$.

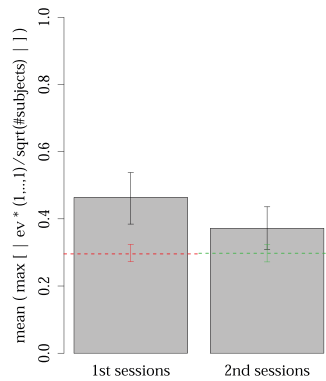


Fig. 6. Average maximum absolute value of the three dot products $\vec{v} \cdot \vec{e}$ between the eigenvectors corresponding to the three largest eigenvalues, and the unit vector \vec{e} , with the corresponding statistical error bars. All orders are considered except for the first time step, for which we expect an artificial bias towards synchronization. The dashed horizontal line around 0.3 corresponds to the null hypothesis of completely uncorrelated actions.

While the largest eigenvalue often corresponds to the “synchronized” mode, it is not clear a priori that this is the dominant mode in our experiments. Therefore we compute the dot product $\vec{v} \cdot \vec{e}$ for the three largest eigenvalues and retain maximum absolute value. We then average this largest absolute dot product over all first sessions and second sessions separately. The resulting values are represented in Fig. 6, and compared with a null-hypothesis benchmark, obtained using 1000 random bootstrap replicates of the experiments.

The dashed lines depict the cases where agents would act completely at random, which would lead to a value of this average maximum overlap at approximately 0.35. We see that the experimental results are clearly larger than the benchmark case: approximately 0.57 for the first sessions and approximately 0.5 for the second sessions (compared to a maximum value of 1 for a fully synchronized activity mode).

The method above allows us to make a quantitative analysis about overall synchronization in our experiment, be it through selling or buying orders, and we find that there is indeed significant synchronisation.

We also consider the separate cases of synchronization for selling orders and for buying orders. In order to do this, we construct an activity correlation matrix as in Eq. (3.1) but change the definition of $\theta_i(t)$ accordingly. This way, when we study the synchronization concerning only buying orders, we define $\theta_i(t) = 0$ if agent i is inactive or sells and $\theta_i(t) = 1$ if he buys. Likewise, in the case where we look into synchronization over selling orders only, we set $\theta_i(t) = 0$ if agent i is inactive or buys and $\theta_i(t) = -1$ if he sells.

In the Online Appendix (Figure A1a and Figure A1b) we show that splitting the data set as explained does yield similar results: the experimental results are larger than the benchmark for each case, with the difference more marked in first sessions than in second sessions. Again, the synchronisation of buy orders appears to be, according to this metric, slightly stronger than that of sell orders.

We therefore conclude that although our subjects cannot directly communicate with one another, there is a significant synchronisation of their activity, in particular during the first sessions and, as the skewness of the distribution reveals, for the buying activity. The mechanism for this synchronisation can only come from the common source of information that the subjects all observe, namely the price time series itself – see below for more about this.

Furthermore, we observe an asymmetry if we repeat the above method conditional on the sign of previous returns, in the sense that synchronisation is stronger for buying orders conditional on negative previous returns and for selling orders conditional on positive previous returns. This indicates some sort of “mean reversion”, which is in line with the findings on subjects’ trading behaviour discussed in Section 3.3 below.

3.2. Risk attitude and activity rate

To measure risk aversion of subjects we use the paired lottery choice instrument of Holt and Laury (2002). The Holt-Laury paired-lottery choice task is a commonly-used individual decision-making experiment for measuring individual risk attitudes. This second experimental task was not announced in advance; subjects were instructed that, if they were willing, they could participate in a second experiment that would last an additional 10–15 min for which they could earn an additional monetary payment. All subjects agreed to participate in this second experiment. In this task subjects choose between a lottery with high variance of payoffs (Option B) and lottery with less variance (Option A). As in Holt and Laury (2002) we use the relative frequency of B-choices (“risky” choices) as a measure for a preference for risk. Moreover, in order to make the amounts at stake comparable to what subjects could earn over an average session of trading periods and to assess whether or not the risk attitudes of subjects depended on the payoff levels involved, we elicited subjects’ choices for two lotteries, corresponding to two times (2×) and ten times (10×) the amounts offered by Holt and Laury (2002) in their baseline treatment. Subjects were ex-ante informed that they would throw a dice to determine which of the two lotteries

