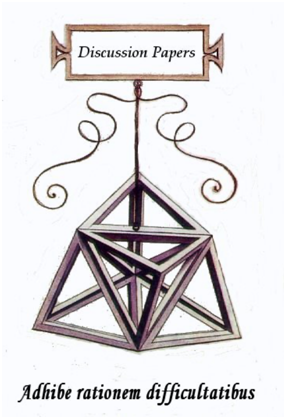
***Discussion Papers***
Collana di
E-papers del Dipartimento di Economia e Management – Università di Pisa




Alexia GAUDEUL and Caterina GIANNETTI

# Trade-offs in the design of financial algorithms

*Discussion Paper* n. 288

2023

*Discussion Paper* n. 288, presentato: **Marzo 2023**

**Indirizzo degli Autori**:

Caterina Giannetti- Dipartimento di Economia e Management, Università di Pisa, via Ridolfi 10, 56100 PISA – Italy. *Email: caterina.giannetti@unipi.it*



Si prega di citare così:
Alexia GAUDEUL and Caterina GIANNETTI (2023), "Trade-offs in the design of financial algorithms", Discussion Papers del Dipartimento di Economia e Management – Università di Pisa, n. 288 (http://www.ec.unipi.it/ricerca/discussion-papers.html).

***Discussion Paper***
n. 288

Alexia GAUDEUL and Caterina GIANNETTI

# Trade-offs in the design of financial algorithms

**Abstract**

We investigate trade-offs when trying to encourage adoption of stock-trading algorithms. We organize an artificial stock market experiment over three weeks where investors experience trading on their own and with the help of a financial algorithm.They then choose whether to adopt it. We vary the algorithm in terms of its trading strategy and whether its decisions can be overriden or not. We find that adoption rates are low, but investors are more likely to adopt an algorithm that trades actively and that they can override. The investor's trading preferences, as revealed by their own trading decisions, does not consistently affect algorithm take-up. Rather, algorithm adoption depends mainly on how succesful a trader was when trading on their own vs. when an algorithm was trading in their place. Analysis of an exit questionnaire matches those observations with the reasons given by individuals for rejecting or adopting a financial algorithm.

**JEL codes:** G11, G40

**Keywords:** algorithm aversion, disposition effect, robo-advisers, sophisticated investors, stock-trading

# Trade-offs in the design of financial algorithms*

Alexia GAUDEUL†and Caterina GIANNETTI‡

March 27, 2023

**Abstract**

We investigate trade-offs when trying to encourage adoption of stock-trading algorithms. We organize an artificial stock market experiment over three weeks where investors experience trading on their own and with the help of a financial algorithm.They then choose whether to adopt it. We vary the algorithm in terms of its trading strategy and whether its decisions can be overriden or not. We find that adoption rates are low, but investors are more likely to adopt an algorithm that trades actively and that they can override. The investor's trading preferences, as revealed by their own trading decisions, does not consistently affect algorithm take-up. Rather, algorithm adoption depends mainly on how successful a trader was when trading on their own *vs.* when an algorithm was trading in their place. Analysis of an exit questionnaire matches those observations with the reasons given by individuals for rejecting or adopting a financial algorithm.

**Keywords:** algorithm aversion, disposition effect, robo-advisers, sophisticated investors, stock-trading

**JEL Classification**: G11, G40

*The authors thank the *Thinking Forward Initiative* for financial support. The authors also thank Paolo Crosetto, as well as participants at the Society for Experimental Finance 2021, the ESA Meetings 2021, and departmental seminars at the universities of Pisa, Siena, Montpellier, and Göttingen.

†Competence Centre on Behavioural Insights, Joint Research Centre of the European Commission, Brussels, Belgium. Email: alexia.gaudeul@ec.europa.eu

‡Department of Economics and Management, University of Pisa, Italy. Email: caterina.giannetti@unipi.it

# 1 Introduction

We present a stock-trading experiment where traders are offered advice from stock-trading algorithms. We vary the design of those algorithms and consider how those variations affect the trading performance of their advisees, and the likelihood that they choose to rely on them in their trades. Our experiment is designed to replicate settings where real investors are offered help from robo-advisers, which provide personalized advice for portfolio management.

Robo-advisers have gone through several stages of development, whereby the most sophisticated now can manage individual portfolios algorithmically depending on the preferences of the investor (Deloitte, 2016). Robo-advisers can now not only help investors maintain portfolios that are better balanced and diversified (Foerster et al. 2017, Uhl and Rohner 2018), but also address their behavioral biases (Bhatia et al., 2020), and correct for the impact of "irrational" factors in their trading decisions (D'Acunto et al., 2019). Their low cost and permanent availability make them particularly helpful for households that have relatively low incomes and capital, as these generally cannot be profitably advised by human advisers (D'Acunto and Rossi, 2020). They can be designed to unambiguously serve the interests of the investor rather than those of the adviser (Brenner and Meyll, 2020) and may therefore be less vulnerable to moral hazard problems than human advisers, and thus face less distrust.

While robo-advisers have many promises, we show the need for some hard choices in terms of design, whereby algorithm performance may need to be sacrificed for the sake of promoting their use. Our experiment is designed to investigate those trade-offs which are about obtaining a robo-adviser that challenges the preferences, habits, and intuitions of traders, but not so much that those traders then choose to ignore its advices. Such preferences, habits and intuitions can only be developed if the trader gains experience on a market over an extended period of time. This is why we innovate on the usual one-hour experiment in a laboratory by letting participants in our experiment trade actively – three times per day — over three weeks on an artificial stock market which they could access from anywhere with an internet connection. They traded one week on their own, another week with an algorithm, and chose whether to trade with the algorithm in the third week.

We systematically varied the type of algorithm we offered participants in order to explore the balance between an algorithm's performance and its attractiveness. The optimal design balances the likelihood the algorithm will be adopted and its performance, as even the best algorithm is useless if it is not adopted.

A first trade-off is between giving full control to an algorithm or having a "human in the loop", whereby users can intervene on a case-by-case basis. The first option is optimal for most people given that algorithms trade better than them on average. The later can however increase rates of adoption, at the cost of letting individuals sometime override algorithmic decisions to their own detriment. This is the performance-control dilemma (Rühr, 2020). We explore this first trade off by varying the possibility to override algorithm's choices.

A second trade-off is between matching the preferences of individuals vs. offering the most efficient trading algorithm. The issue here is that those who follow the worst trading strategies also experience the most discrepancy between their own trades and those made by a good robo-adviser. This is the performance-preference dilemma (Bailey, 1993). We explore this second trade-off by varying the algorithm's trading strategy across treatments and relating the decision to adopt an algorithm with the difference between an individual's own trades and the decisions of the algorithm they had to use.

While investigating adoption of algorithms, we have to take account of a final issue, which is the effect of "mistakes" by the algorithm (Dietvorst et al., 2015). Adoption will depend not only of how an algorithm performs in theory, but of how it performs in practice. This is an important issue whenever one operates in a noisy environment, such as on financial markets, where even the best advice can give bad results. There are indeed many unpredictable factors in stock market movements, so that making the difference between a good decision and a good result is difficult. This issue is seldom covered in the literature on algorithm adoption up to now as it generally focuses on cases where algorithm are relatively precise and one can identify mistakes from the results (see e.g. Asparouhova et al. 2020). This is not the case when trading, and most investors will find it difficult to make the difference between a bad decision and a bad result (König-Kersting et al. 2021). Our experimental setting allows us to make the difference between the two. We are thus able to examine whether users are more sensitive to whether the algorithm trades well in theory ("ex-ante"), or to whether it traded well in practice ("ex-post"). In other words, we can examine if adoption is driven by rational, reflective, strategic considerations, or by such emotions due to the experience of regret (when losing) or joy (when winning).

The experiment we present in this paper allows us to explore those three important factors in algorithm adoption at the same time by monitoring the actual performance of the algorithms and of the individuals. Conclusions from this study are meant to foster adoption of financial algorithms, especially by those traders who need them most. We contribute to a better understanding of what to pay attention to when designing and presenting financial algorithms.

We designed our experiment along Frydman et al. (2014); Frydman and Rangel (2014) whereby stock prices fluctuate randomly over time, but can be in two states, either good, i.e., the trend is generally positive, or bad, i.e. the trend is negative. This type of market is designed to focus on a specific and sub-optimal bias in individual trading decisions, the disposition effect,in which traders tend to sell stocks that go up and keep those that go down. This trading bias is well documented (see Pleßner 2017) and previous research has investigated how to help investors cope with it. For example, Frydman and Rangel (2014) and Frydman and Wang (2020) showed that it is possible to reduce the disposition effect by decreasing the salience of the purchasing price, while Fischbacher et al. (2017) showed that automatic selling devices (i.e. stop-loss and take-gain orders) reduce it. Chang et al. (2016) also showed that the disposition effect could be reversed if the investor could blame someone else for bad trading decisions.

Unlike those previous experiments, we do not manipulate how information is provided to investors (Frydman and Rangel 2014) or whether they have to commit to trade in a certain way (Fischbacher et al., 2017). Rather, we vary across treatment the algorithm's trading strategy and its decision autonomy and look at the decision by investors to rely on the algorithm after experiencing trading with and without its advice (Tse et al. 2022; Filiz et al. 2021). In particular, we vary two levers to ease adoption: whether the algorithm actively trades according to a Bayesian strategy or simply prevents the investor from trading (*Bayes* algorithm *vs. Block* algorithm), and whether individuals can override the decisions of the robo-advisor or not (*Soft* vs *Hard*). The first treatment variation aims at investigating the performance-preference dilemma, the second treatment variation aims at the performance-control dilemma. The *Block* trading algorithm does not sell or buy any stock, while the *Bayes* algorithm buys or sells a stock depending on how likely it is to go up in value in the next period.[1] A *Soft* algorithm is such that participants are free to override the algorithm's choice (i.e the human is in the loop), while a *Hard* algorithm is such that they cannot do so.

Another main difference with previous experiment is that we let participants trade over a long period, three weeks, and using their own devices. This makes our experiment closer to real settings where investors must combine their investment activity with their daily life, whereby they form preferences and habits. Indeed, repeated feedback and gradual development of subjects' experience are necessary conditions to improve learning and confidence (Chacon et al. 2022; Filiz et al. 2022).[2] More specifically, our experiment lasted three weeks, whereby participants could make trading decisions on three independent stocks three times per day. Participants experienced trade on their own for one week, then trade with the help of an algorithm for another week. They then decided whether to trade with the help of that algorithm in the third week.

The results from our experiment show that participants achieved better performance (i.e. they exhibited lower level of the disposition effect) when trading along with an algorithm, even if that algorithm was passive and therefore less efficient, and even if they could override it. However, the majority decided not to adopt an algorithm in the third week. Participants preferred being able to override algorithm choices, and they preferred the optimal, active algorithm to the passive, less efficient one. The sequence in which one experienced trading with the algorithm or on one's own mattered. Participants had short memory, whereby differences between robo-advisers mattered only if they experienced robo-advisers recently, and differences between traders' own trading styles mattered only

[1]Note that of course participants do not observe the probability of the stock being in a good state. See experimental instructions in appendix B.

