Algorithmic Trading, Price Efficiency and Welfare: An Experimental Approach*

Brice Corgnet† Mark DeSantis‡ Christoph Siemroth§

August 29, 2023

Abstract

We develop a novel experimental paradigm to study the causal impact of two classes of trading algorithms on price efficiency, trading volume, liquidity, and welfare. In our design, public information about the asset value is revealed during trading, which gives algorithms a reaction speed advantage. We distinguish market-order (aggressive) and limit-order (passive) algorithms, which replace human traders from the baseline markets. Relative to human-only markets, limit-order algorithms improve welfare, although human traders do not benefit, as the surplus is captured by the algorithms. Market-order algorithms do not change welfare, though they do lower human traders' profits. Both types of algorithms improve price efficiency, lower volatility, and increase the share of profits for unsophisticated human traders. Our results offer unique evidence that non-exploitative algorithms can enhance welfare and be beneficial to unsophisticated traders.

Keywords: Algorithmic Trading, Experimental Markets, High-Frequency Trading, Price Efficiency, News Announcements, Welfare.

JEL Classification: C92, D61, G12, G14, G41.

*We thank Matteo Aquilina, Yashar Barardehi, Andrew Bird, Jean Edouard Colliard, Dan Friedman, Steve Heinke, Chad Kendall, Jason Shachat, Ekaterina Sherstyuk, Sarah Zhang and audiences at Essex, Goethe University Frankfurt, ESA Lyon and EEA-ESEM Barcelona for valuable comments. Siemroth was supported by the Economic and Social Research Council [grant ES/T006048/1]. The experimental design and main statistical tests were pre-registered at `https://www.socialscienceregistry.org/trials/9001`.

†Emlyon Business School, GATE UMR 5824, 23 Avenue Guy de Collongue, 69130 Ecully, France. E-mail: corgnet@em-lyon.com. This research has been performed within the framework of LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon

‡Chapman University, Argyros School of Business and Economics; Economic Science Institute, One University Drive, Orange, CA 92866, USA. E-mail: desantis@chapman.edu

§University of Essex, Department of Economics, Wivenhoe Park, Colchester, CO4 3SQ, UK. E-mail: christoph.siemroth@essex.ac.uk

1 Introduction

One of the major financial transformations in the last twenty years has been the emergence of algorithmic and high-frequency trading (Biais and Foucault (2014); O'Hara (2015)). According to recent estimates, hedge funds use algorithms in 53% of their trades (see Aite (2021)) and the proportion of algorithmic trading in FOREX has reached 58% (BIS (2022)). The growth of algorithmic trading (AT, henceforth) is expected to continue in all markets (SEC (2020)). Despite numerous studies showing a positive impact of AT on both liquidity and price discovery,[1] skepticism remains. Skeptics challenge the welfare benefits of AT (see Kirilenko and Lo (2013); Biais and Foucault (2014); Stiglitz (2014)) and emphasize the negative impact of manipulative algorithms and of the ongoing race for speed in high-frequency trading (HFT, henceforth) (Biais, Foucault, and Moinas (2015); Aquilina, Budish, and O'Neill (2022)). Critics also emphasize the potential threats of HFT to market stability due to periodic illiquidity, increased comovement in returns, and flash crashes (O'Hara (2015); Kirilenko et al. (2017); Malceniece, Malcenieks, and Putniņš (2019)). In the words of Stiglitz (2014, p. 14): "The arguments [...] leave me skeptical about the social value of high-frequency trading."

In this paper, we use experimental markets to conduct what we believe to be the first study of the causal effect of two classes of AT on welfare, thus shedding light on this divisive issue. In our setting, we focus on welfare, measured as gains from trade.[2] Specifically, we define welfare as the sum of payoffs of all traders in the market, which includes the trading algos.[3] Some traders value the asset more than others, so our markets allow for gains from trade and are *not* zero-sum. Thus, the exchange of assets can increase or decrease welfare. This feature makes our setting suitable for welfare analysis. Welfare is thus directly determined by asset allocation and trader preferences, whereas field studies commonly rely on welfare proxies such as price efficiency, trading volume, and liquidity measures.

In our experimental design, subjects participate in a continuous double auction, where they can both submit offers to buy or sell an asset as well as accept others' offers. The asset has a liquidating value that is either high or low. At a specific time during trading, a public news announcement is made, which reveals information about the asset. This information can drive the value of the asset up or down, depending on the public announcement. This feature of the markets is critical, because it is one aspect where critics argue that trading algos have an unfair speed advantage over human traders.

[1] See Hendershott, Jones, and Menkveld (2011); Riordan and Storkenmaier (2012); Brogaard, Hendershott, and Riordan (2014, 2019); Chaboud et al. (2014); Brogaard et al. (2015); Conrad, Wahal, and Xiang (2015); Hu, Pan, and Wang (2017); Chordia, Green, and Kottimukkalur (2018); Korajczyk and Murphy (2019); Chordia and Miao (2020); Boehmer, Fong, and Wu (2021); Chakrabarty, Moulton, and Wang (2022).

[2] To focus our investigation, we deliberately leave aside fixed costs associated with deploying algorithms (see Aldrich and López Vargas (2020) for a discussion).

[3] For a more detailed explanation, see Section 4.1.

We vary exogenously (within subjects) whether the news release time is precisely known in advance, as in the case of a scheduled earnings announcement, or whether it is not precisely known in advance, as in the case of an unexpected event that affects firm profits. Our three main treatments (between subjects) vary whether a large endowment trader (i.e., a trader with more cash and more assets to trade) is played by a human (in the baseline), by a limit-order algo, or by a market-order algo. Thus, we hold the number of traders constant in the market, which allows for clean comparisons of welfare across treatments.

We find that limit-order trading algos—which offer trades and add liquidity to the market, thus acting like market makers—increase welfare in the market, relative to the baseline with only human traders. Hence, limit-order algos can facilitate the transfer of assets to those that value them the most. However, the additional surplus is almost exclusively absorbed by the trading algos, so that the average human trader neither gains nor loses from the introduction of the limit-order algo. The limit-order algo helps the market realize on average 60% of the possible welfare gains, compared to only 35% in the baseline. Structural equation models show that 74% of the positive welfare effect of the limit-order algo is driven by two channels, increased price efficiency (65%)—which helps to ensure traders with high valuations buy from those with low valuations, thereby increasing welfare—and increased trading volume (9%).

We further find that market-order algos—which accept limit orders and take liquidity—do not change welfare relative to human-only markets. On average, markets with the market-order algo realize 43% of possible welfare gains, compared to 35% in the baseline. In these markets, the payoff of the average human trader is lower than in the human-only markets. The market-order algo earns more than the human it replaces. As a result, this type of algorithm causes a redistribution of payoffs from humans to algorithms. The market-order algorithm puts human traders at a disadvantage by utilizing its speed advantage to react to changes in asset value following news releases. This is achieved by accepting human traders' outstanding orders ("stale limit orders") before human traders can react to the news. Hence, market-order algos benefit by exploiting human traders. Market-order algos achieve payoffs (net of initial endowment) that are 49% higher than those of limit-order algos, suggesting that market-order algos are highly profitable and therefore are more likely to be commonly used in markets (see also Asparouhova et al. (2022)). Hence, the individual profit-maximizing objectives of algo funds are not necessarily aligned with maximizing welfare in the market.

Both types of trading algos improve price efficiency, so prices in algo markets better reflect the true asset value than the baseline markets. The limit-order algo improves price efficiency more so than the market-order algo, and both types of algos decrease price volatility. The use of the limit-order algorithm substantially enhances liquidity, as indicated by a narrower bid-ask spread. In contrast, the market-order algorithm does not significantly change liquidity.

The limit-order algorithm generates greater trading volume compared to the market-order algorithm. However, there is no significant difference in trading volume between algo markets and markets populated solely by humans. Finally, the market-order algorithm produces a greater amount of trades immediately following a news release. This outcome can be attributed to the algorithm's speed and ability to take advantage of existing human limit orders.

At the individual level, we find that traders with higher cognitive ability earn more and enhance price efficiency and welfare in markets in line with previous research (see Bosch-Rosa and Corgnet (2022) for a review). Furthermore, the difference in payoffs between traders with varying levels of cognitive ability is reduced in algorithmic treatments. In the baseline, median (average) payoffs of traders with high cognitive skills are 15% (9%) higher than those with low cognitive skills, whereas this difference is only 6% (3%) in the algo treatments. Furthermore, traders with low cognitive skills earn significantly more in algo treatments than in the baseline, whereas the opposite is true for those with high cognitive skills. Finally, traders with more negative attitudes toward robots make more trading mistakes in algo treatments and earn less than in the baseline.

Our contributions encompass several key aspects. Firstly, we design a novel experimental setting that enables us to investigate the causal effect of two distinct types of trading algorithms on welfare. These algorithms, well-established in the field literature, represent two opposite extremes in terms of their level of aggressiveness (i.e., propensity to take liquidity). Secondly, our approach allows us to directly measure welfare rather than using market quality proxies that rely on liquidity and price efficiency measures. We show that although market quality measures tend to improve when both types of trading algorithms are introduced, limit-order algorithms fail to increase human traders' welfare, and market-order algorithms actually reduce it. These findings suggest that standard market quality measures such as volatility and bid-ask spreads may not be reliable indicators of the welfare consequences of algorithms. Moreover, our research emphasizes that trading algorithms are not universally beneficial or detrimental because welfare effects are contingent upon the type of algorithm employed. Finally, our approach allows us to go beyond the estimation of welfare effects and to identify the specific channels through which the two types of algorithms operate.

Our findings result from our methodological choice to use experimental markets, which we discuss further in the next section. We also outline how our experimental setting complements field research.

2 Experimental approach and literature

The experimental approach is new to the AT literature, but it has grown substantially in recent years (see Bao et al. (2022) for a review). The experimental literature has considered a wide range of AT types, including market-making, and liquidity-providing algorithms (Asparouhova et al. (2022)), arbitrage bots (see e.g., Angerer, Neugebauer, and Shachat (2023)) as well as manipulative algorithms (Veiga and Vorsatz (2009); Veiga, Vorsatz et al. (2010); Leal and Hanaki (2018)). In the early studies, a common theme was to compare the performance of human and algorithmic traders, showing that bots usually outperform humans (see e.g., Das et al. (2001); Gjerstad (2007); Feldman and Friedman (2010)). However, Asparouhova et al. (2020) show that humans—who were given the opportunity to partly delegate trading to algorithms using either liquidity-providing ('market-maker') or liquidity-taking ('reactionary') bots—did not increase payoffs despite using bots extensively. Interestingly, Asparouhova et al. (2022) develop a game-theoretic model to study the strategic choice of traders to opt for maker and taker algorithms. They show that, in equilibrium, people will specialize in one of the two algorithms.

Recent experimental studies have also focused on the impact of AT on liquidity and price efficiency, showing that arbitrage bots can increase both trading volume (see Berger, DeSantis, and Porter (2020)) and price efficiency in double auctions with specialized buyers and sellers à la Smith (1962) as well as in continuous double auction markets (Angerer, Neugebauer, and Shachat (2023); Neugebauer, Shachat, and Szymczak (2023)). Recent works have also studied the impact of institutional changes on AT. Aldrich and López Vargas (2020) have shown that introducing first-batch auctions can improve liquidity and price efficiency in the presence of AT. Khapko and Zoican (2021) have found that introducing asymmetric speed-bumps that delay the receipt and execution of trading orders for all market participants but market-maker bots can discourage investments in HFT technology. Despite a growing interest in the use of experimental markets to study AT, the literature has not yet studied the welfare effect of trading algorithms, which is the main contribution of our paper. One exception is the work of Asparouhova et al. (2022) in which human traders could deploy utility-maximizing algorithms in a CAPM experiment (Bossaerts and Plott (2002)). Yet, the main focus of Asparouhova et al. (2022) is on the strategic use of algorithms rather than on welfare effects. In their experiment, algorithms automatically increased welfare, because they were designed to maximize traders' payoffs. Using a particular type of algorithm from computer science, Zero-Intelligence-Plus bots, Peng et al. (2020) show that slower bots can trigger higher profits for human traders than faster bots. Although their design allowed for gains from exchange, welfare was already maximized in human-only markets so that, unlike in our study, the introduction of trading bots could not significantly impact welfare.

Using computerized continuous double auction markets, we study the impact of introducing two common types of AT: A limit-order algo and a market-order algo. We consider these two types of algorithms because, despite the observed heterogeneity in AT strategies, the empirical literature has largely focused on the distinction between algorithms providing (limit-order algos) or withdrawing (market-order algos) liquidity from the market (see Hagströmer and Nordén (2013); Brogaard, Hendershott, and Riordan (2014, 2019)). Our experimental approach complements existing field studies, which are limited by the quality of available data. Empirical works typically use proxies to identify AT and HFT, such as order speed, messages, cancellations, and volumes (e.g., Hendershott, Jones, and Menkveld (2011); Hasbrouck and Saar (2013)), with some risk of misclassification. They also use existing classifications provided by exchanges, and in particular the NASDAQ (e.g., Brogaard, Hendershott, and Riordan (2014)). Nevertheless, exchanges' classifications only focus on a subset of HFT firms, and because of concerns regarding data confidentiality, exchanges such as the CME have restricted access to this data (Reuters (2013)). It is even more challenging to obtain data about AT types, that range from market-making to directional arbitrage (Biais and Foucault (2014)), because their source codes are almost always proprietary and private. To categorize AT, empirical researchers have recently used statistical tools based on principal component analyses (Boehmer, Li, and Saar (2018)). By comparison, in our experiment, we can directly design and thus observe algo types and distinguish their actions from human actions.

The identification of AT and their types is not the only challenge faced by field researchers. As pointed out by Biais and Foucault (2014), AT research would benefit from studies assessing the causal impact of automation because, as argued in Brogaard, Hendershott, and Riordan (2014, p. 2269): "(...) the counterfactual of how other market participants would behave in the absence of HFTs is not known". In our experimental study, this counterfactual is known. As an alternative to the experimental method, empirical researchers have often used institutional changes to alleviate endogeneity issues such as the introduction of automated quoting (Hendershott, Jones, and Menkveld (2011)), microwave technology (Shkilko and Sokolov (2020)), low-latency markets (Malceniece, Malcenieks, and Putniņš (2019)), and colocation services, which allow traders to locate their computer hardware close to an exchange trading center (see e.g., Brogaard et al. (2015); Boehmer, Fong, and Wu (2021)).[4]

Finally, although archival data can be used to study price discovery and weak-form price efficiency, assessing welfare remain elusive, because individual investors' preferences cannot be observed. Using private values in our experimental design, we can induce welfare gains and study welfare directly.

[4]Recent works have also used outages and bad weather episodes that temporarily remove the speed advantages of HFT algorithms to study their causal impact (Shkilko and Sokolov (2020); Eaton et al. (2022)).

The limitations in field AT research are critical, because policymakers are ultimately interested in the impact of HFT on welfare and its distribution (Angel and McCabe (2013); Lewis (2015)), as expressed by Stiglitz (2014, p. 9): "Thus, the real question is, assuming that flash trading improved price discovery, does the information produced lead to better resource allocations?". To answer this question, we use experimental markets, which allow us to exogenously change the presence and the type of AT in a market, thus enabling us to identify the causal effect of algorithms. Furthermore, experimental markets allow us to exogenously vary the valuation of the asset (Frydman et al. (2014)), so we can directly test whether observed asset prices reflect fundamentals (Page and Siemroth, 2021). Our approach thus complements empirical studies that offer an indirect test of the price efficiency of markets, looking at directional evidence and testing implications of inefficient prices (see Biais and Foucault (2014); O'Hara (2015); Menkveld (2016) for reviews).

The experimental approach is sometimes faced with concerns regarding the portability of the laboratory findings. Yet, in the absence of a game-theoretic model of continuous double auctions, experiments have historically played an important role in market design (Smith (1994)) by influencing policymakers to design a variety of auctions in areas such as stock exchanges, spectrum licenses, and electricity markets (Banks et al. (2003); Rassenti, Smith, and Wilson (2003); Smith (2003); Smith (2008)). These empirical successes likely reflect the fact that, despite their simplified structure, experimental markets capture essential features of field markets such as the inherent competition among human traders in continuous double auctions. Moreover, previous research suggests that non-professional subjects generate the same qualitative findings as finance professionals. For example, the findings of asset bubbles arising in experimental markets (e.g., Smith, Suchanek, and Williams (1988)) mirror the bubble patterns observed in the field, and they also occur when finance professionals take part in the same experimental markets (e.g., Weitzel et al. (2020)). Snowberg and Yariv (2021) also show that correlations between behaviors are similar across subject pools. Furthermore, our study controls for the cognitive ability of traders by administering the cognitive reflection test (Frederick (2005)), which has been shown to capture traders' sophistication (Thoma et al. (2015); Corgnet, DeSantis, and Porter (2018)).

