Humans in Charge of Trading Robots: The First Experiment

Elena Asparouhova*, Peter Bossaerts, Kristian Rotaru, Tingxuan Wang,
Nitin Yadav and Wenhao Yang

ABSTRACT

We study engagement of algorithmic trading (robots) in a financial markets experiment when participants have access to a portfolio of robots, which they may deploy, launch, halt and replace at will, while still trading manually. The setting is that of Smith, Suchanek and Williams (1988), where bubbles emerge reliably. Against control sessions where only manual trading is allowed, we observe equally large and frequent bubbles and, in early periods of trading, significantly higher effective bid-ask spreads. Moreover, we find flash crashes/price surges in the first period, which are absent under manual trading. Lastly, the participants who engage in both robot and manual trading perform marginally (but significantly) better.

JEL classification: G12, G41, C92.

*Acknowledgement: Asparouhova and Bossaerts acknowledge funding from the U.S. National Science Foundation (SES-1061824); Bossaerts acknowledges funding from the Australian Research Council (DP180102284). The paper benefited from discussions during the CAFIN-GRU Conference on Market Design and Regulation in the Presence of High-Frequency Trading (2017, Hong Kong).

> I have found that it is seldom a good idea to manually override a model no matter how treacherous the market is looking.
>
> *Dr. E.P. Chan*

I. Motivation

We report results from a series of financial market experiments in a setting where mispricing of a multiperiod security is frequently observed. For the first time, we give participants access to algorithmic trading programs (robots) that execute a number of simple strategies. Participants have the opportunity to deploy any (single) robot of choice but also trade manually, alongside the robot.

There are a number of reasons to believe, ex ante, that access to robots would improve pricing.

First and foremost, algorithms execute a strategy, and thus the choice of a robot implies the choice and possibly the commitment to a strategy. In contrast, when trading manually, one does not have to commit to a strategy; adjustments can be made at any time during the trading process and with virtually no delay. Once deployed, a robot can be stopped, but it takes time relative to the speed with which the robot can submit orders, which in our case is about one every second (a bound imposed by us, the experimenters, so that the orders a robot places before being halted is not too large; the bound can be as small as a millisecond).

Manual trading can thus be thought of as the implementation of what experimental game theorists refer to the "direct response method" while robot trading as the implementation of the "strategy method,"(Brandts and Charness (2011)): rather than directly play a game, one delegates the actual moves to a third party (here: the robot), and the task is limited to choosing the "right" strategy. The strategy method requires participants to think ahead. The opportunity to "learn on the job" (learn while posting offers and executing trades) comes at higher costs when the "wrong" robot is deployed. The strategy method forces the player to consider the strategies others may choose and to respond, in the form of selecting one's own strategy, as to maximize profits.

The strategy method requires that participants think about future prices, about how others react to those prices, and what those reactions imply about one's own best-response strategy. In our context, participants could entertain many future pricing scenarios. But anticipating pricing at fundamental value (equal to the sum of expected payoffs[1] to the asset) is both focal – it is given to the participants during the instruction period – and minimizes losses in case the anticipation is incorrect. In simple games with strategic uncertainty (i.e., multiple equilibria), it has been shown that risk minimization predicts human play (Embrey, Fréchette, and Yuksel (2017)).

As a result, we hypothesized that access to the strategy method would improve pricing: bubbles would cease to emerge, and prices would align with fundamental values (i.e., discounted sum of expected future dividends).

All participants, independent of whether they opted to trade manually or deploy robots, knew about the potential presence of algorithmic traders. Therefore those who traded manually also had to think "strategically." This, we hypothesized, would only reinforce the impact of the strategy method: the mere

availability of algorithmic trading should help reduce mispricing.

Archival data from field markets provide some indication that strategic mis-coordination in the face of strategic uncertainty may play a role in mispricing (Goldstein, Kumar, and Graves (2014)). But such cannot be taken as proof that our hypothesis (that the presence of robots diminishes mis-coordination and, hence, mispricing) is correct, since access to robots in the field studies is not controlled. This is why an experiment is needed, which is what we report on here.

History in field markets has shown that price quality generally improved along with the advent of algorithmic trading. Of course, it is almost impossible to determine the fundamental value of an asset in the field, so evidence is only indirect. The evidence points more directly to a change in liquidity: bid-ask spreads become narrower, and consequently, autocorrelation in transaction price changes is reduced (see, e.g., Carrion (2013); Brogaard, Hendershott, and Riordan (2014); Angel, Harris, and Spatt (2011)). Our experiments could provide further insights.

It is sometimes claimed that the use of robots has increased the incidence of flash crashes, but historical data are contradictory (e.g., Chaboud, Chiquoine, Hjalmarsson, and Vega (2014); Brogaard et al. (2014)). Like Biais, Foucault, and Moinas (2011), our conjecture is that strategic mis-coordination may cause excessive price volatility. We would expect such mis-coordination to mostly occur in early rounds (periods) of our experiments, and hence, we conjectured that more flash crashes (and surges) would emerge when participants are given access to robots.

Interestingly, earnings to traders deploying algorithmic traders have not been abnormal on average (Kissell and Malamut (2005)), suggesting that some “equilibrium” has been reached whereby all traders have adjusted strategies. In experiments like ours, inter-trader profit comparisons can be made meaningfully, which gives us the opportunity to shed light on this controversy.

Importantly, our experiment is not about High Frequency Trading (HFT) per se, but about automated trading. Whether a chosen robot trades frequently depends on its nature (in our case, the reference point relative to which it trades). Early robots-only financial markets experiments have shown that the most profitable algorithms are low-frequency, "parasitic" ones [the so-called "Kaplan sniper," Rust, Miller, and Palmer (1994), and LeBaron, Arthur, and Palmer (1999)].

Our setting is that of Smith, Suchanek, and Williams (1988). There, bubbles are regularly observed, and are robust to a number of interventions, such as ability to short sell, access to futures markets, or inability to re-sell. Professional traders do not generate less mispricing (see King, Smith, Williams, and van Boening (1993) for a study using corporate executives and professional traders, and Palan (2013) for a general survey). There are only a few ways to avoid bubbles altogether: repeat the session with exactly the same participants, expose participants to excruciatingly detailed instructions, or starve the market of cash. (For a review, see, e.g., Powell and Shestakova (2016).)

We ran 4 sessions with robots and 4 manual sessions with students from the University of Melbourne and of Monash University. (We report in the Online Appendix on 4 additional sessions at the University of Utah that we ran as a pilot, to test design and software, and which confirm our findings.)

Participants used robots intensely. Between 2/3 and 4/5 of transactions involved a robot. These numbers are not unlike those generated in advanced field (stock) markets (De Luca and Cliff (2011)).

We find that bubbles are no smaller and no less frequent than in our manual trading control, rejecting our hypothesis that access to algorithmic trading induces the "strategy method" in ways that eliminates mispricing. Effective bid-ask spreads (and hence, autocorrelation) were higher with robots only in early rounds. Consistent with historical data from the field, average earnings of robot users were no different from those of manual traders; however, those deploying robots that traded only infrequently earn marginally more (but statistically significantly so). Finally, robots cause more flash crashes (and sudden price surges) in the first rounds of trading.

The remainder of the paper is organized as follows. Section II details the experimental design. Section III presents the key questions. Section IV discusses the results. Section V concludes.

II. Experimental Design

A. Trading Game

We replicated the design in the original bubble experiment (Smith et al. (1988)).

Eight participants were endowed with cash and a certain number of units of a security called stock. The trading game lasted 15 periods. At the end of each period, a unit of stock paid a random dividend. At the end of the game, the stock expired worthless. The length of a trading period was 5 minutes in early periods, and gradually declined to two minutes (minimum) as trading became less active.

Immediately after the end of each trading period, an online random number generator (`http://www.random.org/integers`) was used to determine the dividend, with equal chance of returning 0, 0.25, 0.5 or 1.25 experimental dollars. As such, the expected dividend per unit stock per period was \$0.50.

Under risk neutrality and no discounting, the fundamental value of the stock was \$7.50 in period 1 (=15*0.5), and declined to \$7, \$6.50, \$6, ..., \$0.50 in periods 2, 3, 4, ..., and 15, respectively.

Half of the participants were initially allocated 20 units of the stock and 100 experimental dollars, while the other half received 12 units of the stocks and 160 experimental dollars. With the stock evaluated at fundamental value, both types of participants started with initial portfolios worth 250 experimental dollars.

B. Incentive Structure

At the end of trading, after all units of stock had paid their dividend and expired worthless, participants took home the cash they were holding, after conversion to Australian dollars (AUD). Participants earned cash through (i) endowment of cash, (ii) dividends on the stock, and (iii) buying stock low and selling high. In the main (robot) experiment, conversion took place at 5:1 (5 experimental dollars = 1 Australian dollar (AUD)). In the much shorter control (manual) experiment, conversion was at 10:1. Participants were given an extra 10 AUD, independent of performance.

As such, participants in the robot experiment could expect a compensation of 60 AUD. The actual remuneration varied depending on performance, but was bounded between 40 and 80 AUD. In the manual control experiment, expected compensation was 35 AUD, with a minimum of 15 and maximum of 55 AUD.

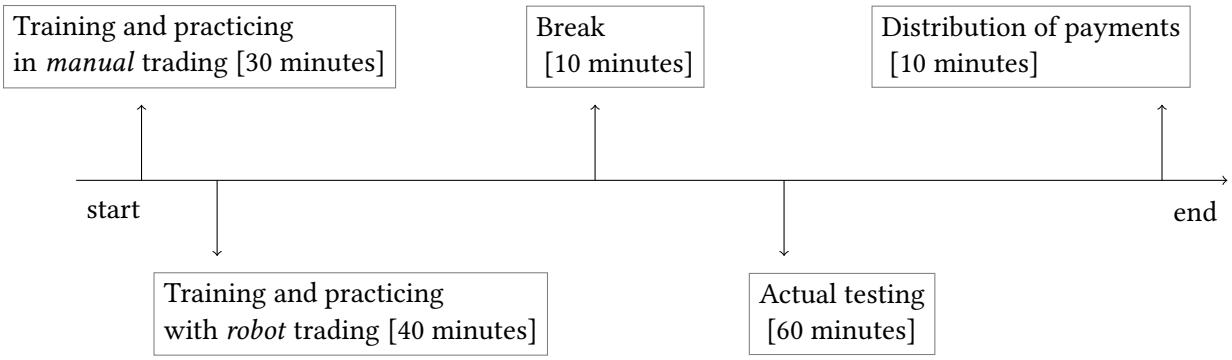


Figure 1. Timeline of the main (robot) experiment.

C. Main (Robot) Sessions (1-4)

The first half of the experiment was dedicated to training the participants, with the actual testing taking place after a brief break in-between. An indicative timeline reflecting the experiment process is depicted in Figure 1.

Training took over one hour, during which we first helped participants familiarize themselves with manual trading on our online trading platform, Flex-E-Markets (see below). After participants demonstrated confidence in manual trading, we instructed them on how to choose, upload, deploy, halt and replace, algorithmic traders (robots). Participants could upload, from their desktop, a variety of robots through an online robot management platform, Algohost (see below). Training in manual and robot trading took place in a setting where a one-period security called stock would pay a random liquidating dividend with expected value 5 dollars. The use of a security different from the one that is used during the actual game is motivated from the findings in the literature that the mispricing disappears when the same situation is repeated with exactly the same participants. Thus, the training only had the purpose to help participants in dealing with the trading system and the algorithmic programs, not with the possible issues of mispricing in the Smith et al. (1988) setup.

D. Control (Manual) Sessions (5-8)

We conducted four control sessions where manual trading was permitted only. The trading environment in these four sessions was identical to the one in the main (robot) experiment, except that participants did not have access to robots. Learning is far quicker, and hence, compensation was reduced (see above) for the manual trading environment.

E. "Session M" (9)

"Session M" was initially intended to be conducted with robot trading, as in Sessions 1-4. But due to an unforeseen technical adversity, the robots were not functioning properly after completion of training. Instead of terminating the experiment early, we asked all participants to complete the trading game with

manual trading only. This accident allowed us to answer an interesting question: what would the market be like when traders are trained to use robots but only allowed to trade manually?

F. Pilot Sessions (10-13)

In the Online Appendix (OA), we report on 4 additional sessions we ran in 2014 at the University of Utah in order to test the software and the experimental design. Results are qualitatively the same (e.g., we recorded a clear bubble in 3 out of 4 sessions), though we had no immediate way to tell whether an order was submitted manually or through a robot. Robots were uploaded into (were collocated with) the market server, which could have been exploited by savvy participants who knew python (the language for the robot scripts), so we ended up moving robots to another server (Algohost) which provided us with the ability to differentiate between manual orders and orders placed by an algorithm.

