DISS. ETH Nr. 27962

We're Living in a Society: Four Studies on Social Information and Decisions Under Uncertainty

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

DÁNIEL KASZÁS,
MSc, University of Zurich
born on 18.09.1990
citizen of Hungary

Accepted on the recommendation of:
Prof. Dr. Christoph Hölscher, Examiner
Prof. Dr. Stefan Zeisberger, Co-examiner
Prof. Dr. Jörg Rieskamp, Co-examiner

2021

Table of Contents

Abstract

In four studies, I investigate how social information shapes people's behavior and attitudes in situations characterized by uncertainty. The first two studies of this dissertation examine how information on comparable individuals' performance and attributes can influence behavior in a retail investment setting. In the first contribution we test the effect an upward social comparison manipulation has on the trading behavior of retail investors in an experimental setting. We document increased risk taking and trading activity, as well as lower satisfaction with one's own performance. In the second study of this dissertation, instead of providing information on how others have performed, overplacement is induced regarding one's level of knowledge of finance and investments. We find that while our experimental manipulation successfully generates overplacement, it does not lead to detectable changes in trading behavior. The third and fourth study of this dissertation investigate the importance of retaining some level of human involvement when it comes to attitudes towards automated investing services and autonomous vehicles. This dissertation provides novel insights into the role social influences play in shaping decisions in various contexts, from trading decisions to technology acceptance.

Zusammenfassung

In vier Studien untersuche ich, wie soziale Informationen das Verhalten und die Einstellungen von Menschen in von Unsicherheit geprägten Situationen beeinflussen. Die ersten beiden Studien dieser Dissertation untersuchen, wie Informationen über die Leistung und Eigenschaften anderer, vergleichbarer Personen das Verhalten von Einzelinvestoren / Kleinanleger beeinflussen können. Im ersten Beitrag wird in einem experimentellen Setting getestet, welchen Effekt eine Manipulation des sozialen Aufwärtsvergleichs auf das Handelsverhalten von Privatanlegern hat. Wir dokumentieren eine erhöhte Risikobereitschaft und Handelsaktivität sowie eine geringere Zufriedenheit mit der eigenen Leistung. In der zweiten Studie dieser Dissertation wird, anstatt über die Leistung anderer zu informieren, eine "Overplacement" in Bezug auf den Kenntnisstand über Finanzen und Investitionen induziert. Wir stellen fest, dass unsere experimentelle Manipulation zwar erfolgreich Overplacement erzeugt, jedoch nicht zu nachweisbaren Veränderungen im Handelsverhalten führt. Die dritte und vierte Studie dieser Dissertation untersuchen, wie wichtig es ist, ein gewisses Mass an menschlichem Engagement bei der automatisierten Anlagedienstleistungen und autonomen Fahrzeugen beizubehalten. Diese Dissertation bietet neue Einblicke in die Rolle sozialer Einflüsse bei der Gestaltung von Entscheidungen in verschiedenen Kontexten, von Handelsentscheidungen bis hin zur Technologieakzeptanz.

Introduction

Since the end of World War II, the question of how social relationships and the real or implied presence of others influence human behavior has led to a proliferation of empirical research in psychology (Cartwright, 1979; Richard, Bond & Stokes-Zoota, 2003; Allport, 1954). In attempting to understand why people behave the way they do, one must always be mindful of the fact that we are ultimately social animals (Aronson & Aronson, 2018; Tomasello, 2014), who over the course of evolutionary history have adapted to survive in tight-knit communities (Dunbar, 1998), integrated in not just a shared physical, but also a shared social environment (Cartwright, 1979). We are deeply embedded in relationships which ultimately form the basis of our social world, and our shared social reality (Searle, 2010; Dunbar, 1998). As Rochat describes it, having "others in mind", and integrating information from our social world is key to the development of self-consciousness (Rochat, 2003). The foundation of our social world lies in social information. Social information can be defined as "knowledge of other people's behaviors, attributes, intentions and preferences" (Hertwig & Herzog, 2009, p. 682). Drawing on social information to make decisions can be an efficient and easy way to aid decisions under resource constraints, but such information can also be less reliable than first-hand, personal information (Hertwig & Hoffrage, 2013). As our world is becoming increasingly complex and interconnected (Bak-Coleman et al., 2021; Backstrom, Boldi, Rosa, Ugander & Vigna, 2012), information on other people's actions, beliefs, attributes etc. is becoming increasingly available. Heuristics such as "imitate the majority" and the "equity heuristic" (Hertwig & Herzog, 2009) might provide a way to make quick decisions when resources are scarce, but can also predispose individuals towards herding on financial markets (Baddeley, 2010), and result in unbalanced investment portfolios due to "naïve diversification" (Benartzi & Thaler, 2007).

The role our expanding social world plays in influencing behavior specifically in the context of financial decision-making will be the focus of the first two studies of this dissertation.

While many domains of human life have been touched by the increasing social interconnectedness and increasing speed and reach of telecommunication, the world of retail investing (where individuals invest their own capital for their own benefit, Mackintosh, 2020) has been transformed substantially over the past years. With the increasing democratization of access to financial investments and financial information, anyone can make highly speculative investments based on information provided by other individual investors. There are already a number of social trading platforms which base their business models on providing such peer information (StockTwits, eToro, Wikifolio etc.). Social media platforms, combined with increased ease of access to the financial markets (Egkolfopoulou, Massa & Melin, 2021) also make it possible for self-organized communities to affect the share price of publicly traded companies in a substantial way (Long, Lucey & Yarovaya, 2021). Given what we know of the effect of peer influence on individuals' financial decisions (Bursztyn & Ederer, 2014; Kaustia & Knüpfer, 2012; Agarwal, Mikhed & Scholnick, 2019; Kuhn, Kooreman, Soetevent & Kapteyn, 2011), a further examination of social influences in financial decision-making among retail investors is well warranted. These recent developments serve as the backdrop for the first two studies of this dissertation.

The technological revolution which has taken place over the past decades has on the one hand helped connect individuals on a heretofore unprecedented level, enabling previously unconnected people to have considerable influence over each other's' lives. On the other hand, the increases in computing power which have partially enabled this development are also contributing to a parallel trend of increasing automation, in domains and decision scenarios which have until recently been exclusive to human participation (Bak-Coleman et al., 2021). As automation is becoming ever more prevalent in domains such as investing and mobility (Martinez-Diaz & Soriguera, 2018; Puschmann, 2017), a thorough investigation of how individuals perceive the loss of human presence in such situations is overdue. Ultimately, as our social world is expanding in many ways due to advances in information technology, it is also contracting, with human interaction and human presence being a lesser part of one's overall experience in a number of ways. People's attitudes towards removing the human element from decision-making, and their readiness to use such automated services provides an interesting and relevant counterpoint to the increasingly important role of social information and social influence. The second half of this dissertation will concentrate on how people perceive the potential loss of human involvement, specifically when it comes to retail finance and mobility – two domains of human decision-making where the transition towards higher automation is already well underway.

The overarching goal of this dissertation is to investigate how our social world impacts us in situations characterized by uncertainty. The majority of the studies are situated in the context of retail finance, and provide insights regarding how individual investors make decisions in today's investment landscape. As information on social others (real or imagined) might impact the investment behaviors of the segment of investors deemed most vulnerable in the marketplace, understanding the potential effects of social information (or lack thereof) in a retail setting is an important contribution. However, as technologies which automate jobs previously predicated on human interaction are becoming widespread, it is also important to investigate if people are ready to get rid of the "human in the loop" – in the domain of retail investing and beyond.

In the following section, the first two studies of this dissertation are described, with a shared focus on how social information affects the behavior and beliefs of retail investors in an experimental setting. This is then followed by a section describing two studies investigating individual behavior and attitudes in the context of automation, and a potential absence of social others. These studies are all connected in that they investigate the effects of our social world on how we make decisions under uncertainty. The last section of this chapter contains a critical reflection and discussion of the four studies, and an attempt to provide a preliminary conclusion.

Social comparison and biased self-assessment in the evaluation of abilities

Study 1: I'll Have What They're Having – The Influence of Upward Social Comparison on Trading Behavior

According to Allport's widely accepted definition of social psychology, it is "the attempt to understand and explain how the thoughts, feelings, and behaviors of individuals are influenced by the actual, imagined, or implied presence of other human beings" (Allport, 1954, p.5). This is a very broad definition, and accordingly, many aspects of human behavior have fallen under the scope of social psychological inquiry. One such phenomenon is social comparison, "the fundamental human tendency to look to others for information about how to think, feel and behave (…)" (Baldwin & Mussweiler, 2018, p.1.). Wood (1996) defines social comparison as "the process of thinking about information about one or more people in relation to the self" (p. 520) It is considered to be one of the most ubiquitous features of human social life (Buunk & Gibbons, 2007; Baldwin & Mussweiler, 2018), enabling us to function in an increasingly complex and interconnected world. Festinger's theory of social comparison process (1954) has laid the foundation for decades of work on the subject (Goethals, 1986; Gerber, Wheeler & Suls, 2018). In his formulation (1954) all humans share a drive to evaluate themselves, and the outcome of such evaluations informs later behavior. In some contexts, one's evaluation of their own ability can be calibrated based on objective benchmarks (an example being athletic skills such as running, where speed and distance covered are objectively measurable). As Festinger notes, most real-world situations do not provide such a clear benchmark, and as such, subjective evaluations of abilities and opinions are often unstable. When information is available concerning the abilities or performance of others, this relative information serves as the basis of evaluation – furthermore, people are more likely to compare themselves to someone who is the most similar to themselves. Among other corollaries, Festinger also theorizes that the existence of discrepancies between one's own self-assessed ability and that of others can lead to behavior change – this is especially true if one comes off as inferior to the target of the comparison. Underlying these social comparisons is a "unidirectional drive upward" (p. 124) when it comes to the evaluation of abilities – meaning that people have a tendency to compare themselves with others who are slightly better off, according to Festinger.

Festinger's pioneering work has only slowly gained influence in the decades following his initial publication (Goethals, 1986), but it is now considered one of the most consequential theories in the fields of social psychology and social cognition. Two important directions of research explored the determinants of social comparisons (Buunk & Gibbons, 2007), and expanded on Festinger's proposed "unidirectional drive upward". Schachter's fear-affiliation theory proposed that feeling threatened (by for example an anticipated low-level electric shock in one experiment, Schachter 1959) triggers social comparison, as individuals are looking for as much information as possible in an uncertain situation. Another expansion of the original theory concerns factors which trigger a downward comparison – either when one feels threatened (Thornton & Arrowood, 1966), or when one is making downward comparisons in order to generate positive affect (Gibbons, 1986). These expansions hint at a more comprehensive view of social comparison, which does not only serve self-evaluation, but also functions as a means of motivated self-enhancement (Wood, 1989; Diel, Grelle & Wilhelm, 2021).

It is clear that social comparison is an important but complex phenomenon, the boundary conditions of which are still being investigated, almost 70 years after Festinger first advanced this concept (Gerber, Wheeler & Suls, 2018). While interest in social comparison has waxed and waned over the years (Goethals, 1986), presently there is a proliferation of research activity on this topic, mostly driven by two factors: the increasing prevalence of social media (resulting in a stream of research of a more descriptive nature), and the increasing use of social comparison nudges to affect behavior change (more often conducted in experimental settings). In their comprehensive overview Buunk and Gibbons (2007) propose that future research should investigate judicious use of social comparison manipulations to shift behavior in a desired direction. Social comparison interventions have been implemented in areas as diverse as encouraging changes in physician prescribing behavior (Linder et al, 2017; Andereck et al., 2019), promoting environmentally conscious behaviors (Myers & Souza, 2020; Allcott & Kessler, 2009) and charitable giving (Bartke, Friedl, Gelhaar & Reh, 2017, Frey & Meier, 2004), among many others. While it has been found that similar interventions can also backfire (Bicchieri & Dimant, 2019), it seems overall that tapping into people's tendency to respond to social information has the potential to steer behavior in a number of domains.

One domain where social comparisons are especially prevalent and highly relevant is finance and investing. As Kirchler, Lindner and Weitzel note, "in the finance industry, rankings, ratings, and awards are the visible hallmarks of a strong culture of relative performance measurement and social competition." (2018, p. 2271). The most illustrative example is the mutual fund industry: A fund's performance relative to benchmark is a central determinant of mutual fund inflows (Sensoy, 2009), and has been also shown to influence subsequent investment decisions by fund managers (Taylor, 2003; Brown, Marlow & Starks, 1993). Social comparison processes are inextricably linked to a number of behavioral phenomena on financial markets, from the development of bubbles (Smith, Suchanek & Williams, 1988), through herding behavior (Cuthbertson, Nitzsche & O'Sullivan, 2016) to excessive risk taking (Dijk, Holmen & Kirchler, 2014; Schoenberg & Haruvy, 2012). As noted at the beginning of the first section, the world is becoming more interconnected, giving individuals access to unprecedented amounts of information without geographical limitations – the world of finance is no exception. Due to the ongoing democratization of investing on financial markets and the proliferation of social networks geared towards exchanging information on trading, investigating the role of upward social comparison in a modern investment setting, and how it might affect the behavior of market participants is overdue. The first study in this dissertation will examine the question of how being informed about the trading success of comparable individuals (retail investors recruited for the experiment from the same online labor market) affects one's trading behavior. This contribution places special emphasis on the relationship between upwards social comparison and trading activity. Trading activity is a facet of trading behavior which until now has not received attention from researchers investigating social comparison processes. A better understanding of how upward social comparison affects risk taking and trading activity in a realistic setting would contribute greatly to our understanding of how retail investors behave on financial markets, and could provide clues as to potential policy steps to support the financial welfare of individual investors.

Abstract: I'll Have What They're Having – The Influence of Upward Social Comparison on Trading Behavior

Authors: Dániel Kaszás, Sandra Andraszewicz, Stefan Zeisberger & Christoph Hölscher

In a study with 807 participants with previous investing experience, we investigated the impact of upward social comparison on risk taking and trading activity. We implemented two trading rounds using a novel experimental trading platform that displays historical market prices as they develop over time, and allows users to make trades based on this price information. Participants were randomly allocated to one of two conditions: they either received feedback on only their own performance after the first round of trading, or received additional information on how three successful prior participants have performed. We find that participants presented with an upward social comparison took more risk and traded more actively in a subsequent round of trading than participants not exposed to an upward social comparison. We also document that participants exposed to an upward social comparison report significantly lower post-task satisfaction with their performance. These findings extend current research by demonstrating the impact of peer performance on trading activity, risk taking and investor satisfaction.

Study 2: Slicker Than Your Average – Relative Overconfidence and Trading Behavior

If one's skills and abilities are not easy to measure objectively, social comparisons are made to obtain information about oneself in relation to the world (Festinger, 1954). The harder it is to observe a matter objectively, the more we rely on social comparisons. However, in the process of self-evaluation of abilities, even when relatively accurate information is available, people still have a hard time correctly evaluating themselves. Examples abound in everyday life – that 8% of a representative sample of US adults believe they can beat a gorilla in a fistfight speaks to the difficulty some of us have in accurately evaluating our own level of skill (YouGov, 2021).

While making accurate self-assessments about one's abilities, knowledge, character and other attributes is central to leading a successful life, not only do people have the tendency to make inaccurate self-assessments, these assessments are often biased in systematic ways (Dunning, Heath & Suls, 2004). Hansford and Hattie (1982) conducted a meta-analysis to survey 1136 correlations between self-reported and actual performance measures published in previous literature. They report a weak mean correlation of r= 0.21, with individual coefficients ranging between -0.77 and 0.96. On a similar note, in a comprehensive meta-analysis of previous literature on the "better-than-average effect" (BTA), Zell, Strickhouser, Sedikides and Alicke (2020) report that the BTA effect is robust across 291 independent samples, with an overall large effect size. In terms of systematic errors in self-assessment, previous literature has documented that optimistic biases can provide an explanation for diverse behavioral phenomena: excess entry into entrepreneurship (Camerer & Lovallo, 1999; Rietveld et al., 2013), health behaviors (Bottemanne, Morlaas, Fossati & Schmidt, 2020; Dillard, McCaul & Klein, 2006) and important decisions in both one's private as well as professional life, including saving, investing, marriage and retirement (Puri & Robinson, 2007), Being overly optimistic about oneself is associated with a number of real-world behaviors, with significant consequences in all walks of life. If self-assessment biases detract from accurate self-evaluation (which Festinger has established as an important human drive), why do they persist? One aspect of flawed self-assessments which has received increased research attention due to its pervasive nature and considerable effect on various aspects of life is overconfidence. Prior theories have proposed that both justified confidence and overconfidence can increase motivation and goal-directed behaviors (Benabou & Tirole, 2002; Bi, Dang, Li, Guo & Zhang, 2016), which in turn has been proposed to yield rewards for even overconfident individuals (Kennedy, Anderson & Moore, 2013), and confer adaptive benefits to either one's social group or the individual in some circumstances (Johnson & Fowler, 2011; Bernardo & Welch, 2004). While some findings hint at overconfidence being a motivated bias, which is exhibited in order to further a self-enhancement motive (Anderson, Brion, Moore & Kennedy, 2012), others find that a more likely explanation is that overconfidence is the product of cognitive biases under ambiguous circumstances (Logg, Haran & Moore, 2018). Certain conditions facilitate the formation of overconfident beliefs especially strongly – situations where feedback is nonexistent, sparse, ambiguous or only available with delay are especially likely to lead to beliefs about abilities or performance which are not in line with reality (Malmendier & Taylor, 2015; Arkes, Christensen, Lai & Blumer, 1987; Russo & Shoemaker, 1992).

According to Plous "no problem in judgment and decision making is more prevalent and more potentially catastrophic than overconfidence" (1993, p. 217). It can be defined as the phenomenon of estimating one's skills, abilities or knowledge as better than it actually is in a given context. Alternatively, Moore and Dev (2018, p.1) state that to be overconfident is simply "to be more confident than one deserves to be". In its most widely-accepted characterization (Moore & Healy, 2008; Moore & Healy, 2007), overconfidence has three facets, which have been found to be interrelated, but empirically distinct (Snowberg & Yariv, 2021): overprecision, overestimation and overplacement. Overplacement refers to the phenomenon where individuals judge their skills or performance to be better in relation to others than it actually is (Moore & Healy, 2008; Benoit Dubra & Moore, 2015). Larrick, Burson and Soll (2007) characterize overplacement as "a common bias in social comparison" (p.76). While absolute estimates of skill also often fall back on social comparison, taking advantage of social information to form one's relative judgment is a core element of overplacement – it is not possible to form an assessment of our relative abilities if we do not try to form an assessment of the abilities of others. Relative overconfidence (an umbrella term describing both overplacement as well as the conceptually similar better-than-average effect) has been the subject of considerable research interest over the years.

The element of social comparison and reliance on social information inherent to overplacement makes it an especially interesting concept to investigate, given the previously noted increase in changes to our social world, and the increasing abundance of social information. One context in which one's self-assessment of skill in relation to others is key is on financial markets. While it is easy to compare one's past performance with peers or a benchmark ex-post if one wants to generate an estimate for one's skill or level of information at a certain point in time, it is much harder to assess one's abilities or level of information as an investor on an ongoing basis. Such evaluations inevitably involve comparisons to peers, benchmarks, and they result in various products and assets to be described as "best-in-breed", "beating the market" or "generating alpha". As established previously, overconfidence is most likely to emerge in ambiguous situations where the consequences of one's actions are not immediately observable, the resulting feedback is ambiguous, and there is no clearly defined frame of reference present for interpreting outcomes. It comes then as no surprise that considerable research has been conducted to investigate the effects of overconfidence on financial markets.

A number of previous studies have reported an association between various facets of overconfidence and trading activity, with overconfidence being the most promising mechanism for explaining the high level of trading volume observed on financial markets (Barberis, 2018).

Previous empirical work has demonstrated a robust relationship between overconfident beliefs and trading activity (Odean & Barber, 2001; Glaser & Weber, 2007; Merkle, 2017; Gervais & Odeon, 2001). Further investigations of this relationship in a laboratory setting reveal a significant association between overconfidence and various metrics of trading activity, as well as the development of bubbles on experimental asset markets (Deaves, Lüders & Luo, 2008; Yang & Zhu, 2016; Bregu, 2020). The one facet of overconfidence which has received comparably little attention is overplacement.

In a series of laboratory experiments, Cheng, Anderson, Tenney, Brion, Moore and Logg (2021) demonstrate not only that overplacement can be transmitted between team members collaborating on a task, but that even observing overconfident individuals can exacerbate this bias. This is in line with Hirshleifer's theory of "social finance", which describes investor traits as the product of cultural evolution (Akcay & Hirshleifer, 2021). Just as the spread of popular narratives about economic events can ultimately drive market prices (Shiller, 2020), beliefs and biases such as overconfidence can be transmitted between individuals to ultimately affect large-scale financial outcomes.

A good example for how such transmission might look like concerns the case of "finance influencers" – a recent analysis found that 1 in 7 videos by such individuals contained misleading information, promising guaranteed high returns on investments into individual stocks and cryptocurrencies (Paxful, 2021). Investigating the emergence of a class of younger self-directed investors who follow social media for investment advice, the UK's Financial Conduct Authority (FCA) finds that these individuals have very high confidence in their investing skill, in many cases coupled with the belief that losing money on speculative investments is not a real possibility for them (FCA, 2021). This is also associated with a lower level of financial resilience, with 59% of those surveyed reporting that a significant loss in their investments would have a "fundamental impact on their current or future lifestyle" (FCA, 2021, p.19).

Changes in how people obtain their information on financial investments, the breadth of information one needs to sift through to make investment decisions, and individuals' increasingly easy access to a wide variety of investment options (including social trading apps) can combine to result in potentially catastrophic effects on individual investors' financial welfare. Trading more riskily and more aggressively as one probably should is now is easier, and has larger consequences than ever – understanding the drivers behind trading behavior in a retail setting is a key question for research in behavioral finance and financial decision-making.

Considering that the overplacement aspect of overconfidence directly incorporates an element of social comparison, it is an especially interesting and relevant construct to investigate at this point in time. A better understanding of how important relative overconfidence is to trading behavior is necessary if we want to support policymakers in search of better tools to ensure the financial welfare of individual investors. Study 1 of this dissertation investigated the causal effect of upward social comparison information regarding previous performance on trading behavior, which then leads people to make inferences about their own attributes in relation to other comparable individuals. In Study 2 we go one step further, and endow people with overconfident beliefs about how knowledgeable they are relative to other comparable individuals, and observe if this has an effect on subsequent trading behavior. As such, both studies focus on the influence new information about others (social information) has on behavior on financial markets.

Abstract: Slicker Than Your Average – Relative Overconfidence and Trading Behavior

Author: Dániel Kaszás

Previous empirical research documents a relationship between relative overconfidence and risk taking and trading activity in retail investors. In two pre-registered studies we investigate the relationship between overplacement regarding performance on a finance and investments quiz, and retail investors' trading behavior on a simulated trading platform. In a first, correlational design we find no support for a relationship between

overconfidence when operationalized as overplacement and trading behavior. In a second, experimental study we successfully manipulate active retail investors' self-perceived knowledge of finance and investment in order to induce overplacement. We fail to find support for a causal effect of overplacement on trading behavior in the subsequent trading task. In an additional analysis we document a positive effect of reporting having performed "better-than-average" on risk taking in our trading task, in line with previous literature. This project makes a crucial contribution in that it is the first to experimentally manipulate overplacement and observe its effect on trading behavior in a realistic setting, with knowledgeable retail investors.

Preference for human decision-makers and technology adoption

Study 3: Robo-investment aversion

The first two studies of this dissertation investigate the important effect information on our peers can have on financial decisions. These studies also demonstrate the importance of social information specifically when it comes to retail investing, and highlight potential avenues through which social information can shift one's performance reference point, lead to biased self-assessments, and ultimately affect one's behavior. Parallel to the trend of increasing human interconnectedness in general, recent years have also seen a trend of decreasing levels of interaction with others in a number of contexts. Relatively inexpensive computing power has made it possible to implement algorithms to augment, and in some cases replace human interaction and human decision-making (Bak-Coleman et al., 2021; Deming, 2021). Many job functions which were previously filled by people are being automated away, replacing human decision-making with algorithms (Deming, 2021; Acemoglu & Restrepo, 2019). Reflecting on peer effects on moral decision-making, Köbis, Bonnefon and Rahwan (2021) highlight that just as humans influence each other's behavior, so can machines impact how individuals think, feel and act. As we are on course to replace human interaction with human-machine interactions in a number of key areas of daily life, investigating how people respond to the lack of human presence can prove to be an important contribution to our understanding of the modern world, and where it's headed. One such area where automation has been progressing at a steady pace is the world of finance.

An illustrative example for how automation is taking an increasingly important role in the world of finance is the development of the retail banking experience over the past few decades. With the advent of online banking and the widespread adoption of mobile banking apps (Jansen & Schaik, 2018; Lee, Tsai & Lanting, 2011), the traditional system of physical bank branches staffed by employees, servicing a local customer base has been transformed. In order to save costs, bank branches are being closed all over the developed world (Tranfaglia, 2018; Leyshon, French & Signoretta, 2008), while in some markets mobile banking is really the only reliable way to access traditional financial services already (Alexander, Shi & Solomon, 2017; Jagtiani & Lemieux, 2018). The digitization of retail financial services results in a reduction of personal connections between customers and bank employees – the decline of traditional relationship banking giving way to transactional banking (Jaksic & Marinc, 2019). Thus, in many cases tasks which were previously performed by people are now automated, changing the nature of the customer interaction in the process.

An extension of this dynamic is now taking place with the proliferation of FinTech companies (Kaja, Martino & Pacces, 2021) offering a suite of services which were traditionally performed by specialized financial advisors. Robo-advisors can best be described as software products offered by investment service providers, which automate both the customer assessment-, as well as the portfolio management aspect of advisory services (Jung, Dorner, Glaser & Morana, 2018). They profile prospective customers based on a questionnaire consisting of a mix of self-reported as well as behavioral measures (Tertilt & Scholz, 2018; Tertilt & Scholz, 2021), in order to find the best possible match between an investor's preferences and a recommended investment portfolio. In most cases this portfolio is then managed passively by the company providing robo-advisory services. Substantial cost savings can result from cutting a human financial advisor out of the loop, which drive the widening selection of robo-advisors in various markets (Tertilt & Scholz, 2021). While these are overall attractive products from a consumer perspective (Schabicki, Quint & Schroeder, 2021), take-up has been slower than expected by industry observers, and not without controversy (SEC, 2018). While an analysis by Deloitte expected the assets under management (AuM) of robo-advisory businesses to reach between $2-4 trillion by 2020 (Moulliet, Majonek, Stolzenbach & Völker, 2016), and a 2015 Statista projection estimated about $8 trillion AuM by 2020 (Statista, 2017) current estimates of AuM are well below these expectations. According to recent data from Statista, the amount of financial wealth managed by robo-advisors is only projected to reach $1.4 trillion globally in 2021, with 2/3 of that amount in the US (Statista, 2021). Overall, while robo-advisory services seem to be suitable and affordable for a large fraction of the investing public, their adoption has been significantly lagging behind expectations for a number of years now. A possible contributor to this relatively slow progress might be an underlying hesitancy on the part of individual investors to entrust financial decisions to machines.

How people respond to information provided by an algorithm has become a question of intense research interest over the past few years (Dietvorst, Simmons & Massey, 2015; Yeomans, Shah, Mullainathan & Kleinberg, 2015), and it is likely to be one of the defining issues of the next decade. One rapidly developing field of research concerns the phenomenon of "algorithm aversion" – people's tendency to prefer the involvement of human decision-makers to algorithmically determined decisions, and to prefer human-generated advice to machine-generated advice under certain circumstances, even when the algorithmically generated information is of higher quality (Dietvorst, Simmons & Massey, 2015; Castelo Bos & Lehmann, 2019). This phenomenon has been demonstrated in a number of contexts: from attitudes towards autonomous vehicles and criminal justice (Bigman & Gray, 2018) to forecasts of student performance (Dietvorst, Simmons & Massey, 2015) and medical decisions (Castelo, Bos & Lehmann, 2019; Longoni, Bonezzi & Morewedge, 2019; Bigman & Gray, 2018). The overall picture is that individuals tend to discount advice from algorithms, and are hesitant towards interacting with machines, compared to interacting with humans. This aversion is highly context-dependent (Castelo, Bos & Lehmann, 2019), and is especially pronounced in domains which have an inherent moral component to them (Gogoll & Uhl, 2018).

The world of finance lends itself especially well to investigating the intersection of automation and individual decision-making. The development of modern financial markets can best be understood in the context of automation – a progression of ever more complex processes being automated, resulting in today's rapidly evolving financial landscape (Pardo-Guerra, 2019; Davis, Kumiega & Van Vliet, 2013). By some estimates, about 40% of stock transactions on the US stock market are completed automatically – with most of this volume generated by high-frequency trading funds (Brogaard, Hendershott & Riordan, 2014). As a parallel trend to the spread of automation in financial services, increasing research attention has been given to both the sustainability, as well as morality of financial investments (Hong & Kacperzyk, 2009; Durand Kohn & Limkriangkrei, 2009). Industry observers document accelerating demand for investment products which meet criteria for environmental sustainability (Choi, 2016), with sustainable investment strategies steadily gaining in popularity, and becoming more accessible through diverse investment vehicles and asset classes (GSIA, 2019). The amount of sustainable AuM has been increasing rapidly over the past decade, increasing by 30% between 2016 and 2020 (Fish, Kim & Venkatraman, 2019). A rapidly increasing number of investment options are on offer to cater to individuals' need to invest in companies which either promote positive societal outcomes (Wallis & Klein, 2015), or restrict investment in companies which are considered controversial from an environmental, societal or governance perspective (Renneboog, Ter Horst & Zhang, 2008).

While the world of retail finance provides a very good proving ground for investigating the intersection of the trend towards increasing delegation of decisions to algorithms and increased attention to less tangible moral aspects of investment decisions, very little previous research has been done on algorithm aversion when it comes to financial investments. In the only similar contribution to date, Germann and Merkle (2020) fail to find robust evidence for- or against the presence of algorithm aversion in a laboratory experiment when it comes to financial decisions. In order to fill this gap in the literature, and gain a better understanding of the challenges facing further automation in the financial sector, we have conducted a series of experiments investigating the boundary conditions for the emergence of algorithm aversion. Just like the previous two studies of this thesis, this third study fits in well with the overarching goal of the present dissertation, which is to understand how the presence (or absence) of social others might influence decisions and attitudes in the domain of retail finance and beyond.

Abstract: Robo-investment aversion

Authors: Paweł Niszczota & Dániel Kaszás

In five experiments (N = 3,828), we investigate whether people prefer investment decisions to be made by human investment managers rather than by algorithms ("robos"). In all of the studies we investigate morally controversial companies, as it is plausible that a preference for humans as investment managers becomes exacerbated in areas where machines are less competent, such as morality. In Study 1, participants rated the permissibility of an algorithm to autonomously exclude morally controversial stocks from investment portfolios as lower than if a human fund manager did the same; this finding was not different if participants were informed that such exclusions might be financially disadvantageous for them. In Study 2, we show that this robo-investment aversion manifests itself both when considering investment in controversial and non-controversial industries. In Study 3, our findings show that robo-investment aversion is also present when algorithms are given the autonomy to increase investment in controversial stocks. In Studies 4 and 5, we investigate choices between actual humans and an algorithm. In Study 4 –which was incentivized–participants show no robo-investment aversion, but are significantly less likely to choose machines as investment managers for controversial stocks. In contrast, in Study 5 robo-investment aversion is present, but it is not different across controversial and non-controversial stocks. Overall, our findings show a considerable mean effect size for robo-investment aversion ($d = -0.39$ [–0.45, –0.32]). This suggests that algorithm aversion extends to the financial realm, supporting the existence of a barrier for the adoption of innovative financial technologies (FinTech).

Study 4: A Driver is Missing – Comfort with Varying Levels of Human Supervision in Self-Driving Cars and Associated Factors

Study 3 establishes the importance of involving human advisers in investment decisions from a retail customer perspective, and demonstrates that individuals show a clear preference for involving human decision-makers in the investment process, especially when the potential investment options at hand have a moral valence to them. While retail finance is unquestionably a key area of inquiry when it comes to perceptions of automation and human involvement, there are a number of domains of daily life where one often has to make decisions which ultimately have a moral character. Study 4 broadens the scope of this dissertation beyond the world of retail finance. A number of recent contributions highlight the intrinsic moral character of decisions made in the context of driving (Awad et al. 2018; Awad, Dsouza, Shariff, Rahwan & Bonnefon, 2020; Bigman & Gray, 2018). Similarly to how developments in modern technology enable the creation of new business models and product offerings in financial services, advances in machine learning and automotive engineering have led to major advances towards the development of highly autonomous vehicles (AVs), versions of which are already being tested around the world. An important question which has emerged recently is the degree to which in-person or remote human supervision of automated vehicles should be retained. Regulatory approaches have been proposed by the European Commission to create a unified legal framework for the regulation of artificial intelligence in applied settings (European Commission, 2021), proposing mandatory human oversight of high-risk automated systems. While the guidelines for classifying certain use cases for artificial intelligence as "high-risk" are still not in place, according to one interpretation of this EU proposal, autonomous vehicles might well fall into this category (Wolff, 2021). At the same time, Germany's government legalized the operation of autonomous vehicles across the country if these vehicles are supervised remotely by human operators (Ewing, 2021). Such human-in-the-loop-systems for now represent the highest level of automation possible in most everyday settings (Gronsund & Aanestad, 2020). However, improvements in technology and economic incentives will almost surely drive autonomous vehicles towards less supervision as time goes on (Sparrow & Howard, 2017; Gogoll & Müller, 2020).

In the middle of this tension between the need for retaining human involvement and the drive for increasing automation are potential users of automated vehicles, whose level of trust in automation, and their comfort with using fully autonomous vehicles will ultimately determine how these technologies will be utilized. While a substantial amount of past research has addressed the individual characteristics associated with attitudes towards using self-driving vehicles in general (Haboucha, Ishaq, & Shiftan, 2017; Becker & Axhausen, 2017), no study has investigated the role these characteristics play across varying levels of human supervision.

The fourth study of this dissertation investigates if demographic- and attitudinal variables predict one's attitude towards traveling in a fully autonomous vehicle, and compares the importance of these characteristics when it comes to predicting comfort with being a passenger in a fully autonomous vehicle with varying levels of human supervision. We analyze recent representative survey data collected by the European Commission, and are not only able to make robust conclusions regarding the predictors of comfort when it comes to self-driving vehicles, but also describe a number of novel relationships which might help shed light on some unresolved questions in present literature, and potentially inform future policy efforts.

Understanding what predicts people's attitudes towards AVs is not only relevant to inform future research and efforts to communicate the potential benefits and dangers of autonomous vehicles in a European context. It is also a contribution to the overarching theme of this dissertation: how does the presence or absence of others in various decision scenarios influence one's beliefs, attitudes and actions? The first two studies of this dissertation investigate the implied presence of others and the influence this has on financial behavior – actively providing our participants with social information in order to examine its effect on their actions when it comes to trading. The third study investigates how people respond to the absence of other humans in an investment setting where one can traditionally rely on others in the process of making investment decisions. Study 4 continues this line of inquiry by investigating how gradually removing the human element, or removing human supervision from the "loop" is perceived by potential future occupants of fully autonomous vehicles.

Abstract: A Driver is Missing – Comfort with Varying Levels of Human Supervision in Self-Driving Cars and Associated Factors

Authors: Dániel Kaszás & Adam Charles Roberts

While numerous studies have investigated attitudes towards self-driving cars in general, less research attention has been focused on individuals' attitudes towards the presence (or absence) of third-party human supervision, and its potential correlates. In the present study we analyze data from a large-scale European survey, and find considerable heterogeneity in both individual attitudes, as well as in country-level attitudes in our

descriptive analysis. We find a trend of decreasing comfort as external human supervision is reduced, although this effect is more pronounced in some countries than others. We then investigate potential drivers of self-reported comfort with varying levels of external human supervision in a regression framework. Gender differences get stronger with decreasing supervision, suggesting a possible resolution to conflicting evidence in previous studies. Following this, we fit an ordinal random forest model to derive variable importance metrics, which enable us to compare the changing role predictor variables might play in shaping self-reported comfort, depending on varying levels of third-party supervision. Data privacy is highlighted as an important variable, regardless of level of supervision. While our findings are largely in line with previously published literature, we also uncover a number of novel associations, providing guidance for future policy-making and research efforts.

Discussion

This dissertation consists of four independent but thematically connected studies. The through-line of all studies is that they attempt to investigate how our relation to the social world influences our attitudes and behavior in today's fast-changing environment. The first two studies focus on specific aspects of this question in a tightly defined experimental setting, with a specific sample – retail investors on an experimental market who are provided information on their competitors. The focus of these studies is on how social information affects one's behavior in a market setting. These studies are especially informative for understanding how recent changes in the investment landscape might affect retail investor behavior. The third study broadens the scope of inquiry and investigates the boundary conditions which might shape attitudes towards new retail investing tools, while keeping the focus on the role the social world plays in our behavior. The fourth study builds on these insights, and extends the thematic scope even further

– exploring the potential drivers for differences in attitudes towards automation, and how they interact with the presence or absence of others. Thus, studies 3 and 4 offer insights into how users might interact with automated services which might completely lack an element of human interaction, across different domains.

Discussion – Studies 1 and 2

Study 1 reports that when individuals are informed about the success of their peers, they adjust their beliefs and behaviors in predictable ways, to become more risk-seeking and more active traders on an experimental market. We find that in support of previous empirical literature and our pre-registered hypotheses, retail investors become more risk-seeking when informed about the relative success of their peers. Individuals in the experimental condition who have received upward social comparison information regarding the performance of their peers allocate a significantly higher amount of their portfolio to the risky asset in a following round of trading compared to the control condition. We document a small but robust treatment effect (d= 0.2). This "keeping up with the Joneses" effect, where people adjust their financial behavior in reaction to their peers has been observed in previous literature on peer effects in various empirical settings (Bursztyn & Ederer, 2014; Agarwal, Mikhed & Scholnick, 2019; Kuhn, Kooreman, Soetevent & Kapteyn, 2011), and has already received substantial attention from experimentalists (Dijk, Holmen & Kirchler, 2014; Schoenberg & Haruvy, 2012).

However, this pattern has not yet been so clearly identified in an experimental setting with a sample of retail investors, especially when it comes to how upward social comparison affects trading activity. In Study 1 we find that in line with our hypothesis, individuals who are provided upward social comparison information on the performance of their peers exhibit higher trading activity than individuals in the control condition who have not been given this information. When decomposing the observed group difference in aggregate trading volume, it seems to be driven equally by the number of transactions executed, and the average size of the executed transactions. The observed group difference is small in terms of effect size (d= 0.13) but similarly to our finding regarding risk-taking, it would be economically highly meaningful if scaled to the overall retail market (resulting in hundreds of millions of additional shares traded on a daily basis by retail investors). This is a novel finding, as we are not aware of previous laboratory experiments which have investigated the relationship between upward social comparisons and trading activity. As an explorative finding, we document lower satisfaction with one's own performance in the experimental condition – this is in line with Schoenberg and Haruvy's similar finding (2012), and highlights a possible unexpected side effect of social comparison messages both in the laboratory, as well as in applied settings.

Overall, we find in our experimental setting that social information provided to our participants contributed to behavior which can ultimately be bad for the financial health of individuals in the long run (Odean, 1999; Barber & Odean, 2000), and possibly detrimental for long-term customer satisfaction. Given recent developments in the landscape of retail finance, renewed research attention should be focused on how social information can affect economic decisions. As the use of marketing materials for financial services and financial products which rely on social comparison- and peer influence messaging is becoming more widespread, it would be worthwhile to investigate this dynamic further.

In Study 2, social information is provided in the form of skewed performance feedback, in order to systematically manipulate the self-assessment of a retail investor sample. We successfully demonstrate that people's perception of relative placement on a previous financial literacy task can be modified using a scale manipulation. By providing participants with skewed performance feedback information, we successfully manipulate their subjective placement regarding performance in a previous quiz task upwards by about 8 percentiles.

Contrary to our hypotheses, we do not find a causal effect of our manipulation on individual trading behavior when examining risk taking, aggregate trading volumes, and the average numberand size of trades executed as dependent variables. We document in two studies that individuals who rank themselves higher in a financial knowledge task compared to others do take on more risk in a subsequent trading task in most model specifications. These exploratory findings are in line with previous research relying on empirical data (Merkle, 2017; Glaser & Weber, 2007; Grinblatt & Keloharju, 2009), and further highlight the importance social comparisons and social information in general play in accurate self-assessment, and ultimately trading behavior. Future studies should investigate how relative beliefs about skills and abilities impact individuals' financial decisions. A better understanding of this relationship could in the future contribute to efforts to inhibit individuals' urge to overtrade and to make overly risky investments.

Apart from advancing our understanding of the overconfidence-trading activity relationship in an experimental setting, Study 2 also contributes a novel method of experimentally manipulating overplacement, which could be employed in other contexts to investigate biased self-assessments. Similarly to Study 1, Study 2 highlights the important role beliefs about others play in self-assessment and ultimately in retail investor behavior.

Our combined insights from Studies 1 and 2 can be summarized as follows: retail investors are sensitive to social information from their peers, and such information can affect them in multiple ways. It can affect self-assessments, as well as trading behaviors and post-trading satisfaction. While these two studies demonstrate the potentially detrimental effect social information can have on individual investors, they are not without their limitations. Both studies rely on retail investors selected from online pool participants who were paid relatively small amounts of money tied to their performance in the trading task. While such experiments can provide an informative "proof of concept", further research is necessary in the context of a field experiment to enable us to make a definitive statement on how social information impacts retail investor behavior both directly and indirectly. Still, already our current findings suggest a need for additional attention from regulators towards the marketing and design of trading services, and highlight the need to educate individual investors about the pitfalls of relying on social information when investing.

Discussion – Studies 3 and 4

The last two studies of this dissertation take a contrasting approach to the first two studies. While in Studies 1 and 2 our main focus is on how information from- and about social others changes our beliefs, attitudes and ultimately our behavior, in Studies 3 and 4 our focus turns to the absence of others in situations where they usually have a supporting or advising role.

Study 3 investigates individuals' attitudes towards investment decisions made by algorithms, compared to investment decisions made by human agents. Relying on previous literature on algorithm aversion and moral decision-making, we formulate the research question that the perceived lack of moral competence in machines could be one of the drivers of the previously observed tendency to discount advice from algorithms. In a series of five studies, we document that individuals prefer to follow investment advice from humans over algorithms, even when they are informed that the expected outcome of both decision scenarios is very similar. This speaks for a reluctance to abandon information provided by our peers (for example financial advisors or other comparable investors) for machine-generated advice. We also find that in accordance with our hypotheses, this effect is stronger in cases where the investment decision at hand has a moral valence. In addition to our pre-registered hypotheses, we also document substantial demographic differences – most interestingly, an internal meta-analysis across all five studies reveals a previously undocumented gender gap, where men are more open to automated financial advice than women are. These findings suggest that while the efficiencies provided by automating various types of services in the financial domain and beyond might be substantial, this disintermediation comes with trade-offs. Similar to the loss of soft information in the transition from relationship banking to transactional banking, taking the interpersonal and social aspect out of the advisory relationship can have unintended consequences on retail investor's investment choices, and their attitudes towards how their money is managed. In addition, this aversion seems to have a demographic element, with women being more conservative in this matter.

Similarly to Studies 1 and 2 of this dissertation, the experiments forming the basis of Study 3 have relied on retail investors drawn from an online labor pool. Such an approach comes with limitations for the external validity of findings, while still being a better fit for our research question than samples of university students often relied on for similar experiments. A further critical point is that only one of the five experiments employs variable monetary incentives contingent on participants' choices. Reflecting on recent developments in retail investing, a planned follow-up study currently in preparation will investigate the intersection of robo-advice, short-selling and moral considerations from an investor perspective.

Given the growing demand for ESG- and impact investments, a hesitancy to delegate sensitive investment decisions to machines is something for industry participants to be mindful of, and something for researchers to investigate further in the future. As suggested by our findings, as well as previous literature (Merkle, 2020), retaining human involvement in some capacity in the advisory process is likely to result in a higher acceptance of algorithmic advice at the present point in time. A hybrid system of human advisors relying on advice generated by algorithms is likely to lead to more adoption and ultimately aid the market penetration of robo-advisory services.

Study 4 examines prospective users' comfort with removing third-party human supervision in a hypothetical scenario involving an autonomous vehicle. Thus, this study is in keeping with Study 3 in investigating how people react to the absence of others in scenarios where they are currently present. We conduct an analysis of representative survey data collected by the European Commission, our variable of interest being one's self-reported comfort with being a passenger in a fully autonomous vehicle with varying degrees of third-party human supervision. Just as in Study 3, we document a number of relevant demographic and psychological factors which are associated with individuals' openness to rely on automation. We find overall that respondents are less comfortable with decreasing human presence. Furthermore, we document substantial gender effects shaping one's reaction to decreasing human presence and human supervision, with women becoming substantially less open to traveling in an autonomous vehicle with decreasing intensity of human supervision than men. In addition to pre-existing demographic differences, such as gender and age, we also document a strong relationship between attitudes towards data sharing and comfort with being a passenger in an autonomous vehicle.

Similar to Study 3, we interpret our findings in Study 4 to mean that individuals are hesitant to try technologies where human interaction is either partially or completely replaced by machines. This study again highlights the importance of hybrid solutions (which might have higher initial acceptance), as well as the need for careful communication of the potential risks and benefits of various forms of autonomous vehicles towards the public. These findings should guide future research to investigate the causal determinants of attitudes towards using autonomous vehicles in an experimental setting, to ultimately inform future policy and communication efforts. In terms of future research within an experimental framework, the results of Study 4 suggest further work to be done on how best to communicate policies regarding the use of personal data when using AVs, and to investigate the underlying attitudinal differences which might be driving the gender differences we have observed. Future research should also investigate attribution of blame between human supervisors and autonomous vehicles, similarly to Awad et al. (2020), but incorporating different levels of human supervision.

While they rely on different methodologies, and are embedded in different real-world contexts, the results from Studies 3 and 4 are similar in many ways. Both document an overall aversion to remove human interaction from a decision scenario characterized by uncertainty, and demonstrate that such attitudes are systematically associated with individual characteristics such as gender. This generalized aversion calls for a more nuanced approach when rolling out automated services across domains – given people's reluctance to give up human advice and human supervision in the two contexts we have investigated (finance and mobility), intermediate approaches retaining human interaction might serve as useful stepping stones towards fully autonomous and unsupervised approaches. Acceptance of such "human-in-the-loop" approaches also provide a promising area of future research.

Conclusion

It is not a simple task to distill a common conclusion from four mostly independent projects, consisting of a combined 9 studies with different methodologies, and with data collections involving more than 5200 individuals spanning almost two years. The four studies discussed previously in brief are included in full detail in the following chapters, and should stand mostly on their own as individual contributions. However, these projects also have clear commonalities, and by discussing their findings together, an interesting picture emerges concerning how living in a highly networked and fast-changing world can affect decisions across different domains.

The first two studies of this dissertation reveal how sensitive individual investors are to information from their social environment, even in somewhat artificial experimental settings. The mere fact that the good performance of unknown individuals affects how retail investors trade and how satisfied they are with their own performance is interesting in itself, and highlights the importance of social information in financial decisions. We also find that relative performance information influences how individuals assess their own abilities and is associated with their trading behavior. As one's performance and compensation are independent from the absolute or relative performance of others in these studies, a purely rational model of economic decision-making, which does not take into account the importance of social information would not do a good job of explaining what we have observed. The outsize importance of other people as a source of information, advice and support is also highlighted in Studies 3 and 4 – we find in both of these studies that individuals are reluctant to replace other individuals as a source of information and support, even when they could potentially benefit from doing so.

Our social world is increasingly complex, and technological and societal developments make it an even bigger challenge to understand and anticipate human behavior. Overall, we find that people are very much influenced by information from their social environment, and this influence can shape beliefs and behaviors in systematic, and often counterproductive ways. However, in scenarios where one has to make important decisions under uncertainty, individuals still show a consistent preference for retaining human involvement. To quote John Donne: "No man is an island" (1624, 2007). Or to paraphrase George Costanza: "We are living in a society!" (David & Seinfeld, 1991). Solutions looking to shape or guide human behavior to achieve certain policy- or business goals have to take this complexity into account.

Bibliography

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, *33*(2), 3-30.

Agarwal, S., Mikhed, V., & Scholnick, B. (2020). Peers' income and financial distress: Evidence from lottery winners and neighboring bankruptcies. *The Review of Financial Studies*, *33*(1), 433-472.

Akçay, E., & Hirshleifer, D. (2021). Social finance as cultural evolution, transmission bias, and market dynamics. *Proceedings of the National Academy of Sciences*, *118*(26).

Allport, G. W. (1954). The historical background of modern social psychology. In G. Lindzey (Ed.), *Handbook of social psychology* (Vol. 1, pp. 3– 56). Cambridge MA: Addison-Wesley.

Alexander, A. J., Shi, L., & Solomon, B. (2017). How Fintech is Reaching the Poor in Africa and Asia: A Start-Up Perspective. EMCompass, no. 34. *International Finance Corporation*. Washington, DC. Retrieved from: https://openknowledge.worldbank. org/handle/10986/30360

Andereck, J. W., Reuter, Q. R., Allen, K. C., Ansari, S., Quarles, A. R., Cruz, D. S., ... & Kim, H. S. (2019). A quality improvement initiative featuring peer-comparison prescribing feedback reduces emergency department opioid prescribing. *The Joint Commission Journal on Quality and Patient Safety*, *45*(10), 669-679.

Anderson, C., Brion, S., Moore, D. A., & Kennedy, J. A. (2012). A status-enhancement account of overconfidence. *Journal of Personality and Social Psychology*, *103*(4), 718.

Arkes, H. R., Christensen, C., Lai, C., & Blumer, C. (1987). Two methods of reducing overconfidence. *Organizational Behavior and Human Decision Processes*, *39*(1), 133-144.

Aronson, E., & Aronson, J. (2018). *The social animal.* Worth Publishers, Macmillan Learning.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., ... & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59-64.

Awad, E., Dsouza, S., Shariff, A., Rahwan, I., & Bonnefon, J. F. (2020). Universals and variations in moral decisions made in 42 countries by 70,000 participants. *Proceedings of the National Academy of Sciences*, *117*(5), 2332-2337.

Awad, E., Levine, S., Kleiman-Weiner, M., Dsouza, S., Tenenbaum, J. B., Shari, A., . . . Rahwan, I. (2020). Drivers are blamed more than their automated cars when both make mistakes. *Nature Human Behaviour*, *4* (2), 134–143.

Backstrom, L., Boldi, P., Rosa, M., Ugander, J., & Vigna, S. (2012, June). Four degrees of separation. In *Proceedings of the 4th Annual ACM Web Science Conference* (*pp.* 33-42).