3 Hypotheses

Based on the extensive theoretical literature on AT and recent empirical evidence, we hypothesize that the limit-order algos will improve liquidity compared to human-only markets by narrowing the bid-ask spread, which will tend to increase trading volume. (see Hasbrouck and Saar (2013); Biais and Foucault (2014); O'Hara (2015); Menkveld (2016)). This prediction directly follows from the design of limit-order algos that place competing orders which improve the available bid-ask spread. By contrast, market-order algos are designed to take away liquidity from the market. We also expect that limit-order algos will lead to higher trading volume than market-order algos (Hagströmer and Nordén (2013)), as they offer more transaction opportunities by populating the order book. Yet, by design, market-order algos will instantly pick off any order that is not in line with the news announcement on the fundamental asset value. We thus expect that market-order algos will increase trading volume immediately after the news announcement.

Hypothesis 1 (Liquidity).

i. Limit-order algos will result in narrower bid-ask spreads and higher trading volume, while market-order algos will result in wider bid-ask spreads and lower trading volume.

ii. Market-order algos will engage in more trades immediately after news announcements.

The existing literature also argues that liquidity-providing algos have a positive impact on price efficiency (see Brogaard, Hendershott, and Riordan (2019)). In particular, this is achieved by algos incorporating public information much faster than humans. Market-order algos are also associated with improved price efficiency (Brogaard, Hendershott, and Riordan (2019)). In our design, this is the case because market-order algos remove orders that are not in line with the fundamental asset value. It follows that transaction prices will reflect publicly available information more accurately. Evidence for this mechanism has been reported in the experimental literature by Berger, DeSantis, and Porter (2020). In the presence of limit-order algos, market participants can learn and extract information from both market transactions and the order book, thus guiding price discovery more effectively than a market-order algo. We thus predict that limit-order algos will achieve higher price efficiency than market-order algos.

In the non-experimental literature, the positive impact of market orders on price efficiency has been shown in a variety of ways. For example, empirical papers have assessed the extent to which and the direction in which prices react to macroeconomic announcements and shocks in the presence of market-order algorithms (e.g., Chordia, Green, and Kottimukkalur (2018); Brogaard, Hendershott, and Riordan (2019)). Other authors have evaluated price efficiency by looking at the potential implications of inefficient prices such as the existence of arbitrage opportunities and the autocorrelation in returns (Chaboud et al. (2014); Conrad, Wahal, and Xiang (2015)).

We also predict that the presence of algorithms will reduce volatility in prices. Indeed, previous empirical studies have shown that AT tends to improve liquidity (see Hypothesis 1), which has been associated with lower volatility levels (e.g., Biais and Foucault (2014); Hasbrouck and Saar (2013); Boehmer, Li, and Saar (2018)). In addition, we anticipate that

AT will reduce informational inefficiencies, resulting in greater stability in asset prices relative to their fundamental values and a decrease in the levels of volatility typically associated with inefficient markets (Shiller (1981, 2014)). Given the existing literature, we make the following predictions regarding the impact of AT on price efficiency and volatility.

Hypothesis 2 (Price efficiency and volatility).

i. Limit-order and market-order algos will improve price efficiency.

ii. Limit-order algos will improve price efficiency more so than market-order algos.

iii. Limit-order and market-order algos will reduce volatility.

Regarding welfare, the comparison between the limit-order and market-order algos is clear-cut, as the former enhance both liquidity (Hypothesis 1) and price efficiency (Hypothesis 2) more so than the latter. As such, we predict that limit-order algos will lead to higher welfare than market-order algos (Hypothesis 3i).

Regarding the welfare impact of AT with respect to human-only markets, theoretical predictions have identified both positive and negative effects. On the one hand, both types of algos will foster price efficiency, thus producing prices that more accurately reflect available information and that can facilitate gains from exchange. On the other hand, market-order algos might lead human traders to fear being exploited by the algo, thus behaving more conservatively and refraining from placing limit orders, especially around news announcements. Market-algos will thus reduce trading, as predicted in Hypothesis 1i, and lead to fewer gains from trade. This tension between information provision (due to increased price efficiency) and fear of exploitation leads to unclear predictions about the welfare impact of market algos (see e.g., Biais, Foucault, and Moinas (2015)). In the case of limit-order algos, the fear of exploitation by the algo still applies. Yet, we have conjectured that limit-orders will improve liquidity and price efficiency, thus leading to higher welfare, as previously conjectured by Menkveld (2016). In sum, the welfare impact of AT remains an empirical question, which can be tackled using experimental markets.

Because algorithms have a speed advantage, they will react to news instantly, thus frontrunning any human traders' orders (see Foucault, Hombert, and Roşu (2016)). This informational advantage implies that algorithms will tend to trade at more advantageous prices than human traders right after news is announced. This suggests that a disproportionate share of any of the welfare gains due to the introduction of algos will be captured by the algorithms themselves. This negative impact on the distribution of welfare will be especially pronounced for market-order algos, which are designed to immediately exploit any existing limit order that is not in line with the news (Hypothesis 3ii).

Hypothesis 3 (**Welfare**).

i. Limit-order algos will achieve higher welfare than market-order algos and human-only markets.

ii. Algos will capture a disproportionate share of the gains from trade, especially in the case of market-order algos.

To study further the impact of algos in markets, we assess the link between individual cognition and market outcomes following the cognitive finance literature (Akiyama, Hanaki, and Ishikawa (2017); Corgnet, DeSantis, and Porter (2018); Bosch-Rosa, Meissner, and Bosch-Domènech (2018); Bosch-Rosa and Corgnet (2022)). In line with the literature, we posit that in human-only markets traders with high cognitive skills will be more likely to trade based on fundamental information about the asset value, thus making fewer mistakes and earning more (Grinblatt, Keloharju, and Linnainmaa (2011, 2012); Breaban and Noussair (2015); Cueva and Rustichini (2015); Noussair, Tucker, and Xu (2016); Corgnet, DeSantis, and Porter (2020)). However, the advantage of traders with high cognitive ability will diminish in the presence of algos because, as predicted in Hypothesis 2, algos will lead to higher levels of price efficiency, which is when cognitive skills play a lesser role in explaining traders' payoffs. This is the case because traders with low cognitive skills, unlike those with high cognitive skills, fail to place orders that reflect the fundamental asset value whenever market prices are inefficient (see Corgnet, DeSantis, and Porter (2018)). We state these predictions in Hypothesis 4i.

Following our reasoning, we expect that markets populated by a greater number of high cognitive skills traders will lead to higher price efficiency in human-only markets (see Bosch-Rosa, Meissner, and Bosch-Domènech (2018); Corgnet, DeSantis, and Porter (2021)). This increase in price efficiency will facilitate gains from trade and enhance welfare (see Hypothesis 4ii, first part). However, in algo markets, the positive impact of high cognitive ability traders on price efficiency and welfare will likely be reduced due to a ceiling effect (Hypothesis 4ii, second part). This is because, as stated in Hypothesis 2, we expect markets with algos to exhibit high levels of price efficiency.

Hypothesis 4 (**Cognitive skills**).

i. Traders with higher cognitive ability will earn more in human-only markets. This difference in earnings across traders will be reduced in the presence of algos.

ii. Markets populated by traders with a higher cognitive ability will exhibit higher price efficiency and welfare. This positive impact will be reduced in the presence of algos.

Finally, we expect the impact of algos to depend on traders' general attitude toward robots as measured using the Negative Attitudes toward Robots Scale (NARS, henceforth) (Syrdal et al. (2009)). Traders with negative attitudes toward robots will be less inclined to interact with them, thus trading less in the presence of algos. The possibility of trading with algos may also produce anxiety in high-NARS traders. Because anxiety tends to impair cognitive ability (e.g., Shields, Sazma, and Yonelinas (2016); Yu (2016)), we expect high-NARS traders to make more mistakes. In sum, high-NARS traders will be less likely to materialize gains from exchange and will make more mistakes when trading in the presence of algos than in human-only markets. We thus expect high-NARS traders to earn less in the presence of algos.

Hypothesis 5 (Negative attitudes toward algorithms).
Traders who express more negative attitudes toward algorithms will trade less, make more mistakes and earn less in the presence of algos.

4 Experimental design

We designed a new experimental paradigm to allow for a direct welfare analysis in markets with trading algorithms.

4.1 Assets, welfare, information and news release

This experiment consists of 20 trading rounds, which we also refer to as markets. During each market, participants trade shares of a single risky asset. The risky asset value to trader i, v_i, has both a common value component π and a private value component θ_i, so that:

$$v_i = \pi + \theta_i.$$

Any share of the asset that is held by trader i at the end of a market will pay off v_i in Experimental Currency Units (ECUs).

The private value component can take two different values, $\theta_i \in \{0, 10\}$. That is, the "low trader type" ($\theta_i = 0$) values the asset less than the "high trader type" ($\theta_i = 10$), throughout the experiment. A possible interpretation is that some traders are more impatient for liquidity reasons than others and hence value holding the asset less. Alternatively, these private values might capture differences in marginal tax rates across trader types. Importantly, the private value component in the asset value allows us to evaluate welfare changes induced by trading. Every time a low type sells to a high type, the sum of payoffs in the market increases by 10, so there are gains from trade. Every time a trade occurs between traders of the same type, there are no gains, and every time the low type buys from the high type, the sum of payoffs decreases. This welfare analysis is not possible with a common value asset, which creates a zero-sum game.

The common value component π can take two different values, $\pi \in \{20, 80\}$. At the beginning of trading, all traders only know that these two common value realizations are possible, and that both are equally likely. Hence, the expected asset value (sum of common and private values) at the beginning of trading is 50 for the low trader type and 60 for the high type.

At a point during trading (detailed below), the realization of the common value π is publicly revealed to all traders. We also call this event "news release". After the news, no uncertainty about the asset value remains, since traders know both their private value and the common value realization. Before the news release, uncertainty exists, because two different asset values may realize. Effectively, the news is either good (if $\pi = 80$ realizes) or bad (if $\pi = 20$ realizes). News release occurs naturally in markets. We specifically add it to our experiment, because news release presents a situation where algos might benefit from their speed advantage against human traders.

The news release shifts the asset value for both trader types in the same way. That is, traders who value the asset less prior to the news release will also be the ones valuing it less after the release. However, the range of prices inducing mutually-beneficial trades changes after the news release. Prior to the news, prices in $p \in (50, 60)$ are mutually beneficial (assuming risk neutrality), whereas after the news release either $p \in (80, 90)$ (in case of good news) or $p \in (20, 30)$ (in case of bad news) is the mutually beneficial price range. Trade is possible at other prices, but those prices make one party worse off compared to holding the asset.

Welfare is the sum of the payoffs of all traders in the market.[5] Based on this utilitarian welfare criterion, the welfare optimum is achieved if all low-type traders sell all of their assets to high-type traders, which generates gains from trade with every transaction. This maximizes the sum of payoffs in the market. Pareto-optimality requires not only that the welfare-maximizing allocation is reached, but also that the transaction prices are between the valuations of the two trader types. We will later define these prices as informationally efficient.

There is no private or asymmetric information about the asset in this experiment, and the structure of information is common knowledge.

4.2 Traders

There are 8 traders per market. All traders know their own type θ_i, and their type is constant throughout the experiment. Traders can be distinguished as follows:

- 5 traders have the high type ($\theta_i = 10$). At the beginning of each market, they are endowed with 4 shares of the asset and 360 ECUs in cash.

[5]Welfare includes not only the payoffs from the asset, but also the cash held.

- 2 traders have the low type ($\theta_i = 0$). At the beginning of each market, they are endowed with 4 shares of the asset and 400 ECUs in cash. The cash endowment of the low-type traders exceeds that of the high-type traders by 40 ECUs. This is because, if held until the end of trading, their 4 shares of the risky asset pay off 40 ECUs less due to the lower private value component.

- 1 *large* trader has the low type ($\theta_i = 0$). At the beginning of each market, this trader is endowed with 12 shares of the asset and 1200 ECUs in cash. That is, this trader holds three times the endowment of the other low-type traders, and this is why we refer to this trader as the *large* trader. In the algo treatments, the algorithms assume the role of the large trader. Because the algo receives a large endowment, it is not easily cash or asset constrained. Replacing a single human trader with an algorithmic trader facilitates comparisons across treatments.

Even though there are only 3 low-type traders, due to the high endowments of the large trader, both the low-type and high-type traders have the same aggregate endowment of 20 shares of the asset. High-type traders are endowed with enough cash to buy all of the assets from low-type traders at any possible asset value.[6] The large trader has an exchange rate of 81 ECUs = \$1 which is three times lower than that of other traders (27 ECUs = \$1) to compensate for the difference in the initial endowment. These exchange rates ensure that all participants have the same expected payoffs. Endowments and types are common information as they are thoroughly explained in the instructions (see Appendix C).

4.3 Double auction trading

Trade in each market is organized as a continuous double auction which lasts for 100 seconds.[7] During this time, all traders can submit sell or buy limit orders that specify a price in ECUs at which they are willing to sell or buy a share of the asset. Traders can also submit market orders to accept active (i.e., not yet accepted or withdrawn) limit orders from others in the order book. Withdrawing unexecuted limit orders is possible at any time. There are no trading fees. For simplicity, there is no short-selling and no borrowing.

The trading screen displays a list of all active limit orders (separate for buy and sell) and a list of executed trade prices. Traders also see the expected asset value given current information and their type. Prior to the news release, the interface displays the two possible future asset values (high and low) as well as the expected value (see Figure 1). This displayed asset value

[6] For this reason, the unique competitive equilibrium price in our setting is equal to the valuation of the high type traders. For any lower price, high value traders' cash endowments create excess demand.

[7] This was more than enough time (5 seconds per trade) for traders to complete at least the 20 trades between low- to high-valuation traders that are necessary to maximize welfare.

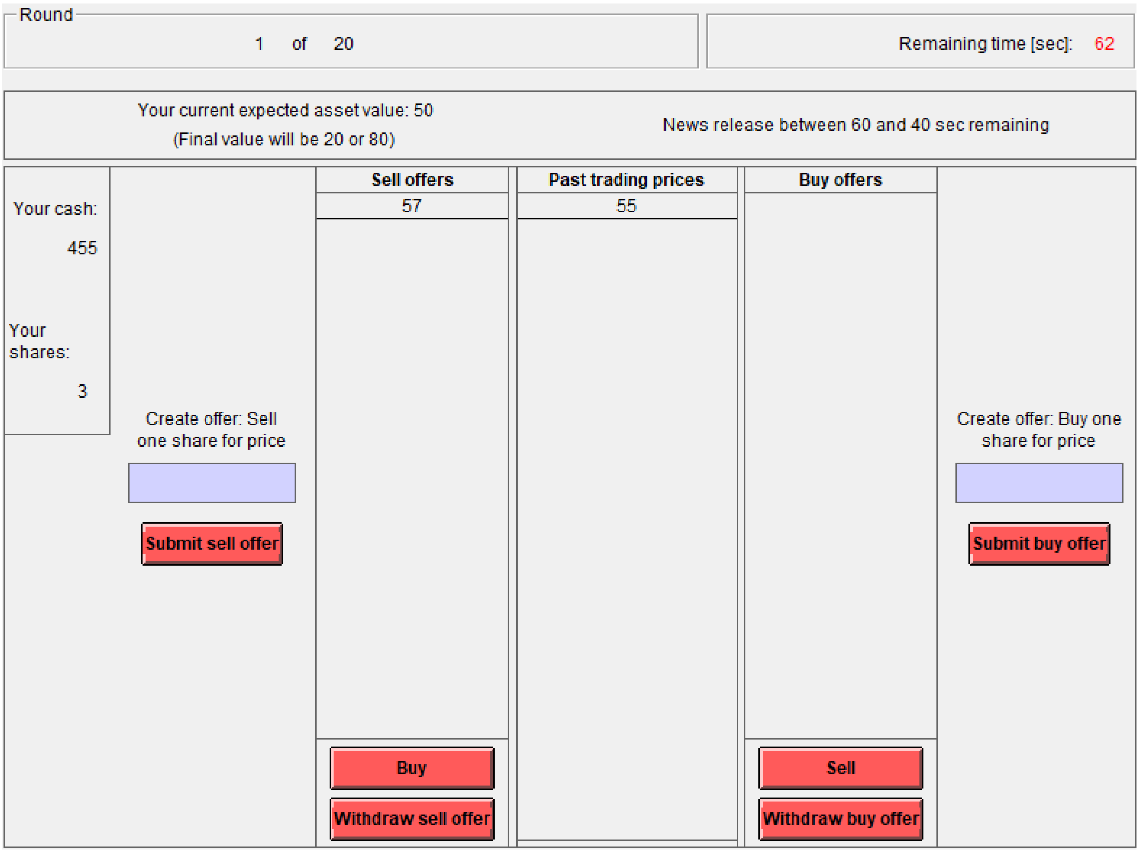


Figure 1: Screenshot of the double auction trading screen. It depicts trading with an unknown news release time. Both the current expected asset value for the given trader type and information about the news release time are displayed at the top. Current holdings are displayed on the left. The columns in the center show active sell limit orders, prices of executed transactions, and active buy limit orders (currently empty).

changes once the news is released. That is, upon news release, the screen displays the asset value that will be effectively paid at the end of the market. Hence, our results cannot be attributed to traders' confusion regarding which asset value is relevant for them.

