TecnoCampus
Escola Superior
de Ciències Socials i de l'Empresa

Centre adscrit a la
upf. Universitat Pompeu Fabra Barcelona

Overconfidence and decision-making in Financial Markets

Student: Annick Torres Stienissen

Tutor: Dr. Mónica Oviedo and Dr. Giovanni Giusti

June 28th, 2020

FINAL THESIS OF THE DEGREE

Course: 2019 - 2020

Studies: Business and Innovation Management (BAIM)

ABSTRACT

This study aims to expand the knowledge of the effects of overconfidence on decision-making in financial markets. Particularly, on the effects of overconfidence levels in general knowledge question and financial markets, the effects of overconfidence and decision-making in asset markets, and the different possible decision mechanisms, such as group decision-making and feedback in individual decision-making, above any other, to possibly reduce the level of overconfidence. To reach this objective, a calibration test to measure levels of overconfidence, and a controlled laboratory experiment, where a double auction of asset markets, has been conducted. A correlation among levels of overconfidence in general knowledge questions and overconfidence levels in financial decision-making is expected, as well as a correlation among overconfidence and riskier financial decisions in asset markets. Additionally, it is expected to identify a reduction in overconfidence with feedback availability, and an increase in confidence and a decrease in accuracy when making decisions in group.

***Keywords**: Overconfidence; financial markets; decision-making; group judgement; feedback*

RESUMEN EJECUTIVO

Este estudio tiene como objetivo ampliar el conocimiento sobre el efecto del exceso de confianza en la toma de decisiones en los mercados financieros. En particular, sobre los efectos de los niveles de exceso de confianza en las preguntas de conocimiento general y los mercados financieros, los efectos del exceso de confianza y la toma de decisiones en los mercados de activos, y los diferentes mecanismos de decisión posibles, como la toma de decisiones grupales y la retroalimentación en la toma de decisiones individuales, por encima de cualquier otro, para posiblemente reducir el nivel de exceso de confianza. Para alcanzar este objetivo, se realizó una prueba de calibración para medir los niveles de exceso de confianza y un experimento de laboratorio controlado, donde se realizó una doble subasta de mercados de activos. Se espera una correlación entre los niveles de exceso de confianza en las preguntas de conocimiento general y los niveles de exceso de confianza en la toma de decisiones financieras, así como una correlación entre el exceso de confianza y las decisiones financieras más riesgosas en los mercados de activos. Además, se espera identificar una reducción en el exceso de confianza con la disponibilidad de retroalimentación, y aumentar la confianza y disminuir la precisión al tomar decisiones en grupo.

ACKNOWLEDGEMENTS

First of all, I want to express my gratitude to the coordinators of this final degree project, Dr. Mónica Oviedo and Dr. Giovanni Giusti, for their unconditional dedication, involvement and all the support they have provided to this work. Also, thank you for the trust offered and for respecting my suggestions and ideas.

Likewise, I thank all the TecnoCampus and ERASMUS students that have participated in this project, as well as, the TecnoCampus institution for the financing and for making this project possible.

Finally, thanks to my friends, with whom I shared all my thoughts and worries during the whole project, and because they have been a great moral and human support. But, above all, thanks to my parents, for their patience, understanding and solidarity with this project.

Table of content

List of Tables

List of Figures

1. INTRODUCTION

Bubbles and crashes in financial markets are not a phenomena that just came up recently in globally interconnected marketplaces or modern financial systems. They appeared very early in history, for example after the disintegration of the tulip price bubble in 1637 in the Netherlands, or the stock prices crash of the South Sea Company in the United Kingdom in 1720.

The impact of bubbles has increased with the today's financial markets' growing interconnectedness. What makes these phenomena so problematic is that bubble-and crash patterns in financial market prices are generally considered detrimental to the economic activity.