[2]Getting decisions in this way is important as financial choices are affected both by an instinctive-affective mechanism (System 1), which may drive short-term decisions, such as what stock to buy or sell, and a deliberative-cognitive mechanism (System 2), which may drive longer-term decisions, such as the way in which one manages one's money (Kahneman 2002, Hirshleifer 2015, Ploner 2017).

if they traded on their own recently. In that case, individuals who were less affected by the disposition effect were more likely to get help from an algorithm. This means that those who needed help the least were also the most likely to get help. Those effects were small compared to the effect of success when trading on one's own, which reduced adoption. These results are robust to a number of additional drivers in adoption, such as whether an investor actively traded stocks when on their own, and their level of risk aversion. An exit survey confirms that adopters believed the algorithm improved their performance, while non-adopters either incorrectly thought algorithms did worse than themselves, or had issues in relinquishing control on their decisions.

## 2 Related literature

As mentioned, we first examine the trade-off between maintaining control on one's trades and letting an algorithm make decisions for oneself (Rühr, 2020). The literature shows that people are generally adverse to delegating decisions, especially so when this is to automated agents. This issue is part of the general problem known as algorithm aversion, and has many explanations beyond the desire of maintaining control over one's decisions (Dietvorst et al., 2015; Prahl and Van Swol, 2017; Dietvorst et al., 2018; Niszczota and Kaszás, 2020; Filiz et al., 2022; Germann and Merkle, 2022). It has been mainly investigated by contrasting human and robo-advice, and the literature suggests that algorithm adoption depends on their design (Burton et al. 2020; Mahmud et al. 2022), their accuracy (e.g. Dietvorst et al. 2018, Hidalgo et al. 2021), the type of task taken up (Morewedge 2022; Castelo et al. 2019), as well as on the characteristics of customers themselves (e.g. Oehler et al. 2022), such as their level of experience, either with the algorithm (e.g. Tse et al., 2022, Filiz et al. 2022) or with the decision domain (D'Acunto et al. 2019). Indeed, D'Acunto et al. (2019) find that those who adopt robo-advisers — a type of algorithm to help financial decisions — are very similar to non-adopters in terms of demographics, but are less prone to behavioral biases and trade more actively. Algorithm aversion sometimes turns into algorithm appreciations, whereby individuals also sometimes prefer automated to human advice (Logg et al., 2019; Chugunova and Sele, 2020; Major and Shah, 2020; Tse et al., 2022; Holzmeister et al., 2022). On the whole, people are more willing to delegate decision to an algorithm when decisions appear more objective (Chugunova and Sele 2020; Burton et al. 2020; Morewedge 2022). Involving individuals and integrating their opinions into the decision process improves take-up (Köbis and Mossink 2021; Kawaguchi 2021). In any case, trust seems to play a crucial role, either in the algorithm or in one's own ability (Chugunova and Sele 2020; Lourenço et al. 2020; Holzmeister et al. 2022; Sharan and Romano 2020). Overall, both algorithm aversion and appreciation do not always benefit individuals, as they are often unable to identify the best algorithm (Bhattacharya et al. 2012, Chang et al. 2016,Tse et al. 2022).

A second trade-off we examine in this paper is between offering the most efficient algorithm and fitting the individual preferences of investors (Bailey, 1993). Indeed, what is "best" may be dependent on the preferences of the individuals, so there are benefits from personalizing, or customizing, the algorithm (Faloon and Scherer, 2017; Capponi et al., 2022). We do not let participants in our experiment design or choose their own algorithm, and we do not ask for their preferences either. Rather, we observe preferences of our participants by letting them trade one week without an algorithm. We can thus compare the trader's own trading decisions with those made by the algorithm. We can then relate their choice of adopting the algorithm in the third week to how the individual traded on their own. This means we can study how algorithm adoption depends not only on whether an algorithm performs better or worse than the individual, but also on whether it is "behaviorally" consistent with their preferences, such as whether its decision when to buy or sell corresponds to what they would have done by themselves, and whether they both traded as actively.

Too much personalization in our setting would lead to us to program an algorithm that would reflect a well-known and detrimental behavioral bias, known as the disposition effect (Shefrin and Statman, 1985), whereby one sells stock that go up and keeps those that go down. One of our proposed algorithm, the Bayesian algorithm, goes squarely against this bias. The issue is that those who are most subject to the disposition effect may also be the least likely to realize the benefits of the optimal, Bayesian algorithm, as it does the opposite of what they would do themselves. This Bayesian algorithm, which is optimal in our setting, may therefore paradoxically be the least likely to be adopted by the most unskilled, even though they need it most. This is a paradox akin to the one identified in Kruger and Dunning (1999). Hence, we also propose a sub-optimal ("Block") algorithm, that simply prevents trades, as a middle-ground. Such a sub-optimal algorithm can ease adoption by less skilled traders, because this algorithm may make decisions that are less inconsistent with their preferences. This trade-off may be worth making because even a sub-optimal algorithm can improve those individuals' performance, as their performance when trading on their own is yet lower than that of the sub-optimal algorithm. On the same line, we also consider whether *Soft* algorithms, which can be overriden, may also ease adoption among those whose strategy differs most from optimal trading.

As mentioned in the introduction, a third issue is whether traders are able to actually evaluate the performance of an algorithm. The literature shows that users react in an exaggerated way to mistakes by the algorithm (Dietvorst et al., 2015). Egocentric advice discounting means that they judge mistakes by the algorithm as worse than their own (Yaniv and Kleinberger, 2000). Combined with a tendency to ignore their own mistakes, this leads them to become overconfident in their own abilities compared to the algorithm. This is particularly concerning in uncertain decision domains such as medical or investement decision making (Dietvorst and Bharti, 2020). Indeed, uncertainty makes it is difficult to distinguish a mistake from bad luck, and thus to make the difference between a good and a successful decision. What is rational in terms of expectations *ex-ante* may differ from what turned out to be best *ex-post*.

From a rational point of view, we would expect that traders would be more likely to adopt the robo-advisers with the best strategies. This would lead to lower adoption of two variations in the algorithms in our experiment which are sub-optimal: the one where the algorithm simply does not trade (*Block*), and the one where the algorithm can be overridden (*Soft*). However, even the best performing, *Bayes* algorithm performs rather badly ex-post (there is only a 55% chance that the recommended trading decision was "correct"). This is a low success rate, and not very different from what one can achieve by simply doing nothing (50% "correct"), which is our *Block* algorithm, or even by doing the opposite of what is optimal (45% correct). Our controlled experimental setting allows us to define both ex-ante and ex-post optimal trade, and relate those to algorithm adoption. We therefore can precisely look at how much ex-post experienced success impacts adoption compared with ex-ante optimality in trading strategy. We also can measure how much a "personal" success compares with an algorithmic success.

# 3 The experiment

We test variations in the design of robo-advisers in a highly controlled experimental environment where our participants could trade on an online artificial stock market, with three independent stocks, three times per day, over three weeks. Our experiment, programmed with *oTree* (Chen et al. 2016), was run online whereby each participants had a personalized web link to access their portfolio and trade. They received reminders to trade via *GMass* every 8 hours at the start of each new market sessions. Participants could take part on their own smartphone or computer while going on with their usual activity. We let participants trade on their own during one week, and with the help of an algorithm in the other week. They then chose whether to use the algorithm in the third week of trading.

**The artificial stock market** Participants in our experiment traded on an online artificial stock market, with three fictitious stocks (A, B and C) independently changing prices three times (i.e. rounds) per day, over three weeks. In the standard treatment, participants traded on their own in the first week, while a robo-adviser made trading decisions for them every two periods in the second week. They then chose whether to trade as in week 1 or in week 2 for the third week. In the reversed treatment, the order of the two first weeks was reversed. Table 1 outlines the chronology of the experiment.

Participants were given 350 ECU at the beginning of each week, 300 ECU originally invested in each of the three stocks at price 100 ECU, and 50 ECU in cash. Participants could hold a maximum of one share of each stock and a minimum of zero (i.e. short-selling was not allowed). The trading decision was therefore reduced to deciding whether to sell a stock (conditional on holding it) or buying a stock (conditional on not holding it). Every eight hours (round), each one of the three stocks had its price randomly updated. As in

Frydman et al. (2014); Frydman and Rangel (2014), the price path of each stock was independently governed by a two-state Markov chain, with a good state and a bad state. The beginning state was drawn randomly and independently for each stock at the beginning of the week, with an equal chances for both. If a stock started in a good state, its price increased with probability 70% and decreased with probability 30%. If a stock started instead in a bad state, its price increased with probability 30% and decreased with probability 70%. The magnitude of the price variation was either 5 ECU, 10 ECU or 15 ECU, each with probability one third. In subsequent rounds, each stock independently remained in the same state with probabilities 80%, and switched state with probability 20%. Each stock thus exhibited positive auto-correlation. In other words, a stock that performed well in the last round was likely to be in a good state in the subsequent round. We predetermined 6 series of price realizations in order to make comparisons easier across participants and treatments.

Participants were paid the value of their portfolio at the end of one of the 3 weeks, selected at random.

**Variations in the design of the algorithm** We consider two ways in which an algorithm can be made less efficient but more attractive to individuals: one ways addresses the issue of keeping control in the hands of the user, the other one addresses the issue of the algorithm precision when giving recommendations.

More precisely, the first variation features the possibility of having the human back in the decision process when the algorithm is in place. That is, the algorithm can be *overriden* (*Soft*) or not (*Hard*) by the human trader. The second variation concerns how *efficient* the algorithm is. We contrast the optimal algorithm (*Bayes*) that sells when the stock goes down, and buys when the stock goes up, with a less efficient and less active algorithm (*Block*), which simply makes no decisions (do not buy or sell, whatever happens). The *Bayes* algorithm was described with the following sentence: "The algorithm is simple and chooses to sell or not to buy a stock whenever the probability it is in a bad state is above 50%, and to buy or keep the stock in the opposite case" while the *Block* algorithm was described as follows: "The algorithm is simple and chooses not to sell and not to buy any stock in the given period".

| | | Week 1 | | | | | | | Week 2 | | | | | | | Week 3 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Day | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Standard treatment | Morning | | | | | | | | | | | | | | | | | | | | | |
| | Afternoon | | | | | | | | | | | | | | | | | | choice, as in week 1 | | | |
| | Evening | | | | | | | | | | | | | | | | | | or as in week 2 | | | |
| Reversed treatment | Morning | | | | | | | | | | | | | | | | | | | | | |
| | Afternoon | | | | | | | | | | | | | | | | | | choice, as in week 1 | | | |
| | Evening | | | | | | | | | | | | | | | | | | or as in week 2 | | | |

Table 1: Chronology of the experiment, standard vs. reversed treatments
Green periods are when the participants trades on their own, red periods are when the robo-adviser is available.