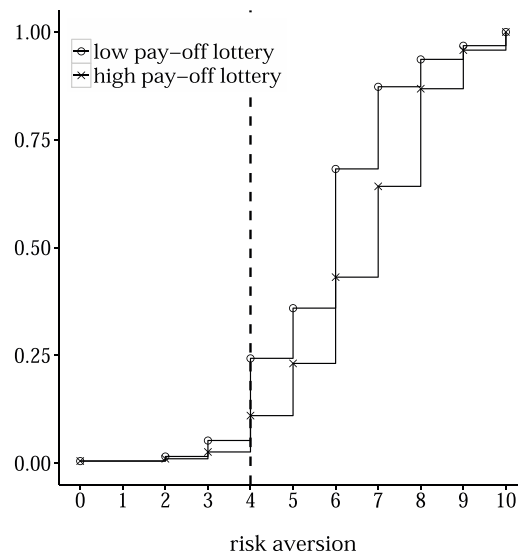


Fig. 7. Risk aversion cumulative distribution function. The x-axis shows the number of A-choices (“safe choices”) in the paired lotteries. The higher this level, the more risk-averse is the agent. The dashed vertical line shows where a risk neutral subject would be.

would determine their payoff from this additional experimental task. The Online Appendix includes the instructions for the Holt-Laury paired-lottery choice experiment.

In Fig. 7 we show the distribution of subjects according to their risk aversion, both for the 2× and the 10× lottery.¹³

The dashed vertical line shows where a risk neutral subject would be, based on the expected payoff differences alone. The fact that the majority of the subjects – 76% for the low-payoff and 89% for the high payoff lotteries – have a number of “safe” A-choices larger than 4 indicates that, overall, the participants in our experiment were risk-averse. Moreover, if we compare the two curves, we see that the subjects tend to safer choices when the lottery payoff is higher, which indicates that the risk aversion of our population not only depends but increases with the payoff level. This is corroborated by the Welch two-sample *t*-test, which states that the mean risk aversion values for low payoff and high payoff are statistically different with a *p*-value of 1.5×10^{-6} .

We now relate individual risk attitude, as elicited by the binary lottery choices, to trading behaviour. Previous experimental studies on the relation between elicited risk attitudes and aggregate market behaviour (Robin et al., 2012; Fellner and Maciejovsky, 2007) show that the higher the degree of risk aversion the lower the observed market activity. On the other hand, Michailova (2010) finds no significant effect of the number of safe choices in paired binary lotteries on the frequency of trading.

Fig. 8 displays average activity rates and levels of final wealth for different risk attitudes in both first and second runs.

Our results clearly show that the activity rate decreases and the final wealth increases with the level of risk aversion, both in first runs (red and orange) and in second runs (green and blue). In other words, preference for risk is an important determinant of excess trading in our experiment. Our findings are in line with the empirical evidence suggesting that risk-loving, overconfident individuals are more willing to invest in stocks (Keller and Siegrist, 2006) and engage in speculative activity (Odean and Barber, 2011; Camacho-Cuena et al., 2012).

3.3. Price forecasts and trading behaviour

The fact that the subjects input their price predictions throughout the experiment allows us to have a glimpse of their frame of mind. In fact, trades only tell about the consequence of the state of mind (the price expectation) of traders when they are active. But traders, both in real life and in experiments, are in fact inactive most of the time. As a consequence, trades alone are unlikely to be able to explain why traders are inactive. Since we have both trades and subjects’ price expectations, we are able to give a consistent picture of activity and inactivity as a consequence of price return expectations.

In both market sessions, the subjects did not input anything in about 7% of the time, as price prediction was not a mandatory activity (although monetarily incentivised); in the following, we restrict our analysis to the subjects that did report their predictions.

The discussion focuses on the predicted log returns. In other words, from subject *i*’s price predictions $\hat{p}_i(t+1)$, we compute the predicted log returns $\hat{r}_i(t+1) = \log[\hat{p}_i(t+1)/p(t)]$. First, we found that the predictions, averaged over all subjects of a given session, are not autocorrelated, which is a first clue that prediction formation is not mostly self-referential. The average

¹³ We discarded from our data set 8 subjects with inconsistent lottery choices.