G. Trading Mechanism

The online marketplace was organised as continuous, double-sided open book system, like most of the real-world electronic financial markets. Only limit orders could be used, and strict time/price priority was adhered to. Order submission and trade reporting was anonymous. Among others, this meant that participants could not know whether their counterparty was a human trader or a robot. An illustration of the user interface is provided in Figure 2 (Left Panel).

The trading mechanism is part of online software called Flex-E-Markets, which provides software as a service (SaaS) platform to organize and manage multiple, simultaneous online marketplaces, and which, at the time of writing, was freely available for academics. See `http://www.flexemarkets.com`.

H. Robot Interface

On the desktop of the participant's computer was folder with python scripts. Each script corresponded to a particular robot that the participant could upload onto an algorithmic trader hosting server called Algohost (`http://algohost.bmmlab.org`). This server facilitated the interaction between the robot and the market platform, Flex-E-Markets. To use a robot, a participant merely had to upload it to Algohost, and once deployed, its launching, pausing and deleting could be controlled by simply clicking on the corresponding button. See Figure 2, Right Panel.

I. Robots

Robots (computer scripts that automatically execute trading algorithms) were explained in detail and the usage was demonstrated with examples, after participants had familiarized themselves with manual trading.

Robots could be classified into several categories, depending on function, role, and benchmark FV. As to function, a robot could either be a market maker or a "reactionary" robot. In market microstructure terminology, the former are liquidity providers; the latter are liquidity takers (Bouchaud, Gefen, Potters, and Wyart (2004)). The roles of a robot could be: buyer, seller, or buyer-and-seller, denoted by BUY, SELL,

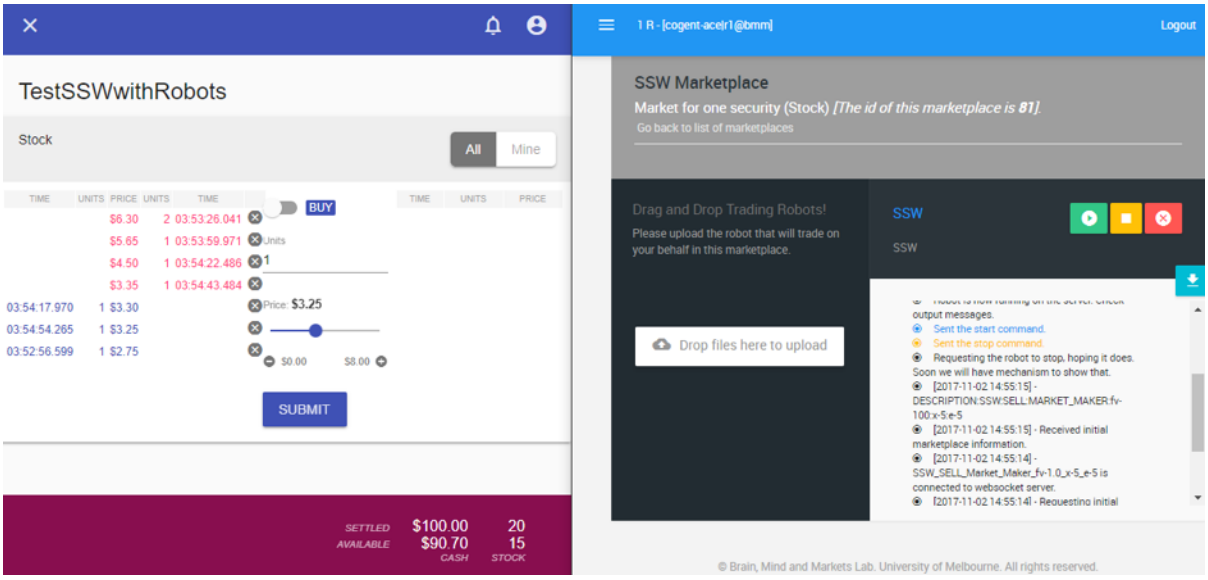


Figure 2. User interfaces. In Flex-e-Markets (Left), participants observe the (open) book of buy (blue) and sell (red) orders, submit/cancel orders, and trade with each other. In Algohost (Right), participants upload one of a range of available robots, start (green button), pause (yellow button) and disengage (red button) the robot, and receive real time information on its communication with the market, including submission of orders and reporting of trades.

or BUY/SELL robots. The benchmark FV is the "fair value" FV that a robot uses to determine what orders to place. FVs ranged from \$0.25, in increments of \$0.25, to \$9.75. Thus, a robot is a triple (function, role, FV).

- *Market maker robots*: Once started, they submit a buy order for one unit at 5 cents below their FV, or a sell order at 5 cents above their FV, or both, depending on their role. For example, a $(market-maker, BUY\&SELL, \$5.00)$ robot would submit a buy order at \$4.95 and a sell order at \$5.05. Once traded, the orders are replenished immediately, until the robot is paused, or runs out of either cash or stock. The $(market-maker, BUY, \$5.00)$ and the $(market-maker, SELL, \$5.00)$ robots submit only the buy/sell orders respectively
- *Reactionary robots* record incoming buy and sell limit orders, and only react to orders that are advantageous compared to their FV. For a reactionary BUY robot, any sell order at FV minus 5 cents or less would trigger the placement of a buy order at its FV. On the opposite trading side, a reactionary SELL robot reacted to any limit buy order at prices equal to FV plus 5 cents or more, by sending a sell order at the FV. E.g., if a reactionary BUY robot with FV equal to 5 dollar observed an incoming ask for 3.15, then it submitted a buy order at 5 dollars. (Trade will take place at 3.15 provided nobody else picks up the order beforehand).

J. Instructions

The version of instructions given to participants in the main (robot) sessions is included in the OA (see B). The instructions for the control (manual) sessions are a reduced version of this, where all information

pertaining to robots is eliminated.

K. Participant Recruitment

In total 40 participants (5*8) were recruited from the University of Melbourne community, through online website and on-site advertisements. An additional four sessions that allowed only manual trading were conducted at Monash University. Eligible participants were healthy and willing people within the age group of 18 to 35. The recruited sample had an average age of 22.57 years (age range: 19 to 34 years, standard deviation = 2.92). Gender composition was reasonably balanced: overall 35% male and 65% female. 62.5% of the participants were enrolled in finance-related courses, and 20% of them were in engineering/computing courses, 10% were studying in other analytically-oriented fields, while the remaining 7.5% were specializing in non-analytically oriented areas. Out of all participants, there were 2 PhD students, 4 in the their honours year of finance (equivalent to Master's) and the rest were evenly spread across bachelor and masters degrees.

The experiment was approved by the Human Research Ethics Committee at the University of Melbourne (Ethics ID: 1749612.1). Rudimentary demographic information for each participant was collected, together with a written informed consent form, in accordance with ethics rules.

III. Key Questions

We focus on four key questions. The first concerns emergence of mispricing.

Will bubbles emerge?

Our expectation is that they may not emerge, because the availability of robots forces all participants to think in terms of the "strategy method." As discussed before, robot use forces participants to commit to a strategy from the moment the robot is launched until it is stopped. As such, participants need to formulate expectations about the evolution of future prices. This is further reinforced by the type of robots made available: all execute order submission strategies relative to a fixed reference value (the FV), which can be changed only when the participant uploads a new robot (with a different reference FV).[2]

What are reasonable prices that participants could entertain as basis for their own or the expectations of others' choices of robot? Participants may expect mispricing, but it is not a priori obvious in which direction (experimentalists know that in the majority of cases, prices start below fundamental values, but it is unlikely that participants have access to this knowledge, being in a setting they have likely never seen before). As a result, correct pricing (prices equal fundamental values) can be argued to be the "focal" benchmark point. In addition, when expecting correct pricing, the resulting position minimizes risk from forecast mistakes; that is, correct pricing constitutes a "risk minimizing" expectation. Therefore, it was reasonable for us to expect that robots would lead to less or even no mispricing, compared to experiments with manual trading only.

The next question concerns liquidity, measured in terms of the effective bid-ask spread.

Will the effective bid-ask spread in the main treatment be lower than under the control, manual trading?

Evidence from historical data from the field seems to indicate that algorithmic trading increases liquidity, and hence, reduces the bid-ask spread. However, there could be many confounding factors. Introduction of algorithmic trading in the field is never exogenous, and reduction of bid-ask spread was already a trend before the advent of algorithmic trading, as a consequence of other causes, such as regulatory changes and increased competition in the brokerage industry, availability of real-time online trading, etc. (Chordia, Roll, and Subrahmanyam (2001)).

There is a fundamental reason that a lower bid-ask spread is not a foregone conclusion, despite the fact that robots can trade faster, and hence, more (in principle). We already conjectured a relationship between choice of robots and the "strategy method" in game theory. Choice of the right strategy (and hence, the right robots) requires participants to think hard about which equilibrium others will play (one with correct pricing, or with under-pricing, or maybe over-pricing?). In simple 2×2 games this is known to lead participants to play the risk-minimizing equilibrium strategies (Embrey et al. (2017)). In complex games, such as our online open-book market, it is not obvious which equilibrium participants will coordinate on, although we did argue before that they could be coordinating on the ones where prices follow the fundamental value. Mis-coordination translates in trades between robots with different reference values, and hence, larger effective bid-ask spreads, until robots are coordinated.

The consequences of mis-coordination are less harmful when there is manual trading only, since participants can change their strategies with no delay. Upon seeing the first order, say a bid at \$10, everyone can immediately adjust expectations and make moves according to any of the strategies consistent with a player making a bid of 10 dollars.[3]

As such, our expectation was that the effective bid-ask spread would actually be higher with robot trading, at least in the early rounds of trading.

Mis-coordination between robot choices would also lead to wild price swings (Biais et al. (2011)), which brings us to our third question.

Will robot usage lead to more flash crashes/surges than manual trading?

The evidence from historical data from the field, as mentioned before, seems to suggest that robots do not increase the incidence of flash crashes (rarely does one deal with surges), though this conclusion is controversial. Based on the coordination argument, our expectations were different–unless the algorithms coordinated on the actual FV, large price swings would be expected in the process of coordinating.

Will traders with higher usage of robots also have higher earnings?

The last question concerns allocational efficiency. If we consider all participants to be identical ex ante, there is little to be said. In fact, one would expect no trade, beyond perhaps initial trade, whereby the (risky) asset is re-allocated towards more risk-tolerant participants.

Here, we focus on impact of earnings because of the use of robots. The literature cited before seems to indicate that, in the field, robot use historically has not increased average earnings for traders who have access to algorithmic trading (though one could expect top earners to significantly outperform others).

We shall compare earnings of participants who used robots against those of participants who chose not to use robots.[4] In addition, we shall look at earnings gains among robot users, in an attempt to identify which type of robots increase earnings.

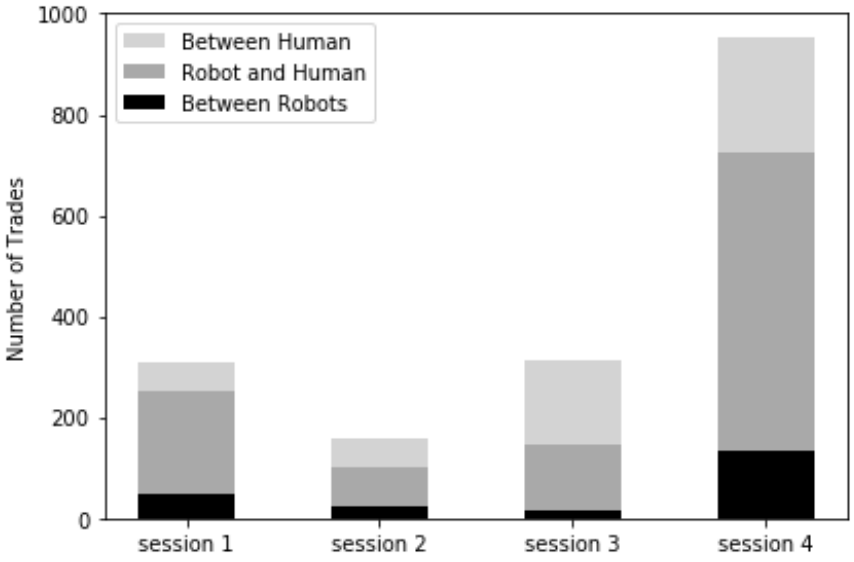


Figure 3. Composition of trades in each session.