The following table summarizes our variations in the design of the robo-adviser:

| | Block | Bayes |
|---|---|---|
| Hard | Hard Block | Hard Bayes |
| Soft | Soft Block | Soft Bayes |

Table 2: Treatments

# 4 Optimal trading, disposition effect, and measures of performance

As in the paper of Frydman et al. (2014); Frydman and Camerer (2016), our set-up induces positive autocorrelation in stock price changes, which implies that a risk-neutral rational trader ought to sell losing stocks and buy winning stocks, thereby exhibiting the *opposite* of the disposition effect. We can define optimal trading more precisely. Let $p_{it}$ be the price of stock $i$ in round $t$ and let $q_{it} = Pr(s_{it} = good|z_{i,t})$ be the probability, from the point of view of a rational (Bayesian) investor, that stock $i$ is in the good state at time $t$, knowing the price of stock $i$ increased ($z_{it} = 1$) or decreased ($z_{it} = -1$). Then, we have

$$q_{i,t} = \frac{Pr(s_{it} = good \cap z_{it})}{Pr(s_{it} = good \cap z_{it}) + Pr(s_{it} = bad \cap z_{it})} \tag{1}$$

$$= \frac{(0.5 + 0.2z_{it})(0.8q_{i,t-1} + 0.2(1 - q_{i,t-1}))}{(0.5 + 0.2z_{it})(0.8q_{i,t-1} + 0.2(1 - q_{i,t-1})) + (0.5 - 0.2z_{it})(0.8(1 - q_{i,t-1}) + 0.2q_{i,t-1})}$$

The optimal strategy is to sell (if holding) or not to buy (if not holding) a stock $i$ when $q_{i,t} < 0.5$, and to keep (if holding) or buy (if not holding) otherwise. The strategy of the Bayes adviser is based on this probability.

Similar to previous experiments (Frydman et al. 2014, Frydman and Rangel 2014), it is difficult for participants to exactly compute this probability. However, it is possible to approximate this optimal strategy with a simple rule of thumb: *i.e. "hold on stocks that have increased in price last period, sell stocks that have decreased in price last period"*.[3]

[3] *As* in Frydman et al. 2014, *the* only exception to this rule is if the stock went in the same direction on at

## The disposition effect

While optimal trading in our setting consists in buying stocks that go up and selling those that go down, this goes against the well studied the tendency by investors whereby they sell rising stocks too early and keep losing ones too long. This is called the disposition effect and is one of the best documented of various market trading anomalies (see Pleßner 2017 for a review). This trading bias leads to portfolios that are over-weighted in loss positions, thus reducing returns. Neural tests show that individuals experience regret due to this behavior (Frydman and Camerer, 2016). This means that investors who are subject to the disposition effect consider this behavior to be sub-optimal ex-post. This is why helping investors overcome the disposition effect is a valid target for behavioral interventions (Thaler and Sunstein, 2008).

The disposition effect has been the subject of both empirical and experimental research (see Pleßner 2017 for a review). A variety of theories have been proposed to explain it: prospect theory (e.g. Li and Yang 2013); regret minimisation (e.g. Bleichrodt et al. 2010); realisation utility (e.g Frydman et al. 2014; Barberis and Xiong 2009). While the underlying causes of the disposition effect are still debated, the evidence on this phenomenon is very robust. In particular, household investors are more affected by the disposition effect than professional investors, and the disposition effect is greater for females, older people and team investors (Dhar and Zhu 2006, Cueva et al. 2019, Rau 2015). National culture also seems to play a role. Indeed, populations that are more focused on the long-term and less bound by strict social norms have lower average levels of the disposition effect (Breitmayer et al. 2019). Vaarmets et al. (2019) find that better cognitive and learning abilities (as measured by the level and type of education) are correlated with lower disposition effects in on a sample of Estonian traders.

## Measures of performance

**Disposition effect: Diff** Performance in our setting is directly and inversely related to the magnitude of the disposition effect. We first compute this as in the original work of

least four preceding periods, but reversed direction in the current period, in which case the signal from the four preceding periods still dominates.

Odean (1998),[4] that is *Diff*, the difference between the proportion of realised gains ("PGR") and losses ("PLR") over all trading decisions during a given time period, that is:

$$Diff = \frac{\#RealizedGains}{\#(RealizedGains + PaperGains)} - \frac{\#RealizedLoss}{\#(RealizedLoss + PaperLoss)} \tag{2}$$

whereby PGR and PLR take account of all trades made in the positive and in the negative domain, respectively, over a given period.

In this paper, we also consider *Diff_Amount*, which takes account of the magnitude of the gains and losses and is calculated similarly:

$$Diff\ Amount = \frac{ECURealizedGains}{ECURealizedGains + ECUPaperGains} - \frac{ECURealizedLoss}{ECURealizedLoss + ECUPaperLoss} \tag{3}$$

Both indicators have a theoretical range going from -1 to +1, where +1 is the value for an investor that sells all his winning positions and holds all losing ones, -1 is the value for an investor that sells all losing positions and holds all winning ones, and 0 is the value for an investor who behaves the same in both cases. The higher the values of these indicators, the more an individual is subject to the disposition effects.

**Trading optimality: OptScore** As mentioned previously, optimal trading decisions consist in selling or not buying stock $i$ at time $t$ when $q_{i,t} < 0.5$, and buying or holding a stock $i$ at time $t$ when $q_{i,t} > 0.5$. Unlike *Diff*, this does not take account of the price at which a stock was bought, but rather simply of its price evolution. In the vast majority of cases, a stock that went down in price will have $q_{i,t} < 0.5$, and conversely. We therefore construct *OptScore*, an alternative measure to *Diff* that measures how optimal a person's trading was, whereby trading optimally scores +1 while doing the opposite scores as -1. Thus, the higher

[4] Odean (1998) assigns trading decisions to four categories:

- ***realised gain***: a stock that is sold at a price that is higher than the purchasing price;
- ***paper gain***: a stock that is *not* sold but whose price is higher than the purchasing price;
- ***realised loss***: a stock that is sold at a price that is lower than the purchasing price;
- ***paper loss***: a stock that is *not* sold but whose price is lower than the purchasing price;

the score, the closer individual trading behaviour is to optimality. This optimality score can be related to *Diff* in that following the optimal Bayesian strategy, whereby the average score is 1, results in an average level of $Diff = -0.5$, while doing the opposite, whereby the average score is -1, results in $Diff = 0.5$.[5] Not trading (as does the *Block* algorithm) results in *OptScore=0* and *Diff=0.*

**Success rate of the algorithm: Algo_success** As mentioned in our introduction, our specific setting makes it particularly difficult to notice that the advice given by the algorithm is good. Indeed, optimal trading as implemented by the *Bayes* algorithm results in a 54.8% success rate (buying or keeping a stock that then goes up, selling or not buying a stock that then goes down).[6] No trading as implemented by the *Block* algorithm results in a 50% success rate. This has to be contrasted to what a person who is subject to the disposition effect would do (selling when a stock goes up and buying when it goes down), which gives a success rate of 45.2%. Therefore, we construct an average measure of ex-post optimality of the trading algorithm, *Algo Success*: whenever the recommendation of the algorithm for a stock was correct (e.g. buying or keeping a stock that went up) the index assumes value equal to 1, otherwise zero. *Algo success* is computed as the share of correct recommendations given by the algorithm in the week.

## Trading experience

Beyond considering the impact of a trader's own trading style on their willingness to adopt an algorithm, we also consider the impact of other, partly independent, but potentially influential aspects of a trader's experience.

[5]Indeed, a stock that is in a good state remains so with probability 80% and goes up with probability 70%, in which case the stock is kept ($Diff = 0$), and goes down with probability 30% in which case the stock is sold so $Diff = -1$. It switches to a bad state with probability 20%, in which case it goes up with probability 30%, whereby $Diff = 0$, and down with probability 70%, whereby $Diff = -1$. On average therefore, a stock that is in a good state has $Diff = (0.8 \times 0.3 + 0.2 \times 0.7) \times (-1) = -0.38$. A stock that is in a bad state has $Diff = (0.8 \times 0.7 + 0.2 \times 0.3) = -0.62$. Since we assigned stocks randomly to a bad or good state in the first period of trading, the optimal Bayesian trader will have an average level of $Diff = 0.5 \times -0.38 + 0.5 \times -0.62 = -0.5$.

[6]With probability 50% the stock is in a good state. With probability 70%, it goes up and you buy or keep it. Then the stock goes up next period if either it remains in a good state in the next period and goes up (probability 80%×70%), or it switches to a bad state but still goes up (proba. 20%×30%). Your success probability is then 56%+6%=62%. With probability 30% it goes down, and you sell or do not buy it, so you are 80%×30%+ 20%×70%=38% likely to be correct. In total therefore you have a 70%×62%+ 30%×38%=54.8% likely to trade "correctly". The same holds if the stock is in a bad state.

**Trading activity.** The first aspect we aim to control is how many stocks a trader buys and sells each period (*Stocks traded*): this would independently lower their willingness to adopt for example the *Block* algorithm, if they like to trade, or the *Bayes* algorithm, if they prefer not to trade.

**Individual own success**. A second aspect is traders' ex-post experience of success or failure (*Own Success*) when trading alone, that is, whether the stocks an individual bought or kept went up or down next period, and conversely. This is partly independent of trading strategy since as we mentioned on page 15, even the best (Bayesian) strategy results in only a 54.8% likelihood to be correct. However, it is likely to affect a trader's self-confidence, *i.e.* whether they think they are doing well on their own or not. We therefore construct an average measure of ex-post optimality of the individual trader's decisions: whenever the trading decision for a stock was correct (*e.g.* buying or keep a stock that went up) the index assumes value equal to 1, otherwise zero. *Own success* is computed as the share of "correct" decisions made by the individual in a the week without the assistance of the algorithm.

**Overriding** Finally, for *Soft* treatments only, we also compute the percentage of algorithmic decisions participants overrode. The variable *Overrode* takes a value of 0 if the participants went along with what the algorithm proposed (no trade if *Block*, optimal trade if *Bayes*), 1/3 if the participant deviated from the recommendation for one of the stocks only, up to 1 if they deviated from the recommendation for all stocks.

We recapitulate the definition of all those variables in table A.11 in the appendix.

# 5 Hypotheses

Based on our review of the literature, and with reference to our experimental design, we make the following four hypotheses:

1. Participants are subject to the disposition effect. We test this hypothesis by considering *Diff* in the first week of trading when participants trade on their own.

2. The level of the disposition effect is lower for participants who get help from a robo-adviser than for those who trade on their own. We test this hypothesis by comparing *Diff* in the first week in the "standard" treatment, without assistance of a robo-adviser

in week 1, vs. the "reverted" treatment, with assistance of a robo-adviser in week 1.

3. Participants who experienced *Soft* and less active (*Block*) robo-advisers are more likely to adopt them compared to those who experienced *Hard* and more active (*Bayes*) robo-advisers. This is because it is the most predictable and devolves only minimal control to the algorithm. We test this hypothesis by considering the level of adoption of different robo-advisers in the third week.

4. Participants that are the most affected by the disposition effect are also the least likely to adopt robo-advisers in the third week. We test this hypothesis by correlating the level of adoption of the robo-advisers in the third week with the individual performance in the week when a participant traded on his or her own.

The first two hypotheses could be seen as a confirmation and generalization of previous evidence on the disposition effect. They are extended in our novel experimental setting, which spans longer and more realistic timescales (Frydman and Rangel 2014; Fischbacher et al. 2017). As stated above, this longer time horizon is essential to the development of subjects' understanding of the algorithm and to the development of their own ability (Tse et al. 2022; Filiz et al. 2021), in a way that is at the same time coherent with their preferences and habits.

