Computer algorithms are less permitted to invest in controversial stocks than humans: experimental evidence on robo-fund aversion

[Extended abstract]

Paweł Niszczota¹, Dániel Kaszás²

¹ Poznań University of Economics and Business; Department of International Finance
² ETH Zürich; Chair of Cognitive Science

Keywords: quantitative funds, robo advice, algorithm aversion, sin stocks, social norms

Correspondence to: Paweł Niszczota, pawel.niszczota@ue.poznan.pl. This project has been supported by grant 2018/02/X/HS4/01703 from the National Science Centre, Poland.

1. Introduction

The growth of algorithmic (systematic) investment funds (Harvey et al., 2017) and the robo-advisory industry (D’Acunto et al., 2019) means that more and more investment decisions are made not by humans, but by computer algorithms. What this implies is that we already delegate to computers (machines) a controversial decision: how much they should invest in morally controversial companies. Delegating this decision to machines means that they have (or will have) the power to autonomously exclude controversial companies (e.g., companies that produce tobacco or firearms) from portfolios or invest more heavily in them (e.g., if there was a financial advantage of doing so, as demonstrated by Hong and Kacperczyk (2009)). In this study we investigate whether people find that it is permissible for computer algorithms (“robo” fund managers) to do so.

Recent evidence suggests that individuals should show an aversion towards computer algorithms, especially if it concerns decisions from the moral domain. When given the possibility to choose between advice provided by a human or an algorithm, people show a preference for the former and thus demonstrate algorithm aversion (Castelo et al., 2019; Dietvorst et al., 2015; Longoni et al., 2019). Bigman and Gray (2018) have demonstrated via a series of experiments that decisions in the domain of morality appear to be the domain of humans and not autonomous computer algorithms (or machines operated by these algorithms), showing an algorithm aversion in the moral domain. Other studies also suggest that people do not generally feel that machines have the competence to make moral decisions (Gogoll & Uhl, 2018). This led us to hypothesize that there is an aversion towards computer algorithms making investment decisions in morally controversial companies.

While the study does not deal with investment mechanisms currently available to the average investor, progress in the robo-advisory industry suggests that more and more autonomy will be given to machines, with perhaps some of the funds being managed (almost) fully autonomously. Thus, this is a subject worthy of investigation, somewhat similar to the case of autonomous vehicles: even though they are still in a phase that requires human-assistance (many years from the full automation of road vehicles), researchers find that it is important to know what the autonomous vehicles should do when facing a moral dilemma (Bonnefon et al., 2016).

2. Methodology

We conducted three pre-registered experiments, in which our main purpose was to test whether people exhibit an aversion towards computer algorithms making investment decisions for them. To test this hypothesis, we used a between-subjects design, where participants were assigned to a condition where the fund manager was either a human or a computer algorithm. While all of the experiments were meant to test the main hypothesis, each experiment was meant to also investigate the conditions that could affect the strength of the proposed effect.
In Experiment 1 (N = 466), we investigated whether there is an aversion towards computer algorithms making investment decisions (“robo-fund aversion”) by investigating the right of a fund manager to exclude stocks from 14 controversial industries from the funds’ investment portfolio (Trinks & Scholtens, 2017). In this experiment we also wanted to assess whether permissibility ratings might be different if participants knew that excluding stocks might lead to a weaker performance of the portfolio. This is consistent with research from the financial literature, showing that there is a performance benefit from investing in controversial stocks (“sin stocks”; Hong & Kacperczyk, 2009). If this was the case, then the exclusion of some controversial stocks from the portfolio might hurt portfolio performance (returns).

In Experiment 2 (N = 1231), we tested whether robo-fund aversion is present both when contemplating the right to exclude companies from controversial industries and uncontroversial industries. In Experiment 3 (N = 683), we further investigated how generalizable robo-fund aversion is by simultaneously looking at the permissibility to exclude (just as in Experiment 1 and 2) but also invest more heavily in controversial companies.

We recruited all participants via Mechanical Turk. Participants had to imagine that a fund manager or a computer algorithm (depending on which condition they were assigned to) has the autonomy to select companies to stock portfolios for moral reasons, i.e. has the autonomy to exclude companies from a selection of controversial industries (Studies 1-3) or invest more heavily in them (Study 3). These were presented in the first part of the study, for participants to have a reference point. Participants then rated – on a scale of 1 (strongly disagree) to 5 (strongly agree) – three statements adapted from Bigman and Gray (2018):

1. It is appropriate for a [fund manager/computer algorithm] to make these decisions.
2. A [fund manager/computer algorithm] should be the one to make these decisions.
3. A [fund manager/computer algorithm] should be forbidden from making these decisions.

The dependent variable – the permissibility score – was computed as the mean of the three items (after recoding the third item). To test the hypotheses, we performed a series of ANOVAs.

3. Results

In all of our studies a robo-fund aversion emerged (see Figure 1): participants rated the permissibility of computer algorithms making exclusions (Studies 1-3) and heavier investment decisions (Study 3) to be lower compared to a human fund manager. The size of the effects was considerable: after pooling the data (as shown in Figure 1) the effects were d = 0.25 [95% CI: 0.07, 0.48] in Study 1, d = 0.58 [95% CI: 0.47, 0.69] in Study 2, and d = 0.81 [95% CI: 0.70, 0.92] in Study 3. Altogether, the size of the effects should be interpreted as large (Funder & Ozer, 2019) and economically meaningful.

In Study 1, robo aversion was not different for participants that were informed that excluding controversial stocks might have an adverse effect on returns. In Study 2, robo aversion was present both when fund managers had the autonomy to exclude controversial stocks and uncontroversial stocks, and was, surprisingly, higher in the latter case. In Study 3 robo aversion manifested itself both when the fund manager had the autonomy to exclude and invest more heavily in controversial stocks.

4. Conclusions

In three experiments (N = 2380) we show that people find it more permissible when allocations in controversial industries are made by human fund managers as opposed to when they are made by computer algorithms (“robos”). While this finding is consistent with the literature on algorithm aversion, it should not be treated as an obvious one, considering that researchers have shown algorithm appreciation under certain conditions (Logg et al., 2019), and were unable to show algorithm aversion in financial investment (Germann & Merkle, 2019).

Our results do not suggest that robos might be used as a “laundering machine” of morals, exploiting the potential of machines (“robo” fund managers or advisors) to shield from responsibility (Bigman et al., 2019) and omit unfavorable human judgments (Bursztyn & Jensen, 2017; Cohn, Gesche, & Maréchal, 2018).
Figure 1. Permissibility to exclude or invest more heavily in stocks

Note: Horizontal bars show means.

References