News about the asset value is released at one of the following times $t \in \{40, 45, 50, 55, 60\}$ in seconds after the start of trading (recall that trading lasts for 100 seconds). The exact time at which news is released is either known or unknown (see Section 4.5).

At the end of each market, participants are shown a summary of their payoffs for that market. Trader payoffs per market are equal to their cash holdings at the end of a market plus any liquidating value from shares held.

4.4 Trading algorithms

The academic finance literature broadly distinguishes two kinds of algorithmic traders: Liquidity providing (passive) and liquidity-taking (aggressive), which rely respectively on limit and market orders.[8]

One defining characteristic of trading algorithms is that they are very fast; faster than any human. In our setting, algorithms can react to news immediately, which gives them an advantage over human traders. We use two different trading algorithms. There is at most one algorithmic trader in the market, which plays the role of the large trader previously described. So, algorithms transact against human traders, but not against other algorithmic traders. This allows us to cleanly identify the effect of specific kinds of algos on market outcomes.

Market-order algorithm. The market-order algo is liquidity-taking, as it only uses market orders and never submits limit orders. Its design is intentionally simple and structured to answer the following question: How would an algo maximize expected profits only using market orders in our setting? In pseudo-code, it operates as follows:

1. Every time a new limit order appears in the order book, check if accepting it implies a positive expected profit. If so, accept it, otherwise, ignore it.

2. As soon as news is released, check the entire order book and accept all limit orders that imply a positive profit.

In sum, the market-order algorithm accepts any active limit order that implies a positive expected profit. Given our setting, it only has to check limit orders when they are submitted and when news is released (which might make older limit orders suddenly profitable due to the change in asset value). The market-order algorithm has the ability to exploit stale limit orders, which is one of the main concerns in the debate surrounding algorithmic traders. Therefore, it was critical for us to incorporate this feature into our algorithm, which could be referred to as "sniper" (Li, Wang, and Ye (2021)).

Limit-order algorithm. The limit-order algo is liquidity-making and functions as an algo market maker. It always posts a buy order and a sell order as long as it has the necessary shares to sell and cash to buy. In short, the algo sets a sell order price that is equal to the current (expected) asset value plus a small markup (δ) to make a profit, but undercuts competing prices as long as this implies positive expected profits. It functions in a similar manner for buy orders.

[8]For example, Boehmer, Li, and Saar (2018) identify most HFT activity as market making (our limit-order algo), short-term directional speculation (our market-order algo), or cross-venue arbitrage (not applicable here, since there is only one trading venue). Hagströmer and Nordén (2013) make the same distinction. Hautsch, Noé, and Zhang (2017) distinguish mostly market-making and directional strategies for HFTs. In experimental works, this dichotomy is also used in Aldrich and López Vargas (2020) and Asparouhova et al. (2020, 2022).

Let the current expected asset value for the algo be V_A and let the minimum competing sell limit order price be $\underline{p}$ (if there is no other sell limit order, then let $\underline{p} = \infty$). Similarly, $\overline{p}$ is the maximum competing buy order ($\overline{p} = -\infty$ if there is no other buy limit order). The limit-order algo operates as follows in pseudo-code:

1. At any time, as long as the algo has assets to sell, offer to sell via limit orders at price:

$$p_s = \min(V_A + \delta, \max(\underline{p} - 1, V_A + 1)).$$

2. At any time, as long as the algo has enough cash to buy, offer to buy via limit order at price:

$$p_b = \max(V_A - \delta, \min(\overline{p} + 1, V_A - 1)).$$

This implies that the algo potentially updates its active limit orders any time a competing trader posts or cancels a limit order and any time a limit order is accepted because $(\underline{p}, \overline{p})$ might change. It will also update its orders when news is released because V_A changes.

We set the markup or "trading profit absent competition" at $\delta = 5$ in all sessions.[9] Taking $\delta = 5$ implies that the limit-order algo offers to sell at the midpoint of the valuations of the two trader types if there are no competing limit orders, given that the algo is always the low type. In other words, it splits the surplus evenly with a high-type trader if there is no competing limit order. This choice of δ is consistent with the maximum arbitrage profit of one-half the difference in fundamental values of two perfectly correlated assets in Angerer, Neugebauer, and Shachat (2023).[10]

The algo sets its sell order price, p_s, to be the minimum of the current expected asset value plus the markup, $V_A + \delta$, and the max term. The max term is the maximum between the lowest competing sell order $\underline{p}$ minus 1 and the asset value V_A plus 1 (any integer lower than $V_A + 1$ implies no or negative profits, so the algo does not price lower). In other words, the max term takes the maximum between the price that undercuts the competition ($\underline{p} - 1$) and the price that implies a profit ($V_A + 1$). So this term undercuts the competition but prevents prices that imply (expected) losses. The minimum between the asset value plus markup and the max term ensures the algo does not lower the sell order price unless the competition is undercutting it. Prices in the experiment are integers, so 1 is the minimum increment to undercut the competition and is the minimum positive profit. The buy order price formula is similar, except the price is outbidding rather than undercutting the competition.

[9] In future research, it might be interesting to investigate how the markup δ, which determines profits, affects market outcomes. We did not vary it in this experiment, because we needed enough statistical power for treatment comparisons.

[10] Our choice of delta, about 10% of the asset's pre-news expected value, is consistent with Asparouhova et al. (2020) who set their delta to 10% of the asset's expected dividend per period.

4.5 Treatments

4.5.1 Algorithmic trading

We have a baseline and two between-subject treatments that differ in the large trader type. All other traders are humans. The treatments are:

1. Baseline: The large trader is a human.

2. MarketAlgo: The large trader is the market-order algo.

3. LimitAlgo: The large trader is the limit-order algo.

In the two algo treatments, instructions inform participants that the large trader is played by the computer, but it does not explain the algo itself. Hence, traders are not told how the trading algorithms operate. We made this choice purposefully to mimic the situation of field traders who do not know the code of trading algorithms, which are kept private. Not describing the pseudo-code of the algorithm also simplified the instructions. Importantly, we gave traders numerous opportunities to learn about the design of the algo as sessions lasted for 20 trading rounds, preceded by one non-paid practice round, which allowed participants to familiarize themselves with the trading environment.

The baseline provides a valuable counterfactual that enables the study of the causal impact of algorithms. This is achieved by comparing our algorithmic treatments to baseline asset markets, where algorithms are completely absent (these markets no longer exist in the field).

In each market, the asset value realization and the news release time are drawn anew. Trader types are assigned at the beginning of the experiment and remain fixed throughout the experiment to facilitate learning. We use the same randomly pre-drawn sequence of liquidating values of the asset and news release times in all sessions, so that there are no differences between sessions and treatments.

4.5.2 Knowledge of news timing

In addition to the baseline and the two algo treatments, we vary the knowledge about the time of the news release within-subject on two levels: known and unknown. Specifically, in the known news release time treatment, which is modeled after scheduled firm-specific and macroeconomic announcements (e.g., earnings, unemployment rate, interest rates, etc.), the trading screen displays the exact time at which news will be released. In the unknown news release time treatment, which is modeled after unexpected events such as bankruptcies, pandemics or technological discoveries, the trading screen shows the time interval during which news will be released (see Figure 1). Traders are also reminded before a market starts whether the news release time is known or unknown. At the time of the news release and for a few seconds thereafter the screen displays "**Update:** Good news" or "**Update:** Bad news" in red in the upper right. This ensures the news release is salient to traders.

We administer the treatment so that for the baseline and the two algo treatments (MarketAlgo and LimitAlgo), the news release time is known in the first 10 trading rounds and unknown in the last 10 in half the sessions, whereas the reverse order applies to the other half of the sessions. This design allows us to control for order effects.

In sum, we conduct a mixed design with three between-subject dimensions (no algo, a market-order algo, or a limit-order algo) and two within-subject dimensions (known or unknown news release time).[11]

4.6 Post-experiment questionnaire

After the 20 rounds of trading, we administered a questionnaire, which can be found in full in Appendix B. To test Hypothesis 4, we administered 4 standard cognitive reflection test (CRT) questions adapted from Frederick (2005) and Toplak, West, and Stanovich (2014).

To test Hypothesis 5, we also included a battery of 14 questions on a 5-point Likert scale eliciting negative attitudes toward robots (NARS) (Syrdal et al. (2009)). For example, one question is "I would feel relaxed talking with robots", with responses from strongly disagree to strongly agree. We combine these responses into one robot aversion scale by taking the mean of the 14 responses (after inverting the scale for positively phrased items and keeping the scale for negatively phrased items).

4.7 Implementation

All sessions were run at Chapman University in the Spring and Fall terms of 2022. We had 10 groups of traders for the baseline and for each of the two algo treatments. A group consists of 7 human participants for the algo treatments and 8 for the baseline. Some sessions had two groups of traders participating in the same treatment simultaneously in the lab. In these sessions, participants did not interact across groups and remained matched to their own group.

[11] We chose a mixed design, because we wanted the presence of a specific algorithm to be between subjects. This prevents inattentive participants from mistakenly believing they were interacting with humans instead of algos, and simplifies the experiment and instructions for subjects, as they only have to learn one configuration. We incorporated the news release time variation within subjects, because it is a small change that was easy to learn within-subject.

In total, we had 220 participants and 600 markets. The experimental design, the sample size, and the main statistical tests were pre-registered based on a power calculation.[12]

The experiment was run in z-tree (Fischbacher, 2007). Sessions typically lasted just under 90 minutes. The payoff of one of the 20 markets was randomly chosen at the end of the experiment to compute participants' payments. The average payment per participant was \$28.06, including a show-up payment of \$7.

5 Results

As noted in the Introduction, a key advantage of the experimental approach is that it allows us to analyze the impact of trading algorithms on welfare, which is typically not observable in the field. Thus, we start with our welfare results, and proceed with our results on standard market outcomes: trading volume, liquidity, price efficiency, and volatility.

5.1 The effect of trading algorithms on welfare

In this subsection, we focus on welfare as the main outcome at the market level. As documented in the pre-registration, we use random effects regressions to estimate treatment effects in order to take advantage of the panel structure of the data[13]. Table 1 displays the estimates.[14]

Welfare. A primary question is how the presence of trading algorithms in the market affects welfare. Welfare is defined as the sum of payoffs of all traders, and welfare increases with the number of mutually beneficial trades. Welfare includes the payoff of the algorithmic trader in the algorithmic trading treatments, since it replaces a human trader from the baseline. Column 1 of Table 1 shows that the limit-order algorithm significantly improves welfare over markets with only human traders (see LimitAlgo coefficient), thus supporting Hypothesis 3i. Hence, the *right kind* of trading algorithm not only does not hurt, but can actually improve welfare. In line with Hypothesis 3i, LimitAlgo also significantly improves welfare compared to MarketAlgo (Wald-test, $\chi^2(1) = 5.67$, $p = 0.017$). Finally, welfare in MarketAlgo does not significantly differ from markets with only human traders.

Using expected welfare at the initial endowments as a benchmark and expected welfare once all Pareto-optimal trades are made as the maximum achievable value, the LimitAlgo treatment realizes about 60% of possible welfare gains on average compared to 43% and 35%

[12] See `https://www.socialscienceregistry.org/trials/9001` for the pre-registration.

[13] We cannot use fixed effects regressions to estimate the treatment effects, since the algo treatments are assigned between subjects. We ran Hausman tests comparing OLS as the consistent model and RE to test the validity of RE; the validity is never rejected even at the 10% level.

[14] The summary statistics for all market outcomes are in Appendix A.1.

Table 1: Treatment effects on welfare, random effects regressions

	(1)	(2)
Dependent var.	Welfare	Small Trader Welfare
MarketAlgo	16.950	-85.690***
	(19.322)	(32.687)
LimitAlgo	51.850***	1.895
	(16.266)	(29.785)
KnownNewsTime	0.200	8.773
	(9.513)	(13.856)
Constant	6308.900***	4466.725***
	(17.601)	(35.990)
Control order	Yes	Yes
Observations	600	600
Clusters	30	30

Note: Welfare is the sum of payoffs among all traders in the market. Small Trader Welfare is the sum of payoffs among all traders who are not the large trader. MarketAlgo, LimitAlgo, and KnownNewsTime are dummies for markets in which there is a market-order algo, a limit order-algo, and a known news release time, respectively. The base category is the baseline with only human traders. Control order is a dummy which equals 1 for groups who started with a known news release time. The unit of observation is a single market. Standard errors are displayed below the coefficients and are clustered at the group level. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

for MarketAlgo and the baseline.[15] In sum, the limit-order algo increases welfare substantially compared to a market with only human traders and compared to the no-trade benchmark (i.e., initial endowments).

Small (human) trader welfare. To assess the impact of algorithmic trading on the welfare of human traders, we compute the variable Small Trader Welfare, which is the sum of payoffs of all traders except the large trader. These *small* traders are humans in all treatments. Consequently, this variable allows us to cleanly compare the same number of human traders across treatments. Column 2 in Table 1 displays the results.

Even though the LimitAlgo treatment significantly increases overall welfare, it does not significantly increase the welfare of human traders compared to the baseline (see LimitAlgo in Column 2). Hence, the additional welfare that the limit-order algo generates is almost entirely captured by the algo itself. A field translation of these results is as follows. The market is

[15] The expected welfare if all traders kept their initial endowments is $5*360+5*400+5*4*56+5*4*66 = 6240$ (cash low types+cash high types+expected liquidating value low types+expected liquidating value high types). The expected welfare if all Pareto-optimal trades are made is $5*360+5*400+10*4*66 = 6440$. The expected asset value for the low type, given the random draws used over the 20 trading rounds, is 56, not 50, since there were slightly more good news realizations than bad news realizations.

initially populated by small human retail traders and one large human trader, who trades for a hedge fund; this is our baseline. The hedge fund then replaces its human trader with a limit-order algo; this is our limit-order algo treatment. The hedge fund now makes more profit, without hurting or helping the average retail investor.

Column 2 in Table 1 also shows that market-order algos (see MarketAlgo in Column 2) significantly decrease the payoffs of human traders, compared to the baseline. The drop in profits among human traders due to the market-order algo (−86 ECUs, that is −2% of the payoffs compared to the baseline) is larger than the welfare gains generated by the limit-order algo (+52 ECUs or less than +1% of payoffs). This noticeable drop in Small Trader Welfare with the market-order algo is a strong indication that the market-order algo exploits human traders. Consequently, market-order algos achieve payoffs (net of initial endowment) that are 49% higher than limit-order algos (Wilcoxon Rank Sum Test, $p = 0.036$). Because the decrease in Small Trader Welfare is larger than the increase in Welfare for the market-order algo, we find experimental support for Hypothesis 3ii.

In addition, welfare among human traders is significantly lower in MarketAlgo than in LimitAlgo (Wald-test, $\chi^2(1) = 16.65$, $p < 0.001$). Consequently, using the field translation again, if a hedge fund operating in a market with only human traders switches from employing a human trader to a market-order algo, it will extract higher profits at the expense of retail investors, without increasing welfare in the market. The introduction of trading algorithms to a market might therefore hurt retail investors or leave them as well off, depending on the type of trading algo. Thus, while passive trading algos do not, in general, harm human traders, aggressive algos might.

Result 1 (**Welfare effects**).

- *Compared to human-only markets, the introduction of a limit-order algo increases welfare, from about 35% to 60% of potential welfare gains. The surplus is earned almost exclusively by the algo, and human traders on average are as well off as in human-only markets.*

- *Compared to human-only markets, the introduction of a market-order algo does not change welfare. However, human traders are worse off, and the algo makes higher profits than the human it replaces.*

5.2 The effects of trading algorithms on liquidity, price efficiency, and price volatility

In this subsection, we investigate the effect of the trading algorithms on liquidity, as measured via bid-ask spreads and trading volume, price efficiency, and volatility. These financial variables are potential candidates to help us explain how the trading algorithms affect welfare. Except for price volatility, all of these regressions were pre-registered.

Bid-Ask Spread. Column 1 of Table 2 shows the treatment effects on the bid-ask spread, which is calculated as the time-weighted average of the bid-ask spread in a market. The market-order algo does not lead to significantly smaller bid-ask spreads compared to human-only markets, whereas the limit-order algo does so (see MarketAlgo and LimitAlgo in Column 1). In fact, the limit-order algo reduces the bid-ask spread to practically zero. Indeed, the limit-order algo is programmed to offer a spread of 10 ECUs absent competition, and this spread can only decrease if human traders compete, which explains the tight spreads in this treatment. In addition, the limit-order algo significantly reduces bid-ask spreads compared to the market-order algo (Wald-test, $\chi^2(1) = 160.58$, $p < 0.001$). These findings offer partial support for Hypothesis 1i, where the limit-order algo increases bid-ask spreads, whereas the market-order algo does not reduce them.