The third hypothesis explores how the two types of algorithm variations impact take-up rates. First of all, regarding trading style, we expect the differences between trading by the algorithm and trading by the individual to matter. In particular, we expect that a better fit between the decisions made by an algorithm and the decisions one would make on one's own to positively affect the take-up rate. Since we expect most traders to suffer from the disposition effect, most will not like trades made by the *Bayes* algorithm. Individuals will judge this algorithm badly as it deviates from their behavior, most often without understanding why it does so, which reinforces their unwillingness to adopt it. Second, we expect the *Soft* algorithm to be preferred, as this allows the investor to correct for differences in trading preferences with the algorithm. This is in line with the general literature on the willingness to adopt commitment devices (e.g.Bryan et al. 2010, Fischbacher et al. 2017). From a rational poiint of view, traders with low skills should make a Ulysses pact and commit not to override the algorithm, but individuals often overestimate their own skill and thus prefer to be free to override. The least skilled individuals will be the most likely to reject hard commitments, because they do not understand they are overconfident (Beshears et al. 2018; Dupas and Robinson 2013; Royer et al. 2015; Beshears et al. 2015; Burke et al. 2018; Duckworth et al. 2016;Bryan et al. 2010). This is also in line with the recent literature on algorithm aversion which suggest that individuals are more willing to adopt algorithm if they are able to slightly modify them (Dietvorst et al. 2018) or able to keep (the feeling of) control (Chugunova and Sele 2020; Burton et al. 2020; Morewedge 2022). Trading style and ability to override interact, in the sense that one may be ready to fully relinquish control to an algorithm that does nothing *(Block)*, but not to one that makes decisions on its own (*Bayes*).

The fourth hypothesis explores the possibility that individuals are not fully aware of their bias when trading and do not understand that the algorithm improves their performance (they are "unskilled and unaware of it", cf. Kruger and Dunning 1999). Indeed, a person who is fully subject to the disposition effect will have an average $Diff = 0.5$, while the Block algorithm has $Diff = 0$ and the Bayes algorithm has an average $Diff = -0.5$. Thus, both algorithms deviate from the behavior of traders that are the most subject to the disposition effect. We expect therefore that the majority of people suffering from the disposition effect will not adopt an algorithm.

In line with previous research (Dietvorst et al. 2015, Dietvorst and Bharti 2020), we expect that investors judge algorithms consequentially, meaning that they do not consider whether its "intentions" (strategy) were good, bur rather whether the consequences of those decisions were good. As a result, we expect *Algo Success* to positively affect algorithm take-up. On the contrary, the individual variable *Own Success* may have an ambiguous effect, whereby traders who trade according to the Bayes algorithm will on average obtain higher success rates, which would encourage them to trade on their own, but also experience the least discrepancy with the Bayes robo-adviser, which encourages them to adopt it.

# 6 Participants and experimental protocol

Data collection started in June 2020 (right after the end of Covid19 lockdown restrictions in Italy) and ended in Mid-August 2020, before the start of a new round of restrictions. Therefore, all our data was collected while the public health situation in Italy was quite stable (e.g discos were open and people could easily travel outside Italy for leisure).[7] Participants were randomly selected out of a pool of about 3000 students from 20 departments of the University of Pisa. Participants were invited to the LES laboratory online with *Microsoft Teams* where they received instructions and a *personal weblink* they could use to play (either on their computer or on their mobile phone). There was a trial session with two periods of trading and participants could ask for clarifications. Although always available online, instructions were also read aloud on *Teams* during the explanatory session. Participants also received instructions in a PDF version (see the english translation in appendix B). In addition to having a general description of the experimental market, participants were told that at the end of the second and third week they would receive additional information about variations that would be introduced in the game.

Once the experiment started, participants received every 8 hours an email reminding of them the beginning of a new trading session, as well as their personal link to play. They were finally told that, at the end of the third week, one randomly selected week would be selected for payment. At this time, they had to fill an exit-questionnaire concerning their experience with the algorithm and perception of (relative) performance. After that, they needed to write an email to the experimenter in order to receive their payment by bank transfer (or cash if preferred).

The following table shows the distribution of participants by treatment:

| | | Block | Bayes |
|---|---|---|---|
| Standard treatment | Hard | 58 | 51 |
| | Soft | 57 | 46 |
| Reversed treatment | Hard | 48 | 45 |
| | Soft | 51 | 53 |

Table 3: Distribution of participants by treatment

[7] Collecting the entire data during the same health policy conditions is important. Indeed Ben-David and Sade (2021) observe a change in adoption rates after Covid-19 compared to pre-Covid-19.

# 7 Results

As stated above, data collection started in June 2020 (right after the end of lockdown restrictions in Italy) and ended in Mid-August. Slightly more than 450 participants took part in the online experiment. The drop out rate was low (about 9%), resulting in a sample of 409 participants who went through all phases, i.e. played all three weeks and claimed payment at the end of the experiment. A large majority of participants were students at the University of Pisa, and were studying engineering or economics. The average age of participants was 25 and 46% were male (see table A.10). The average payment for participation was 17.80 Euro, including a show-up fee of 5 Euro.[8]

We collected information about participants' cognitive ability and level of concern for the future, as well as their financial literacy, locus of control and risk-aversion. On average, participants were able to answer correctly two out of three CRT questions (Cognitive Reflection Test, Frederick 2005), slightly more than two out of three basic financial literacy questions (Lusardi and Mitchell 2007, 2011), and scored about 37 (min 0, max 94) in the consideration of future of consequence scale (Consideration Future Consequences, 12-item scale developed by Strathman et al. 1994). Participants scored 2.6 on average on a 4-level Likert scale of general risk-aversion (see table A.10).

Conditional on being in the sample, participants' activity rate was quite high and stable during all three weeks. Participants actively traded (i.e. either sold or bought at least one stock) in about 57% to 74% of all possible trading periods, i.e. on average about twice per day. We did not observe significant differences in activity levels across treatments.

[8]While having university students in our pool may first appear to be a drawback, younger individuals are particularly worth studying for our research question. Indeed, given their inexperience, they are most likely to benefit from an algorithm (Isaia and Oggero 2022). Furthermore, they are likely to be more comfortable with fully automated algorithm interactions than others (D'Acunto and Rossi 2020). Finally, their typically low level of wealth makes them unattractive for human advisers, so that robo-advisers are an unique opportunity for them. Studying populations that would benefit less, who are already advised, or who are less positively inclined is certainly also worth of study, but we think that investigation should start exactly with our type of sample. Indeed, a robo-adviser that is not adopted in our case is even less likely to be adopted in other cases. Furthermore, as ever, there is no reason to think treatment differences would affect different population groups in different ways. It is therefore likely that our results will hold for other populations, if not in level, at least in nature.

### 7.1 Do people suffer from the disposition effect and do they benefit from using an algorithm?

We test our first two hypotheses in this part. As stated in section 4, our setting implies that a risk-neutral expected value maximizer would exhibit the opposite of the disposition effect – that is, negative *Diff* values.

We report the average disposition effect for our participants in Table (4) for week 1, depending on whether they had access to a robo-adviser or not.[9] As Table (4) highlights, the value of *Diff* (0.088) and *Diff_Amount* (0.109) was greater than zero in the first week if there was no robo-advisers. This was the opposite if participants got the help of a robo-adviser in the first week, whereby *Diff* (-0.116) and *Diff Amoun*t (-0.133) were lower than 0 overall. The disposition effect was particularly low in the *Hard Bayes* treatment, where trade was made according to a Bayesian algorithm every two period.

We find that participants overrode 21.9% of decisions by the Bayesian algorithm in week 1, which explains in part why the disposition effect was not as low in that treatment than in the *Hard Bayes* treatment. They overrode the *Block* algorithm less often (12.0%), so the disposition effect was similar to its level in the *Hard Block* treatment. Note that this pattern of overrides is the opposite of what an optimal trader would be doing, whereby he or she would override the *Block* algorithm often, while he or she would never override the Bayesian algorithm.

Finally, we find as expected that the average number of stocks traded per period was lower in the *Block* treatments than if there was no robo-adviser, and higher in the *Bayes* treatments. Allowing participants to override the algorithm led to higher stock trading in the *Block* treatment, and lower stock trading in the *Bayes* treatment, which is consistent with going against algorithmic recommendations.

Overall, we find that the Bayes robo-adviser improved trading the most, but that even the *Block* adviser improved performance. However, help came at the cost of trading less (*Block*) or more (*Bayes)* than what participants would have wanted. This explains why participants were quite likely to override algorithms when they could do so.

[9] To make the cleanest and simplest comparisons, in the following we only focus on between treatment differences. Within treatment analyses, not reported but available upon request, provide consistent evidence.

Table 4: AVERAGE DISPOSITION EFFECT AND OPTIMALITY OF TRADE IN WEEK 1, NORMAL VS. REVERSED TREATMENT

| | TREATMENT | Diff | DiffAmount | OptScore | Stocks traded | Overrides | Obs |
|---|---|---|---|---|---|---|---|
| Reverted treatment | Hard Block | -0.100 | -0.130 | 0.132 | 0.33 | NA | 48 |
| (robo adviser in week 1) | Hard Bayes | -0.409 | -0.486 | 0.554 | 1.14 | NA | 45 |
| (by treatment) | Soft Block | -0.109 | -0.116 | 0.166 | 0.44 | 12.0% | 51 |
| | Soft Bayes | -0.217 | -0.259 | 0.320 | 0.92 | 21.9% | 53 |
| | Average | -0.116 | -0.133 | 0.288 | 0.70 | NA | 197 |
| Standard treatment (no robo adviser in week 1) | Average | 0.088 | 0.109 | 0.019 | 0.48 | NA | 212 |

To get a better sense of the effect of each type of algorithm on performance in the first week of our experiment, we run the following regression

$$Y_{i,T} = \beta_0 + \mathbf{Treatment_T}\beta' + \mathbf{Controls_i}\gamma' + \epsilon_{i,T} \tag{4}$$

where $Y_{i,T}$ is the individual level of *Diff*, *Diff_Amoun*t, and *OptScore* in the first week of the experiment in treatment $T$, and *Treatment* is $1 \times T$ vector of dummy variables for our treatments: Hard Bayes, Soft Bayes, Hard Block, and Soft Block. $Controls_i$ is a $1 \times K$ vector of variables including a set of $K$ individual characteristics collected through the questionnaire, such as age, gender, financial literacy, future attitude (*CFC*), cognitive reflection (*CRT*), locus of control (*LoC)* and risk aversion. Results are reported in Table (5). The intercept $\beta_0$ represents the level of the dependent variable in the baseline group (i.e. the average level of $Y$ in the standard treatment where individuals were not exposed to any algorithm in the first week).