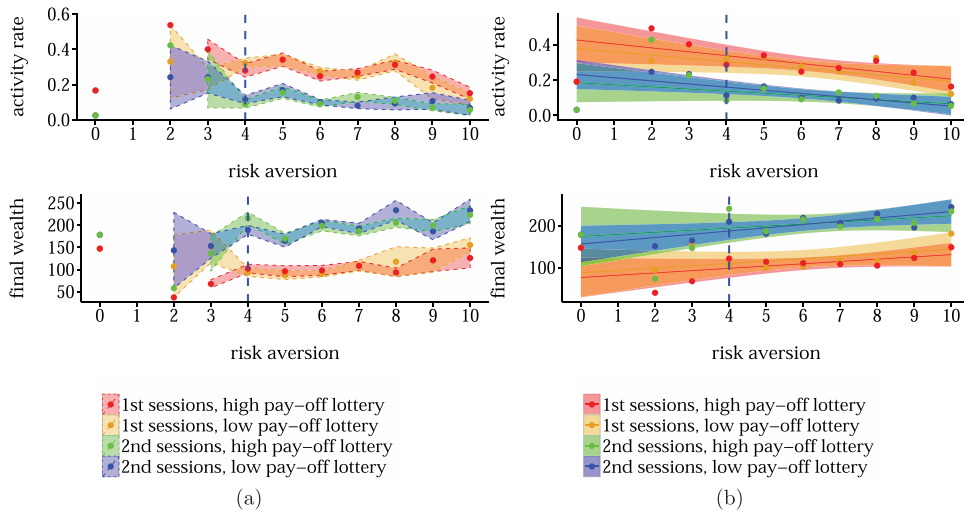


Fig. 8. Risk attitudes, trading activity and final wealth. Plot (a) displays average activity rate and final wealth for each level of risk aversion. Plot (b) shows linear regressions with 95% intervals of confidence of the average activity rate and final wealth on risk aversion. Clearly, more risk averse subjects trade less and end up with a higher final wealth. (For interpretation of the references to color in the text, the reader is referred to the web version of this article.)

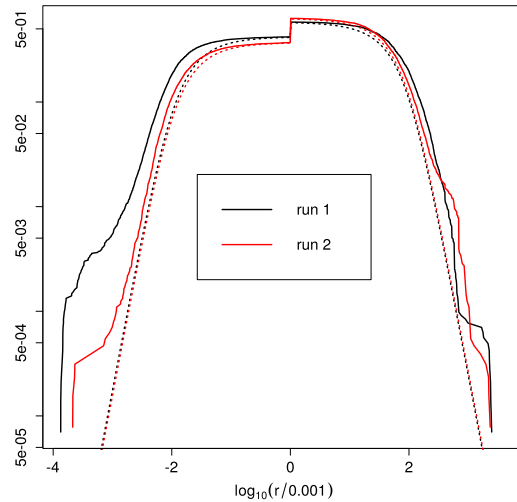


Fig. 9. Reciprocal empirical cumulative distribution functions of negative (left) and positive (right) expected price returns during the first and second sessions (black and red lines, respectively). The baseline return distribution is plotted in dashed lines. The jump at $r=0$ indicates the respective fraction of positive and negative predictions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

prediction in the first market sessions is -0.01 , and $+0.02$ in the second market sessions (exactly in line with the average return $m = 2\%$ in absence of trading). The percentage of positive predicted returns is 54% in the first run, and 58% in the second run. Fig. 9 illustrates the full empirical cumulative distribution functions of expectations for both runs and for positive and negative return separately.

The starting point for each of the positive and negative distributions represents the fraction of positive and negative expectations, thus the higher jump at $r=0$ for the second session reflects the increase in the fraction of positive expected returns.

We then check how the distributions of positive and negative expected returns are related to the baseline return distribution, determined by the Student noise term $s\eta_t$ in the price updating rule. The baseline return distribution is plotted in dashed lines in Fig. 9; comparison between the latter and empirical distribution of expected returns gives a first clue about the type of extrapolation rules from past returns that the agents use. The asymptotic power-law tails of Student's t -distributed variables, such as the noise η_t , are unchanged under summation, by virtue of the central limit theorem. This implies that linear expectations yield expectation distributions with power-laws tails that have the same exponent. On the other hand, panic or euphoria may lead to non-linear extrapolations and thus may modify either the tail exponent of these distributions, or even the nature of the tails.

Table 1

Fits of the power-law part of return expectations; r_{min} and α denote respectively the most likely starting point and the exponent of the power-law.

	r_{min}	α
Run 1 $r < 0$	0.21	3.5
Run 2 $r > 0$	0.10	2.7
Run 1 $r < 0$	0.13	2.6
Run 2 $r > 0$	0.18	3.5