IV. Results

A. Descriptive Statistics

A.1. Robot Activity

We observe ample interactions between humans and robots in all four experimental sessions. The percentage of trades involving robot on one side and human on the other are 71%, 56%, 45% and 67%, for sessions 1 to 4 respectively. See Figure 3. These numbers closely resemble the situation in major European and US equity exchanges, where the proportion of robot-executed trades is known to be within the range of 30% to 70% (De Luca and Cliff (2011); Goldstein et al. (2014)).

A.2. Earnings

The final performances of participants in the robot sessions (sessions 1 - 4), the manual control sessions (sessions 5 - 8) and the "M" session (session 9) are plotted in Figure 4, the performances in manual sessions (number 5 - 8) are appended on the right for comparison. The ex-ante expected payoff (value of initial holdings plus cash) is represented by the dashed line. Performance ranges from a maximum of 370 and a minimum of 51 experimental dollars, with mean equal to 248 and standard deviation 70. The lowest individual performance score is recorded in one of the robot sessions, while the highest one appears in one of the manual sessions. The standard deviations for robot sessions separately are 33, 42, 73 and 97. Visual inspection of performance ranges reveals that no distinguishable pattern could be discerned across robot and manual (control) sessions.

A measure of statistical dispersion, the Gini coefficient (Gini (1936)), is used here to indicate the degree of inequality in final wealth distribution among participants. The average Gini coefficient (standard error) for robot sessions is 0.139 (0.034), slightly, but insignificantly ($p > 0.10$) higher than its value of 0.127 (0.016)

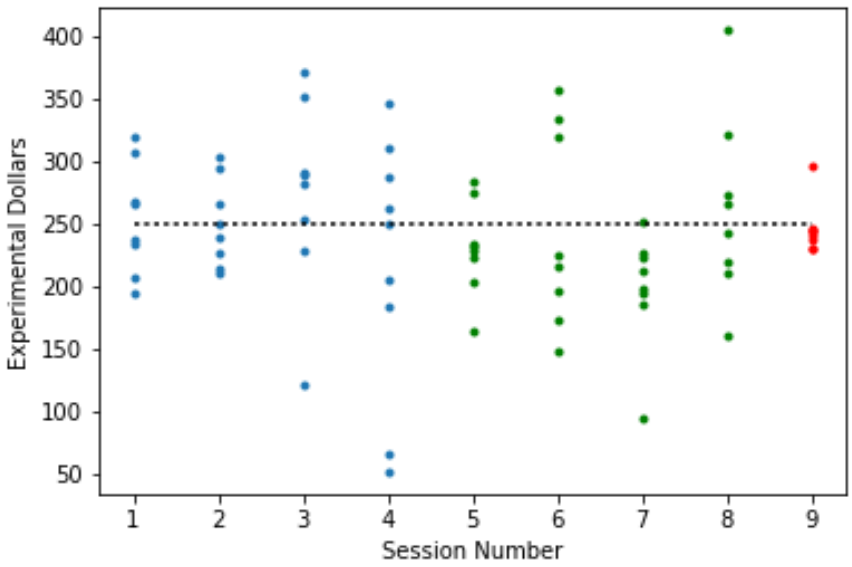


Figure 4. Individual final performance for each session. Total experimental dollars earned in four main (robot) sessions (blue dots), four control (manual) sessions (green dots), and in "session M" (red dots). Dashed line indicates expected payoff.

for the manual sessions. The lowest dispersion of final wealth is recorded in Session "M;" this may have been caused by the fact that this was the only session where prices closely followed fundamental values (see later).

A.3. Choices of Robots

Substantial heterogeneity emerged in participants' choices of robots. Figure 5 provides details.

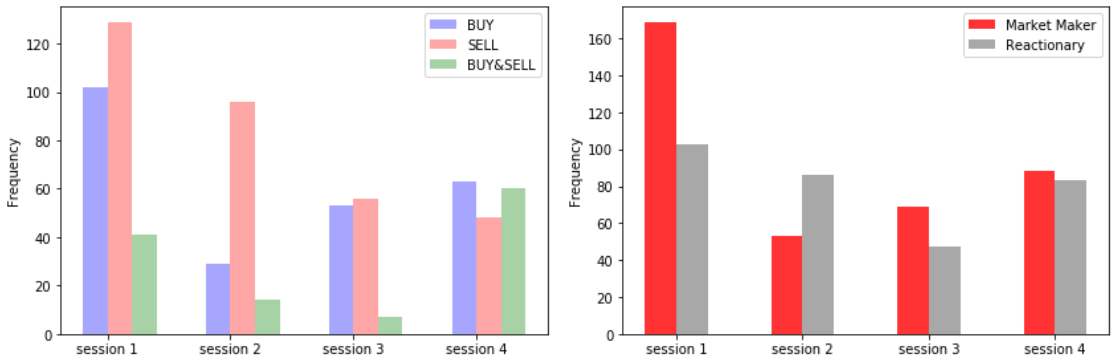
Overall, seller robots and market-making robots were favoured in the majority of the sessions. This means that robots tended to provide liquidity on the buy side. Double-sided robots, which submit both buy and sell offers, were used the least.

The total order submissions were very similar between robot and manual-only sessions ($p > 0.10$), thus the use of robots did not increase the depth of the book or the number of executed transactions.

By correlating robot usage with mispricing magnitude, two interesting observations can be made. First, Session 2 was the only session where reactionary robots (liquidity takers) were used more frequently than market-making robots (liquidity providers). Also, in Session 2, the proportion of seller robots was unusually high compared to other sessions. Session 2 happened to have the lowest level of mispricing among the robot sessions. Second, Session 4 had the highest usage of double-sided robots, and the least frequent use of single-side seller robots. This session also produced the biggest price bubble.

Robot usage, defined by number of robots launched in a period, or average number of active robots in a period, is higher in the first half of the experiment in comparison to the second half.

We now turn to address the four questions we hoped to answer with our experiments.



(a) Robot use by role.

(b) Robot use by function.

Figure 5. Number of times a robot of a particular role and function is launched.

B. Result 1: Bubble Are No Less Frequent with Robots

The average mispricing across the four robot sessions is depicted by a solid red line in Figure 6; dashed red lines delineate the region of two standard errors relative to the mean. We observe substantial under-pricing in the starting period; at 2.91 dollars, the average transaction price was just over half its fundamental value (7.5 dollars). Starting from around the 7th period, average mispricing became positive, while the standard errors increased as well. Over-pricing on average remained flat across periods 8 to 15, at about 1 dollar per share. In the 15-th period, the average overpricing reached 1.35, which exceeded the maximum possible value of the single dividend still due (1.25 dollars).

The average mispricing of the four manual sessions (black solid line in Figure 6) exhibited more moderate under-pricing in the earlier periods, but the bubble emerged earlier (on average around period 5). Average mispricing peaked in period 10, to subsequently drop to close to zero by the last trading period. There was substantial heterogeneity across sessions, so that the main difference between robot and manual session was limited to more underpricing in early periods in robot sessions. Once the bubble emerged, no more significant differences were discernible.

C. Result 2: Effective Bid-Ask Spreads Are Larger in Early Periods for Robot Sessions

The effective bid-ask spread is measured from the autocovariance of consecutive transaction price changes (Roll (1984)). We use the effective bid-ask spread since volatility tended to be higher with robot trading, especially in early rounds (see OA), and with higher volatility, there is a corresponding need to continuously adjust quoted bids and asks, and the resulting synchronicity issues cause the quoted bid-ask spread to over-estimate the true bid-ask spread. Underlying Roll's estimate is the assumption that the fundamental value of the asset is a martingale, and that there is no asymmetric information. Both assumptions are appropriate in our setting: fundamental values are a trivial martingale (they are constant within a period), and if participants entertained other price expectations, it is hard to imagine that they would consider anything but a martingale; likewise, the assumption of no asymmetric information is satisfied as well, since everyone was given the same information about dividends, and this was common knowledge.

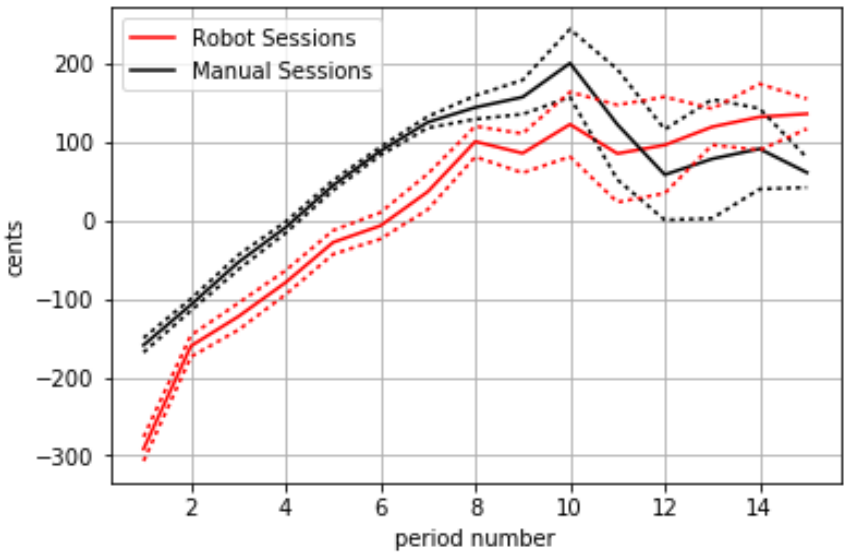


Figure 6. Mispricing. After initial under-pricing, bubbles emerge in both robot sessions (red) and manual sessions (black). Dashed lines delineate 95% confidence intervals.

Figure 7 displays the per-period evolution of the average effective bid-ask spread (in cents). Asterisks indicate whether bid-ask spreads are significantly different ($p = 0.05, 0.01$) between robot and manual sessions.[5] At $p = 0.01$, the effective bid-ask spreads are not significantly different across the two treatments, except in periods 1, 2 and 7. Economically, the effective bid-ask spread is substantially higher in early rounds with robots compared to manual trading. Effective bid-ask spreads are higher in later periods in both treatments, coinciding with increased probability of a substantial price adjustment (during the inevitable crash).

We attribute the higher effective bid-ask spreads in early rounds of the robot session to the "strategy method" setup. When deploying a robot, the participant is committing, at least temporarily (until the robot can be halted), to a strategy. Because participants could entertain different expectations about future price evolution, they would choose different strategies, and hence, different robots. Mis-coordination of strategy choices, and the resulting need to halt robots to re-align strategies, leads to higher bid-ask spreads. As we shall see next, it also leads to a higher number of flash crashes/surges in early rounds.

D. Result 3: Robots Cause Flash Crashes/Surges in the First Trading Period

Figure 8 displays, per period, the frequency of flash crashes (transaction price decreases) and flash surges (transaction price increases or bubbles), stratified by Treatment (robot/manual-only). Flash crashes/surges were defined as outliers in log price changes (returns) that were more than two standard deviations away from the sample mean for the entire session.

Curiously, the number of flash crashes/surges we identified in this way, 48 (for the robot sessions) and 38 (for the manual session), were less than half the frequency we expected under the Gaussian distribution.[6] If we extend the cut-off in the definition of an outlier beyond two standard deviations, typical features of

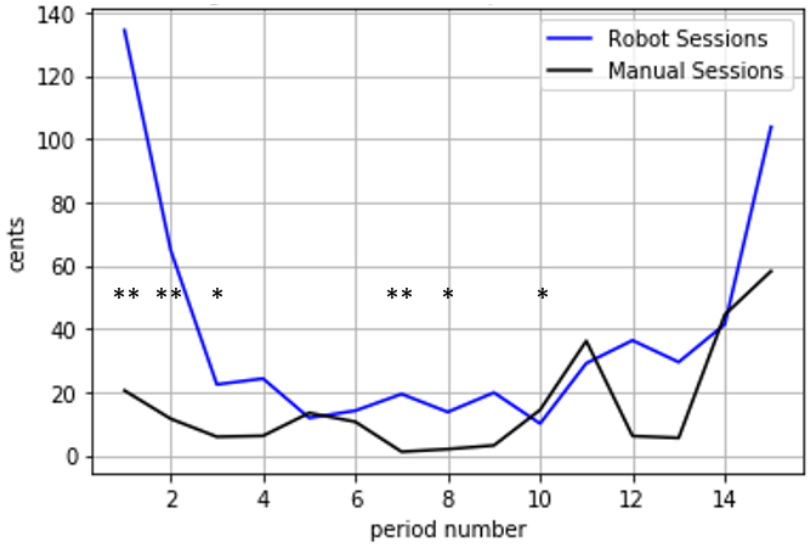


Figure 7. Average Effective Bid-Ask Spreads. Computed from estimated autocovariances of trade prices. Significance levels based on two-sample test of difference in autocorrelations (*: $p < 0.05$; **: $p < 0.01$; two-sided)

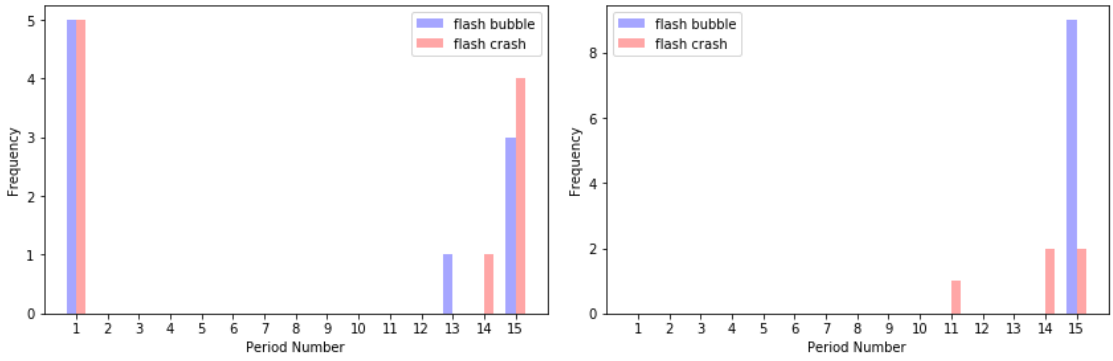


Figure 8. Frequency of flash crashes/surges (bubbles).

a leptokurtic distribution emerged. See the boxplots in the OA.