Table 5: Performance in the first week

| | (1) Diff | (2) Diff Amount | (3) Opt Score | (4) Diff | (5) Diff Amount | (6) Opt Score |
|---|---|---|---|---|---|---|
| Hard Block | $-0.100^{*}$ | $-0.130^{**}$ | $0.113^{**}$ | $-0.088^{*}$ | $-0.119^{*}$ | $0.109^{*}$ |
| | (0.042) | (0.049) | (0.042) | (0.042) | (0.049) | (0.043) |
| Soft Block | $-0.109^{**}$ | $-0.116^{*}$ | $0.147^{***}$ | $-0.098^{*}$ | $-0.105^{*}$ | $0.145^{***}$ |
| | (0.041) | (0.047) | (0.041) | (0.041) | (0.048) | (0.042) |
| Hard Bayes | $-0.409^{***}$ | $-0.486^{***}$ | $0.535^{***}$ | $-0.398^{***}$ | $-0.475^{***}$ | $0.532^{***}$ |
| | (0.043) | (0.050) | (0.043) | (0.044) | (0.051) | (0.044) |
| Soft Bayes | $-0.217^{***}$ | $-0.259^{***}$ | $0.301^{***}$ | $-0.203^{***}$ | $-0.246^{***}$ | $0.299^{***}$ |
| | (0.040) | (0.047) | (0.041) | (0.041) | (0.047) | (0.041) |
| Male | | | | 0.002 | 0.008 | −0.017 |
| | | | | (0.028) | (0.032) | (0.028) |
| Age | | | | 0.001 | 0.001 | −0.002 |
| | | | | (0.003) | (0.004) | (0.003) |
| CRT | | | | −0.017 | −0.014 | 0.005 |
| | | | | (0.013) | (0.015) | (0.013) |
| CFC | | | | 0.001 | 0.001 | −0.000 |
| | | | | (0.003) | (0.003) | (0.003) |
| Financial literacy | | | | −0.022 | −0.032 | −0.005 |
| | | | | (0.020) | (0.023) | (0.020) |
| LoC | | | | 0.005 | 0.003 | 0.007 |
| | | | | (0.008) | (0.009) | (0.008) |
| Risk_aversion | | | | −0.012 | −0.007 | 0.012 |
| | | | | (0.017) | (0.019) | (0.017) |
| Constant | $0.088^{***}$ | $0.109^{***}$ | 0.019 | 0.082 | 0.117 | 0.018 |
| | (0.018) | (0.021) | (0.018) | (0.160) | (0.185) | (0.161) |
| Observations | 409 | 409 | 409 | 409 | 409 | 409 |

Standard errors in parentheses

Dependent variables: Diff, Diff_amount and Optimality Score

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

To begin with, we can notice in column (1) and (2) that *Diff* and *Diff_amount* (0.088 and 0.109 respectively) are both significantly different from zero 0 in the baseline group: individuals who were not exposed to an algorithm in the first week suffered from the disposition effect. The *Hard Block* algorithm results in a significant reduction in *Diff* (-0.100 significant at 5% level), and a significant increase (0.130 significant at 1% level) in the optimality score (see column 3). Similarly, the *Hard Bayes* algorithm significantly reduces the level of *Diff* (-0.409 significant at 0.1% level), and increases the level of optimal trading (0.147 significant at 0.01% level). Moreover, the *Hard Bayes* algorithm improves on the *Hard Block algorithm* (a significant decrease in *Diff*, F-test=0.000, a significant increase in optimality, F-test=0.000). Noticed that all these effects are also economically significant: *Diff* varies theoretically from -1 to +1, and its average level in the base sample (without help of an algorithm) is 0.088 (*Diff*). The effect sizes therefore range from 134% to 465% of the baseline.[10]

Looking at the *Soft* treatments, we notice that this variation does not impacts results for the *Block* algorithm (no significant diffence with the *Hard Block* treatment). On the contrary, the *Soft* variation on the *Bayes* algorithm significantly affects *Diff* compared to the *Hard Bayes* treatment (F-test=0.029). Finally, the *Soft* variation on the *Bayes* treatment still is better than the Hard Block treatment, but less significantly so (F-test=0.029 for *Diff*, and F-test=0.0004 for *Optscore*). No significant results emerge if we look at the alternative indicator of *Diff, Diff_Amount* (see column 2). The last three columns further show that none of the individual control variables have a significant impact on any of our performance measures, while treatment effects remain significant.

## 7.2 Who adopts robo-advisers, which are preferred and why?

We test our hypotheses 3 and 4 in this section. In the third and last week of our experiment, participants could decide how to play the remaining rounds of the game, *i.e.* whether to play with the assistance of an algorithm or not. The overall adoption rate was quite low, as only 36.7% of our participants decided to rely on a trading algorithm in the third week (see Table 6). This is less than the 55% of them who suffered from the disposition effect (*Diff*

[10]These effect sizes are calculated as $\frac{\hat{\beta}_{treated}}{Control\ group\ mean}$.

>0) in the week when they traded on their own, meaning that they would have benefited even from the block algorithm, and this is of course less than the 100% who would have benefited from using the Bayes algorithm.

Table 6: ADOPTION OF ROBO-ADVISER, IN %, BY TREATMENT.

| | **Block** | | **Bayes** | | Total |
|---|---|---|---|---|---|
| | HARD | SOFT | HARD | SOFT | |
| Standard treatment (robo in week 2) | 19.0% | 31.6% | 41.2% | 50.0% | 34.4% |
| Reverted treatment (robo in week 1) | 37.5% | 43.1% | 37.8% | 37.8% | 39.1% |
| Overall adoption | 27.4% | 37.0% | 39.6% | 43.4% | 36.7% |

If we consider now differences in adoption across treatments, we notice that there are minimal treatment differences if the algorithm was experienced in week 1, and the level of adoption in that case is rather high overall at 39.1%. The opposite holds if the algorithm was experienced in week 2, whereby we observe large differences across treatments (*Soft* algorithms are preferred to *Hard*, and *Bayes* are preferred to *Block*), while the overall level of adoption is lower at 34.4%. This points towards a recency effect, whereby differences between robots are diluted with time as their memory fades, while overall attitudes to them improves.

Beyond treatment differences, we also consider how adoption rates depend on behavior *when trading alone*: individual trading style (as measured with average *Diff*, *Diff_amount* and *Optscore*), own success rate (*Own Success*), and average number of stocks traded per period (*Stock traded*). We also take account of average success of the algorithm in periods when it traded for them (*Algo Success*). As mentioned on page 15, own success rates are computed in terms of the percentage of one's own trades that resulted in gains or avoided losses next period, and algorithm success rates are the percentage of the trades recommended or made by the algorithm in the week with a robo-adviser that turned out to be right (see also table A.11 for a recapitulation of the definition of our variables). We thus observe adoption rates depending on the quartiles of those variables.

Table 7: ADOPTION RATE BY QUARTILE OF THE COLUMN VARIABLE

| QUARTILE | DIFF | DIFF AMOUNT | OPTIMALITY SCORE | OWN SUCCESS RATE | STOCKS TRADED | ALGO SUCCES RATE |
|---|---|---|---|---|---|---|
| 1 | 41% (-0.35) | 41% (-0.42) | 33% (-0.22) | 46% (0.43) | 40% (0.11) | 25% (0.42) |
| 2 | 40% (-0.03) | 40% (-0.04) | 34% (-0.01) | 36% (0.50) | 29% (0.33) | 37% (0.50) |
| 3 | 33% (0.09) | 34% (0.13) | 41% (0.19) | 28% (0.53) | 38% (0.58) | 45% (0.57) |
| 4 | 32% (0.35) | 31% (0.43) | 39% (0.52) | 31% (0.59) | 41% (0.95) | 41% (0.66) |

This table shows adoption rate for each quartile of the variable in column. Quartiles are computed based on the distribution of that variable over all individuals. We show in parenthesis the average value of the variable for individuals in the respective quartiles.

Individuals in the first and second quartiles of *Diff* - *i.e.* those less subject to the disposition effect - are more likely to adopt an algorithm than those in the third and fourth quartiles of *Diff* (about 40% vs about 32% respectively, see table 7). The same pattern holds, but reverted, for the optimality score, whereby those who trade more optimally are also more likely to adopt an algorithm.

However, a yet stronger pattern emerges in terms of the influence of one's success rate when trading on one's own, whereby those in the bottom quartile, who have an average 43% success rate, adopt the robo-adviser 46% of the time, while those in the top quartile, who have an average success rate of 59%, adopt the robo-adviser only 31% of the time. The same pattern, but reversed, emerges in terms of experience of success with the algorithmic recommendations. On the one hand, individuals who are less affected by the disposition effect are more likely to adopt a robo-adviser as it is consistent with choices they would have made by themselves. On the other hand, their better trades lead them to having higher success rates,[11] which should make it less likely that they would adopt the robo-adviser. There is thus a possible interaction between one's own style of trading and one's success when trading, whereby those who trade optimally but are not successful are the most likely to then adopt a robo-adviser.

To get a synthetic view of how each variable affects the adoption rate we run a set of logistic regressions, separately for the reverted and standard treatments (as well as on the full sample), as follows:

$$\begin{aligned} Logit(Adoption)_{week\,3,i} = \beta_1 Bayes_i + \beta_2 Soft_i + \beta_3 Diff_i + \beta_4 Stocks\,traded_i + \\ \beta_5 Own\,success_i + \beta_6 Algo\,success_i + controls\gamma' + \epsilon_i \end{aligned}$$

where the dependent variable is the dummy *Adoption* equal to 1 if an individual opted for getting assistance from an algorithm in the third week (and zero otherwise). *Diff, Diff amount*, *OptScore*, *Own Success*, *Stocks traded* are averages in the week where an individual traded on his own, and *Algo success* is an average measured in the other week. The [1xK] vector *controls* includes a series of individual variables we collected through the entry and

[11] from 46% success rate in the first quartile of the optimality score to 54% in the fourth quartile.

exit questionnaire. In particular, we control for cognitive reflection ability (*CRT*), financial literacy (*Financial Literacy*), consideration for future consequences (*CFC*), locus of control (*LoC*), and risk aversion (*Risk Aversion*) — see page 20 for more details on those variables.

Results (average marginal effects) are reported in Table (8), whereby we split results depending on whether the participants experienced the robo-adviser in week 1 (Reverted) or in week 2 (Standard). In columns (1) and (5), we confirm that the trading style of robo-advisers matters mainly only if robo-advising was experienced recently, i.e. in the Standard treatment. In that case, the more sophisticated type of algorithm (i.e. *Bayes)* is preferred (20 percentage points ("pp"), significant at the 1‰ level). The soft robo advisers (i.e. *Soft)* are preferred, significantly so in the standard treatment (11 pp, significant at the 1‰ level). In columns (2) and (6), we control for one's own style of trading, as measured by *Diff* and *Stocks traded*. We find that *Diff* only matters if one experienced trading on one's own recently, i.e. in the Reverted treatment. In that case, those who are most affected by the disposition effect (high *Diff*) are also those least likely to adopt an algorithm (13 pp). In columns (3) and (7), we find that experiencing success on one's own (*Own Success*) decreases the likelihood of adopting the algorithm. This effect is both statistically and economically significant.[12] The value of the estimated parameter is consistent whether trading one one's own was recent (reverted treatments, 104 pp) or far in time (standard treatment, 71 pp), although we observe differences in terms of statistical significance. Similarly, experiencing success when trading with the algorithm (*Algo Success*) increases the likelihood of adopting it. The effect is both economically and statistically significant both in the reverted (71 pp) and the standard treatment (80 pp).