Baddeley, M. (2010). Herding, social influence and economic decision-making: socio-psychological and neuroscientific analyses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1538), 281-290.

Bak-Coleman, J. B., Alfano, M., Barfuss, W., Bergstrom, C. T., Centeno, M. A., Couzin, I. D., ... & Weber, E. U. (2021). Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences*, *118*(27).

Baldwin, M., & Mussweiler, T. (2018). The culture of social comparison. *Proceedings of the National Academy of Sciences*, *115*(39), E9067-E9074.

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, *55*(2), 773-806.

Barber, B, Odean T. (2001) Boys will be boys: gender, overconfidence, and common stock investment. *Quarterly Journal of Economics, 141*(2001), pp. 261-292

Barberis, N. (2018). Psychology-based models of asset prices and trading volume. In *Handbook of Behavioral Economics: Applications and Foundations 1* (Vol. 1, pp. 79-175). North-Holland.

Bartke, S., Friedl, A., Gelhaar, F., & Reh, L. (2017). Social comparison nudges—Guessing the norm increases charitable giving. *Economics Letters*, *152*, 73-75.

Becker, F., & Axhausen, K. W. (2017). Literature review on surveys investigating the acceptance of automated vehicles. *Transportation*, *44*(6), 1293-1306.

Benartzi, S., & Thaler, R. (2007). Heuristics and biases in retirement savings behavior. *Journal of Economic Perspectives*, *21*(3), 81-104.

Benoît, J. P., Dubra, J., & Moore, D. A. (2015). Does the better-than-average effect show that people are overconfident?: Two experiments. *Journal of the European Economic Association*, *13*(2), 293-329.

Bernardo, A. E., & Welch, I. (2001). On the evolution of overconfidence and entrepreneurs. *Journal of Economics & Management Strategy*, *10*(3), 301-330.

Bi, Y., Dang, Q., Li, S., Guo, J., & Zhang, B. (2016). The Effect of Overconfidence on Persistent Behavior: The Mediation Effect of “I Think I Can Do It” Rather Than “I'm Attracted To It”. *Psychological Reports*, *118*(1), 138-153.

Bicchieri, C., & Dimant, E. (2019). Nudging with care: The risks and benefits of social information. *Public Choice*, 1-22

Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, *181*, 21-34.

Bottemanne, H., Morlaàs, O., Fossati, P., & Schmidt, L. (2020). Does the coronavirus epidemic take advantage of human optimism bias? *Frontiers in Psychology*, *11*.

Bregu, K. (2020). Overconfidence and (Over) Trading: The Effect of Feedback on Trading Behavior. *Journal of Behavioral and Experimental Economics*, 88, 101598.

Brogaard, J., Hendershott, T., & Riordan, R. (2014). High-frequency trading and price discovery. *The Review of Financial Studies*, *27*(8), 2267-2306.

Brown, K. C., Harlow, W. V., & Starks, L. T. (1996). Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *The Journal of Finance*, *51*(1), 85-110.

Bursztyn, L., Ederer, F., Ferman, B., & Yuchtman, N. (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, *82*(4), 1273-1301.

Buunk, A. P., & Gibbons, F. X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes*, *102*(1), 3-21.

Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American economic review*, *89*(1), 306-318.

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

Cartwright, D. (1979). Contemporary social psychology in historical perspective. *Social Psychology Quarterly*, 82-93.

Choi, A. (2016). Morgan stanley perspectives on sustainable investing: Acceleration and integration. *Journal of Applied Corporate Finance*, *28*(2), 62-65.

Cuthbertson, K., Nitzsche, D., & O'Sullivan, N. (2016). A review of behavioural and management effects in mutual fund performance. *International Review of Financial Analysis*, *44*, 162-176.

David, L, Seinfeld, J. (Writers) (1991, May 23.) "The Chinese Restaurant" (Season 2, Episode 11) [TV series episode] In David L., et al. *Seinfeld*, Castle Rock Entertainment, 1991

Davis, M., Kumiega, A., & Van Vliet, B. (2013). Ethics, finance, and automation: A preliminary survey of problems in high frequency trading. *Science and Engineering Ethics*, *19*(3), 851-874.

Deaves, R., Lüders, E., & Luo, G. Y. (2008). An experimental test of the impact of overconfidence and gender on trading activity. *Review of Finance*, *13*(3), 555-575.

Deming, D. J. (2021). *The Growing Importance of Decision-Making on the Job* (No. w28733). National Bureau of Economic Research.

Diel, K., Grelle, S., & Hofmann, W. (2021). A motivational framework of social comparison. *Journal of Personality and Social Psychology*.

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

Dijk, O., Holmen, M., & Kirchler, M. (2014). Rank matters–The impact of social competition on portfolio choice. *European Economic Review*, *66*, 97-110.

Dillard, A. J., McCaul, K. D., & Klein, W. M. (2006). Unrealistic optimism in smokers: Implications for smoking myth endorsement and self-protective motivation. *Journal of Health Communication*, *11*(S1), 93-102.

Donne, J., & In Dickson, D. R. (2007*). John Donne's poetry: Authoritative texts, criticism*. W. W Norton

Dunbar, R. I. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *22*(6), 469-493.

Dunbar, R. I. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, *6*(5), 178-190.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5*(3), 69-106.

Durand, R. B., Koh, S., & Limkriangkrai, M. (2013). Saints versus Sinners. Does morality matter? *Journal of International Financial Markets, Institutions and Money*, *24*, 166-183.

Ewing, J (2021). How Germany Hopes to Get the Edge in Driverless Technology. *NYTimes.com*. Retrieved from: https://www.nytimes.com/2021/07/14/business/germany-autonomous-driving-new-law.html

Egkolfopoulou, M, Massa, A., and Melin, A. (2021, 21. April) How Robinhood Made Trading Easy—and Maybe Even Too Hard to Resist. Bloomberg Businessweek. Retrieved from: https://www.bloomberg.com/features/2021-robinhood-stock-trading-design/

European Commission (2021) Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS COM/2021/206 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206

Financial Conduct Authority, UK (2021) *Understanding self-directed investors.* Retrieved from: https://www.fca.org.uk/publication/research/understanding-self-directed-investors.pdf

Fish, A. and Kim, DH and Venkatraman, S. The ESG Sacrifice (November 17, 2019). Available at SSRN: https://ssrn.com/abstract=3488475 or http://dx.doi.org/10.2139/ssrn.3488475

Frey, B. S., & Meier, S. (2004). Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment. *American Economic Review*, *94*(5), 1717-1722.

Gerber, J. P., Wheeler, L., & Suls, J. (2018). A social comparison theory meta-analysis 60+ years on. *Psychological bulletin*, *144*(2), 177.

Germann, M., & Merkle, C. (2020). Algorithm Aversion in Delegated Investing. *Available at SSRN 3364850*.

Gervais, S., & Odean, T. (2001). Learning to be overconfident. *The Review of Financial Studies*, *14*(1), 1-27.

Glaser, M., & Weber, M. (2007). Overconfidence and trading volume. *The Geneva Risk and Insurance Review*, *32*(1), 1-36.

Global Sustainable Investment Alliance, 2019. http://www.gsi-alliance.org/wp-content/uploads/2019/03/GSIR_Review2018.3.28.pdf

Goethals, G. R. (1986). Social comparison theory: Psychology from the lost and found. *Personality and Social Psychology Bulletin*, *12*(3), 261-278.

Gogoll, J., & Uhl, M. (2018). Rage against the machine: Automation in the moral domain. *Journal of Behavioral and Experimental Economics*, *74*, 97-103.

Gogoll, J., & Müller, J. F. (2017). Autonomous cars: in favor of a mandatory ethics setting. *Science and engineering ethics*, *23*(3), 681-700.

Grinblatt, M., & Keloharju, M. (2009). Sensation seeking, overconfidence, and trading activity. *The Journal of Finance*, *64*(2), 549-578.

Grønsund, T., & Aanestad, M. (2020). Augmenting the algorithm: Emerging human-in-the-loop work configurations. *The Journal of Strategic Information Systems*, *29*(2), 101614.

Hattie, J. A., & Hansford, B. C. (1982). Self-measures and achievement: Comparing a traditional review of literature with a meta-analysis. *Australian Journal of Education*, *26*(1), 71-75.

Haboucha, C. J., Ishaq, R., & Shiftan, Y. (2017). User preferences regarding autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, *78*, 37-49.

Hertwig, R., & Herzog, S. M. (2009). Fast and frugal heuristics: Tools of social rationality. *Social Cognition*, *27*(5), 661-698.

Hertwig, R., & Hoffrage, U. (2013). Simple heuristics: the foundations of adaptive social behavior. In *Simple heuristics in a social world* (pp. 3-36). Oxford University Press.

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of financial economics*, *93*(1), 15-36.

Jagtiani, J., & Lemieux, C. (2018). Do fintech lenders penetrate areas that are underserved by traditional banks?. *Journal of Economics and Business*, *100*, 43-54.

Jakšič, M., & Marinč, M. (2019). Relationship banking and information technology: The role of artificial intelligence and FinTech. *Risk Management*, *21*(1), 1-18.

Jansen, J., & Van Schaik, P. (2018). Testing a model of precautionary online behaviour: The case of online banking. *Computers in Human Behavior*, *87*, 371-383.

Johnson, D. D., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature*, *477*(7364), 317-320.

Kaustia, M., & Knüpfer, S. (2012). Peer performance and stock market entry. *Journal of Financial Economics,* 104(2), 321-338.

Kennedy, J. A., Anderson, C., & Moore, D. A. (2013). When overconfidence is revealed to others: Testing the status-enhancement theory of overconfidence. *Organizational Behavior and Human Decision Processes*, *122*(2), 266-279.

Köbis, N., Bonnefon, J. F., & Rahwan, I. (2021). Bad machines corrupt good morals. *Nature Human Behaviour*, *5*(6), 679-685.

Kuhn, P., Kooreman, P., Soetevent, A., & Kapteyn, A. (2011). The effects of lottery prizes on winners and their neighbors: Evidence from the Dutch postcode lottery. *American Economic Review*, *101*(5), 2226-47.

Lee, K. W., Tsai, M. T., & Lanting, M. C. L. (2011). From marketplace to marketspace: Investigating the consumer switch to online banking. *Electronic Commerce Research and Applications*, *10*(1), 115-125.

Leyshon, A., French, S., & Signoretta, P. (2008). Financial exclusion and the geography of bank and building society branch closure in Britain. *Transactions of the Institute of British Geographers*, *33*(4), 447-465.

Lindenfors, P., Wartel, A., & Lind, J. (2021). 'Dunbar's number' deconstructed. *Biology Letters*, *17*(5), 20210158

Linder, J. A., Meeker, D., Fox, C. R., Friedberg, M. W., Persell, S. D., Goldstein, N. J., & Doctor, J. N. (2017). Effects of behavioral interventions on inappropriate antibiotic prescribing in primary care 12 months after stopping interventions. *JAMA*, *318*(14), 1391-1392.

Logg, J. M., Haran, U., & Moore, D. A. (2018). Is overconfidence a motivated bias? Experimental evidence. *Journal of Experimental Psychology: General*, *147*(10), 1445.

Long, C., Lucey, B. M., & Yarovaya, L. (2021). " I Just Like the Stock" versus" Fear and Loathing on Main Street": The Role of Reddit Sentiment in the GameStop Short Squeeze. *SSRN Electronic Journal*, 31.

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, *46*(4), 629-650.

Mackintosh, P. (2020). Who Counts as a Retail Investor? https://www.nasdaq.com/articles/who-counts-as-a-retail-investor-2020-12-17. Website: https://www.nasdaq.com/articles/who-counts-as-a-retail-investor-2020-12-17. Last retrieved: 16.07.2021

Martínez-Díaz, M., & Soriguera, F. (2018). Autonomous vehicles: theoretical and practical challenges. *Transportation Research Procedia*, *33*, 275-282.

Merkle, C. (2017). Financial overconfidence over time: Foresight, hindsight, and insight of investors. *Journal of Banking & Finance*, *84*, 68-87.

Merkle, C. (2020). Robo-advice and the future of delegated investment. *Journal of Financial Transformation*, *51*, 20-27.

Moore, D. A., & Dev, A. S. (2018). Individual differences in overconfidence. In Zeigler-Hill and

T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences*. New York: Springer.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, *115*(2), 502.

Moulliet D, Majonek A, Stolzenbach J, Völker T. *The Expansion of Robo-Advisory in Wealth Management*. Deloitte. 2016. Available: https://www2.deloitte.com/content/dam/Deloitte/de/Documents/financial-services/Deloitte-Robo-safe.pdf

Myers, E., & Souza, M. (2020). Social comparison nudges without monetary incentives: Evidence from home energy reports. *Journal of Environmental Economics and Management*, *101*, 102315.

Odean, T. (1999). Do investors trade too much? *American Economic Review*, *89*(5), 1279-1298.

Paxful (2021)- https://paxful.com/insights/influencer-investors/#views-and-likes. Last accessed: 09.11.2021

Pardo-Guerra, J. P. (2019). *Automating finance: infrastructures, engineers, and the making of electronic markets*. Cambridge University Press.

Plous, S. (1993). *The psychology of judgment and decision making*. New York, NY: McGraw-Hill.

Puri, M., & Robinson, D. T. (2007). Optimism and economic choice. *Journal of financial economics*, *86*(1), 71-99.

Puschmann, T. (2017). Fintech. *Business & Information Systems Engineering*, *59*(1), 69-76.

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, *32*(9), 1723-1742.

Richard, F. D., Bond Jr, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*(4), 331-363.

Rietveld, N., Groenen, P., Koellinger, P., van der Loos, M., & Thurik, R. (2013). *Living forever: Entrepreneurial overconfidence at older ages* (No. ERS-2013-012-STR).

Rochat, P. (2009). *Others in mind: Social origins of self-consciousness*. Cambridge University Press.

Russo, J. E., & Schoemaker, P. J. (1992). Managing overconfidence. *Sloan Management Review*, *33*(2), 7-17.

Schabicki, T., Quint, Y., & Schroeder, S. (2021). (Re-) Launching a Robo-Advisor as a Bank. In *Robo-Advisory* (pp. 163-186). Palgrave Macmillan, Cham.

Schoenberg, E. J., & Haruvy, E. (2012). Relative performance information in asset markets: An experimental approach. *Journal of Economic Psychology*, *33*(6), 1143-1155.

Scholz, P., & Tertilt, M. (2021). Robo-Advisory: The Rise of the Investment Machines. In *Robo-Advisory* (pp. 3-19). Palgrave Macmillan, Cham.

Searle, J. (2010). *Making the social world: The structure of human civilization*. Oxford University Press.

SEC (2018) *SEC Charges Two Robo-Advisers With False Disclosures*. Retrieved from: https://www.sec.gov/news/press-release/2018-300

Sensoy, B. A. (2009). Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics*, *92*(1), 25-39.

Shariatmadari, D. (2015, July 18). Daniel Kahneman: 'What would I eliminate if I had a magic wand? Overconfidence'. *The Guardian*. Retrieved from https://www.theguardian.com/books/2015/jul/18/daniel-kahneman-books-interview

Shiller, R. J. (2020). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.

Smith, V. L., Suchanek, G. L., & Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica: Journal of the Econometric Society*, 1119-1151.

Snowberg, E., & Yariv, L. (2021). Testing the waters: Behavior across participant pools. *American Economic Review*, *111*(2), 687-719.

Sparrow, R., & Howard, M. (2017). When human beings are like drunk robots: Driverless vehicles, ethics, and the future of transport. *Transportation Research Part C: Emerging Technologies*, *80*, 206-215.

Statista (2017*) Total value of assets under management by robo-advisors worldwide from 2015 to 2020*. Retrieved from: https://www.statista.com/statistics/741512/aum-of-robo-advisors-globally/

Statista (2021) *Robo-Advisors*, Retrieved from: https://www.statista.com/outlook/dmo/fintech/personal-finance/robo-advisors/worldwide

Taylor, J. (2003). Risk-taking behavior in mutual fund tournaments. *Journal of Economic Behavior & Organization*, *50*(3), 373-383.

Tertilt, M., & Scholz, P. (2018). To advise, or not to advise—how robo-advisors evaluate the risk preferences of private investors. *The Journal of Wealth Management*, *21*(2), 70-84.

Tomasello, M. (2014). The ultra-social animal. *European journal of social psychology*, *44*(3), 187-194.

Thornton, D. A., & Arrowood, A. J. (1966). Self-evaluation, self-enhancement, and the locus of social comparison. *Journal of Experimental Social Psychology*, *1*, 40-48.

Tranfaglia, A. (2018). Shrinking networks: A spatial analysis of bank branch closures. (Federal Reserve Bank of Philadelphia, Working Paper 18-12)

Wolff, J. (2021) The EU's New Proposed Rules on A.I. Are Missing Something. *Slate*. https://slate.com/technology/2021/04/eu-proposed-rules-artificial-intelligence.html

Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological bulletin*, *106*(2), 231.

Yang, X., & Zhu, L. (2016). Ambiguity vs risk: An experimental study of overconfidence, gender and trading activity. *Journal of Behavioral and Experimental Finance*, *9*, 125-131.

Yeomans, M., Shah, A., Mullainathan, S., & Kleinberg, J. (2019). Making sense of recommendations. *Journal of Behavioral Decision Making*, *32*(4), 403-414.

YouGov (2021) *Rumble in the jungle: what animals would win in a fight?* Retrieved from: https://today.yougov.com/topics/lifestyle/articles-reports/2021/05/13/lions-and-tigers-and-bears-what-animal-would-win-f

Zell, E., Strickhouser, J. E., Sedikides, C., & Alicke, M. D. (2020). The better-than-average effect in comparative self-evaluation: A comprehensive review and meta-analysis. *Psychological Bulletin*, *146*(2), 118.

Study 1: I'll Have What They're Having – The Influence of Upward Social Comparison on Retail Trading Behavior

Dániel Kaszás[1], Sandra Andraszewicz[1], Stefan Zeisberger[2,3], Christoph Hölscher[1]

[1]ETH Zurich

[2]Radboud University Nijmegen

[3]Unversity of Zurich

Abstract

In a study with 807 participants with previous investing experience, we investigated the impact of upward social comparison on risk taking and trading activity. We implemented two trading rounds using a novel experimental trading platform that displays historical market prices as they develop over time, and allows users to make trades based on this price information. Participants were randomly allocated to one of two conditions: they either received feedback on only their own performance after the first round of trading, or received additional information on how three successful prior participants have performed. We find that participants presented with an upward social comparison took more risk and traded more actively in a subsequent round of trading than participants not exposed to an upward social comparison. We also document that participants exposed to an upward social comparison report significantly lower post-task satisfaction with their performance. These findings extend current research by demonstrating the impact of peer performance on trading activity, risk taking and investor satisfaction.

Keywords: social comparison, trading behavior, self-assessment, post-task satisfaction

Investing is an inherently social activity (Shiller, 1984). This is truer than ever now, as we are experiencing a historical increase in retail participation and trading volumes (McCrank, 2021; CBOE, 2021) This shift is taking place in the context of an increasingly widespread use of social media to exchange information on trading opportunities and sharing investing experiences. Similar to early chat rooms and online message boards during the dotcom boom (Lewis, 2001), a variety of online forums sprung up over recent years to facilitate exchange between actual and prospective retail investors (such as r/WallStreetBets, StockTwits and Finance TikTok). In conjunction with the increasing retail interest in finance and investment, a number of "social trading" platforms have been launched recently, bringing a social media element into the investment process. According to a website providing information and advice on social trading, 22 designated social trading platforms were active in 2019 (SocialTradingGuru, 2021).

A key development which enabled the proliferation and fast adoption of these new trading platforms was a fundamental shift in the business model of how retail trading platforms operate. Compared to the earlier, commission-based business model, services such as Robinhood, WeBull and eToro enable commission-free trading. In this new paradigm, the main profit driver for service providers is attracting novice investors to their service and maximizing their trading activity. In a sense the market signals generated by the mostly novice investors on these platforms become the product itself, which gets sold to the highest-bidding third party (Pisani, 2021; SEC, 2000). Commission-free trading broadens the circle of potential investors, and incentivizes higher trading activity – according to some observers, potentially to the detriment of consumer financial welfare (Galvin, 2020; Nicodemus, 2020).

Platform providers rely on a variety of measures to increase user engagement. Such measures, informed by insights from cognitive science include efforts to "gamify" investing through user experience design elements (Massa, McDonald & Alexander, 2020). A common element found across the board in most trading platforms is appeals to people's inherently competitive nature, by presenting leaderboards of the most successful or most active users

(Galvin, 2020), or organizing competitions, where the most successful participants and their results are displayed prominently. Industry insiders report the use "peer marketing strategies" to drive engagement on their platforms in their desired demographic (Wursthorn & Choi, 2020). Information on the extraordinary success of other users is not only present in marketing materials – in the case of a number of social trading platforms they are the most salient user interface elements. For example, the social trading platforms eToro and ZuluTrade prominently list their top traders, who can be "followed" by other users.

Since these changes in the retail investment space are taking place at dizzying pace, many aspects of how investment decisions are potentially influenced by social comparison have not received adequate research attention. Generating trading activity is of central importance to the success of modern trading platforms – and since active trading has been associated with negative outcomes for retail investors in past research (Odean, 1999; Barber & Odean, 2000; Barber, Lee, Liu & Odean, 2009), gaining a clear understanding of how widely implemented social comparison appeals impact retail investors' trading behavior is a crucial question. It is especially important given the increasing popularity of social trading platforms, where providing information on how others have performed is a key component of such services (Oehler, Horn & Wendt, 2016; Röder & Walter, 2019).

According to Suls, Martin and Wheeler (2002) "social comparison consists of comparing oneself with others in order to evaluate or to enhance some aspects of the self" (p.159.). The concept of social comparison has been at the center of psychological research since Festinger's pioneering work (1954). According to Festinger's theory of the process of social comparison (1954), individuals have a need to accurately assess their skills and abilities. While they prefer objective information, if such information is noisy or absent, individuals will fall back on social comparisons. Such comparisons can be directed either upward or downward (Buunk & Gibbons, 2007) – comparing one's own skill, performance or abilities to people who are superior or inferior in some respect. While Festinger hypothesized the existence of an overall an "upward drive" to compare one with a superior reference group when it comes to the self-evaluation of abilities, more recent research finds the choice of reference group (better or worse) is a complex and dynamic process with a number of moderators (Buunk, Gibbons, 2007; Wheeler, 1966; Wheeler & Miyake, 1992; Gruder, 1971).

Previous efforts to investigate the role of peer influence and social comparisons when it comes to investment decisions have been mostly experimental in nature, due to the increased level of control required to cleanly identify the effect of social comparison on one's behavior (Baghestanian, Gortner & van der Weele, 2015). Outside an experimental setting, it is very challenging to ascertain that all individuals have the same information and beliefs about the market, their peers, and their relative skill or success. Compounding this problem is the fact that in most real-world settings there is often not just one clearly identified peer group one could compare themselves with – social comparisons are often made in the presence of incomplete information, and reference groups can shift depending on the situation (Festinger, 1954; Gruder, 1971). Clearly communicating social comparison information in an experimental setting enables the experimenter to control participant's beliefs and provide them with a uniform and clearly identified peer group.

Kirchler, Lindner and Weitzel (2018) conducted laboratory investment experiments with both students and professionals, where they investigated the influence of ranking the participants on their risk taking. They found that social influence induced via a tournament incentive scheme increases risky investment. A tournament is characterized by two components: a monetary reward that depends on one's position in the rank and a non-monetary incentive to outperform the peers. The authors found that both components increase risk taking in underperforming professionals while only monetary component affected students. This finding speaks to the increased importance of non-monetary incentives social reference points when it comes to the domain of finance and investments.

In a laboratory experiment Dijk, Holmen and Kirchler (2013) asked their participants to choose between more or less risky portfolios. They concluded that people change their portfolio choices as a response to their relative rank, such that relative underperformance results in higher risk taking and vice versa. Most importantly, they find that presenting social comparison information leads to higher risk taking, even when it is not payoff-relevant and attached to tournament incentives. This further highlights the importance of social status concerns in the finance domain. Similarly, Schoenberg and Haruvy (2012) report that providing participants in a 15-round experimental asset market with an upward performance comparison leads to higher overall market prices and bubble metrics. Similar findings are reported by Deck, Hao, Xu and Yeager (2020). Conversely, Baghestanian, Gortner and van der Weele (2015) find that exposing their participants to others' performance on an experimental asset markets leads to less risky portfolio allocations, independently of how others have performed. In a novel laboratory experiment investigating the impact of social influence on social trading platforms on the users' risk taking, Apesteguia, Oechssler and Weidenholzer (2020) report that information about the success of other participants leads to increased risk-taking.

When it comes to trading activity, existing findings on the effect of social influence and peer effects are much more sparse and quasi-experimental in nature. Escobar and Pedraza (2019) conducted an ingenious field experiment, in which participants in a financial education initiative were randomly allocated to groups containing others who either do or do not already trade on their own. They document the existence of a social transmission channel, where those randomly allocated to a group with a member who is already an active trader were more likely to become active traders as well. This is especially strong in cases where the already active traders report achieving high returns. Escobar and Pedroza also observe systematic underperformance on the part of those novice investors whose group members have reported very high returns – they propose biased communication of positive results by already active traders as a probable mechanism driving this observation. Individuals selectively communicating only positive information (and thus creating a high social reference point) for various reasons is a well-documented phenomenon in general (Leary & Kowalski, 1990), and has been found to cause higher stock market participation rates in a quasi-experimental setting (Kaustia & Knüpfer, 2012). In a study on the behavior of professional investors (mutual fund managers), Pütz and Ruenzi (2011) report that both the highest- and lowest- performing fund managers in their class increase their turnover ratio over the following period. In the case of the lowest-performers, this points to the possibility that laggards undertake an overhaul of their investment strategy in response to relative underperformance.

One common feature of most previous experimental studies investigating the relationship between social comparison and risk taking is that participants are usually presented with decision problems including a fixed number of investments. In each of these investments, participants choose among choice options with well described features such as possible outcomes, probabilities of these outcomes, expected return etc. However, in real life, investment decisions do not occur at strictly defined times and at a predefined frequency. Therefore, static investment tasks with a fixed number of decision possibilities might well capture risk-taking in investment as the product of a deliberative decision-making process (Ferri, Ploner & Rizzoli, 2016), but they are less successful in reproducing the sometimes hectic and emotionally-driven behaviors a more realistic experimental task would capture. In addition, complex and static round-based investment tasks cannot capture participants' investment activity such as trading volume or number of transactions in a naturalistic way that is representative of the retail investing experience. This dynamic element differentiates the current experimental paradigm from previous experimental efforts exposing participants to historical price paths (for example, Lejarraga, Woike & Hertwig, 2016; Papadovasilaki, Guerrero, Sundali & Stone, 2015), as well as previously discussed studies implementing experimental asset market designs, and makes it an especially good fit for investigating research questions regarding trading activity.

Despite the substantial literature devoted to social comparison and trading behavior (especially regarding risk taking), experimental research focusing on the relation between upward social comparisons and trading activity is still lacking. Given the increasing relevance of retail trading activity and social aspects of investing over the past years, a better understanding of how social comparisons affect retail investor's engagement with the market is of high economic- and social relevance. Demonstrating a causal relationship between upward social comparison and trading activity can fill a gap in our understanding when it comes to social information and investment decisions, which has up to now received relatively little research attention. It can also be the first step in developing and implementing interventions aiming to decrease individuals' trading activity, thereby potentially contributing to better individual-level financial outcomes. As a number of previous studies show, individual investors who trade actively tend to underperform professional investors, according to some studies even before accounting for trading costs (Barber, Lee, Liu & Odean, 2009; Barber, Lee, Liu & Odean, 2014; Ryu, 2012).

In order to investigate the relationship between social comparisons and trading behavior, we have constructed an experiment with a between-groups design. We selectively provide upward performance comparison information to one group of participants after the first round of a trading task, and then observe the resulting behavioral differences that emerge in a subsequent round of the trading task. We measure various facets of trading activity (i.e., number of transactions, average size of transactions and aggregate volume) and risk taking (i.e., percentage of a risky asset in one's portfolio) in the second round of trading. We chose to decompose "trading activity" into three facets we can compute based on the information available to us from our experimental software, and investigate its individual components following previous literature. Previous research relying on empirical data also differentiates between various operationalizations of trading activity such as the overall volume traded, trading frequency, and turnover ratio (e.g. Grinblatt & Keloharju, 2009; Glaser & Weber, 2007; Merkle, 2017).

The aim of this study is twofold. First, we aim to replicate and extend previous findings regarding the relationship between upward social comparison and risk taking with a novel experimental paradigm. Second, our goal is to investigate the influence of social comparison on trading activity. We have derived our four pre-registered hypotheses from the literature discussed above.

Hypothesis 1 *(Risk taking)*: Upward social comparison will lead to increased risk taking, where risk taking (c.f., *Risk*) is defined as the mean fraction of the risky asset in one's portfolio throughout one experimental round.

Hypothesis 2a *(Trading activity):* Upward social comparison will lead to an increased number of transactions, where the number of transactions (c.f., *Transactions*) refers to a number of separate executed trades in one experimental round.

Hypothesis 2b *(Trading activity):* Upward social comparison will lead to an increased volume of each individual trade, where the volume (c.f., *Vol_Avg*) corresponds to the mean size of individuals transactions executed.

Hypothesis 2c *(Trading activity):* Upward social comparison will lead to an increased overall trading volume, where the overall trading volume (c.f., *Vol_Sum*) refers to the cumulative value of all shares traded.

Methods

The study was preregistered on the Open Science Framework (https://osf.io/we48d/) and approved by the ETH Ethics Committee (2020-N-120).

Participants

We recruited 1622 workers from the Amazon Mechanical Turk labor pool between 17.11.2020 – 07.12.2020, with the aim to collect full data from 800 participants based on an a-priori power analysis. 1262 individuals entered the experiment and accepted the informed consent form. After a pre-screening, 807 people (300 female, M_{age} = 40.62, $Md_{Age:}$ 40.62, $Range_{Age}$: 18-83) completed the full study. This results in a 36% screen-out rate for consenting participants. Each prospective participant received a $0.3 base payment (regardless of passing the initial screening or not), while everyone who completed the experiment received a variable performance-dependent bonus ranging between $1.38 and $1.66 ($M_{Bonus}$ = $1.51, Md_{Bonus} =$1.51). Our main selection criterion for recruiting participants was that participants had to be individual investors currently holding direct investments in the US stock market. We have only reached out to potential participants whose investor status has been previously validated by CloudResearch (formerly TurkPrime, Litman, Robinson & Abberbock, 2016). Further criteria include completing the experiment using Chrome (to ensure the best possible compatibility with our experimental platform), being a resident of the United States, and being physically present in the United States at the time of data collection. We have only invited participants who have been identified as "high-quality" by CloudResearch based on their earlier survey participation, and those who have completed more than 100 HITs (Human Intelligence Tasks) with an at least 95% acceptance rate on Mechanical Turk. These criteria were set in order to reach active retail investors on the US stock market, in an effort to increase the generalizability of our findings, and to create a clearly identifiable reference group.

Materials and Procedure

In 13 batches, we sent out invitations to a pool of workers fitting our previously described recruitment criteria. After a worker accepted an invitation, they were re-directed to a landing page implemented on the Qualtrics survey platform (Qualtrics, Provo, UT) and were randomly assigned to one of two conditions: the *control* condition and the *experimental* condition, where participants selectively received upward social comparison information. Both conditions consisted of four parts: 1) a pre-screening questionnaire validating investor status and financial literacy, 2) a self-reported risk attitude measure, 3) a trading task implemented in the "Zurich Trading Simulator" (c.f., ZTS) software (Andraszewicz, Kaszás & Hölscher, 2021) and 4) a short post-task questionnaire including items regarding participants' perception of others' performance, as well as satisfaction with one's own performance on the previous trading task and a battery of basic demographics questions.

The pre-screening questionnaire consisted of 11 questions. Three of these were "filler questions" (implemented to make our relevant screener criteria less salient) and the remaining eight questions were meant to screen out people with insufficient knowledge and experience in finance (Appendix A). Three additional questions (Van Rooij, Lusardi & Alessie, 2012; FINRA, 2012) were asked after the screening was complete, in order to ascertain that passing the screener is indeed associated with higher knowledge. We have implemented the screening protocol to ensure that the individuals recruited for this experiment did not misrepresent their investor status, and possessed adequate literacy regarding finance and investments. As overclaiming (falsely stating the possession of certain skills or attributes) is a serious concern when it comes to the validity of data collected on online labor markets (Chandler & Paolacci, 2017), we posed participants a number of questions where we can detect if participants are attentive and truthful in their answers (inspired by Bentley, 2018). An example question prompts prospective participants to indicate if they currently hold or recently have held a direct investment in any of the companies listed as answer options. However, all the widely-known consumer brands listed were actually privately held at the time of data collection, making it very unlikely that a participant held a direct investment in one of these companies. Thus, selecting any option other than "None" would be a strong signal for overclaiming. Another item asks participants to identify which brokerage, app, or other trading service they use to manage their investments – this list contains a number of fictional companies, as well as companies which are not accessible to US-based retail investors. If one selects such a company, that is also an indicator for overclaiming. A pilot analysis (n= 164) reveals that participants who passed our screener exhibited significantly higher objective financial literacy than those who were screened out (as measured by two validation questions from previous literature; Van Rooij, Lusardi & Alessie, 2012; (t(141)=-7.15, p<0.001).

Once a participant passed the pre-screening questionnaire, they were asked to read a participant information sheet and sign an informed consent form (Appendix B.). If they consented to participate in the study, their self-reported attitude towards risk was elicited using the SOEP general risk item (Dohmen, Falk, Huffman, Sunde, Schupp & Wagner, 2011).

Next, participants were re-directed from Qualtrics to the trading task (Figure 1B displays the user interface of the task) which consisted of one practice round explaining functionality of the software and the task, and two experimental rounds. The practice round presented an artificial price pattern accompanied by an interactive introduction to the trading task, while the two experimental rounds included historical closing prices from the Swiss Market Index (SMI). The stimuli were selected such that they are broadly representative of the SMI's performance and volatility profile over the past 20 years, without containing extreme events which might induce bias, inform participants about the future development of the price paths, or induce wealth effects. These historical price paths were standardized to uniformly start at an arbitrary value (141.7), in a further effort to ease comprehension and avoid any participant correctly identifying the source of the price path.

Participants in the experimental condition received information about the performance of three other participants in the study (see Figure 1A). In fact, these three participants were the best performers from the first batch of data collection, which only included participants from the control condition, who were not shown comparative performance information. In each round of trading in both conditions, participants were endowed with 10 000 units of experimental currency evenly split between a risky asset (i.e., shares of the market index) and cash. After the second round of the trading task, the amount of money participants made in both rounds separately was added up, and participants were compensated for their cumulative performance. The accumulated earnings were converted from the experimental currency to USD at the exchange rate of $1 for every 14 000 in experimental currency earned, and paid out as a bonus. The choice of paying out accumulated earnings instead of incentivizing a randomly selected round was made in order to maintain comparability with previous experimental efforts with a similar design (Dijk, Holmen & Kirchler; 2013; Schoenberg & Haruvy, 2012), and in general with the existing empirical literature on mutual fund tournaments (Brown, Harlow & Starks, 1996; Pütz & Ruenzi, 2011).

After completing the trading task, participants were directed back to Qualtrics to complete the post-trading questionnaire. The post-trading questionnaire measured participants' perception of the average performance of all participants and participants' satisfaction with their own performance (Schoenberg & Haruvy, 2012). The first question served as a manipulation check, testing whether our experimental manipulation successfully shifted participants' performance reference point upwards in the experimental condition. The question regarding post-task satisfaction was included in order to replicate- and extend on Schoenberg and Haruvy's findings (2012) on the effect of upward social comparisons on investor satisfaction. The post-trading questionnaire also included demographics questions, as well as questions regarding the participant's experience with- and knowledge of investments and finance. At the end of the study, each participant was provided with their unique mTurk exit code that entitled them to receive payment and was then thanked for their participation. The median time required for completing the experiment was 15 minutes.

Results

In order to investigate the influence of upward social comparison on risk taking (*H1*), we tested the difference between our two conditions in Round 2 (i.e., following the presentation of the peer information) using an independent-samples t-test. The choice of a parametric test was pre-registered in the case of *H1*, as previous studies conducted with the same experimental paradigm found that our metric for risk seeking will likely be amenable to parametric tests. As such tests provide more power than non-parametric alternatives, the choice was made for a parametric means-comparison in this specific case. We found that participants in the experimental condition have significantly riskier portfolios ($t(771.35)=-3.77$, $p<0.001$, $d = 0.2$, Fig. 2.). Therefore, we conclude that the data supports our first hypothesis.

When pre-registering our hypotheses, the choice was made to conduct our hypothesis tests concerning trading activity using non-parametric methods (Wilcoxon rank-sum tests, similar to a Mann-Whitney U-test). This decision was motivated by findings from previous studies conducted with our experimental platform, which have shown that our trading activity variables are not amenable to parametric tests due to the presence of substantial outliers. In order to provide additional context, the results of exploratory independent-samples t-tests are also reported for the hypotheses regarding trading activity. In case of conflicting findings between parametric and non-parametric methods, the more robust estimates yielded by non-parametric methods is preferred. In addition, our effect size estimates were calculated using the *WRS2* package for R to ensure robustness to the presence of outliers (Mair & Wilcox, 2020), yielding effect size estimates which are less influenced by unequal variances between groups (Yuen, 1974).

When testing *H2a* (*Transactions*), we find no support for the hypothesis that participants in the experimental condition would trade more frequently, according to a Wilcoxon rank-sum test ($W = 75591$, $p = 0.11$) – we observe a small effect size ($d = 0.07$). The difference between conditions was marginally statistically significant according to an exploratory t-test ($t(561.33) = -1.97$, $p = 0.05$).

In a non-parametric test of hypothesis *H2b* (*Vol_Avg*) concerning the average size of transactions, we do not find a significant group difference (W = 75512, p = 0.06) and identify a small effect (d = 0.1). In an additional parametric test we document a marginally significant group difference (t(797.51) = - 2.09, p = 0.04). To summarize, we find no support for Hypothesis 2b using our pre-registered methodology.

Our key dependent variable for trading activity, the combined value of all shares bought and sold by a participant (*H2c, Vol_Sum*) was significantly higher in the experimental condition than in the control condition (*Fig. 3.*, W = 72374, p = 0.01, d = 0.13). A secondary parametric means comparison supports comes to a similar conclusion (t(692.98) = -2.55, p= 0.01). We can thus state that the data supports our key hypothesis regarding aggregate trading volume. While an effect size of 0.13 is considered to be "small" (Sawilowsky, 2009), in terms of real-world economic impact, a comparable increase in US retail volume would correspond to about 600 million additional shares traded each day (CBOE, 2021). Descriptive statistics relating to our four dependent variables are presented in *Table 1*.

Robustness Checks

In order to test whether the treatment effects observed in Round 2 were not attributable to randomization failure (where pre-existing group differences might be responsible for our observed effects), we tested group differences in risk taking and trading activity in Round 1, before the upward social comparison was induced. Three Wilcoxon tests reveal no pre-existing differences in of our dependent variables for trading activity: *Transactions* (W = 79793, p=0.85), *Vol_Avg* (W = 82020, p = 0.71) and *Vol_Sum* (W = 79730, p = 0.74). Similarly, parametric comparison for first-round risk taking between conditions does not yield a significant difference (t(790.58), p=0.73).

Also, we found no relationship between *Risk* and return on investment (*ROI*), as indicated by Pearson's correlation coefficients for Round 1: $r = -0.01$, $t = -0.41$, $df = 804$, $p = 0.68$ and Round 2: $r = -0.06$, $t = -1.72$, $df = 804$, $p = 0.08$. However, in both rounds, we found a positive relationship between *ROI* and *Vol_Sum*. This is indicated by Spearman's correlation coefficients for Round 1: $\rho = 0.52$, $S = 43027126$, $p < 0.001$ and for Round 2: $\rho = 0.45$, $S = 48247247$, $p < 0.001$, for data aggregated in both conditions. (Parametric tests were conducted in the case of investigating the *ROI* and *Risk* relationship, as both have been found to satisfy assumptions for parametric methods).

We also conducted tests to determine if there was a difference in earnings between the control- and the experimental condition, either before the experimental manipulation was implemented, or after. We found that participants' earnings in Round 1 and Round 2 were not significantly different between conditions (ROI_1 : $t = -1.40$, $df = 776.89$, $p = 0.16$, ROI_2 : $t = -0.99$, $df = 758.6$, $p = 0.32$). The effect of experimental condition as a binary variable on *Risk* and *Vol_Sum* in Round 2 remains significant in additional regression analyses, after controlling for performance in Round 1 (ROI_1) (unstandardized estimate for the effect of "*Condition*": B_{Vol_Sum}: 3506, p=0.03; β_{Risk}: 0.05, p<0.01). This analysis tells us that our observed treatment effect is not attributable to pre-existing differences between conditions.

A possible alternative explanation for the observed differences in trading activity between the two conditions (*H2c*) could be that this effect is completely driven by differences in risk taking (*H1*). This would happen for example if in the experimental condition all additional trading activity would consist exclusively of buying shares of the risky asset, in order to increase portfolio risk – in such a case, an observed increase in trading volume would be mechanistically driven by an increase in risk appetite. In order to test this, we regressed *Vol_Sum* in Round 2 on *Risk* in Round 2 and a dummy variable for experimental condition. Due to the distribution of *Vol_Sum*, we again implemented a rank-based estimation method for linear models from the *rfit* R-package (Kloke, 2012). We found that the effect of "*Condition*" on trading activity is substantially unchanged after accounting for concurrent differences in risk taking (B= 4328, p =0.02). This implies that the higher observed trading activity in the experimental condition cannot be interpreted as an artefact, or a byproduct of increased risk taking.

In additional robustness checks we include controls for the influence of gender, age, education level and self-reported willingness to take risks (Dohmen, Falk, Huffman, Sunde, Schupp & Wagner, 2011) in a regression model, and observe how the inclusion of these predictors affects our estimates for the treatment effect. We find overall that the effect of "*Condition*" does not diminish after accounting for these potential confounding variables. When it comes to a rank-based regression predicting aggregate trading volumes (*Vol_Sum*), we document an unstandardized treatment estimate of $B_{Condition}$ =4967.93 (p<0.001), indicating an even larger positive point estimate than in our previous analyses. Similarly, when it comes to one's level of risk taking (*Risk*), we also find a significant effect of experimental condition, following the inclusion of our demographic and attitudinal control variables in the model ($\beta_{\text{Condition}}$=0.049, p<0.001).

Overall, these findings indicate that the induced social comparison significantly and robustly increased risk taking and trading activity.

Manipulation Check

After the trading task, we asked participants to estimate the mean performance of all participants. This served as a manipulation check to verify that participants in the experimental condition have a higher reference point regarding typical earnings in the task. If participants in the experimental condition indicated higher average earnings than participants in the control condition, we could assume that our manipulation worked. According to a Wilcoxon rank sum test (W = 51600, p < 0.001), participants in the experimental condition perceived the average earnings of all participants expressed in experimental currency to be significantly higher than participants in the control group have ($Md_{Control}$= 500, $Md_{Experimental}$= 1000, d=0.46). This is in line with our expectation and provides further validation for the effectiveness of our experimental manipulation.

Post-task satisfaction

The question of investor satisfaction has received increased research attention over the past years (Grosshans & Zeisberger, 2018; Schwaiger, Kirchler, Lindner & Weitzel, 2020). As Statman (2017) notes, reaping financial rewards from a successful investment is not the only way people can benefit from investing. Feeling that one is successful can bring the emotional benefit of satisfaction, which can in turn influence risk preferences, and future trading decisions. Given the widespread use of marketing messages triggering upward social comparisons when it comes to retail investments, it is especially important to understand the potential negative side effects such messaging might have on individual investors' satisfaction.

In an early contribution, Schoenberg and Haruvy (2012) find that experimental participants presented with an upward social comparison report lower satisfaction with their performance, compared to other participants who have been presented with a downward social comparison. In a pre-registered additional analysis, we posed a version of Schoenberg and Haruvy's original question adapted to our experimental design (*"How do you feel about your performance in the trading task?",* Appendix E). We find that participants in the experimental condition, who were presented with the performance of highly successful peers are less satisfied with their own performance than participants in the control condition ($Me_{Control}$= 4.65, $Me_{Experimental}$=4, d=0.36, W = 1022889, p < 0.01; *Fig. 5.*). Being exposed to the good performance of others leads to lower satisfaction with one's own performance in our experiment, lending support to previous findings by Schoenberg and Haruvy (2012).

Discussion

Our experiment investigated the influence of upward social comparison on risk taking on risk taking and on trading activity. Our findings are in line with previous research documenting a relationship between upward social comparison and risk taking (see for example, Apesteguia, Oechssler & Weidenholzer, 2020; Dijk, Holmen & Kirchler, 2013, Schoenberg & Haruvy, 2012). Our findings regarding upward social comparison and aggregate trading volume complement previous research by demonstrating that in addition to risk taking, trading activity (when operationalized as aggregate trading volume) is also substantially influenced by upward social comparison. Overall, our findings help us understand some of the drivers behind the historically high level of retail trading activity we can observe on today's financial markets. In addition, these findings provide valuable context to better understand the practices proprietors of retail investment services might engage in, in order to drive trading volumes on their platforms.

Our finding regarding lower satisfaction in the experimental condition provides an interesting addition to our conclusions previously derived from the trading task – it seems that our participants responded to an upward peer comparison by engaging with the market more actively, but were ultimately less satisfied with their performance. This is especially intriguing in light of the fact that the two conditions did not differ in terms of final earnings, and one's compensation was independent of that of other participants. This finding highlights the importance for social information for not just affecting beliefs and trading behavior, but also post-task attitudes regarding one's own success.

While our decision to expose our participants to identical historical market conditions contributed to our ability to ensure all our participants have similar beliefs about the market, and has enabled us to implement a meaningful peer-performance manipulation, we cannot completely rule out the possibility that our observed effects are specific to the price paths used. It could be that our participants in the experimental condition changed their behavior in the hypothesized direction only because of a belief that if high trading activity was profitable in the first round of trading (as demonstrated by the existence of other high performing participants and a significant activity-returns correlation in Round 1), it makes sense to trade more actively in the second round as well. However, our finding that risk taking in the first round did not correlate with returns in in the same round, yet we could still observe substantial differences in risk taking in Round 2 between conditions would speak against such an alternative explanation. It is also worth noting, that while there might exist an overall positive relationship between certain trading behaviors and returns on a specific price path, it is far from clear that these relationships are apparent to the individual participant. Overall, presenting our participants with identical pre-determined price paths might limit the external validity of our findings to a degree. However, it also makes it possible for us to state that since individuals in both conditions have seen the exact same price information, our treatment effect cannot be attributed to differential information participants might have received in the two conditions regarding the properties of the risky asset itself. Nonetheless, an eventual follow-up should investigate the systematic effect various price patterns might have on trading activity.

A valid criticism of our experimental manipulation could be made along the lines of Eriksen and Kvaløy (2017): if we selectively informed only one group of participants about the existence of other, similar individuals who have already participated, that might have triggered a "tournament mindset" selectively in the experimental condition, shifting risk taking for a reason other than upward social comparison. In order to avoid this potential confound, we made a conscious effort when designing our experiment to communicate multiple times to participants in both conditions clearly and explicitly the fact that other participants exist, and one might or might not see their results, depending on one's own experimental condition.

A limitation of the current study is that we did not investigate the effect of downward social comparisons as others have done in the literature (Dijk, Holmen & Kirchler, 2013; Schoenberg & Haruvy, 2012). This was a conscious decision motivated by the fact that success in the domain of finance and investments is not only often more salient than failure, but success stories are communicated much more readily than stories of loss (Kaustia & Knüpfer, 2012; Escobar & Pedraza, 2019). Furthermore, early research on social comparison suggests an overall tendency of people for upward social comparisons in performance situations (Festinger, 1954; Wheeler, 1966). In addition, a short survey of existing social trading platforms shows the central role "successful others" play in the marketing and design of such services (Wursthorn & Choi, 2020), as well as the direction of investment flows to portfolio managers on social trading platforms (Röder & Walter, 2019) – relatively unsuccessful peers receive less attention in such contexts. While we have chosen to investigate the effect of upward comparison due to its higher prevalence and relevance in applied contexts, studies investigating the effect of downward social comparisons with a similar experimental paradigm in the future could further expand our understanding of the dynamics of retail investing.

Conclusion

We demonstrate that upward social comparisons cause active retail investors recruited from on online labor pool to invest more actively, and take on more risk on an experimental market. We draw these conclusions by employing a novel experimental paradigm, which we find to be especially well-equipped to investigate questions regarding trading activity. This aspect of our design enables us to contribute to the literature on social comparisons in finance and investments in a valuable way. Aside from a theoretical contribution, we believe our study shows the way for future studies wishing to investigate the determinants of retail investor behavior and intending to inform future regulatory efforts. While recent European regulation (MiFID II, Directive 2014/65) aims to provide guidelines for disseminating only "fair, clear and not misleading" materials to prospective investors, it does not provide explicit guidelines for presenting performance comparison information and other similar appeals either as part of marketing messages, or integrated into a software's user interface. Further experimental research on how people respond to social influences in the domain of finance and investing should inform future policy efforts to maximize consumer welfare, while allowing individual investors to participate safely in the next wave of retail financial innovation.

References

Andraszewicz, S., Kaszás, D., & Hölscher, C. (2021). Zurich Trading Simulator – a dynamic trading task for oTree. *Working Paper*

Apesteguia, J., Oechssler, J., & Weidnholzer, S. (2020). Copy Trading. *Management Science, 66*(12), 5608-5622, doi: 10.1287/mnsc.2019.3508

Baghestanian, S., Gortner, P. J., & Van der Weele, J. J. (2015). *Peer effects and risk sharing in experimental asset markets* (No. 67). SAFE Working Paper.

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, *55*(2), 773-806.

Barber, B. M., Lee, Y. T., Liu, Y. J., & Odean, T. (2009). Just how much do individual investors lose by trading? *The Review of Financial Studies*, *22*(2), 609-632.

Barber, B. M., Lee, Y. T., Liu, Y. J., & Odean, T. (2014). The cross-section of speculator skill: Evidence from day trading. *Journal of Financial Markets*, *18*, 1-24.

Bentley, J. W. (2018). Challenges with Amazon Mechanical Turk research in accounting. *Available at SSRN 2924876*.

Brown, K. C., Harlow, W. V., & Starks, L. T. (1996). Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *The Journal of Finance*, *51*(1), 85-110.

CBOE, (2021), U.S. Equities Market Volume Summary. https://www.cboe.com/us/equities/market_share/. Last accessed: 09.03.2021

Chandler, J. J., & Paolacci, G. (2017). Lie for a dime: When most prescreening responses are honest but most study participants are impostors. *Social Psychological and Personality Science*, *8*(5), 500-508

Deck, C., Hao, L., Xu, W., & Yeager, T. J. (2020). Social Comparison and Wealth Inequality in a Leveraged Asset Market. *Journal of Behavioral Finance*, 1-21.