We expected to record wild price behaviour in later periods, coinciding with the bursting of the bubbles. Indeed, in manual and robot sessions alike, many flash crashes/surges occur after period 10. Notice that bubbles don't necessarily simply crash; prices at times drop and then increase again. That is, crashes are accompanied with high volatility. These wild gyrations are not unique to robot sessions (in which case they may be attributed to participants inability to switch off buy robots that still have higher reference points). In period 15, we recorded 9 flash surges across all manual sessions.

The robot sessions stand out, however, in that they generated many flash crashes/surges in period 1. We attribute this, once again, to mis-coordination between robots. Figure 9 illustrates the phenomenon for the first robot session. Shown are: (i) the evolution of transaction prices, fundamental values, and maximum possible values (based on the maximum amount of dividends remaining to be paid) (Top); (ii) Total number of robots *launched* each period (Bottom Left); (iii) Average number of *active* (i.e., engaged)

robots per second (Bottom Right). The average number of active robots fluctuated in narrow bounds between 3 and 5. However, at 35, the number of robots started was highest in trading period 1, and dropped subsequently, to settle between 10 and 15 after period 5. Hence, in early periods, many robots were started, but stopped very quickly, and replaced, so that only 5 were active at any point in time. The robot switches led to huge price volatility (see Top).

Evidently, participants struggled to deploy the right robots in early periods. This leads to huge price swings and hence, flash crashes/surges (Figure 9). We attribute this struggle to strategic uncertainty. Participants' robots were sufficiently mis-coordinated (reflected very different pricing expectations) to cause huge volume at wildly varying prices. By the second period, robot choices were better coordinated, and flash crashes/surges disappeared. By period 5, many of the 15 robots deployed remained active for longer periods of time. The result is that no fewer robots (5) were active on average.

These results provide suggestive evidence of our main conjecture, that robot choice is akin to asking participants to play the trading game using the "strategy method," and hence, that strategy coordination failures cause anomalous price behaviour.

E. Result 4: Robot Use Increases Performance Only Marginally and Only for Participants who Also Trade Manually

Finally, we investigate earnings. We look at two measures of robot use: (i) Number of robots used, calculated as number of times a participant started a robot (including re-launching a robot previously deployed); (ii) Percentage of trades completed by robots, calculated as the per-participant frequency that a trade originated with the participant's robot rather than manually.

No significant correlations were found between either metric and performance. However, a regression where participants' total number of orders submitted is controlled for, both measures contributed significantly to explaining earnings. In particular:

- Number of robots used increased earnings: starting an extra robot led to a 2.3 cents increase in final performance ($p < 0.05$);
- Percentage of trades completed by robots reduced earnings: a 50 percentage points increase in the fraction of individual trades completed by robots rather than manually decreased earnings by 78.6 cents ($p = 0.01$).

Altogether, this implies that robot use does increase earnings, but only if robots are used alongside manual trading. Evidently, participants who know when to deploy robots and when to trade manually generate higher earnings.

Consistent with historical data from the field, robot use per se does not increase earnings; only judicious use of robots does. One way to interpret our results is that skill that allows one to deploy the right robots at the right time also makes one perform better when trading manually. Alternatively, the results could be interpreted as suggesting that low-frequency robots (robots that don't trade significantly more frequently than manual traders) are superior. The latter is reminiscent of results obtained in early experiments in financial markets experiments populated solely with robots (Rust et al. (1994) and LeBaron et al.

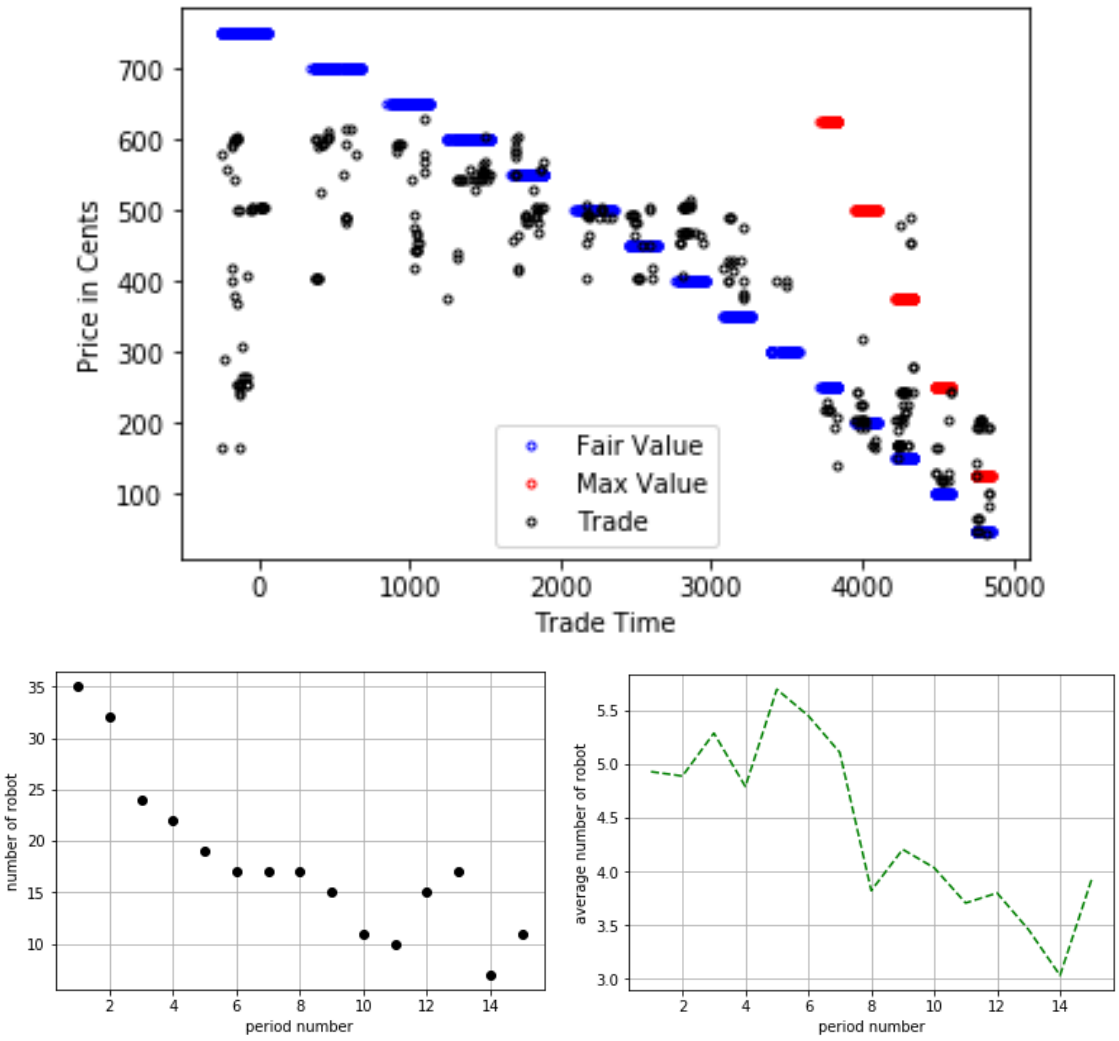


Figure 9. Prices and robot use in first robot session.

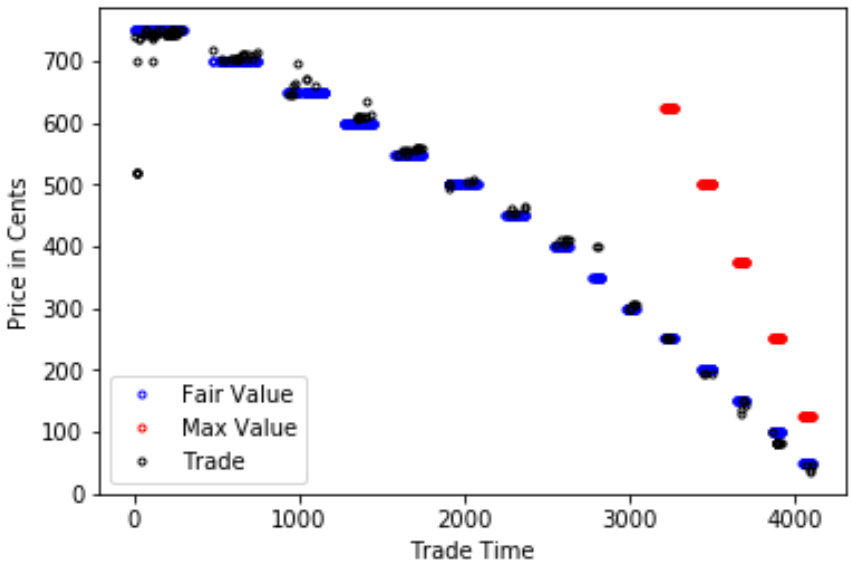


Figure 10. Prices in Session "M".

(1999)).

F. Session "M"

Figure 10 displays the evolution of prices, fundamental values and maximal values in Session M. Here, participants were fully trained in the use of robots, but just before the actual session started, the robot-markets interaction failed, and we switched to manual trading only.

Surprisingly, pricing remained remarkably close to fundamental values throughout the session. Only a few trades, early on in period 1, where off. Notice that these trades occurred at prices far below fundamental value, as in all other sessions.

This finding is suggestive of the following. We conjectured that the availability of robot trading forces participants to think more strategically, i.e., to entertain possible behavioral strategies rather than single moves (the "strategy method"). While a strategy built around the expectation that prices follow fundamental values is risk-minimizing, it is possible that many participants entertained other price expectations, such as ones where prices start significantly below fundamental values (which is consistent with reality). In sessions where robots could be launched, mis-coordination between strategy choices leads to extreme volatility, as evidenced in Figure 9. In Session M, however, participants could re-evaluate strategies after each move: since they did not have to commit to a robot, but had to trade manually, every move in the marketplace allowed one to revise one's opinion of future price evolution, eventually coordinating on the risk-minimizing strategy, which is to trade expecting prices to remain close to fundamental values.

Because prices remain closer to fundamental values, there is less dispersion of final wealth, as reported earlier. The lower wealth dispersion also illustrates that the equilibrium where everyone expects prices to be correct is risk minimizing.

V. Conclusion

We studied robot choice and pricing in a canonical multi-period markets experiment which reliably generates mispricing when participants can only trade manually. We hypothesized that the availability of a range of trading robots would force participants to explicitly consider future price scenarios. Participants who select to deploy robots would choose algorithms that are consistent with fundamental values, and participants who opt to trade manually would expect robots in the marketplace to trade with reference to the fundamental values. Hence, we expected less mispricing when participants have access to algorithmic trading.

Rejecting this conjecture, no reduction in mispricing was observed. We find that while robots were used extensively when available, robot usage was a substitute, not a complement to manual trading. While average earnings did not increase for participants who used robots, they were significantly higher for participants who deployed robots in ways that kept manual trading a significant fraction of total trading.

We find that the effective bid-ask spread, based on Roll's measure, is significantly higher in the early rounds of trading in the robot sessions. This is consistent with our conjecture that the robot trading leads to mis-coordination in the early periods.