Columns (4) and (8) show that results are robust to the inclusion of individual characteristics, among which only risk aversion appears to play a role (it increases take-up rate by by 7 and 10 pp). Age also seems to impact the decision to adopt the algorithm but only in the standard treatment. Being older decreases the probability of adopting the financial algorithm. The effect, however, is not economically significant (0.1 pp). Results are substantially analogous when considering the full sample together, or when considering measures of trading style other than *Diff,* such as *Diff_amount* and *OptScore* (not reported

[12]Remember the variable *Own success* varies between 0 and 1, and the average level in the sample is 0.50

but available upon request).[13]

Overall, these results suggest that participants prefer a robo adviser that trades optimally (*Bayes*), but this preference only appears if experience with algorithms was recent. Own trading style also matters for adoption, but only if one traded on one's own recently. Those differences depending on how the sequence in which the robo-adviser was experienced make sense in our experiment, which was run over several weeks, whereby experience made in the past week would be more pregnant than experiences made two weeks ago. Traders generally reject the Hard types of algorithm. They prefer being able to intervene when they do not like the decision implemented by the algorithm. Experience of success or failure when trading and when the algorithm trades also matter. The more the individual is successful, the lower the probability to take-up the algorithm. The higher the success-rate of the algorithm, the higher the probability to adopt the algorithm. In the next section, we look at the data collected in our exit questionnaires, specifically to the stated reasons participants gave for adopting (or not) an algorithm. This allows us to better understand the reasons behind individual choices to adopt the algorithm

### 7.3 Were adopters more sophisticated?

As shown in table 7, individuals who were more prone to the disposition effect were also the least likely to adopt a robo-adviser. However, results in table 8 suggest that those who experienced the algorithm recently were sensitive to the advantage of the optimal, Bayes robo-adviser. Experiencing success or failure was however the main relevant variable in driving adoption. This means that individuals may have judged the benefit of adoption not based on their own view of what a good trading strategy is, but rather based on their experience. This even though that experience is only weakly related to whether their decisions were good or bad. This result is likely to hold quite generally in very noisy environments such as trading, where success or failure depends on many factors other than one's

[13]We also considered whether own trading style impacted adoption in a different way depending on whether one had to adopt the Bayes or the Block algorithm. Indeed, if the fit between the individual and the algorithm matters, then low *Diff* or high *OptScore* would mainly ease adoption of the *Bayes* algorithm, while low *Stocks traded* would mainly ease adoption of the *Block* algorithm. However, differences in adoption rates depending on the algorithm occur only if one experienced algorithmic trade in the second week. There are then not enough observations left to reliably assess if, in that case, those two variables have such a different impact on adoption of those two algorithms.

Table 8: ADOPTION IN THE THIRD WEEK, DEPENDING ON WHEN THE ROBO-ADVISER WAS IN PLACE

| | Reverted (Robo in 1st week ) | | | | Standard (Robo in 2nd Week) | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Bayes | -0.027 | -0.019 | 0.003 | -0.009 | 0.204*** | 0.215*** | 0.146*** | 0.159*** |
| | (0.069) | (0.013) | (0.008) | (0.019) | (0.020) | (0.022) | (0.019) | (0.025) |
| Soft | 0.028 | 0.030** | 0.035*** | 0.036*** | 0.109*** | 0.104*** | 0.097*** | 0.088*** |
| | (0.070) | (0.010) | (0.009) | (0.004) | (0.020) | (0.018) | (0.015) | (0.015) |
| Diff | | -0.132** | -0.141** | -0.172*** | | 0.046 | 0.024 | 0.062 |
| | | (0.042) | (0.052) | (0.020) | | (0.124) | (0.093) | (0.110) |
| Stocks traded | | -0.047 | -0.027 | -0.072 | | 0.117 | 0.139** | 0.137+ |
| | | (0.127) | (0.104) | (0.094) | | (0.072) | (0.051) | (0.070) |
| Own Success | | | -1.042*** | -0.928*** | | | -0.709 | -0.753 |
| | | | (0.108) | (0.276) | | | (0.757) | (0.776) |
| Algo success | | | 0.707*** | 0.643** | | | 0.798* | 0.737* |
| | | | (0.168) | (0.224) | | | (0.385) | (0.327) |
| Male | | | | -0.005 | | | | -0.036+ |
| | | | | (0.095) | | | | (0.020) |
| Age | | | | 0.001 | | | | -0.009* |
| | | | | (0.015) | | | | (0.003) |
| CRT | | | | 0.018 | | | | 0.001 |
| | | | | (0.070) | | | | (0.027) |
| CFC | | | | -0.008 | | | | -0.007 |
| | | | | (0.006) | | | | (0.011) |
| Financial Literacy | | | | -0.066 | | | | -0.018 |
| | | | | (0.081) | | | | (0.051) |
| Control | | | | 0.007 | | | | -0.023 |
| | | | | (0.009) | | | | (0.027) |
| Risk_aversion | | | | 0.099*** | | | | 0.071*** |
| | | | | (0.021) | | | | (0.015) |
| Observations | 197 | 197 | 197 | 197 | 212 | 212 | 212 | 212 |

Clustered errors in parentheses, + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

own actions. We also found that individuals generally preferred soft algorithms, meaning that they were not only not sophisticated enough to distinguish sound investment strategy and good trading results, but also not sophisticated enough to recognize that, if the robo-adviser follows a sounder strategy than themselves, then they might want not to be able to override algorithmic decisions.

We further examine those results by looking at the reasons participants gave for adopting or rejecting the algorithm, as expressed in the exit questionnaire (Table A.12).[14] Each participants could mention only one reason. We find that performance played a bigger role for adopters than non-adopters: 48% of adopters cited improved performance as the main reason, while only 25% of non-adopters thought the algorithm reduced their performance. This difference is both economically and statistically significant at the 1% level (test for proportions). A desire to keep control played a role for 40% of non-adopters, while its converse among adopters, such as a desire to free one's time or mind from trading, was also important. Indeed, 17% of adopters mentioned how the algorithm freed their mind, and 14% mentioned how the algorithm freed their time. Enjoyment played a larger role for non-adopters than for adopters: only 1% of adopters said they did not like trading, while 18% of non-adopters said they had fun trading. Finally, only 5% of the adopters reported how the algorithm helped them reduce their temptation to trade, while 6% of the non-adopters specifically mentioned they knew best how to trade. Being conscious of self-control issues was therefore probably not a main driver for adoption.

To get a concise view about the role of each of these motivations on algorithm adoption, we run the following regression

$$\begin{aligned} Logit(Adoption)_{week\,3,i} &= \beta_1 Bayes_i + \beta_2 Soft_i + \beta_3 Performance_i + \beta_4 Selfconfidence_i + \\ &\quad + \beta_5 Control_i + \beta_6 Fun_i + controls\gamma' + \epsilon_i \end{aligned}$$

for the standard and reverted treatments, where *Performance*, *Self-Control*, *Control* and

[14] Individuals adopting the algorithm were asked "Why did you select an algorithm"? Possible answers were: 1a) I believe the algorithm improves my performance; 2a) There are trades that I should not do; 3a) It frees my time 4a) I do not like to make transactions 5a) It frees my mind 6a) I do not know 7a) other reasons. Individual not adopting the algorithm were asked "Why did you not select an algorithm"? Possible answers were: 1b) I believe the algorithm reduces my performance; 2b) I know when to do trades; 3b) I have fun trading 4b) I like to keep control 5b) I do not know 6b) other reasons.

*Fun* are variables measuring the individual stated reasons to adopt the algorithm or not,[15] while *Bayes*, *Soft* and *controls* are defined as in 7.2.

Table 9: Adoption in the third week and reasons to adopt

| | REVERTED (ROBO 1ST WEEK) (1) | (2) | STANDARD (ROBO 2ND WEEK) (3) | (4) |
|---|---|---|---|---|
| Bayes | -0.037* | -0.039* | 0.131*** | 0.136*** |
| | (0.018) | (0.019) | (0.006) | (0.014) |
| Soft | 0.037* | 0.055** | 0.012 | 0.000 |
| | (0.015) | (0.018) | (0.014) | (0.012) |
| Performance algo | 0.203*** | 0.188*** | 0.213*** | 0.206*** |
| | (0.026) | (0.025) | (0.008) | (0.013) |
| Self-confidence | -0.191 | -0.212* | -0.149** | -0.144*** |
| | (0.118) | (0.090) | (0.047) | (0.043) |
| Fun | -0.132 | -0.141+ | -0.182 | -0.189+ |
| | (0.084) | (0.086) | (0.116) | (0.102) |
| Control | -0.341*** | -0.386*** | -0.256*** | -0.262*** |
| | (0.022) | (0.023) | (0.014) | (0.016) |
| Male | | -0.013 | | -0.007 |
| | | (0.043) | | (0.042) |
| Age | | -0.001 | | -0.008* |
| | | (0.004) | | (0.004) |
| CRT | | 0.037+ | | 0.006 |
| | | (0.022) | | (0.016) |
| CFC | | 0.007+ | | -0.001 |
| | | (0.004) | | (0.007) |
| Financial Literacy | | -0.094** | | -0.001 |
| | | (0.036) | | (0.025) |
| Control | | 0.018 | | -0.006 |
| | | (0.019) | | (0.015) |
| Risk aversion | | 0.107*** | | 0.082*** |
| | | (0.028) | | (0.021) |
| Observations | 194 | 194 | 201 | 201 |

**Standard errors in parentheses,** $^{+}$ $p < 0.1$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Overall these results confirm that those individuals who adopted the algorithm attached significantly more importance to the performance aspect of this decision. However, and in line with the recent analysis of Rossi and Utkus (2020), these results also suggest

[15] With reference to table A.12, *performance* equals 1 (-1) if the answer is 1a (1b); *self-confidence* equals 1 (-1) if the answer is 2b (2a); *fun* equals 1 (-1) if the answer is 3b (4a) ; *control* equals 1 (-1) if the answer is 4b (3a and 5a). The residual category is 1 if the answer is 6a, 7a, 5b or 6b.

that some individuals decided to rely on financial advisers not so much for portfolio return maximization but rather with the aim of having peace of mind and saving time by delegating financial decisions. Conversely, not adopting the algorithm was associated with the need to keep control of transactions, in other words, non-adopters did not feel secure in leaving decisions to an algorithm.

# 8 Discussion and conclusion

The results from our research shed light on how robo advisers can help investors overcome behavioral issues, such as, in our case, the disposition effect. Our experiment, conducted online over 3 weeks, highlights drivers of adoption of financial algorithms along three dimensions: control left to the investor, skill level of the investor, and success rate of the investor and of the algorithm. Along with the literature, we find that investors are not willing to restrict their freedom to trade in order to achieve better outcomes. They do not realize that committing not to trade on their own (as implemented in our "Hard" treatments) would allow them to improve their performance. This explains why robo-advisers that can be overridden are more likely to be adopted. However, we also found that letting participants override algorithmic decisions did not fully negate the benefits of having an algorithmic adviser. Unlike expected, we found that investors were more likely to adopt a robo-adviser that trades actively rather than one that simply does not trade. This seems to contradict their preference for keeping control and thus limiting delegation to an algorithm.