Dijk, O., Holmen, M., & Kirchler, M. (2014). Rank matters–The impact of social competition on portfolio choice. *European Economic Review*, *66*, 97-110.

Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU (recast). Retrieved from: http://data.europa.eu/eli/dir/2014/65/2020-03-26

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522-550.

Eriksen, K. W., & Kvaløy, O. (2017). No guts, no glory: An experiment on excessive risk-taking. *Review of Finance*, *21*(3), 1327-1351.

Escobar, L., & Pedraza, A. (2019). Active trading and (poor) performance. *Development Research*.

Ferri, G., Ploner, M., & Rizzolli, M. (2016). *Count to ten before trading: Evidence on the role of deliberation in experimental financial markets* (No. wpC07). CERBE Center for Relationship Banking and Economics.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117-140, doi: 10.1177/001872675400700202

FINRA (2012) Investor Knowledge Quiz. https://www.finra.org/sites/default/files/InvestorDocument/p014495.pdf. Last accessed: 22.03.2021

Galvin, W. (2020) *Administrative complaint*. Office of the Secretary of the Commonwealth, Securities Division, https://www.sec.state.ma.us/sct/current/sctrobinhood/MSD-Robinhood-Financial-LLC-Complaint-E-2020-0047.pdf

Glaser, M., & Weber, M. (2007). Overconfidence and trading volume. *The Geneva Risk and Insurance Review*, *32*(1), 1-36.

Grinblatt, M., & Keloharju, M. (2009). Sensation seeking, overconfidence, and trading activity. *The Journal of Finance*, *64*(2), 549-578.

Grosshans, D., & Zeisberger, S. (2018). All's well that ends well? On the importance of how returns are achieved. *Journal of Banking & Finance*, *87*, 397-410.

Gruder, C. L. (1971). Determinants of social comparison choices. *Journal of Experimental Social Psychology*, *7*(5), 473-489.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*(1), 34

Kaustia, M., & Knüpfer, S. (2012). Peer performance and stock market entry. *Journal of Financial Economics*, *104*(2), 321-338.

Kloke JD, McKean JW (2012). “Rfit: Rank-based estimation for linear models” *The R Journal*, **4**(2), 57–64.

Kirchler, M., Lindner, F., & Weitzel, U. (2018). Ranking and risk-taking in the finance industry, *The Journal of Finance, 73*(5), 2271-2302, doi 10.1111/jofi.12701

Lejarraga, T., Woike, J.K., & Hertwig, R. (2016). Description and experience: How experimental investors learn about booms and busts affects their financial risk taking. *Cognition*, 157, 365-383, doi: 10.1016/j.cognition.2016.10.001

Lewis, M. (2001, February 25) Jonathan Lebed's Extracurricular Activities. *New York Times*. Retrieved from: https://www.nytimes.com/2001/02/25/magazine/jonathan-lebed-s-extracurricular-activities.html

Litman, L., Robinson, J., & Abberbock, T. (2016). TurkPrime. com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods,* 1-10.

Mair P, Wilcox R (2020). “Robust Statistical Methods in R Using the WRS2 Package.” *Behavior Research Methods*, *52*, 464–488.

Massa, A., McDonald, M., Alexander, S. (2020, December 16.) *Robinhood is accused of "gamification" by Massachussetts*. Bloomberg.com, https://www.bloomberg.com/news/articles/2020-12-16/robinhood-accused-of-gamification-by-massachusetts-regulator

McCrank(2021, January 21) Factbox: The U.S. retail trading frenzy in numbers. *Reuters*. https://www.reuters.com/article/us-retail-trading-numbers-idUSKBN29Y2PW

Merkle, C. (2017). Financial overconfidence over time: Foresight, hindsight, and insight of investors. *Journal of Banking & Finance*, *84*, 68-87.

Nicodemus, A. (2020, December 17.) FinTech darling Robinhood fined $65M for misleading customers. *ComplianceWeek*. Retrieved from: https://www.complianceweek.com/regulatory-enforcement/fintech-darling-robinhood-fined-65m-for-misleading-customers/29852.article

Odean, T. (1999). Do investors trade too much?. *American Economic Review*, *89*(5), 1279-1298, doi: 10.1257/aer.89.5.1279

Oehler, A., Horn, M., & Wendt, S. (2016). Benefits from social trading? Empirical evidence for certificates on wikifolios. *International Review of Financial Analysis, 46*, 202-210, doi: 10.1016/j.irfa.2016.05.007

Papadovasilaki, D., Guerrero, F., Sundali, J., & Stone, G. (2015). How important are early investment experiences on subsequent investment decisions? A laboratory experiment on asset allocation. *Managerial Finance* . 41(6), 582-590, doi: 10.1108/MF-09-2014-0246

Pisani, B. (2021, February 18) Robinhood-GameStop hearing will scrutinize how brokerages get paid for trades. *CNBC.com*. https://www.cnbc.com/2021/02/18/payment-for-order-flow-the-controversial-wall-street-practice-to-draw-scrutiny-at-robinhood-hearing.html

Puetz, A., & Ruenzi, S. (2011). Overconfidence among professional investors: Evidence from mutual fund managers. *Journal of Business Finance & Accounting*, *38*(5-6), 684-712.

Röder, F., & Walter, A. (2019). What drives investment flows into social trading portfolios?. *Journal of Financial Research*, *42*(2), 383-411.

Ryu, D. (2012). The profitability of day trading: An empirical study using high-quality data. *Investment Analysts Journal,* 41(*75*), 43-54.

Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, *8*(2), 26.

Schoenberg, E. J., & Haruvy, E. (2012). Relative performance information in asset markets: An experimental approach. *Journal of Economic Psychology*, *33*(6), 1143-1155.

Schwaiger, R., Kirchler, M., Lindner, F., & Weitzel, U. (2020). Determinants of investor expectations and satisfaction. A study with financial professionals. *Journal of Economic Dynamics and Control*, *110*, 103675.

SEC, Office of Compliance Inspections and Examinations (2000, December) *Special Study: Payment for Order Flow and Internalization in the Options Markets.* Securities and Exchange Commission. Retrieved from: https://www.sec.gov/news/studies/ordpay.htm

Shiller, R. J. (1984). Stock prices and social dynamics. *Brookings papers on economic activity*, *1984*(2), 457-510.

SocialTradingGuru (2021, March 16*). List and overview of social Trading Networks* Retrieved March 25, 2021, from https://socialtradingguru.com/networks

Statman, M. (2017). Financial advertising in the second generation of behavioral finance. *Journal of Behavioral Finance*, *18*(4), 470-477.

Van Rooij, M. C., Lusardi, A., & Alessie, R. J. (2012). Financial literacy, retirement planning and household wealth. *The Economic Journal*, *122*(560), 449-478.

Wheeler, L. (1966). Motivation as a determinant of upward comparison. *Journal of Experimental Social Psychology*, *1*, 27-31.

Wheeler, L., & Miyake, K. (1992). Social comparison in everyday life. *Journal of Personality and Social psychology*, *62*(5), 760.

Wursthorn, M., Choi, E. (2020, August 20) Does Robinhood Make It Too Easy to Trade? From Free Stocks to Confetti. *Wall Street Journal.* Retrieved from:

https://www.wsj.com/articles/confetti-free-stocks-does-robinhoods-design-make-trading-too-easy-11597915801

Yuen, K. K. (1974). The two-sample trimmed t for unequal population variances. *Biometrika*, *61*(1), 165-170.

Tables

Variable	Mean		Median		Min		Max	
	C	**E**	**C**	**E**	**C**	**E**	**C**	**E**
Risk	0.39	0.44	0.37	0.42	0.01	0.04	0.99	0.99
Transactions	21.53	25.00	18	20	0	0	159	453
Vol_Avg	13.92	14.67	15.67	16.13	0	0	20	20
Vol_Sum	41081	47860	34644	42974	0	0	321102	385437

Table 1. *Descriptive statistics by experimental condition (C for Control and E for Experimental). Interpretation: Risk = Mean portfolio risk, Transactions =Aggregate number of transactions executed, Vol_Avg = Average size of transactions, Vol_Sum = Aggregate value of all shares traded.*

Figures

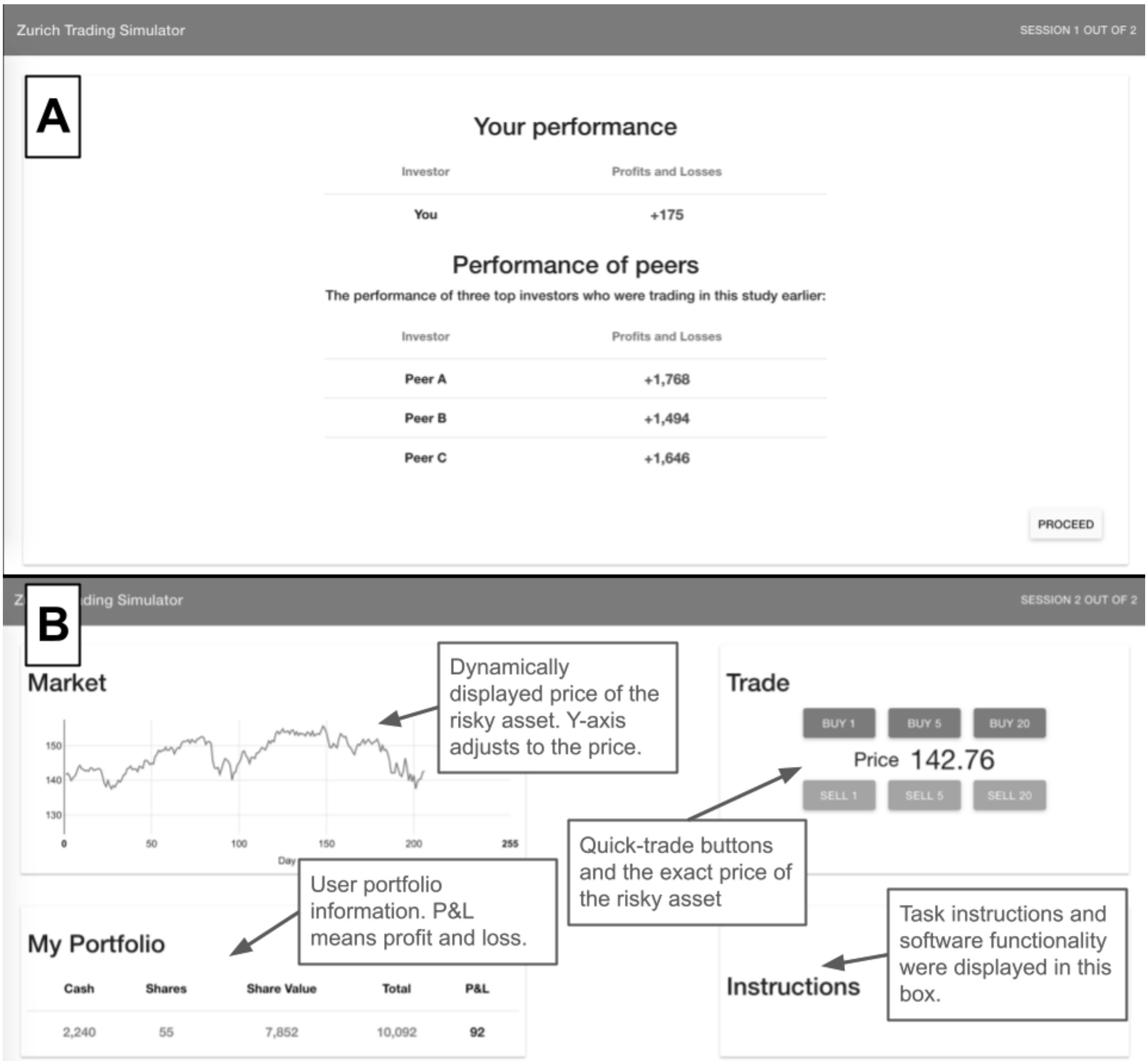


Figure 1. A: Screenshot of the social comparison information provided to participants in the experimental condition at the end of Round 1. The information contained the actual performance of the three top participants from the first batch of data collection. B: User interface of the trading task. The text in the text boxes describes the functionality of the software features.

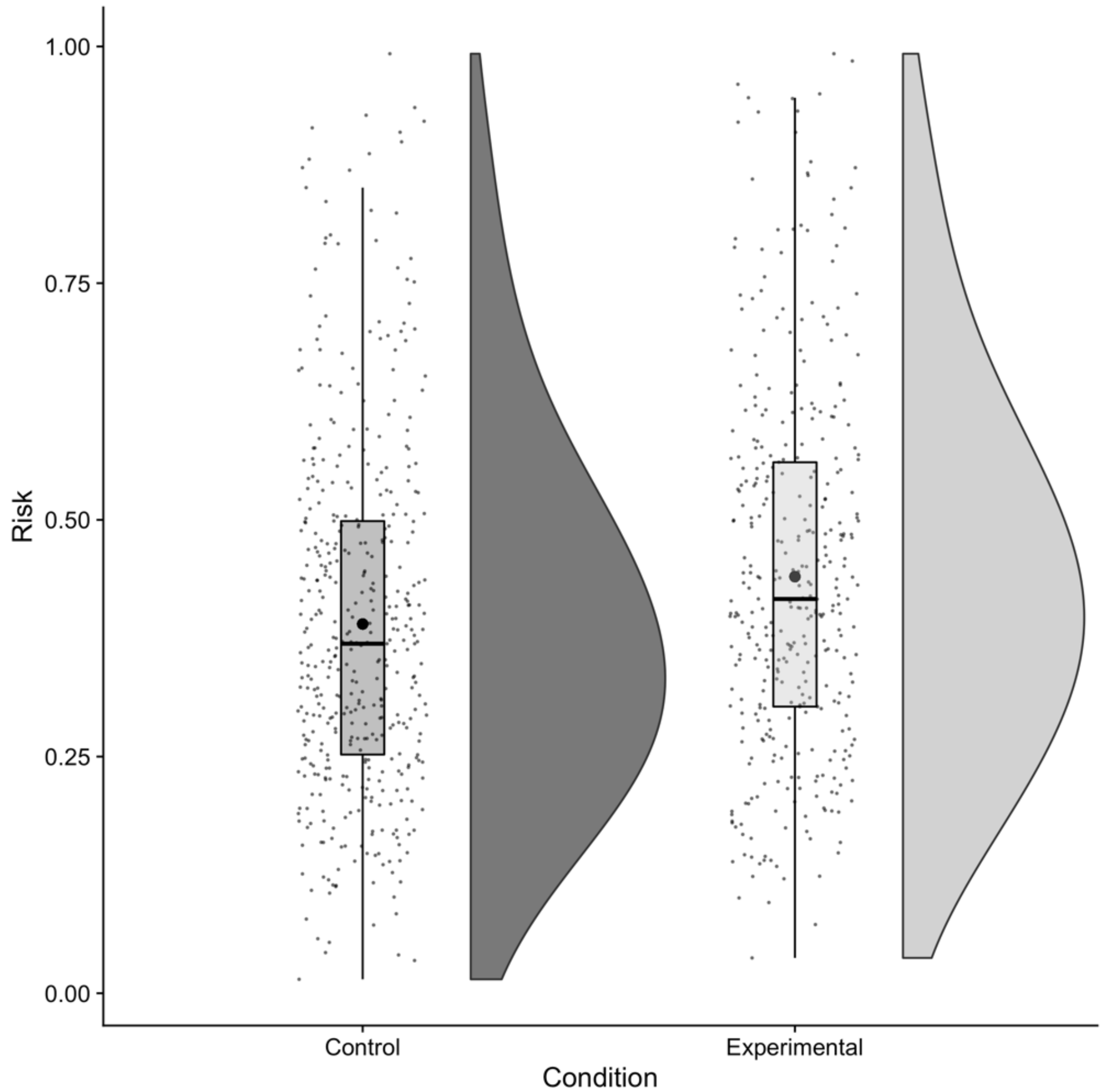


Figure 2. The distribution of portfolio risk (*Risk*) between conditions. Horizontal bars indicate the group median, while larger colored dots indicate group means. Participants in the control condition exhibit substantially lower risk taking following the experimental manipulation.

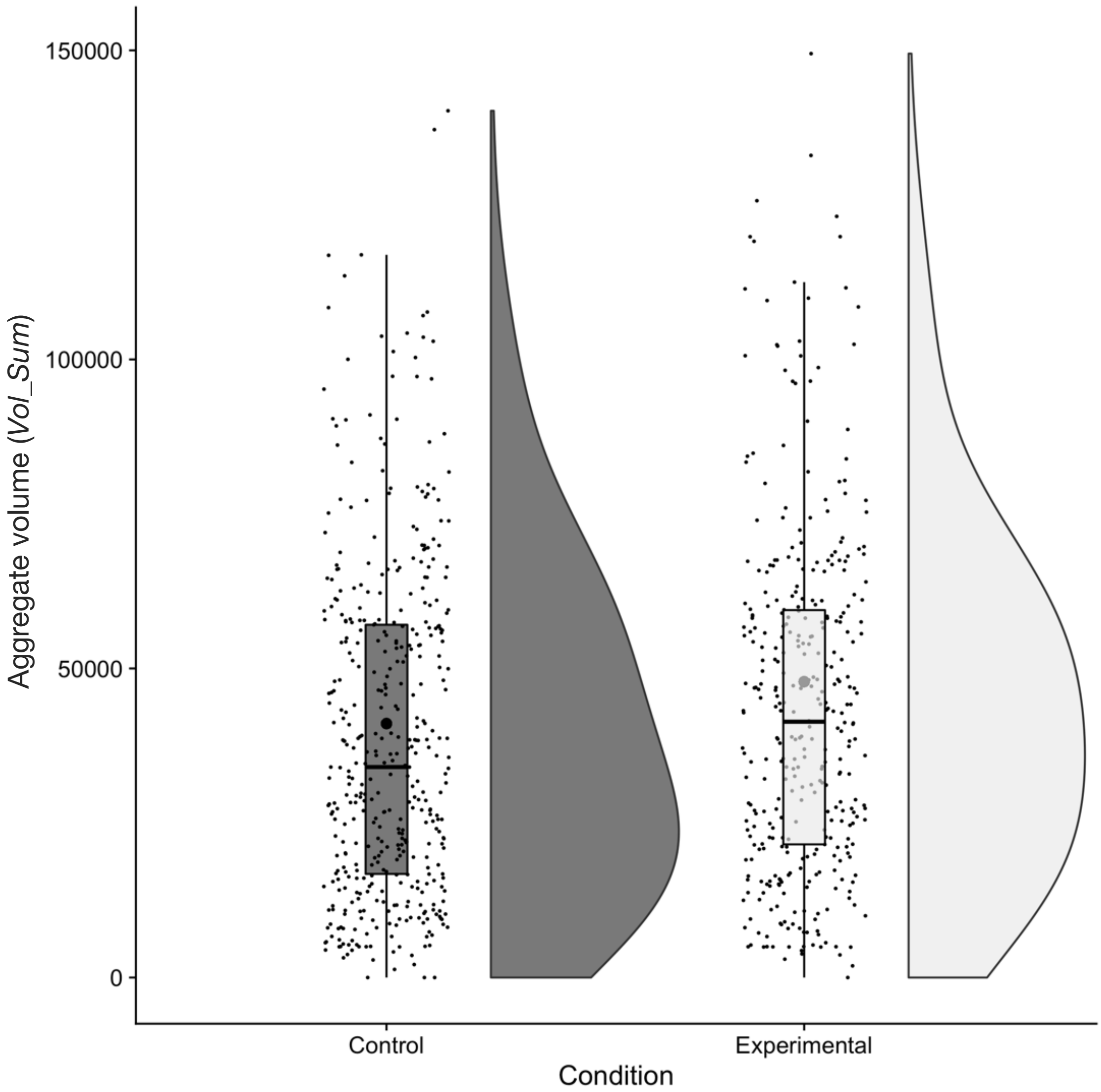


Figure 3. Distribution of trading activity (*Vol_Sum*, aggregate value of shares traded) following the experimental manipulation, by condition. Horizontal bars indicate group medians, while larger colored dots indicate group means. Due to the presence of outliers the y-axis is bounded at 150 000 Experimental Currency Units.

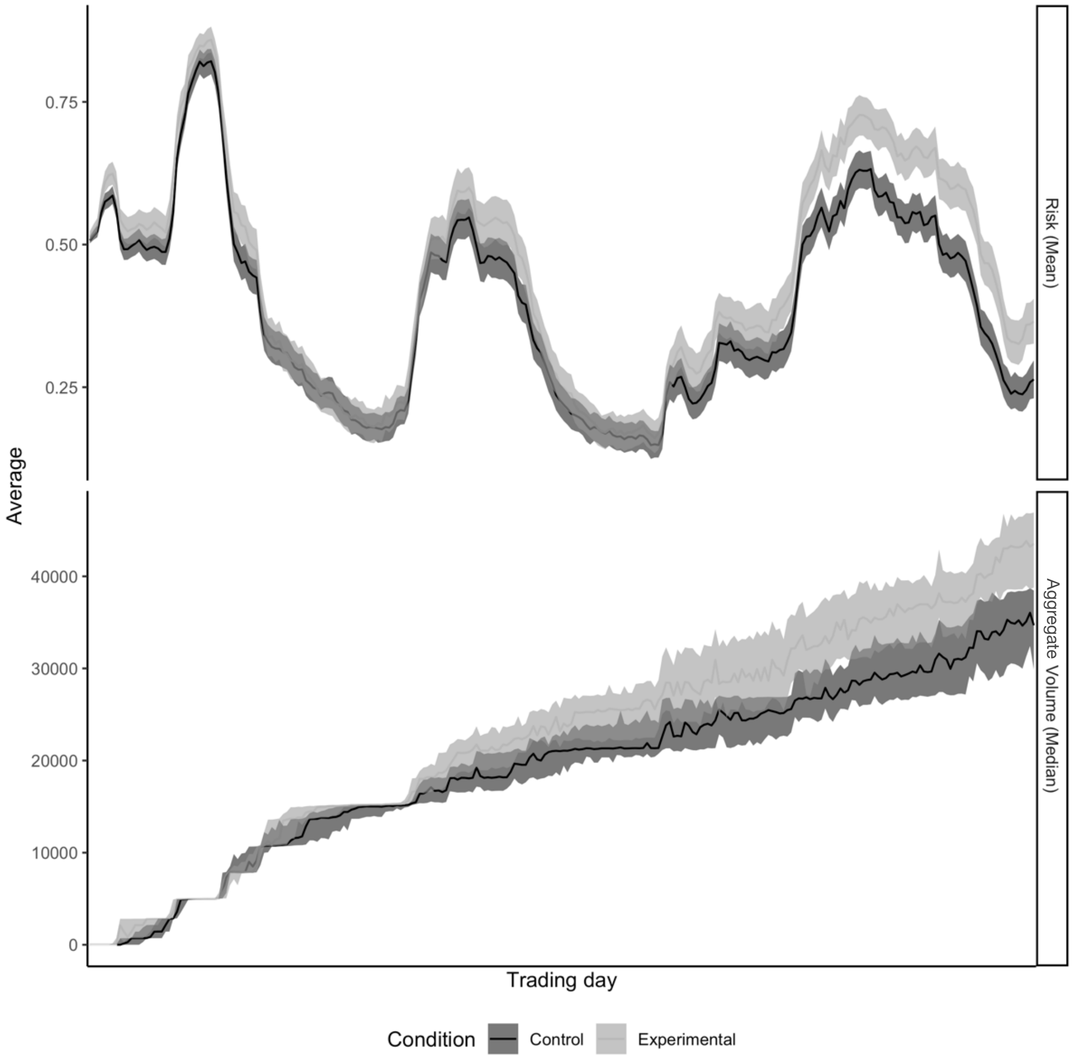


Figure 4. The development of mean portfolio risk (*Risk*) and median volume of transactions (*Vol_Sum*, cumulative) through each trading day, aggregated across participants in the control condition and experimental condition. The solid line corresponds to the bootstrapped average for each measure. The shaded area around the solid line corresponds to the 95% bootstrapped confidence interval.

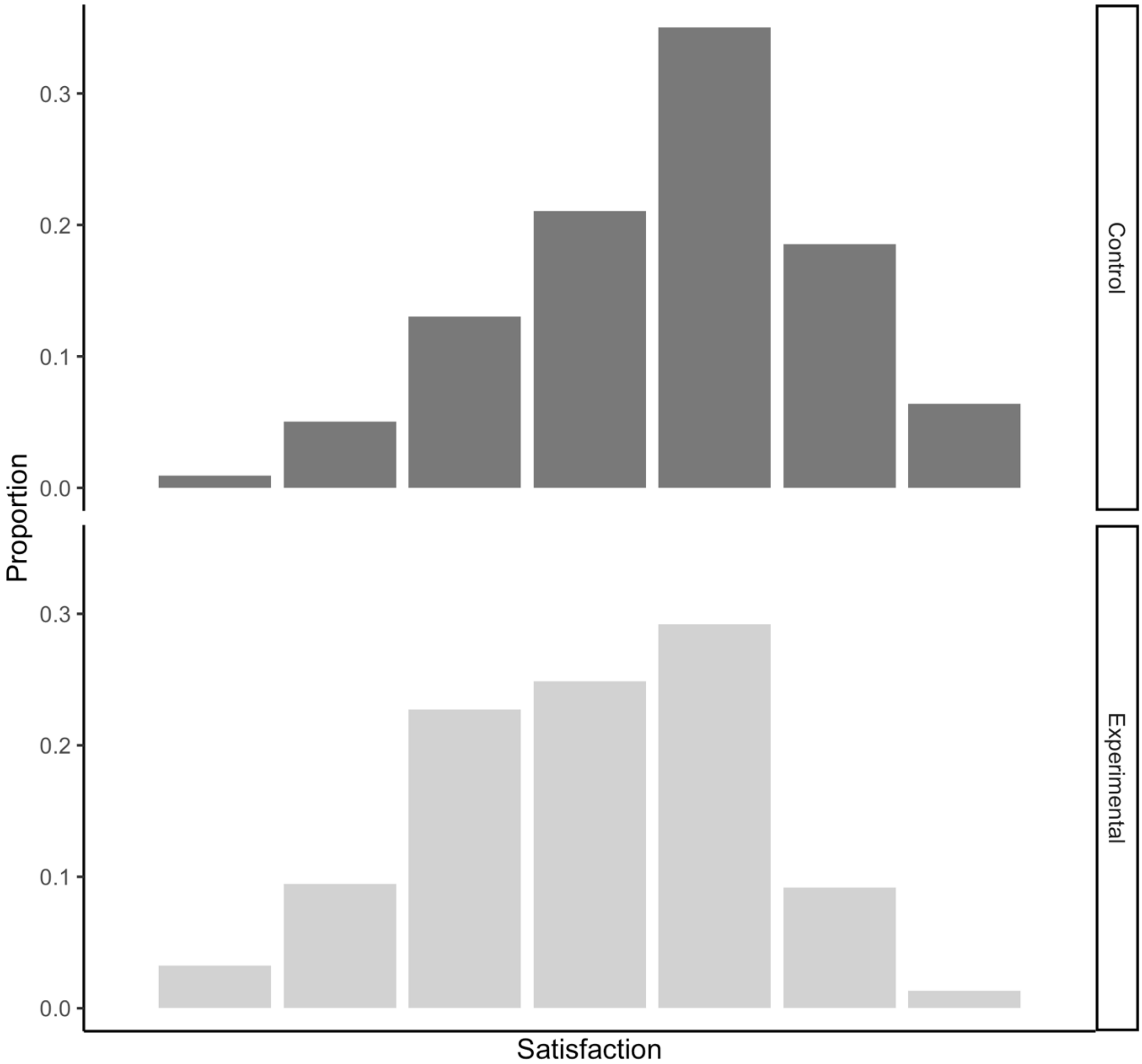


Figure 5. Histograms depicting the proportion of responses to the post-task satisfaction question, from 1 (Not satisfied at all) to 7 (Very satisfied). We observe a significant mean difference between conditions, mostly due to deviations at the lower- and upper ends of the distribution.

Appendix

Experimental Materials

Listed in the order as they were presented to participants

A: ***Screening Questionnaire*** *with correct answers included in [the square brackets]. Filler questions marked with (F), the pre-screening questions are marked with (S), while the financial literacy are marked with (L).*

Filler questions are not used for the pre-screening and any answer in the filler questions is correct. The point of the pre-screening questions was to screen out participants who are not actively investing in the stock market. The point of the financial literacy questions

The following questions were presented in a randomized order. In order to pass the pre-screening questionnaire, a participant had to correctly respond to all pre-screening and financial literacy questions.

1. How many hours a week do you exercise? (F) [all answers are correct]
    a. I generally don't exercise
    b. 20 minutes to 1 hour
    c. 1 hour – 2 hours
    d. 2 hours – 3 hours
    e. More than 3 hours
    f. I'd prefer not to say

2. Which of the following best describes your current status? (F) [all answers are correct]
    a. Student
    b. Employee
    c. Home maker
    d. Business owner
    e. Unemployed
    f. Retired
    g. Other
    h. I'd prefer not to say

3. Are you registered to vote? (F)
    a. Yes
    b. No

4. Do you currently hold a US passport? (S)
    a. Yes
    b. No

5. Do you personally invest in the stock market? (S) [“a” is the correct answer]
    a. Yes
    b. No

6. Do you currently hold an investment in the stock market? (S) [“a” is the correct answer]
    a. Yes
    b. No

7. Please, name one asset that is currently in your investment portfolio! (S)

{open question}

8. Please, tell us how your {asset name entered in question 4} investment performed over the past month in percent terms (rounded to the closest 10%) (S) [all answers between e and m, are coded as “correct”]
    a. < -70%
    b. -70%
    c. -60%
    d. -50%
    e. -40%
    f. -30%
    g. -20%
    h. -10%
    i. 0%
    j. 10%
    k. 20%
    l. 30%
    m. 40%
    n. 50%
    o. 60%
    p. 70%
    q. <70%

9. Which of the following companies do you own stock in directly, or have done so in the past 3 months? (S) [“j” is the correct answer]
    a. Panera Bread
    b. Mars
    c. State Farm
    d. Liberty Mutual
    e. Pilot Travel Centers
    f. Mass Mutual
    g. Penske
    h. Publix

   i. Albertsons
   j. None of the above

10. Please select which brokerage service provider you use to invest in stocks! (multiple answers are possible) (S) [“a” or “g” are the correct answers]
    a. Interactive Brokers
    b. TradeFox
    c. E-Stox
    d. Forex.com
    e. Wall Street Associates
    f. Comdirect
    g. Other

11. Please name the brokerage service that you are using to invest in stocks! (S) [any answer is correct]

    {open question}

12. What has been the average annual return on a very broad US stock index market investment over the last two decades, in percentages? (S) [any value in the range from 5 to 10 is correct]

    {open question}

13. Which of the following statements describes the main function of the stock market? (S) [“c” is the correct answer]
    a. The stock market helps predict stock earnings
    b. The stock markets result in an increase in the price of stocks
    c. The stock market brings people who want to buy stocks together with people who want to sell stocks
    d. None of the above

14. Which of the following statements is correct? If somebody buys a bond of firm B: (S) [b is the correct answer]
    a. He owns a part of firm B
    b. He has lent money to firm B
    c. He is liable for firm B’s debts
    d. None of the above

15. Please complete the Captcha below to show that you are not a robot! (S) [the captcha is automatically generated by Qualtrics]

B: ***General study information for prospective participants and informed consent***

A study on how individuals behave on a simulated trading platform

The goal of the present study is to investigate how individual investors manage their investments.

The study consists of a trading task, followed by a short questionnaire.

You will receive 10 000 Experimental Currency at the beginning of each round of a realistic trading simulation.

Your performance in each of the two rounds will be added up, and paid out in real money.

For every 14 000 in Experimental Currency you earn, you receive $1 in real money.

The more you make in the task, the more real money you earn.

A person earning 0% in both rounds can expect to earn a bonus of $1.43.

You will also receive a guaranteed base payment of $0.30 for having participated, regardless of your performance.

The median expected completion time of the study is **14 minutes**.

If you encounter problems submitting this HIT, please email dKaszás@ethz.ch.

Further terms and information:

It will not be possible to connect your identity to the information you provide us in your survey answers.

Only the responsible investigators and/or the members of the Ethics Commission will have access to this anonymous data under strictly observed rules of confidentiality.

Please be aware that participating in this study does not present any known risks to participants, physical, psychological or otherwise.

You are not obliged to complete the study, and can quit anytime without having to justify it to the requester.

However, due to the design of our study, only participants who have completed the whole study can claim their compensation.

Possible damage to your health, which is directly related to the study and demonstrably the fault of ETH Zurich, is covered by the general liability insurance of ETH Zurich (Insurance Policy No. 30/4.078.362 of the Basler Versicherung AG).

Beyond the aforementioned conditions, health insurance and accident insurance is the responsibility of the participant.

The study is financed by the Dr. Donald C. Cooper-Fonds.

Consent form

I have been informed about the aims and procedures of the study, the advantages and disadvantages, as well as potential risks of participating.

I have read and understood the information sheet for volunteers (included above).

I was given sufficient time to make a decision about participating in the study.

I agree that the responsible investigators and/or the members of the Ethics Committee of ETH

Zurich have access to the original data under strictly observed rules of confidentiality.

I participate in this study on a voluntary basis and can withdraw from the study at any time.

I recognize that participants who exit the study before its conclusion forfeit their right to compensation.

By clicking "Proceed with the study", I agree to all of the statements above, and agree to participate in the study:

a. Proceed with the study
b. Exit study

C: ***Self-reported risk attitudes***

How do you see yourself? Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

1. Not at all willing
2.
3.
4. Moderately willing
5.
6.
7. Very willing

D: ***Instructions to the ZTS trading task***

Instructions

The upcoming trading task will be conducted using a browser-based trading interface (see the image below for an illustration).

The task consists of three rounds: one practice, and two main rounds.

Your combined performance in the two main rounds determines your compensation.

At the beginning of each round you will be endowed with an initial portfolio consisting of shares of a risky asset (Shares) and a safe asset (Cash), worth a combined 10 000 currency units.

50% of this endowment will be in the form of the risky asset (Shares), and 50% will consist of the safe asset (Cash).

In each round, you can buy or sell Shares by clicking ”*Buy*” or “*Sell*” buttons of the corresponding size.

Shares are characterized by higher volatility than Cash.

The prices are historical daily closing prices of a real-world market index.

Prices are predetermined, and your actions do not have an influence on market prices.

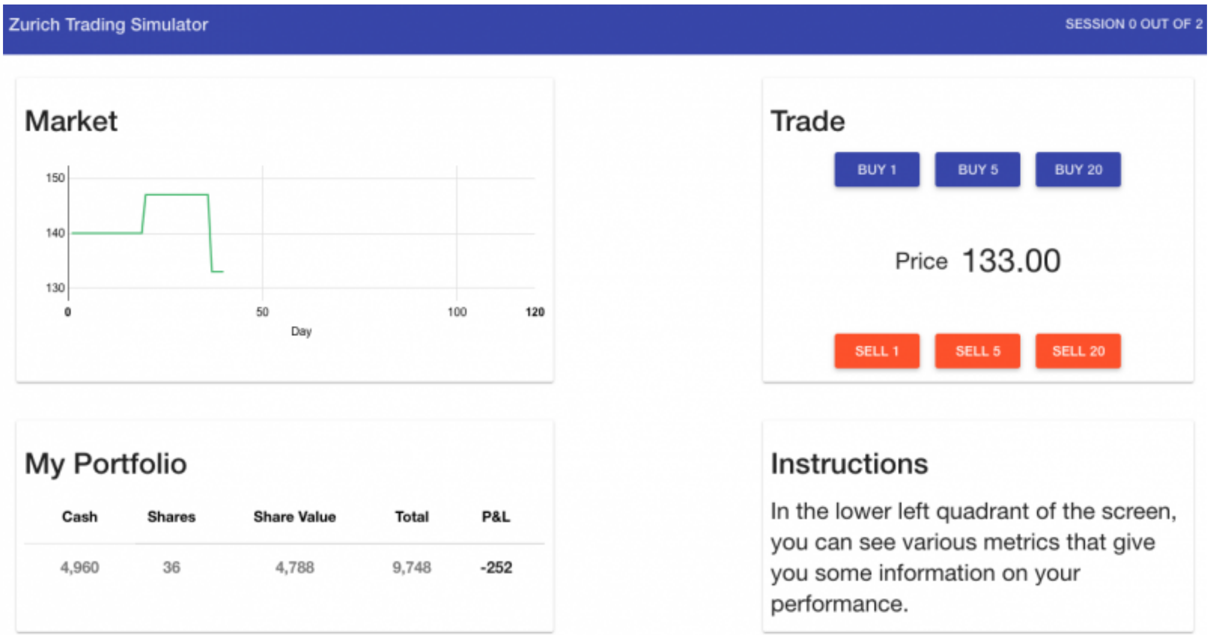
Any trade that you make will be executed instantly at the current market price.

You can buy or sell as many shares as you want, under the condition that you have sufficient funds for the transaction.

No short-selling is possible.

The y-axis of the price chart adjusts automatically to the current market price.

These changes only reflect the current price, and are not informative for future price changes.



Other people also participate in this study, but you do not interact with them.

Depending on which study condition you are in, you might see information on how some of them have performed.

G: *Explanation of the ZTS task*

This is the practice round.

Its goal is to familiarize you with the trading software.

Your performance in this round does not have an influence on your compensation.

Please pay attention to the instructions appearing in the box in the lower right corner!

The following information was displayed in the "Instructions" box

1. In order to achieve the best performance possible in the upcoming experimental round, pay attention to the instructions.

2. In order to make trades, you can use the six buttons above to buy or sell different amounts of stock.
3. The current price of the asset will be presented between the two rows of buy and sell buttons.
4. In the lower left quadrant of the screen, you can see various metrics that give you some information on your performance.
5. The first metric is Cash. It tells you how much liquidity you currently have that you can use to buy shares.
6. The second metric is Shares. It tells you how many shares are currently in your portfolio.
7. The third metric is Share Value. It tells you how much the shares in your portfolio are currently worth.
8. The fourth metric is Total. It is the combined value of your Cash and Shares.
9. The fifth metric is Profits and Losses. It tells you how much money you have made or lost, compared to the amount of money you started the round with.
10. Try buying and then selling 20 shares to demonstrate you have understood the instructions.
11. On the screen following the first round, you can find an overview of how you performed.

This is the first round.

Your performance from now on will determine your compensation.

There are other people participating in this study.

After this round, you will receive performance feedback.

You will either see only your performance, or your performance and the performance of others.

E: ***Post-trading task questionnaire***

1. Please, give us an estimate for how much money you think other participants earned in the **first main round** on average! (in P&L term, relative to the starting portfolio value of 10 000) The mean participant earned…

   {open question}

   Experimental Currency Units

2. How do you feel about your performance in the trading task? (Overall, across both main rounds)

    a. 1 - Very negatively
    b. 2
    c. 3
    d. 4 – Neutral
    e. 5
    f. 6
    g. 7 – Very positively
3. Gender
    a. Male
    b. Female
    c. Other
4. Age

    {Open question}

5. Highest level of education completed:
    a. Less than high school diploma
    b. High school diploma or GED
    c. Some college, but no degree
    d. Associates degree
    e. Bachelor's degree
    f. Master's degree
    g. Professional degree
    h. Doctorate
    i. Other: {Open question}
6. How would you judge your own level of experience in investing?
    a. 1 - No experience

b. 2

c. 3

d. 4 – Moderate experience

e. 5

f. 6

g. 7 – Very experienced

7. How would you judge your knowledge and understanding of finance and investments?

a. 1 – No knowledge and understanding

b. 2

c. 3

d. 4 – Moderate

e. 5

f. 6

g. 7 – Very high knowledge and understanding

8. Please, share any observations you have regarding the study! (optional)

{Open question}

F: ***Final information provided to participants***

Checkout

You have finished the study. Thank you for taking the time! In order to receive your payment you must copy and paste the following code back to Amazon Mechanical Turk:

{9-digit mTurk code}

Your payment will be processed within the next 48 hours. If you encounter problems submitting this HIT, please contact dkaszas@ethz.ch and report the problem there.

Thank you!

G: *Price patterns displayed in the ZTS trading task*

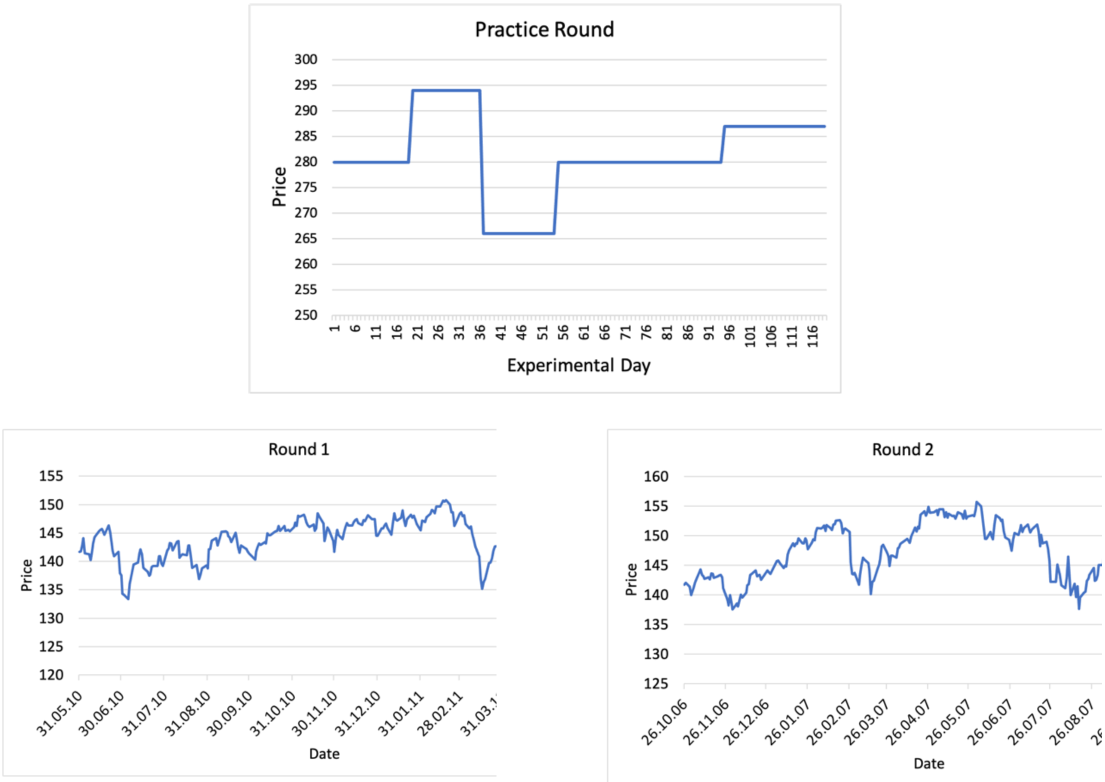


Figure A1. Price charts used in the practice round, the first and the second experimental round. The price data for the practice round were artificially generated such that they do not prime participants with any price patterns. Data in experimental rounds 1 and 2 are historical closing prices from the Swiss Market Index (SMI) 31.05.2010 – 18.05.2011 and 21.10.2006 – 26.10.2007.

Study 2: Slicker Than Your Average – Relative Overconfidence and Trading Behavior

Dániel Kaszás

Chair of Cognitive Science

ETH Zurich

Correspondence address:

Dániel Kaszás
RZ E 12.1
Clausiusstrasse 59
8092 Zürich
Switzerland

dkaszas@ethz.ch

Abstract

Previous empirical research documents a relationship between relative overconfidence and risk taking and trading activity in retail investors. In two pre-registered studies we investigate the relationship between overplacement regarding performance on a finance and investments quiz, and retail investors' trading behavior on a simulated trading platform. In a first, correlational design we find no support for a relationship between overconfidence when operationalized as overplacement and trading behavior. In a second, experimental study we successfully manipulate active retail investors' self-perceived knowledge of finance and investment in order to induce overplacement. We fail to find support for a causal effect of overplacement on trading behavior in the subsequent trading task. In an additional analysis we document a positive effect of reporting having performed "better-than-average" on risk taking in our trading task, in line with previous literature. This project makes a crucial contribution in that it is the first to experimentally manipulate overplacement and observe its effect on trading behavior in a realistic setting, with knowledgeable retail investors.

Keywords: overconfidence, overplacement, trading activity, risk taking

The importance of overconfidence in the context of judgment and decision making is hard to overstate. Kahneman has alternatively referred to it as "the most pervasive bias" (Shariatmadari, 2015), and "the most significant cognitive bias" (Kahneman, 2011). DeBondt and Thaler (1995) state that it is "one of the most robust findings in the psychology of judgment", while according to Statman, Thorley and Vorkink (2006, p. 1531), "overconfidence in one's own talent is a pervasive behavioral norm".

A general definition of overconfidence is provided by Moore and Dev (2018, p.1): "to display overconfidence is to be more confident than one deserves to be". Put another way, overconfidence is the phenomenon when one overestimates one's skill, knowledge or performance in a certain domain, either in absolute or relative terms. Classic examples of overconfidence (specifically in the form of the so-called better-than-average effect) include Svenson's finding (1981), that 80% of drivers believe they have above average driving skills, or Mele's finding (1997) that 94% of academics consider themselves above average in their field of research. Shefrin (2000, p. 48) in his survey of behavioral finance notes: "Inexperienced investors are more confident that they will beat the market than are experienced investors. Given the difficulty that many investors actually have beating the market, novice investors may be not just confident, but overconfident". Such unwarranted confidence often has significant real-world consequences, including excess entry into entrepreneurship (Cooper, Woo & Dunkelberg, 1988; Camerer & Lovallo, 1999; Moore & Cain, 2007) corporate investment distortions (Malmendier & Tate, 2005) or diagnostic decisions by physicians (Berner & Graber, 2008; Saposnik, Redelmeier, Ruff & Tobler, 2016).

Following Moore and Healy's widely accepted classification (Moore & Healy, 2008; Healy & Moore, 2007), there are three major facets of overconfidence: *overprecision, overestimation* and *overplacement*. Overprecision is defined as "an excessive faith in the quality of our judgment" (Moore & Tenney, 2015, p. 4.): This type of overconfidence can also be referred to as miscalibration, where one is overly certain regarding the accuracy of his or her beliefs. According to Moore and Healy (2008), about 31% of all empirical studies on overconfidence operationalize it as overprecision. It is operationalized most frequently via the "interval production task" (Langnickel & Zeisberger, 2016). This facet (and operationalization) of overconfidence has been mostly examined in the past in the experimental finance literature to investigate the relationship between overconfidence and risk taking and trading activity in an investment context, with mixed success (e.g., Deaves, Lüders & Luo, 2008; Glaser & Weber, 2003; Phan, Rieger & Wang, 2018; Fenton-O'Creevy, Nicholson, Soane & Willman, 2003). The second facet of overconfidence is overestimation, where individuals overestimate their "actual ability, performance, level of control or chance of success" (Moore & Healy, 2008, p. 3). The most common way to operationalize overestimation is to compare one's estimated results on a general knowledge test with one's actual results. According to Moore and Healy (2008), 64% of empirical papers published by 2008 examined overconfidence in the form of overestimation.

Moore and Healy's third facet of overconfidence is overplacement (Moore & Healy, 2008) – a sense that one is superior in some way when compared to others, while this better performance, knowledge etc. does not actually correspond to reality. That people have overly positive impressions of themselves relative to others (both in terms of knowledge, skill, previous performance or future performance), has been demonstrated in a number of different contexts (Svenson, 1981, Taylor & Brown, 1998, Klar & Giladi, 1997; Cooper, Woo & Dunkelberg, 1988). This so-called better-than-average (BTA) effect has been critized from a statistical and Bayesian learning perspective (Larrick, Burson & Soll, 2007; Benoit, Dubra & Moore, 2015; Benoit & Dubra, 2011; Merkle & Weber, 2011). While it is an apt measure to describe a population – if more than half is better than average, there is overconfidence in the sample – it cannot accurately characterize individual-level beliefs on relative placement in a population. To address these issues, Larrick, Burson and Soll (2007) introduced the concept of overplacement. By comparing a person's perceived percentile (or rank) in a population completing a specific task to his or her actual population percentile or rank, one can derive an individual measure of relative over- or underconfidence. Larrick, Burson & Soll (2007) demonstrate that both BTA beliefs and overplacement correlate with measures of overestimation in a number of different experiments. It is worth noting that according to Moore and Healy (2008), only 5% of empirical overconfidence studies looked at relative overconfidence as of 2008. Compared to the other two aspects already presented, overplacement is an under-researched form of overconfidence which has the potential to contribute to our understanding of behavior in a number of different domains.

Overall, the three aspects of overconfidence seem to be interrelated, but empirically distinct facets of a general construct. Chapman, Dean, Ortoleva, Snowberg and Camerer (2018) report that when clustering 21 behavioral or cognitive biases and preferences, they can find six distinct factors. The three aspects of overconfidence, as defined by Moore and Healy (2008) add up to one factor, with the three facets showing moderately high intercorrelations (coefficients between .49 and .62). Stango, Yoong and Zinman come to a similar conclusion (2017).

While its effects are present in many domains of life, overconfidence has probably received the most research interest in the context of financial decisions, specifically as a construct to explain trading behavior. It is the second most-researched behavioral bias in the investment and decision-making domain, following the disposition effect (Kumar & Goyal, 2015). Explaining historically higher trading volumes on financial markets than what would be justified just by liquidity needs alone has been the topic of decades of research, and most explanations have converged on overconfidence as overprecision as the main driver of trading volume (DeBondt & Thaler, 1995; Daniel, Hirshleifer & Subrahmanyam, 1998; Odean, 1998; Luo, Subrahmanyam & Titman, 2018). In his chapter reviewing psychology-based models of trading volume, Barberis (2018) comments that models of trading volume that incorporate overconfidence predict substantial volumes, similar to what can be observed on real markets. Most theories explaining the effect of overconfidence on trading activity rely on the overprecision facet of overconfidence: investors overestimate their ability to value securities accurately, leading to an underestimation of the potential error in their forecasts (Daniel, Hirshleifer & Subrahmanyam, 1998). Systematically overestimating the value of one's own information, and underestimating the precision of the information of other market participants (Eyster, 2018; Langnickel, 2018) is a recipe for disagreement on the real value of an asset between investors, generating trading volume.

The theoretical potential of overconfidence to explain trading volumes has led to great interest in empirical research to investigate the effect of specific aspects of overconfidence on trading activity, both on the individual- and market level. Past research has approached the question in two ways: First, by utilizing observational data and using diverse approaches to identify or proxy overconfidence in real market participants. Second, using experimental data mostly collected via experimental asset markets and framed decision-making tasks. However, most of these contributions don't attempt to experimentally manipulate specific aspects of overconfidence (with the notable exception of Deaves, Lüders & Luo, 2008 and Bregu, 2020) to establish the presence of a causal relationship, the empirical gold standard in the social sciences (List, 2011).