Trading volume. Column 2 in Table 2 estimates the effect of the trading algorithms on trading volume. The market-order algo leads to a slightly lower trading volume (23.8) than the baseline, but the difference is not statistically significant (see MarketAlgo in Column 2). The limit order algo increases transactions relative to the baseline by 10%, from 25 to 27.5 trades, but not significantly so (see LimitAlgo in Column 2). Although consistent with Hypothesis 1i, the evidence is only directional and not statistically significant. However, trading volume is more than 20% higher in LimitAlgo than in MarketAlgo, and this difference is significant (Wald-test, $\chi^2(1) = 7.43$, $p = 0.006$). Thus, different kinds of trading algorithms can cause differences in trading volume.

Trading volume due to news. Number of News Trades in Column 3 of Table 2 measures the number of trades within 1 second of news release. While some human traders might be fast enough to react within a second to the new asset value, the speed advantage of the market-order algo is evident within this time window. The market-order algo immediately accepts limit orders that fail to reflect the news and imply a profit.[16]

As a consequence, the regression in Column 3 shows that the market-order algo generates 2.5 more trades on average just after news release, which is three times more new trades than in human-only markets. This difference is statistically significant (see MarketAlgo in Column 3). Moreover, these news trades are almost exclusively stale human limit orders accepted at the human traders' expense, as the increase of 2.5 trades immediately after news release (relative to the baseline) barely changes to an increase of 2.4 trades once only stale limit orders are counted (not displayed in the table). In most cases where the market-order algo accepts a stale

[16]The limit-order algo does not pick off stale limit orders, but only updates its own offers.

Table 2: Treatment effects for main outcome variables, random effects regressions

	(1)	(2)	(3)	(4)	(5)
Dependent var.	Bid-Ask Spread	Trading Volume	Number of News Trades	Price Efficiency	Price Volatility
MarketAlgo	1.991	-2.200	2.545***	0.140**	-74.390***
	(2.640)	(2.377)	(0.384)	(0.070)	(28.610)
LimitAlgo	-23.878***	2.485	-0.145	0.246***	-100.007***
	(2.271)	(2.199)	(0.165)	(0.055)	(27.124)
KnownNewsTime	0.638	1.207	-0.177	-0.031	12.137
	(1.029)	(1.127)	(0.132)	(0.033)	(12.762)
Constant	21.404***	25.062***	0.772***	0.506***	156.557***
	(2.568)	(2.265)	(0.189)	(0.057)	(25.038)
Control order	Yes	Yes	Yes	Yes	Yes
Observations	600	600	600	600	600
Clusters	30	30	30	30	30

Note: Bid-Ask Spread is the time weighted bid-ask spread in the market, where the lower (upper) bound is set to the minimum (maximum) asset valuation of 20 (90) in the absence of a bid (ask). Number of News Trades is the number of trades within 1 second of news release. Price Efficiency is the share of transaction prices between the current asset valuation of the low trader type and the high trader type. Price Volatility is the variance of prices around their pre-news and post-news means, respectively, i.e., the price variance controlled for the shift in means due to news release. MarketAlgo, LimitAlgo, and KnownNewsTime are dummies for markets in which there is a market-order algo, a limit order-algo, and a known news release time, respectively. The base category is the baseline with only human traders. Control order is a dummy which equals 1 for groups who started with certain news release time. The unit of observation is a single market. Standard errors are displayed below the coefficients and are clustered at the group level. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

limit order, it implies a loss for the human trader and a profit for the market-order algo.[17] In Appendix A.2, we show that human traders learn to avoid this exploitation over time, so that the difference in the number of news trades between the MarketAlgo treatment and the baseline shrinks in later rounds of the experiment.

Finally, the market-order algo generates significantly more news trades than the limit-order algo (Wald-test, $\chi^2(1) = 51.4$, $p < 0.001$). This result provides strong support for Hypothesis 1ii. In addition, the number of news trades is not significantly different between the limit-order algo and human-only markets (see LimitAlgo in Column 3).

[17] For example, a trader offers to buy a share of the asset for 50 ECUs before news release. The algo values the asset at 50 ECUs so that accepting the buy offer implies no profit and the offer remains in the order book. After the release of bad news, the algo valuation of the asset drops to 20 ECUs so that accepting the now stale buy limit order implies a profit of 30 ECUs. If the offering human trader has the same trader type, then they incur a loss of 30 ECUs, otherwise they incur a loss of 20 ECUs.

Result 2 (**Liquidity**).

- *Compared to human-only markets, the introduction of a limit-order algo lowers bid-ask spreads while a market-order algo does not impact the bid-ask spread.*

- *Compared to human-only markets, the introduction of a limit-order algo neither significantly increases trading volume nor the number of news trades. The limit-order algo generates more trading volume than the market-order algo.*

- *Compared to human-only markets, the introduction of a market-order algo does not change trading volume, but it increases the number of news trades.*

Price efficiency. Unlike in common value settings, our setting is such that there is not a single informationally efficient price under the assumption of risk neutrality.[18] Instead, there is a range of efficient prices between the valuation of the low and high trader types.[19] The variable Price Efficiency in Column 4 of Table 2 measures the share of transaction prices in this informationally efficient interval. In the human-only markets, only about 51% of transactions are at efficient prices. Both trading algos increase this share significantly, by 14 percentage points to about 65% for the market-order algo and by 25 percentage points to about 75% for the limit-order algo, and these differences are significant (see MarketAlgo and LimitAlgo in Column 4). These findings support Hypothesis 2i.

The difference in Price Efficiency between the limit-order algo and market-order algo is also statistically significant (Wald-test, $\chi^2(1) = 4.13$, $p = 0.042$), which provides support for Hypothesis 2ii. The limit-order algo generates the highest price efficiency, because it is programmed to offer prices in the efficient interval. The market-order algo, on the other hand, accepts (human-made) limit orders as long as they imply a profit. These profitable trades are not necessarily informationally efficient, as is the case when market-order algos pick off stale limit orders. As a result, price efficiency tends to be lower in MarketAlgo than in LimitAlgo.

Price volatility. Column 5 of Table 2 estimates the treatment effects on price volatility.[20] In line with Hypothesis 2iii, the regression shows that the limit-order algo significantly decreases price volatility by about 64% compared to human-only markets (see LimitAlgo in Column 5).

[18] Prior to news release, we assume risk neutrality to define informationally efficient prices. After news release, there is no uncertainty left, so no assumptions about risk preferences are necessary.

[19] These are also the prices that generate Pareto-efficient trades.

[20] Since news release shifts the asset value, prices also shift after news release. To control for this exogenous shift in mean prices, we calculate Price Volatility $= \sum_i (p_i - \bar{p}_i)^2/N_i + \sum_j (p_j - \bar{p}_j)^2/N_j$, where i indexes all transactions prior to news release and j indexes transactions after news release, $\bar{p}_i, \bar{p}_j$ are the mean prices before and after news release, respectively, and p_i, p_j are the transaction prices. Hence, our Price Volatility variable takes into account the shift in mean prices due to the news release. In the field, the change in fundamental value is typically not observed, whereas our lab measure can control for it and isolate price volatility from fundamental changes.

This result follows from the fact that the limit-order algo is programmed to offer the same prices absent competition and news release. It will only change the offered prices to react to competing offers and to incorporate news.

In line with Hypothesis 2iii, the market-order algo also significantly decreases price volatility by approximately 47% compared to human-only markets (see MarketAlgo in Column 5). The market-order algo does not set prices itself and merely accepts offers in the book, which are set by humans. Hence, the large magnitude of the decrease in price volatility might appear surprising. However, as we show in the learning regressions (see Appendix A.2), the market-order algo exploits human traders, who adapt by setting more reasonable prices with experience (especially around news time), whereas there is no similar pressure to adapt in human-only markets. Moreover, the market-order algo is very systematic in that it only accepts buy (sell) orders that are higher (lower) than the fundamental value. By contrast, humans can make mistakes in accepting offers, driving up the variance of observed prices. These two explanations can partially account for the lower price volatility in market-order algo markets. Price volatility is lower in limit-order algo markets compared to market-order algo markets, but not significantly so (Wald-test, $\chi^2(1) = 3.47$, $p = 0.063$). Both algos reducing volatility is consistent with Hypothesis 2iii.

Result 3 (**Price efficiency and volatility**).

- *Compared to human-only markets, the introduction of a limit-order algo significantly increases price efficiency and reduces volatility.*
- *Compared to human-only markets, the introduction of a market-order significantly increases price efficiency (less than the limit-order algo) and reduces price volatility.*

In Appendix A.3, we complete our analysis of the main outcome variables by showing that news release times (known vs unknown) do not have a significant effect. Appendix A.2 analyzes whether there is learning in the experiment.

5.3 How do limit-order algorithms increase welfare?

Our main welfare result is that limit-order algos significantly increase welfare compared to the baseline. Our goal in this section, which was not pre-registered, is to uncover the mechanisms underlying this key finding. The standard market quality measures, i.e., price efficiency, trading volume, bid-ask spread, and price volatility, are natural candidates to explain welfare differences. Regressing welfare on these four variables, using data from the LimitAlgo and baseline treatments, we find that only price efficiency and trading volume are significant predictors, regardless of whether we add the LimitAlgo dummy or omit it (see Table 3). These exploratory

Table 3: Welfare and market quality, random effects regressions

	(1)	(2)
Dependent var.	Welfare	Welfare
Trading Volume	2.939***	2.944***
	(0.456)	(0.456)
Price Efficiency	151.751***	154.013***
	(22.322)	(21.932)
Bid-Ask Spread	0.306	0.210
	(0.231)	(0.239)
Price Volatility	<0.001	-0.002
	(0.023)	(0.023)
LimitAlgo	14.541	
	(15.379)	
KnownNewsTime	3.636	3.723
	(4.352)	(4.318)
Constant	4710.257***	4716.893***
	(25.819)	(26.197)
Control order and π-realization	Yes	Yes
Observations	400	400
Clusters	20	20

Note: Welfare is the sum of payoffs among all traders in the market. MarketAlgo, LimitAlgo, and KnownNewsTime are dummies for markets in which there is a market-order algo, a limit order-algo, and a known news release time, respectively. The base category is the baseline with only human traders. Control order is a dummy which equals 1 for groups who started with a known news release time. The unit of observation is a single market. Standard errors are displayed below the coefficients and are clustered at the group level. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

regressions are also associated with a high $R^2 = 0.998$, thus suggesting that our market measures predict welfare with great accuracy, and that we are not missing critical explanatory variables.

Therefore, price efficiency and trading volume are reasonable candidates for explaining welfare differences across treatments. Price efficiency captures the informativeness of prices, which will permit the completion of transactions that are mutually beneficial. By definition, any transaction that is completed at an informationally efficient price is mutually beneficial, as long as the buyer is a high type and the seller is a low type. By contrast, transactions at prices that are not informationally efficient cannot be mutually beneficial. It follows that price efficiency measures the propensity with which a transaction is welfare-enhancing. We can thus refer to price efficiency as an indicator of the *quality* of a transaction. Ultimately, markets will increase welfare when they facilitate a high *quantity* of high *quality* transactions. We thus expect trading volume to have a positive impact on welfare, because it will tend to increase the number of welfare-enhancing transactions. Given the field literature, it might be surprising that the bid-ask spread plays less of a role in explaining welfare.

To ensure that price efficiency and trading volume are the only two relevant mechanisms through which the LimitAlgo treatment impacts welfare, we start by estimating a structural equation model using as mediators our four measures of market quality: price efficiency, trading volume, bid-ask spread, and price volatility. However, this model specification (see Figure 3 in Appendix A.4) underperforms the two-mediator model in Figure 2 in terms of information criteria, because the slightly improved fit does not compensate for the increased number of estimated parameters. Hence, we use the two-mediator model with price efficiency and trading volume.

In Figure 2, we report the standardized estimates of the structural model along with significance levels using bootstrapped standard errors (5,000 replications) and Maximum Likelihood (ML).[21] To alleviate any potential omitted variable bias that would lead to an overestimation of the mediation of welfare (see MacKinnon and Pirlott (2015); Rohrer et al. (2022)), we introduce price volatility and bid-ask spread as controls when evaluating the direct welfare effect of the two mediators (price efficiency and trading volume), and the LimitAlgo dummy. In line with previous regressions (see Table 2), we also add a control variable for news timing order and KnownNewsTime as regressors. Following the comprehensive structural equation modeling approach (MacKinnon and Pirlott (2015)), our model incorporates the main relevant variables that have been previously identified in the market microstructure literature to characterize market quality (see Foucault et al. (2013)).[22,23]

Our estimates show that limit-order algos positively and significantly impact both price efficiency and trading volume. Furthermore, price efficiency and trading volume have a positive and significant direct impact on welfare, thus confirming they are valid mediators of the impact of limit-order algos. We find that 74% of the impact of limit-order algos is mediated by the two mediators, so the direct effect of the LimitAlgo treatment on welfare is not significant at

[21] We obtain similar results when introducing random effects using multilevel structural equation modeling techniques following Krull and MacKinnon (2001) and Preacher, Zyphur, and Zhang (2010).

[22] Our experimental method makes this approach especially convincing, because our data includes all features of the trading interface. It is difficult to think of a variable that would affect price efficiency, trading volume, and welfare separately and that is left outside of the current statistical model.

[23] Furthermore, following Rohrer et al. (2022), we conduct a robustness estimation in which we add the interaction terms between the two moderators (price efficiency and trading volume), volatility, and bid-ask spreads as controls. We also estimate the model with fixed effects for groups in line with the recommendation of Rohrer et al. (2022). All these robustness checks produce similar results. Finally, we conduct sensitivity tests (Cinelli and Hazlett (2020)) that suggest an omitted variable would have to correlate substantially (more than 40% [20%]) with welfare to explain away our mediation findings via price efficiency [trading volume]. Consequently, we expect the omitted variable bias to be negligible in our setup and the mediation analysis to be valid.

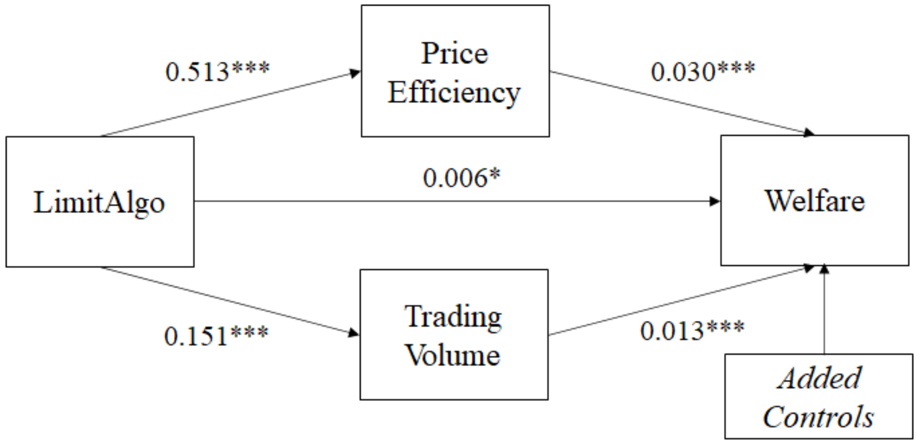


Figure 2: Structural equation model with standardized coefficient estimates for LimitAlgo and the baseline. *Added Controls*: Bid-Ask Spread and Price Volatility. All regressions control for KnownNewsTime, order and π-realization. Arrows represent direct effects. Standard criteria (see Williams, Vandenberg, and Edwards (2009)) show adequate fit for valid measurements given our moderate sample size and the use of ML estimation: Likelihood Ratio Test, $\chi^2(1) = 116.28, p < 0.001$, SRMR (Standardized Root Mean Residual) = 0.055 [< 0.100], CFI (Comparative Fit Index) = 0.977 [>0.950] (see Shah and Goldstein (2006)). AIC = 16994.650 and BIC = 17082.462. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level. The estimated equations are as follows, where i is a market in the LimitAlgo treatment or baseline:

$$
\begin{aligned}
\text{Price Efficiency}_i &= \alpha_1 + \beta_1 \text{LimitAlgo}_i + \delta_1 \text{KnownNewsTime}_i + e_{1,i} \\
\text{Trading Volume}_i &= \alpha_2 + \beta_2 \text{LimitAlgo}_i + \delta_2 \text{KnownNewsTime}_i + e_{2,i} \\
\text{Welfare}_i &= \alpha_3 + \beta_{3,1} \text{LimitAlgo}_i + \beta_{3,2} \text{Price Efficiency}_i \\
&\quad + \beta_{3,3} \text{Trading Volume}_i + \beta_{3,4} \text{Bid-Ask Spread} \\
&\quad + \beta_{3,5} \text{Price Volatility} + \delta_3 \text{KnownNewsTime}_i + e_{3,i}
\end{aligned}
$$

standard levels. Price efficiency accounts for 88% of the mediation and trading volume accounts for 12%.