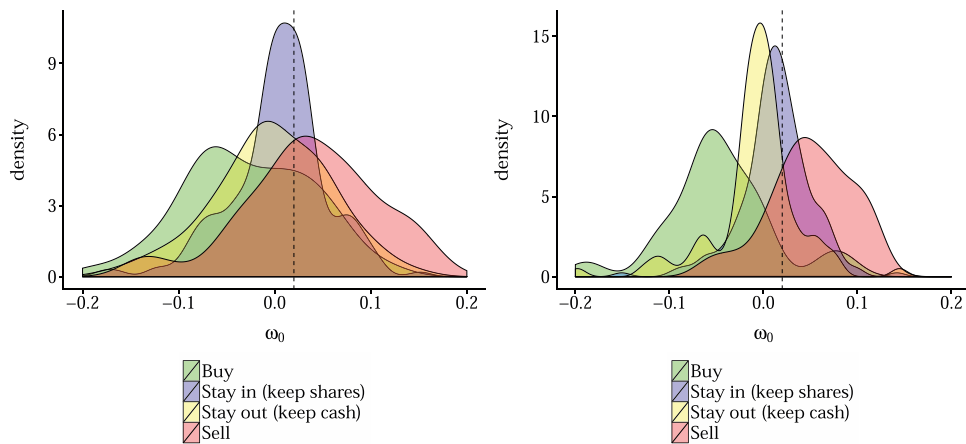


Fig. 10. Densities of trader average return prediction ω_0 during the first sessions (left plot) and second sessions (right plot) for the four types of decisions. Dashed vertical lines refer to the baseline return of 2%.

The most obvious finding is that the actions of the agents increase the volatility of the baseline signal (in dashed lines) as the empirical distribution functions are above the baseline signal for both sequences. The amplification of the noise for positive expectations is almost the same in the two sequences, while for negative expectations there are marked differences between the two runs as the scale of negative expectations was much larger during the first run. We used robust power-law tail fitting (Clauset et al., 2009; Gillespie, 2015) and determined the most likely starting point of a power-law r_{min} and the exponent α (see Table 1). Quite remarkably, the parameters of the positive and negative tails are simply swapped between the two runs: thus not only the scale of negative expectations changes, but the nature of largest positive and negative expectations also changes. The fitted tail exponent is not far from 3, the one of the Student noise showing once again the absence of destabilizing feedback loops.

The fact that the subjects have heavy-tailed predictions suggests that they form their predictions by learning from past returns, which do contain heavy tails because of the Student's t -distributed noise. We thus hypothesize some relationship between predicted returns and past returns. This is in line with the best established fact about real investors, which is the contrarian nature of their trades: their net investment over a given period is anti-correlated with past price returns (Jackson, 2003; Kaniel et al., 2008; Grinblatt and Keloharju, 2000; Challet, 2013). In addition, previous experiments (Hommes et al., 2005) have demonstrated that four simple classes of linear predictors using past returns are usually enough to reproduce the observed price dynamics.

Based on the considerations outlined above, we fit the return predictions of each subject with a linear model.

$$\hat{r}(t+1) = \omega_0 + \omega_1 r(t), \quad (3.2)$$

where ω_0 and $\omega_1 \in \mathbb{R}$. Price return predictions are fitted separately for each trader, each session, and each possible action. We fit ω_0 and ω_1 simultaneously. Section 3.3.1 discusses results for ω_0 while Section 3.3.2 is devoted to ω_1 .

3.3.1. Average predictions (ω_0)

In order to give a picture of traders' movements on the market we compute return expectations conditionally on the actions of the subjects. There are four possible actions: buying, selling, holding shares and holding cash. Fig. 10 reports the conditional distributions of price return predictions, for both sessions.

The results are qualitatively the same for both market sequences: the conditional distributions are clearly separated, as shown in Table 2 below; the main difference between the two runs is that the variance of expectations among the population is much reduced in the second run.

Let us break down the results for each possible action:

- (1) When the subjects hold assets, their expectations are in line with the baseline return of 2%.
- (2) When the subjects hold cash, their expectations are significantly lower (essentially zero).

Table 2

Tests of the difference of distributions of ω_0 among the subjects, conditional on two given actions. The table reports the p -values of Mann–Whitney tests for each possible pair of actions.

ω_0 , 1st run	Sell	Hold cash	Hold shares
Buy	$1.5 \cdot 10^{-8}$	$3.9 \cdot 10^{-3}$	$1.2 \cdot 10^{-5}$
Sell		$3.8 \cdot 10^{-5}$	$8.2 \cdot 10^{-5}$
Hold cash			$1.5 \cdot 10^{-1}$
ω_0 , 2nd run	Sell	Hold cash	Hold shares
Buy	$2.1 \cdot 10^{-7}$	$1.5 \cdot 10^{-4}$	$1.0 \cdot 10^{-7}$
Sell		$6.1 \cdot 10^{-7}$	$5.9 \cdot 10^{-5}$
Hold cash			$6.6 \cdot 10^{-6}$