Building on Odean's model of how overconfidence as overprecision leads to higher trading volumes (1998), Barber and Odean (2001) utilized the finding that gender differences in overconfidence are the strongest in domains and tasks which are stereotypically considered masculine. They use gender as a proxy for overconfidence, and partition their brokerage dataset according to gender. They find that men trade 45% more actively, and this significantly reduces their net returns by about 1% compared to women. While gender is a very rough proxy for overconfidence, this study makes it clear that there is considerable inter-individual variability in trading activity. A closer examination of the role of overconfidence in trading activity and risk taking in the financial domain was conducted by Grinblatt and Keloharju (2009). Using data from Finnish investors who have completed a standard psychological assessment upon entering mandatory military service, they find that their measure of relative overconfidence (which they characterize as being "far closer to better-than-average effect than to a miscalibration bias", p. 553) is a significant predictor of a dummy variable for trading at all, and the number of trades executed, while it is not significantly related to turnover. They document an interesting trend where the most overconfident and most underconfident individuals underperform.

Using data from a large-scale panel of online brokerage clients, Merkle (2017) investigated the relationship between overplacement, overprecision and overestimation. He finds that overplacement (measured by comparing one's own predicted portfolio returns to predicted market returns, essentially a better-than-average expectation) robustly predicts the number of trades, the overall volume, and portfolio turnover, while overprecision doesn't. As opposed to expectations based on earlier literature, overprecision predicted a lack of diversification and higher risk taking, but did not have a relationship with trading activity. In a similar design, Glaser and Weber (2007) surveyed 3079 individual investors to estimate the effect of different manifestations of overconfidence (as overprecision and the better-than-average effect) on trading activity. They find that while overprecision is not predictive, individuals who report possessing above-average skill or performance do trade significantly more. Dorn and Huberman (2005) describe similar findings, with individuals reporting a high better-than-average belief exhibiting more portfolio churn. To summarize the state of research on overconfidence and trading activity based on observational data: While overprecision is proposed as the mechanism in the theoretical literature, relative overconfidence (most often operationalized as a better-than-average belief) seems to be a better predictor of trading activity in the real world than overprecision.

While most empirical and theoretical work has focused on the effect of overconfidence on trading activity, a number of contributions have also examined the relationship between overconfidence and risk taking in the domain of investments. Odean and Barber (2001) find that men not only execute more trades than women, they also take more risk with their portfolio. Grinblatt and Keloharju (2009) find that more overconfident men are more likely to trade stocks in general, an inherently risky activity. Puri and Robinson (2007) find that more optimistic people (those, whose self-reported expected life expectancy is higher than one indicated by actuarial tables) engage more in holding individual stocks, while in an investment experiment Pikulina, Renneboog and Tobler (2017) find that participants who overestimate their financial knowledge relative to their peers invest more in a risky investment proposition. Fenton-O'Creevy, Nicholson, Soane and Willman (2003) report that risk management skill is negatively correlated with illusion of control, which is sometimes considered as another facet of overconfidence (Taylor & Brown, 1988). In an asset market experiment, Michailova and Schmidt (2016) find that markets consisting of overconfident participants generate larger bubbles, indicating a higher propensity to take risks. Chuang and Lee (2006) systematically test if the overconfidence hypothesis also has an effect on the riskiness of the stocks investors trade. They find that after making a profit (leading to an increase in overconfidence; Gervais & Odean, 2001), American investors shift onto trading riskier assets more actively, thereby increasing the riskiness of their portfolio. In addition, Merkle (2017) also finds that relative overconfidence significantly predicts some aspects of diversification and risk taking in an investment context.

To summarize the present state of research: a number of contributions relying on observational data find a positive relationship between relative overconfidence (most often operationalized as BTA effect) and trading activity. There are also a number of contributions (both relying on observational- as well as experimental data) which document a relationship between relative overconfidence and risk taking in market settings. One important issue with the previously presented findings is that they cannot be meaningfully interpreted on the level of the individual, as overplacement measures could be. As about half of individuals can accurately describe themselves as "better than average" on any arbitrary metric, BTA beliefs are only informative when characterizing groups of people. Thinking one is better than average is justified in half of all cases, and this belief might be associated with other individual characteristics. Given recent critiques of BTA as a measure of overconfidence (Benoît & Dubra, 2011; Benoît, Dubra &

Moore, 2015) and the the paucity of literature on this specific aspect of overconfidence, investigating the relationship between relative overconfidence operationalized as overplacement and trading behavior is a much-needed addition to the literature. Providing insights into the overconfidence-trading activity relationship on the level of the individual is the general goal of the following two studies. Study 1 takes a correlational approach, attempting to establish the existence of a correlation between overplacement (operationalized with an individual-level measure grounded in performance on a thematically relevant knowledge task) and various measures of trading activity and risk taking. This lays the foundation for Study 2, which builds on the previous study, to attempt a manipulation of individual beliefs regarding overplacement, and documents the potential causal effect of such an intervention on trading behavior. Providing an accurate and domain-specific measure of relative overconfidence which is interpretable on the level of the individual, and then successfully manipulating this measure to increase individuals' trading activity and risk taking in a market setting would help put previously surveyed empirical findings relying on the better-than-average effect on solid empirical footing, while offering a potential avenue for future interventions aiming to contribute to consumer financial welfare.

Study 1

Our first study has two overall goals. The primary goal is to examine the relationship between baseline overplacement regarding performance on a finance and investments quiz, and trading behavior in a subsequent experimental trading task. As most previous research operationalized relative overconfidence as a question of being better-than-average, not much is known about the relationship between relative overconfidence and trading behavior on an individual level. Thus, any attempt to experimentally manipulate overplacement should be preceded by an effort to examine this dynamic in a baseline "unmanipulated" setting first. The secondary goal of this study is to establish an objective ranking system for performance on the finance and investment quiz task, specific to our participant pool. Being able to determine a participant's correct performance ranking based on objective scores on the quiz task relative to others automatically will serve as a key component of the experimental manipulation in Study 2.

Previous research relying on empirical data (Glaser & Weber, 2007; Merkle, 2017; Grinblatt & Keloharju, 2009) describes a mostly positive but differentiated relationship between various facets of trading activity (such as the number of trades executed, portfolio turnover and the decision to trade at all) and relative overconfidence. In order to test if these relationships hold up in the context of our experimental paradigm, and with a precise and domain-specific measure of overplacement (as opposed to the previously documented better-than-average beliefs), we pre-registered the following three hypotheses regarding trading activity (AsPredicted, 55102):

- **Hypothesis 1a**: The number of transactions executed shows a positive relationship with overplacement
- **Hypothesis 1b**: The average size of individual transactions shows a positive relationship with overplacement
- **Hypothesis 1c**: The aggregate volume traded shows a positive relationship with overplacement

This differentiation enables us to make nuanced statements about the relationships of these three facets of trading activity and overplacement – two aspects being more specific (H1, H2) and one more global (H1c). Hypotheses 1a-1cbconcern more granular measures of trading activity (trading frequency and the average size of individual transactions), with the aggregate volume traded (Hypothesis 1c) being an overall measure of activity, and thus our key dependent variable of interest.

As described previously, a number of previous contributions also hint at a positive relationship between relative overconfidence and risk taking (Pikulina, Renneboog & Tobler, 2017; Puri & Robinson, 2007; Fenton-O'Creevy, Nicholson, Soane & Willman, 2003). We also wanted to investigate if these associations hold up when we implement a domain-specific measure of overplacement as a predictor for risk taking in the context of a realistic simulated market. In order to examine this proposed relationship, we formulated and pre-registered the following hypothesis:

Hypothesis 2: Portfolio risk shows a positive relationship with overplacement

Hypothesis 2 concerns risk taking in the trading task, and the relevant dependent variable is operationalized as the average proportion of one's portfolio composed of the risky asset. This study received approval from the Ethics Committee of ETH Zurich (*2019-N-115*).

Methods

This study has a correlational design, meaning that no experimental manipulation takes place – our focus is testing if our participant's perception of their performance on a finance and investment quiz in relation to their peers (other individuals selected from the same participant pool) is predictive of their behavior in a subsequent one-round trading task.

Participants

Participants were recruited from the online labor pool Amazon Mechanical Turk (MTurk) in 3 batches between 29.12.2020 and 31.12.2020. As the focus of the current study was on investigating the behavior of actual retail investors in an experimental market setting, our main criterion for recruitment and inclusion in the study was as follows: only participants who are currently directly investing in the US stock market should participate. An additional criterion for participation was that only permanent US residents, who are in the United States at the time of survey completion are eligible for the task. Furthermore, we asked prospective participants to complete the study using the Chrome browser, in order to minimize compatibility issues with our experimental software (Andraszewicz, Kaszás & Hölscher, 2021). As an additional criterion, only workers who have completed at least 100 Human Intelligence Tasks (HITs), and with a success rate of more than 95% were eligible to participate.

Given the widespread issue of overclaiming in the context of study participation on MTurk (Chandler & Paolacci, 2017; Bentley, 2021; Thomas & Clifford, 2017), a screening mechanism was implemented at the beginning of this study. The process is identical to that described in Chapter 1, with the change that the score from the three validation items used there (Appendix I), which did not affect one's screener score, does flow into the screening procedure in the present iteration. This additional level of strictness is warranted by the more open nature of data collection on base MTurk, compared to data collections utilizing a pre-vetted pool of participants via CloudResearch (Litman, Robinson & Abberbock, 2017). One additional change was the removal of three "distractor" questions at the beginning of the screener survey – as participants were directly asked to only participate if they were active US retail investors fulfilling all other criteria, adding such questions would not have contributed to the validity of our findings.

In addition to the screening process, a Captcha (Google, Mountain View, CA) was also implemented, to avoid data quality issues due to bots and prospective participants with extremely low attentiveness. From the 643 individuals who have attempted to complete the study and have accepted the consent form for participating in the study (Appendix II.), 265 individuals finished with a complete dataset (Female: 77, M_{Age}=43.71, $Range_{Age}$=18-83*),* comprising our final sample. This is slightly higher than the 250 participants pre-registered – this is due to how MTurk data collections tend to overshoot to achieve minimum recruitment numbers to ensure successful data collection. While our screening process results in a high screening-out rate of about 58%, it also ensures much-needed control over individual participants' tendency to overclaim their expertise or experience, and contributes to a higher baseline level of financial literacy and attention (results of a pilot comparison to this effect are described in Chapter 1 in more detail). All participants who have attempted the current experiment have received a $0.5 base payment. Those who have passed the screening procedure, accepted the informed consent form and completed the trading task received additional performance-based variable compensation, based on how they have performed in the trading task (*M*=$2.5, *Range*= $2.32 - $2.75).

Procedure and Materials

Participants could indicate their interest in participating in the study, by reading the description for the HIT they have been invited to. All participants who accepted these terms and clicked the survey link provided by CloudResearch were automatically forwarded to the survey platform Qualtrics (Qualtrics, Provo, UT), where the pre-survey was implemented, containing the screener items and a finance and investments quiz. Participants first completed the screener questionnaire, and depending on their answers, were either then routed to an end-of-experiment screen with details on their compensation, or were then forwarded to a screen displaying a participant information screen and the opportunity to accept the consent form provided, or opt out of the study altogether. Participants who have opted to participate were asked to read detailed instructions to the trading task (Appendix II.).

After reading the instructions, the study consists of two main parts: a quiz on finance and investments, and a subsequent incentivized trading task. The quiz consists of a mix of 13 single-choice, multiple choice and open-ended questions (for a full list see Appendix III.). These questions have either been developed by the authors, or have been developed as part of previously published measures of financial literacy and investment knowledge (Alessie, Van Rooij & Lusardi, 2011; FINRA, 2012). Correct answers yield positive points, while incorrect answers result in deductions. The lowest score achieved by a study participant was 0.99, while the highest was 60.5 (*M*= 37.91). Participants have been informed in advance that they will be provided with news during the trading task. As the content of quiz questions relates mostly to the US stock market, and the subsequent trading task incorporates price-relevant news events, we expect that participants will find their performance on the quiz task informative to their actual performance on the trading task.

After the quiz concluded, participants were asked to provide an indication of their relative rank when it comes to performance on the quiz, compared to 100 other participants. We have provided answer options on a slider from Rank 1 (best) to Rank 100 (worst). While some of the previous overconfidence literature was based on percentile comparisons, we have decided to present our information by way of ranks for ease of comprehension (Bowman, 2002). As a manipulation check (Appendix IV.) we have asked participants to provide their expected rank when it comes to their performance on the trading task – the values correlate to a high degree ($r_S = 0.6793$, $p<0.001$, $n=265$), indicating that participants perceived a connection between how well they have performed in the quiz task and how well they expected to perform in the upcoming task.

Following the pre-survey block, participants were asked to click a link to proceed to the trading task. They were forwarded to our experimental trading platform (Appendix 5.*)*, developed in-house at the ETH Zurich (further information on the trading platform is available in Chapter 1). The platform displays the development of a pre-determined historical price path in dynamic fashion – market participants are price takers, and can buy or sell shares of the risky asset from an endowment provided to them at the start of each round of trading. The trading task consists of two rounds. The first is a practice round, where participants receive an interactive overview of the trading software, in order to refresh their knowledge from previous text instructions. Following the practice round, participants are warned that the next round (lasting about 5 minutes) will be payment-relevant. Participants received an endowment of $2.5 (or 30 000 in Experimental Currency Units, ECU), invested at the start of the experimental round 50% in the risky asset, and 50% in cash. The portfolio value they ended up with at the end of the experimental round was paid out at the exchange rate of $1 for 12 000 ECU. The median participant took 21 minutes to complete the experiment (SD= 10.71).

Participants were informed at the outset that they will be observing the development of an unnamed US stock index, from an undisclosed 1-year period. A key feature of this study is the use of a real-world historical price path and news events. A time series of adjusted daily closing prices of the S&P 500 market index (06.01.1984-28.12.1984) was displayed in the experimental round. This time series was selected in order to provide participants with a stimulus that is not exhibiting substantial upward- or downward trends, and for which enough contemporaneous news data was available. In order to ensure that participants do not easily recognize the exact asset and time period, prices were standardized to start at a randomly selected value (338.56 ECU). In a post-task survey question, participants were asked to identify the asset and time period, and no participant provided an approximately correct answer to both questions. The news events presented were selected to provide information on market developments and the economy one would encounter in real life, and were all sourced from contemporaneous news reports (Appendix V.) Some materials were lightly edited or condensed in order to avoid biasing participants with period-specific details.

As participants traded in the experimental round, data on their behavior (buying, selling and holding) was collected, in order to derive our four independent variables of interest: the number of transactions one executed over the experimental round (*Transactions*, H1a); the average size of individual transactions over the course of the experimental round (*Vol_Avg*, H1b); the combined volume traded by an individual across the experimental round (*Vol_Sum*, H1c), and the mean level of portfolio risk displayed by an individual over the course of the experimental round (*Risk*, H2). Following the conclusion of the trading task, we have implemented a post-task survey (Appendix VI.).

To derive our main predictor variable (overplacement in the quiz task), we first calculated each participant's objective rank on the quiz task from 1 (best) to 100 (worst), compared to all other study participants. We then subtracted each participants' subjectively reported rank from their objective rank, yielding an overplacement measure which is positive in the presence overplacement, and negative in the presence of underplacement. We find that our sample is well-calibrated overall, with a small mean overplacement value of 0.55 rank (for summary statistics on our key variables see Table 1.). A histogram shows a symmetrical distribution with considerable dispersion in individual values (Fig. 1.). A one-sample t-test confirms that the mean of the distribution is not significantly different from 0 ($t(273) = 0.12$, $p = 0.91$).

Results

Following our pre-registered analysis plan, we calculated non-parametric correlation coefficients for the trading activity measures (H1a-c), due to the presence of substantial outliers. Parametric correlation tests were conducted to examine the relationship between risk taking and overplacement. All hypotheses tests were two-tailed with an alpha value of 5%. When it comes to our first measure of trading activity (*Transactions*, H1a), we find no significant correlation between the number of transactions executed over the course of the experimental round, and overplacement on the preceding quiz task ($r_S = 0.056$, $p = 0.2345$, n=265). This implies that our data does not support Hypothesis 1a, concerning the relationship of the number of trades executed and overplacement. A test of the correlation coefficient between the average size of individual trades and overplacement (*Vol_Avg*, H1b) similarly reveals no relationship ($r_S = 0.003$, p=0.96, n=265). Examining the bivariate relationship of our key dependent variable for trading activity (*Vol_Sum*, the aggregate value of shares traded) and overplacement (H1c), we similarly report no statistically significant relationship (Fig. 2, $r_S = 0.07$, $p = 0.23$, n=265). While the coefficient estimates of all of our hypothesized relationships concerning trading activity are in the expected positive direction, none of our hypothesis tests report a significant relationship. Our last hypothesis, concerning the positive relationship between risk taking (*Risk*, operationalized as the mean share of the risky asset in one's portfolio over the experimental round) and overplacement similarly reveal the absence of a significant parametric correlation (Fig. 3, r=-0.04, $t(272)$ = -0.59, p= 0.55).

In an additional analysis driven by previous research on the association between "better-than-average" beliefs and trading behavior, we create a dummy variable to identify if individuals have indicated that their quiz performance was better or worse than the median participant. A means comparison test reveals that participants who have indicated being better than average took on significantly more risk in the trading task than participants who have indicated performing worse than average (Figure 4; $t(231.89)$, p = 0.035), yielding a robust explanatory measure of effect size of 0.2 (Mair & Wilcox, 2020). To further investigate this finding, we constructed four regression models, controlling for basic demographic variables which have in the past been associated with risk taking (Charness, Eckel, Gneezy & Kajackaite, 2018; Schildberg-Hörisch, 2018; Mata, Josef & Hertwig, 2016; Josef, Richter, Samanez-Larkin, Wagner, Hertwig, & Mata, 2016; Dohmen, Falk, Huffman, Sunde, Schupp & Wagner, 2011). The results of our first model reveal that controlling for key socio-demographic and attitudinal measures attenuates the effect of BTA (Table 2.). In addition, we find that the effect of gender, age and self-reported risk attitudes overlap with findings from previous literature regarding them being correlates of risk taking. This is an important finding, indirectly supporting the external validity of behavioral differences observed using our experimental trading platform. In a series of three further regressions, we find no significant effect of one's objective score, subjectively perceived rank or level of overplacement on risk taking.

Discussion

In an incentivized correlational study with retail investors, we find no support for the existence of a relationship between overplacement on a thematically relevant quiz task and various facets of trading behavior on an experimental market. We document a weak relationship between one's belief in having performed better than average on the quiz task and risk taking, but this relationship is attenuated when accounting for age, gender and self-reported risk preference. These overall null findings can be explained in two different ways: either they are the result of methodological inadequacies (due for example to lack of statistical power, incentives or other related factors), or they are due to the fact that the hypothesized relationship between overplacement (and more generally relative overconfidence) either does not exist or is very weak in our abstract correlational setting.

As for the first explanation, the current study had a preregistered minimum sample size of 250, which enables the identification of small (0.2) correlation effects with about 90% power. We do document significant relationships of this magnitude, but not when it comes to our pre-registered hypotheses. Another potential issue with the present design which might have contributed to the absence of our hypothesized effects could be a perceived lack of realism or verisimilitude in the trading task. While our software was intended to provide a simplified but realistic interface for executing trades, experienced retail investors might have found the task presented uninteresting or not immersive enough, which could have then attenuated our hypothesized effects. We have made steps to counteract this by ensuring that the price paths and news events presented feel as real as possible, to encourage engagement with the experimental task.

While it is not possible to make a definitive statement with the data currently available to us, a text analysis of an optional open-ended post-task question asking participants to provide feedback on the study yields encouraging results. Of the 176 inputs provided, and after the exclusion of commonly used English stop words (Benoit, Muhr & Watanabe, 2020) "fun" was one of the 5 most frequently used words (preceded by "stock", "study", "time" and "buy"). The overall mean sentiment of words used according to the AFINN sentiment lexicon (Nielsen, 2011) from -5 (negative) to +5 (positive) was 0.79 – trending positive. While there is an obvious issue of self-selection when it comes to optional text inputs, we interpret this finding as a signal that our task was engaging enough for a substantial number of our participants.

Another potential issue is that the investor status of individuals we have recruited on the MTurk platform was not previously validated by a third party, such as CloudResearch, as was done for the first study of this dissertation. We have opted to refrain from using third-party validation due to the fact that the pool of active individuals vetted by CloudResearch previously is relatively small, and has been partially exhausted by a previous research effort (described in more detail in Chapter 1). Data from a previously conducted pilot provides an insight into the financial literacy of participants vetted by CloudResearch, recruited via MTurk and then screened with the same screening survey implemented for this study. We report no significant differences between externally validated and screened pilot participants, and participants of the present study in terms of a two-item index of standard financial literacy questions (Alessie, Van Rooij & Lusardi, 2011; $W = 20976$, $p = 0.8323$). This highlights the fact that our post-screening sample of self-described retail investors are not distinguishable from participants validated by third-party service providers widely used by other researchers with regard to financial literacy.

When it comes to incentives, based on recent experience with collecting data in an online setting, we have set a relatively high \$0.5 base compensation, and a similarly high ~\$2.5 variable bonus for this study. Given the 18-minute planned running time of this study, this equates to a roughly \$10 hourly wage – well above present average rates for workers on MTurk and current US federal minimum wage. We believe this amount is meaningful enough for our participants to take our study seriously.

While there are potential design factors which might have contributed to our null findings, our previous observation of a partial BTA effect and its attenuation after controlling a select few background variables implies an alternative explanation. It might well be that some previously documented correlations between relative overconfidence measures and trading behavior (Merkle, 2017; Glaser & Weber, 2007; Grinblatt & Keloharju, 2009) could in part be accounted for by omitted variable bias. If certain unobserved features of individual investors are systematically associated with overplacement or BTA beliefs, that might result in documenting a spurious relationship between various operationalizations of relative overconfidence and trading behavior. Short of a robust quasi-experimental design, the best way to address these concerns is to implement a manipulation of relative overconfidence in a realistic experimental setting. This enables the experimenter to cleanly identify the causal effect of overplacement on trading behavior if one exists, while avoiding potential confounding from unobserved third variables. In Study 2 we attempt to do exactly this. Potentially finding that successfully manipulating overplacement leads to higher trading activity (in one of its three distinct operationalizations) or to higher risk taking could provide the first causal evidence for the positive relationship between relative overconfidence (operationalized as overplacement) and trading behavior – this relationship was previously documented in settings which did not enable causal inference. An experimental approach can account for potential confounding a purely correlational approach could never do, even when controlling for a number of relevant background variables, as was done in Study 1. Conversely, documenting the absence of a causal effect in an experimental setting would raise questions about the robustness of the relationship between relative overconfidence and trading behavior documented in previous literature.

Study 2

In their 2000 paper titled "Trading Is Hazardous to Your Wealth", Barber and Odean show that individual stock investors who trade actively significantly underperform the market, by as much as 6% per year. Odean (1999) finds that discount brokerage customers trade so much that the most active traders end up underperforming the market substantially after trading costs are considered. Barber & Odean (2001) demonstrate that male investors, who on average trade more actively than women, reduce their returns via higher trading costs more than female investors, who tend to invest more passively on average. In their 2008 study ("Just How Much

Do Individual Investors Lose by Trading?") Barber, Lee, Liua and Odean (2008) use detailed transaction-level data from the Taiwanese stock market, and find that individual investors are responsible for most of the losses generated to the Taiwanese stock exchange, amounting to 2.2% of Taiwanese GDP in the period where data was collected.

The issue of overtrading is as relevant today as ever, as wealth invested in easily tradeable vehicles like exchange-traded funds (ETFs) is continually increasing: since they enable intraday trading of shares of investment funds at a relatively low cost, some observers believe they might counterintuitively give rise to overtrading (Bogle, 2016). As noted by Weber and Balchunas (2018), with the increasing ease of trading in today's digital environment, some trading platform providers themselves have started taking steps to warn their clients in case they're trading too actively, to prevent them from racking up excessive losses and transaction fees. Other significant changes have recently taken place in the retail investment landscape, with the democratization of online investing services, booming retail trading activity and the elimination of transaction fees (these developments are reviewed in more detail in Chapter 1.)

Various regulatory approaches to curbing trading volumes and the volatility associated with speculative behavior have been discussed in the past. The best-known idea, a transaction tax to counteract the "excessive international (…) mobility of private financial capital" (Tobin, 1978, p.2.) was proposed by Tobin (1978, 1996). Subsequent empirical studies have reported significant negative externalities associated with implementing such rules in realistic settings (Hanke, Huber, Kirchler & Sutter, 2010; Palley, 1999; Mannaro, Marchesi & Setzu, 2006; McCulloch & Pacillo, 2011). A popular initiative currently proposed in Switzerland would go one step further, and would tax all financial transactions – the specific goal of these microtaxes would be to curb speculation, and to stabilize financial markets by making speculative trades more expensive, and thereby less profitable (Braun, 2018). There have been recent developments worldwide in increasing stamp duties and taxes on financial transactions with the dual goal to raise tax revenues and discourage speculation (Robertson, 2021). It seems there is a need to find effective but lighttouch alternatives to curb speculative trading, and such an approach could have positive welfare effects both on a micro-, as well as on a macro-level. As relative overconfidence in various operationalizations has been found to be associated with trading activity and risk taking in past empirical research, it would be worthwhile to investigate its potential causal effect on trading activity. If the existence of a causal relationship is established via experiment, overplacement might then be targeted for later debiasing efforts, in order to influence individual investors' trading activity.

A number of experimental contributions have investigated the causal relationship between various facets of overconfidence and trading behavior in the past. Deaves, Lüders and Luo (2008) conducted an experiment where they used a measure of calibration-based overconfidence to elicit participants' overprecision, and provide them with increasingly accurate private information, in increasing overconfidence (asymmetric information design). Following this manipulation, they sorted participants into low/high overconfidence asset markets. They were thus able to make market- and individual-level conclusions on the relationship between overprecision and trading activity. They find that overprecision indeed explains trading activity both on the individual and market level, but overplacement incrementally predicts trading activity as well (this approach has received numerous critiques: Phan, Rieger & Wang, 2018; Fellner-Röhling & Krügel, 2014; Glaser & Weber, 2007).

Bregu (2020) used Deaves, Lüders and Luo's (2008) experimental setting, but distributed private information according to participants' level of overestimation on an unrelated task, such that individuals who exhibited higher overestimation ended up with more accurate signals. He finds that in a similar market setting like Deaves, Lüders and Luo (2008), both overestimation and overplacement regarding previous performance will positively impact trading activity. However, Bregu (2020) also does not manipulate overplacement directly. In an asset market experiment, Michailova and Schmidt (2016) sort participants into experimental asset markets according to level of overconfidence measured via an overprecision task. They find higher deviation from fundamental value and higher trading volumes in overconfident markets, as well as higher average price levels.

In another experimental effort Yang and Zhu (2016) measure three forms of overconfidence: overprecision, the better-than-average effect and illusion of control (Taylor & Brown, 1988). Following this, they assign participants randomly to 10-person 10-round continuous experimental asset markets (Smith, Suchanek & Williams, 1988) with symmetric information. There are two types of markets: one where the value of the dividend for each period is randomly determined (with a 25% probability of a specific dividend value taking place in a given round – *risk*), while in the *ambiguity* market the value of the four possible dividends was the same, but their probabilities were unknown. They find that the type of market one participates in changes the relationship between different forms of overconfidence and trading activity. The authors' expectation was that since the ambiguity condition represents a more complex (and more realistic) asset pricing problem, participants' trading will be more strongly affected by overconfidence. Yang and Zhu (2016)'s two main findings are that different forms of overconfidence don't correlate with each other significantly in this context, and only better-than-average belief correlates positively with the number of trades in the ambiguity condition, while miscalibration and illusion of control do not.

In summary, there do exist a number of efforts which attempt to investigate the causal relationship between various aspects of overconfidence and behavior in experimental market settings – however, none of them have attempted to implement an overplacement manipulation, to study the effect of relative overconfidence on the level of the individual. Building on the first study in this chapter, our second study implements a novel experimental manipulation of overplacement, in an attempt to test the existence of a causal relationship between overplacement and trading behavior (trading activity and risk taking). Given the inconclusive findings of Study 1 it is especially important to further investigate this proposed relationship.

As in Study 1, we have four preregistered hypotheses (AsPredicted, 56895), relating to our four dependent variables (Trading activity: *Number of transactions (Transactions)*, *Average size of transactions (Vol_Avg)*, *Aggregate trading volume (Vol_Sum)*. Risk taking: *Mean level of portfolio risk (Risk)*).

- **Hypothesis 1a**: The number of transactions (*Transactions*) executed is higher in the presence of induced overplacement
- **Hypothesis 1b**: The average size of individual transactions (*Vol_Avg*) is higher in the presence of induced overplacement
- **Hypothesis 1c**: The aggregate volume traded (*Vol_Sum*) is higher in the presence of induced overplacement
- **Hypothesis 2**: Individuals take on more risk (*Risk*) in the presence of induced overplacement

Methods

To examine if overplacement has a positive effect on trading activity and risk taking in the domain of investments and trading, an experimental manipulation of overplacement was developed and implemented in Study 2. The present study has a between-groups design with two conditions: an experimental condition with induced overplacement and a control condition where no overplacement is induced. In order to investigate our hypotheses, we implement a scale manipulation (Ockenfels & Werner, 2014; Schwarz, 1985) to induce overplacement regarding one's performance in a finance and investments quiz. Following this, participants complete a trading task identical to the one presented in Study 1. Noisy and skewed performance feedback on the first experimental task (finance and investments quiz) is expected to induce overplacement, while a second experimental task (trading task) is implemented in order to measure the behavioral effects of the overconfidence induction, and measure our four dependent variables.

Participants.

Participants were recruited from the same online labor pool, and using the same recruitment criteria as was the case in Study 1. They had to hold direct investments in the US stock market, be permanent US residents, be physically present in the US at the time of data collection and satisfy our criteria for MTurk experience and previous work quality. The only deviation in terms of recruitment procedure was that in addition to completing the study using Google Chrome, we have also added the requirement that only workers using Windows machines can participate. This change was implemented due to the sudden revocation of an intermediate server certificate used by our experimental software, shortly before the data collection for Study 2 was supposed to begin (Digicert QuoVadis, 2021). While updates which fix this certificate issue take place automatically on Windows PCs, users on the MacOS operating system would have to manually update Chrome, otherwise they might receive a security error message when entering our experimental software. The decision therefore has been made to only include Windows users in the experiment.

Data collection lasted from 29.01.2021 until 27.02.2021, in 21 batches. A combined 2046 workers have accepted our invitation and provided informed consent. Of these individuals 735 completed the trading task, for a screening-out rate of 64% – somewhat higher but generally in line with Study 1. Due to the fact that our scale manipulation has limited effectiveness at the extremes of the quiz performance distribution, we have restricted our main analysis to individuals who rank between 20 and 90 from 100 in terms of their quiz performance. Thus, the final number of observations in our analysis is 447 (Female= 155, M_{Age}= 41.25, $Range_{Age}$= 19-79). Participants were incentivized identically to Study 1, resulting in mean variable bonus of $2.5 ($Range_{Bonus}$= $2.14 - $2.71), in addition to the base compensation of $0.5 offered to everyone who attempted the experiment. The median participant completed the experiment in 21 minutes (*SD*= 6.16).

Our pre-registered target sample size was 580. This is a substantially higher number than the number of participant pre-registered for Study 1. The reason for this is that while in Study 1 we tested a correlation hypothesis, in Study 2 we are comparing two groups of participants, thus requiring a higher number of participants to reliably observe a small-to-moderate effect (Sawilowsky, 2009). This higher number of participants is also warranted by the fact that about 30% of participants are excluded from our main analyses.

However, our data collection was terminated before reaching this pre-determined target. The decision was made before analyzing any behavioral or self-reported data, and was driven by a substantial drop in the pace at which new data points were collected (Appendix. VIII). The most likely reason for this was the exhaustion of qualifying participants at our predetermined level of compensation. The decision was overall motivated by concern for the cost of further data collection (both in terms of time and money) and a concern for data quality. According to an ex-post power analysis, with the number of participants collected, our analysis has 85% power to find an effect of *d*=0.3 (McKenzie & Ozier, 2019). At the current number of participants, we have 65% power to find an effect size of 0.2. Given previous experimental contributions (Yang & Zhu, 2016; Bregu, 2020; Deaves, Luo & Lüders, 2008) a small effect between 0.2-0.3 is a plausible estimate for the size of the expected treatment effects, and our experiment is therefore not seriously underpowered.

Procedure and Materials.

The procedure and materials for Study 2 are mostly identical to Study 1, with one key exception. While in Study 1 participants were asked directly after the quiz to estimate their rank from 1 to 100, in Study 2 participants first receive noisy feedback on their potential position before giving us their estimate. In order to determine participants' true rank "on the fly", we based our calculations on quiz performance data from Study 1. Taking the subset of participants who have completed Study 1 on a Windows system, and determining which objective score on the quiz task corresponds to which rank from 1 to 100 (where 1 is the best and 100 is the worst) enabled us to calculate Study 2 participants' rank compared to individuals drawn from the same participant pool, and screened using identical criteria.

Having determined each participant's "true rank" compared to our reference sample from Study 1 following the quiz task, the experimental manipulation takes place next: using a scale manipulation (Ockenfels & Werner, 2014; Schwarz, 1985, Bertrand & Mullainathan, 2001; Payne, 1951), participants in the experimental condition were presented with a scale where an interval covers 20 potential ranks. They are truthfully told that the presented interval estimate contains their actual rank. These participants' actual ranking is at the very bottom of the scale, and possible values extend 20 ranks higher (Fig. 4.). Based on previous literature on survey design and scale manipulations, we expected participants to treat the midpoint of a proposed scale as an informative anchor for their actual rank. By skewing the midpoint of the presented interval upwards we thus expect participants in the experimental condition to estimate their position compared to their peers higher than it is in reality. Participants in the control group have also been presented with a 20-rank interval, but their true placement was in the middle of the 20-rank span – they are therefore expected to have an unbiased estimate of their placement. As in the experimental condition, they are also truthfully told that the presented interval contains their actual rank.

The manipulation was expected to result in one group whose members systematically overestimate their skills and knowledge compared to others by about 10 ranks on average, and where about 60% consider themselves "better-than-average". In the control group we expected to find better calibration, with ~50% of participants reporting themselves as better-than-average. A downside of such an approach is that because of floor- and ceiling effects, the intervention is weakened for participants approaching the extreme ends of the spectrum, and some participants will have to be excluded from the analysis altogether. Participants in the top 20 will be excluded from the analysis, as the experimental manipulation is weakened from rank 20 upwards. Similarly, participants worse than the 90th rank will be excluded, as the scale offered in the control condition cannot extend below 100 (the lowest rank possible). These sample selection choices have been pre-registered and have been accounted for in an a priori power analysis.

Ockenfels and Werner (2014) successfully used a scale manipulation to influence participants' economic behavior in the context of a dictator game. Ockenfels and Werner (2014) show scale manipulations to be a potentially "effective tool to induce systematic and predictable shifts in beliefs, which otherwise could not be easily induced without deception (…)"(p. 141) , a potential concern with other potential methods of inducing overplacement. As deception is frowned upon in experimental economics- and finance, making sure that any manipulation used in the present study does not constitute deception is of key importance. While explicitly supplying participants with false information regarding their placement would be more likely to reach the treatment effect of interest, contributions using such a manipulation would have questionable validity, and would not get published in the target domain of experimental finance (Hertwig & Ortmann, 2008; Bonetti, 1998; Hey, 1998; Hersch, 2015). As Hey (1998) noted in a widely cited response to Bonetti's critique of how researchers in experimental economics approach deception: "There is a world of difference between not telling subjects things and telling them the wrong things. *The latter is deception, the former is not."* (Hey, 1998, p.1.). In the present study no false information is provided, and it has been approved by the Ethics Committee of ETH Zurich (2019-N-115). A further benefit of our method is that it avoids presenting two groups of participants a different set of questions during the quiz task, a potential confound that would be associated with a manipulation using the hard-easy effect (Larrick, Burson & Soll, 2007; Moore & Healy, 2008; Dafoe, Zheng & Caughey, 2018).

Following our experimental manipulation, participants were asked to provide an exact estimate for their ranking on the quiz task, as well as a prediction for how well they will perform in the trading task (also in relative terms). The trading task proceeded exactly as it has in Study 1 (Appendix V.). After completing the trading task, we asked participants to complete a similar battery of post-task survey questions as in Study 1 (Appendix VI.). To close the experiment, we debriefed our participants by explaining our experimental manipulation, their true rank on the quiz task, and which experimental condition they were assigned to (Appendix VII.).

Results

In order to assess if our scale manipulation achieved its intended effect, we have asked our participants two manipulation check questions (Appendix IV.). After being presented with the 20-rank interval containing their true rank, we asked participants to give us an exact estimate for their actual rank. If this point estimate is better in the experimental condition (higher rank), that is a sign that our manipulation has achieved its intended effect. Firstly, we can report that we find no statistically significant difference between participants when it comes to actual rank ($t(443.46) = 0.43$, $p = 0.67$). Most importantly, we find a significant effect of the experimental manipulation on self-reported rank: participants in the experimental condition reported significantly higher point estimates for their rank ($t(444.7) = 4.45$, $p < 0.001$). The mean participant in the experimental condition reported being at rank 40.13, while the mean participant in the control condition reported a rank of 48.94 – yielding a robust effect size estimate of 0.26 (Mair & Wilcox, 2020). As in Study 1, we then asked our participants to give us a prediction for how well they will perform in the trading task, compared to other participants. We again find significantly higher expectations in the experimental group ($t(445) = 2.46$, , p-value = 0.01, d=0.14). Overall, these checks tell us that our manipulation successfully generated overplacement in the context of the quiz task, and this carried over to expectations regarding later task performance (Fig. 6.).

To test our hypotheses we conducted four pre-registered means comparisons. Using a Wilcoxon rank-sum test, we find no support for our hypothesis that overplacement leads to a higher number of transactions (*Transactions*, H1a, $W = 27104$, $p = 0.12$). Our data similarly does not support the existence of between-condition differences in the average size of transactions (*Vol_Avg*, H1b, $W = 25946$, $p = 0.42$). Our key dependent variable, the aggregate volume of shares traded was likewise not significantly different between conditions (*Vol_Sum*, H1c, $W = 26770$, $p = 0.19$). The distribution of values by condition is illustrated in Fig. 7.

Overall, we can conclude that in the present experiment inducing overplacement did not lead to higher trading activity. When we examine the medians for each dependent variable (Table 3.), we see that there actually seems to be a trend in the opposite of the hypothesized direction. To test our hypothesis that induced overplacement causes higher risk taking (H2), we have conducted a pre-registered parametric means comparison. We again find no significant group difference (*Risk*, Fig. 7, $t(444.32) = -1.02$, $p = 0.31$). We thus can conclude that our overplacement manipulation did not shift participants' behavior in the expected direction.

Similarly to Study 1, we conduct a series of additional regression analyses while controlling for factors, which in previous literature have been documented to correlate with risk taking (age, gender and self-reported risk attitude). We systematically vary key independent variables across models (Table 4.), to observe their effect on portfolio risk in the trading task. In a regression framework we do not observe a significant effect of experimental condition (Model 1.), nor do we observe an effect overplacement when operationalized as a continuous variable (Model 2.). We find that one's self reported rank on the quiz positively predicts risk taking (Model 3.). Most interestingly, if we create a dichotomized variable to label individuals who have indicated performing better or worse than the average participant, this binary predictor has a substantial effect on risk taking – implying that those who consider themselves above average trade more riskily (Model 4). This relationship shifts in significance above the 5% critical threshold (p=0.07) after combining most predictors in one model (Model 7). The coefficient for

BTA is basically unchanged compared to the restricted model – combining somewhat collinear variables (BTA and Objective Score, $VIF_{BTA}=2$, $VIF_{Obj.\ Score} = 2.21$) leads to an inflation in standard errors, explaining why the coefficient becomes not significant on a 5% significance level. This analysis shows that beyond the effect of our experimental manipulation, demographic factors, and actual knowledge and understanding of finance and investments, the most important predictor of trading behavior is a binary variable for one's self-assessment being above-average or not. It is important to note that this analysis is specific to risk taking – these patterns are not present in models where the dependent variable is various facets of trading volume.

Discussion

We have found that our method successfully induces overplacement regarding performance on a finance and investments quiz, relative to other participants from the same participant pool. We have also demonstrated that these differences extend to one's expected performance on a subsequent trading task as well. However, we find no differences in trading behavior between our two conditions following a series of pre-registered means comparison tests. There are multiple possible explanations for this – some pertaining to the design and execution of the experiment, and some relating to more fundamental conceptual issues.

One potential issue as discussed earlier is the question of statistical power. Our experiment had enough power to observe small- to medium effects with high probability, and additional regression analyses provide support for this position. In addition, when examining the results from our means comparisons it becomes clear that if any meaningful effects are present, they are too small to have been detected even with our higher pre-registered sample size. An alternative reason could be sample selection – in order to examine the effect of our manipulation at its full strength, we have decided to exclude participants at the extremes of our quiz ranking from our analyses. A secondary analysis finds no group differences or substantial changes regarding aggregate trading volumes (W = 70386, p = 0.31) or risk taking (t(717.22)= -0.06, p = 0.95) between conditions when taking the complete sample of participants who have completed the experiment.

Another explanation could be that our peer group (other self-reported retail investors who decided to participate in this experiment, and passed the screening) is a highly synthetic one. We have no information indicating that individuals found the rankings meaningless or contrived in some way – it is nonetheless true that a similar manipulation implemented with a naturally existing group might have led to observable behavior change. One more potential issue, as has been the case for Study 1 is a question of verisimilitude. Did trading in our experiment bear a resemblance to trading in the real world? While we only have indirect evidence, the data does seem to support the view that behavior on our task has parallels with real-world trading behavior. Our regression analyses from Study 2 show systematic relationships between factors such as gender, age and self-reported risk preferences and risk taking, in line with previous literature. Furthermore, we document a relationship between gender and aggregate trading activity in a rank-based regression framework (Kloke & McKean, 2012), with women trading significantly less actively– this is in line with previous empirical findings on gender differences in trading behavior (Barber & Odean, 2001). These findings show overall that individual characteristics associated with behavior in risky settings in the real world show similar associations in our experimental setting. Even so, a potential replication and extension intending to explain the null effects documented here might still benefit from drawing from a different pool of retail investors – while we overall find the attentiveness and performance of our screened-in participants to be overall quite satisfactory, the robustness of our findings could be improved upon by relying on a sample coming from a different source, and validated according to different criteria.

One key finding from our additional regression analyses is that one's position relative to the average participant seems to carry substantial weight. Given the importance of outperforming the mean in diverse financial contexts (Kirchler, Lindner & Weitzel, 2018; Kempf & Ruenzi, 2008; Gärling, Fang & Holmen, 2019), it seems likely that one's perception of being better than average is much more important than one's level of overplacement for determining subsequent trading behavior. Put simply, someone who objectively ranks as the 90^{th} from 100 individuals, and shows a large overplacement effect of 35 ranks would still consider themselves to be worse than average. We had only 29 participants in the experimental condition who have been "upgraded" from actually being below-average in their quiz performance to a better-than average belief by our experimental manipulation. If one's relative position to the midpoint is all that matters, it is understandable that we do not find a significant effect with only 29 individuals treated.

While there are certainly a number of competing explanations for the null findings in our pre-registered analyses, one central finding from our additional analyses is that we document a better than average effect when it comes to risk taking. In contrast to our findings from Study 1, this effect remains significant on a 10% level of significance, even after controlling for a number of background variables, including one's actual score on the finance and investments quiz. While most previous empirical work on relative overconfidence has documented a relationship between BTA and trading behavior, as a concept, BTA beliefs are only meaningfully interpreted in group contexts. One would hope that by calculating an individual-level overplacement measure, and systematically manipulating one group's placement upwards by about 10 positions, we would be able to quantify how overplacement might result in different levels of trading activity and risk taking across the spectrum of quiz performance. What we observe here is that as long as one's belief about one's performance on the quiz is that they are worse than average, one is less inclined to take risks – this is independent of one's actual rank and experimental condition. This highlights the salience of the midpoint in our rankings, and in the domain of finance and investments in general – people seem to take their perceived position relative to the midpoint as a relevant reference point to guide their subsequent behavior.

Conclusion

Overall the two studies presented in this manuscript represent substantial contributions both in terms of furthering our understanding of the relationship between overconfidence and trading behavior, as well as in terms of methodology.

We first tested the baseline relationship between overplacement and trading behavior in a correlational setting. We found no significant overplacement-trading behavior relationship in our preregistered analyses. This is in conflict with previous empirical findings, which documented a positive relationship between better-than-average beliefs and various facets of trading behavior. Such overconfident beliefs are only meaningfully interpretable on a group level, and potentially confounded on the level of the individual. Our overplacement measure relies on actual performance on a thematically relevant knowledge task, and is thus a more internally valid measure of overconfidence. In an additional analysis we did document a significant difference in group means when comparing risk taking between participants with above- and below-average self-perceived performance. This group difference did not remain statistically significant in a regression analysis after the inclusion of control variables derived from previous literature. This finding also deviates from previous literature describing substantial BTA effects on risk taking in correlational settings, and raises the possibility that earlier reports of a BTA-trading activity relationships might have been partially due to confounding.

In Study 2 we set out to investigate a potential causal relationship between overplacement and trading behavior using a novel experimental manipulation, to address the issues highlighted in Study 1. We come to the conclusion that induced overplacement regarding performance on a previous quiz containing thematically relevant questions has no effect on trading behavior. In further analyses we document a significant association between BTA beliefs and risk taking, the point estimate for which remains unchanged and significant at a $p<0.1$ level of significance, even after controlling for a number of relevant demographic and attitudinal covariates, as well as additional collinear metrics. The fact that this effect is not driven by our experimental treatment leaves the question of causality open – people who believe they are better than average on some metric might have this belief for various reasons, most of which are not observable in our experimental setting.

Our overall conclusion from this project when it comes to our contribution to the literature is that overplacement might matter for risk taking only when it results in shifting one's beliefs from being below-average to being above-average. We do not document a relationship between any operationalization of relative overconfidence and any aspect of trading activity in our experimental setting – this is certainly a surprising finding, requiring further investigation in future research efforts.

When it comes to our methodological approach, we have demonstrated an effective way to shift individuals' self-assessment relative to their peers. While this method comes with obvious drawbacks (such as the necessity to exclude individuals at the extremes of the distribution, and to give participants a general impression of their underlying performance), it also provides a cleaner manipulation of overconfidence than alternative approaches would (for example Larrick, Burson & Soll, 2007). Apart from the manipulation, our use of a realistic incentivized trading task, based on historical prices and contemporaneous news events represents a significant addition to the methodological toolkit for investigating the behavior of retail investors.

Further research should examine conditions under which overplacement might actually cause behavior change in a finance and investment setting. Specifically targeting better-than-average beliefs for manipulation in an incentivized framework would also represent a substantial addition to the literature. More generally, future projects should be conducted in a more naturalistic setting, preferably as part of a field experiment with selective performance feedback, to provide insights which might be more generalizable to real-world trading decisions made by real-world retail investors. We believe that this contribution is a valuable step in the right direction.

References

Agnew, J. R., & Szykman, L. R. (2005). Asset allocation and information overload: The influence of information display, asset choice, and investor experience. The Journal of Behavioral Finance, 6(2), 57-70.

Alessie, R., Van Rooij, M., & Lusardi, A. (2011). Financial literacy and retirement preparation in the Netherlands. Journal of Pension Economics & Finance, 10(4), 527.

Barber, B. M., Lee, Y. T., Liu, Y. J., & Odean, T. (2014). The cross-section of speculator skill: Evidence from day trading. *Journal of Financial Markets*, *18*, 1-24.

Barber, B. M., Lee, Y. T., Liu, Y. J., & Odean, T. (2008). Just how much do individual investors lose by trading? *The Review of Financial Studies*, *22*(2), 609-632.

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, *55*(2), 773-806.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, *116*(1), 261-292.

Barberis, N. C. (2018). *Psychology-based Models of Asset Prices and Trading Volume* (No. w24723). National Bureau of Economic Research.

Benoît, J. P., Dubra, J., & Moore, D. A. (2015). Does the better-than-average effect show that people are overconfident? Two experiments. *Journal of the European Economic Association*, *13*(2), 293-329.

Benoît, J. P., & Dubra, J. (2011). Apparent overconfidence. *Econometrica*, *79*(5), 1591-1625.

Benoit K, Muhr D, Watanabe K. stopwords: Multilingual Stopword Lists. 2020. https://CRAN.R-project. org/package=stopwords

Bentley, J. W. (2021). Improving the Statistical Power and Reliability of Research Using Amazon Mechanical TurkImproving the Statistical Power and Reliability of MTurk Research. *Accounting Horizons. doi: https://doi.org/10.2308/HORIZONS-18-052*

Berner, E. S., & Graber, M. L. (2008). Overconfidence as a cause of diagnostic error in medicine. *The American Journal of Medicine*, *121*(5), S2-S23.

Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review*, *91*(2), 67-72.

Bonetti, S. (1998). Experimental economics and deception. *Journal of Economic Psychology*, *19*(3), 377-395.

Bowman, M. L. (2002). The perfidy of percentiles. *Archives of Clinical Neuropsychology, 17*(3), 295-303.

Bregu, K. (2020). Overconfidence and (Over) Trading: The Effect of Feedback on Trading Behavior. *Journal of Behavioral and Experimental Economics*, 88, 101598.

Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, *89*(1), 306-318.

Chandler, J. J., & Paolacci, G. (2017). Lie for a dime: When most prescreening responses are honest but most study participants are impostors. *Social Psychological and Personality Science*, *8*(5), 500-508

Charness, G., Eckel, C., Gneezy, U., & Kajackaite, A. (2018). Complexity in risk elicitation may affect the conclusions: A demonstration using gender differences. *Journal of Risk and Uncertainty*, *56*(1), 1-17.

Chuang, W. I., & Lee, B. S. (2006). An empirical evaluation of the overconfidence hypothesis. *Journal of Banking & Finance*, *30*(9), 2489-2515.

Cooper, A. C., Woo, C. Y., & Dunkelberg, W. C. (1988). Entrepreneurs' perceived chances for success. *Journal of Business Venturing*, *3*(2), 97-108.

Dafoe, A., Zhang, B., & Caughey, D. (2018). Information equivalence in survey experiments. *Political Analysis*, *26*(4), 399-416.

Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty*, *41*(3), 219-243.

Deaves, R., Lüders, E., & Luo, G. Y. (2008). An experimental test of the impact of overconfidence and gender on trading activity. *Review of Finance*, *13*(3), 555-575.

De Bondt, W. F., & Thaler, R. H. (1995). Financial decision-making in markets and firms: A behavioral perspective. *Handbooks in operations research and management science*, *9*, 385-410.