Result 4 (**Welfare channels**).

- *Price efficiency and trading volume are the only significant predictors of welfare out of the set of commonly-used market quality variables.*
- *About 74% of the limit-order algo welfare effect can be explained by its positive effect on price efficiency (88%) and by its positive effect on trading volume (12%).*

We used a similar approach to explain why the market-order algo has a negative effect on Small Trader Welfare relative to the baseline (see Figure 4 in Appendix A.4). We show that the increase in the number of news trades due to the market-order algo explains 93% of the decrease in Small Trader Welfare, whereas the decrease in the number of no-news trades (i.e., Trading Volume minus News Trades) accounts for only 7% of the negative effect. Increased price efficiency is a partially offsetting positive effect.

5.4 Trader characteristics: Cognitive ability and robot aversion

In this section, we investigate how trading algorithms interact with traders of different individual characteristics.

Table 4, shows descriptive statistics for individual trader measures across CRT levels. In line with Hypothesis 4i, we observe that high-CRT traders, defined as scoring at least one correct answer (the median score) on the CRT, earn more during trading than low-CRT traders, defined as scoring below the median (Rank Sum Test, $p < 0.001$). In line with Hypothesis 4ii, the relative difference in payoffs is higher in the baseline (9%) than in the algo treatments (3%).

Interestingly, low-CRT traders complete 36% more trades (Mean = 6.13) than high-CRT traders (Mean = 4.52) in the baseline (Rank Sum Test, $p < 0.001$). But, in algo treatments, low-CRT traders decrease their trading volume by 26% (Mean = 4.53) (Rank Sum Test, $p < 0.001$), whereas high-CRT traders slightly increase it (Mean = 4.75) (Rank Sum Test, $p = 0.625$). This implies that high-CRT traders are often transacting with algos (38% of the time), so their trades are on average less profitable in the algo treatments than in the baseline (see Trade ratio with Algo in Table 4). By contrast, low-CRT traders transact less (Mean = 1.60 vs Mean = 2.15) and are less often involved in trades with algos (32% of the time) than high-CRT traders (Rank Sum Tests, $p < 0.001$ for all comparisons).

In line with these results, the Mistakes Index of high-CRT traders, which measures the share of a trader's transactions that implies an expected loss given their information, increases from 23% in the baseline to 36% (27%) in the MarketAlgo (LimitAlgo), and the difference is significant (see Rank Sum Tests, $p < 0.001$ for both tests). By contrast, low-CRT traders transact more cautiously in algo treatments, thus not making more mistakes in the presence of limit-order (41%) and market-order (40%) algos than in the baseline (42%) (see Rank Sum Tests, $p = 0.267$ and $p = 0.203$). By doing so, low-CRT traders avoided being exploited by algos.

In Table 5, our aim is to test Hypothesis 4i using random effects regressions, following the specification stated in our pre-registration plan.[24] In line with the first part of Hypothesis 4i, traders with higher CRT scores earn more (see positive and significant CRT coefficient in Column 1) in human-only markets. Consistent with the second part of Hypothesis 4i, the interaction term between CRT and MarketAlgo, and between CRT and LimitAlgo, is negative

[24]We obtain similar results when controlling for gender and risk attitudes (see post-experimental questionnaire in Appendix B) in our regression analyses.

Table 4: Descriptive statistics individual measures across CRT levels

	High-CRT	Low-CRT	Rank Sum Tests	Total	High-CRT	Low-CRT	Rank Sum Tests	Total
	Payoffs		(p-value)		Trading Volume		(p-value)	
[1] Baseline	656.60	603.63	(<0.0001)	634.66	4.52	6.13	(<0.0001)	5.19
[2] LimitAlgo	646.52	623.33	(<0.0001)	634.93	5.33	4.91	(0.092)	5.12
[3] MarketAlgo	627.62	609.41	(<0.0001)	622.42	4.34	3.86	(0.359)	4.21
Total	642.30	613.22	(<0.0001)	630.67	4.68	5.08	(0.0006)	4.84
[1] vs [2]	(0.214)	(0.073)	-	(0.875)	(0.0002)	(<0.0001)	-	(0.937)
[1] vs [3]	(<0.0001)	(0.656)	-	(0.003)	(0.0003)	(<0.0001)	-	(<0.0001)
[2] vs [3]	(0.0006)	(0.139)	-	(0.005)	(<0.0001)	(<0.0001)	-	(<0.0001)

	High-CRT	Low-CRT	Rank Sum Tests	Total	High-CRT	Low-CRT	Rank Sum Tests	Total
	Mistakes Index		(p-value)		Trade ratio with Algo		(p-value)	
[1] Baseline	0.23	0.42	(<0.0001)	0.31	-	-	-	-
[2] LimitAlgo	0.27	0.41	(<0.0001)	0.34	0.42	0.36	(<0.0001)	0.38
[3] MarketAlgo	0.36	0.40	(0.084)	0.37	0.33	0.28	(0.0007)	0.34
Total	0.29	0.41	(<0.0001)	0.34	0.38	0.32	(<0.0001)	0.36
[1] vs [2]	(0.076)	(0.267)	-	(0.129)	-	-	-	-
[1] vs [3]	(<0.0001)	(0.203)	-	(0.0003)	-	-	-	-
[2] vs [3]	(<0.0001)	(0.675)	-	(0.042)	(0.0001)	(0.002)	-	(0.0001)

Note: High-CRT traders are defined as scoring one or more on the CRT, which corresponds to a score above the median. Low-CRT traders are defined as scoring zero on the CRT. For the sake of comparability across treatments, the trader with the large endowment is excluded from Baseline. Mistakes Index measures the share of a trader's transactions that imply a loss given their information. Trade ratio with Algo is the proportion of transactions a trader completes with an algo. The rows '[1] vs [2]' ('[1] vs [3]') '[2] vs [3]' report the p-values for Rank Sum Tests comparing Baseline and LimitAlgo, Baseline and MarketAlgo, LimitAlgo and MarketAlgo treatments.

and significant for payoffs (see Column 1). This shows that high-CRT traders earn less in algo treatments than in the baseline. This might be explained by the fact that traders with higher CRT scores complete more transactions in the presence of algos (see MarketAlgo $\times$ CRT, and LimitAlgo $\times$ CRT in Column 2 and CRT in Column 4). This implies that traders with higher CRT scores transact more often with algos (see the significantly positive CRT coefficient in Column 4). In line with this finding, traders with higher CRT scores make more mistakes in MarketAlgo than in the baseline (see positive and significant MarketAlgo $\times$ CRT in Column 3). The variable LimitAlgo $\times$ CRT in Column 3 is positive yet not significant, so it is mostly the market-order algo that induces more mistakes. This is consistent with the evidence above that the market-order algo exploits human traders just after news release.

Overall, participants with low CRT levels trade more cautiously in algo treatments, thus avoiding being exploited by algos. Note that this cautious behavior cannot be explained by negative attitudes toward robots (NARS), since it is included as a control variable in all the regressions in Table 5.[25]

[25] The correlation between CRT and NARS scores is close to zero ($\rho = -0.045$, $p = 0.504$).

Using the regressions in Table 5, we can also test Hypothesis 5 on robot aversion. In Column 1, we observe that LimitAlgo $\times$ NARS is negative and significant, whereas MarketAlgo $\times$ NARS is negative, yet not significant. It follows that traders who have a more negative attitude toward robots, that is a higher NARS score, earn less in algo treatments than in the baseline. The reason why they earn less is that they trade slightly more and make more mistakes (see LimitAlgo $\times$ NARS and MarketAlgo $\times$ NARS in Columns 2 and 3). That is, people who fear robots earn less in algo treatments, not due to their excessive cautiousness, but rather because they are more likely to be exploited by trading algorithms. Overall, our findings provide support for Hypothesis 5 except for the fact that negative attitudes toward robots do not discourage trading in algo treatments.

Table 5: Individual characteristics and main outcome variables, random effects regressions

	(1)	(2)	(3)	(4)
Dependent var.	Payoffs	Trading Volume	Mistakes Index	Trade ratio with Algo
MarketAlgo	54.096	-5.740**	-0.119	0.266
	(36.884)	(2.598)	(0.132)	(0.258)
MarketAlgo × CRT	-18.392***	0.866**	0.052**	-0.006
	(6.346)	(0.396)	(0.026)	(0.043)
CRT	26.065***	-0.642**	-0.094***	0.045**
	(5.563)	(0.297)	(0.018)	(0.021)
LimitAlgo	81.794**	-6.051**	-0.419**	
	(31.765)	(2.292)	(0.172)	
LimitAlgo × CRT	-14.205**	1.039***	0.020	
	(6.438)	(0.382)	(0.037)	
MarketAlgo × NARS	-15.573	1.176*	0.042**	-0.093
	(10.301)	(0.659)	(0.036)	(0.068)
NARS	15.061	-1.583***	-0.067**	0.009
	(6.659)	(0.569)	(0.031)	(0.037)
LimitAlgo × NARS	-20.516**	1.519*	0.128**	
	(9.634)	(0.829)	(0.061)	
KnownNewsTime	1.253	0.177	0.004	0.009
	(1.975)	(0.205)	(0.017)	(0.019)
Constant	563.883***	11.457***	0.589***	0.305**
	(24.883)	(2.262)	(0.100)	(0.126)
Control order	Yes	Yes	Yes	Yes
Observations	4,200	4,200	4,007	2,656
Clusters	210	210	210	140

Note: Payoffs is the total amount earned by a trader in a given market in ECU. Trading Volume is the number of trades completed by a given trader in a given market. Mistakes Index is the share of a trader's transactions that imply a loss given their current information. Trade ratio with Algo is the proportion of transactions a trader completed with an algo. MarketAlgo, LimitAlgo, and KnownNewsTime are dummies for participants who faced in a given trading round a market-order algo, a limit order-algo, and a known news release time, respectively. The base category is the baseline with only human traders. CRT is the score (between 0 and 4) on the Cognitive Refection Test. Similar results are obtained using the dummy variable High-CRT instead of CRT. Control order is a dummy which equals 1 for groups who started with certain news release time. The unit of observation is a trader in a given market. Standard errors are displayed below the coefficients and are clustered at the group level. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

Finally, we test Hypothesis 4ii by studying the impact of average CRT scores in a given market (Market CRT) on welfare and price efficiency. Thus, the analysis moves from the trader-level to the market-level. In line with Hypothesis 4ii, we show in Table 6 that baseline markets populated by traders with higher CRT scores produce higher welfare and price efficiency (see positive and significant variable Market CRT).

As conjectured in Hypothesis 4ii, the increase in welfare associated with Market CRT is reduced both in LimitAlgo and MarketAlgo (see negative and significant coefficient for LimitAlgo $\times$ CRT and negative yet not significant coefficient for MarketAlgo $\times$ CRT in Column 1). Furthermore, we cannot reject the hypothesis that the impact of CRT in both MarketAlgo and LimitAlgo is null (Coefficient Tests, MarketAlgo $\times$ CRT + CRT = 0, $p = 0.281$, and LimitAlgo $\times$ CRT + CRT = 0, $p = 0.429$). Hence, trading algorithms in the market neutralize the ability of high-CRT traders to boost Welfare. Regarding Price Efficiency, the increase associated with Market CRT is also reduced in LimitAlgo (see negative and significant coefficient for LimitAlgo $\times$ CRT in Column 2). Furthermore, we cannot reject the hypothesis that the impact of CRT on Price Efficiency in LimitAlgo is null (Coefficient Test, LimitAlgo $\times$ CRT + CRT = 0, $p = 0.499$). The coefficient for MarketAlgo $\times$ CRT is negative yet not significant, so the positive effect of CRT on Price Efficiency remains positive in the MarketAlgo treatment (Coefficient Test, MarketAlgo $\times$ CRT + CRT = 0, $p = 0.024$).[26]

We summarize our findings as follows.

Result 5 (**Cognitive ability**).

- *Traders with higher cognitive ability earn more, but this advantage is significantly reduced in the algo treatments.*
- *Markets populated by traders with higher cognitive ability exhibit higher price efficiency and welfare in the baseline, but this effect tends to vanish in the algo treatments.*

Result 6 (**Negative attitudes toward robots**).

- *Traders with more negative attitudes toward robots make more mistakes and earn less in the algo treatments than in the baseline, but they do not trade less.*

[26]Those same interactions with NARS instead of CRT are all statistically insignificant, which is why we did not add them in Table 6.

Table 6: CRT, welfare and price efficiency, random effects regressions

	(1)	(2)
Dependent var.	Welfare	Price Efficiency
MarketAlgo	82.632	0.242
	(60.596)	(0.174)
MarketAlgo × Market CRT	-80.509*	-0.168
	(47.211)	(0.157)
Market CRT	110.893***	0.355***
	(34.713)	(0.122)
LimitAlgo	170.603***	0.597***
	(45.153)	(0.131)
LimitAlgo × Market CRT	-129.766***	-0.381***
	(42.820)	(0.133)
KnownNewsTime	0.200	-0.031
	(9.537)	(0.033)
Constant	6,207.921***	0.176
	(40.384)	(0.111)
Control order	Yes	Yes
Observations	600	600
Clusters	30	30

Note: Welfare is the sum of payoffs among all traders in the market. Price Efficiency is the share of transaction prices between the current asset valuation of the low trader type and the high trader type, assuming risk neutrality. Market CRT is the average CRT score of human traders in a given market. MarketAlgo, LimitAlgo, and KnownNewsTime are dummies for markets in which there is a market-order algo, a limit order-algo, and a known news release time, respectively. The base category is the baseline with only human traders. Control order is a dummy which equals 1 for groups who started with certain news release time. The unit of observation is a single market. Standard errors are displayed below the coefficients and are clustered at the group level. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

6 Concluding remarks

Previous empirical research has not directly measured the causal effect of AT on welfare and its redistributive effects, thus leaving the issue open to heated debates. We use a novel experimental paradigm to provide the first estimate of the welfare gains and distributive consequences of AT. By replacing humans with market-order algorithms, welfare gains are largely absent, and markets redistribute profits in a way that benefits algorithms at the expense of human traders. By contrast, limit-order algos increase overall welfare without improving or harming the welfare of human traders on average. We identify price efficiency as the main driver of the positive impact of limit-order algos on welfare, whereas increased trading volume plays a more limited role. Interestingly, by enhancing price efficiency, algos reduce earnings differences between sophisticated traders and unsophisticated ones. This finding offers a novel argument for why AT may be advantageous for certain types of traders.

Our results show that standard measures of market quality are not necessarily good indicators of the welfare implications of AT for retail investors. In our experiments, market-order algos enhance price efficiency and lower price volatility, while reducing the welfare of small human traders.

Future research could also examine additional factors that are relevant to welfare analysis and that were not included in our study, such as sunk costs in technology enabling faster algorithm reaction times (see e.g., Biais, Foucault, and Moinas (2015)).

Our results also indicate that different types of trading algorithms can have very different consequences for retail investors in terms of welfare. To maintain focus, we have run separate markets for different types of algorithms. However, future research should investigate market outcomes when multiple types of algorithms interact in the market. Our study suggests that exchange rules and AT regulations should aim at increasing the limit-order to market-order algo ratio to materialize the welfare gains of AT. To increase this ratio, exchanges have used maker-taker pricing that requires market-order traders to pay fees while granting rebates to limit-order traders. This pricing scheme might be especially effective in attracting market-maker HFTs (see e.g., O'Hara (2015)).

References

Aite. 2021. Algorithmic trading survey: Hedge fund. *Report* 1–20.

Akiyama, E., N. Hanaki, and R. Ishikawa. 2017. It is not just confusion! strategic uncertainty in an experimental asset market. *Economic Journal* .

Aldrich, E. M., and K. López Vargas. 2020. Experiments in high-frequency trading: comparing two market institutions. *Experimental Economics* 23:322–52.

Angel, J. J., and D. McCabe. 2013. Fairness in financial markets: The case of high frequency trading. *Journal of Business Ethics* 112:585–95.

Angerer, M., T. Neugebauer, and J. Shachat. 2023. Arbitrage bots in experimental asset markets. *Journal of Economic Behavior & Organization* 206:262–78.

Aquilina, M., E. Budish, and P. O'Neill. 2022. Quantifying the high-frequency trading "arms race". *The Quarterly Journal of Economics* 137:493–564.

Asparouhova, E. N., P. Bossaerts, K. Rotaru, T. Wang, N. Yadav, and W. Yang. 2020. Humans in charge of trading robots: The first experiment. *SSRN 3569435* .