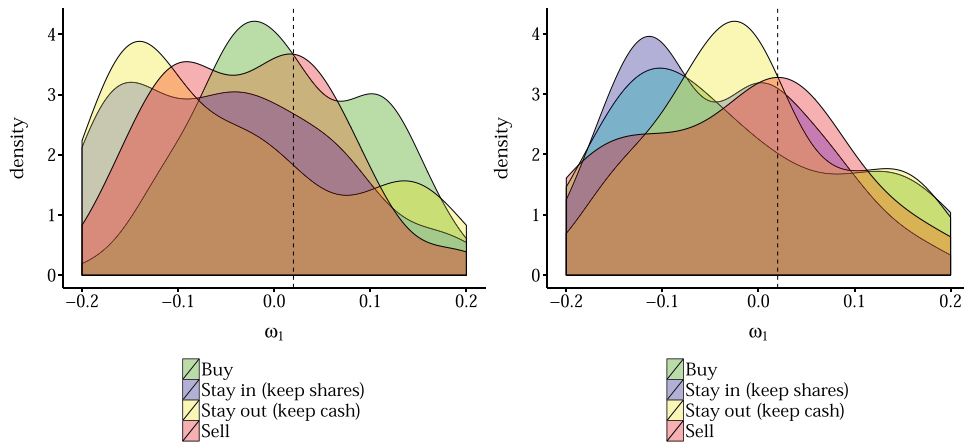


Fig. 11. Densities of trader conditional return prediction ω_1 during the first sessions (left plot) and second sessions (right plot) for the four types of decisions. Dashed vertical lines refer to the baseline return of 2%.

Table 3

Tests of the difference of distributions of ω_1 among the subjects, conditional on two given actions. The table reports the p -values of Mann–Whitney tests for each possible pair of actions.

ω_1 , 1st run	Sell	Hold cash	Hold shares
Buy	$9.6 \cdot 10^{-2}$	$2.3 \cdot 10^{-5}$	$8.8 \cdot 10^{-10}$
Sell		$3.7 \cdot 10^{-2}$	$1.7 \cdot 10^{-4}$
Hold cash			$2.9 \cdot 10^{-2}$
ω_1 , 2nd run	Sell	Hold cash	Hold shares
Buy	$6.2 \cdot 10^{-1}$	$7.6 \cdot 10^{-1}$	$4.1 \cdot 10^{-2}$
Sell		$5.6 \cdot 10^{-1}$	$8.6 \cdot 10^{-2}$
Hold cash			$2.9 \cdot 10^{-4}$

(3) When the subjects make a transaction, however, their expectations of the next returns are anti-correlated with their actions. In other words, they buy when they expect a negative price return and vice versa.

Thus, the actions of the subjects are fully consistent with their expectations: they do not invest when they do not perceive it as worthwhile and they keep their shares when they have a positive expectation of future gains. The actions of trading subjects are instead consistent with a “buy low, sell high” strategy. In fact, knowing that transactions will be executed at the price realized in the next period, subjects submit buy orders when they expect negative returns and sell orders when they expect positive returns.

3.3.2. Predictions and past price returns (ω_1)

We find that the importance of the coefficient ω_1 , which encodes the linear extrapolation of the past return on future returns, is very weak. Fig. 11 reports the conditional distributions of past returns’ impact coefficient for both runs, while Table 3 reports the p -values of the Mann–Whitney tests between coefficients ω_1 between all state pairs for all subjects.

The plot shows that during the first run, this coefficient was negative when the agents did not act and zero when they did trade. The second run is different: the coefficients do not seem to depend much on the state, the only clear difference is between holding cash and holding shares.

The lack of influence of this coefficient is confirmed when one measures the average predicted return conditional on the action of the subjects, which gives results very close to ω_0 .

4. Conclusion

We presented the results of a trading experiment in which the pricing function favours early investment in a risky asset and no posterior trading. In our experimental markets the subjects would make an almost certain gain of over 600% if they all bought shares in the first period and held them until the end of the experiment.

However, market impact as defined in Eq. (2.4) acts *de facto* as a transaction cost which erodes the earnings of the traders. Our subjects are made well aware of this mechanism. Still, when they participate in the experiment for the first time, their trading activity is so high that their profits average to almost zero. They are however found to fare much better when they repeat the experiment as they earn 105% on average – which is still much below the performance of the simple risk-neutral rational strategy outlined above. We therefore find that, echoing Odean (1999) and Odean and Barber (2011), *investors trade too much*, even in an environment where trading is clearly detrimental and buy-and-hold is an almost certain winning strategy (at variance with real markets where there is nothing like a guaranteed average return of 2% per period). In line with previous experimental results, we do observe some learning when the experiment is repeated within the same market environment. However, we remark that financial markets are evolving environments (for example, due to the continuous arrival of new market participants), which are far more complex than our artificial setup (for example, assets' returns are unknown). This presumably makes learning in real conditions much more difficult than in our stylized world. Therefore we believe that our results are potentially quite important when translated into the real world: unwarranted individual decisions can lead to a substantial loss of collective welfare when mediated by the mechanics of financial markets, in particular price impact. We leave the analysis of whether and how fast our artificial markets converge to the optimal outcome and what would happen if shocks hit the system to future research.