In the presence of robots, flash crashes/surges were recorded in the first trading period (across the four robot sessions). We attributed this to failed attempts to coordinate robots to similar reference values in early stages of the trading game.

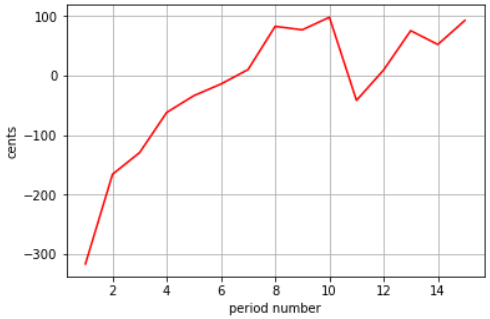
We were most surprised to find that mis-pricing was not reduced with robots. The finding is significant because participants had to choose robots (or had to anticipate others' choosing robots) that submitted orders relative to a pre-determined reference point. The most obvious reference point to choose was the fundamental value.

Our experiments demonstrate that it is possible to study robot use in financial markets in a controlled setting. This is important, because studies that focus on archival data are limited to correlation analysis, and may be biased because of a host of confounding factors that cannot reasonably be taking into account.

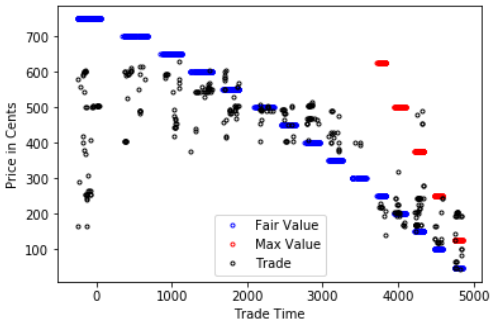
Appendix A. Online Appendix

Appendix A. Trade Prices and Mispricing in Robot Sessions [session 1 - 4]

[Session 1]

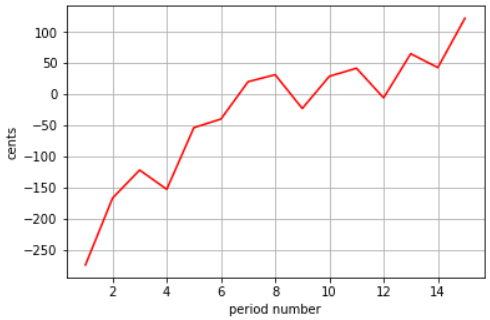


(a) Trade prices, Fair values and Max values.

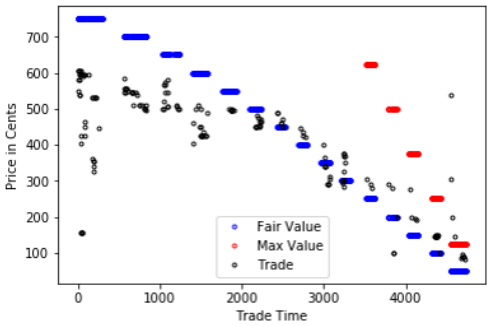


(b) Average mispricing for each period.

[Session 2]

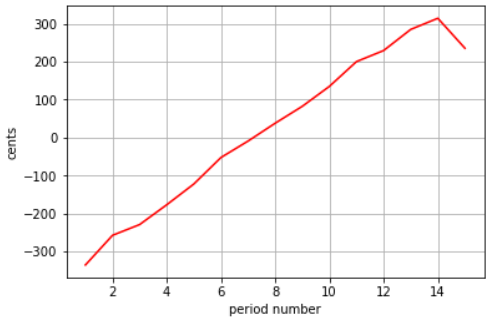


(c) Trade prices, Fair values and Max values.

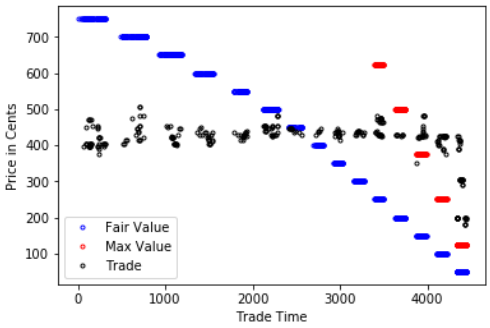


(d) Average mispricing for each period.

[Session 3]

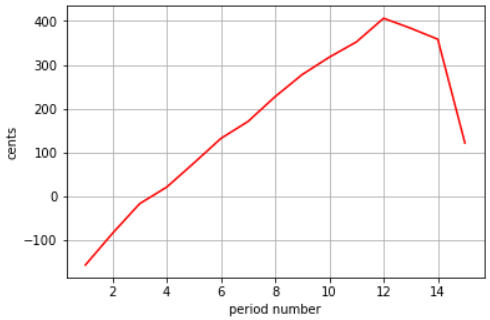


(e) Trade prices, Fair values and Max values.

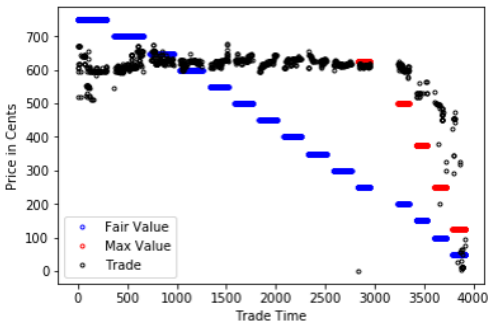


(f) Average mispricing for each period.

[Session 4]



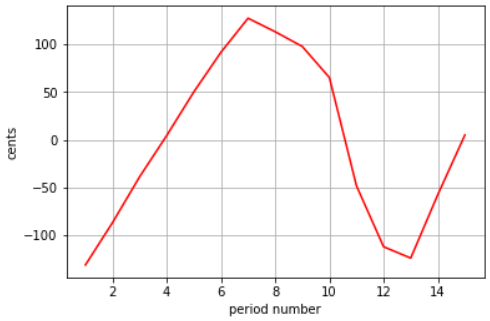
(g) Trade prices, Fair values and Max values.



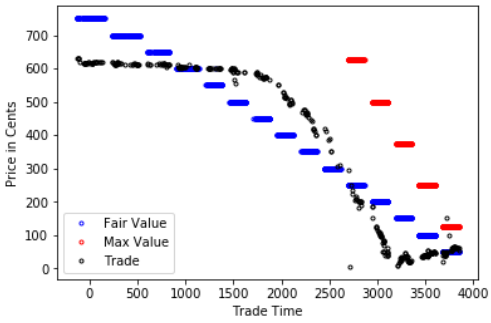
(h) Average mispricing for each period.

Figure 11. Trade prices and mispricing for robot sessions.

[Session 5]

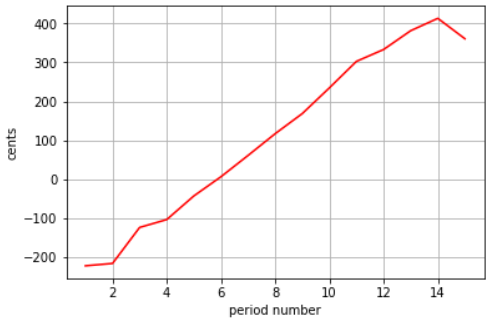


(a) Trade prices, Fair values and Max values.

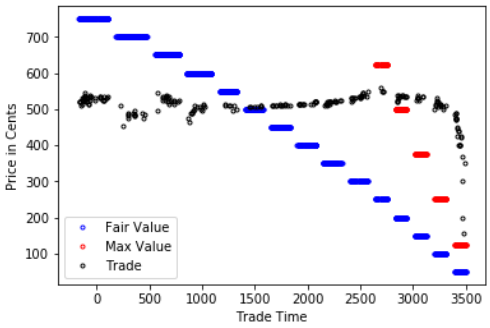


(b) Average mispricing for each period.

[Session 6]

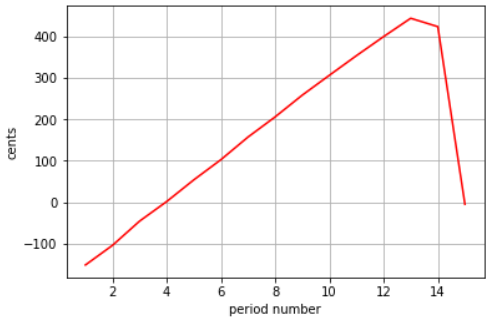


(c) Trade prices, Fair values and Max values.

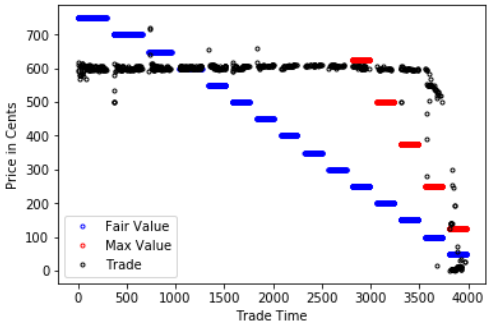


(d) Average mispricing for each period.

[Session 7]

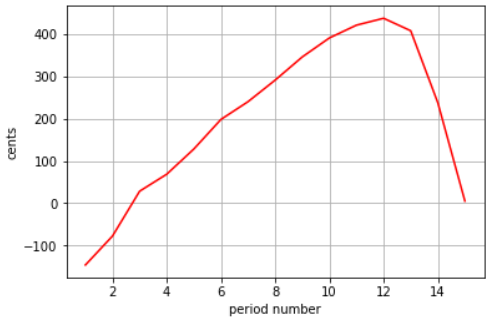


(e) Trade prices, Fair values and Max values.

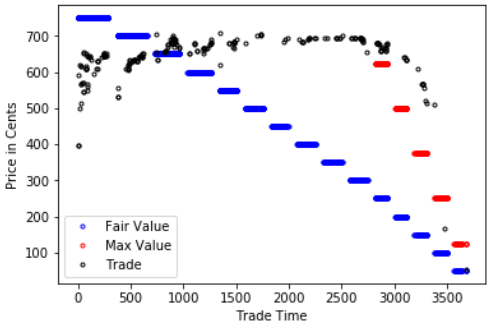


(f) Average mispricing for each period.

[Session 8]



(g) Trade prices, Fair values and Max values.

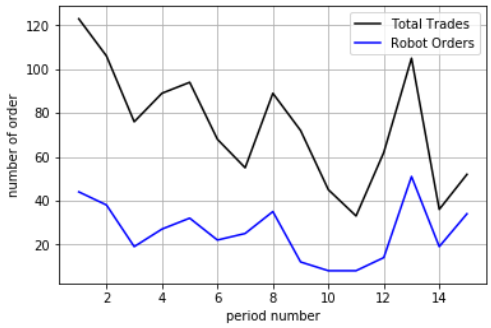


(h) Average mispricing for each period.

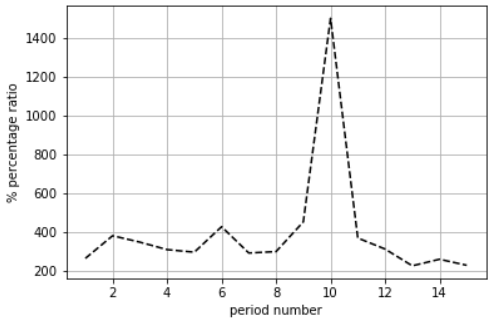
Figure 12. Trade prices and mispricing for manual sessions.

Appendix C. Total Order, Robot Order and Order-to-Trade Ratio for each period [session 1 - 4]

[Session 1]

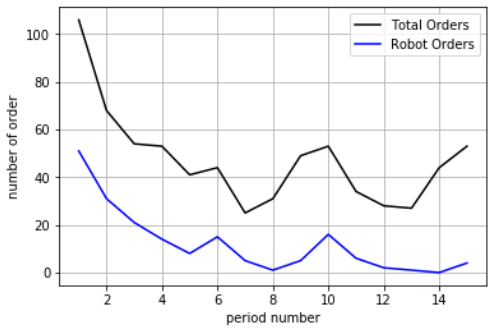


(a) Total orders vs robot orders for each period.

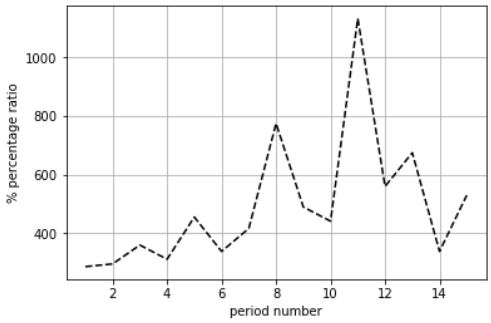


(b) Order to trade ratio for each period.