Digicert QuoVadis (2021, January 14.) *Intermediate CA revocations*. Retrieved from: https://knowledge.digicert.com/quovadis/ssl-certificates/Intermediate-CA-Revocations-January-2021.html

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522-550.

Dorn, D., & Huberman, G. (2005). Talk and action: What individual investors say and what they do. *Review of Finance*, *9*(4), 437-481.

Fellner-Röhling, G., & Krügel, S. (2014). Judgmental overconfidence and trading activity. *Journal of Economic Behavior & Organization*, *107*, 827-842.

Fenton-O'Creevy, M., Nicholson, N., Soane, E., & Willman, P. (2003). Trading on illusions: Unrealistic perceptions of control and trading performance. *Journal of Occupational and Organizational Psychology*, *76*(1), 53-68.

FINRA (2012) Investor Knowledge Quiz. https://www.finra.org/sites/default/files/InvestorDocument/p014495.pdf. Last accessed: 22.03.2021

Gärling, T., Fang, D. and Holmen, M. (2019). Review of behavioral explanations of how rank-based incentives influence risk taking by investment managers in mutual fund companies, *Review of Behavioral Finance*, *12*(2), 136-150. https://doi.org/10.1108/RBF-01-2019-0013

Gervais, S., & Odean, T. (2001). Learning to be overconfident. *The Review of Financial Studies*, *14*(1), 1-27.

Glaser, M., & Weber, M. (2007). Overconfidence and trading volume. *The Geneva Risk and Insurance Review*, *32*(1), 1-36.

Grinblatt, M., & Keloharju, M. (2009). Sensation seeking, overconfidence, and trading activity. *The Journal of Finance*, *64*(2), 549-578.

Guiso, L., Sapienza, P., & Zingales, L. (2018). Time varying risk aversion. *Journal of Financial Economics*, *128*(3), 403-421.

Hanke, M., Huber, J., Kirchler, M., & Sutter, M. (2010). The economic consequences of a Tobin tax—an experimental analysis. *Journal of Economic Behavior & Organization*, *74(*1-2), 58-71.

Healy, P., & Moore, D. (2007). Bayesian overconfidence. Working paper.

Hey, J. D. (1998). Experimental Economics and Deception: Comment. *Journal of Economic Psychology*, *19*, 397-401.

Hersch, G. (2015). Experimental economics' inconsistent ban on deception. *Studies in History and Philosophy of Science Part A*, *52*, 13-19.

Hertwig, R., & Ortmann, A. (2008). Deception in experiments: Revisiting the arguments in its defense. *Ethics & Behavior*, *18*(1), 59-92.

Hirshleifer, D., & Luo, G. Y. (2001). On the survival of overconfident traders in a competitive securities market. *Journal of Financial Markets*, *4*(1), 73-84.

Hong, H., & Stein, J. C. (2007). Disagreement and the stock market. *Journal of Economic Perspectives*, *21*(2), 109-128.

Josef, A. K., Richter, D., Samanez-Larkin, G. R., Wagner, G. G., Hertwig, R., & Mata, R. (2016). Stability and change in risk-taking propensity across the adult life span. *Journal of Personality and Social Psychology*, *111*(3), 430.

Kahneman, D. (2011, October 24). Bias, Blindness and How We Truly Think (Part 1). *Bloomberg View*. Retrieved from https://www.bloomberg.com/opinion/articles/2011-10-24/bias-blindness-and-how-we-truly-think-part-1-daniel-kahneman

Kempf, A., & Ruenzi, S. (2008). Tournaments in mutual-fund families. *The Review of Financial Studies*, *21*(2), 1013-1036.

Kirchler, M., Lindner, F., & Weitzel, U. (2017). Rankings and Risk-Taking in the Finance Industry. *The Journal of Finance*. *73*(5) 2271-2302.

Klar, Y., & Giladi, E. E. (1997). No one in my group can be below the group's average: A robust positivity bias in favor of anonymous peers. *Journal of Personality and Social Psychology*, *73*(5), 885.

Kloke, J.D., McKean, J.W. (2012). Rfit: Rank-based estimation for linear models, *The R Journal*, *4*(2), 57–64.

Kumar, S., & Goyal, N. (2015). Behavioural biases in investment decision making–a systematic literature review. *Qualitative Research in Financial Markets*, *7*(1), 88-108.

Langnickel, F., & Zeisberger, S. (2016). Do we measure overconfidence? A closer look at the interval production task. *Journal of Economic Behavior & Organization*, *128*, 121-133.

Langnickel, F. (2018). Naïve News Trading: Experimental Evidence. Available at SSRN 3165379.

Larrick, R. P., Burson, K. A., & Soll, J. B. (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes*, *102*(1), 76-94.

List, J. A. (2011). Why economists should conduct field experiments and 14 tips for pulling one off. *Journal of Economic Perspectives*, *25*(3), 3-16.

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime. com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433-442.

Luo, J., Subrahmanyam, A., Titman, S. (2018, August 9). *Stock Price Dynamics With Overconfident Investors* Available at SSRN: https://ssrn.com/abstract=2997001or http://dx.doi.org/10.2139/ssrn.2997001

Mair P, Wilcox R (2020). "Robust Statistical Methods in R Using the WRS2 Package." *Behavior Research Methods*, *52*, 464–488.

Mannaro, K., Marchesi, M., & Setzu, A. (2008). Using an artificial financial market for assessing the impact of Tobin-like transaction taxes. *Journal of Economic Behavior & Organization*, *67*(2), 445-462.

Mata, R., Josef, A. K., & Hertwig, R. (2016). Propensity for risk taking across the life span and around the globe. *Psychological Science*, *27*(2), 231-243.

McCulloch, N., & Pacillo, G. (2011). The Tobin tax: a review of the evidence. *IDS Research Reports*, *2011*(68), 1-77.

McDaniel, T., & Starmer, C. (1998). Experimental economics and deception: A comment. *Journal of Economic Psychology*, *19*(3), 403-409.

McKenzie, D. & Ozier, O. (2019 May 16) *Why ex-post power using estimated effect sizes is bad, but an ex-post MDE is not.* Retrieved from: https://blogs.worldbank.org/impactevaluations/why-ex-post-power-using-estimated-effect-sizes-bad-ex-post-mde-not

Mele, A. R. (1997). Real self-deception. *Behavioral and Brain Sciences*, *20*(1), 91-102.

Merkle, C., & Weber, M. (2011). True overconfidence: The inability of rational information processing to account for apparent overconfidence. *Organizational Behavior and Human Decision Processes*, *116*(2), 262-271.

Merkle, C. (2017). Financial overconfidence over time: Foresight, hindsight, and insight of investors. *Journal of Banking & Finance*, *84*, 68-87.

Michailova, J., & Schmidt, U. (2016). Overconfidence and bubbles in experimental asset markets. *Journal of Behavioral Finance*, *17*(3), 280-292.

Moore, D. A., & Cain, D. M. (2007). Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition. *Organizational Behavior and Human Decision Processes*, *103*(2), 197-213.

Moore, D. A., & Dev, A. S. (2018). Individual differences in overconfidence. In V. Zeigler-Hill and T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences.* New York: Springer.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, *115*(2), 502.

Moore, D. A., & Small, D. A. (2007). Error and bias in comparative judgment: on being both better and worse than we think we are. *Journal of Personality and Social Psychology*, *92*(6), 972.

Moore, D. A., Tenney, E. R., & Haran, U. (2015). Overprecision in judgment. *The Wiley Blackwell handbook of judgment and decision making*, 182-209.

Nielsen, F. Ä, (2011), "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages* (2011) 93-98.

Ockenfels, A., & Werner, P. (2014). Scale manipulation in dictator games. *Journal of Economic Behavior & Organization*, *97*, 138-142.

Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *The Journal of Finance*, *53*(6), 1887-1934.

Odean, T. (1999). Do investors trade too much? *American Economic Review*, *89*(5), 1279-1298.

Palley, T. I. (1999). Speculation and Tobin taxes: Why sand in the wheels can increase economic efficiency. *Journal of Economics*, *69*(2), 113-126.

Payne, S. L. (1951). *The art of asking questions.* Oxford, England: Princeton University Press.

Phan, T. C., Rieger, M. O., & Wang, M. (2018). What leads to overtrading and under-diversification? Survey evidence from retail investors in an emerging market. *Journal of Behavioral and Experimental Finance.*, *19*, 1176-1180.

Pikulina, E., Renneboog, L., & Tobler, P. N. (2017). Overconfidence and investment: An experimental approach. *Journal of Corporate Finance*, *43*, 175-192.

Saposnik, G., Redelmeier, D., Ruff, C. C., & Tobler, P. N. (2016). Cognitive biases associated with medical decisions: a systematic review. *BMC Medical Informatics and Decision Making*, *16*(1), 138.

Puri, M., & Robinson, D. T. (2007). Optimism and economic choice. *Journal of Financial Economics*, *86*(1), 71-99.

Rieger, M. O., Wang, M., & Hens, T. (2014). Risk preferences around the world. *Management Science*, *61*(3), 637-648.

Robertson, B. (2021, March 21) *Who Taxes Financial Transactions and Will Others Join?*. Bloomberg. Retrieved from: https://www.bloomberg.com/news/articles/2021-03-11/who-taxes-financial-transactions-and-will-others-join-quicktake

Schildberg-Hörisch, H. (2018). Are Risk Preferences Stable? *Journal of Economic Perspectives*, *32*(2), 135-54.

Schwarz, N., Hippler, H. J., Deutsch, B., & Strack, F. (1985). Response scales: Effects of category range on reported behavior and comparative judgments. *Public Opinion Quarterly*, *49*(3), 388-395.

Shariatmadari, D. (2015, July 18). Daniel Kahneman: 'What would I eliminate if I had a magic wand? Overconfidence'. *The Guardian*. Retrieved from https://www.theguardian.com/books/2015/jul/18/daniel-kahneman-books-interview

Shefrin, H. (2000). *Beyond Greed and Fear: Understanding Behavioral Finance and the Psychology of Investing.* Oxford University Press.

Smith, V. L., Suchanek, G. L., & Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica: Journal of the Econometric Society*, 1119-1151.

Statman, M., Thorley, S., & Vorkink, K. (2006). Investor overconfidence and trading volume. *The Review of Financial Studies*, *19*(4), 1531-1565. p. 1531

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, *47*(2), 143-148.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin*, *103*(2), 193.

Thomas, K. A., & Clifford, S. (2017). Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. *Computers in Human Behavior, 77*, 184-197.

Tobin, J. (1978). A proposal for international monetary reform. *Eastern Economic Journal*, *4*(3/4), 153-159.

Tobin, J. (1996). A currency transactions tax, why and how. *Open Economies Review*, *7*(1), 493-499.

Weber, J., & Balchunas, E. (Hosts). (2018, October 3). *Four ETF nerds debate 20 topics* [Audio podcast]. Retrieved from https://www.bloomberg.com/news/audio/2018-10-03/four-etf-nerds-debate-20-topics

Yang, X., & Zhu, L. (2016). Ambiguity vs risk: An experimental study of overconfidence, gender and trading activity. *Journal of Behavioral and Experimental Finance*, *9*, 125-131.

Tables

Table 1.

Summary statistics for key variables (rounded)

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Risk	0.72	0.18	0.09	0.59	0.72	0.87	1
Transactions	17.72	24.75	0	6	11	22	304
Vol_Avg	16.05	6.25	0	11.31	15.33	22.37	25
Vol_Sum	79,579.79	96,298.79	0	26,833.02	46,589.60	96,881.56	790,023.00
Subjective rank	49.6	23.1	1	31	50	67	100
True rank	49.85	29.34	1	25	49	76	100
Overplacement	0.55	32.06	-89	-19	1	25	80
Objective score	37.91	11.72	1	30	38.5	46.5	60.5
Post-trading perf. rank	58.33	18.31	6	46	56	70	100

Table 2.

Regression table for additional analysis

	Dependent variable:			
	Risk			
	(1)	(2)	(3)	(4)
Age	-0.001	-0.002*	-0.002*	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Gender_Female	-0.096***	-0.092***	-0.093***	-0.101***
	(0.026)	(0.026)	(0.026)	(0.025)
SOEP	0.028***	0.027***	0.027***	0.028***
	(0.010)	(0.010)	(0.010)	(0.010)
BTA_Below	-0.025			
	(0.023)			
Objective score		0.001		
		(0.001)		
Subjective rank			-0.001	
			(0.0005)	
Overplacement				0.00001
				(0.0004)
Constant	0.721***	0.674***	0.750***	0.711***
	(0.061)	(0.068)	(0.065)	(0.061)
Observations	236	236	236	236
R^2	0.127	0.128	0.132	0.123
Adjusted R^2	0.108	0.109	0.113	0.104
Residual Std. Error (df = 230)	0.173	0.173	0.173	0.174
F Statistic (df = 5; 230)	6.715***	6.765***	7.005***	6.455***
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 3.

Summary statistics for Study 2

Variable	Condition	Mean	St. Dev.	Min	Median	Max
Risk	Control	0.7	0.19	0.01	0.7	0.99
	Experimental	0.71	0.18	0.15	0.72	0.99
Transactions	Control	20.28	18.74	2	14	135
	Experimental	20.25	28.09	0	12	294
Vol_Avg	Control	13.8	4.92	1.43	14.58	20
	Experimental	13.52	4.73	0	14.75	20
Vol_Sum	Control	88,167.04	101,790.90	8,308.50	51,312.60	786,515.90
	Experimental	80,453.14	108,219.10	0	45,329.68	828,688.00
True rank	Control	53	21.26	20	55	90
	Experimental	52.12	21.86	20	51	90
Subj. rank	Control	48.95	21	1	50	100
	Experimental	40.13	20.88	1	38	83
Task perf. exp.	Control	46.06	18.16	1	45	100
	Experimental	41.91	17.57	1	41	90

Table 4.

Exploratory regression analyses – Study 2

	Dependent variable:						
	Risk						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	-0.001* (0.001)	-0.001* (0.001)	-0.002** (0.001)	-0.002** (0.001)	-0.001* (0.001)	-0.002** (0.001)	-0.002** (0.001)
Gender_Female	-0.037* (0.019)	-0.033* (0.019)	-0.022 (0.020)	-0.022 (0.020)	-0.029 (0.020)	-0.022 (0.020)	-0.017 (0.020)
SOEP	0.021*** (0.007)	0.020*** (0.007)	0.021*** (0.007)	0.021*** (0.007)	0.022*** (0.007)	0.019*** (0.007)	0.020*** (0.007)
Condition	0.020 (0.018)						0.007 (0.019)
Overplacement		0.001 (0.001)					-0.00000 (0.001)
Subjective rank			-0.001*** (0.0004)				
BTA_Below				-0.064*** (0.018)			-0.064* (0.035)
Objective score					0.003** (0.001)		-0.002 (0.003)
Task perf. expectation						-0.001*** (0.001)	-0.001 (0.001)
Constant	0.673*** (0.049)	0.676*** (0.048)	0.737*** (0.051)	0.713*** (0.048)	0.559*** (0.074)	0.759*** (0.054)	0.822*** (0.160)
AIC	-235.02	-235.02	-236.27	-242.63	-246.04	-238.58	-242.33
Observations	427	427	427	427	427	427	427
R^2	0.057	0.059	0.073	0.081	0.064	0.073	0.087
Adjusted R^2	0.043	0.046	0.060	0.067	0.051	0.059	0.065
Residual Std. Error	0.182 (df = 420)	0.182 (df = 420)	0.180 (df = 420)	0.180 (df = 420)	0.181 (df = 420)	0.180 (df = 420)	0.180 (df = 416)
F Statistic	4.192*** (df = 6; 420)	4.409*** (df = 6; 420)	5.526*** (df = 6; 420)	6.132*** (df = 6; 420)	4.813*** (df = 6; 420)	5.473*** (df = 6; 420)	3.944*** (df = 10; 416)
Note:							*p<0.1; **p<0.05; ***p<0.01

Figures

Histogram of overplacement scores, Study 1

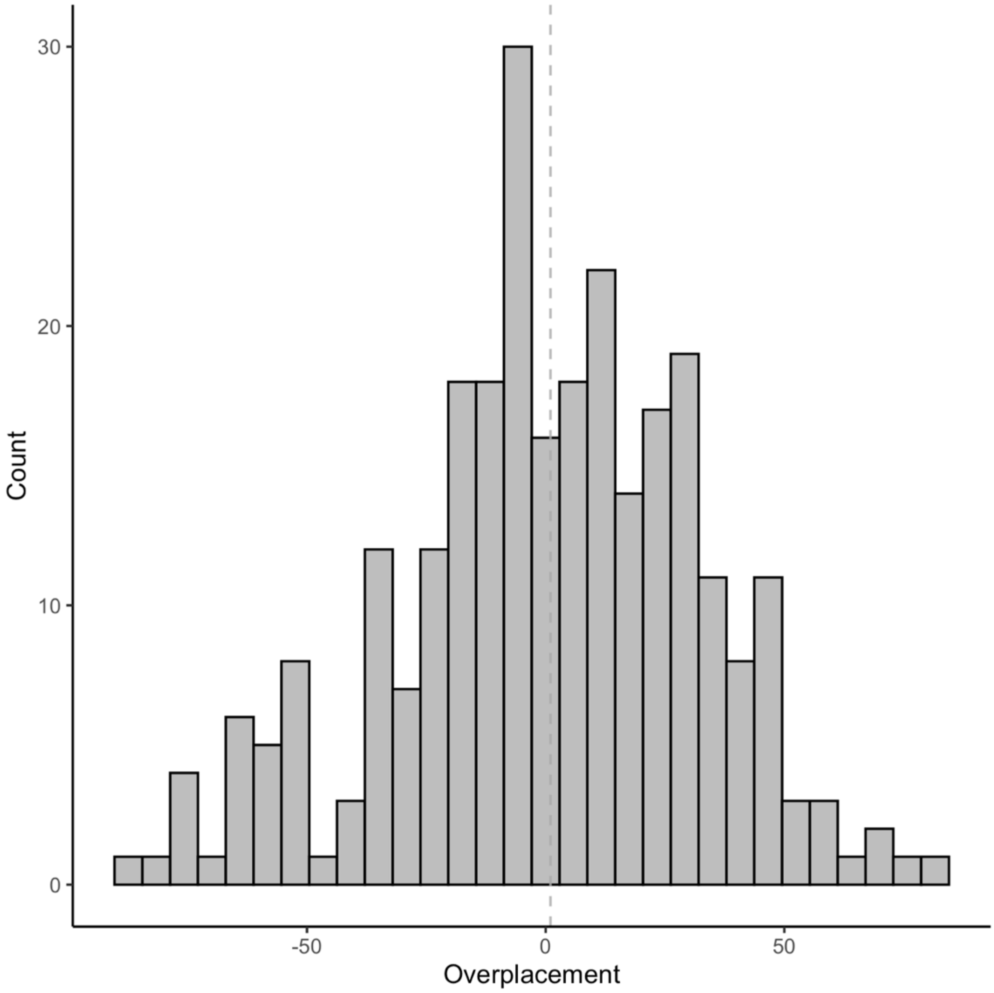


Figure 1. A histogram showing the distribution of overplacement in ranks. The dashed line marks the mean overplacement of 0.55.

Scatterplot of overplacement and aggregate volume in Study 1

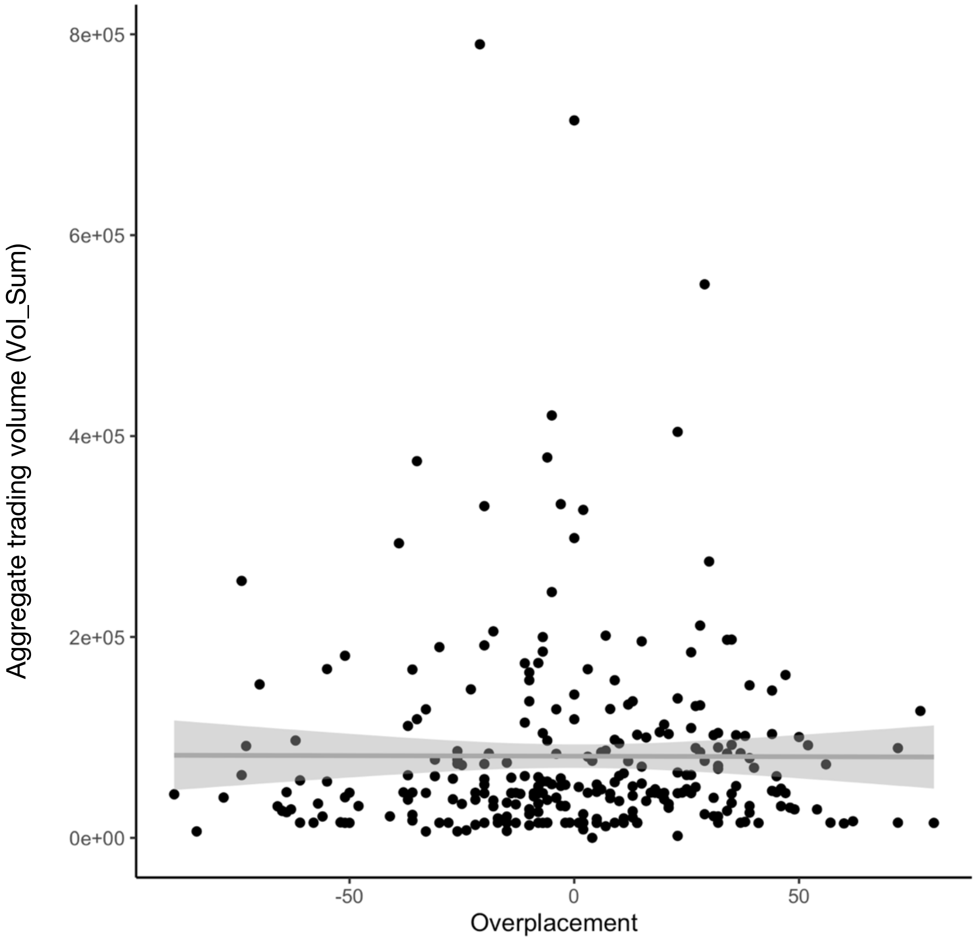


Figure 2. A scatterplot illustrating the relationship between overplacement and the aggregate volume of trades (*Vol_Sum*). A regression line is fitted to illustrate the bivariate relationship.

Scatterplot of overplacement and risk taking in Study 1

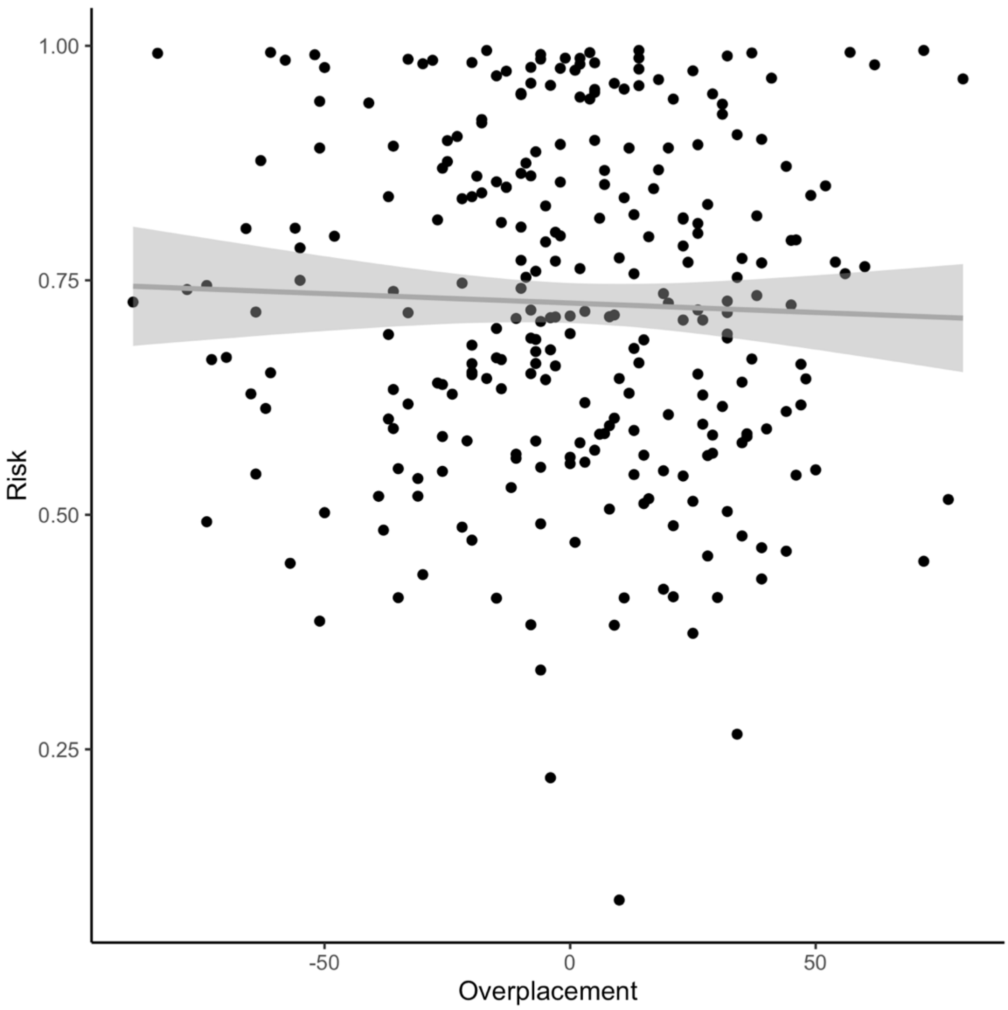


Figure 3. A scatterplot illustrating the relationship between overplacement and risk taking. The grey line is a regression line fitted to better illustrate the bivariate relationship.

Comparing better- and worse-than average participants in Study 1 – Risk

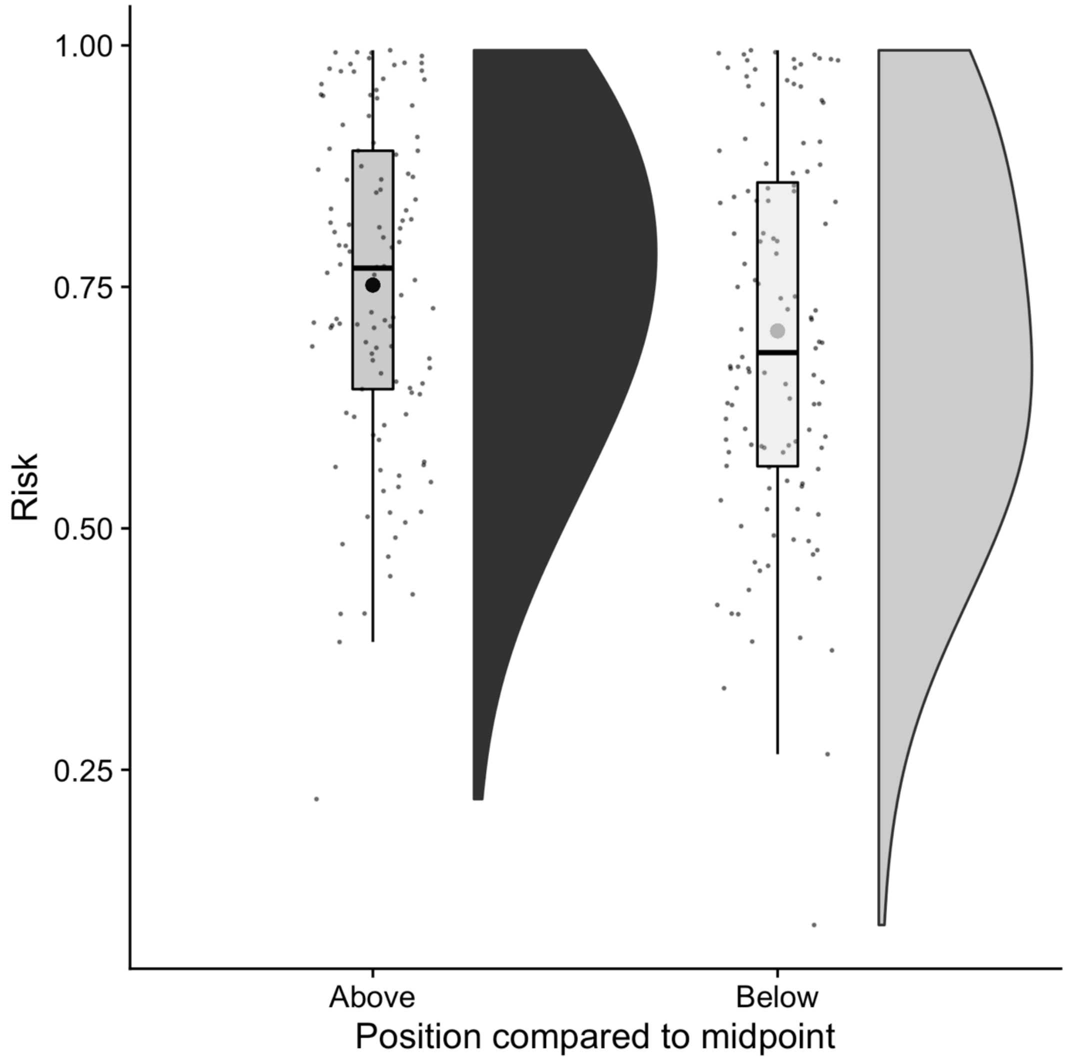


Figure 4. The distribution of values regarding *Risk* between individuals who believe they have performed better than average (higher than Rank 50) and those who believe they have performed worse (below Rank 51) shows substantial differences both in terms of means and medians.

Manipulation for Study 2

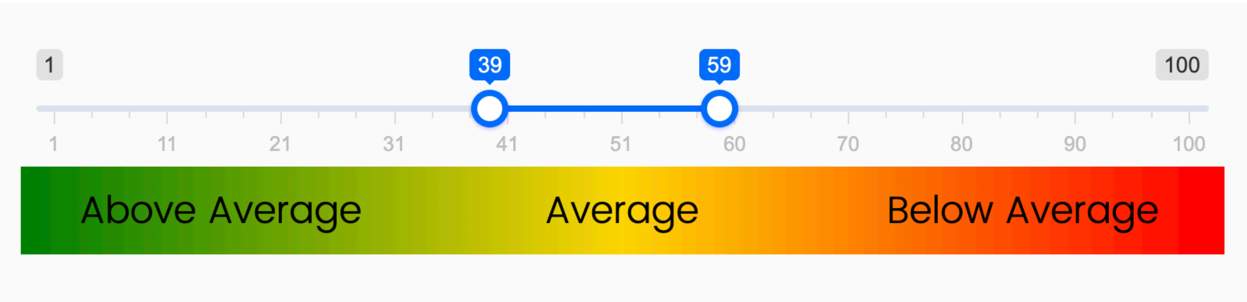


Figure 5: The interval feedback participants receive as part of the scale manipulation. The true rank of the participant receiving this feedback in the experimental condition would be 59. The true rank of a participant receiving this feedback in the control condition would be 49.

Distribution of true quiz performance rank and manipulation checks between conditions – Study 2

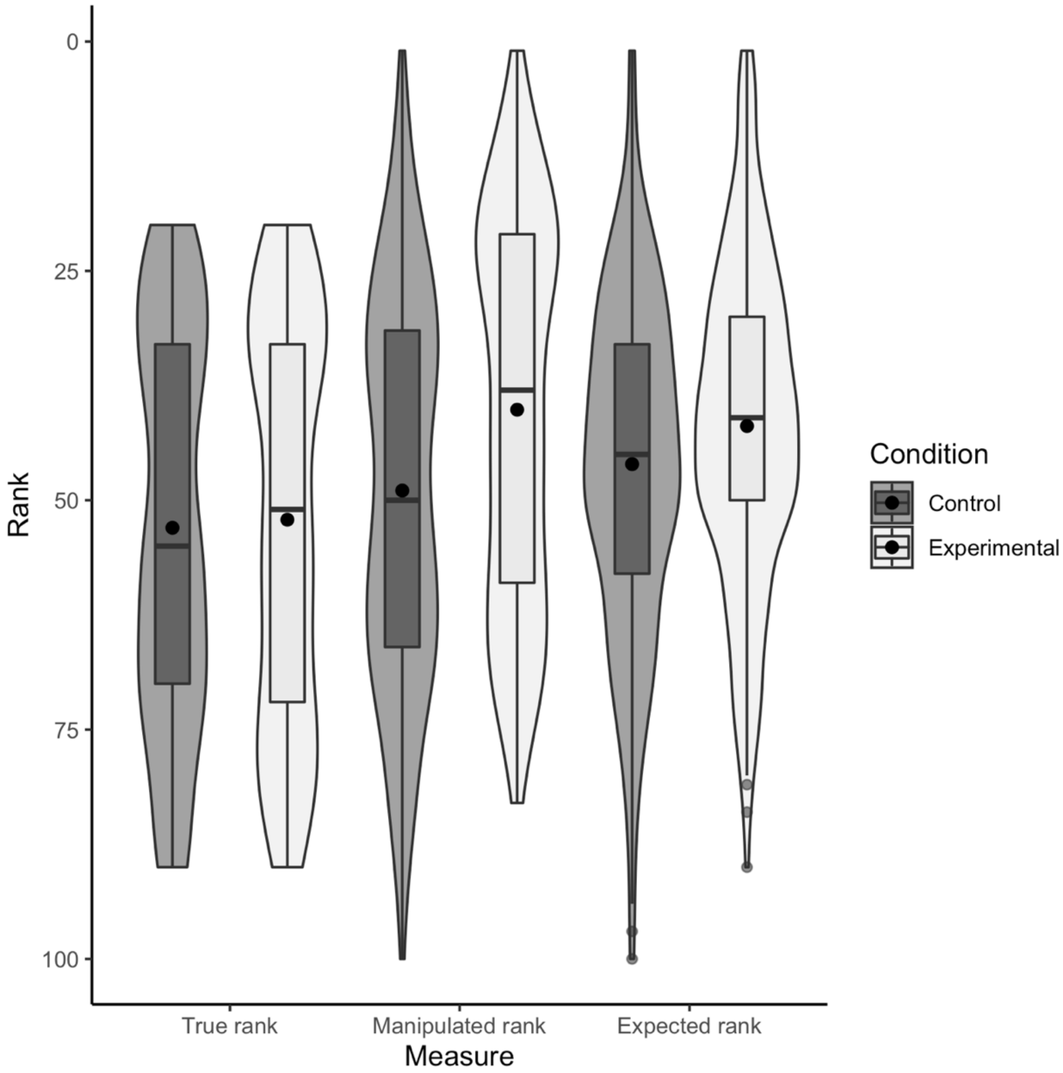


Figure 6. This figure illustrates the distribution of three variables by condition. *True rank* stands for a participant's actual position in the quiz task compared to our reference group (participants from Study 1). After providing our participants with our scale manipulation, we ask them to give us their exact position in the quiz task, as a manipulation check (*Manipulated rank*, or *Subjective rank*). Following this, we asked participants to indicate the relative performance they expect to attain in the trading task as a second manipulation check (*Expected rank* or *Task performance expectation*). We find significant differences between our manipulation checks in the expected direction (*Manipulated rank* & *Expected rank*).

Distribution of risk taking by condition – Study 2

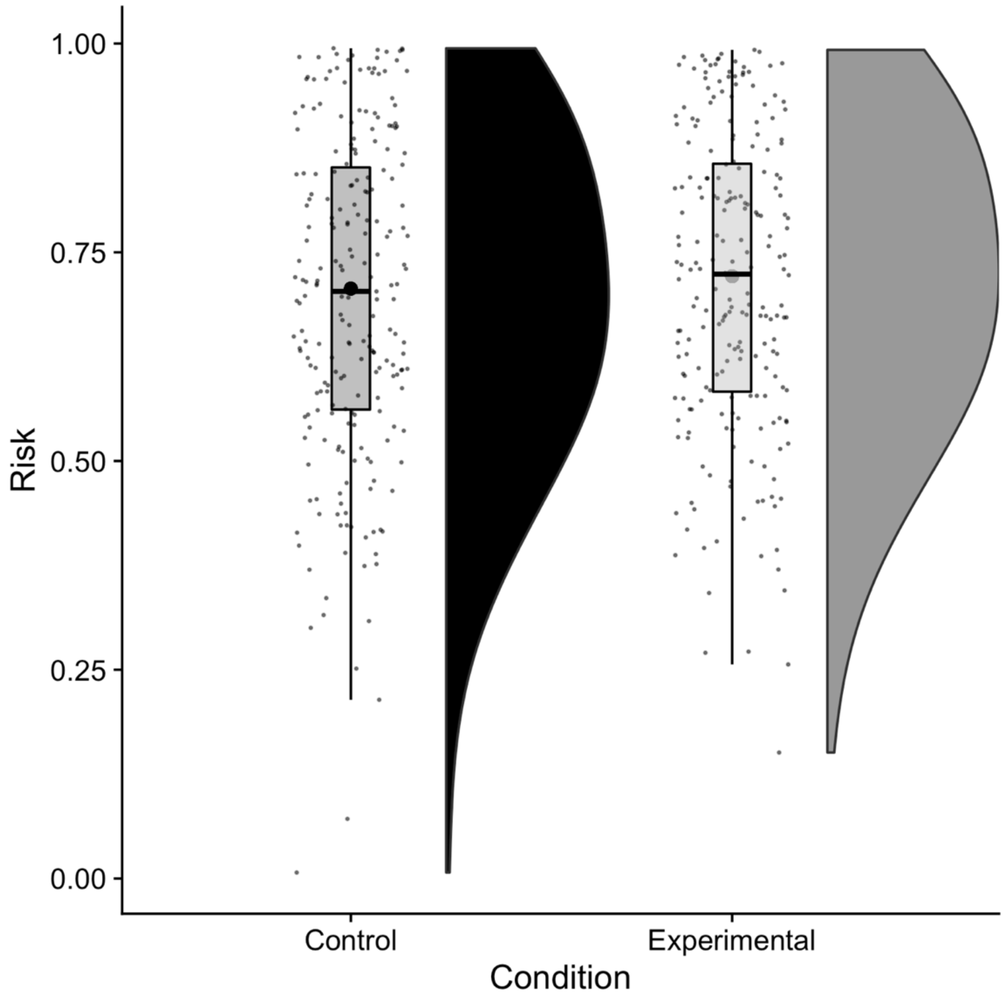


Figure 7. Plot illustrating the distribution of risk taking (*Risk*) in the trading task between conditions. The distributions overlap considerably and we do not document a significant mean difference.

Distribution of trading activity by condition – Study 2

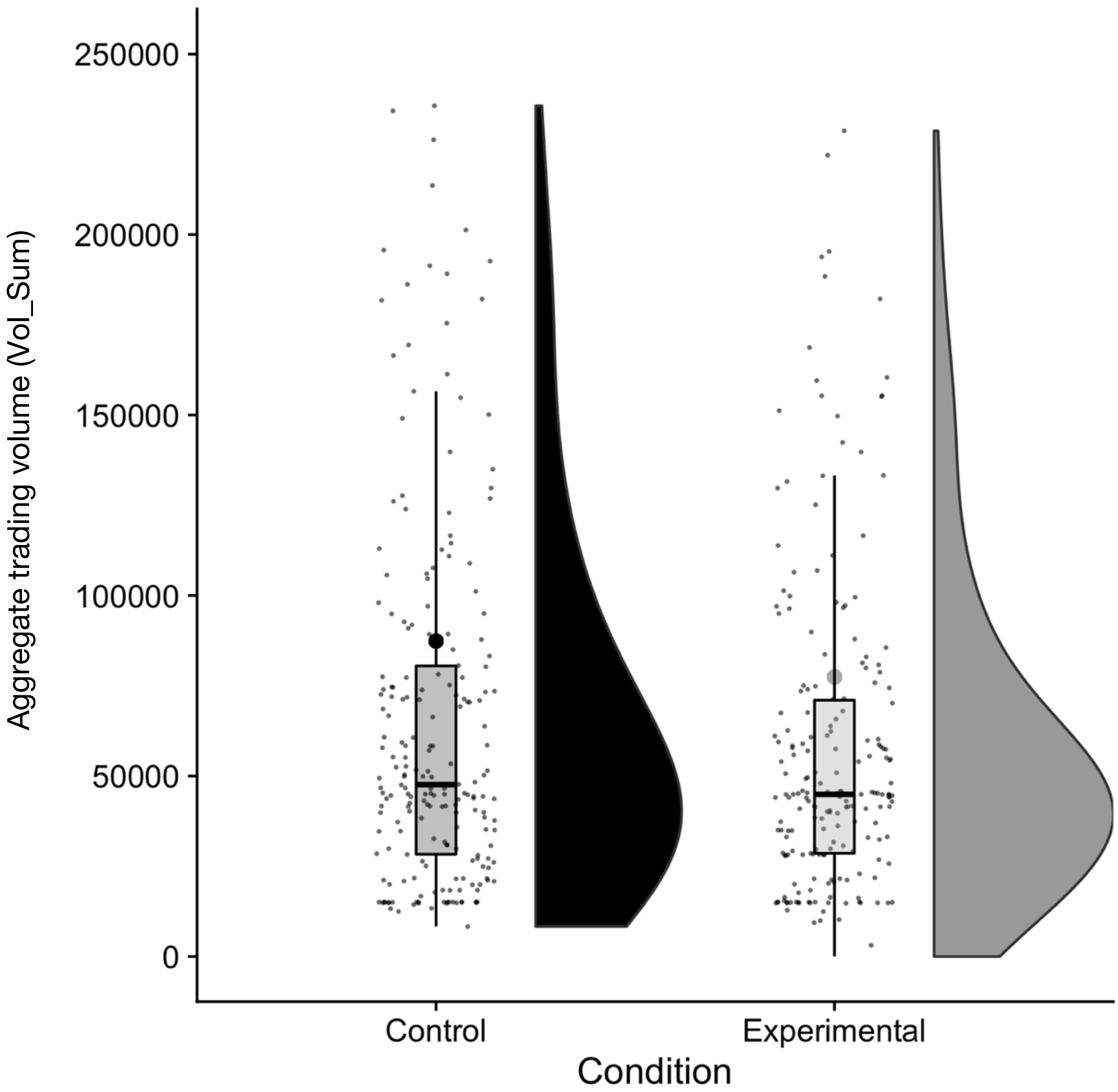


Figure 8. Plot illustrating the distribution of aggregate volume of shares traded (*Vol_Sum*, our focal dependent variable for trading activity) in the trading task between conditions. The distributions overlap considerably and we do not document a significant mean difference. The y-axis is censored at 250 000 due to the presence of substantial outliers in both conditions, which were included in all data analyses.

Appendix

Appendix I. – Screener survey

Please name one stock that is currently in your portfolio!: ___

Please tell us how your *ChoiceTextEntryValue* stock has performed over the past month in percent terms (rounded to the closest 10%)!

For example if the stock is now up 46% compared to 1 month ago, you should select 50% on the scale below!

<-70%

-70%

-60%

-50%

-40%

-30%

-20%

-10%

0%

10%

20%

30%

40%

50%

60%

70%

>70%

Which of the following companies do you own stock in directly, or have done so in the past 3 months? (Multiple answers possible)

Publix

Mars

Pilot Travel Centers

Liberty Mutual

Penske

Mass Mutual

Albertsons

State Farm

Panera Bread

None of the above

Please select which brokerage service provider you use to invest in stocks! (multiple answers are possible)

TradeFox

E-Stox

Forex.com

Wall Street Associates

Comdirect

Other

Please name the brokerage service you are using to trade stocks!: ___

Do you personally invest in the stock market?

Yes

No

Do you currently hold an investment in the stock market?

Yes

No

Which of the following statements describes the main function of the stock market?

The stock market results in an increase in the price of stocks

The stock market brings people who want to buy stocks together with people who want to sell stocks

None of the above

Which of the following statements is correct? If somebody buys a bond of firm B:

He owns a part of firm B

He has lent money to firm B

He is liable for firm B's debts

None of the above

Please complete the Captcha below to show that you are not a robot!

(Captcha dynamically generated)

Appendix II. – Information sheet, consent form and instructions to trading task

a.) Participant information sheet:

A study on how individuals perform in a simulated trading task

The goal of the present study is to investigate the role individual characteristics of retail investors play in how they manage their investments.
The study consists of a short quiz on financial literacy, followed by a trading task where you will receive an **endowment of 30 000 Experimental Currency (≈ $2.5)** to manage.
For every 12 000 in Experimental Currency you earn, you receive $1 in real money.
The more you make in the task, the more real money you earn.
In **addition, you will receive a flat fee of $0.5** for having participated in this short study.
After the task concludes, you will be asked to answer a few questions on your demographic characteristics and investment experience. The expected completion time of the study is **18 minutes.**
If you encounter problems submitting this HIT, please write an email to dkaszas@ethz.ch and report your problem there.

Further terms and information:

It will not be possible to connect your identity to the information you provide us in your survey answers. Only the responsible investigators and/or the members of the Ethics Commission will have access to this anonymous data under strictly observed rules of confidentiality.

Participating in this study does not present any known risks to participants. However, the trading software presents flashing images – if you have been diagnosed with epilepsy in the past, or have a family history of epileptic episodes, please do not participate in the study.

You are not obliged to complete the study, and can quit anytime without having to justify it to the requester. However, due to the design of our experiment, only participants who have completed the whole study can claim their compensation.

Possible damage to your health, which is directly related to the study and demonstrably the fault of ETH Zurich, is covered by the general liability insurance of ETH Zurich (Insurance Policy No. 30/4.078.362 of the Basler Versicherung AG). Beyond the aforementioned conditions, health insurance and accident insurance is the responsibility of the participant. The study is financed by the Dr. Donald C. Cooper-Fonds. If you have questions regarding the study, please contact the primary investigator: Dániel Kaszás, ETH Zurich, Chair of Cognitive Science, dkaszas@ethz.ch

b.) **Consent form:**

I have been informed about the aims and procedures of the study, the advantages and disadvantages, as well as potential risks of participating.

I have read and understood the information sheet for volunteers.

I was given sufficient time to make a decision about participating in the study.

I agree that the responsible investigators and/or the members of the Ethics Committee of ETH Zurich have access to the original data under strictly observed rules of confidentiality.

I participate in this study on a voluntary basis and can withdraw from the study at any time.

I recognize that participants who exit the experiment before its conclusion forfeit their right to compensation. By clicking "Proceed with the experiment", I agree to all of the statements above, and agree to participate in the study:

Proceed with the experiment

Exit experiment

c.) Instructions to trading task:

Please read the instructions for the trading task carefully!

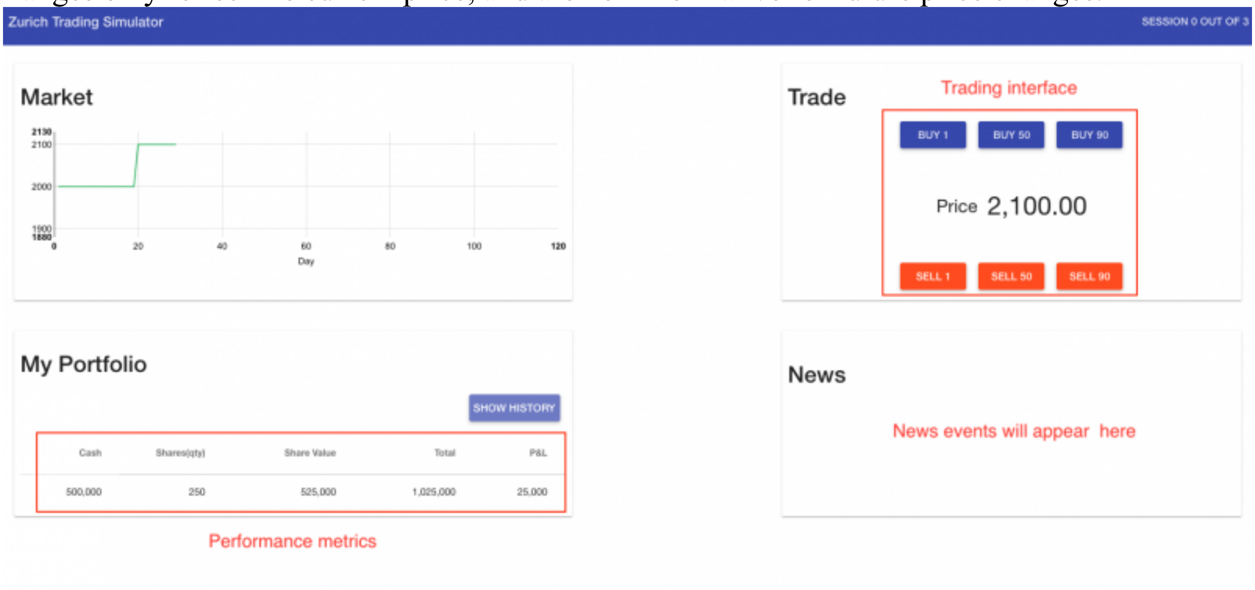
The trading task will be conducted using a browser-based trading interface (see the image below for an illustration). The task consists of two rounds: one practice round and one experimental round.

At the beginning of each round you will be endowed with an initial portfolio of shares of a

risky asset (Shares) and a safe asset (Cash), worth a combined 30 000 currency units.
50% of this endowment will be in the form of the risky asset (Shares), and 50% will consist of the safe asset (Cash). In each round, you can buy or sell Shares by clicking ”*Buy*” or “*Sell*” buttons of the corresponding size. Shares can fluctuate in value, while the value of Cash does not change.

The price development of the shares is based on historical daily closing prices of a real US market index. The news events presented are based on real contemporaneous news reports. Prices are predetermined, and your actions do not have an influence on market prices. Any trade that you make will be executed instantly at the current market price. You can buy or sell as many shares as you want, under the condition that you have sufficient funds for the transaction. No short-selling is possible.

The y-axis of the price chart adjusts automatically to the current market price.
These changes only reflect the current price, and are not informative for future price changes.



Appendix III. – Pre-task survey questions (incl. quiz)

How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

1 Not at all willing ; 2 ; 3; 4 Moderately willing; 5; 6; 7 Very willing

Please answer the following quiz questions.

Some questions allow multiple choices, and some questions are open-ended.

If you do not know the answer to an open-ended question, just type **NA** in the text box.

The scoring for the open-ended questions is case-sensitive.

Please do not use any external sources to answer these questions – doing so would harm the validity of the study, and given the limited time allotted to answer the questions, it would not contribute to your performance.

You have a maximum of 7 minutes to answer all questions.

Please select all the options that are TRUE.

Higher reported revenues always mean higher profits.

Reporting an increase in quarterly earnings is always followed by an increase in the share price of a company.

All else being equal, an increase in earnings per share (EPS) is a good sign for the price of a stock

Generally, a high Price-to-Earnings (PE) ratio means that investors are anticipating higher growth in the future.

Please select all the options that are TRUE.

Increasing consumer confidence is a positive indicator for the economy.

The VIX Index measures volatility expectations

A moderate increase in the price of consumer goods is often associated with an improving economy.

Please select all the options that are TRUE.

Disney is in the Russell 2000 index for small-cap stocks.