At the end of the experiment we collect data on traders' risk attitude by means of paired lottery choices à la Holt and Laury (2002). We observe that, overall, our subjects are risk-averse and their relative risk aversion increases with the payoff level in a way that is quantitatively similar to the results reported in Holt and Laury (2002). We then relate individual risk attitude with the results of the trading experiment and we observe that the activity rate increases and the final wealth decreases with subjects' preference for risk.

Moreover, in each period we also ask subjects to predict the next price of the asset. This provides us with additional information about our controlled experimental market, since we have access not only to the decisions of each trader but also to their expectations. It is important to emphasize that this information would not be available in broker data. In fact, knowing the expectation behind each decision of each trader – including the decisions to do nothing – is one of the advantages of laboratory experiments compared to empirical analysis of real data. Using this information, we confirm in Section 3.3 that the traders in our experiment have a contrarian nature, which, together with the pattern of excessive trading, is one of the known features of individual traders in real financial markets, as discussed in de Lachapelle and Challet (2010) and Challet (2013). In particular, it seems that our subjects actively engaged in trading in the attempt to “beat the market”, as they tried to make profits by buying the asset when they expected the price to be low and selling the asset when they expected the price to move upwards.

Contrary to what happens in real financial markets, we have not observed any “leverage effect” (increase of volatility after down moves (Bouchaud et al., 2001)). Although there is a clear detrimental collective behaviour in all sessions of this experiment, we do not witness any big crash or avalanche of selling orders that would result from a panic mode. As we discuss in Section 3.1.2, the coordination amongst traders was actually slightly stronger when buying than when selling, resulting in a positive skewness of the returns. In order to induce crashes and study their dynamics, the duration of the experiment could be expanded while the time available for each decision could be reduced. The former would increase the probability of a sudden event by increasing the number of trials and the amounts at stake, while the latter could contribute to higher stress levels amongst subjects and increased sensitivity to price movements. However, we believe that a more efficient way to generate panic and herding behaviour would be to reduce the “normal” volatility level while and increasing the amplitude of “jumps” in the bare return time series. Within the current setting, large fluctuations do not seem surprising enough to trigger panic among our participants. Another idea, perhaps closer to what happens in financial markets, would be to increase the impact of sell orders and reduce the impact of buy orders when the price is high, mimicking the fact that buyers are rarer when the price is high (on this point, see the recent results of Donier and Bouchaud (2015a,b) on the Bitcoin). Other natural extensions of this experiment would include the possibility of fraction orders and hedging, as well as short selling. We leave these extensions and additional experimental designs to future research.

Acknowledgments

We would like to thank the associate editor and two anonymous referees. This work was partially financed by the EU “CRISIS” project (grant number: FP7-ICT-2011-7-288501-CRISIS), Fundação para a Ciência e Tecnologia, and the Ministry of Education, Universities and Research of Italy (MIUR), program *SIR* (grant n. RBSI144KWH). None of the above are responsible for errors in the paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [10.1016/j.jebo.2017.05.013](https://doi.org/10.1016/j.jebo.2017.05.013).