[Session 2]

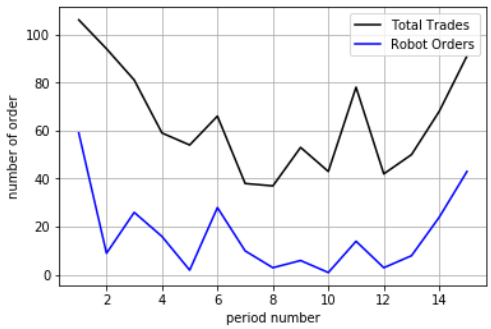


(c) Total orders vs robot orders for each period.

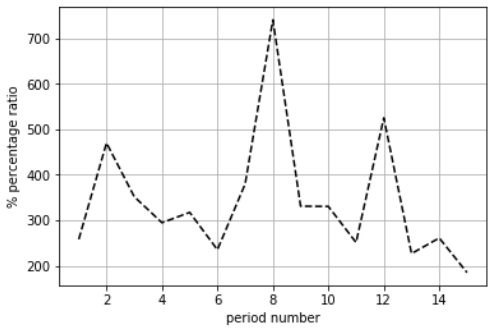


(d) Order to trade ratio for each period.

[Session 3]

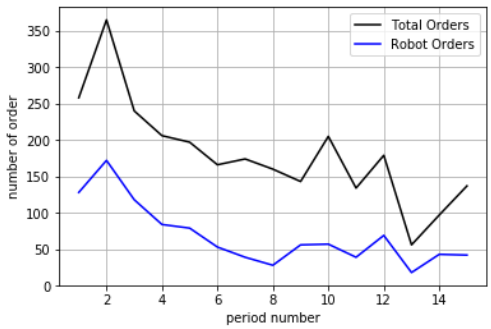


(e) Total orders vs robot orders for each period.

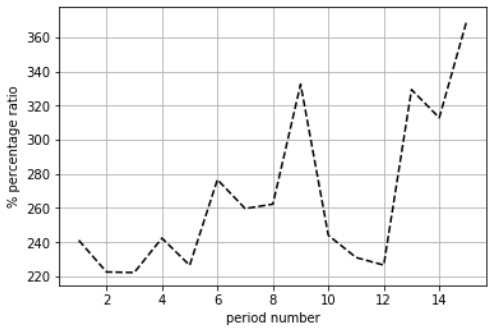


(f) Order to trade ratio for each period.

[Session 4]



(g) Total orders vs robot orders for each period.

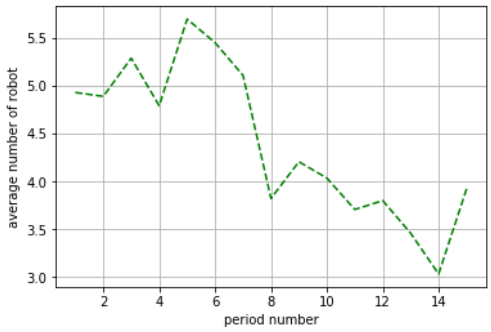


(h) Order to trade ratio for each period.

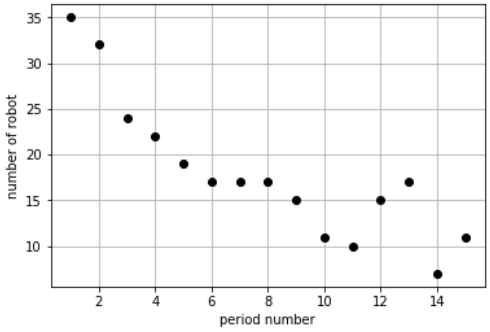
Figure 13. Trade prices and mispricing for manual sessions.

Appendix D. Robot Usage in each Period [session 1 - 4]

[Session 1]

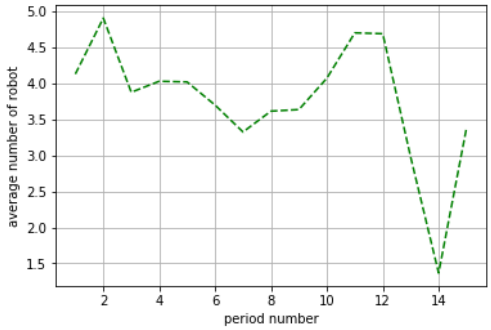


(a) Average active robots per second for each period.

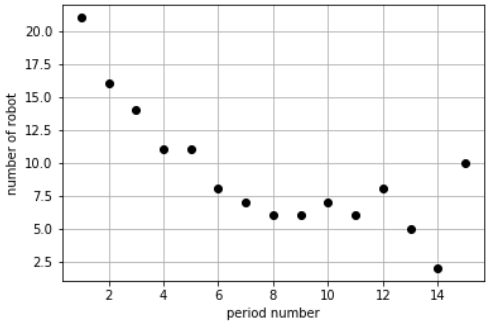


(b) Total robots started in each period.

[Session 2]

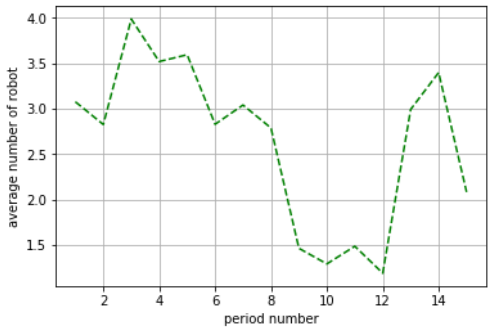


(c) Total orders vs robot orders for each period.

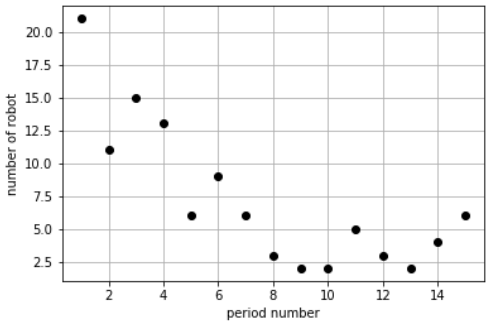


(d) Order to trade ratio for each period.

[Session 3]

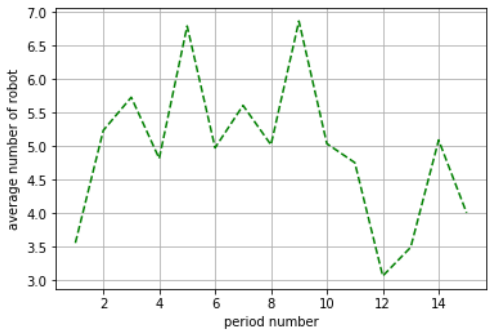


(e) Total orders vs robot orders for each period.

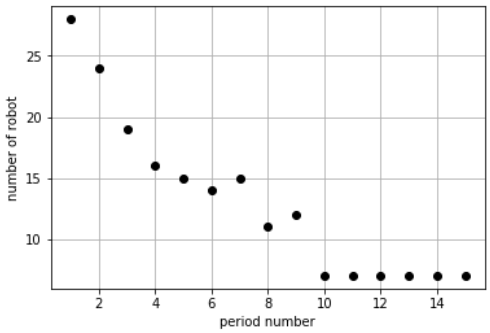


(f) Order to trade ratio for each period.

[Session 4]



(g) Total orders vs robot orders for each period.

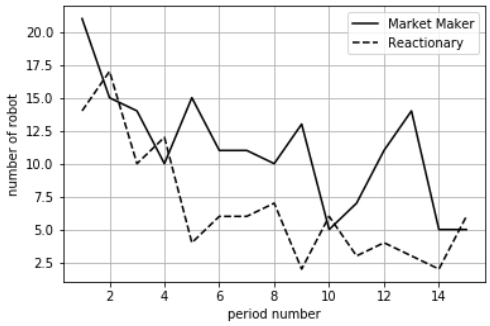


(h) Order to trade ratio for each period.

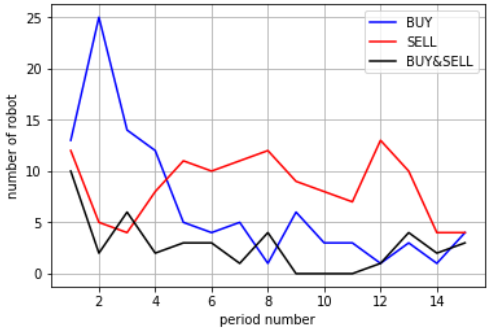
Figure 14. Robot usage in sessions.

Appendix E. Robot Usage by Types and Roles in each Period [session 1 - 4]

[Session 1]

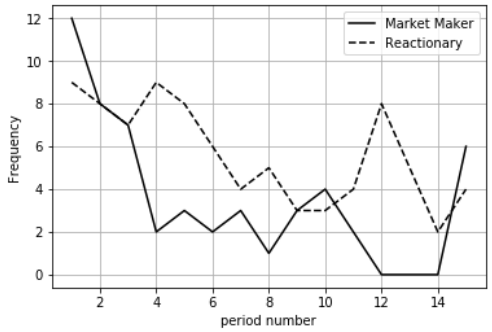


(a) Type of robots started in each period.

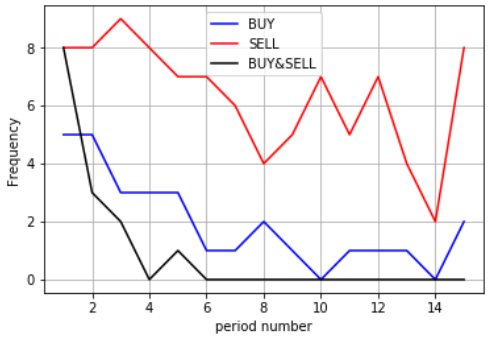


(b) Number of Buy, Sell, Buy&Sell robots in each period.

[Session 2]

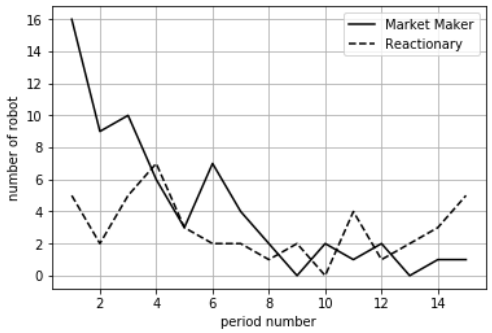


(c) Type of robots started in each period.

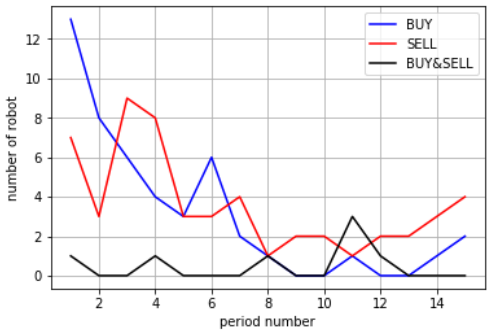


(d) Number of Buy, Sell, Buy&Sell robots in each period.

[Session 3]

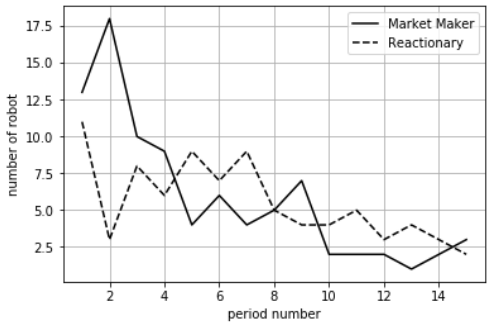


(e) Type of robots started in each period.

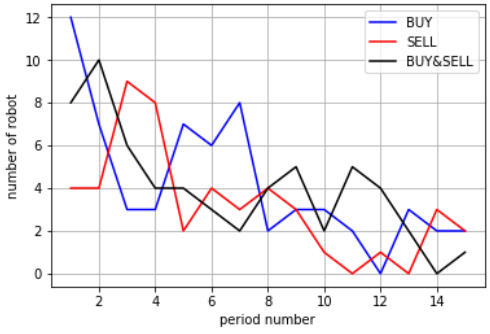


(f) Number of Buy, Sell, Buy&Sell robots in each period.

[Session 4]



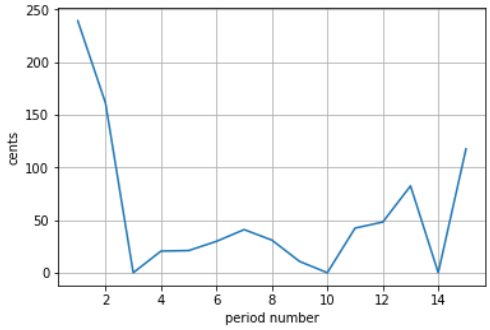
(g) Type of robots started in each period.