Asparouhova, E. N., X. Cai, D. Meloso, J. Nielsen, C. A. Parlour, and W. Yang. 2022. For better or for worse: Algorithmic choice in experimental markets. *SSRN* .

Banks, J., M. Olson, D. Porter, S. Rassenti, and V. Smith. 2003. Theory, experiment and the federal communications commission spectrum auctions. *Journal of Economic Behavior & Organization* 51:303–50.

Bao, T., E. Nekrasova, T. Neugebauer, and Y. E. Riyanto. 2022. Algorithmic trading in experimental markets with human traders: A literature survey. *Handbook of Experimental Finance* 302–22.

Berger, N., M. DeSantis, and D. Porter. 2020. The impact of high-frequency trading in experimental markets. *The Journal of Investing* 29:7–18.

Biais, B., and T. Foucault. 2014. Hft and market quality. *Bankers, Markets & Investors* 128:5–19.

Biais, B., T. Foucault, and S. Moinas. 2015. Equilibrium fast trading. *Journal of Financial Economics* 116:292–313.

BIS. 2022. Triennial central bank survey otc foreign exchange turnover. *BIS Monetary and Economic Department* .

Boehmer, E., K. Fong, and J. J. Wu. 2021. Algorithmic trading and market quality: International evidence. *Journal of Financial and Quantitative Analysis* 56:2659–88.

Boehmer, E., D. Li, and G. Saar. 2018. The competitive landscape of high-frequency trading firms. *The Review of Financial Studies* 31:2227–76.

Bosch-Rosa, C., and B. Corgnet. 2022. Cognitive finance. In *Handbook of Experimental Finance*, 73–88. Edward Elgar Publishing.

Bosch-Rosa, C., T. Meissner, and A. Bosch-Domènech. 2018. Cognitive bubbles. *Experimental Economics* 21:132–53.

Bossaerts, P., and C. Plott. 2002. The capm in thin experimental financial markets. *Journal of Economic Dynamics and Control* 26:1093–112.

Breaban, A., and C. N. Noussair. 2015. Trader characteristics and fundamental value trajectories in an asset market experiment. *Journal of Behavioral and Experimental Finance* 8:1–17.

Brogaard, J., B. Hagströmer, L. Nordén, and R. Riordan. 2015. Trading fast and slow: Colocation and liquidity. *The Review of Financial Studies* 28:3407–43.

Brogaard, J., T. Hendershott, and R. Riordan. 2014. High-frequency trading and price discovery. *The Review of Financial Studies* 27:2267–306.

———. 2019. Price discovery without trading: Evidence from limit orders. *The Journal of Finance* 74:1621–58.

Chaboud, A. P., B. Chiquoine, E. Hjalmarsson, and C. Vega. 2014. Rise of the machines: Algorithmic trading in the foreign exchange market. *The Journal of Finance* 69:2045–84.

Chakrabarty, B., P. C. Moulton, and X. F. Wang. 2022. Attention: How high-frequency trading improves price efficiency following earnings announcements. *Journal of Financial Markets* 57:100690–.

Chordia, T., T. C. Green, and B. Kottimukkalur. 2018. Rent seeking by low-latency traders: Evidence from trading on macroeconomic announcements. *The Review of Financial Studies* 31:4650–87.

Chordia, T., and B. Miao. 2020. Market efficiency in real time: Evidence from low latency activity around earnings announcements. *Journal of Accounting and Economics* 70:101335–.

Cinelli, C., and C. Hazlett. 2020. Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 82:39–67.

Conrad, J., S. Wahal, and J. Xiang. 2015. High-frequency quoting, trading, and the efficiency of prices. *Journal of Financial Economics* 116:271–91.

Corgnet, B., M. DeSantis, and D. Porter. 2018. What makes a good trader? on the role of intuition and reflection on trader performance. *The Journal of Finance* 73:1113–37.

———. 2020. The distribution of information and the price efficiency of markets. *Journal of Economic Dynamics and Control* 110:103671–.

———. 2021. Information aggregation and the cognitive make-up of market participants. *European Economic Review* 133:103667–.

Cueva, C., and A. Rustichini. 2015. Is financial instability male-driven? gender and cognitive skills in experimental asset markets. *Journal of Economic Behavior & Organization* 119:330–44.

Das, R., J. E. Hanson, J. O. Kephart, and G. Tesauro. 2001. Agent-human interactions in the continuous double auction. In *International joint conference on artificial intelligence*, vol. 17, 1169–78.

Eaton, G. W., T. C. Green, B. S. Roseman, and Y. Wu. 2022. Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of Financial Economics* 146:502–28.

Feldman, T., and D. Friedman. 2010. Human and artificial agents in a crash-prone financial market. *Computational Economics* 36.

Fischbacher, U. 2007. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10:171–8.

Foucault, T., J. Hombert, and I. Roşu. 2016. News trading and speed. *The Journal of Finance* 71:335–82.

Foucault, T., M. Pagano, A. Roell, and A. Röell. 2013. *Market liquidity: theory, evidence, and policy*. Oxford University Press.

Frederick, S. 2005. Cognitive reflection and decision making. *Journal of Economic perspectives* 19:25–42.

Frydman, C., N. Barberis, C. Camerer, P. Bossaerts, and A. Rangel. 2014. Using neural data to test a theory of investor behavior: An application to realization utility. *The Journal of Finance* 69:907–46.

Gjerstad, S. 2007. The competitive market paradox. *Journal of Economic Dynamics and Control* 31:1753–80.

Grinblatt, M., M. Keloharju, and J. Linnainmaa. 2011. Iq and stock market participation. *The Journal of Finance* 66:2121–64.

Grinblatt, M., M. Keloharju, and J. T. Linnainmaa. 2012. Iq, trading behavior, and performance. *Journal of Financial Economics* 104:339–62.

Hagströmer, B., and L. Nordén. 2013. The diversity of high-frequency traders. *Journal of Financial Markets* 16:741–70.

Hasbrouck, J., and G. Saar. 2013. Low-latency trading. *Journal of Financial Markets* 16:646–79.

Hautsch, N., M. Noé, and S. S. Zhang. 2017. The ambivalent role of high-frequency trading in turbulent market periods. *CFS Working paper* .

Hendershott, T., C. M. Jones, and A. J. Menkveld. 2011. Does algorithmic trading improve liquidity? *The Journal of Finance* 66:1–33.

Hu, G. X., J. Pan, and J. Wang. 2017. Early peek advantage? efficient price discovery with tiered information disclosure. *Journal of Financial Economics* 126:399–421.

Khapko, M., and M. Zoican. 2021. Do speed bumps curb low-latency investment? evidence from a laboratory market. *Journal of Financial Markets* 55:100601–.

Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun. 2017. The flash crash: High-frequency trading in an electronic market. *The Journal of Finance* 72:967–98.

Kirilenko, A. A., and A. W. Lo. 2013. Moore's law versus murphy's law: Algorithmic trading and its discontents. *Journal of Economic Perspectives* 27:51–72.

Korajczyk, R. A., and D. Murphy. 2019. High-frequency market making to large institutional trades. *The Review of Financial Studies* 32:1034–67.

Krull, J. L., and D. P. MacKinnon. 2001. Multilevel modeling of individual and group level mediated effects. *Multivariate behavioral research* 36:249–77.

Leal, J. S., and N. Hanaki. 2018. Algorithmic trading, what if it is just an illusion? evidence from experimental financial markets. *Working Paper* .

Lewis, M. 2015. *A wall street revolt*. WW Norton & Company.

Li, S., X. Wang, and M. Ye. 2021. Who provides liquidity, and when? *Journal of Financial Economics* 141:968–80.

MacKinnon, D. P., and A. G. Pirlott. 2015. Statistical approaches for enhancing causal interpretation of the m to y relation in mediation analysis. *Personality and Social Psychology Review* 19:30–43.

Malceniece, L., K. Malcenieks, and T. J. Putniņš. 2019. High frequency trading and comovement in financial markets. *Journal of Financial Economics* 134:381–99.

Menkveld, A. J. 2016. The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics* 8:1–24.

Neugebauer, T., J. Shachat, and W. Szymczak. 2023. A test of the modigliani-miller theorem, dividend policy and algorithmic arbitrage in experimental asset markets. *Journal of Banking & Finance* 106814.

Noussair, C. N., S. Tucker, and Y. Xu. 2016. Futures markets, cognitive ability, and mispricing in experimental asset markets. *Journal of Economic Behavior & Organization* 130:166–79.

O'Hara, M. 2015. High frequency market microstructure. *Journal of Financial Economics* 116:257–70.

Page, L., and C. Siemroth. 2021. How much information is incorporated into financial asset prices? experimental evidence. *The Review of Financial Studies* 34:4412–49.

Peng, Y., J. Shachat, L. Wei, and S. S. Zhang. 2020. Speed traps: Algorithmic trader performance under alternative market structures. *ESI Working Paper 20-39* .

Preacher, K. J., M. J. Zyphur, and Z. Zhang. 2010. A general multilevel sem framework for assessing multilevel mediation. *Psychological methods* 15:209–.

Rassenti, S. J., V. L. Smith, and B. J. Wilson. 2003. Controlling market power and price spikes in electricity networks: Demand-side bidding. *Proceedings of the National Academy of Sciences* 100:2998–3003.

Reuters. 2013. Cme group sparked shutdown of cftc's academic research program `https://www.reuters.com/article/us-cftc-cme-research-idUSBRE93N0YN20130424`.

Riordan, R., and A. Storkenmaier. 2012. Latency, liquidity and price discovery. *Journal of Financial Markets* 15:416–37.

Rohrer, J. M., P. Hünermund, R. C. Arslan, and M. Elson. 2022. That'sa lot to process! pitfalls of popular path models. *Advances in Methods and Practices in Psychological Science* 5:25152459221095827–.

SEC. 2020. Staff report on algorithmic trading in u.s. capital markets. *SEC Report* .

Shah, R., and S. M. Goldstein. 2006. Use of structural equation modeling in operations management research: Looking back and forward. *Journal of Operations management* 24:148–69.

Shields, G. S., M. A. Sazma, and A. P. Yonelinas. 2016. The effects of acute stress on core executive functions: A meta-analysis and comparison with cortisol. *Neuroscience & Biobehavioral Reviews* 68:651–68.

Shiller, R. J. 1981. Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review* 71:421–36. ISSN 00028282.

———. 2014. Speculative asset prices. *American Economic Review* 104:1486–517.

Shkilko, A., and K. Sokolov. 2020. Every cloud has a silver lining: Fast trading, microwave connectivity, and trading costs. *The Journal of Finance* 75:2899–927.

Smith, V. L. 1962. An experimental study of competitive market behavior. *Journal of Political Economy* 70:111–37.

———. 1994. Economics in the laboratory. *Journal of economic perspectives* 8:113–31.

———. 2003. Constructivist and ecological rationality in economics. *American economic review* 93:465–508.

———. 2008. Discovery: A memoir .

Smith, V. L., G. L. Suchanek, and A. W. Williams. 1988. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica* 1119–51.

Snowberg, E., and L. Yariv. 2021. Testing the waters: Behavior across participant pools. *American Economic Review* 111:687–719.

Stiglitz, J. 2014. Tapping the brakes: Are less active markets safer and better for the economy? *Federal Reserve Bank of Atlanta, Financial Markets Conference* .

Syrdal, D. S., K. Dautenhahn, K. L. Koay, and M. L. Walters. 2009. The negative attitudes towards robots scale and reactions to robot behaviour in a live human-robot interaction study. *Adaptive and emergent behaviour and complex systems* .

Thoma, V., E. White, A. Panigrahi, V. Strowger, and I. Anderson. 2015. Good thinking or gut feeling? cognitive reflection and intuition in traders, bankers and financial non-experts. *PloS one* 10:e0123202–.

Toplak, M. E., R. F. West, and K. E. Stanovich. 2014. Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & reasoning* 20:147–68.

Veiga, H., and M. Vorsatz. 2009. Price manipulation in an experimental asset market. *European Economic Review* 53:327–42.

Veiga, H., M. Vorsatz, et al. 2010. Information aggregation in experimental asset markets in the presence of a manipulator. *Experimental Economics* 13:379–.

Weitzel, U., C. Huber, J. Huber, M. Kirchler, F. Lindner, and J. Rose. 2020. Bubbles and financial professionals. *The Review of Financial Studies* 33:2659–96.

Williams, L. J., R. J. Vandenberg, and J. R. Edwards. 2009. 12 structural equation modeling in management research: A guide for improved analysis. *The Academy of Management Annals* 3:543–604.

Yu, R. 2016. Stress potentiates decision biases: A stress induced deliberation-to-intuition (sidi) model. *Neurobiology of Stress* 3:83–95.

A Appendix: Additional results

A.1 Market level summary statistics

	Mean	SD	N
Baseline			
Welfare	6300.10	1176.69	200
Small Trader Welfare	4442.60	882.58	200
Trading Volume	22.95	8.88	200
Number of News Trades	0.84	0.87	200
Price Efficiency	0.48	0.24	200
Bid-Ask Spread	23.00	10.32	200
Price Volatility	171.89	141.92	200
Limit-order algo treatment			
Welfare	6351.95	1155.57	200
Small Trader Welfare	4444.50	915.98	200
Trading Volume	25.43	7.32	200
Number of News Trades	0.70	0.76	200
Price Efficiency	0.72	0.17	200
Bid-Ask Spread	-0.88	9.80	200
Price Volatility	71.88	69.38	200
Market-order algo treatment			
Welfare	6317.05	1149.27	200
Small Trader Welfare	4356.91	864.84	200
Trading Volume	20.75	9.40	200
Number of News Trades	3.39	2.52	200
Price Efficiency	0.62	0.24	200
Bid-Ask Spread	24.99	8.83	200
Price Volatility	97.50	95.25	200

A.2 Learning effects

Did human traders improve during the experiment, as they learned from their interaction with the trading algos? Table 7 displays regressions with all pre-registered outcome variables that test for time trends, by interacting the treatment dummies with the number of rounds played. The un-interacted Round estimate is the time trend in the human-only markets. A significantly positive interaction term indicates that the outcome variable increased more over time in the respective algo treatment, relative to the baseline. We also control for the asset value realization.[27]

In the pre-registration, we hypothesized that learning might be stronger in the market-order algo markets compared to the limit-order algo markets, because the market-order algo is more exploitative. Thus, we expected human traders to learn to avoid the exploitation over time, whereas such exploitation does not occur with the limit-order algo. Column 1 in Table 7 presents the estimates for welfare in the market. Welfare in the market order markets increases by about 3 ECUs per round more than in the baseline, a highly statistically significant difference. By comparison, there is no significant difference to the baseline in the limit order market time trends. Thus, there is significantly more learning in the market-order algo markets than in the limit-order algo markets (Wald-test, $\chi^2(1) = 9.54$, $p = 0.002$), and our hypothesis that the market-order algo forces more learning finds experimental support. Welfare increases by about 3.3 ECUs per round in the baseline on average, so as traders gain experience, welfare on average increases over time in all treatments, but the increase is strongest in the market-order algo markets. In Column 2 in Table 7 with welfare only among human traders, there is similarly stronger learning in the market-order algo vs limit-order algo markets (Wald-test, $\chi^2(1) = 6.94$, $p = 0.008$), but the difference in learning between market-order algo markets and human-only markets is not statistically significant.

There are no interesting learning patterns regarding the number of trades. But in Column 4 in Table 7 with the number of news trades, we see a significant decrease in the number of news trades over time in the market-order algo markets relative to the baseline. This relative decrease of about 0.07 news trades per round indicates that human traders realize their stale limit orders are picked off at their expense, and they limit this exposure with experience. This is more direct evidence how traders improve more in the market-order algo than limit-order

[27]Welfare is the sum of payoffs in the market, and payoffs are higher in rounds with a high asset value realization. If we did not control for the asset value realization, then it would be a confound with learning effects. For this reason, we add the realization dummy variable in the time trend regressions in Table 7, even though we did not pre-register including this control variable. This control variable is not needed in the other tables that do not estimate time trends, because all treatments used the same sequence of realizations, hence the realizations are independent of treatment.

algo treatments, by limiting their exposure around news time they learn to avoid part of the exploitation by the market-order algo.

Result 7 (Learning).

- *Welfare increases with experience in all treatments. Welfare increases more with experience in the market-order algo markets compared to the limit-order algo markets and the human-only markets.*
- *The time trend of the number of news trades is significantly smaller in market-order algo markets compared to human-only markets. Both results together indicate that human traders partly learn to avoid some of the exploitation by the market-order algo around news time with experience.*

Column 5 in Table 7 shows a significant positive learning effect regarding price efficiency in the human-only markets, where the share of transactions at efficient prices increases by about 1.4 percentage points per round. This learning trend is not significantly different in the trading algo treatments. In Column 6 for bid-ask spreads, none of the treatments have a time trend that is significantly different from zero.[28]

[28]The limit-order algo time trend is significantly larger than the baseline time trend at the 10% level, but when adding the negative Round coefficient the sum is statistically zero (Wald-test, $\chi^2(1) = 1.58$, $p = 0.209$).