References

- Barber, B.M., Lee, Y.-T., Liu, Y.-J., Odean, T., 2009. Just how much do individual investors lose by trading? *Rev. Financ. Stud.* 22 (2), 609–632.
- Barber, B.M., Odean, T., 2000. Trading is hazardous to your wealth: the common stock investment performance of individual investors. *J. Finance*, 773–806.
- Barber, B.M., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors? *Rev. Financ. Stud.* 21 (2), 785–818.
- Barber, B.M., Odean, T., 2013. The behavior of individual investors. In: *Handbook of the Economics of Finance*, vol. 2., pp. 1533–1570.
- Barberis, N., Thaler, R., 2003. A survey of behavioral finance. In: *Handbook of the Economics of Finance*, vol. 1., pp. 1053–1128.
- Bondt, W.F., Thaler, R., 1985. Does the stock market overreact? *J. Finance* 40 (3), 793–805.
- Bouchaud, J.-P., 2010. Price impact. In: *Encyclopedia of Quantitative Finance*. Wiley Online Library.
- Bouchaud, J.-P., 2011. The endogenous dynamics of markets: price impact, feedback loops and instabilities. In: Berd, A. (Ed.), *Lessons from the Financial Crisis*. Risk Publications, pp. 346–374.
- Bouchaud, J.-P., Farmer, J.D., Lillo, F., 2009. How markets slowly digest changes in supply and demand. In: Hens, T., Schenk-Hopp, K.R. (Eds.), *Handbook of Financial Markets: Dynamics and Evolution*, Handbooks in Finance, North-Holland, San Diego., pp. 57–160.
- Bouchaud, J.-P., Matacz, A., Potters, M., 2001. Leverage effect in financial markets: the retarded volatility model. *Phys. Rev. Lett.* 87 (22), 228701.
- Bouchaud, J.-P., Potters, M., 2003. *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*. Cambridge University Press.
- Breaban, A., Noussair, C., 2014. *Fundamental Value Trajectories and Trader Characteristics in an Asset Market Experiment*. Discussion paper. Tilburg University, Center for Economic Research.
- Caccioli, F., Shrestha, M., Moore, C., Farmer, J.D., 2014. Stability analysis of financial contagion due to overlapping portfolios. *J. Bank. Finance* 46, 233–245.
- Camacho-Cuena, E., Requate, T., Waichman, I., 2012. Investment incentives under emission trading: an experimental study? *Environ. Resour. Econ.* 53 (2), 229–249.
- Challet, D., de Lachapelle, D.M., 2013. A robust measure of investor contrarian behaviour. In: *Econophysics of Systemic Risk and Network Dynamics*. Springer, pp. 105–118.
- Cipriani, M., Guarino, A., 2014. Estimating a Structural Model of Herd Behavior in Financial Markets? *Am. Econ. Rev.* 104 (1), 224–251.
- Clauset, A., Shalizi, C.R., Newman, M.E., 2009. Power-law distributions in empirical data? *SIAM Rev.* 51 (4), 661–703.
- Cont, R., Wagalath, L., 2014. Fire sales forensics: measuring endogenous risk. *Math. Finance*.
- Crockett, S., Duffy, J., 2013. A Dynamic General Equilibrium Approach to Asset Pricing Experiments. Working Paper.
- de Lachapelle, D.M., Challet, D., 2010. Turnover, account value and diversification of real traders: evidence of collective portfolio optimizing behavior. *New J. Phys.* 12 (7), 075039.
- Donier, J., Bouchaud, J.-P., 2015a. From Walras' Auctioneer to Continuous Time Double Auctions: A General Dynamic Theory of Supply and Demand, [Arxiv:1506.03758](https://arxiv.org/abs/1506.03758).
- Donier, J., Bouchaud, J.-P., 2015b. Why Do Markets Crash? Bitcoin Data Offers Unprecedented Insights, [Arxiv:1503.06704](https://arxiv.org/abs/1503.06704).
- Dufwenberg, M., Lindqvist, T., Moore, E., 2005. Bubbles and experience: an experiment. *Am. Econ. Rev.*, 1731–1737.
- Fellner, G., Maciejovsky, B., 2007. Risk attitude and market behavior: evidence from experimental asset markets? *J. Econ. Psychol.* 28 (3), 338–350.
- Gervais, S., Odean, T., 2001. Learning to be overconfident? *Rev. Financ. Stud.* 14 (1), 1–27.
- Gillespie, C.S., 2015. Fitting heavy tailed distributions: the powerLaw package? *J. Stat. Softw.* 64 (2), 1–16.
- Giusti, G., Jiang, J.H., Xu, Y., 2012. Eliminating Laboratory Asset Bubbles by Paying Interest on Cash. Working Paper GATE., pp. 2012–2035.
- Gopikrishnan, P., Meyer, M., Amaral, L.N., Stanley, H.E., 1998. Inverse cubic law for the distribution of stock price variations? *Eur. Phys. J. B-Condens. Matter Complex Syst.* 3 (2), 139–140.
- Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. *Rev. Financ. Stud.*, 714–746.
- Grinblatt, M., Keloharju, M., 2000. The investment behavior and performance of various investor types: a study of finland's unique data set? *J. Financ. Econ.* 55 (1), 43–67.