We were able to disentangle the impact of trading strategy *vs.* trading success by considering whether likelihood to adopt an algorithm depended on a trader's own trading strategy, while controlling for the success rate of the algorithm and that of the investor. We found that investor's own trading style affected adoption, but only if they had to trade on their own recently. In that case, those who traded well (low disposition effect) were also more likely to adopt the algorithm. The effect of own success when trading was robust, whereby higher success rate lowered rates of adoption of an algorithm. Those who were the most likely to adopt an algorithm were therefore those with good trading strategy but low success rate when trading on their own. Overall, our results suggest that individuals judge the benefit of adoption mainly based on the algorithm's success rate compared to theirs, even though this success rate is only weakly related to whether a strategy is good or bad in our experiment. An analysis of the stated reasons for adopting a robo-adviser corroborates this analysis, further suggesting that individuals like to adopt advisers for "peace of mind" (see Rossi and Utkus 2020).

In line with previous evidence (e.g. D'Acunto et al. 2019), we also found that individual characteristics - such as financial literacy (Bhattacharya et al. 2012) - were irrelevant for adoption, the only exception being the level of risk aversion, as more risk-averse individuals were more willing to adopt an algorithm (see also Oehler et al. 2022; Kawaguchi 2021). We also have suggestive evidence that younger individuals were more willing to adopt the algorithm.

Our research, together with evidence emerging from related studies, suggests important directions for encouraging delegation of financial decisions to algorithms, especially for young and less wealthy investors. In line with other research (Tse et al. 2022; Filiz et al. 2022; Chacon et al. 2022), the benefits of robo-advisers should be made more obvious for those investors who perform worst on the stock market, as they also may not realize how badly they are performing and how much they could benefit from the use of an adviser. This means that they should not only get trading experience both with and without a robo-adviser, but the difference in their performance in both conditions should be conveyed to them in very simple and transparent terms (*e.g.* by promoting algorithm literacy). Second, we confirm that giving people the ability to override the adviser increases adoption while still improving trading performance. Simply offering the option to not approve trades by the robo-adviser is a simple and effective way to enhance the feeling of being in control of decisions, and thus overcome algorithm aversion.

Further research could look into whether adoption is encouraged by giving people more leeway in the design of their advisers, that is, by letting them determine themselves what algorithm to use, rather than simply whether to adopt the one they are offered. For example, one could let them vary the strength of their commitment to follow the adviser, e.g. by putting a price on overrides. It would also be interesting to consider the impact of a "mixed model" of robo-advising, where a human conveys the advice of the algorithm. This could enhance adoption, in particular among older people and/or those who may be less technically savvy and thus more reluctant to interact with robo-advisers.

# References


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

Bailey, R. W. (1993, October). Performance vs. Preference. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 37*(4), 282–286.

Barberis, N. and W. Xiong (2009). What drives the disposition effect? an analysis of a long-standing preference-based explanation. *Journal of Finance 64*(2), 751–784.

Ben-David, D. and O. Sade (2021). Robo-advisor adoption, willingness to pay, and trust — before and at the outbreak of the covid-19 pandemic. SSRN Working Paper, Hebrew University of Jerusalem - Department of Finance.

Beshears, J., J. J. Choi, C. Harris, D. Laibson, B. C. Madrian, and J. Sakong (2015). Self control and commitment: can decreasing the liquidity of a savings account increase deposits? NBER Working Paper.

Beshears, J., J. J. Choi, D. Laibson, and B. C. Madrian (2018). Potential vs. realized savings under automatic enrollment. *TIAA Institute. Research Dialogue* (148).

Bhatia, A., A. Chandani, and J. Chhateja (2020). Robo advisory and its potential in addressing the behavioral biases of investors — a qualitative study in an indian context. *Journal of Behavioral and Experimental Finance 25*, 100281.

Bhattacharya, U., A. Hackethal, S. Kaesler, B. Loos, and S. Meyer (2012). Is unbiased financial advice to retail investors sufficient? answers from a large field study. *The Review of Financial Studies 25*(4), 975–1032.

Bleichrodt, H., A. Cillo, and E. Diecidue (2010). A quantitative measurement of regret theory. *Management Science 56*(1), 161–175.

Breitmayer, B., T. Hasso, and M. Pelster (2019). Culture and the disposition effect. *Economics Letters 184*, 108653.

Brenner, L. and T. Meyll (2020). Robo-advisors: A substitute for human financial advice? *Journal of Behavioral and Experimental Finance 25*, 100275.

Bryan, G., D. Karlan, and S. Nelson (2010). Commitment devices. *Annu. Rev. Econ. 2*(1), 671–698.

Burke, J., J. Luoto, and F. Perez-Arce (2018). Soft versus hard commitments: a test on savings behaviors. *Journal of Consumer Affairs 52*(3), 733–745.

Burton, J. W., M.-K. Stein, and T. B. Jensen (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making 33*(2), 220–239.

Capponi, A., S. Olafsson, and T. Zariphopoulou (2022). Personalized robo-advising: Enhancing investment through client interaction. *Management Science 68*(4), 2485–2512.

Castelo, N., M. W. Bos, and D. R. Lehmann (2019). Task-dependent algorithm aversion. *Journal of Marketing Research 56*(5), 809–825.

Chacon, A., E. E. Kausel, and T. Reyes (2022). A longitudinal approach for understanding algorithm use. *Journal of Behavioral Decision Making 35*(4). e2275.

Chang, T. Y., D. H. Solomon, and M. M. Westerfield (2016). Looking for someone to blame: Delegation, cognitive dissonance, and the disposition effect. *The Journal of Finance 71*(1), 267–302.

Chen, D. L., M. Schonger, and C. Wickens (2016). oTree: an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance 9*, 88–97.

Chugunova, M. and D. Sele (2020). We and it: An interdisciplinary review of the experimental evidence on human-machine interaction. *Max Planck Institute for Innovation & Competition Research Paper* (20-15).

Cueva, C., I. Iturbe-Ormaetxe, G. Ponti, and J. Tomás (2019). An experimental analysis of the disposition effect: Who and when? *Journal of Behavioral and Experimental Economics 81*, 207–215.

D'Acunto, F., N. Prabhala, and A. G. Rossi (2019). The promises and pitfalls of robo-advising. *The Review of Financial Studies 32*(5), 1983–2020.

D'Acunto, F. and A. G. Rossi (2020). Robo-advising. CESifo Working Paper 8225, Boston College and Georgetown University.

Deloitte (2016). The expansion of Robo-Advisory in Wealth Management. https://www2.deloitte.com/de/de/pages/financial-services/articles/the-expansion-of-robo-advisory-in-wealth-management.html.

Dhar, R. and N. Zhu (2006). Up close and personal: Investor sophistication and the disposition effect. *Management Science 52*(5), 726–740.

Dietvorst, B. J. and S. Bharti (2020). People reject algorithms in uncertain decision domains because they have diminishing sensitivity to forecasting error. *Psychological science 31*(10), 1302–1314.

Dietvorst, B. J., J. P. Simmons, and C. Massey (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General 144*(1), 114.

Dietvorst, B. J., J. P. Simmons, and C. Massey (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science 64*(3), 1155–1170.

Duckworth, A. L., T. S. Gendler, and J. J. Gross (2016). Situational strategies for self-control. *Perspectives on Psychological Science 11*(1), 35–55.

Dupas, P. and J. Robinson (2013). Why don't the poor save more? evidence from health savings experiments. *American Economic Review 103*(4), 1138–71.

Faloon, M. and B. Scherer (2017, April). Individualization of Robo-Advice. *The Journal of Wealth Management 20*(1), 30–36.

Filiz, I., J. R. Judek, M. Lorenz, and M. Spiwoks (2021). Reducing algorithm aversion through experience. *Journal of Behavioral and Experimental Finance 31*, 100524.

Filiz, I., J. R. Judek, M. Lorenz, and M. Spiwoks (2022). Algorithm aversion as an obstacle in the establishment of robo advisors. *Journal of Risk and Financial Management 15*(8), 353.

Fischbacher, U., G. Hoffmann, and S. Schudy (2017). The causal effect of stop-loss and take-gain orders on the disposition effect. *The Review of Financial Studies 30*(6), 2110–2129.

Foerster, S., J. T. Linnainmaa, B. T. Melzer, and A. Previtero (2017). Retail financial advice: does one size fit all? *The Journal of Finance 72*(4), 1441–1482.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives 19*(4), 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*(2), 907–946.

Frydman, C. and C. Camerer (2016). Neural evidence of regret and its implications for investor behavior. *The Review of Financial Studies 29*(11), 3108–3139.

Frydman, C. and A. Rangel (2014). Debiasing the disposition effect by reducing the saliency of information about a stock's purchase price. *Journal of Economic Behavior & Organization 107*, 541–552.

Frydman, C. and B. Wang (2020). The impact of salience on investor behavior: Evidence from a natural experiment. *The Journal of Finance 75*(1), 229–276.

Germann, M. and C. Merkle (2022). Algorithm aversion in delegated investing. *Journal of Business Economics*, 1–37.

Hidalgo, C. A., D. Orghian, J. A. Canals, F. De Almeida, and N. Martín (2021). *How humans judge machines*. MIT Press.

Hirshleifer, D. (2015). Behavioral finance. *Annual Review of Financial Economics 7*, 133–159.

Holzmeister, F., M. Holmén, M. Kirchler, M. Stefan, and E. Wengström (2022). Delegation decisions in finance. *Management Science*. Online in advance of print.

Isaia, E. and N. Oggero (2022). The potential use of robo-advisors among the young generation: Evidence from Italy. *Finance Research Letters*, 103046.

Kahneman, D. (2002). Maps of bounded rationality: A perspective on intuitive judgment and choice. *Nobel Prize Lecture*.

Kawaguchi, K. (2021). When will workers follow an algorithm? A field experiment with a retail business. *Management Science 67*(3), 1670–1695.

Köbis, N. and L. D. Mossink (2021). Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in Human Behavior 114*, 106553.

König-Kersting, C., M. Pollmann, J. Potters, and S. T. Trautmann (2021). Good decision vs. good results: Outcome bias in the evaluation of financial agents. *Theory and Decision 90*, 31–61.

Kruger, J. and D. Dunning (1999, December). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology 77*(6), 1121–1134.

Li, Y. and L. Yang (2013). Prospect theory, the disposition effect, and asset prices. *Journal of Financial Economics 107*(3), 715–739.

Logg, J. M., J. A. Minson, and D. A. Moore (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes 151*, 90–103.

Lourenço, C. J., B. G. Dellaert, and B. Donkers (2020). Whose algorithm says so: The relationships between type of firm, perceptions of trust and expertise, and the acceptance of financial robo-advice. *Journal of Interactive Marketing 49*, 107–124.

Lusardi, A. and O. S. Mitchell (2007). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics 54*(1), 205–224.

Lusardi, A. and O. S. Mitchell (2011). Financial literacy around the world: an overview. NBER Working Paper 17107.

Mahmud, H., A. N. Islam, S. I. Ahmed, and K. Smolander (2022). What influences algorithmic decision-making? a systematic literature review on algorithm aversion. *Technological Forecasting and Social Change 175*, 121390.

Major, L. and J. Shah (2020). *What to Expect when You're Expecting Robots: The Future of Human-robot Collaboration*. Hachette UK.