Table 7: Tests for learning during the experiment, random effects regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent var.	Welfare	Small Trader Welfare	Bid-Ask Spread	Trading Volume	Number of News Trades	Price Efficiency	Price Volatility
MarketAlgo	-15.395	-121.515***	2.561	-7.142**	3.289***	0.101*	-70.024*
	(14.219)	(45.660)	(3.270)	(3.349)	(0.461)	(0.060)	(36.305)
MarketAlgo × Round	3.080***	3.412	-0.054	0.471*	-0.071**	0.004	-0.416
	(0.911)	(2.167)	(0.214)	(0.246)	(0.032)	(0.004)	(2.051)
LimitAlgo	51.289***	25.702	-28.299***	0.665	0.246	0.233***	-88.346**
	(16.562)	(39.165)	(3.591)	(3.079)	(0.244)	(0.050)	(35.353)
LimitAlgo × Round	0.053	-2.267	0.421*	0.173	-0.037**	0.001	-1.111
	(1.104)	(1.998)	(0.234)	(0.148)	(0.019)	(0.004)	(1.971)
KnownNewsTime	0.200	8.773	0.638	1.207	-0.177	-0.031*	12.137
	(3.850)	(9.361)	(0.970)	(0.946)	(0.112)	(0.016)	(8.385)
Round	3.315***	4.214***	-0.080	0.030	0.051***	0.014***	-5.089***
	(0.726)	(1.409)	(0.187)	(0.098)	(0.016)	(0.003)	(1.753)
Constant	4856.258***	3354.691***	20.698***	26.300***	0.282	0.396***	199.507***
	(15.020)	(49.349)	(3.443)	(2.935)	(0.285)	(0.044)	(26.427)
Control order	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control π-realization	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	600	600	600	600	600	600	600
Clusters	30	30	30	30	30	30	30

Note: Welfare is the sum of payoffs among all traders in the market. Small Trader Welfare is the sum of payoffs among all traders who are not the large trader. Bid-Ask Spread is the time weighted bid-ask spread in the market, where no bid is set to the minimum asset valuation of 20 and no ask is set to the maximum asset valuation of 90. Number of News Trades is the number of trades within 1 second of news release. Price Efficiency is the share of transaction prices between the current asset valuation of the low trader type and the high trader type, assuming risk neutrality. Price Volatility is the variance of prices around their pre-news and post-news means, respectively, i.e., the price variance controlled for the shift in means due to news release. MarketAlgo, LimitAlgo, and KnownNewsTime are dummies for markets in which there is a market-order algo, a limit order-algo, and a known news release time, respectively. The base category is the baseline with only human traders. Control order is a dummy which equals 1 for groups who started with certain news release time. The unit of observation is a single market. Standard errors are displayed below the coefficients and are clustered at the group level. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

A.3 The effect of news timing on market outcomes

In every treatment, we have markets where the news release time is known in advance, and other markets where only the possible interval of news release times is known ("unknown news release time"). Table 8 estimates whether market outcomes differ with known vs unknown news release time in the baseline (coefficient KnownNewsTime), and whether these differences are different in the algo treatments (the interaction terms). In short, it makes no difference to any of our seven main outcome variables whether traders know the exact time of news release. This may be surprising. In the pre-registration we hypothesized that the unknown news time environment should give the market-order algo more power to exploit human traders via its speed advantage, compared to human-only markets where news timing should not matter due to a level playing field. But since none of the interaction terms show a significant effect, this theoretical advantage is clearly not enough to make a difference in the data. In practice, these results suggest that trading algos do not appear to especially benefit from unscheduled news events such as natural disasters, compared to scheduled news events such as earnings announcements.

Table 8: The effect of certain vs uncertain news timing, random effects regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent var.	Welfare	Small Trader Welfare	Bid-Ask Spread	Trading Volume	Number of News Trades	Price Efficiency	Price Volatility
MarketAlgo	55.150**	-56.160	6.666	-6.010*	2.095***	0.320***	-100.963***
	(22.886)	(53.983)	(4.573)	(3.584)	(0.374)	(0.060)	(34.984)
MarketAlgo × KnownNewsTime	6.900	-43.200	0.319	4.520	-0.370	0.023	-6.274
	(25.618)	(38.236)	(2.468)	(3.030)	(0.370)	(0.084)	(31.326)
LimitAlgo	57.450**	-21.095	-18.458***	0.585	-0.215	0.348***	-95.815***
	(24.219)	(53.999)	(3.892)	(3.185)	(0.254)	(0.054)	(32.797)
LimitAlgo × KnownNewsTime	6.300	-15.290	-3.730	1.950	-0.070	0.044	-18.645
	(17.927)	(25.149)	(2.651)	(1.312)	(0.257)	(0.077)	(34.443)
KnownNewsTime	-4.200	28.270	1.775	-0.950	-0.030	-0.053	20.444
	(12.678)	(21.344)	(2.141)	(0.695)	(0.229)	(0.055)	(26.659)
Constant	6294.300***	4464.545***	18.040***	26.965***	0.945***	0.412***	164.018***
	(20.009)	(51.786)	(3.454)	(3.052)	(0.212)	(0.039)	(29.450)
Control order	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	600	600	600	600	600	600	600
Clusters	30	30	30	30	30	30	30

Note: Welfare is the sum of payoffs among all traders in the market. Small Trader Welfare is the sum of payoffs among all traders who are not the large trader. Bid-Ask Spread is the time weighted bid-ask spread in the market, where no bid is set to the minimum asset valuation of 20 and no ask is set to the maximum asset valuation of 90. Number of News Trades is the number of trades within 1 second of news release. Price Efficiency is the share of transaction prices between the current asset valuation of the low trader type and the high trader type, assuming risk neutrality. Price Volatility is the variance of prices around their pre-news and post-news means, respectively, i.e., the price variance controlled for the shift in means due to news release. MarketAlgo, LimitAlgo, and KnownNewsTime are dummies for markets in which there is a market-order algo, a limit order-algo, and a known news release time, respectively. The base category is the baseline with only human traders. Control order is a dummy which equals 1 for groups who started with certain news release time. The unit of observation is a single market. Standard errors are displayed below the coefficients and are clustered at the group level. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

A.4 Mechanisms underlying the welfare effect of the LimitAlgo treatment

Figure 3 displays estimates from a similar SEM model as in the main part (Figure 2), but with 4 instead of 2 mediators. Information criteria favor the 2-mediator model in the main part of the paper. Also in this model, as in the one in the main part, price efficiency explains most of the positive causal effect of the limit-order algo on welfare.

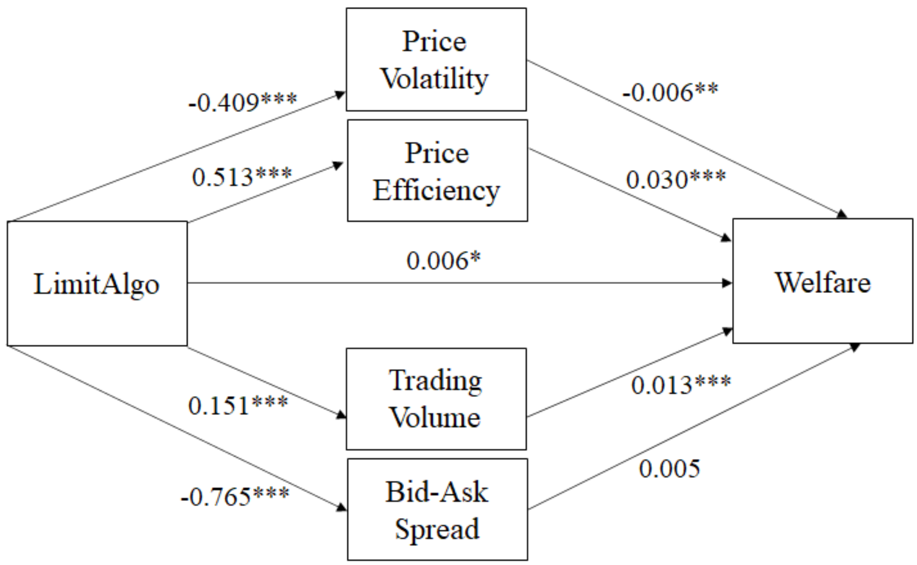


Figure 3: Structural equation model with standardized coefficient estimates for LimitAlgo and the baseline. All regressions control for KnownNewsTime, order and π-realization. Arrows represent direct effects. Standard criteria (see Williams, Vandenberg, and Edwards (2009)) show adequate fit for valid measurements given our moderate sample size and the use of ML estimation: Likelihood Ratio Test, $\chi^2(1) = 116.28, p < 0.001$, SRMR (Standardized Root Mean Residual) = 0.055 [< 0.100], CFI (Comparative Fit Index) = 0.967 [>0.950] (see Shah and Goldstein (2006)). AIC= 17018.731 and BIC=17154.441. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level. The estimated equations are as follows, where i is a market in the LimitAlgo treatment or baseline:

$$\begin{aligned}
\text{Price Efficiency}_i &= \alpha_1 + \beta_1 \text{LimitAlgo}_i + \delta_1 \text{KnownNewsTime}_i + e_{1,i} \\
\text{Trading Volume}_i &= \alpha_2 + \beta_2 \text{LimitAlgo}_i + \delta_2 \text{KnownNewsTime}_i + e_{2,i} \\
\text{Price Volatility}_i &= \alpha_3 + \beta_3 \text{LimitAlgo}_i + \delta_3 \text{KnownNewsTime}_i + e_{3,i} \\
\text{Bid-Ask Spread}_i &= \alpha_4 + \beta_4 \text{LimitAlgo}_i + \delta_4 \text{KnownNewsTime}_i + e_{4,i} \\
\text{Welfare}_i &= \alpha_5 + \beta_{5,1} \text{LimitAlgo}_i + \beta_{5,2} \text{Price Efficiency}_i + \beta_{5,3} \text{Trading Volume}_i \\
&\quad + \beta_{5,4} \text{Bid-Ask Spread} + \beta_{5,5} \text{Price Volatility} + \delta_5 \text{KnownNewsTime}_i + e_{5,i}
\end{aligned}$$

Figure 4 displays estimates from a SEM model for Small Trader Welfare, where our goal is to disentangle the channels of the negative market-order algo effect. Unlike the analysis of Welfare, we separate trading volume into news trades and no-news trades (all trades that are not categorized as news trades). We do so to highlight the unique impact of the market-order algo on news trades as a potential explanatory channel for the decrease in Small Trader Welfare. Moreover, this distinction illustrates that more no-news trades have positive effects on welfare, whereas more news trades have a negative effect, as they exploit the small human traders. We do not include a direct effect of MarketAlgo on Small Trader Welfare, because of potential collinearity issues, given that the correlation between MarketAlgo and News Trades is $\rho = 0.56$.

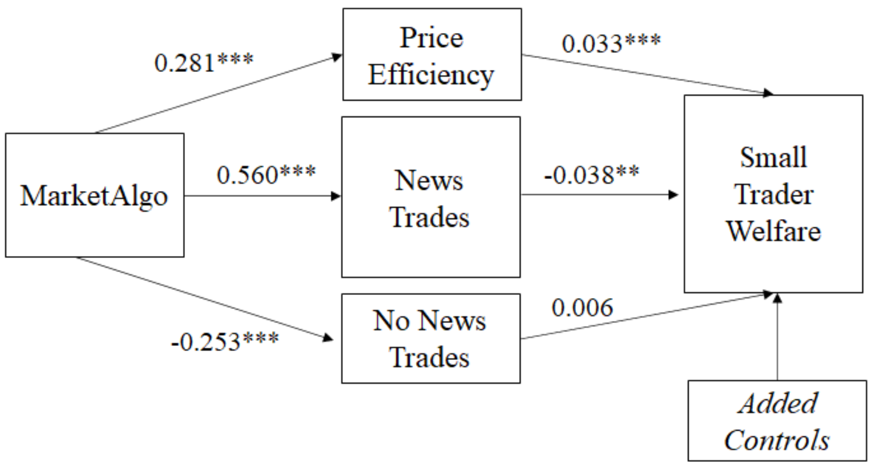


Figure 4: Structural equation model with standardized coefficient estimates for MarketAlgo and the baseline. All regressions control for KnownNewsTime, order and π-realization. Arrows represent direct effects. Standard criteria (see Williams, Vandenberg, and Edwards (2009)) show reasonable fit for valid measurements given our moderate sample size and the use of ML estimation: Likelihood Ratio Test, $\chi^2(1) = 130.98, p < 0.001$, SRMR (Standardized Root Mean Residual) = 0.058, CFI (Comparative Fit Index) = 0.929. AIC= 19918.810 and BIC=20030.571. ***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level. The estimated equations are as follows, where i is a market in the MarketAlgo treatment or baseline:

$$\begin{aligned}
\text{Price Efficiency}_i &= \alpha_1 + \beta_1 \text{MarketAlgo}_i + \delta_1 \text{KnownNewsTime}_i + e_{1,i} \\
\text{News Trades}_i &= \alpha_2 + \beta_2 \text{MarketAlgo}_i + \delta_2 \text{KnownNewsTime}_i + e_{2,i} \\
\text{No News Trades}_i &= \alpha_3 + \beta_2 \text{MarketAlgo}_i + \delta_3 \text{KnownNewsTime}_i + e_{3,i} \\
\text{Small Trader Welfare}_i &= \alpha_4 + \beta_{4,1} \text{Price Efficiency}_i + \beta_{4,2} \text{News Trades}_i \\
&\quad + \beta_{4,3} \text{No News Trades}_i + \beta_{4,4} \text{Bid-Ask Spread} \\
&\quad + \beta_{4,5} \text{Price Volatility} + \delta_4 \text{KnownNewsTime}_i + e_{4,i}
\end{aligned}$$

The SEM shows that while the market-order algo increases price efficiency, and this increases Small Trader Welfare, the other two channels more than offset it, explaining the negative treatment effect. Among the negative channels, News Trades explains $0.56 \times 0.038/(0.56 \times$

$0.038 + 0.253 \times 0.006) \approx 0.93$, i.e., 93%, of the negative effect. The reduction in no news trades explains the remaining 7%.

B Appendix: Post-experimental questionnaire

1. **Negative attitudes toward robots scale**

   Please indicate the extent to which you agree or disagree with each of the following statements. [5-point Likert scale from strongly disagree to strongly agree]

   (a) I would feel uneasy if robots really had emotions.

   (b) Something bad might happen if robots developed into living beings.

   (c) I would feel relaxed talking with robots.

   (d) I would feel uneasy if I was given a job where I had to use robots.

   (e) If robots had emotions, I would be able to make friends with them.

   (f) I feel comforted being with robots that have emotions.

   (g) The word "robot" means nothing to me.

   (h) I would feel nervous operating a robot in front of other people.

   (i) I would hate the idea that robots or artificial intelligences were making judgments about things.

   (j) I would feel very nervous just standing in front of a robot.

   (k) I feel that if I depend on robots too much, something bad might happen.

   (l) I would feel paranoid talking with a robot.

   (m) I am concerned that robots would be a bad influence on children.

   (n) I feel that in the future society will be dominated by robots.

2. Please answer the following socio-demographic questions.

   (a) Gender [male, female, other]

   (b) Age (in years)

   (c) In general, how willing or unwilling are you to take risks? Please use a scale from 0 to 10, where 0 means you are completely unwilling to take risks and 10 means you are very willing to take risks. You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

3. [**Algo-treatments only**] Please indicate the extent to which you agree or disagree with each of the following statements. [5-point Likert scale from strongly disagree to strongly agree]

   (a) The algorithmic trader was very active in the market.

   (b) The algorithmic trader helped human traders make more transactions.

   (c) The algorithmic trader hurt the profits of human traders.

   (d) The algorithmic trader helped human traders to trade close to the value of the asset.

4. **Cognitive Reflection Test (CRT)**

   (a) If Mario can eat one entire pizza in 5 minutes, and John can eat one pizza in 20 minutes, how long would it take them to eat an entire pizza together? [number]

   (b) Mary received both the 18th highest and the 18th lowest mark in the class. How many students are in the class? [number]

   (c) An investor buys a company for $80M, sells it for $90M, buys it back for $95M, and sells it finally for $110M. How much has he earned in $M? [number]

   (d) Warren bought $10,000 of stocks in 2018. Six months after he invested, on August 11, the stocks he had purchased were down 40%. Fortunately for Warren, from August 11 to November 11, the stocks he had purchased went up 60%. At this point, Warren has: [Neither earned nor lost money since 2018; Earned money since 2018, Lost money since 2018]

C Appendix: Instructions for algo treatments with known news release time first

- The following pages display the instructions for the algo treatments, in the order where the news release time was known first.
- The instructions are in the form of slides, which are displayed here as two slides per page.
- The instructions in the baseline are identical, except there is no mentioning of the computer playing (available on request, not included here to preserve space).
- The instructions for the treatments with “unknown news release time first” are identical, except the reversed order is listed and explained.