- Grinblatt, M., Keloharju, M., 2001. What makes investors trade? *J. Finance* 56 (2), 589–616.
- Haruvy, E., Lahav, Y., Noussair, C.N., 2007. Traders' expectations in asset markets: experimental evidence? *Am. Econ. Rev.* 97 (5), 1901–1920.
- Holt, C.A., Laury, S.K., 2002. Risk aversion and incentive effects? *Am. Econ. Rev.* 92 (5), 1644–1655.
- Hommes, C., 2013. Reflexivity expectations feedback and almost self-fulfilling equilibria: economic theory, empirical evidence and laboratory experiments. *J. Econ. Methodol.* 20 (4), 406–419.
- Hommes, C., Sonnemans, J., Tuinstra, J., Van de Velden, H., 2005. Coordination of expectations in asset pricing experiments. *Rev. Financ. Stud.* 18 (3), 955–980.
- Hommes, C., Sonnemans, J., Tuinstra, J., van de Velden, H., 2008. Expectations and bubbles in asset pricing experiments. *J. Econ. Behav. Organ.* 67 (1), 116–133.
- Huck, S., Normann, H., Oechssler, J., 2004. Two are few and four are many: number effects in experimental oligopolies. *J. Econ. Behav. Organ.* 53, 435–446.
- Hussam, R.N., Porter, D., Smith, V.L., 2008. That she blows: can bubbles be rekindled with experienced subjects? *Am. Econ. Rev.* 98 (3), 924–937.
- Jackson, A., 2003. *The Aggregate Behaviour of Individual Investors*. Working Paper ssn 536942.
- Jondeau, E., Rockinger, M., 2003. Testing for differences in the tails of stock-market returns? *J. Empir. Finance* 10 (5), 559–581.
- Kaniel, R., Saar, G., Titman, S., 2008. Individual investor trading and stock returns? *J. Finance* 63 (1), 273–310.
- Keller, C., Siegrist, M., 2006. Investing in stocks: the influence of financial risk attitude and values-related money and stock market attitudes? *J. Econ. Psychol.* 27 (2), 285–303.
- King, R.R., 1991. Private information acquisition in experimental markets prone to bubble and crash? *J. Financ. Res.* 14 (3), 197–206.
- Lei, V., Noussair, C.N., Plott, C.R., 2001. Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. actual irrationality. *Econometrica*, 831–859.
- Lei, V., Vesely, F., 2009. Market efficiency: evidence from a no-bubble asset market experiment? *Pac. Econ. Rev.* 14 (2), 246–258.
- Michailova, J., 2010. *Overconfidence, Risk Aversion and (Economic) Behavior of Individual Traders in Experimental Asset Markets*. MPRA Paper 26390. University Library of Munich, Germany.
- Michailova, J., Schmidt, U., 2016. Overconfidence and bubbles in experimental asset markets? *J. Behav. Finance* 17 (3), 280–292.
- Noussair, C.N., Powell, O., 2010. Peaks and valleys: price discovery in experimental asset markets with non-monotonic fundamentals? *J. Econ. Stud.* 37 (2), 152–180.
- Odean, T., 1999. Do investors trade too much? *Am. Econ. Rev.* 89 (5), 1279–1298.
- Odean, T., Barber, B., 2011. Boys will be boys: gender, overconfidence, and common stock investment. *Q. J. Econ.* 116 (1), 261–292.
- Oechssler, J., 2010. Searching beyond the lamppost: let's focus on economically relevant questions? *J. Econ. Behav. Organ.* 73 (1), 65–67.
- Palan, S., 2013. A review of bubbles and crashes in experimental asset markets? *J. Econ. Surv.* 27 (3), 570–588.
- Powell, O., Shestakova, N., 2016. Experimental asset markets: a survey of recent developments. *J. Behav. Exp. Finance* 12, 14–22.
- Reigneron, P.-A., Allez, R., Bouchaud, J.-P., 2011. Principal regression analysis and the index leverage effect? *Phys. A: Stat. Mech. Appl.* 390 (17), 3026–3035.

- Robin, S., Straznicka, K., Villevall, M.C., 2012. *Bubbles and Incentives: An Experiment on Asset Markets*. MPRA Paper no. 37321.
- Smith, V., 2010. *Theory and experiment: what are the questions?* *J. Econ. Behav. Organ.* 73 (1), 3–15.
- Smith, V.L., Suchanek, G.L., Williams, A.W., 1988. *Bubbles, crashes, and endogenous expectations in experimental spot asset markets*. *Econometrica* 56 (5), 1119–1151.
- Solnik, B., Zuo, L., 2012. *A global equilibrium asset pricing model with home preference?* *Manag. Sci.* 58 (2), 273–292.
- Stöckl, T., Huber, J., Kirchler, M., 2015. *Multi-period experimental asset markets with distinct fundamental value regimes?* *Exp. Econ.* 18 (2), 314–334.
- Van Boening, M.V., Williams, A.W., LaMaster, S., 1993. *Price bubbles and crashes in experimental call markets?* *Econ. Lett.* 41 (2), 179–185.
- Xie, H., Zhang, J., 2012. *Bubbles and Experience: An Experiment with a Steady Inflow of New Traders*. Cirano Scientific Series. CIRANO, Montreal.