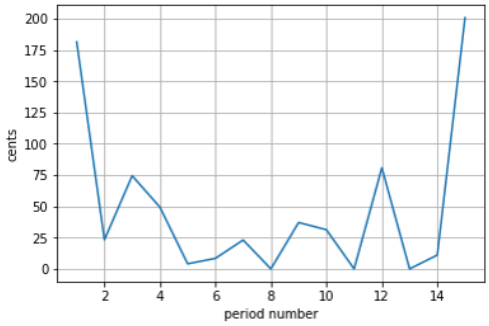
(h) Number of Buy, Sell, Buy&Sell robots in each period.

Figure 15. Robot Usage by Types and Roles in robot sessions.

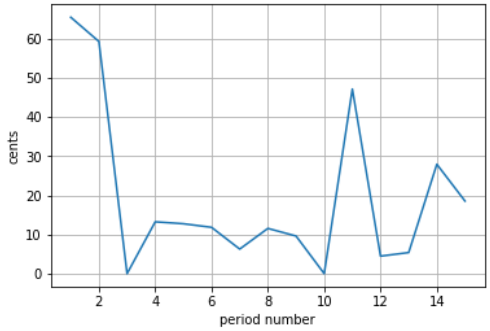
Appendix F. Average Effective Bid-Ask Spread in each Period [session 1 - 8]



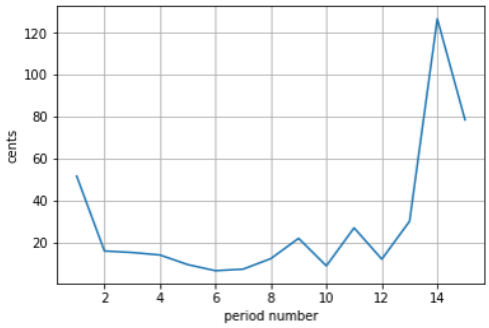
(a) Average bid ask spread for each period [session 1].



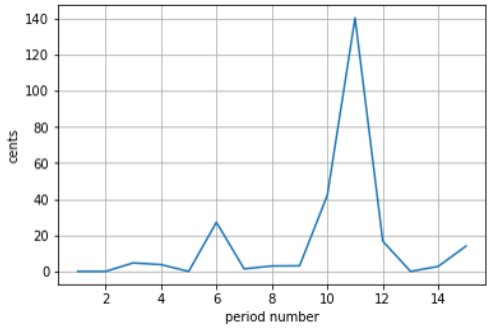
(b) Average bid ask spread for each period [session 2].



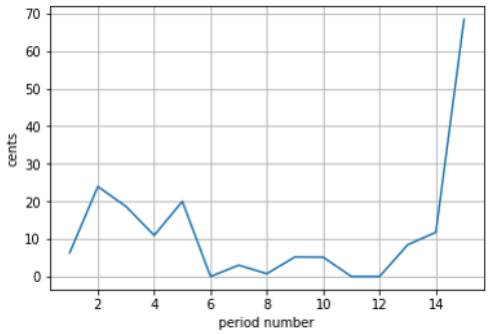
(c) Average bid ask spread for each period [session 3].



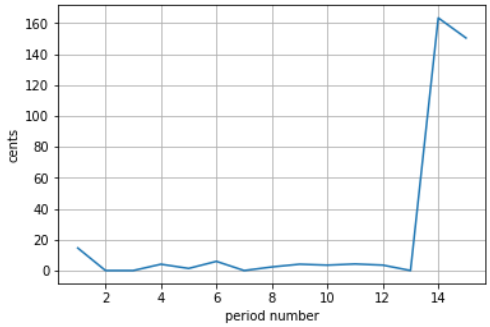
(d) Average bid ask spread for each period [session 4].



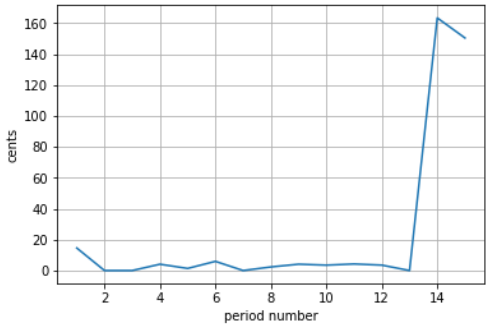
(e) Average bid ask spread for each period [session 5].



(f) Average bid ask spread for each period [session 6].



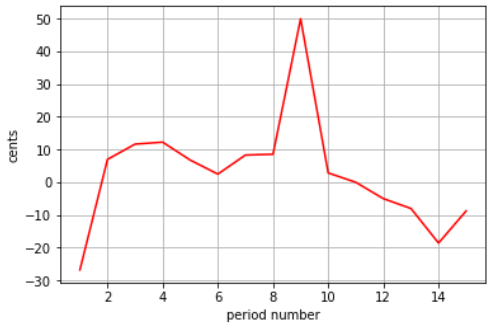
(g) Average bid ask spread for each period [session 7].



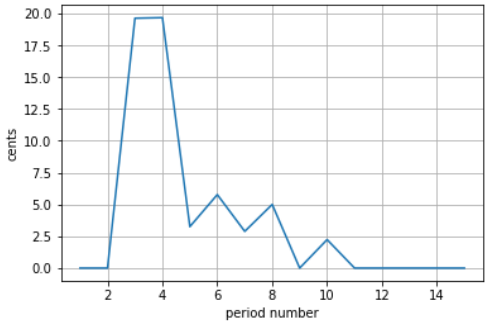
(h) Average bid ask spread for each period [session 7].

Figure 16. Average Effective Bid-Ask Spread in each Period [session 1 - 8].

Appendix G. Trade prices and mispricing for Special Case [Session M]



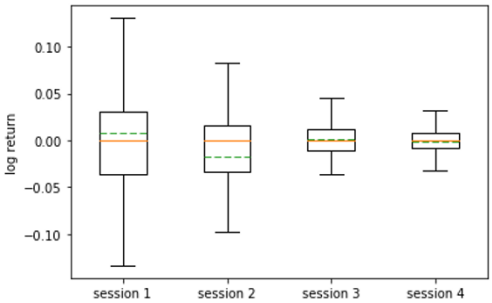
(a) Average mispricing for each period.



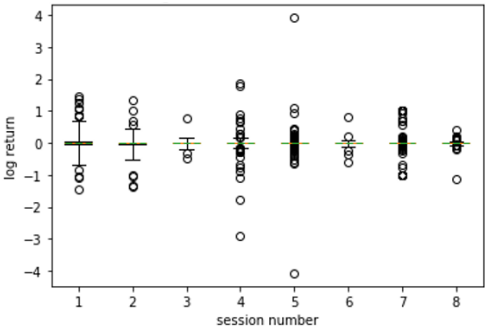
(b) Average bid ask spread for each period.

Figure 17. Average Effective Bid-Ask Spread in each Period for Session M.

Appendix H. Log Returns and Outliers



(a) Box plot of log returns for each robot session.



(b) Outlier of log return for all sessions.

Figure 18. Log returns and outliers.

Appendix I. Results from four (4) pilot experiment sessions.

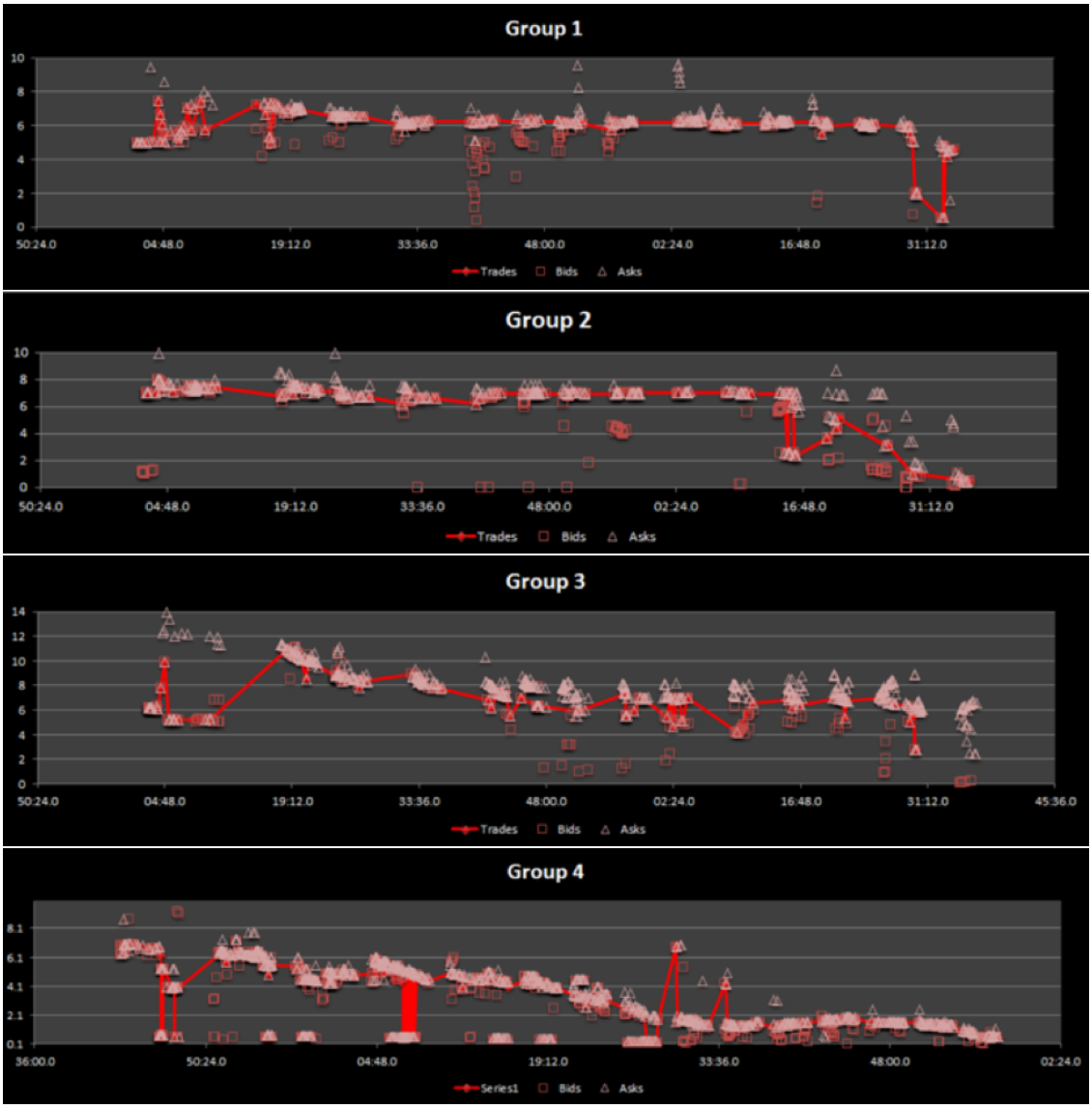


Figure 19. Plots of trade prices (red diamonds, connected with solid red line), asks (triangles) and bids (squares) for four pilot experiments. The design was identical to that of the main experiment. The yellow straight line depicts the downward trend in the fundamental value, from 7.5 dollars in period 1, to 0 dollar after the end of period 15. Bubbles emerge in 3 out of 4 replications. Flash crashes/surges are evident even in the replication with the best pricing (replication 4).

Appendix B. Instruction (Robot Sessions)

INSTRUCTIONS

For the Experiment on Human-Robot Interaction in Financial Markets

Summary

Your task will be to complete a robot-based trading game on our online trading platform. There will be a training session, to provide ample opportunity to familiarise yourself with the online marketplace and the use of the various robots available to you.

In the testing session, you have the opportunity to trade a single, 15-period-lived security called "stock." The stock pays a random dividend at the end of each period. The expected dividend each period is 0.50 (experimental) dollar. You start with an endowment of cash and stock. You can sell stocks for cash, and/or use cash to buy stocks.

You trade either manually or with a robot, or both. The choice of which robot to use and the timing of robot use are completely at your discretion.

Your goal is to maximise your performance, measured by dividends received from holding stocks, plus cash accumulation through trading. You will be given an extra 10 (real) dollars for participation in the Training Session.

Online platforms

The trading games take place in an online trading platform called Flex-E-Markets. This online marketplace can be accessed through the following link: www.flexemarkets.com. Log on to the account and with the email and password given to you.

Your robots will be in a folder on the desktop of your computer. To upload and engage a robot, log onto the following platform with the same account, email and password as for Flex-E-Markets: algohost.bmmlab.org.