Facebook is headquartered in California.

Apple’s stock price has more than doubled since October 1st, 2018.

Verizon is a publicly traded company.

NYSE stands for: ___

Select the smallest company by market capitalization, as of today!

Apple

Netflix

Bed Bath & Beyond

Which is the best definition of "selling short"?

Selling shares of a stock shortly after buying it

Selling shares of a stock before it has reached its peak

Selling shares of a stock at a loss

Selling borrowed shares of a stock

Don't know / Not sure

Which financial instrument(s) would you use to gain exposure to a single company's performance?

Index ETFs

Options

Treasuries

Individual stocks

Please select all the options that are TRUE

Considering a long time period (e.g. 10 or 20 years) from the three following assets, bonds historically give the highest return: savings accounts, bonds or stocks.

From the following 3 assets, stocks represent the highest risk: savings accounts, bonds or stocks.

Please select all the options that are TRUE

When an investor spreads money among different unrelated assets, the risk of losing money increases.

If you were to invest $1000 in a stock fund, it would be possible to have less than $1000 when you decide to withdraw or move it to another fund.

High yield bond funds are invested in bonds with strong credit ratings.

If you buy a company's stock...

You own a part of the company.

You have lent money to the company.

You are liable for the company's debts.

The company will return your original investment to you with interest.

Don't know / Not sure

Which of the following organisations insures all investors against losses in the stock market that take place during ordinary trading?

FDIC (Federal Deposit Insurance Corporation)

FINRA (Financial Industry Regulatory Authority)

SEC (Securities and Exchange Commission)

None of the above

Don't know / Not sure

What has been the average annual return on a very broad US stock index market investment over the last two decades, in percentages?

-20% or less

Between -20% and -15%

Between -15% and -10%

Between -10% and -5%

Between -5% and 0%

Between 0% and 5%

Between 5% and 10%

Between 10% and 15%

Between 15% and 20%

+20% or more

You hear rumors that Italy is planning a surprise exit from the Eurozone tomorrow morning. How do you expect Italian financial markets to react? (Mark all options which are TRUE)

The price of Italian government bonds will decline.

The interest rate on Italian government bonds will increase.

The Italian stock market index FTSE MIB will increase.

You get news that a new Coronavirus vaccine protected equally well against the virus as a placebo in Phase III trials. How do you think the shares of the company developing the vaccine will react?

Shares will go up

Shares will go down

The news is irrelevant

Appendix IV. Items eliciting ranking and expected performance in Study 1

Please provide an estimate for your performance on the **previous quiz task**, compared to other participants:

Rank 1 means that you expect to have the **best performance from 100 participants.**

Rank 100 means that you expect to have the **worst performance from 100 participants.**

Best **Average** **Worst**

1 11 21 31 41 51 60 70 80 90 100

Please provide an estimate for your performance on the **upcoming trading task**, compared to other participants:

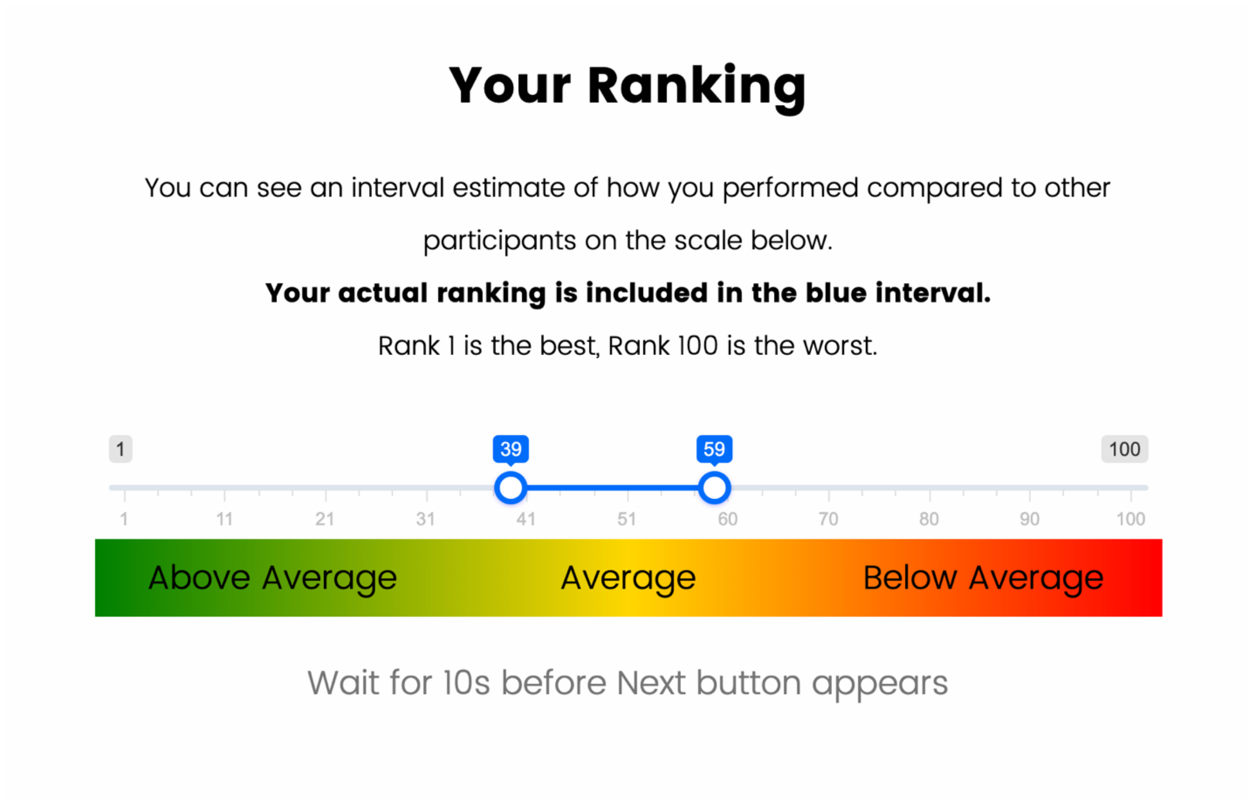
Rank 1 means that you expect to have the **best performance from 100 participants.**

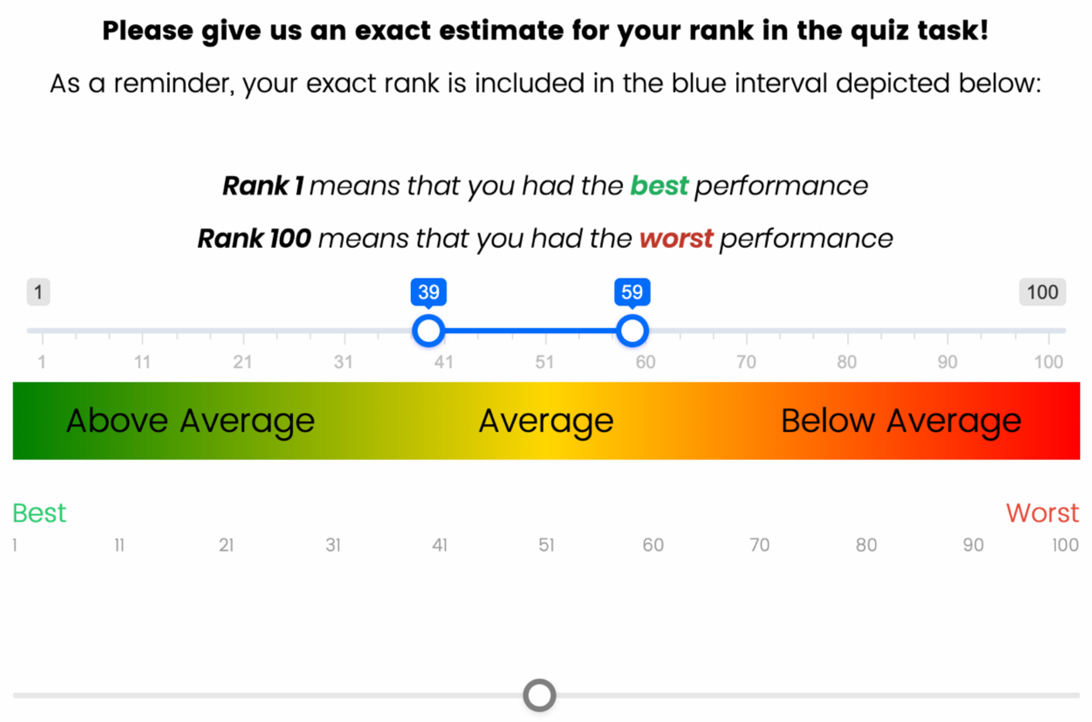
Rank 100 means that you expect to have the **worst performance from 100 participants.**

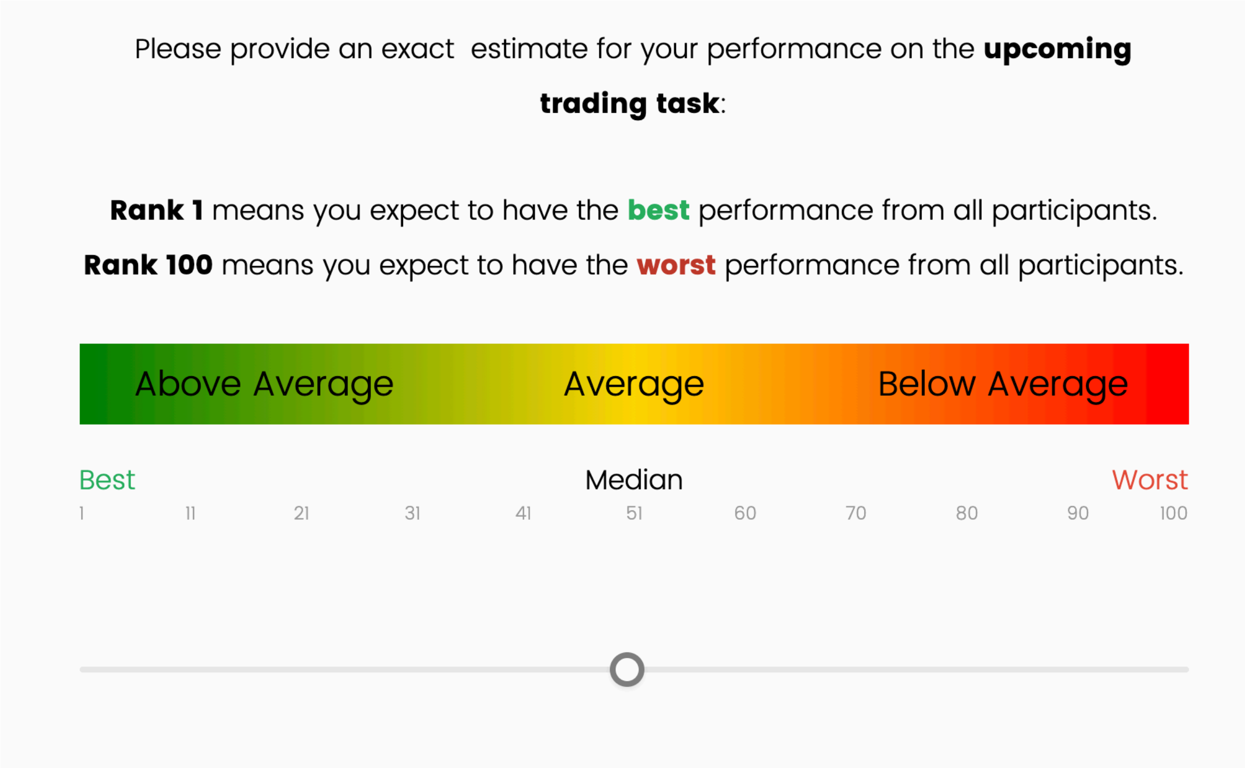
Best **Average** **Worst**

1 11 21 31 41 51 60 70 80 90 100

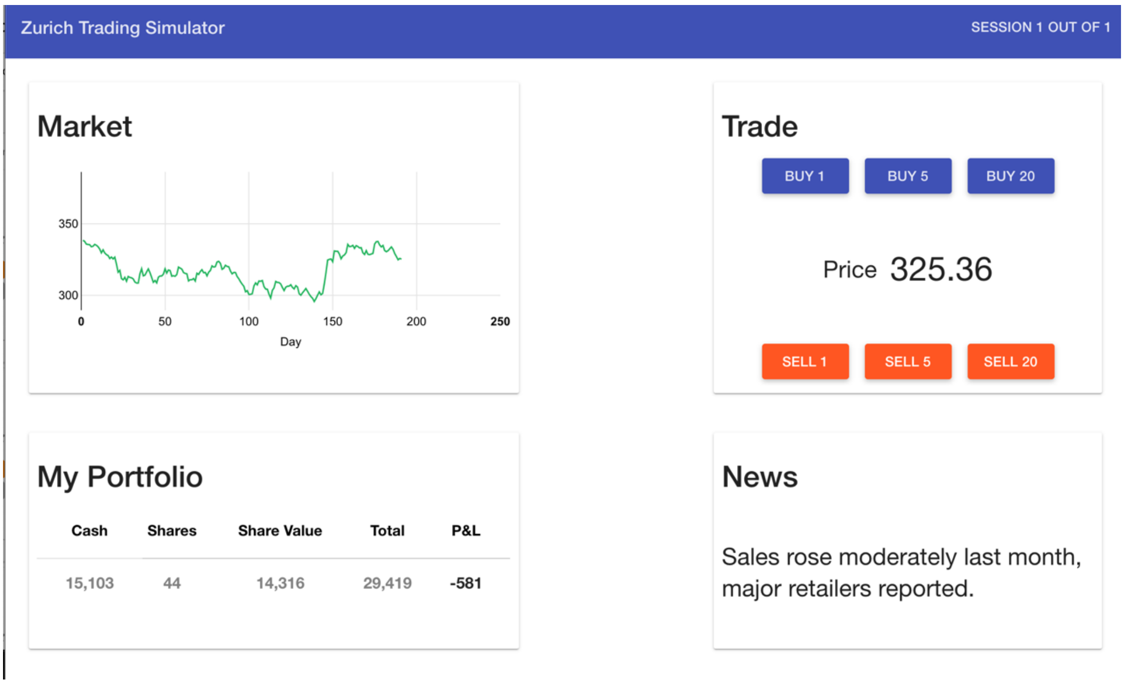
Appendix IV. – Manipulation for Study 2



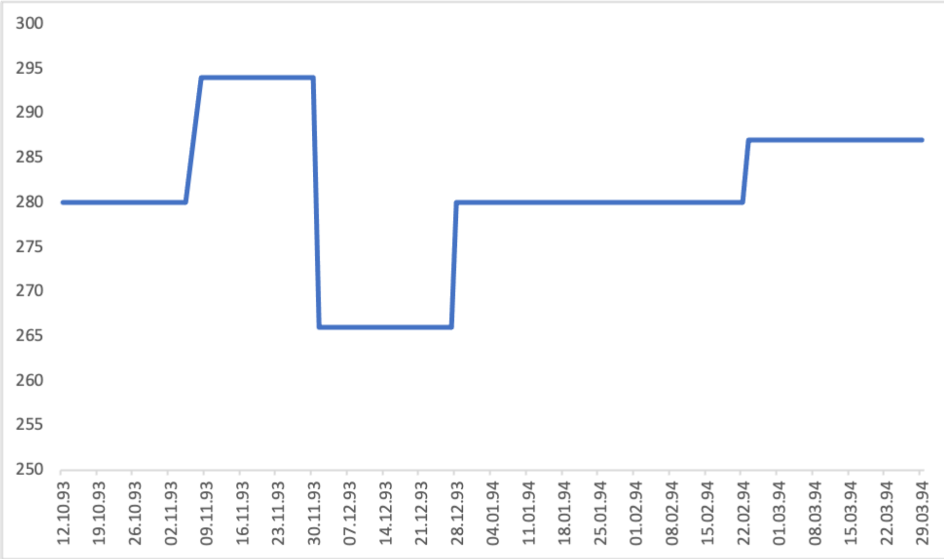




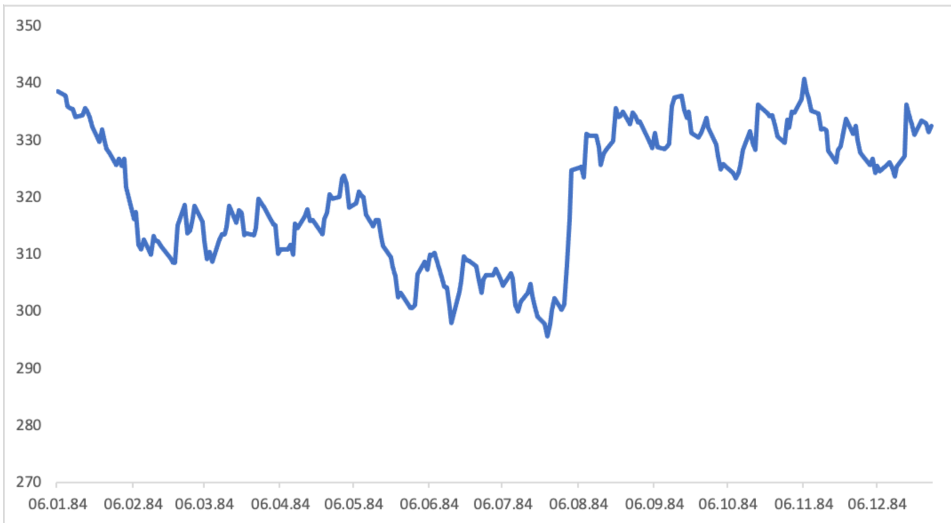
Appendix V. – Picture of software, price paths and news



Price path for practice round:



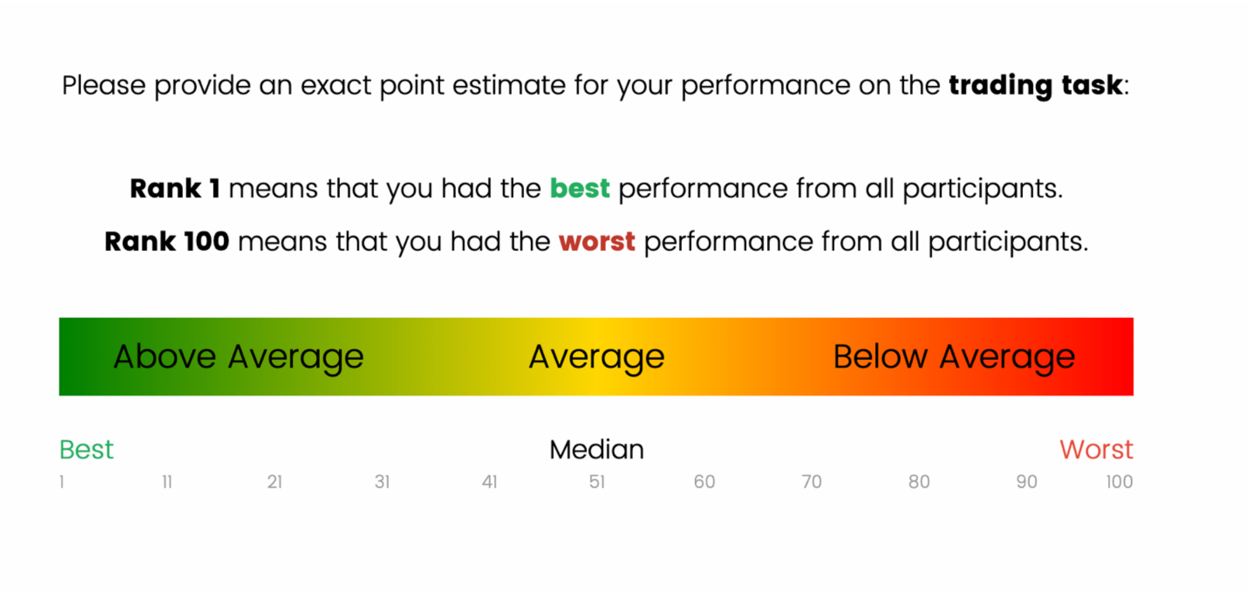
Price path for experimental round:



List of news events displayed in the experimental round

Date	News	Source
09.01.1984	Stock prices edged downward, with many investors cashing in on recent profits.	New York Times Business Digest
18.01.1984	New-home construction fell 5 percent last month, but for all of the past 12 months it soared 60 percent, the Commerce Department reported.	New York Times Business Digest
06.02.1984	Having slipped sharply, the stock market may fall farther in the days ahead. But with an eye on underlying value and encouraging economic signals, a turnaround still seems possible.	Christian Science Monitor - Nexis Uni
24.02.1984	Factory orders for durable goods rose 1.1 percent in January, helped by a large jump in demand for steel and other metals	New York Times Business Digest
12.03.1984	The expansion continues in all Federal Reserve districts, but with variation in vigor among regions and sectors.	Beige Book – Minn. Fed
29.03.1984	The index of leading economic indicators rose 0.7 percent last month. This was less the preceding month's 1 percent gain, but it was still the 17th rise in 18 months	New York Times Business Digest
18.04.1984	AT&T, the US telecommunications group cuts annual earnings forecast.	Financial Times - Nexis Uni
01.05.1984	Stock prices rose sharply in the heaviest trading in six weeks. Institutional buyers came to the market in what was described as bargain-hunting	New York Times Business Digest
21.05.1984	Stocks fell broadly, putting the Dow average at its lowest in more than a year.	New York Times Business Digest
04.06.1984	The economy expanded at a slower pace during the past month, the National Association of Purchasing Management reported	New York Times Business Digest
13.06.1984	A Chemical Weekly survey shows that Wall Street analysts have revised upward their estimates for chemical company earnings.	Chemical Week – Nexis Uni
25.06.1984	Auto sales surged 22.6 percent this month, compared with a year earlier. But some analysts believe sales will begin to slow	New York Times Business Digest
12.07.1984	Recent market moves reflect the market's low tolerance for surprises.	
01.08.1984	The stock market scored its biggest gain in six weeks, based on hopes interest rates will ease.	UPI Archives
03.08.1984	Stock prices skyrocketed in record trading volume as institutional investors jumped in on the action in the belief interest rates were headed lower.	San Diego Union Tribune – Nexis Uni
20.08.1984	The optimists see the new stock-market high soon. More-cautious analysts warn that a continued upward surge of prices will depend on whether interest rates drop as investors expect.	U.S. News and World Report – Nexis Uni
04.09.1984	Major manufacturers' spending plans rose 38 percent in the past quarter, to a new record.	New York Times Business Digest
20.09.1984	Construction of new houses fell 12.8 percent in the previous month, the lowest in almost two years.	New York Times Business Digest
04.10.1984	September sales rose moderately, major retailers reported.	New York Times Business Digest
19.10.1984	Alcoa posts slight gain in earnings.	Washington Post – Nexis Uni
09.11.84	In the past month, almost 75 percent of corporate insiders were sellers, notes the newsletter on the topic, The Insiders.	San Diego Union Tribune – Nexis Uni
07.12.1984	Wall Street closed moderately lower on the final day of a dismal week of trading as investors remained unwilling to bid up even the price of blue chip stocks because of economic and monetary uncertainty.	Financial Times – Nexis Uni

Appendix VI. – Post-task survey



How do you feel about your performance in the trading task?

1 Very positively; 2; 3; 4 Neutral; 5; 6; Very negatively

Please provide an estimate for how much money you think the average participant ended up with after the trading task! Remember, you start with an endowment of 30 000 Experimental Currency Units: ___

Please provide an estimate for how much money the following two participants ended up with!

Remember, you start with an endowment of 30 000 Experimental Currency Units.

A participant **ranked number 10** from 100 earned:

A participant **ranked number 90** from 100 earned:

Gender:

Male

Female

Other

Prefer not to say

Household income:

Less than $20,000

$20,000 to $34,999

$35,000 to $49,999

$50,000 to $74,999

$75,000 $99,999

Over $100,000

Prefer not to say

Highest level of education completed:

Less than high school diploma

High school degree or equivalent (e.g. GED)

Some college, no degree

Associate degree (e.g. AA, AS)

Bachelor's degree (e.g. BA, BS)

Master's degree (MA, MS, Med)

Professional degree (e.g. MD, DDS, DVM)

Doctorate (e.g. PhD, EdD)

Prefer not to say

Employment status:

Employed full time (40 or more hours per week)

Employed part time (up to 39 hours per week)

Unemployed and currently looking for work

Unemployed and currently not looking for work

Student

Retired

Homemaker

Self-employed

Unable to work

Prefer not to answer

Marital status

Single (never married)

Married, or in a domestic partnership

Widowed

Divorced

Separated

Prefer not to answer

Could you identify which asset was presented to you in the previous task, AND around what year?

Yes

No

If Yes:

Name of asset: __

Which year did the presented market come from? (format: yyyy): ___

Appendix VII. – Debriefing for Study 2

Debriefing

The goal of this experiment was to investigate the effect of relative overconfidence on trading behavior. Relative overconfidence means that one has a higher opinion of oneself compared to his or her peers as one should, based on available evidence. In order to examine this question, participants have been randomly assigned to one of two conditions, where they received noisy feedback on their quiz performance:

Participants in one condition (Control condition) were provided with a span of potential performance rankings that was centered on that participant's actual position. For example, if a participant in this condition ranked 50 from 100 participants in his or her performance on the quiz task, this participant was presented with a span from rank 40 to rank 60.

In the other condition (Experimental condition), participants also received feedback on their performance on the quiz task, but this feedback was skewed. As stated in the instructions, the presented span of potential rankings contained each participants' true ranking, but in the experimental condition this true ranking constituted the lowest possible ranking in this span. For example, if a participant in this condition ranked 50 from 100 participants in his or her performance on the quiz task, this participant was presented a span from rank 50 to rank 30. As individuals have a tendency to consider the middle of a distribution as most informative, this method was intended to create a feeling of overconfidence in the experimental condition.

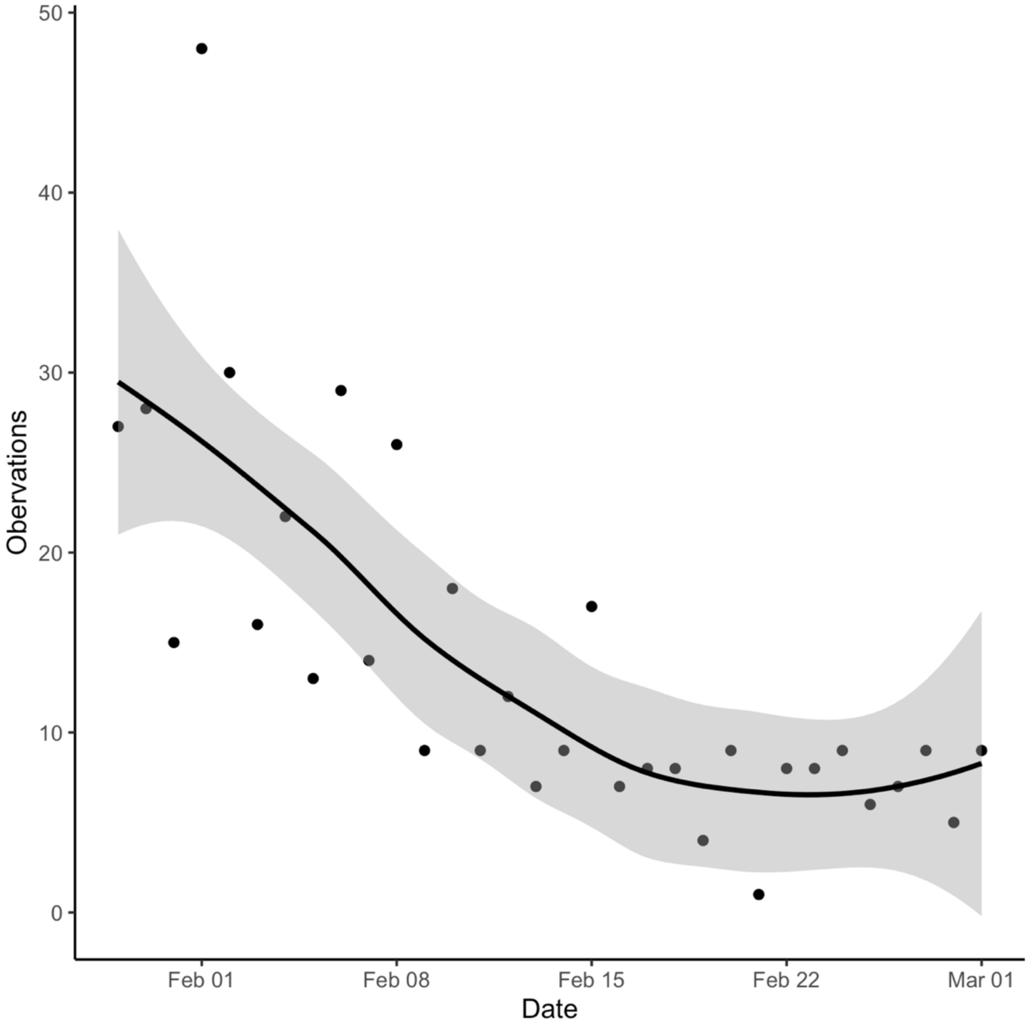
You were in the {control} {experimental} condition.

Your true rank in the quiz task, where Rank 1 is the best and Rank 100 is the worst: X

Click on the button below to finish the experiment and receive your Secret Key for MTurk

Appendix VIII. Data collection over time – Study 2

Data collection over time – Study 2



The development of data collection over time. Due to a substantial slowing down in the pace of data collection in the second half of February the decision was made to halt data collection before the pre-registered sample size was reached.

Study 3: Robo-investment aversion

Paweł Niszczota[1,*], Dániel Kaszás[2]

[1] Poznań University of Economics and Business, Department of International Finance, Poznań, Poland

[2] ETH Zürich, Chair of Cognitive Science, Zürich, Switzerland

* Corresponding author: Paweł Niszczota, Poznań University of Economics and Business, Department of International Finance, al. Niepodległości 10, 61-875, Poznań, Poland; E-mail address: pawel.niszczota@ue.poznan.pl; ORCID: 0000-0002-4150-3646

Open Science Practices

All studies have been pre-registered. The pre-registration documents, data and materials are available at: https://osf.io/gdp42/

Acknowledgments

This research has been supported by grants 2018/31/D/HS4/01814 and 2018/02/X/HS4/01703 from the National Science Centre, Poland.

This paper has been published in PLOS One

Niszczota, P., & Kaszás, D. (2020). Robo-investment aversion. *PLOS One*, *15*(9), e0239277.

Abstract

In five experiments (N = 3,828), we investigate whether people prefer investment decisions to be made by human investment managers rather than by algorithms ("robos"). In all of the studies we investigate morally controversial companies, as it is plausible that a preference for humans as investment managers becomes exacerbated in areas where machines are less competent, such as morality. In Study 1, participants rated the permissibility of an algorithm to autonomously exclude morally controversial stocks from investment portfolios as lower than if a human fund manager did the same; this finding was not different if participants were informed that such exclusions might be financially disadvantageous for them. In Study 2, we show that this robo-investment aversion manifests itself both when considering investment in controversial and non-controversial industries. In Study 3, our findings show that robo-investment aversion is also present when algorithms are given the autonomy to increase investment in controversial stocks. In Studies 4 and 5, we investigate choices between actual humans and an algorithm. In Study 4 – which was incentivized – participants show no robo-investment aversion, but are significantly less likely to choose machines as investment managers for controversial stocks. In contrast, in Study 5 robo-investment aversion is present, but it is not different across controversial and non-controversial stocks. Overall, our findings show a considerable mean effect size for robo-investment aversion (d = –0.39 [–0.45, –0.32]). This suggests that algorithm aversion extends to the financial realm, supporting the existence of a barrier for the adoption of innovative financial technologies (FinTech).

Keywords: robo-advice; algorithm aversion; sin stocks; socially responsible investing; FinTech

Introduction

We are seeing a strong growth of algorithmic (systematic) investment funds (Abis, 2017; Harvey, Rattray, Sinclair & Hemert, 2017) and the robo-advisory industry (D'Acunto, Prabhala & Rossi, 2019; Kaya, Schildbach & Schneider, 2017; Moulliet, Majonek, Stolzenbach & Völker, 2016). Computer algorithms run funds that account for 35% of the US stock market, and are responsible for 60% of the trading that happens on it (The Economist, 2019). Investment decisions are also increasingly perceived to have moral undertones. In January 2020, the world's biggest asset management fund declared that it would not invest in companies that receive more than a quarter of its revenues from the production of coal (Pettifor, 2020). This move is in line with the notion that fossil fuel producers listed on stock markets might now be considered "sin stocks": morally controversial companies whose activities cause much social harm (Hong & Kacperczyk, 2009; Trinks, Scholtens, Mulder & Dam, 2018).

In this paper, we investigate whether people are averse to "robo" investment managers making investment decisions, and explore whether machines have a "moral mandate" that would give them legitimacy to perform investment decisions that concern controversial companies. While the study does not deal with investment mechanisms currently available to the average individual 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 (Moulliet, Majonek, Stolzenbach & Völker, 2016). A number of automated investment services already offer "ethical" portfolios – however, at a considerable price premium due to the necessity of keeping human fund managers in the loop (Fantato, 2018). 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 (i.e., we are still many years from the full automation of road vehicles), researchers find that it is important to know what autonomous vehicles should do when facing a moral dilemma (Awad et al., 2018; Bonnefon, Shariff & Rahwan, 2016). Given the increasingly important role ethical and environmental considerations play in today's investment landscape (Choi, 2016; van Duuren, Plantinga & Scholtens, 2016), it is worth considering if the trend towards integrating moral considerations into investment strategies and the increasing automation of investments are potentially on a collision course (Beioley, 2018).

Hypothesis development

Recent evidence suggests that individuals show an aversion towards computer algorithms under certain conditions. When given the possibility to choose between advice provided by a human or an algorithm, people tend to display a preference for the former and thus exhibit algorithm aversion (Castelo, Bos & Lehmann, 2019; Dietvorst, Simmons & Massey, 2015; Dietvorst, Simmons & Massey, 2016; Longoni, Bonezzi & Morewedge, 2016). This algorithm aversion in the original sense can emerge in various contexts, but a common component is that people have to witness the algorithm making a mistake. In another strand of literature, Bigman and Gray (2018) have demonstrated via a series of experiments that decisions with a moral component are perceived to be the domain of humans and not "machines", showing machine-aversion in the moral domain – a tendency that was also documented by Gogoll and Uhl (2018). This effect is not predicated on the presence of previous errors, but is specific to scenarios which have a significant moral component, including the possibility of serious negative externalities (medical- and military decisions, for example).

When considering if there are further domains where decisions are often perceived to have a moral character, the world of finance and investments comes to mind. Recent evidence on how moral considerations can affect asset prices (Hong & Kacperczyk, 2009; Durand, Koh & Limkriangkrai, 2013) and experimental work on the intersection of morality and markets (Sutter, Huber, Kirchler, Stefan & Walzl, 2019) suggest that markets and investments have an inherent moral dimension to them. The increasing demand for investments in companies representing environmentally sustainable and ethical business practices (van Duuren, Plantinga & Scholtens, 2016) shows that investors are increasingly mindful of the potential negative externalities associated with their choices. Given the parallel trend for morally sound investment practices and the automation of investment decisions (D'Acunto, Prabhala & Rossi, 2019; Kaya, Schildbach & Schneider, 2017), questions emerge on the significance of algorithm aversion in investment. Our general hypothesis is that when given the decision to take investment advice from either a human or an algorithm, we expect to observe algorithm aversion. Due to the inherent moral character of investment decisions, furthermore, we hypothesize that this effect will be stronger in cases where the investment opportunities have a strong negative moral valence ("sin stocks").

In contrast to the definition of algorithm aversion which is widely accepted in the literature (Dietvorst, Simmons & Massey, 2015; Dietvorst, Simmons & Massey, 2016), we do not link aversion to the success or failure of the algorithms in question. Our approach is conceptually related to Bigman and Gray's (2018) and Castelo and colleagues' findings (Castelo, Bos & Lehmann, 2019) on people exhibiting a generalized aversion to machines making decisions in tasks which require subjective judgments, in our case moral judgments.

In essence, we are asking whether individual investors will be comfortable ceding control over their investments to algorithms, and if considerations of morality play a role.

Overview of experimental studies

In Studies 1-3, our main goal was to test whether people exhibit an aversion towards computer algorithms making investment decisions for them. To test this hypothesis, we implemented 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 these experiments were intended to test the main hypothesis, each experiment was also meant to investigate the conditions that could affect the strength of the proposed effect. In Studies 4-5, we investigate whether algorithm aversion emerges in a simulated investment scenario, where participants have to choose between real human advisors recruited via Mechanical Turk and an actual (artificial neural network) algorithm.

In Study 1, we investigated whether there is an aversion towards computer algorithms making investment decisions (which, for simplicity, we henceforth call "robo-investment aversion"), by investigating the right to exclude stocks from controversial industries from an investment portfolio. In this study, we also wanted to assess whether permissibility ratings might be different if participants are informed that excluding stocks might lead to worse 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 indeed hurt portfolio performance (returns).

In Study 2, we tested whether robo-investment aversion is present both when contemplating the right to exclude companies from controversial industries and non-controversial industries.

In Study 3, we further investigated how generalizable robo-investment aversion is, by simultaneously looking at the permissibility to exclude (just as in Studies 1 and 2), but also to invest more heavily in controversial companies.

In Studies 4 and 5, we pitted human advisors against an algorithm, by asking participants to choose between relatively knowledgeable humans and an artificial neural network algorithm. Crucially, we used a between-subjects design that randomly assigned participants to either controversial or non-controversial stocks, allowing us to more fully test whether an aversion towards algorithms is exacerbated in the case of controversial stocks (consistent with an attenuation in the use of algorithms for investments with moral undertones). In Study 4, we tested the preference for human versus algorithm investment advice in an incentivized experiment, where participants received a bonus if they chose the advisor that was more accurate. In Study 5, we tested whether people have a preference for knowledgeable humans over an algorithm, after finding out that they had similar overall performance in a prior task.

In all experiments, we recruited subjects through Mechanical Turk, and it is worth discussing some issues relating to this design choice. Firstly, a number of studies support the suitability of Mechanical Turk samples for experimentation (in general, Berinsky, Huber & Lenz, 2012; Thomas & Clifford, 2017; Peer, Brandimarte, Samat & Acquisti, 2017; and in finance-related issues as well; König-Kersting, Pollmann, Potters & Trautmann, 2017; Gerhard, Hoffmann & Post, 2017). Secondly, Mechanical Turk samples seem to be more representative of the general population than student samples (Paolacci, Chandler & Iperiotis, 2015), and thus seem more suitable when representing the preferences of people that are owners of companies via pension funds or mutual funds.

Data, materials and pre-registration documents for all studies are available on OSF at https://osf.io/gdp42/?view_only=f6eeccb5e9d34f84928f58c601a7c66f.

Study 1: Exclusion of controversial stocks

In the first study, we hypothesized that people have an aversion towards computer algorithms excluding controversial companies from stock portfolios, compared to human fund managers.

Method

Participants

We recruited 522 participants via Mechanical Turk, in order to obtain 90% power to detect a small-to-medium effect (f = .15). We excluded 56 participants who did not correctly answer an attention check question. One participant did not answer one (out of three) statement, and thus we computed the dependent variable based on the available answers. We analyzed the data of 466 participants, 261 (56%) of which were female. The mean age of participants was 39.2 years (*SD* = 11.9).

Procedure

The study had a 2 (first between-subjects factor: portfolio choice made by human fund manager vs portfolio choice made by a computer algorithm) × 2 (second between-subjects factor: absence of penalty resulting from exclusion vs presence of penalty) design. Participants had to imagine that a fund manager or a computer algorithm (depending on which condition they are assigned to) has the autonomy to select companies to stock portfolios for moral reasons, i.e. has the autonomy to exclude (sell or not invest in) companies from a selection of controversial industries. These were presented in the first part of the study, for participants to have a reference point, using the following text:

> Would you feel comfortable if your pension fund invested in companies from the following industries? Please rate on a scale of 1 (not at all) to 7 (completely).

Participants assigned to the "penalty present" condition additionally read that the exclusion of controversial stocks might lead to a slight reduction in the expected return of the portfolio. The exact instructions are presented below, with the text shown in italics being varied in human vs robo condition, and text in the square brackets being presented only to participants that were in the "penalty present" condition:

> Imagine that a *fund manager/computer algorithm* had the ability to exclude (sell or not invest in) companies from some of the industries that were listed on the last page. In other words, the *fund manager/computer algorithm* would have the freedom to take into account moral issues while investing. Thus, *he/it* would have the autonomy to exclude companies with various ethical, social or environmental issues. [Also imagine that this could potentially lead to a slightly smaller expected return (e.g., slightly less money in your pension fund).]

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 two-way ANOVA, that was meant to assess: (1) the main effect of fund manager type (first between-subjects factor) on permissibility to exclude stocks, and (2) the interaction effect between fund manager type and effect of stock exclusion (first and second between-subjects factors).

Participants in the "robo" condition were informed that "an algorithm is a sequence of computational steps that transform inputs into outputs, similar to a recipe" (Corme, Leierson, Rivest & Stein, 2010). The full instructions are available on OSF (https://osf.io/gdp42/?view_only=f6eeccb5e9d34f84928f58c601a7c66f).

This procedure has been pre-registered (https://aspredicted.org/blind.php?x=k8wm7t) and approved by the Committee of Ethical Research conducted with participation of humans at the Poznań University of Economics and Business (Resolution 5/2019). Informed consent was obtained from all participants. Participants rated how comfortable they would feel if their pension fund invested in the controversial industries (instead of moral appropriateness; however, we use moral appropriateness in Study 2). This was the only amendment to the pre-registration (we decided to change the dependent variable, but erroneously did not incorporate this ultimate change to the pre-registration document).

Results

Main analysis

In line with the main hypothesis, participants rated the permissibility to exclude stocks as higher if the fund manager was human rather than a computer algorithm ($M_{human} = 2.95$ vs $M_{robo} = 2.64$, $d = -0.25$; $F(1, 462) = 7.83$, $\eta^2 = .02$, $p = .005$). An analysis of variance determined that there was no interaction between the two factors ($F(1, 462) = 0.92$, $\eta^2 = .002$, $p = .34$). There were no significant differences in the permissibility ratings of people that were assigned to conditions where there was no mention of a penalty from excluding controversial stocks or a mention of such a penalty ($M_{penalty\ absent} = 2.79$ vs $M_{penalty\ present} = 2.80$, $d = 0.004$; $F(1, 462) = 0.00$, $\eta^2 < .001$, $p = .96$).

Additional analyses

As an addition to the pre-registered analyses, we performed a number of complementary tests relating to the general credibility of Mechanical Turk participants for carrying out tasks relating to investment in the stock market. In order to see if differences in investment knowledge affect robo-investment aversion, we collected data on participants' subjective investment knowledge, but also constructed a six-item test to verify their objective investment knowledge, drawing on previous research (van Rooij, Lusardi & Alessie, 2011; Agnew & Szykman, 2005). The former was assessed using one statement (*"My investment knowledge is good"*), whereas the latter was assessed using answers to questions such as *"Normally, which asset displays the highest fluctuations over time: savings accounts, bonds or stocks?"*. The mean subjective investment knowledge of participants (on a scale of 1 to 7) was 4.11 ($SD = 1.60$) and was not significantly different in the human and robo condition ($t(464) = 1.50$, $p = .14$). The mean score on the investment test was 4.52 ($SD = 1.24$), and was similarly not different in the human and robo condition ($t(459) = 0.41$, $p = .68$). Interestingly, the subjective and objective investment knowledge were very weakly correlated ($r = .07$, $t(464) = 1.56$, $p = .12$). While low levels of correlation between subjective and objective measures may be surprising, this is not out of line with previous research (Carlson, Vincent, Hardesty & Bearden, 2009). We independently investigated potential differences in robo-investment aversion for individuals with different levels of subjective (declared) and objective (actual) investment knowledge. There was no interaction between the between-individuals factor and neither subjective (b = –0.03, t(462) = –0.38, p = .70) nor objective (b = 0.07, t(462) = 0.83, p = .41) investment knowledge. Additionally, we split participants in two groups according to their level of investment experience (participants were classified as experienced if they rated themselves or scored 5 or higher; 42% (54%) of participants were in the high knowledge group based on subjective (objective) knowledge). Analyses of variance did not support the existence of an interaction between fund manager type with neither subjective (F(1, 462) = 0.61, η^2 =.001, p = .43) nor objective F(1, 462) = 0.96, η^2 = .002, p = .33) investment knowledge groups.

Discussion

The results of the study are consistent with the existence of robo-investment aversion, that is, a preference for human fund managers to make investment decisions, in this case via the exclusion of companies from portfolios. This is consequential, as exclusion serves as one form of socially responsible investment (Trinks, Scholtens, Mulder & Dam, 2018). Although the existence of a penalty resulting from excluding morally controversial stocks – performing negative screening [34] – appeared to have some effect on the size of the aversion, the difference was not statistically significant. Hence, our analysis did not support the notion that people judge fund managers differently in a hypothetical setting if it is made explicit that their decision to exclude companies from portfolios might produce lower rewards (expected returns) for investors (Hong & Kacperczyk, 2009). Note, however, that it is not clear what the effects of socially responsible and irresponsible investing are, that is if the former or the latter produce higher returns or there is the lack of such differences (Blitz & Fabozzi, 2017; Lobe & Walkhäusl, 2016; Revelli & Viviani, 2015).

While some studies suggest that people might show an aversion towards machines making decisions in the moral domain (Bigman & Gray, 2018; Gogoll & Uhl, 2028), this does not imply that there will be an absence of a bias against robo-fund managers when investing in conventional (i.e., non-sin) stocks. Therefore, in order to test whether robo-investment aversion generalizes to conventional stocks we designed Study 2, in which participants had to rate either controversial or non-controversial (conventional) companies.

Study 2: Exclusion of controversial vs non-controversial stocks

In our earlier study, we hypothesized that people have an aversion towards computer algorithms excluding morally controversial companies from portfolios (in relation to human fund managers doing the same). In Study 2, we wanted to replicate Study 1's main finding, and also test whether the aversion towards computer algorithms is stronger in the case of controversial than non-controversial stocks. Moreover, we amended what participants were asked in the first stage of the study – instead of being asked about how *comfortable* they would feel if their pension fund invested in stocks from the shown industries, they were asked how *morally appropriate* such an investment would be.

Method

Participants

We recruited 1,444 participants via Mechanical Turk, in order to obtain 90% power to detect a small effect of similar size to the more conservative estimate in our previous study (f = .09). We excluded 211 participants who did not correctly answer an attention check question, and additionally two participants that have answered none of the three permissibility questions used to compute the permissibility score (for five participants we computed the scores based on the available ratings). We analyzed the data of 1,231 participants, 569 (46%) of whom were female. The mean age of participants was 38.7 years (*SD* = 12.4).

Procedure

The study had a 2 (first between-subjects factor: portfolio choice made by a human fund manager vs portfolio choice made by a computer algorithm) × 2 (second between-subjects factor: controversial stocks vs non-controversial stocks) design. Participants rated how morally appropriate it would be if their pension fund invested in companies from the list of industries that they were shown, using the following text:

> "Would it be morally appropriate if your pension fund invested in companies from the following industries? Please rate on a scale of 1 (not at all) to 7 (completely)."

Participants were then told to imagine the investment scenario. The exact instructions are presented below, with the text shown in italics being varied in the human vs robo condition, and text in the square brackets being presented only to participants that rated controversial stocks:

> "Imagine that a *fund manager/computer algorithm* had the ability to exclude (sell or not invest in) companies from some of the industries that were listed on the last page. [In other words, the fund manager would have the freedom to take into account moral issues while investing. Thus, *he/it* would have the autonomy to exclude companies with various ethical, social or environmental issues.]"

Afterwards, they rated the permissibility of either a human or computer algorithm to exclude companies from some of the industries. We used the same dependent variable as in Study 1. The controversial stocks we used were the seven industries that participants felt the least comfortable with in Study 1, whereas the non-controversial stocks were selected by us via a pilot study.

The procedure has been pre-registered (https://aspredicted.org/blind.php?x=cs8xe3) and approved by the Committee of Ethical Research conducted with participation of humans at the

Poznań University of Economics and Business (Resolution 6/2019). Informed consent was obtained from all participants.

Results

Main analysis

Prior to analyses, we performed a manipulation check, which confirmed that participants found that it would be significantly more appropriate if their pension fund invested in non-controversial stocks than in controversial stocks ($M_{\text{controversial}}$ = 3.25 vs $M_{\text{non-controversial}}$ = 5.55; $t(1217) = 27.0$, $p < .001$, $d = 1.54$).

Once again, consistent with our main hypothesis, participants rated the permissibility of excluding stocks as higher if the fund manager was human rather than an algorithm (M_{human} = 3.45 vs M_{robo} = 2.83; $F(1, 1227) = 109.71$, $\eta^2 = .08$, $p < .001$, $d = –0.58$). An analysis of variance determined that there was an interaction between fund manager type and stock type ($F(1, 1227) = 4.51$, $\eta^2 = .004$, $p = .034$), suggesting a larger difference in the permissibility ratings of human and robo fund managers for non-controversial stocks than for controversial stocks. The effect of stock type was significant, suggesting that people, in general, judge the exclusion of non-controversial stocks to be more permissible than the exclusion of controversial stocks ($M_{\text{controversial}}$ = 2.96 vs $M_{\text{non-controversial}}$ = 3.33; $F(1, 1227) = 39.22$, $\eta^2 = .03$, $p < .001$, $d = 0.34$).

Additional analyses

Similarly to what was done in Study 1, we investigated whether the degree of robo-investment aversion depended on subjective and objective investment experience. Analyses of variance provided limited support for the existence of an interaction between fund manager type and subjective investment experience group ($F(1, 1227) = 3.73$, $\eta^2 = .003$, $p = .054$), suggesting that robo-investment aversion was stronger for participants with lower investment experience.

However, there was no similar interaction with objective investment experience group ($F(1, 1227) = 0.77$, $\eta^2 = .001$, $p = .38$).

Discussion

Whereas robo-investment aversion in Study 1 was equal to $d = -0.25$, in participants of Study 2 that rated controversial stocks this aversion was noticeably greater ($d = -0.47$). However, surprisingly, robo-investment aversion was also present in participants that rated non-controversial stocks ($d = -0.74$). A plausible reason for this increase in effect size is that participants in Study 2 were primed about moral norms when asked to rate how *morally appropriate* it would be if their pension fund excluded investments in companies from the seven industries presented to them. In contrast, in Study 1 they had to rate how *comfortable* they would feel, which might not necessarily prime moral norms. Also, in Study 2 we used a subset of controversial industries with which participants felt the least comfortable with and had a high potential to violate social norms, which further facilitated the moral aspect of allocating funds in companies listed on stock markets. While there are relatively clear-cut moral concerns in the case of controversial industries, this is less so the case when it comes to traditionally non-controversial companies. In order to judge the moral appropriateness of investing in such companies, one might need a higher level of moral competence – one that machines are not expected to possess. Finally, participants were not informed that the exclusion of stocks might lead to worse returns (as were half the participants in Study 1).