Welcome

The experiment will take place through computer terminals at which you are seated. If you have any questions during the instruction phase, raise your hand and a person will come by to answer your question. If any difficulties arise after the experiment has begun, raise your hand, and someone will assist you.

This is an experiment in market decision making. You will be paid in cash for your participation at the end of the experiment. Different participants may earn different amounts. What you earn depends on your decisions and the decisions of others.
The earnings for this experiment are calculated in ECUs, which is an experimental currency that will be converted into cash for payments at the end of the experiment.
27 ECUs are worth **$1**.

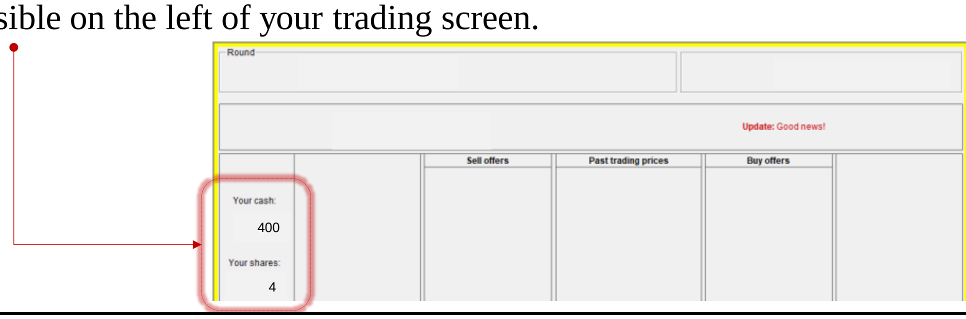
Endowment and share values

In this experiment you will be able to buy and sell **shares** of an asset from one another.
All 8 traders will be given cash and shares. At the beginning of the experiment, you will be randomly assigned one of the two possible endowments of cash and shares:

- 5 traders will receive 360 ECUs in cash and 4 shares
- 2 traders will receive 400 ECUs in cash and 4 shares

The 8^{th} trader will be the computer and will receive 1200 ECUs in cash and 12 shares.

You will start each round with this same endowment. In the example below, you are given an endowment of 400 in cash and 4 shares. Your cash and the number of shares you hold will always be visible on the left of your trading screen.



Endowment and share values

The **shares** last for exactly **1** round of trading, and trading in each round lasts 100 seconds. There will be a total of 20 rounds. At the beginning of each new round, your cash and shares will always be reset at their initial values. Note that you will receive the **same** amount of cash and shares at the beginning of each round.

At the end of the round, each share you hold will give you a payout. You will not know the exact value of the payout at the beginning of the round. But, you know that the payout can be LOW or HIGH with equal chances.

The LOW and HIGH values will differ across traders. For the 5 traders with the 360 ECU endowment the LOW and HIGH values will be 30 and 90. For the remaining traders, the LOW and HIGH values will be 20 and 80. Your LOW and HIGH values will not change from one round to the next. So, if your LOW and HIGH values are 30 and 90 in round 1, then they will be 30 and 90 in **all** rounds.

Endowment and share values

- ✓ If you were assigned the values 30 and 90 and you hold a share at the end of the round, you will earn either 30 ECUs per share if the asset payout turns out to be LOW or 90 ECUs if it turns out to be HIGH. Thus, the expected value of the asset, which is the average of the two possible values, is equal to 60. In this case, your cash endowment will be 360 ECUs.

- ✓ If you were assigned the values 20 and 80 and you hold a share at the end of the round, you will earn either 20 ECUs per share if the asset payout turns out to be LOW or 80 ECUs if it turns out to be HIGH. Thus, the expected value of the asset is equal to 50. In this case, your cash endowment will be 400 ECUs.

News timing

Although the exact value of the payout will not be known at the beginning of the round, it will be announced on every trader's screen during trading.

In the first 10 rounds, the time of the news will be announced in the beginning of the round. It will also be displayed at the top of your trading screen. In the example below, the news will be announced when there are exactly 50 seconds remaining in the round.

The information about the news release time will also be shown to you in a separate screen that you will face before a trading round starts. This screen will also inform you about your endowment and about your valuation of the asset (20, 80) or (30, 90).

News timing

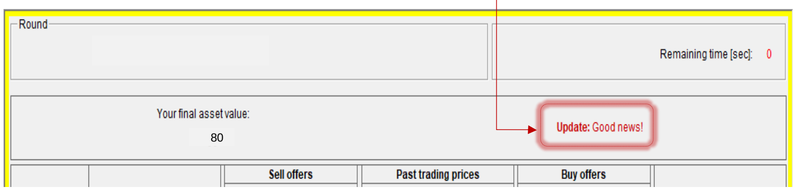
The possible remaining times at which the news will be released will vary across rounds. These values can be: 40, 45, 50, 55 or 60.

In the last 10 rounds, the time at which the news will be released will not be exactly known at the beginning of the round. However, the possible times for the news release will be the same as in the first 10 rounds. That is, the news will be released when the time remaining in a round is 40, 45, 50, 55 or 60. In a given round, news will have an equal chance to occur at each of these times.

Before the news, the information about the timing of the news will be displayed on your screen as follows.

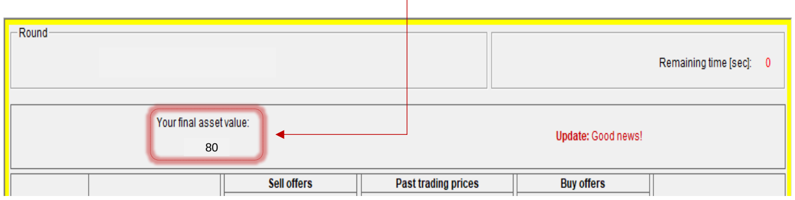
News announcement

If the share earns a HIGH payout, then 'Good news!' will appear at the top right corner of your screen.



News announcement

At the same time, the top left corner of your screen will show the updated asset value.

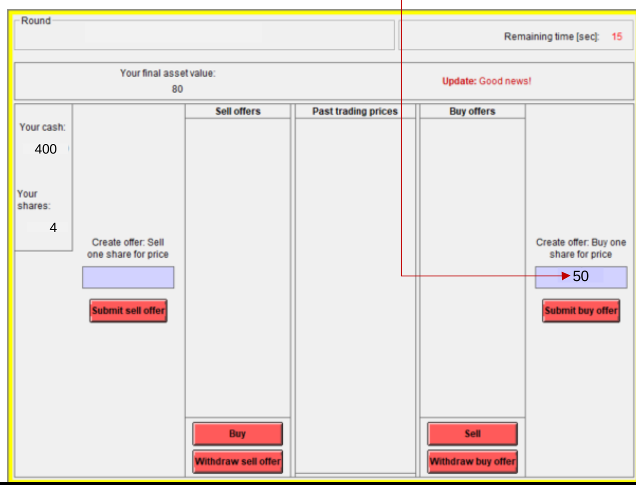


If the share earns a LOW payout, then 'Bad news!' will appear at the top right corner of your screen. At the same time, the top left corner of your screen will show the updated asset value. You can trade before and after the news was released.

For each of the 20 rounds, the payout value (LOW or HIGH) has been randomly drawn by a computer. That is, the value of the payout can change across rounds.

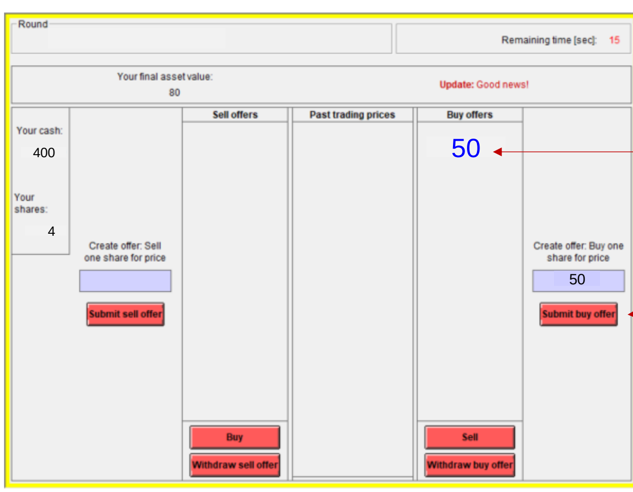
Trading (buy offers)

During a round, traders can buy or sell shares from one another by making Buy offers or Sell offers. To create a new Buy offer for a share, type in the price at which you would like to buy on the right side of the trading screen. Note that you will not be able to create a Buy offer with a higher price than your available cash, because you would not have enough money to buy the share. In the example below, a new Buy offer of 50 ECUs has been entered.



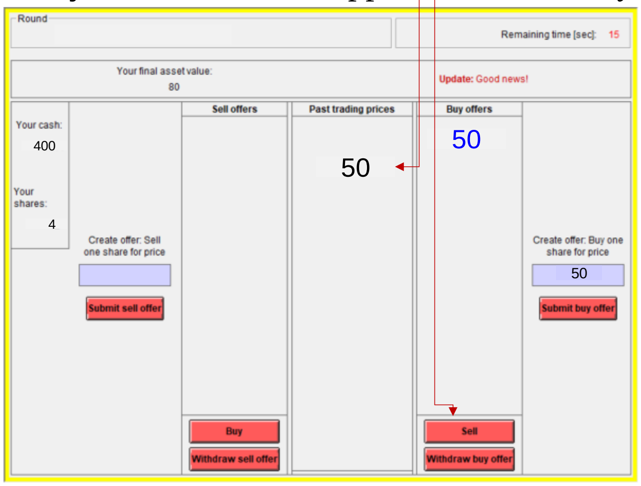
Trading (buy offers)

To confirm the new Buy offer of 50 ECUs, you have to click on ‘Submit buy offer’ right below your offer. The new Buy offer will then appear in the ‘Buy offers’ column.



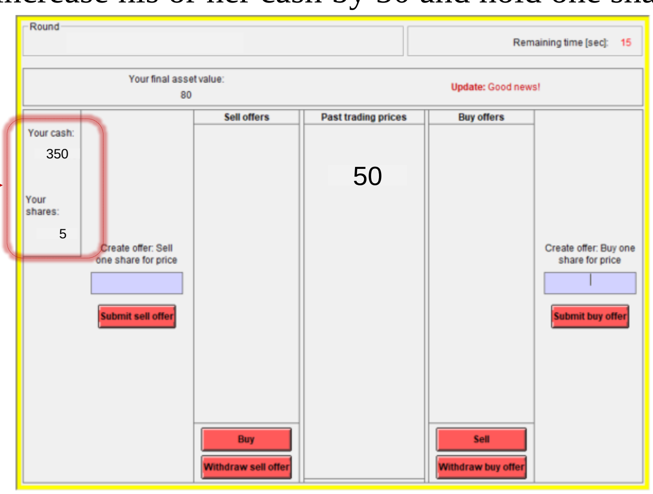
Trading (buy offers)

Another trader can now accept the new Buy offer of 50 that you have just made by selecting your offer and clicking on the ‘Sell’ button at the bottom of the screen. The transaction price will then appear in the middle column ‘Past trading prices’. The Buy offer is not available anymore and will disappear from the ‘Buy offers’ column.



Trading (buy offers)

The Buy offer of 50 has been accepted meaning you have bought one share at 50 from another trader. Your available cash will thus go down by 50 (from 400 to 350) and the number of shares you hold will increase by one (from 4 to 5). The trader who sold the share to you will increase his or her cash by 50 and hold one share less.

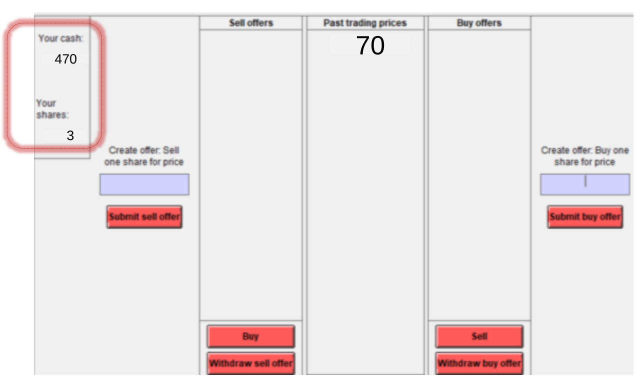


Trading (sell offers)

Creating a new Sell offer for a share works the same way as creating a new Buy offer, except that you have to place your offers on the left hand side of the screen. Note that you will not be able to create a Sell offer if you do not hold any shares.
In the example below, you created a Sell offer for 70 that you confirmed by clicking on the 'Submit sell offer' button. This offer was accepted by a buyer who clicked on the 'Buy' button.

Trading (sell offers)

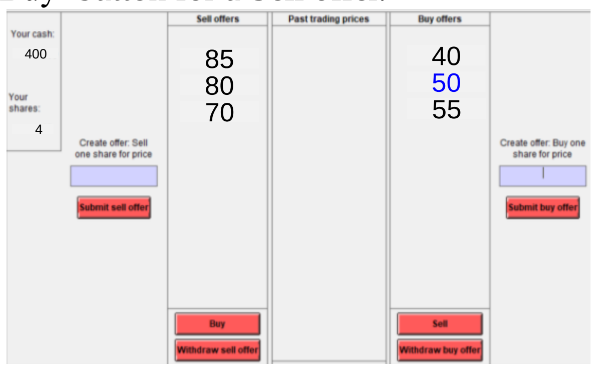
This transaction increases your cash by 70 and decreases the number of shares you hold by one. The buyer holds one more share but has 70 ECUs less in cash.



Trading (order book)

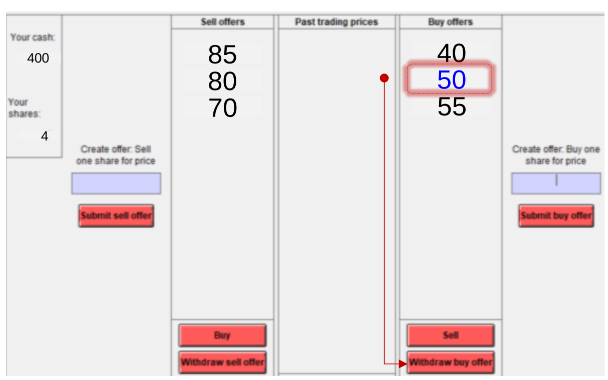
Several Buy offers or Sell offers can be available at the same time. In the example below, there are three Buy offers (85, 80, 70) and three Sell offers (40, 50, 55) available. Note that the lowest (best) Sell offer and the highest (best) Buy offer appear at the bottom. The Buy offer of 50 has been created by yourself so it appears in blue.

You can click on any of the Buy and Sell offers in the list to sell or buy a share. Once you pick the offer it will be highlighted and you will just have to confirm by clicking the 'Sell' button for a Buy offer or the 'Buy' button for a Sell offer.



Trading (order book)

You can cancel any of your offers (those that appear in blue) by clicking on the offer and then on the 'Withdraw sell offer' button if it is a Sell offer or on the 'Withdraw buy offer' button if it is a Buy offer. However, your offer can no longer be withdrawn once it has been accepted by someone else.



Trading (multiple offers)

You will make offers to buy or sell for one share at a time. You will need to make multiple offers or accept multiple offers if you want to trade multiple shares.

Earnings

Your round earnings equal the amount of cash you have at the end of a round plus the sum of the payouts you collect from your shares.

For example, imagine that you have 462 in cash and you hold 3 shares at the end of the round. Suppose the payout for each share is 80. This means you collect 3×80 = 240 from all your shares.

Your earnings in this round will be: Your cash earnings (462) + Your shares earnings (240) = 702 ECUs. Your round earnings are always displayed at the end of the round.

Remember that **27 ECUs = $1**

Out of the 20 rounds, 1 round will be picked at random for payment. Your final earnings are the earnings from the randomly selected round.

Summary

1. You will be given an initial amount of cash and shares.
2. To trade shares, you can submit Buy and Sell offers and accept available offers.
3. Every share generates a payout at the end of the round. Trading lasts for 100 seconds.
4. The value of the payout is not known at the beginning of the round. But, you know that the payout can be LOW or HIGH with equal chances. The LOW and HIGH values can be either (20 and 80) or (30 and 90).
5. News revealing the final payout value will be released to all traders at the same time, and this time will be known at the beginning of the first 10 rounds and unknown in the remaining 10 rounds.
6. There will be 20 rounds. At the beginning of each round, your cash amount and your shares are reset to their initial values.
7. One of the 8 traders will be played by the computer and will receive 1200 ECUs in cash and 12 shares.
8. Your actual experiment earnings will depend on your earnings in 1 randomly selected round. Earnings in a round is the sum of the cash you have available at the end of the round and the sum of payouts you obtain from your shares at the end of the round.

Practice

✓ Before starting with the 20 rounds, you will play 1 practice round that will not be selected for payment.