Once you open these two links, you will be ready to go.

Robots

Robots are scripts that automatically submit orders (and trade) according to simple rules or "algorithms." One example is a robot that submits a sell order whenever the price is at least 5 cents above a certain value called *Fair Value*. Another example is a buy robot that submits a buy order whenever the price is 5 cents below a specific Fair Value. During the training session, we will introduce participants to a limited number of robots, distinguished by Fair Value, whether they buy or sell (or both), and whether they are pro-active ("Market Maker") or simply react to incoming order ("Reactionary").

A. Training Session

A.1 Setting

During the training session, participants will be given the opportunity to trade a single asset called stock with a random final payoff at the end of the session. The final payoff will be one of

0, 4, 6 and 10 (experimental) dollars, with equal chance. This means that the expected payoff for each stock is (0 + 4 + 6 + 10)/4 = 5 dollar.

A.2 Manual Trading

The trading platform is organised as a continuous open book system, in which you can submit orders to buy (bids) and orders to sell (asks) at any time when the market is open. When your bid is at a price higher than the best standing ask, you will trade immediately with the originator of that ask, at the ask price. When you submit an ask at a price below that of the best standing bid, you trade immediately with the bid originator, at the bid price. If there are more than one order at the same price, earlier orders get executed first. Order submission and transactions are anonymous.

The order submission area is located at the middle of the trading screen of Flex-E-Markets. There, you can switch between buying and selling using the "Buy/Sell" button, specify the quantity under "Units" and set your price using the slider below "Price". At any moment while the market is open, you can see outstanding bids and asks as blue (bids) and pink (asks) entries on the left of the screen. You can cancel your own orders by clicking the cross button to the right of your orders. By toggling the "All/Mine" button on the top right corner, you can switch between viewing all the orders in market or the orders originated by you. On the right of your screen you can observe the list of transaction history including information on time, price and (cumulative) quantity.

After you familiarised yourself with the manual trading interface, you should try robots. How to do this is explained next.

A.3 Robots

We have prepared a set of robots for you to use. These robots trade whenever prices are advantageous compared to a reference point called *Fair Value*. Robots come in two main types. The *Market Making* robots submit orders in order to be first to trade when others submit orders. The *Reactionary* robots trade in reaction to incoming orders from other traders. The Market Maker robots are pro-active; the Reactionary robots are re-active.

Choose the robot you want to use from the folder on the desktop of your computer, upload it to algohost.bmmlab.org, wait until the period starts and launch it by clicking on the green button. You can stop the robot whenever you'd like, by clicking on the yellow button. But remember that robots are fast – by the time you hit "stop," your robot may have executed a few trades, and will complete those before coming to a complete stop.

There is no problem with uploading a different robot during trading if you like to do so. However, first stop your current robot before you upload your next one! Then delete the current robot by clicking on the red button in the algohost.bmmlab.org interface. Then choose another robot from your folder, upload, and engage by clicking the green button.

All robots trade with reference to a Fair Value. Robots come with different Fair Values, anywhere from 0.25 to 9.75 experimental dollars, in increments of 25 cents. Robots are also distinguished by the way they take action.

- ***Market Maker* robots** do not wait for the right orders to come in; they immediately submit a buy at Fair Value MINUS 5 cents or a sell order at Fair Value PLUS 5 cents, and will do so again whenever the order is filled (i.e., whenever a trade takes place). Because of the rules of the market, orders will be executed at the prices submitted by the robot. E.g., if the robot has a Fair Value equal to $5.00, then it will submit a buy order at $4.95 (provided it is a BUY or BUY & SELL robot) and this order will execute at $4.95. The market making robots submit an order for one (1) unit at a time.
- ***Reactionary* robots** wait for the right orders to come in; they submit a buy order whenever there is a sell order at Fair Value – 5 cents or less (if a BUY robot), or a sell order whenever there is a buy order at Fair Value + 5 cents or more (if a SELL robot). Because of the rules of the market, these orders will execute at the price of the order the robot reacts to. E.g., if there is a sell order for $3.15 and Fair Value equals $5.00, then the robot will buy (provided it is a BUY or BUY & SELL robot) at $3.15. The reactionary robots submit orders for one (1) unit at a time. Reactionary robots are aggressive, and will submit their orders at Fair Value (unlike the Market Makers) in order to ensure they get maximum priority in case competing orders become available.

The robots sit in the folder "ROBOTS" on your desktop. You can recognise the type of robot and its Fair Value by looking at its name. The naming convention is as follow:
SSW_<**role_name**>_<**type**>_fv-<**fair_value**>_x-<x>_e-<e>, where:
<**role_name**> = BUY, SELL, BUY-SELL
<**type**> = MARKET_MAKER, REACTIONARY
<**fair_value**> = 0.25, 0.50, 0.75,… 9.75
(Ignore <x> and <e>; these numbers are always the same)
Robot **E**xample **1**:

SSW_**BUY-SELL_REACTIONARY**_fv-**4.0**_x-5_e-5

This is a Reactionary double-sided (buy & sell) robot that submits orders around a Fair Value equal to $4.00.
Robot **E**xample **2**:

SSW_**BUY-SELL_Market_Maker**_fv-**5.50**_x-5_e-5

This is a double-sided Market Making robot with Fair Value equal to $5.50, which sends buy orders (for 1 unit) at $5.45 to capture incoming sell orders, and sell orders (for 1 unit) at $5.55 to capture incoming buy orders.
Robot **E**xample **3**:

SSW_**SELL_REACTIONARY**_fv-**2.25**_x-5_e5

This is a single-side reactionary SELL robot with Fair Value equal to $2.25, which means that it will sell whenever there is a buy order at 2.30 dollars (=2.25+0.05) or higher.

B. Test Session

In the test session, you will again be trading a single security called stock, against cash. There will be fifteen (15) trading periods. After each of them, the stock pays a random dividend. After paying the last dividend in period 15, the stock expires worthless. Periods will last 5 minutes or less.

At the beginning of the first period, about half of the participants start with 20 units of stock and 100 experimental dollars. The remaining participants will start with 12 units of stock and 160 experimental dollars.

During each period, the market will be open and you will be able to trade. When the market closes we use random number generator in MATLAB to determine the dividend. The command being used is "randi([1 100])", generating uniformly distributed random integer from 1 to 100.[5] If the generated number is in the range of [1 25], the dividend will be $0. If the generated random number falls in [26 50], the dividend will be $0.25. If the generated number falls in [51 75], the dividend will be $0.50, and if the generated number falls in [76 100], the dividend will be $1.25. As such, the expected period dividend equals $0.50 (=1/4*(0+0.25+0.5+1.25)).

The dividend will be distributed to you in the form of experimental cash, and paid for each unit of stock you own at the end of the period. For example, if in period 1 you end with 25 units of stock and the dividend is $0.25, then $6.25 (=25*0.25) will be added to your experimental cash. In period 2 you will thus start with the same 25 units of the stock, but your cash will have increased relative to the end of period 1 by $6.25.

Each period, the dividend on each stock is $0.50 on average. Since there are 15 periods, the sum of all expected dividends in period 1 is $7.50 (=15*0.50) for each stock. In the second period the stock will have paid its first dividend, so there are only 14 more payments left, each worth $0.50 in expectation. So, the sum of all remaining expected dividends is $7 (=14*0.50). In the third period there will be 13 more payments left, with an expected total dividend of $6.50, etc.

How much you are willing to buy or sell your stock in any given period will depend on the horizon you plan to hold this stock for and what you foresee other participants to be willing to buy and sell the stock for.

[5] The four random outcomes, defined as ranges of [1 25], [26, 50], [51 75] and [76 100], have equal probability of 25%.

C. Compensation

The cash you are holding at the end of the 15th period is yours to take home. It will be converted to real dollars at a 5:1 exchange rate (meaning that 5 experimental dollars convert into 1 real dollar). In addition, you will be given 10 (real) dollars for the Training Session. The expected amount of compensation for each participant in this experiment is $60. The actual amount can be higher or lower depending on your individual performance, but will not be above $80

REFERENCES

Angel, James J, Lawrence E Harris, and Chester S Spatt, 2011, Equity trading in the 21st century, *The Quarterly Journal of Finance* 1, 1–53.

Biais, Bruno, Thierry Foucault, and Sophie Moinas, 2011, Equilibrium high frequency trading, in *International Conference of the French Finance Association (AFFI)*.

Bouchaud, Jean-Philippe, Yuval Gefen, Marc Potters, and Matthieu Wyart, 2004, Fluctuations and response in financial markets: the subtle nature of 'random'price changes, *Quantitative finance* 4, 176–190.

Brandts, Jordi, and Gary Charness, 2011, The strategy versus the direct-response method: a first survey of experimental comparisons, *Experimental Economics* 14, 375–398.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2014, High-frequency trading and price discovery, *The Review of Financial Studies* 27, 2267–2306.

Carrion, Allen, 2013, Very fast money: High-frequency trading on the nasdaq, *Journal of Financial Markets* 16, 680–711.

Chaboud, Alain P, Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega, 2014, Rise of the machines: Algorithmic trading in the foreign exchange market, *The Journal of Finance* 69, 2045–2084.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2001, Market liquidity and trading activity, *The journal of finance* 56, 501–530.

De Luca, Marco, and Dave Cliff, 2011, Human-agent auction interactions: Adaptive-aggressive agents dominate, in *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, volume 22, 178.

Embrey, Matthew, Guillaume R Fréchette, and Sevgi Yuksel, 2017, Cooperation in the finitely repeated prisoner's dilemma, *The Quarterly Journal of Economics* 133, 509–551.

Gini, Corrado, 1936, On the measure of concentration with special reference to income and statistics, *Colorado College Publication, General Series* 208, 73–79.

Goldstein, Michael A, Pavitra Kumar, and Frank C Graves, 2014, Computerized and high-frequency trading, *Financial Review* 49, 177–202.

King, R.R., V.L. Smith, A.W. Williams, and M.V. van Boening, 1993, The robustness of bubbles and crashes in experimental stock markets, in Richard Hollis Day, and Ping Chen, eds., *Nonlinear Dynamics and Evolutionary Economics*, chapter 13, 183–200 (Oxford University Press, Oxford).

Kissell, Robert, and Roberto Malamut, 2005, Understanding the profit and loss distribution of trading algorithms, *Algorithmic Trading* 41–49.

LeBaron, Blake, W Brian Arthur, and Richard Palmer, 1999, Time series properties of an artificial stock market, *Journal of Economic Dynamics and control* 23, 1487–1516.

Palan, Stefan, 2013, A review of bubbles and crashes in experimental asset markets, *Journal of Economic Surveys* 27, 570–588.

Powell, Owen, and Natalia Shestakova, 2016, Experimental asset markets: A survey of recent developments, *Journal of Behavioral and Experimental Finance* 12, 14–22.

Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *The Journal of finance* 39, 1127–1139.

Rust, John, John H Miller, and Richard Palmer, 1994, Characterizing effective trading strategies: Insights from a computerized double auction tournament, *Journal of Economic Dynamics and Control* 18, 61–96.

Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams, 1988, Bubbles, crashes, and endogenous expectations in experimental spot asset markets, *Econometrica* 56, 1119–1151.

Notes

[1]Formally, the fundamental value should be the sum of the discounted future payoffs, however in a short experimental session the discount factor is 1.

[2]Because robots can be stopped and replaced, some more sophisticated history-dependent strategies can be implemented but not automated. E.g., if a participant would like to exploit momentum and trade based on the average price over the last minute, she could replace her robot every minute with a new one, with reference price equal to the average traded price over the previous minute.

[3]In principle, robots could be programmed to act under strategic thinking. We are not aware of any real-world robots that come even close to implementing protocol to resolve strategic uncertainty. Obviously, the robots in our experiment did not have this sophistication.

[4]Note that the total earnings in experimental sessions with robots (in experimental dollars) is exactly the same as in manual trading experimental sessions.

[5]Significance is determined, not by comparing the estimated effective bid-ask spreads (which were set equal to zero if the estimated autocovariances were positive), but by comparing estimated autocorrelations. For each period and session we first computed the autocorrelations of transaction price changes. When multiplied by the square root of the time series length, the resulting scaled autocorrelations are standardized estimates of the autocovariances, meaning that, in large samples, they are normally distributed with mean zero and variance 1. Next, we ran a simple t-test with known variances on the differences in the mean standardized autocovariances across the two treatments (robots; manual trading). We set the variance of each observation (each standardized autocovariance) equal to 1.

[6]Under the Gaussian distribution, and with our definition of outliers, we expected 4.2% of all transaction price changes to be flash crashes/surges.