Morewedge, C. K. (2022). Preference for human, not algorithm aversion. *Trends in Cognitive Sciences 26*(10), 824–826.

Niszczota, P. and D. Kaszás (2020). Robo-investment aversion. *Plos one 15*(9), e0239277.

Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of Finance 53*(5), 1775–1798.

Oehler, A., M. Horn, and S. Wendt (2022). Investor characteristics and their impact on the decision to use a robo-advisor. *Journal of Financial Services Research 62*(1), 91–125.

Pleßner, M. (2017). The disposition effect: a survey. *Management Review Quarterly 67*(1), 1–30.

Ploner, M. (2017). Hold on to it? An experimental analysis of the disposition effect. *Judgment & Decision Making 12*(2), 118–127.

Prahl, A. and L. Van Swol (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting 36*(6), 691–702.

Rau, H. A. (2015). The disposition effect in team investment decisions: Experimental evidence. *Journal of Banking & Finance 61*, 272–282.

Rossi, A. G. and S. P. Utkus (2020). Who benefits from robo-advising? Evidence from machine learning. *SSRN Working Paper 3552671*. https://ssrn.com/abstract=3552671.

Royer, H., M. Stehr, and J. Sydnor (2015). Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a Fortune-500 company. *American Economic Journal: Applied Economics 7*(3), 51–84.

Rühr, A. (2020, June). Robo-Advisor Configuration: An Investigation of User Preferences and the Performance-Control Dilemma. *ECIS 2020 Research Papers*.

Sharan, N. N. and D. M. Romano (2020). The effects of personality and locus of control on trust in humans versus artificial intelligence. *Heliyon 6*(8), e04572.

Shefrin, H. and M. Statman (1985). The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence. *The Journal of Finance 40*(3), 777–790.

Strathman, A., F. Gleicher, D. S. Boninger, and C. S. Edwards (1994). The consideration of future consequences: weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology 66*(4), 742–752.

Thaler, R. and C. R. Sunstein (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Tse, T., N. Hanaki, and B. Mao (2022). Beware the performance of an algorithm before relying on it: Evidence from a stock price forecasting experiment. *ISER Discussion Paper 1194*.

Uhl, M. W. and P. Rohner (2018). Robo-advisors versus traditional investment advisors: An unequal game. *The Journal of Wealth Management 21*(1), 44–50.

Vaarmets, T., K. Liivamägi, and T. Talpsepp (2019). How does learning and education help to overcome the disposition effect? *Review of Finance 23*(4), 801–830.

Yaniv, I. and E. Kleinberger (2000, November). Advice Taking in Decision Making: Egocentric Discounting and Reputation Formation. *Organizational Behavior and Human Decision Processes 83*(2), 260–281.

# A Additional data

Table A.10: PARTICIPANTS' CHARACTERISTICS

| | MEAN | SD | RANGE | OBS |
|---|---|---|---|---|
| Age | 24.95 | 4.03 | 19-49 | 409 |
| Male | 0.46 | | 0-1 | 409 |
| CRT | 2.01 | 1.08 | 0-3 | 409 |
| Future Attitude (CFC) | 36.69 | 4.44 | 23-46 | 409 |
| Financial Literacy | 2.39 | 0.72 | 0-3 | 409 |
| Locus of Control | 8.21 | 1.74 | 3-12 | 409 |
| Risk Aversion | 2.62 | 0.82 | 1-4 | 409 |

Table A.11: DESCRIPTION OF STATISTICS ON TRADING ACTIVITY

| | RANGE | DESCRIPTION |
|---|---|---|
| Disposition effect (Diff) | -1 to 1 | average PGR - average PLR, see formula 2 |
| Disposition effect (amount) | -1 to 1 | average PGR - average PLR, weighted by the value of the gains and losses, see formula 3. |
| Optscore | -1 to 1 | Weighted sum of trading decisions made according to Bayesian updating (+1) and against it (-1). |
| Stocks traded | 0-3 | Number of stocks bought or sold in a period |
| Overrides | 0-1 | % of algorithmic recommendations overriden (in Soft treatments) |
| Adoption | 0-1 | Equal to 1 if the individual adopted the algorithm in the third week. |
| Own success | 0-1 | % of trading decisions that were successful(1) ex-post |
| Algo Success | 0-1 | % of tradings decisions or recommendations by the algorithm that were successful(1) ex-post |

(1) Success is defined as buying or holding a stock that goes up in the next period, or selling or not buying a stock that goes down in the next period

Table A.12: STATED REASONS FOR ALGORITHM ADOPTION / NO ADOPTION

| | REASONS TO ADOPT | % | | REASONS NOT TO ADOPT | % |
|---|---|---|---|---|---|
| 1a | Improves my performance | 48 | 1b | Reduces my performance | 25 |
| 2a | Trades should not do | 5 | 2b | Knows when to trade | 6 |
| 3a | Frees my time | 14 | 3b | Has fun trading | 18 |
| 4a | Dislikes trading | 1 | 4b | Likes to keep control | 40 |
| 5a | Frees my mind | 17 | 5b | | |
| 6a | Does not know | 3 | 6b | Does not know | 4 |
| 7a | Other reasons | 12 | | Other reasons | 7 |
| | | N=145 | | | N=250 |

This table reports the share of each stated reasons gave by participants in the exit questionnaire. In particular, participants answer the following questions depending on adoption. If they had adopted the algorithm, then we asked: "Why did you select an algorithm"? Possible answers were 1a) I believe the algorithm improves my performance; 2a) There are trades that I should not do; 3a) It frees my time 4a) I do not like to make transactions 5a) It frees my mind 6a) I do not know 7a) other reasons. If they had not adopted the algorithm then we asked: "Why did not you select an algorithm"? Possible answers were 1b) I believe the algorithm reduces my performance; 2b) I know when to do trades; 3b) I have fun trading 4b) I like to keep control 5b) I do not know 6b) other reasons. Only one reason could be given.

# B Instructions

Welcome! The experimental session today will last about 30 minutes and you will receive 5 Euro for your participation. You can then participate in a second phase of the experiment that will last 21 days. You will be able to earn additional money depending on the choices you will make during these 21 days.

Please read those instructions carefully. This first session will take place in this virtual room (including the demo). At the end of this session, the real experiment will began, taking place on your device and lasting 21 days.

**IMPORTANT**: We remind you that your participation will remain anonymous to the other participants as well as to the experimenters. You will receive an identification number, automatically assigned by the computer, which will be used for payments.

## Description of the game

In this experiment you will be given 350 ECU to invest in three different stocks. One ECU corresponds to 0.04 Euro (that is 50 ECU= 2 Euro).

Your job is to choose when to buy and sell each stock, so that you earn the most money by the end of the experiment. Throughout the experiment, you will see the price of each stock changing (more detail below), and you will use this information to decide when to buy and sell. When you sell a stock, you receive an amount of cash equal to the price of the stock. When you buy a stock, you receive one unit of the stock, but you must give up an amount of cash equal to the current price of the stock.

The three stocks you can buy or sell are simply called Stock A, Stock B, and Stock

At the beginning of the experiment each one of three stocks will be automatically assigned to you and each one costs 100 ECU. Therefore, at the beginning of the experiment you will have the following situation:

| **Stock** | **Quantity** | **Current Price** | ECU Value | Euro Value |
|---|---|---|---|---|
| A | 1 | 100 | 100 | 4 |
| B | 1 | 100 | 100 | 4 |
| C | 1 | 100 | 100 | 4 |
| **Cash** | | 50 | 50 | 2 |
| Total value | | | 350 | 14 |

For the entire duration of the experiment, you can hold **one unit at most of each stock.** You cannot hold negative quantity (that is you cannot sell stocks that are not at your disposal). Nevertheless, you might have a negative amount of cash. That will happen should you buy a stock at a price that is higher than the amount of cash you have at the moment of the purchase. This negative amount will be subtracted from your earnings at the end of the experiment.

## Structure of the market

In this experiment, every day you will be able to buy and sell stocks in three different time window that we call "market sessions":

1. **Morning session**: from 4:00 a.m. to 12:00 p.m;
2. **Afternoon session**: from 12:00 p.m. to 8:00 p.m;
3. **Night session**: from 8:00 p.m. to 4:00 a.m.

In particular, during each market sessions,

- the **price of each stock will be updated** and you will be informed whether the price increased or decreased, and of which amount;
- at the new price you will have the possibility to sell each stock (should you hold it) or buy it (should you not).

You will be able to make your choice at any moment during the opening of the market sessions but you can not make a choice once the session is closed.

## Repetition over 21 days

The game will be repeated with each three market sessions over 21 days. Small variations will be introduced after 7 and 14 days. Those will be notified directly on your screen as well as by email (see further below " **Earnings"**).

In particular, **at the beginning of the second and of the third week you will receive a notification of the changes that will intervene during each week.** This notification will remain visible on your screen for at least 16 hours. Only once this time expires will you be able to play again. You will receive a reminder by email as well.

## How stock prices change

Each stock changes price according to the exact same rule. Each stock is either in a **good state** or in a **bad state**. In the **good state**, the stock goes up with 70% chance, and it goes down with 30% chance. In the bad state, the stock goes down with 70% chance and it goes up with 30% chance.

Once it is determined whether the price will go up or down, the *size* of the change is random, and is either be ECU 5, ECU 10, or ECU 15. For example, in the bad state, the stock will go down with 70% chance, and the amount it goes down by is ECU 5, ECU 10, or ECU 15 with equal chances. Similarly, in the good state, the stock will go up with 70% chance, and the amount it goes up by will either be ECU 5, ECU 10, or ECU15.

The stocks will all randomly start in either the good state or the bad state, and after each price update, there is a 20% chance the stock switches state.

The tables below summarise these information

**Price changes**

| | *Good state* | *Bad state* |
|---|---|---|
| **+ (UP)** | 70% | 30% |
| **- (DOWN)** | 30% | 70% |

**State changes**

| | *Good state today* | *Bad state today* |
|---|---|---|
| *Good state tomorrow* | 80% | 20% |
| *Bad state tomorrow* | 20% | 80% |

## Your earnings

You will play this game 21 days in total, divided into three phases of each 7 days each. In particular, at the beginning of each new phase (i.e. after 7 and 14 days) you will be able to buy again each stock at 100 ECU and the state of each stocks will be restarted, i.e. randomly drawn again as at the beginning of the experiment.

You earnings will be restarted as well at the beginning of each new phase (i.e. after 7 and 14 days) and will be computed for each phase at the end of the experiment. More precisely, the earnings corresponding to each phase will be equal to the amount of cash you accrued over the two scanning sessions from buying and selling stocks, plus the current price of any stocks that you own.

*Earnings=cash + price A*(Hold A) + Price B*(Hold B) + Price C*(Hold C)*

Finally, one phase of 7 days out of three will be randomly selected for payment (i.e. you will be paid according to the total earning of a randomly selected week).

**Your final earnings will be converted in Euro at the exchange rate of 1 ECU=0.04 Euro.**

For payment you will have two options:

1. by IBAN

2. in cash at the Department of Economics (but only if compatible with the current health norms)

In any case, you will have to send an email to **caterina.giannetti@gmail.com**