Study 3: Exclusion vs heavier investment in controversial stocks

In both previous studies, we hypothesized that people have an aversion towards computer algorithms excluding companies from stock portfolios compared to human fund managers. In our third study, we wanted to test whether apart from an aversion to exclude controversial stocks (as earlier), there is also an aversion towards computer algorithms to invest more heavily in controversial stocks. Therefore, we used a mixed design, where the type of fund manager (human vs robo) was the between-subjects factor (similarly to Study 1 and 2) and decision type (exclusion or heavier investment) was the within-subjects factor. While one might consider exclusion and inclusion opposite sides of the same coin, recent research on the psychological differences between buying and selling decisions (Akepanidtaworn, Di Mascio, Imas & Schmidt, 2019; Barber & Odean, 2013) hints at the potential existence of asymmetries between how investors approach positive- and negative screening. This might especially be the case when it comes to morally controversial decisions (Woollard, 2015).

Method

Participants

We recruited 722 participants via Mechanical Turk. Although this was a significantly lower sample size than the one used in the previous study, one of the factors was within-subjects, which allowed us to retain 90% power from the former studies much more efficiently (with less participants). This change was also intended to test of the robustness of our effect to changes in the design, similar to other studies on machines (Bigman & Gray, 2018). We excluded 39 participants who did not correctly answer an attention check question. We analyzed the data of 683 participants, of which 387 (57%) were female. The mean age of participants was 39.3 years (*SD* = 12.0).

Procedure

Participants assessed the moral appropriateness of their mutual fund holding 14 controversial industries (the same industries that we used in Study 1). We changed the investment fund type to account for the eventuality that this might be more believable to some participants (and, at the same time, to test the robustness of our earlier findings). After rating the industries, participants assessed the permissibility of their mutual fund excluding and investing more heavily in companies from some of the controversial industries (the order of presentation was counter-balanced). The design and wording of this study was similar to Studies 1 and 2, but as part of the newly added within-factor, participants were asked not just about the permissibility of excluding certain stocks, but also about investing more heavily in certain stocks.

We used the same dependent variable as in Studies 1 and 2. We performed a mixed ANOVA, that was meant to assess: (1) the main effect of fund manager type (between-subjects factor) on permissibility, (2) the main effect of decision type (the within-subjects factor) on permissibility, and (3) the interaction effect between fund manager type and decision type.

This procedure has been pre-registered (https://aspredicted.org/blind.php?x=d5fg7a) and approved by the Committee of Ethical Research conducted with participation of humans at the Poznań University of Economics and Business (Resolution 6/2019). Informed consent was obtained from all participants.

Results

Main analysis

An analysis of variance showed that, in general, participants rated computer algorithms to be less permitted to make investment decisions than humans (M_{human} = 3.48 vs M_{robo} = 2.58; $F(1, 681) = 130.38$, $\eta^2 = .14$, $p < .001$, $d = -0.81$). Robo-investment aversion (see Fig. 1C) was present both when fund managers had the autonomy to exclude controversial stocks (M_{human} = 3.46 vs M_{robo} = 2.64, $t(671) = 9.72$, $p < .001$, $d = -0.75$), and when they had the autonomy to invest more heavily in such stocks (M_{human} = 3.50 vs M_{robo} = 2.53, $t(665) = 11.50$, $p < .001$, $d = -0.88$). The type of operation (exclusion vs heavier investment) did not have a statistically significant effect on permissibility ratings ($F(1, 682) = 0.97$, $p = .33$). However, the interaction between the between-subjects factor (human vs robo fund manager) and within-subjects factor (autonomy to exclude stocks vs autonomy to invest more heavily in stocks) was statistically significant ($F(1, 681) = 4.77$, $\eta^2 = .001$, $p = .029$), consistent with robo-investment aversion being stronger when fund managers had the autonomy to invest more heavily in controversial stocks than when they had the autonomy to exclude them.

Additional analyses

In Study 3 we collected data on whether participants held stocks directly and via pension- or mutual funds, in order to see if robo-investment aversion manifests itself in experienced participants. Robo-investment aversion was present both in participants that held stocks directly (d_{stocks} = –0.73) and those that didn't ($d_{\text{no stocks}}$ = –0.85). Similarly, robo-investment aversion appeared in participants with or without mutual- ($d_{\text{mutual funds}}$ = –0.77 vs $d_{\text{no mutual funds}}$ = –0.83) or pension- ($d_{\text{pension funds}}$ = –0.73 vs $d_{\text{no pension funds}}$ = –0.86) fund holdings. In none of the cases was there an interaction between fund manager type and the holding of stocks or funds (ps > .43). Separate analyses of variance showed that there was no interaction between fund manager type and subjective ($F(1, 679) = 0.86$, $\eta^2 = .001$, $p = .35$) or objective investment knowledge group ($F(1, 679) = 0.05$, $\eta^2 < .001$, $p = .82$).

Discussion

In Study 3 we investigated whether robo-investment aversion generalizes to investment decisions, in which fund managers decide to allocate *more* in companies from controversial industries. Consistent with Studies 1 and 2, participants exhibited an aversion towards algorithms making decisions about increased investments in controversial stocks. In fact, the aversion was higher when the fund managers had heavier-investment autonomy instead of exclusion autonomy. Overall, this supports the robustness of the investigated effect, suggesting that algorithmic investment funds might be considered a less legitimate way of investing in companies from controversial industries. This suggests that computerized funds might *also* have lower legitimacy to perform socially responsible investing (e.g., excluding or investing more heavily in "saint" stocks; [22]). This prediction was not tested by us but might be of interest to other researchers.

We summarize permissibility ratings in Studies 1-3 in Table 1.

Table 1. Permissibility to make investment decisions

Study	Second factor	Human	Robo	*d*
Study 1	Penalty absent	3.01 (1.22)	2.59 (1.16)	–0.35 [–0.62, –0.08]
	Penalty present	2.90 (1.24)	2.69 (1.19)	–0.17 [–0.42, 0.08]
	Overall	**2.95 (1.23)**	**2.64 (1.17)**	**–0.25 [–0.44, –0.07]**
Study 2	Controversial	3.22 (1.07)	2.71 (1.11)	–0.47 [–0.62, –0.31]
	Non-controversial	3.73 (1.01)	2.96 (1.05)	–0.74 [–0.91, –0.57]
	Overall	**3.45 (1.07)**	**2.83 (1.09)**	**–0.58 [–0.69, –0.47]**
Study 3	Exclusion	3.46 (1.09)	2.64 (1.14)	–0.75 [–0.90, –0.59]
	Heavier investment	3.50 (1.05)	2.53 (1.14)	–0.88 [–1.04, –0.72]
	Overall	**3.48 (1.07)**	**2.58 (1.14)**	**–0.81 [–0.92, –0.70]**

Notes: All studies reported in this table have a 2 × 2 design, with human vs robo investment manager being the first factor, and the second factor reported in the second column. *SD*s are reported in parentheses. *d* denotes Cohen's *d;* 95% confidence intervals are reported in the square brackets.

Studies 1-3 relied solely on hypothetical choices. In order to test actual decisions – as it is often performed in the literature on algorithm aversion (Dietvors, Simmons & Massey, 2015; Logg, Minson & Moore, 2018) – in Studies 4 and 5 participants had to decide whether they would prefer humans or an algorithm as advisors in an investment task that mimicked investment reality.

Study 4: Incentivized choice between human and robo investment managers

In Study 4 we pitted humans against algorithms as investment advisors. In this study we tested in an incentive-compatible manner whether people exhibit algorithm aversion when asked to decide whether they would want actual humans or an actual algorithm to serve as advisors in a simplified investment task. Additionally, the design of the study allows us to once again directly compare whether participants' choices differ for controversial and non-controversial stocks.

Study 4 is conceptually similar to studies on algorithm aversion by Dietvorst and colleague (Dietvorst, Simmons & Massey, 2016). The design has also been guided by Önkal and colleagues (Önkal, Goodwin, Thomson, Gönül & Pollock, 2009).

Methods

Participants

We recruited 838 participants via Mechanical Turk, allowing us to obtain 80% power detect a 10% percentage point difference in choice (between 55 and 65%, set arbitrarily). We excluded 133 participants who did not correctly answer an attention check question. We analyzed the data of 705 participants, of which 288 (41%) were female. The mean age of participants was 39.4 years (SD = 12.3).

Procedure

Participants saw a series of historical price paths (Borsboom & Zeisberger, 2020; Grosshans & Zeisberger, 2018) and were asked to decide whether they would prefer their investments in these stocks to be managed by a group of humans with investment knowledge, or by an algorithm (using the *nnetar* function from the *forecast* package in *R* (Hyndman & Khandakar, 2008), which trains a neural network on the historical price paths to forecast future prices). The knowledgeable humans were the 11 most effective performers in a pilot study (N = 101), carried out earlier on Mechanical Turk: we assessed their investment knowledge using a very brief test (with a maximum score of 3; see Experimental Materials on *OSF* for details), and then told them to predict the price of the stocks two months after the shown 1-year period (252 trading days). We used actual stock prices, but not from the most recent year, to make it harder to identify the price paths from memory. Participants were told that if they chose the advisor that was more accurate (on average) in predicting the stock prices, they would get a bonus, doubling the compensation for participating in the task.

The experiment had a 2-cell design, where participants were shown the price paths of controversial or non-controversial stocks. This procedure has been pre-registered (https://aspredicted.org/blind.php?x=ak6th4).

Results

A similar proportion of participants chose human (48%) over algorithm (52%) advice, and thus – in contrast to the earlier studies – there was no evidence of an aversion towards algorithms in the entire sample ($\chi^2(1) = 1.19$, $p = .27$; $d = 0.08$). However, the choice of advisor differed in controversial stocks (where 47.8% chose the algorithm; $d = -0.09$) and non-controversial stocks (where 56.5% chose the algorithm; $d = 0.26$), showing a greater reliance on human advice for controversial stocks ($\chi^2(1) = 5.37$, $p = .021$; $d = 0.18$).

Additional analyses

The effect was of similar magnitude for participants that held stocks directly ($d_{\text{stocks}} = 0.10$) and those that didn't ($d_{\text{no stocks}} = 0.07$), but the preference for machines was not statistically significant at conventional levels. The same pattern appeared in participants with or without pension fund holdings ($d_{\text{pension funds}} = 0.16$ vs $d_{\text{no pension funds}} = 0.05$). However, algorithm (machine) appreciation was significant in participants that had mutual fund holdings, in contrast to participants that did not have such holdings ($d_{\text{mutual funds}} = 0.26$ vs $d_{\text{no mutual funds}} = -0.11$).

Discussion

In contrast to the results of the earlier studies, in Study 4 we found no general aversion towards algorithms, but a difference in the utilization of human advice depending on whether advice concerned morally controversial investments or not. This is consistent with the notion that people might be more reluctant to use robo-investment funds for morally controversial investments. A potential explanation for why we could not produce general aversion towards algorithms in Study 4 might be due to differences in the framing and design of the experimental task. While Studies 1-3 concern the morality of investment decisions (with a potential impact on returns), Study 4 concerns performance in a forecasting task between human and algorithmic agents.

A plausible reason for the lack of algorithm aversion is that perhaps a considerable number of people consider algorithms to have superior performance over (knowledgeable) humans, at least in this particular task. While we do not believe that such beliefs are responsible for driving the observed effects, we nevertheless decided to conduct an additional study to account for this possibility. To address this potential alternative explanation, in Study 5 we equalize the expected outcomes of humans and the artificial neural network algorithm, allowing us to investigate whether people will prefer humans over algorithms after making it explicit that there are no differences – on average – between them.

Study 5: Choice between human and robo investment managers after outcomes are equalized

In Study 5 we utilized the same design that we used in Study 4, with a key difference: participants were truthfully informed that the (knowledgeable) humans – that scored a maximum (3 out of 3) score on the investment knowledge test in the same pilot study that we describe in Study 4 – performed similarly to the (artificial neural network) algorithm. However, participants were also informed that humans and the algorithm might differ in their accuracy to predict future prices of certain industries or particular stocks. This was intended to set the conditions for the final test of the hypothesis that the reliance on humans as investment managers will be stronger for controversial than non-controversial stocks, assuming that the industry itself is a strong enough cue to guide participants' moral compasses.

Methods

Participants

We recruited 882 participants via Mechanical Turk. Similarly to Study 4, we wanted to detect a 10 percentage point difference, but collected more data to account for a higher (15%) expected exclusion rate of inattentive participants. We excluded 139 participants who did not correctly answer an attention check question. We analyzed the data of 743 participants, of which 335 (45%) were female. The mean age of participants was 40.3 years (*SD* = 12.5).

Procedure

Participants were informed that the performance of humans and computers is similar (for all stocks), but that humans and computers can have a different level of accuracy for some stocks (or industries). This experiment had a 2-cell design: participants were either allocated to the controversial stocks condition or the non-controversial stocks condition, in which they saw the historical price paths of five (sin or conventional) stocks. We used the past performance of stocks that represented each of these industries. The stocks were selected so that the past performance of the portfolios of sin and conventional stocks was similar.

This procedure has been pre-registered (https://aspredicted.org/blind.php?x=nt9n9j).

Results

Participants showed a preference for human advice, choosing them in 57.3% of cases, which was significantly different from 50% ($\chi^2(1) = 16.0$, $p < .001$). This is consistent with robo-investment aversion ($d = -0.30$).

Participants showed an aversion of similar magnitude for controversial stocks (human advice was preferred by 57.0% of participants) and non-controversial stocks (human advice was preferred by 57.6% of participants; $\chi^2(1) = 0.03$, $p = .87$).

Additional analyses

Pooled analysis

Considering the very similar structure of Studies 4 and 5, we also performed an analysis on the pooled dataset. Overall, 47.0% of participants in these studies preferred the algorithm over humans, which is consistent with robo-investment aversion ($\chi^2(1) = 4.42$, $p = .036$). This aversion was not different across controversial and non-controversial stocks ($\chi^2(1) = 1.95$, $p = .16$).

Effect of differences in investment experience

In Study 5, robo-investment aversion was present both in participants that did not hold stocks or mutual funds ($d_{\text{no stocks}} = -0.45$; $d_{\text{no mutual funds}} = -0.54$), but not in those that did ($d_{\text{stocks}} = -0.15$; $d_{\text{mutual funds}} = -0.10$). However, robo-investment aversion was of similar magnitude for both participants that did or did not hold pension funds ($d_{\text{pension funds}} = -0.29$ vs $d_{\text{no pension funds}} = -0.35$).

Text analysis

In addition to testing our hypotheses, we wanted to further investigate the beliefs motivating our participants to choose humans or an algorithm. As part of a pre-registered analysis, we gave our participants the option to explain their decision, and incentivized them by offering a lottery bonus of \$20. Anyone submitting an original and intelligible contribution participated in the lottery, regardless of their stated reasons for selecting one option or the other. Six-hundred and twenty-two participants opted to participate, yielding a combined output of 23,721 words. Excluding English stop words (Benoit, Muhr & Watanabe, 2020) and words which were contained in the original question formulation (*"Please tell us why you chose the sophisticated (neural network) algorithm / Please tell us why you chose the knowledgeable humans"*) resulted in 10,741 words.

We find that individuals choosing the algorithm use the terms "bias" and "emotions" eleven- and three times more often than the human-choice group, respectively. This indicates that individuals who chose the algorithm might have done so in part due to a belief that algorithms are less influenced by factors extraneous to the investment decision. In a subsequent bigram analysis for this subgroup, "human" is the word most often associated with "emotion" or "emotions" ($N = 19$), supporting this interpretation. Similarly, the word most often associated with "bias" or "biases" is "human" ($N = 8$). On the other hand, participants who preferred to follow human investment advice made four times as many references to "industries", presumably as a reference to the type of industries presented in the respective conditions. Further terms more associated with choosing humans were "knowledge", "experience" and "factor" – all most often preceded by the term "human" in a bigram analysis. What this analysis reveals is that even though participants were explicitly told to expect similar prediction performance from both human- and algorithmic decision-makers, they still express diverse beliefs when it comes to the process leading to these similar outcomes.

Discussion

In contrast to Study 4, in Study 5 the results point to the existence of robo-investment aversion, but an aversion that is not different across controversial and non-controversial stocks. A pooled analysis for Study 4 and 5 supports the existence of robo-investment aversion, but does not suggest the existence of differences in this aversion between controversial and non-controversial stocks. Note, however, that Studies 1-3 and Studies 4-5 are different in their nature. In the former, participants judged the permissibility of either a human or robo to "call the shots" while excluding (or investing more heavily) in companies. In contrast, in Studies 4-5, participants are asked whether they would prefer (knowledgeable) humans or a (neural network) algorithm for predictions of future stock prices. If people show a preference for a human over a robo investment fund manager (a long-term, complex task, requiring a series of judgments of various nature), it is not necessarily the case that this preference will manifest itself in a much simple task of predicting stock prices based on past stock prices alone.

General discussion

The dawn of algorithm-generated ("robo") advisors warrants us to ask how people judge the perspective of machines making investments for them ("robo-investments"), some of them in morally controversial stocks (D'acunto, Prabhala & Rossi, 2019; Hong &Kacperczyk, 2009). Previous research usually documents an aversion towards algorithm use (Dietvors, Simmons & Massey, 2015; 2016), yet under some circumstances the opposite is true (Logg, Minson & Moore, 2019), This suggests that algorithm aversion should not be universally expected. In our case, we test whether it extends to investments, focusing mainly on investments in morally controversial stocks. In five experiments (N = 3,828) – summarized in Table 2 – we document a considerable robo-investment aversion (d = –0.39 [–0.45, –0.32] in internal meta-analysis (Goh, Hall & Rosenthal, 2016).

Comparisons of the strength of the aversion for controversial and non-controversial stocks – defined based on the industry in which they operate – suggest that it might even be stronger for the latter (Study 2), or that the two are not different (Study 5). However, this does not necessarily rule out the role moral concerns play when choosing between a human and robo investment manager or advisor. Note that this finding might simply reflect the fact, that it is more difficult to make judgments in the case of non-controversial than controversial industries. Robos can be easily programmed to exclude companies from some (or all) sin industries if this is the wish of the investor, who will surely not question the robos' ability to follow such a simple rule. Put differently, some participants might perceive the industry in which the company operates as a sufficient "moral cue", which makes further investigation of a controversial company by a competent moral examiner (a human; Bigman & Gray, 2018) redundant. For companies from non-controversial industries, however, choosing the "sinners" is a more difficult task. An investor that cares about moral issues will want the investment manager (advisor) to have the additional ability to identify the "bad apples" in industries that are not known to be controversial, at least relative to other industries. For example, it is probably fair to say that the car industry produces – relative to the tobacco industry – less harmful (more morally acceptable) goods. In such a case, people might prefer their money to be managed by a human fund manager, who will probably be perceived to be more competent to handle less conventional cases. A counter-argument of increasing prominence is that although moral issues seem to be more "fuzzy" (more difficult to quantify) in general, it is now easier for algorithms to monitor the "moral risk" of potential investments, taking into account scores for sustainability (Filbeck, Filbeck & Zhao, 2019) or corporate governance, and synthesizing it with the emotional valence of news or social media content (Capelle-Blancard & Petit, 2019; Chen, Liao & Hsieh, 2019; Tetlock, 2007). A number of startups have recently started offering such services (Fantato, 2020; Beioley, 2018). Yet, the availability of "hard" data on responsible business practices in itself is insufficient to solve the problem. There are multiple agencies that rate what environmental, social and governance policies companies have in place (or lack). The ratings provided by competing agencies are diverging (Berg, Kölbel & Rigobon, 2020). Choosing the "right" agency is a task where, once again, people will probably gravitate towards human decision-makers (and not robos).

Understanding the personal and demographic variables associated with robo-investment aversion is a further step towards understanding the role such differences might play in technology adoption in the domain of finance. In a secondary analysis we find that robo-investment aversion is present among males and females (Figure 1, Panel A), in people with high or low investment knowledge (both objectively- and subjectively-measured; Panels B-C), and regardless of whether a person holds stocks in their portfolios (either directly or via pension or mutual funds; Panels D-F). Interestingly, the (exploratory) analysis revealed that robo-investment aversion was stronger for female participants, consistent with a multitude of studies documenting gender differences in investment behavior (e.g. Lusardi & Mitchell, 2008; Barber & Odean, 2001; Niszczota & Białek, 2020). The aversion was also stronger in participants with lower investment knowledge, and in participants that do not hold stocks directly or via mutual funds (ps $< .05$). This is perhaps unfortunate, as one of the benefits of robo-advice is that is lowers the barriers-of-entry into financial investments, making it easier to enjoy the benefits of stock market participation.

A limitation of our paper is that we have incentivized the choice of investment advisor in only one of our studies (Study 4; in Study 5 the incentives were used only to promote thoughtful responses in an exploratory study of the motivations behind participants' choices), and it happens to be the only study where there was no bias against algorithms. However, we believe this is most likely due to the belief of some participants that algorithms produce superior outcomes, in accordance with reality (Dietvorst, Simmons & Massey, 2015). Bias against algorithms re-emerges once participants are informed that humans and algorithms (in our case: an artificial neural network) produce similar outcomes, (presumably) taking away the algorithms' perceived edge (Castelo, Bos & Lehmann, 2019).

Overall, our study suggests the existence of a barrier for the adoption of innovative financial technologies (FinTech; e.g., Goldstein, Jiang & Karolyi, 2019; however, see Gogoll & Uhl, 2018 for an alternative account). While investors today enjoy an unprecedented ease of access to financial products and services, they might still be reluctant to use them. Our findings suggest the future proliferation of hybrid systems (Kaya, Schildbach & Schneider, 2017), considering mounting pressures to reduce funds' costs of operation. The need for supervision by humans can be seen as a pitfall for the robo-advisory services (D'Acunto, Prabhala & Rossi, 2019). A potential problem that might arise in the future would be the overreporting of human fund manager presence, to lower costs. However, finding out that investment funds have been managed by machines might undermine trust, given the evidence of the emergence of a bias against robots in certain scenarios (Ishowo-Oloko, Bonnefon, Soroye, Crandall, Rahwan, Rahwan, 2019).

To summarize, we find that people display an aversion to investing with algorithms. Our findings can be interpreted as follows: people perceive algorithms as being less effective than humans at tasks which require making subjective judgments (Castelo, Bos & Lehmann, 2019). Future research could investigate the drivers of algorithm aversion in the domain of investments, including a more systematic exploration of the role of moral considerations and situational factors, and a further examination of the role of individual differences.

Table 2. Robo-investment aversion across studies

Study	*N*	Design	Dependent variable	Effect size < 0 indicates robo-investment aversion		
				Controversial stocks	**Non-controversial stocks**	**Pooled**
1	466	2 (between-subjects: human vs robo) × 2 (between-subjects: penalty vs no penalty)	Permissibility to exclude (Studies 1-3) or invest more heavily (Study 3) in stocks *1 = strongly disagree* *5 = strongly agree*	–0.25 [–0.44, –0.07]	–	–0.25 [–0.44, –0.07]
2	1,231	2 (between-subjects: human vs robo) × 2 (between-subjects: controversial vs non-controversial)		–0.47 [–0.62, –0.31]	–0.74 [–0.91, –0.57]	–0.58 [–0.69, –0.47]
3	683	2 (between-subjects: human vs robo) × 2 (within-subjects: exclusion vs inclusion)		–0.81 [–0.92, –0.70]	–	–0.81 [–0.92, –0.70]
4	705	2-cell (between-subjects: controversial vs non-controversial)	Choice of investment manager *=1 robo* *=0 human*	–0.09 [–0.30, 0.12]	0.26 [0.05, 0.47]	0.08 [–0.07, 0.23]
5	743	2-cell (between-subjects: controversial vs non-controversial)		–0.31 [–0.51, –0.11]	–0.28 [–0.49, –0.07]	–0.30 [–0.44, –0.15]
1-5	**3,828**					**Mean effect** fixed effects model **–0.39 [–0.45, –0.32]** **Mean effect** random effects model **–0.36 [–0.64, –0.08]**

Notes: Mean effect sizes are Cohen's *d*s, with 95% confidence intervals reported in square brackets. Computed using the *R* package *metaviz*, based on the *metafor* package (Kossmeier, Tran & Voracek, 2020; Viechtbauer, 2010). Random effects are computed using the *REML* method.

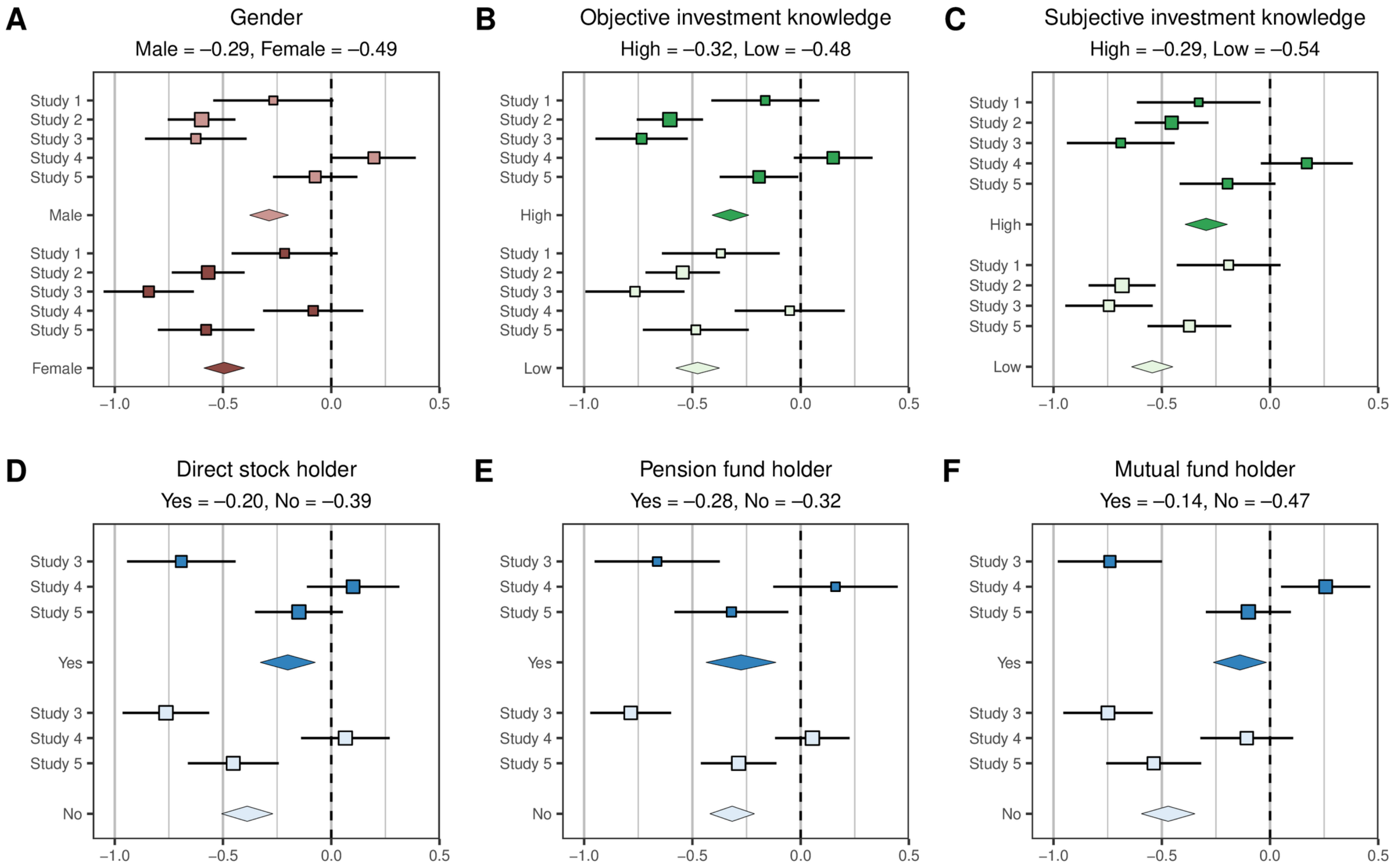

A
Gender
Male = −0.29, Female = −0.49
B
Objective investment knowledge
High = −0.32, Low = −0.48
C
Subjective investment knowledge
High = −0.29, Low = −0.54
D
Direct stock holder
Yes = −0.20, No = −0.39
E
Pension fund holder
Yes = −0.28, No = −0.32
F
Mutual fund holder
Yes = −0.14, No = −0.47
Study 1
Study 2
Study 3
Study 4
Study 5
Male
Female
High
Low
Yes
No
−1.0
−0.5
0.0
0.5

Figure 1. Robo-investment aversion across subsamples

Notes: Effects sizes are presented as Cohen's *d*s. Mean effect sizes are fixed effects, computed using the *R* package *metaviz* (Kossmeier, Tran & Voracek, 2020; Viechtbauer, 2010). Error bars correspond to 95% confidence intervals. All mean subgroup effects are significant at the 5% level.

References

Abis S. Man vs. machine: Quantitative and discretionary equity management. Unpublished working paper Columbia Business School. 2017.

Agnew JR, Szykman LR. Asset Allocation and Information Overload: The Influence of Information Display, Asset Choice, and Investor Experience. Journal of Behavioral Finance. 2005;6: 57–70. doi:10.1207/s15427579jpfm0602_2

Akepanidtaworn K, Di Mascio R, Imas A, Schmidt L. Selling Fast and Buying Slow: Heuristics and Trading Performance of Institutional Investors. Rochester, NY: Social Science Research Network; 2019 Sep. Report No.: ID 3301277. Available: https://papers.ssrn.com/abstract=3301277

Awad E, Dsouza S, Kim R, Schulz J, Henrich J, Shariff A, et al. The Moral Machine experiment. Nature. 2018;563: 59–64. doi:10.1038/s41586-018-0637-6

Barber BM, Odean T. Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment. Quarterly Journal of Economics. 2001;116: 261–292.

Barber BM, Odean T. The Behavior of Individual Investors. In: Constantinides GM, Harris M, Stulz RM, editors. Handbook of the Economics of Finance. Elsevier; 2013. pp. 1533–1570. doi:10.1016/B978-0-44-459406-8.00022-6

Beioley K. Robo advisers short-circuit over ethical investing. In: Financial Times [Internet]. 13 Sep 2018 [cited 19 Jan 2020]. Available: https://www.ft.com/content/1dabbce8-b5d8-11e8-b3ef-799c8613f4a1

Benoit K, Muhr D, Watanabe K. stopwords: Multilingual Stopword Lists. 2020. Available: https://CRAN.R-project.org/package=stopwords

Berinsky AJ, Huber GA, Lenz GS. Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. Political Analysis. 2012;20: 351–368 doi:10.1093/pan/mpr057

Berg F, Kölbel JF, Rigobon R. Aggregate Confusion: The Divergence of ESG Ratings. Rochester, NY: Social Science Research Network; 2020 May. Report No.: ID 3438533. doi:10.2139/ssrn.3438533

Bigman YE, Gray K. People are averse to machines making moral decisions. Cognition. 2018;181: 21–34. doi:10.1016/j.cognition.2018.08.003

Blitz D, Fabozzi FJ. Sin Stocks Revisited: Resolving the Sin Stock Anomaly. J Portf Manag. 2017;44: 105–111. doi:10.3905/jpm.2017.44.1.105

Bonnefon J-F, Shariff A, Rahwan I. The social dilemma of autonomous vehicles. Science. 2016;352: 1573–1576. doi:10.1126/science.aaf2654

Borsboom C, Zeisberger S. What makes an investment risky? An analysis of price path characteristics. Journal of Economic Behavior & Organization. 2020;169: 92–125. doi:10.1016/j.jebo.2019.11.002

Capelle-Blancard G, Petit A. Every Little Helps? ESG News and Stock Market Reaction. J Bus Ethics. 2019;157: 543–565. doi:10.1007/s10551-017-3667-3

Castelo N, Bos MW, Lehmann DR. Task-Dependent Algorithm Aversion. Journal of Marketing Research. 2019;56: 809–825. doi:10.1177/0022243719851788

Carlson JP, Vincent LH, Hardesty DM, Bearden WO. Objective and subjective knowledge relationships: A quantitative analysis of consumer research findings. Journal of Consumer Research. 2009;35: 864–876. doi:10.1086/593688

Chen M-Y, Liao C-H, Hsieh R-P. Modeling public mood and emotion: Stock market trend prediction with anticipatory computing approach. Computers in Human Behavior. 2019 [cited 17 Jun 2019]. doi:10.1016/j.chb.2019.03.021

Choi A. Morgan Stanley Perspectives on Sustainable Investing: Acceleration and Integration. Journal of Applied Corporate Finance. 2016;28: 62–65. doi:10.1111/jacf.12176

Cormen TH, Leiserson CE, Rivest RL, Stein C. Introduction to Algorithms. MIT Press; 2009.

D'Acunto F, Prabhala N, Rossi AG. The Promises and Pitfalls of Robo-Advising. Rev Financ Stud. 2019;32: 1983–2020. doi:10.1093/rfs/hhz014

Dietvorst BJ, Simmons JP, Massey C. Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General. 2015;144: 114–126. doi:10.1037/xge0000033

Dietvorst BJ, Simmons JP, Massey C. Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. Management Science. 2016;64: 1155–1170. doi:10.1287/mnsc.2016.2643

Durand RB, Koh S, Limkriangkrai M. Saints versus Sinners. Does morality matter? Journal of International Financial Markets, Institutions and Money. 2013;24: 166–183. doi:10.1016/j.intfin.2012.12.002

Fantato D. Robo-adviser launches ethical portfolios. Financial Adviser. 15 Aug 2018. Available: https://www.ftadviser.com/your-industry/2018/08/15/robo-adviser-launches-ethical-portfolios/. Accessed 21 Jan 2020.

Filbeck A, Filbeck G, Zhao X. Performance Assessment of Firms Following Sustainalytics ESG Principles. The Journal of Investing. 2019;28: 7–20. doi:10.3905/joi.2019.28.2.007

Gerhard P, Hoffmann AOI, Post T. Past performance framing and investors' belief updating: Is seeing long-term returns always associated with smaller belief updates? Journal of Behavioral and Experimental Finance. 2017;15: 38–51. doi:10.1016/j.jbef.2017.07.007

Goh JX, Hall JA, Rosenthal R. Mini Meta-Analysis of Your Own Studies: Some Arguments on Why and a Primer on How. Social and Personality Psychology Compass. 2016;10: 535–549. doi:10.1111/spc3.12267

Gogoll J, Uhl M. Rage against the machine: Automation in the moral domain. Journal of Behavioral and Experimental Economics. 2018;74: 97–103. doi:10.1016/j.socec.2018.04.003

Goldstein I, Jiang W, Karolyi GA. To FinTech and Beyond. Rev Financ Stud. 2019;32: 1647–1661. doi:10.1093/rfs/hhz025

Grosshans D, Zeisberger S. All's well that ends well? On the importance of how returns are achieved. Journal of Banking & Finance. 2018;87: 397–410. doi:10.1016/j.jbankfin.2017.09.021

Harvey CR, Rattray S, Sinclair A, Hemert OV. Man vs. Machine: Comparing Discretionary and Systematic Hedge Fund Performance. The Journal of Portfolio Management. 2017;43: 55–69. doi:10.3905/jpm.2017.43.4.055

Hong H, Kacperczyk M. The price of sin: The effects of social norms on markets. J Financ Econ. 2009;93: 15–36. doi:10.1016/j.jfineco.2008.09.001

Hyndman RJ, Khandakar Y. Automatic Time Series Forecasting: The forecast Package for R. Journal of Statistical Software. 2008;27: 1–22. doi:10.18637/jss.v027.i03

Ishowo-Oloko F, Bonnefon J-F, Soroye Z, Crandall J, Rahwan I, Rahwan T. Behavioural evidence for a transparency–efficiency tradeoff in human–machine cooperation. Nat Mach Intell. 2019;1: 517–521. doi:10.1038/s42256-019-0113-5

Kaya O, Schildbach J, AG DB, Schneider S. Robo-advice–a true innovation in asset management. Deutsche Bank Research, August, available at https://www dbresearch com/PROD/DBR_INTERNET_EN-PROD/PROD0000000000449010/Robo-advice_-_a_true_innovation_in_asset_managemen pdf. 2017.

König-Kersting C, Pollmann M, Potters J, Trautmann ST. Good decision vs. good results: Outcome bias in the evaluation of financial agents. Working paper, Tilburg University; 2017.

Kossmeier M, Tran US, Voracek M. metaviz: Forest Plots, Funnel Plots, and Visual Funnel Plot Inference for Meta-Analysis. 2020. Available: https://CRAN.R-project.org/package=metaviz

Lobe S, Walkshäusl C. Vice versus virtue investing around the world. Rev Manag Sci. 2016;10: 303–344. doi:10.1007/s11846-014-0147-3

Logg JM, Minson JA, Moore DA. Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes. 2019;151: 90–103. doi:10.1016/j.obhdp.2018.12.005

Longoni C, Bonezzi A, Morewedge CK. Resistance to Medical Artificial Intelligence. J Consum Res. 2019;46: 629–650. doi:10.1093/jcr/ucz013 153–163. doi:10.1016/j.jesp.2017.01.006

Lusardi A, Mitchell OS. Planning and Financial Literacy: How Do Women Fare? The American Economic Review. 2008;98: 413–417.

Moulliet D, Majonek A, Stolzenbach J, Völker T. The Expansion of Robo-Advisory in Wealth Management. Deloitte. 2016. Available: https://www2.deloitte.com/content/dam/Deloitte/de/Documents/financial-services/Deloitte-Robo-safe.pdf

Niszczota P, Białek M. Women oppose sin stocks more than men do. 2020 [cited 19 Mar 2020]. doi:10.31234/osf.io/7a4cq

Önkal D, Goodwin P, Thomson M, Gönül S, Pollock A. The relative influence of advice from human experts and statistical methods on forecast adjustments. Journal of Behavioral Decision Making. 2009;22: 390–409. doi:10.1002/bdm.637

Paolacci G, Chandler J, Ipeirotis PG. Running experiments on Amazon Mechanical Turk. Judgment and Decision Making. 2010;5: 411–419.

Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. Journal of Experimental Social Psychology. 2017;70: 28.

Pettifor A. BlackRock gets praise for coal divestment. What it really needs is regulation | Ann Pettifor. The Guardian. 16 Jan 2020. Available: https://www.theguardian.com/commentisfree/2020/jan/16/blackrock-coal-divestment-regulation-fund-manager. Accessed 18 Jan 2020.

Revelli C, Viviani J-L. Financial performance of socially responsible investing (SRI): what have we learned? A meta-analysis. Business Ethics: A European Review. 2015;24: 158–185. doi:10.1111/beer.12076

Sutter M, Huber J, Kirchler M, Stefan M, Walzl M. Where to look for the morals in markets? Exp Econ. 2019 [cited 3 Oct 2019]. doi:10.1007/s10683-019-09608-z

Tetlock PC. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. The Journal of Finance. 2007;62: 1139–1168. doi:10.1111/j.1540-6261.2007.01232.x

The Economist. The rise of the financial machines. The Economist. 2019. Available: https://www.economist.com/leaders/2019/10/03/the-rise-of-the-financial-machines. Accessed 24 Jan 2020.

Thomas KA, Clifford S. Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. Computers in Human Behavior. 2017;77: 184–197. doi:10.1016/j.chb.2017.08.038

Trinks A, Scholtens B, Mulder M, Dam L. Fossil Fuel Divestment and Portfolio Performance. Ecol Econ. 2018;146: 740–748. doi:10.1016/j.ecolecon.2017.11.036

Trinks PJ, Scholtens B. The Opportunity Cost of Negative Screening in Socially Responsible Investing. J Bus Ethics. 2017;140: 193–208. doi:10.1007/s10551-015-2684-3

van Duuren E, Plantinga A, Scholtens B. ESG Integration and the Investment Management Process: Fundamental Investing Reinvented. J Bus Ethics. 2016;138: 525–533. doi:10.1007/s10551-015-2610-8

van Rooij M, Lusardi A, Alessie R. Financial literacy and stock market participation. J Financ Econ. 2011;101: 449–472. doi:10.1016/j.jfineco.2011.03.006

Woollard F. Doing and allowing harm. Oxford University Press, USA; 2015.

Viechtbauer W. Conducting meta-analyses in R with the metafor package. Journal of Statistical Software. 2010;36: 1–48. doi:10.18637/jss.v036.i03

Study 4: A Driver is Missing – Comfort with Varying Levels of Human Supervision in Self-Driving Cars and Associated Factors

Dániel Kaszás[1], Adam Charles Roberts[1,2]

[1]Chair of Cognitive Science, ETH Zürich

[2]Singapore ETH Center, Future Resilient Systems

Abstract

While numerous studies have investigated attitudes towards self-driving cars in general, less research attention has been focused on individuals' attitudes towards the presence (or absence) of third-party human supervision, and its potential correlates. In the present study we analyze data from a large-scale European survey, and find considerable heterogeneity in both individual attitudes, as well as in country-level attitudes in our descriptive analysis. We find a trend of decreasing comfort as external human supervision is reduced, although this effect is more pronounced in some countries than others. We then investigate potential drivers of self-reported comfort with varying levels of external human supervision in a regression framework. Gender differences get stronger with decreasing supervision, suggesting a possible resolution to conflicting evidence in previous studies. Following this, we fit an ordinal random forest model to derive variable importance metrics, which enable us to compare the changing role predictor variables might play in shaping self-reported comfort, depending on varying levels of third-party supervision. Data privacy is highlighted as an important variable, regardless of level of supervision. While our findings are largely in line with previously published literature, we also uncover a number of novel associations, providing guidance for future policy-making and research efforts.

Keywords: "human supervision; third-party supervision; technology acceptance; autonomous vehicles "

Recent advances in the field of machine learning and computer vision have rapidly accelerated the development of autonomous vehicles (AVs). While some level of automation has been present in commercially available vehicles since the introduction of cruise control (Teetor, 1950), the component parts of a car capable of fully autonomous driving in urban environments are only now all coming into place (Badue et al., 2020). Considerable public interest is following the successes and obstacles industry-leaders such as Alphabet subsidiary Waymo or GM subsidiary Cruise experience in the process of testing vehicles on different automation levels (Awad et al., 2020). There are a number of different classification systems already in place to characterize autonomous vehicles (NHTSA, 2013; Sousa, Almeida, Coutinho-Rodrigues, Natividade-Jesus, 2018). For example, the UNECE (United Nations Economic Commission for Europe) defines six Functional Categories of Automated Function, which emphasise generic functional requirements for policymaking and safety. In contrast, the most commonly used Level of Automation (LoA) classification system is the SAE (Society of Automotive Engineers, 2014), which specifies five Automation Levels. This emphasises division of authority between the driver and the system, and the highest level is reached by a vehicle which can autonomously navigate under all circumstances, and is in no need of any form of human supervision (Sousa, Almeida, Coutinho-Rodrigues, & Natividade-Jesus, 2018).

What all taxonomies have in common is that the highest level is reached when vehicle is in no need of human supervision – either in the form of supervision by a passenger, or by a third party (another person in the car or in a remote location). The vehicle currently at the most advanced level of automation accessible to the public is Waymo's Level 4 driverless car (according to the SAE's classification system, Ackerman, 2021), currently capable of operating in a geofenced area of Phoenix, AZ under ideal weather conditions. This vehicle is supervised by so-called "remote operators" (Matousek, 2019; Hawkins, 2020), whose job involves intervening in scenarios where the vehicle is faced with an uncertain or novel scenario. As there will ultimately be no need for passengers to take control of these vehicles, lobbying efforts are already underway by industry participants to remove regulations requiring the inclusion of traditional control elements such as pedals and steering wheels from autonomous vehicles (Hawkins, 2019). While currently no vehicle can be described as having reached full automation under all circumstances according to the Society of Automotive Engineers (SAE, 2014) or National Highway Traffic Safety Administration (NHTSA, 2013) frameworks, most industry observers agree that eventually progressing to the highest level of automation is only a matter of time. While technology develops enough for Level 5 automation to become widespread, there will be a time period where varying levels of human-supervised autonomous vehicles share the road with traditional vehicles. Sparrow and Howard (2017) argue that autonomous vehicles relying on human supervision or intervention to any degree will prove a safety risk when used at a large enough scale by the driving public. Merat, Jamson, Lai, Daly and Carsten (2014) report in a driving simulation experiment that on average a driver takes 15 seconds to regain control of a vehicle, with up to 40 seconds to stabilize control – expecting drivers or remote operators to take over a vehicle in a dangerous scenario given these response times is not without risks (Mirnig, Stadler & Tscheligi, 2017). Similarly, remote human supervision (as in the case of Waymo or Cruise at the present moment) will simply not be economically feasible with widespread adoption of vehicles at Level 4 automation. Some researchers go even further, and propose eventually banning manual driving, once reliable-enough Level 5 automation has been established (Müller & Gogoll, 2020; Sparrow & Howard, 2017). It seems that the industry is headed to full automation without any human supervision in the medium- to long-term future, driven by technological developments, economic incentives and ethical considerations. Even though such developments are still in the future, how people currently feel about their risks and benefits now might be indicative of future readiness to adopt these technologies (Davis, Bagozzi & Warshaw, 1989).

Understanding what drives people's current level of comfort with decreasing (or eventually losing) third-party human supervision in autonomous vehicles would go a long way towards devising effective public communication- and information campaigns when the time comes (Bögel et al., 2018). Similarly, being able to characterize attitudes on a higher, macro-level can help inform policymakers when targeting communication strategies on a national level. Given the high level of economic and political integration in the European Union, and the diverse attitudes member states represent regarding a number of relevant issues including the acceptance of robots, or GMOs for example (Gnambs & Appel, 2019, Torgersen & Seifert, 1997), it is important to understand differences between member states. Additionally, identifying similarities across member states could aid policymaking and communication on an international level. Being able to characterize attitudinal differences and similarities when it comes to decreasing human supervision in automotive mobility on a country-level might contribute to tailoring policy and communication in a European context. We start the present study by surveying previous findings on the determinants of attitudes towards self-driving in general, and literature on individuals' preference for human involvement when it comes to automated decision processes. We derive our research questions from this existing literature.

Previous research on attitudes towards self-driving

Understanding people's level of comfort with increasing levels of automation and decreasing human supervision has generated a considerable body of research over the past years (Gkartzonikas & Gkritza, 2019; Alawadhi, Almazrouie, Kamil, & Khalil, 2020). Technology Adoption Models (TAM) (Davis, 1985; Davis, Bagozzi & Warshaw, 1989) are often used to understand factors driving individuals' intentions to adopt new technologies, and predict their eventual use. These models build on Fishbein and Ajzen's Theory of Reasoned Action (Fishbein & Ajzen, 1975; Fishbein & Ajzen, 1980), which states that attitudes and subjective norms drive the behavioral intention to use a certain new product or technology, which in turns drives actual behavior. The potential of attitudes when it comes to predicting technology acceptance has been validated in a number of settings (for example when it comes to internet usage – Porter & Donthu, 2006; online banking – Lai & Li, 2005; or hydrogen fuel cells – Bögel et al. 2018). TAMs have been used specifically to understand attitudes towards innovation in cars and autonomous vehicles (Hewitt, Politis, Amanatidis & Sarkar, 2019; Koul & Eydgahi, 2018). At their core, the main function of TAMs is to characterize the factors that play a role in decisions to ultimately use a technology or not. An intermediate step in this process is identifying the determinants of one's attitudes towards a novel technology. Understanding what drives attitudes towards being a passenger in an autonomous vehicle can contribute to predicting who will actually use such a vehicle once they become widely available.

A number of individual characteristics have been found to be associated with attitudes towards using vehicles at various levels of autonomy, and have been the subject of considerable previous research (Haboucha, Ishaq, & Shiftan, 2017; Becker & Axhausen, 2017). When it comes to gender differences, Bansal, Kockelman and Singh (2016) find women have a lower willingness than men to pay for higher levels of automation, similarly to Pettigrew, Talati, and Norman (2018) and Sener, Zmud and Williams (2019). However, Rödel, Stadler, Meschtscherjakov, and Tscheligi (2014) find no gender difference in attitudes towards higher levels of autonomy. Kyriakidis, Happee, & Winter, (2015) find no consistent gender effect in a large-scale survey when it comes to attitudes, although men do show a higher willingness to pay for automation. Hohenberger, Spörrle, & Welpe (2016) report gender differences in willingness to use autonomous vehicles – this is explained by the mediating effect of affective responses.

Another common finding is that older individuals show more negative attitudes towards autonomous vehicles, and exhibit lower behavioral intentions to use (Rahman, Deb, Strawderman, Burch, & Smith, 2019; Brechin, Farr, Hurley, King, & Willis, 2017). Kyriakidis, Happee and Winter (2015) find that age correlates positively with worries about safety, and they report that individuals are less comfortable without the presence of a wheel in an autonomous vehicle. Pettigrew, Talati and Norman (2018) find that individuals above 60 years of age exhibit more negative attitudes towards self-driving cars, while Bansal, Kockelman and Singh (2016) document lower willingness to pay for higher levels of automation, similarly to Sener, Zmud and Williams (2019). In a nationwide US survey, Owens, Antin, Doerzaph and Willis (2015) found older respondents to be less enthusiastic for advanced vehicle technologies. Hulse, Xie, and Galea (2018) also find that younger and male individuals have a more positive perception of autonomous vehicles, and Hudson, Orviska and Hunady (2019) report significant age and gender effects in the 2014 Eurobarometer data on attitudes about autonomous vehicles.

The potential benefits in mobility caused by the widespread adoption of autonomous vehicles for those living with a disability or medical conditions are significant, as estimated by Harper, Hendrickson, Mangones and Samaras (2016). Currently there is no consensus on how those living with a disability or other medical conditions (which make driving difficult) view autonomous vehicles. From one perspective, those with medical conditions and disabilities often are on average older, and thus might be expected to be more skeptical when it comes to autonomous vehicles (Bansal, Kockelman & Singh, 2016; Kyriakidis, Happee & Winter, 2015). On the other hand, they are also the individuals who could benefit the most from these technologies (Fortunati, Lugano & Manganelli, 2019; Cavoli, Phillips, Cohen & Jones, 2017). Zmud, Sener and Wagner (2016) find a small but significant positive effect on intent to use autonomous vehicles once they become available. The relative lack of evidence regarding this question might be due to the low level of representation of individuals with disabilities in the general population, yielding low statistical power when analyzing a representative sample.

Preference for human decision-makers

Overall, previous findings show that people are not comfortable with AVs yet, and these attitudes have a number of demographic and psychological correlates. Previous literature also finds that increasing levels of automation are associated with increasingly negative attitudes and behavioral intentions for adoption. As Rödel, Stadler, Meschtscherjakov and Tscheligi (2014) and Hewitt, Politis, Amanatidis and Sarkar (2019) report, with increasing levels of automation (progressing along the SAE classification system from 0 through 5), anxiety with using an autonomous vehicle increases, and perceived safety diminishes. This is especially the case when it comes to the transition from Level 4 (requiring supervision) to Level 5 (unsupervised).

Importantly, these previous studies follow the SAE's previously existing system for categorizing the level of automation a vehicle exhibits (SAE, 2014), without taking into account the potential presence or absence of third-party human supervision in each scenario. As all currently circulating autonomous vehicles either have a human operator physically present in the car, or one remotely supervising and intervening when necessary, understanding how the presence of third-party human supervision (or lack thereof) influences potential passengers' comfort with using an autonomous vehicle is essential to anticipate acceptance. Given the fact that Level 4 autonomy with remote human supervision is presently the most advanced use case at the forefront of AV development, and future developments will with a high likelihood bring fully autonomous cars lacking supervision by third parties, it is especially worthwhile to investigate what happens to consumer attitudes when external human supervision goes away completely. And more importantly, what are the individual-level determinants of comfort with such a scenario?

Aside from previously documented correlates of attitudes towards self-driving cars in general, the present research is also motivated by previous literature on how individuals perceive it when machines make decisions for them, especially when it comes to decisions with moral implications, as can be the case when driving (Awad et al. 2018; Awad, Dsouza, Shari, Rahwan & Bonnefon, 2020). Previous literature on the permissibility of letting machines make life-or-death decisions in a driving context shows that people have strong preferences for human involvement (Bigman & Gray, 2018) – participants overwhelmingly prefer that humans make such decisions. Castelo, Bos and Lehmann (2019) report in a series of studies that people show a strong preference for human drivers versus algorithms, even in a situation where they are informed that the algorithm shows superior performance. Additionally, recent findings on assigning blame to a human passenger or to a self-driving car when someone is harmed shows that blame is apportioned unequally between human and machine. Awad et al. (2020) show that the most of the blame still falls on the passenger/operator of an autonomous vehicle. Given this dynamic, potential passengers' reluctance to get third-party human supervision (be it in-car or remote) out of the loop would be understandable.

Outside the automotive context, Dietvorst, Simmons and Massey (2018) find that individuals prefer algorithms over which they can exert some control – a potential for including human influence on an algorithm's decision seems to put people at ease. There are a number of real-world applications where so-called human-in-the-loop systems are present (Nahavandi, 2017; Berman, 2019). In such scenarios a machine is working autonomously, while being supervised by a human operator who provides an input in case the machine is uncertain how to interpret a situation. The advanced driving frameworks by Alphabet subsidiary Waymo and GM subsidiary Cruise are further examples for such a system.

To summarize, people show a strong preference for involving human decision-making. This is also true when it comes to making decisions in the context of driving. In addition, as we have seen in our previous literature review on the individual correlates and attitudes towards AVs, people's attitudes towards being a passenger in an autonomous vehicle can be at least partially explained by their individual characteristics. Given that external human supervision of autonomous vehicles is expected to be present in the short- and medium term, and is expected to decline as the technology becomes more widespread, this is a topic that deserves increased research attention. We need a better understanding of the determinants for attitudes towards keeping humans in the loop when it comes to autonomous vehicles.

Our main research questions are as follows:

- What are people's attitudes towards fully automated vehicles at different levels of third-party human supervision?
- What are the individual-level predictors for comfort at different levels of supervision?
- Does the importance of previously identified predictors for comfort with autonomous vehicles change, depending on the level of third-party human supervision present?

Data

In order to investigate these questions, we rely on Eurobarometer survey data (European Commission, 2020), collected during September 2019 (*n*=27565). This cross-sectional dataset contains information on individual demographic characteristics, and attitudes when it comes to autonomous vehicles and other relevant variables. The data was collected in-person by trained interviewers in the respondent's household and is freely available for purposes of academic research. The Eurobarometer covers all current member states of the European Union at the time of data collection. About one thousand randomly selected individuals (ages 15 and older) are contacted by Kantar Public in most countries, with about 1500 in Germany, and 500 in Cyprus and Malta and Luxembourg. The survey is representative on the level of individual countries. The three questions included in the survey which constitute our dependent variables are framed as seen in Table 1.

Methods

The content and structure of the three survey questions makes it possible for us to investigate not just what determines answers to these individual items in isolation, but also gives us the opportunity to investigate changing influences when it comes to systematically decreasing levels of third-party human supervision (for an overview of our dependent variables see Table 1.)

We will start by profiling the development of responses on the level of the individual, and then on an aggregate level, to characterize country-level patterns. This descriptive analysis is meant to give a high-level overview of our data, before attempting to identify the most important predictors behind these response patterns.

Then we will analyze individual responses in a regression framework, as respondents state their level of comfort in three hypothetical self-driving scenarios with decreasing levels of external human supervision. This is done by constructing three ordered logistic regression models, as our dependent variables are ordinal (comfort ratings ranging from "*Not at all comfortable*" to "*Totally comfortable"*). As our predictors, we have selected a wide range of variables to include in our models, guided by previous literature: most demographic variables which have been found in past research to be predictive of attitudes towards self-driving vehicles, and are available to us have been included. These variables include age, gender, marital status, household size, education level, occupation, internet and telephone use. In addition, we have included a number of attitudinal measures from the original survey, which are also expected to be predictive of comfort with autonomous driving at different levels of human supervision. These measures include for example, items surveying comfort with sharing data with other road users, governments or private companies.

In order to better understand the drivers of comfort across varying levels of human supervision, we have constructed three ordinal random forest models (Hornung, 2020), where we leverage our rich survey dataset to predict comfort with our three scenarios/dependent variables. Similar to our earlier regressions, our approach respects the ordinality of our dependent variables – not assuming equal distance between answer options as an ordinary least squares approach would, while still retaining the hierarchical nature of the answer options. This analysis does not rely on stringent model assumptions due to its nonparametric nature, while also being less prone to be influenced by multicollinearity. Furthermore, due to the iterative nature of the algorithm, it is able to account for interactions without including previously specified interaction effects. Most importantly, this method provides a permutation variable importance metric to quantify the overall influence certain variables have in a model, and which then enables us to compare how this influence changes from one model to the next (in our case from dependent variable to dependent variable). Interpreting changes in variable importance provides valuable insight into how the effect of various predictors (for example gender, or having a driver's license) might have a different effect on one's comfort level, depending on the level of third-party human supervision present. Our models were fitted using default hyperparameters, as they were all in a suitable range for the size and complexity of our dataset (Hornung, 2020). One downside of random forest models is that the time required to fit them increases exponentially with the complexity of the dataset. As one of our key predictor variables (*Country*) has 28 levels, iterating through each value and selectively including them or not including them would lead to a large computational burden, and ultimately would not reflect the role of country-differences in our dataset accurately. One-hot encoding a variable with 28 values runs into similar constraints. In order to address this issue, we first collapsed our complete dataset into 4 factors (as suggested by a scree-plot) using Multiple Correspondence Analysis (Kassambara & Mundt, 2017; Sebastien Le, et al., 2008). This method is conceptually similar to principal component analysis, but it needs categorical data (which we have) to reduce dimensionality. After generating factor scores for each individual, we aggregated these scores by country, to represent country-level variability in our models without having to include 28 additional variables. Following this, we performed k-means clustering, in order to create 5 clusters of countries, based on their similarity to each other in our dataset. The resulting permutation variable importance rankings for each model will enable us to visualize and interpret the changing role variables such as age, gender etc. play across our three models.

Results

Descriptive analysis of dependent variables.

When examining the distribution of responses from question to question with decreasing levels of human supervision ("*Comfort with in-car human supervision*", "*Comfort with remote human supervision*", and "*Comfort with no human supervision*"), we can clearly observe a trend of decreasing comfort. While 33.5% are completely comfortable with being a passenger in the physical presence of a human operator, this value decreases to 10.4% when it comes to remote supervision, and further to 6.6% in the case of no human supervision. (Table 1.). The data clearly shows that respondents feel more comfortable with more human supervision.

Interestingly, we observe that this shift in individual attitudes is not monotonous. We find that a number of individuals exhibit more positive attitudes towards remote supervision compared to in-car human supervision (Figure 1.). Similarly, a smaller number of individuals even prefer full autonomy to supervised options. Our survey dataset enables us to characterize individuals who react to decreasing third-party human supervision with increasing comfort. There are a number of countries where this response pattern is two-to-three times more common than in the median country. Overall, Romania, Hungary and Poland give 24% of those who prefer less supervision at some point to more supervision, while constituting only 10.8% of the complete sample. Fisher's exact tests show that the distribution of respondents who either prefer remote human supervision to in-car supervision ($p<0.001$) or who prefer no third-party human supervision to remote supervision ($p<0.001$) is not independent from country of origin. This pattern can be interpreted as follows: while there seems to be a general aversion to decreasing third-party supervision, for a substantial number of respondents in some countries, lower levels of human supervision are associated with higher comfort. While it is beyond the scope of the present paper to investigate the underlying cause of such preferences in depth, a potential explanation could be lower interpersonal trust levels in some of these countries (Pellegrini, De Cristofaro, Salvati, Giacomantonio & Leone, 2021).

At a higher level of aggregation (on the level of countries), we find that samples in various member states have divergent views on human supervision of autonomous vehicles. Some Nordic countries such as Sweden, Denmark and Finland overall show higher comfort with in-car human supervision and remote supervision than most other countries, while Eastern-European countries such as Poland and Romania are also relatively comfortable with remote supervision (Figure 2. – in all plots and analyses higher values indicate higher comfort). However, when it comes to unsupervised self-driving vehicles, comfort in Nordic countries for example drops considerably below that of Eastern European countries. A further interesting case is that of the Netherlands, where respondents are some of the most open to in-car supervision, while they are among the least comfortable with no third-party human supervision. In order to understand what might be contributing to these observed shifts in comfort, and how these effects might be different depending on the level of human supervision at hand, we implement three ordinal logistic regression models, predicting our three dependent variables separately.

Results from regression models and permutation variable importance.

In this section we will describe results from our regression models first, then to interpret the changing role of these variables from model to model (in decreasing third-party human supervision), we characterize the dynamics of our permutation variable importance rankings. Table 2. contains detailed information on the regression models, while Figure 3. depicts variable importance rankings from our random forest models. The random forest approach provides a very similar overall picture when it comes to the effect of individual variables, which serves as an additional validation of our regression models.

In line with previous literature, we document a significant effect of gender on comfort with being a passenger in a fully autonomous vehicle. Males report being significantly more comfortable, independently of the level of third-party human supervision present. When comparing the coefficient estimates across our regression models, we find a growing gender gap with decreasing third-party human supervision. Similarly, our variable importance rankings show that the effect of gender increases substantially across our three random forest models. While gender has low variable importance when it comes to predicting comfort with in-car human supervision (Model 1), it becomes one of the more important variables when it comes to predicting comfort with remote human supervision (Model 2). Gender emerges as the fifth most important variable in Model 3, predicting comfort with unsupervised driving.

When it comes to the effect of age, we find that younger respondents are more comfortable with being a passenger in a fully autonomous vehicle, independently of the level of third-party human supervision present. This is also reflected in the relatively stable permutation variable importance rankings from our random forest models. We find respondent age to be consistently one of the most important variables in all three random forest models.

The Eurobarometer includes a proxy for education level, namely the age at which one has finished their education. Higher age of exiting education corresponds to higher level of educational attainment overall, and it has been found by Hudson, Orviska & Hunady (2019) to be a good predictor for attitudes towards self-driving cars and trucks. Our regression models show that individuals who exited the education system at age 17 or before tend to be substantially less open to traveling in an autonomous vehicle. Conversely, those who have ended their education at the age of 22 or older are significantly more open at all levels of third-party human supervision. This finding might be partially due to a cohort effect, where older respondents also report lower educational attainment in most EU countries (Eurostat, 2020). When it comes to one's occupational status, it seems those currently classified as unemployed are consistently and significantly less comfortable with the prospect of traveling in a fully autonomous vehicle.

While phone access seems to have substantial predictive power at lower levels of third-party human supervision (such that those with mobile phone access are significantly more comfortable with AVs), this effect disappears in Model 3., where no human supervision is present. Our importance ranking shows this variable to not contribute much to our models. A key finding is that one's self-reported frequency of Internet use overall predicts comfort in our first two models. It seems that even after accounting for associated variables such as age, education and social status, those who use the Internet daily report higher comfort with being a passenger in a fully autonomous vehicle. However, our variable importance metric shows that the contribution of this variable declines substantially from Model 1 to Model 3.

Variables such as internet use and phone access are most likely proxies for an interest or capacity to keep up with technological developments. Our respondents were also asked if they recall having heard of autonomous vehicles over the past 12 months. This variable had a significant effect in the first model, and no significant effect in the second and third models. While having heard of AVs recently positively influences people's comfort with traveling in an AV with in-person human supervision, it has no effect on comfort with lower levels of human supervision. Similarly to our two variables discussed previously (phone access and internet use), the importance ranking of this variable decreases consistently from Model 1 to Model 3. It seems that previous knowledge of autonomous vehicles only affects one's attitudes towards actually traveling in one if there is a human operator physically present – when this element is not the case, one's level of previous information does not seem to make a difference.

We find that having a driver's license does not have a large effect – while in our first model there is a small significant negative effect on comfort, there are no differences in Models 2 and 3. While previous studies report that having a driver's license is associated with more positive attitudes towards AVs, we do not document an influence of this factor. Interestingly, living with a disability or reduced mobility seems to have a negative effect on one's comfort with traveling in a fully autonomous vehicle at lower levels of supervision. This status seems to have a significant negative effect on comfort in our first two models, which gets attenuated in the third model.

What immediately becomes clear when surveying our three regression models (Table 2.) and our permutation variable importance rankings (Figure 3.) is the key role attitudes towards sharing data with third parties play in predicting one's comfort with different self-driving scenarios. In accordance with the overall high importance placed on data privacy present in Europe (with considerable heterogeneity in attitudes and practices across member states, Lusoli et al., 2012), self-reported level of comfort with sharing data with other road users, one's government or even public companies is predictive of one's comfort level when it comes to AVs. Individuals who are more comfortable with data sharing are more comfortable with using self-driving vehicles at all levels of supervision. The importance of data privacy is highlighted by our variable importance measure, where these are consistently among the variables with the highest importance. Interestingly, we find that sharing data with private companies becomes more important from Model 1 to Model 3 – this observation is also supported when surveying the coefficient estimates for different levels of this variable in our regression analyses.

When observing the odds ratios for country effects in our regression, we find a similar pattern to our earlier descriptive analyses – the positive effect from countries such as Sweden, Finland and the Netherlands drop considerably in Models 2 and 3, while the effect from Romania, Poland for example remain mostly unchanged compared to other countries. Our variable importance ranking shows a slight increase in the importance of our country cluster variable from Model 1 to Model 3.

Discussion & Conclusion

Overall, our combined analyses give us some key insights. Age, education and gender have a significant effect in the expected direction in all models.

The importance of gender increases as third-party supervision decreases. It seems that the varying effect of gender on self-driving attitudes in the literature (Hohenberger, Spörrle & Welpe, 2016; Rödel, Stadler, Meschtscherjakov & Tscheligi, 2014) can at least partly be explained by attitudes towards third-party human supervision – gender differences get stronger with decreasing supervision, with men being increasingly comfortable in such a scenario.

Attitudes towards data sharing with third parties are also consistently significant across all models. The importance of attitudes towards sharing data with private companies increases consistently from Model 1 to Model 3. Internet and telephone use, and having recently heard of AVs also has an effect in Models 1 and 2. However, all three of these variables decrease in importance from Model 1 to Model 3. Living with a disability or medical conditions seems to have a slight negative effect on comfort with traveling in an autonomous vehicle in Models 1 and 2. We also find that after controlling for individual characteristics, we still have country effects which broadly reflect our descriptive analysis.

Overall, we have documented that respondents to the Eurobarometer survey are less comfortable with lower levels of third-party human supervision in a fully autonomous vehicle. While there is considerable heterogeneity in how individuals responded to these three questions, the overall negative trajectory is very clear. When it comes to country effects, we find both in our descriptive- as well as our regression analyses that there are marked differences on a national level – not just within each individual model, but also from one model to the next. This implies that the dynamics of how comfortable inhabitants of different countries are with varying levels of human supervision can be quite different. These country effects are robust when accounting for a large number of individual characteristics. This heterogeneity should inform those trying to shape policy and approaches to communication in a European context when it comes to the future of autonomous vehicles – while ultimately there is a need for a harmonized framework across the EU when it comes to these issues, a one-size-fits all approach seems challenging in the face of such a diversity of attitudes. In addition, further study of cultural aspects when it comes to acceptance of unsupervised self-driving technologies is called for in the future.

When it comes to individual characteristics, our findings are largely in line with existing literature (c.f. Haboucha, Ishaq, & Shiftan, 2017; Becker & Axhausen, 2017; Bansal, Kockelman & Singh, 2016; Sener, Zmud & Williams, 2019). Younger, male and more educated individuals would be more comfortable overall with traveling in a fully autonomous vehicle. We find that level of education plays a similarly important role across all models: individuals who spent more time in formal education are in general more comfortable with self-driving at all levels of supervision. In what might be a phenomenon unique to a European context, we find that some of the strongest and most consistent predictors of comfort with being a passenger in a self-driving car is comfort with sharing data with other road users, authorities and private companies. Brell, Philipsen and Ziefle (2018) similarly find in a survey with 516 German respondents that potential users of autonomous vehicles have concerns regarding how their personal data is handled in such a setting – furthermore, these concerns are not abated by increased experience, as is the case for other perceived risks. These observations highlight the need for transparent regulation and communication when it comes to the user data AV operators have access to, and how this is shared with governments and other road users. We also document that the importance of attitudes towards sharing data with private companies increases with lower supervision.

We additionally find that the effect of gender depends on the level of human supervision present to a large extent. Female respondents become substantially less comfortable with being in a fully automated vehicle only when in-car human supervision is replaced by remote human supervision. A further shift to complete lack of external human supervision drives an even larger deviation in comfort. This finding might help explain some previously described contradictory findings in the literature when it comes to gender as predictors for AV acceptance (Hohenberger, Spörrle & Welpe, 2016; Rödel, Stadler, Meschtscherjakov & Tscheligi, 2014). It seems that the effect of gender on attitudes towards travelling in autonomous vehicles is moderated by the level of third-party human supervision present.

While delivering interesting new insights, the current study is not without its limitations. Most importantly, it relies on a cross-sectional dataset, which does not enable us to establish the existence of causal relationships between our predictors and our dependent variables. Furthermore, while the data is representative on the level of individual countries, certain segments of the public (for example those living with disability or reduced mobility) might still be underrepresented in terms of statistical power. In addition, one might wonder about the hypothetical nature of our dependent variables, when there is currently no fully autonomous vehicle on the road without human supervision.

However, given the potentially transformative role such future developments will have, and the fact that attitudes towards the risks and benefits of technological innovation have been shown in a number of studies to predict their eventual acceptance (c.f. Porter & Donthu, 2006; Bögel et al., 2018), we believe the current formulation of our study is well-justified. While attitudes will most certainly change as technologies develop, having a baseline understanding of the various factors at play is essential. We hope that future research will delve deeper into the findings we document here. Furthermore, we hope that the causal relationships driving the changes we observed in our correlational setting will eventually be investigated in an experimental fashion, enabling a more targeted approach when it comes to guiding public policy and communications.

References

Ackerman, E. (2021 March 04.) What Full Autonomy Means for the Waymo Driver. *IEEE Explore*. Retrieved from: https://spectrum.ieee.org/cars-that-think/transportation/self-driving/full-autonomy-waymo-driver

Alawadhi, M., Almazrouie, J., Kamil, M., & Khalil, K. A. (2020). A systematic literature review of the factors influencing the adoption of autonomous driving. *International Journal of System Assurance Engineering and Management*, 1–18.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shari, A., . . . Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59–64.

Awad, E., Dsouza, S., Shari, A., Rahwan, I., & Bonnefon, J.-F. (2020). Universals and variations in moral decisions made in 42 countries by 70,000 participants. *Proceedings of the National Academy of Sciences*, *117* (5), 2332–2337.

Awad, E., Levine, S., Kleiman-Weiner, M., Dsouza, S., Tenenbaum, J. B., Shari, A., . . . Rahwan, I. (2020). Drivers are blamed more than their automated cars when both make mistakes. *Nature Human Behaviour*, *4* (2), 134–143.

Badue, C., Guidolini, R., Carneiro, R. V., Azevedo, P., Cardoso, V. B., Forechi, A., . . . others. (2020). Self-driving cars: A survey. *Expert Systems with Applications*, 113816.

Bansal, P., Kockelman, K. M., & Singh, A. (2016). Assessing public opinions of and interest in new vehicle technologies: An austin perspective. *Transportation Research Part C: Emerging Technologies*, *67*, 1–14.

Becker, F., & Axhausen, K. W. (2017). Literature review on surveys investigating the acceptance of automated vehicles. *Transportation*, *44* (6), 1293–1306.

Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, *181*, 21–34.

Bögel, P., Oltra, C., Sala, R., Lores, M., Upham, P., Dütschke, E., . . . Wiemann, P. (2018). The role of attitudes in technology acceptance management: Reflections on the case of hydrogen fuel cells in europe. *Journal of Cleaner Production*, *188*, 125–135. https://doi.org/https://doi.org/10.1016/j.jclepro.2018.03.266

Brechin, S., Farr, M., Hurley, J., King, P., & Willis, N. (2017). Age and safety perceptions in autonomous vehicles. *University of Christchurch, New Zealand.*

Brell, T., Philipsen, R., & Ziefle, M. (2019). sCARy! Risk perceptions in autonomous driving: The influence of experience on perceived benefits and barriers. *Risk analysis*, *39*(2), 342-357.

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

Cavoli, C., Phillips, B., Cohen, T., & Jones, P. (2017). *Social and behavioural questions associated with Automated Vehicles A Literature Review*. UCL Transport Institute January.

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (PhD thesis). Massachusetts Institute of Technology.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982–1003.

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

European Commission, B. (2020). Eurobarometer 92.1 (2019). GESIS Datenarchiv, Köln. ZA7579 Datenfile Version 2.0.0, https://doi.org/10.4232/1.13541. https://doi.org/10.4232/1.13541

Eurostat (2020) *Level of educational attainment by age*. Retrieved from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Educational_attainment_statistics#Level_of_educational_attainment_by_age

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention, and behavior reading. *MA: Addison-Wesley*, 913–927.

Fishbein, M., & Ajzen, I. (1980). Predicting and understanding consumer behavior: Attitude-behavior correspondence. *Understanding Attitudes and Predicting Social Behavior*, 148–172.

Fortunati, L., Lugano, G., & Manganelli, A. M. (2019). European perceptions of autonomous and robotized cars. *International Journal of Communication*, *13*, 2728–2747.

Gkartzonikas, C., & Gkritza, K. (2019). What have we learned? A review of stated preference and choice studies on autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, *98*, 323–337.

Gnambs, T., & Appel, M. (2019). Are robots becoming unpopular? Changes in attitudes towards autonomous robotic systems in europe. *Computers in Human Behavior*, *93*, 53–61.

Haboucha, C. J., Ishaq, R., & Shiftan, Y. (2017). User preferences regarding autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, *78*, 37–49.

Harper, C. D., Hendrickson, C. T., Mangones, S., & Samaras, C. (2016). Estimating potential increases in travel with autonomous vehicles for the non-driving, elderly and people with travel-restrictive medical conditions. *Transportation Research Part C: Emerging Technologies*, *72*, 1–9.

Hawkins, A. J. (2019). Self-driving carmakers urge regulators to whi the steering wheel out the window. Retrieved from https://www.theverge.com/2019/8/30/20840631/self- driving-carmakers-federal-safety-rules-nhtsa-steering-wheels-pedals-waymo-cruise

Hawkins, A. J. (2020). Coronavirus shows there's still no such thing as a totally human-free self-driving car. Retrieved from https://www.theverge.com/2020/3/18/21183421/self- driving-cars-coronavirus-covid19-human-drivers-waymo-cruise-zoox

Hewitt, C., Politis, I., Amanatidis, T., & Sarkar, A. (2019). Assessing public perception of self-driving cars: The autonomous vehicle acceptance model. In *Proceedings of the 24th international conference on intelligent user interfaces* (pp. 518–527).

Hohenberger, C., Spörrle, M., & Welpe, I. M. (2016). How and why do men and women dier in their willingness to use automated cars? The influence of emotions across dierent age groups. *Transportation Research Part A: Policy and Practice*, *94*, 374–385.

Hornung, R. (2020). *OrdinalForest: Ordinal forests: Prediction and variable ranking with ordinal target variables*. Retrieved from https://CRAN.R-project.org/package=ordinalForest

Hudson, J., Orviska, M., & Hunady, J. (2019). People's attitudes to autonomous vehicles. *Transportation Research Part A: Policy and Practice*, *121*, 164–176.

Hulse, L. M., Xie, H., & Galea, E. R. (2018). Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety Science*, *102*, 1–13.

Kassambara, A., & Mundt, F. (2017). Package 'factoextra'. Extract and visualize the results of multivariate data analyses, 76.

Koul, S., & Eydgahi, A. (2018). Utilizing technology acceptance model (tam) for driverless car technology adoption. *Journal of Technology Management & Innovation*, *13* (4), 37–46.

Kyriakidis, M., Happee, R., & Winter, J. C. de. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Trac Psychology and Behaviour*, *32*, 127–140.

Lai, V. S., & Li, H. (2005). Technology acceptance model for internet banking: An invariance analysis. *Information & Management*, *42* (2), 373–386.

Lusoli, W., Bacigalupo, M., Lupiáñez-Villanueva, F., Andrade, N. N. G. D., Monteleone, S., & Maghiros, I. (2012). Pan-European survey of practices, attitudes and policy preferences as regards personal identity data management. *JRC Scientific and Policy Reports, EUR*, *25295*.

Matousek, M. (2019). Waymo is telling customers it will start offering rides in its autonomous cars without safety drivers. Retrieved from https://www.businessinsider.com/waymo-says-it-will-start-giving-rides-without- safety-drivers-2019-10?r=US&IR=T

Merat, N., Jamson, A. H., Lai, F. C., Daly, M., & Carsten, O. M. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Trac Psychology and Behaviour*, *27*, 274–282.

Mirnig, A. G., Stadler, S., & Tscheligi, M. (2017). Handovers and resumption of control in semi-autonomous vehicles: What the automotive domain can learn from human-robot-interaction. In *Proceedings of the companion of the 2017 acm/ieee international conference on human-robot interaction* (pp. 207–208).

Müller, J. F., & Gogoll, J. (2020). Should manual driving be (eventually) outlawed? *Science and Engineering Ethics*, 1–19.

Nahavandi, S. (2017). Trusted autonomy between humans and robots: Toward human-on-the-loop in robotics and autonomous systems. *IEEE Systems, Man, and Cybernetics Magazine*, *3*(1), 10–17.

National Highway Traffic Safety Administration. (2013). *Preliminary statement of policy concerning automated vehicles*. Washington, DC, 1, 14.

Owens, J. M., Antin, J. F., Doerzaph, Z., & Willis, S. (2015). Cross-generational acceptance of and interest in advanced vehicle technologies: A nationwide survey. *Transportation Research Part F: Trac Psychology and Behaviour*, *35*, 139–151.

Pellegrini, V., De Cristofaro, V., Salvati, M., Giacomantonio, M., & Leone, L. (2021). Social Exclusion and Anti-Immigration Attitudes in Europe: The mediating role of Interpersonal Trust. *Social Indicators Research*, 1-28.

Pettigrew, S., Talati, Z., & Norman, R. (2018). The health benefits of autonomous vehicles: Public awareness and receptivity in australia. *Australian and New Zealand Journal of Public Health*, *42*(5), 480–483.

Porter, C. E., & Donthu, N. (2006). Using the technology acceptance model to explain how attitudes determine internet usage: The role of perceived access barriers and demographics. *Journal of Business Research*, *59* (9), 999–1007.

Rahman, M. M., Deb, S., Strawderman, L., Burch, R., & Smith, B. (2019). How the older population perceives self-driving vehicles. *Transportation Research Part F: Trac Psychology and Behaviour*, *65*, 242–257.

Rödel, C., Stadler, S., Meschtscherjakov, A., & Tscheligi, M. (2014). Towards autonomous cars: The effect of autonomy levels on acceptance and user experience. In *Proceedings of the 6th international conference on automotive user interfaces and interactive vehicular applications* (pp. 1–8).

SAE. (2014). Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. In *Society of Automotive Engineers: On-Road Automated Vehicle Standards Committee*. SAE (Society of Automotive Engineers) Pub. Inc.

Sebastien Le, Julie Josse, Francois Husson (2008). FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software, 25(1), 1-18. 10.18637/jss.v025.i01

Sener, I. N., Zmud, J., & Williams, T. (2019). Measures of baseline intent to use automated vehicles: A case study of texas cities. *Transportation Research Part F: Trac Psychology and Behaviour*, *62*, 66–77.

Sousa, N., Almeida, A., Coutinho-Rodrigues, J., & Natividade-Jesus, E. (2018). Dawn of autonomous vehicles: Review and challenges ahead. In *Proceedings of the institution of civil engineers-municipal engineer* (Vol. 171, pp. 3–14). Thomas Telford Ltd.

Sparrow, R., & Howard, M. (2017). When human beings are like drunk robots: Driverless vehicles, ethics, and the future of transport. *Transportation Research Part C: Emerging Technologies*, *80*, 206–215.

Teetor, R. R. (1950). Speed control device for resisting operation of the accelerator. *Google Patents*.

Torgersen, H., & Seifert, F. (1997). Aversion preceding rejection: Results of the eurobarometer survey 39.1 on biotechnology and genetic engineering in austria. *Public Understanding of Science*, *6*(2), 131–142.

Zmud, J., Sener, I. N., & Wagner, J. (2016). Self-driving vehicles: Determinants of adoption and conditions of usage. *Transportation Research Record*, *2565* (1), 57–64.

Tables

Table 1. Distribution of dependent variables

	Level	N	%
With the supervision of a human operator in the vehicle	Totally comfortable	9223	33.5
	Fairly comfortable	9988	36.2
	Not very comfortable	3976	14.4
	Not at all comfortable	3669	13.3
	Don't know / NA	709	2.6
With the remote supervision of a human operator	Totally comfortable	2855	10.4
	Fairly comfortable	7219	26.2
	Not very comfortable	8706	31.6
	Not at all comfortable	7997	29
	Don't know / NA	788	2.9
Without the supervision of a human operator	Totally comfortable	1815	6.6
	Fairly comfortable	4200	15.2
	Not very comfortable	7240	26.3
	Not at all comfortable	13406	48.6
	Don't know / NA	904	3.3

Note: Answers to the following question: «To what extent would you feel comfortable or not travelling in a fully automated vehicle under the following conditions:

Table 2. Ordinal logistic regression models

		In-car supervision	Remote supervision	No supervision
Gender	Male	1.057 ** (0.049)	1.178 *** (0.055)	1.315 *** (0.084)
	Female	0.946 ** (0.044)	0.849 *** (0.039)	0.761 *** (0.048)
Age	15-24	1.031 (0.190)	1.271 *** (0.162)	1.290 *** (0.161)
	25-34	1.115 ** (0.107)	1.126 *** (0.109)	1.094 * (0.124)
	35-44	1.038 (0.077)	1.027 (0.077)	1.093 ** (0.090)
	45-54	1.020 (0.081)	1.004 (0.074)	1.013 (0.075)
	55-64	1.005 (0.089)	0.904 *** (0.063)	0.870 *** (0.069)
	65 years and older	0.816 *** (0.082)	0.750 *** (0.081)	0.736 *** (0.087)
Marital status	(Re-)Married	0.907 (0.156)	1.014 (0.190)	1.036 (0.206)
	Divorced/Separated	0.866 * (0.145)	1.011 (0.208)	1.024 (0.234)
	Other	1.160 (0.486)	1.066 (0.465)	1.050 (0.552)
	Refusal	1.468 (0.961)	0.966 (0.883)	0.847 (0.807)
	Single	0.925 (0.168)	1.046 (0.196)	1.092 (0.236)
	Single living with partner	0.992 (0.132)	1.045 (0.245)	1.076 (0.279)
	Widow	0.814 ** (0.162)	0.868 (0.199)	0.904 (0.230)
Education ended at age	14 years or earlier	0.806 ** (0.152)	0.715 *** (0.121)	0.775 ** (0.176)
	15	0.828 * (0.190)	0.846 * (0.189)	0.827 ** (0.168)
	16	0.918 (0.142)	0.839 ** (0.132)	0.838 * (0.165)
	17	1.020 (0.162)	0.878 ** (0.103)	0.822 ** (0.151)
	18	1.036 (0.135)	0.958 (0.098)	0.955 (0.103)
	19	1.005 (0.123)	0.998 (0.099)	1.026 (0.112)
	20	1.000 (0.175)	1.026 (0.151)	0.932 (0.168)
	21	1.092 (0.191)	1.099 (0.172)	0.979 (0.169)
	22+	1.188 *** (0.124)	1.209 *** (0.105)	1.091 * (0.105)
	DK	1.057 (0.367)	1.025 (0.284)	1.042 (0.353)
	No full-time education	0.733 *** (0.184)	1.017 (0.390)	1.049 (0.360)
	Refusal	1.054 (0.948)	1.135 (0.910)	1.763 ** (1.138)
	Still studying	1.446 *** (0.370)	1.455 *** (0.307)	1.206 ** (0.233)
Occupation	Employed	1.023 (0.092)	1.016 (0.061)	1.000 (0.070)
	Not working	0.973 (0.060)	0.909 *** (0.059)	0.899 *** (0.076)
	Self-employed	1.003 (0.110)	1.083 * (0.093)	1.112 ** (0.111)
Phone	Landline and Mobile	1.231 *** (0.153)	1.041 (0.114)	0.958 (0.107)
	Landline only	0.742 ** (0.194)	0.881 (0.162)	0.913 (0.162)
	Mobile only	1.169 *** (0.128)	1.138 ** (0.137)	1.058 (0.110)
	No telephone	0.936 (0.160)	0.959 (0.191)	1.081 (0.192)
Internet use index	Every day / almost every day	1.331 *** (0.180)	1.233 *** (0.197)	1.064 (0.168)
	Once a week	0.962 (0.241)	1.007 (0.224)	1.067 (0.239)
	Less often	1.156 (0.307)	1.285 * (0.364)	1.346 ** (0.371)
	Never	0.941 (0.120)	0.946 (0.093)	0.911 (0.137)
	No access	0.983 (0.279)	0.833 (0.240)	0.811 (0.298)
	Two or three times a month	0.720 (0.361)	0.789 (0.345)	0.937 (0.506)
	Two or three times a week	1.015 (0.167)	1.010 (0.149)	0.945 (0.156)
Country	Austria	0.794 *** (0.031)	1.117 *** (0.049)	1.273 *** (0.054)
	Belgium	0.956 *** (0.036)	0.915 *** (0.031)	1.076 *** (0.034)
	Bulgaria	1.040 (0.062)	1.303 *** (0.096)	1.155 *** (0.084)
	Cyprus	1.236 *** (0.056)	0.597 *** (0.025)	0.713 *** (0.030)
	Czech Republic	1.526 *** (0.099)	0.864 *** (0.042)	0.736 *** (0.046)
	Germany	0.943 *** (0.044)	0.843 *** (0.050)	0.768 *** (0.046)
	Denmark	0.987 (0.072)	1.084 ** (0.082)	0.869 *** (0.059)
	Estonia	0.919 *** (0.051)	0.720 *** (0.038)	0.831 *** (0.049)
	Spain	0.532 *** (0.039)	0.733 *** (0.037)	0.878 *** (0.050)
	Finland	1.518 *** (0.110)	0.856 *** (0.047)	0.579 *** (0.032)
	France	1.061 ** (0.054)	0.999 (0.050)	0.953 ** (0.044)
	Great Britain	0.945 * (0.057)	0.939 ** (0.052)	1.020 (0.066)
	Greece	1.042 (0.063)	0.823 *** (0.041)	0.904 *** (0.039)
	Croatia	0.869 *** (0.049)	1.043 ** (0.046)	1.209 *** (0.056)
	Hungary	1.114 *** (0.050)	1.530 *** (0.078)	1.590 *** (0.090)
	Ireland	0.805 *** (0.028)	1.346 *** (0.023)	1.715 *** (0.079)
	Italy	0.941 *** (0.042)	0.955 ** (0.040)	1.116 *** (0.065)
	Lithuania	1.114 *** (0.055)	0.829 *** (0.034)	0.826 *** (0.043)
	Luxembourg	0.883 *** (0.050)	0.899 *** (0.062)	0.854 *** (0.069)
	Latvia	1.101 *** (0.046)	0.730 *** (0.035)	0.544 *** (0.040)
	Malta	0.774 *** (0.056)	1.181 *** (0.088)	1.392 *** (0.116)
	Netherlands	1.493 *** (0.120)	1.131 *** (0.104)	0.746 *** (0.065)
	Poland	1.204 *** (0.073)	1.699 *** (0.130)	1.434 *** (0.096)
	Portugal	0.669 *** (0.038)	1.240 *** (0.077)	1.608 *** (0.140)
	Romania	1.272 *** (0.095)	1.411 *** (0.122)	1.495 *** (0.123)
	Sweden	1.655 *** (0.130)	1.161 *** (0.083)	0.742 *** (0.060)
	Slovenia	0.874 *** (0.052)	0.787 *** (0.042)	0.951 ** (0.042)
	Slovakia	0.687 *** (0.040)	1.126 *** (0.057)	1.364 *** (0.074)
Community size	Rural area	0.959 (0.079)	0.914 *** (0.066)	0.910 ** (0.076)
	City or large urban area	1.047 (0.075)	1.122 *** (0.071)	1.126 ** (0.114)
	Towns and suburbs or small urban area	0.996 (0.077)	0.976 (0.056)	0.976 (0.067)
Driver's licence	No	0.961 * (0.051)	0.974 (0.053)	1.000 (0.055)
	Yes	1.041 * (0.055)	1.027 (0.056)	1.000 (0.055)
Reduced mobility or disability	No	1.094 *** (0.076)	1.061 ** (0.063)	1.038 (0.059)
	Yes	0.914 *** (0.063)	0.943 ** (0.056)	0.963 (0.055)
Heard about AV within last 12 months	No	0.864 ** (0.118)	0.915 (0.157)	0.918 (0.136)
	DK	0.847 * (0.180)	0.958 (0.277)	1.148 (0.307)
	Yes	1.366 *** (0.146)	1.140 * (0.169)	0.949 (0.146)
Sharing data, road users	Totally comfortable	1.630 *** (0.247)	1.731 *** (0.387)	1.819 *** (0.457)
	DK	0.989 (0.215)	1.032 (0.291)	0.932 (0.229)
	Fairly comfortable	1.046 (0.102)	1.197 *** (0.145)	1.252 *** (0.121)
	Not at all comfortable	0.693 *** (0.102)	0.552 *** (0.082)	0.546 *** (0.120)
	Not very comfortable	0.855 *** (0.070)	0.847 *** (0.102)	0.865 *** (0.079)
Sharing data, private companies	Totally comfortable	1.134 * (0.199)	1.742 *** (0.370)	1.939 *** (0.631)
	DK	1.031 (0.273)	0.837 * (0.175)	0.769 ** (0.201)
	Fairly comfortable	0.929 (0.105)	1.202 *** (0.128)	1.313 *** (0.144)
	Not at all comfortable	0.985 (0.127)	0.634 *** (0.107)	0.553 *** (0.119)
	Not very comfortable	0.934 (0.118)	0.900 ** (0.096)	0.923 * (0.100)
Sharing data, public authorities	Totally comfortable	2.094 *** (0.536)	1.673 *** (0.363)	1.298 *** (0.246)
	DK	0.672 *** (0.220)	0.691 *** (0.223)	0.858 (0.254)
	Fairly comfortable	1.314 *** (0.171)	1.369 *** (0.146)	1.176 *** (0.107)
	Not at all comfortable	0.603 *** (0.122)	0.645 *** (0.114)	0.774 ** (0.170)
	Not very comfortable	0.896 * (0.109)	0.978 (0.139)	0.987 (0.128)
	logLik	-31637.302	-30701.948	-27373.196
	AIC	63441	61570	54912
	BIC	64121	62250	55592
	n	26846	26768	26652

Notes: All effects are deviation effect coded (comparing coefficient to grand mean across variable). Standard errors clustered by countries.*** p < 0.001; ** p < 0.01; * p < 0.05

Figures

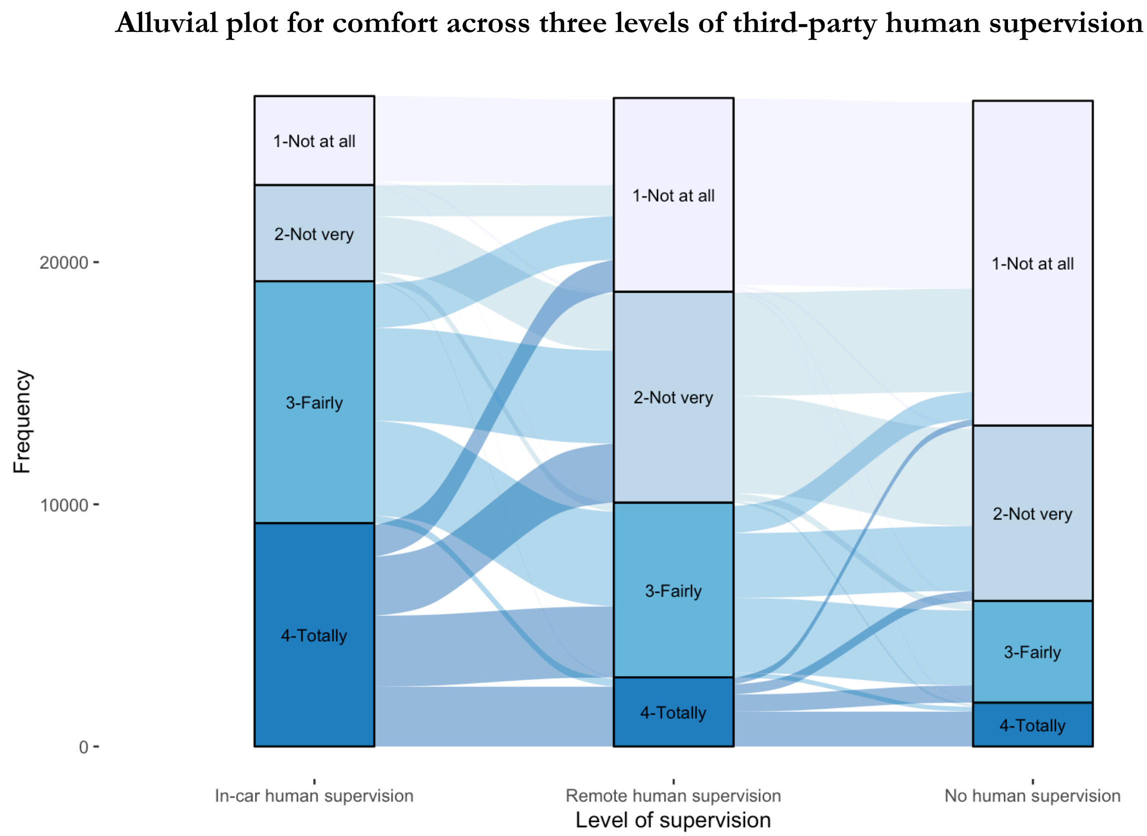


Figure 1. The flow of responses from the highest level of third-party human supervision (*In-car human supervision*), to the lowest level (*No human supervision*).

Changes in comfort with varying levels of third-party human supervision

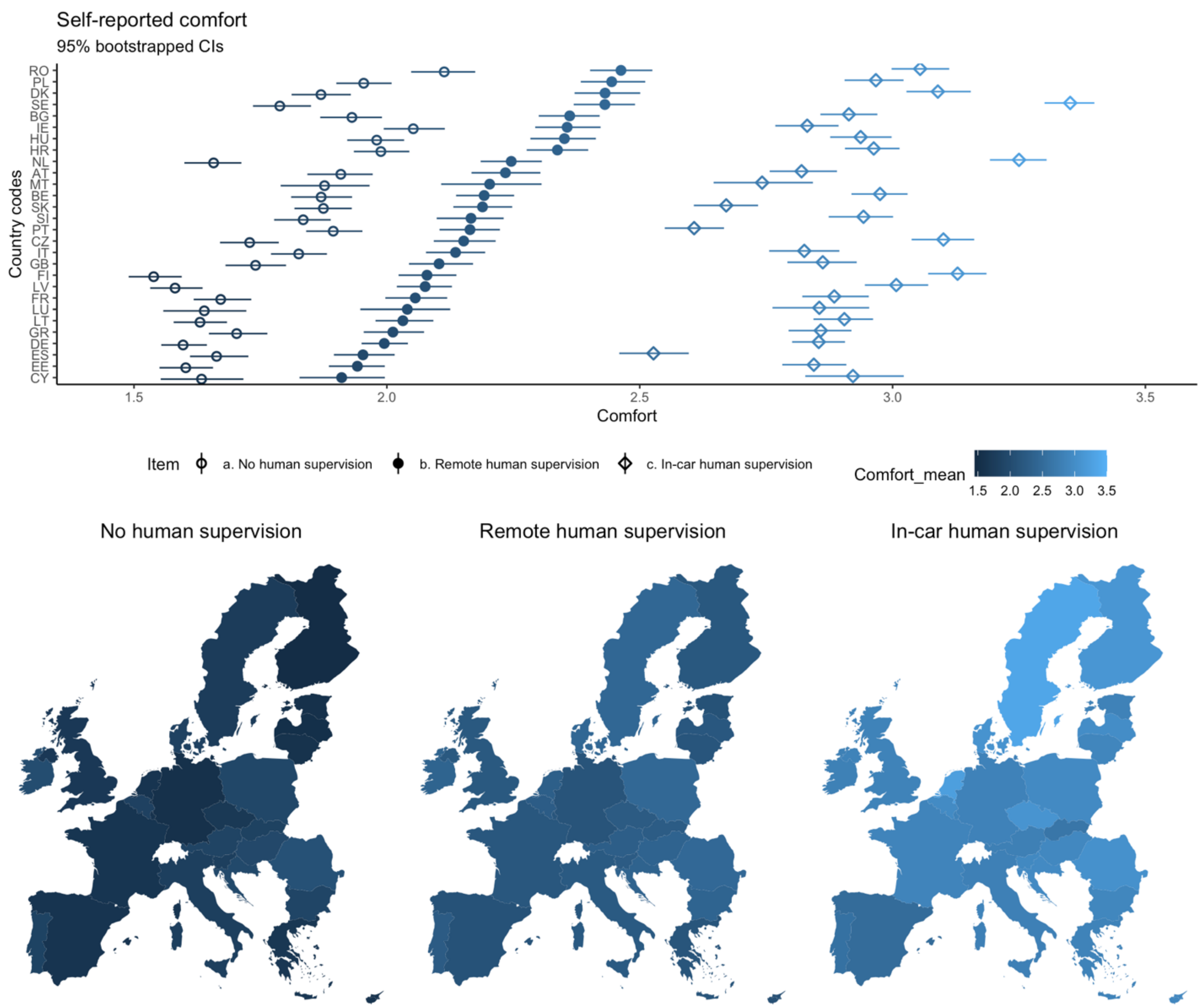


Figure 2. Country means and bootstrapped confidence intervals for each of our three dependent variables. Higher values on the x-axis indicate higher self-reported comfort. Dot plot ordered by "Remote human supervision". Mean comfort values are calculated after assigning a numerical value to the available answer options as follows: 1 – *Not at all comfortable*; 2 – *Not very comfortable*; 3 – *Fairly comfortable*; 4 – *Totally comfortable*.

Permutation variable importance of predictors with decreasing third-party human supervision

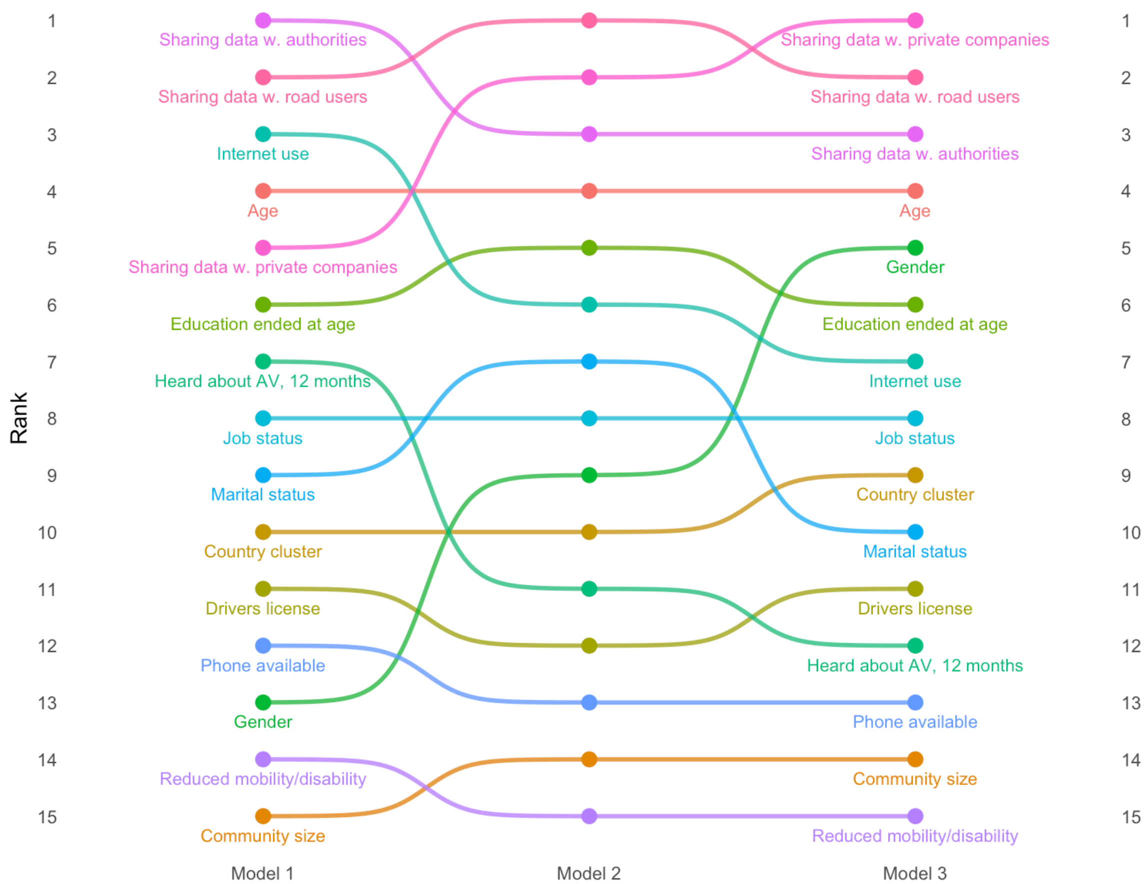


Figure 3.: Variables presented in descending order of variable importance. The dependent variable for Model 1 is the item "*In-car human supervision*". The dependent variable in Model 2 is "*Remote human supervision*", and in Model 3 it is "*No human supervision*".