As an example of how problematic it can be for the economy, we take a look to the Great Depression in the 1930s which clearly demonstrated the danger of the effects of price bubbles in financial markets constitute for the economy since the real output and prices fell precipitously, the industrial production in the United States declined 47%, and the real GDP fell 30%. The wholesale price index declined 33%, and although there is some discussion about the reliability of the statistics, it is widely agreed that the unemployment rate surpassed 20% at its highest point. The seriousness of the Great Depression in the U.S. became clearer when it was compared to the Great Recession of 2007-09. During this recession the country's real GDP declined 4.3% and the unemployment rate increased less than 10%.

According to some literature, some characteristics of trading behaviour, are related with the occurrence of market bubbles. At the same time, behavioural finance researches have used this literature to support the claim that people are overconfident and that is the reason why people make errors when making decisions in financial markets (Barber and Odean 2001; Glaser, Nöth, and Weber 2008; Odean 1998). For that reason, it is very important to understand how people make decisions in financial markets and which factors can influence those decisions. Therefore, among different factors, this project focuses on the effects of overconfidence in asset markets.

Overconfidence is a decision bias by which we tend to overestimate our intuitive capacity to reason and make predictions. It leads us to believe that we have more knowledge or information than the rest to make decisions more accurately - known as certainty overconfidence, and in the case of having to make predictions, this bias leads us to think that the probability of failure is less than the real one, just as we give more probability of hitting than would correspond - known as the overconfidence prediction (Natalia Cassinello Plaza, 2016).

Overconfidence shows up in many areas of our daily life. It can easily affect aspects of our private life such as being too confident about the amount of food that we need to buy for the whole week, about an exam result by thinking that we performed better than we actually did, as well as, aspects of our working life, such as the risk of overconfidence at work which is an important factor in the generation of accidents since it can lead us to adopt an attitude of accommodation in the face of risk and non-perception of it. Of all possible behavioural biases that may affect decision making, overconfidence is one of the biases that is considered to have a more negative influence on decision making and this is the reason why it has become one of the main topics in financial decision-making.

There are different definitions of overconfidence. For example, the term overconfidence has been used to describe a situation in which people can be miscalibrated[1] and suffer from 'positive illusions' such as better than the average effect[2], illusion of control[3], and unrealistic optimism[4].

Nevertheless, different authors consider that these forms of overconfidence are related to each other. Moore and Healy (2008) differentiated three different kinds of overconfidence: over-placement, overestimation, and over-precision, and presented experimental evidence to demonstrate that these kinds of overconfidence can be dissociated. Other authors such as Griffin and Brenner (2008) classified overconfidence at a more specific level: motivational perspective and other perspectives such as: confirmatory bias, ecological probability, case-base judgement, and error in judgement.

Overconfidence can cause serious consequences because of its generality and importance which caused a huge influence on topics related to financial markets. Many researches, such as Barber and Odean (1999); Camerer and Lovallo (1999); Glaser and Weber (2007); Malmendier and Tate (2005), used overconfidence to explain events such as wars, stock market bubbles, strikes, entrepreneurial failures and litigation. The fact that overconfidence may lead to serious problems, led researchers in psychology to analyse different possible mechanisms to reduce individual overconfidence such as feedback availability and group decisions making.

Several authors that studied the effects of feedback on overconfident people found that people who were initially overconfident could learn to be better calibrated after making many decisions while receiving constant performance feedback (Slovic,

[1] Which means that they overestimate the probability that their decisions are correct (Alpert and Raiffa, 2013).
[2] i.e., people overestimate their abilities and achievements with respect to others (Svenson, 1981).
[3] i.e., people overestimate their control over situations, or events (Langer and Roth, 1975).
[4] i.e., they belief that bad events are more likely to happen to others than to themselves (Weinstein, 1982).

Fischhoff, and Lichtenstein 1982). However, although group judgement is another typical mechanism studied to reduce overconfidence, it ended up increasing confidence levels and decreasing the accuracy when making decisions (Russo and Schoemaker 1992).

Despite the extensive research on overconfidence, there are no studies aimed to demonstrate that there is a correlation between overconfident people in general knowledge questions and financial decision making and that, at the same time, studies possible mechanisms to lessen overconfidence, such as, group-decision and feedback. Therefore, it seems relevant to investigate whether overconfident people in general knowledge questions are also overconfident when making financial decisions, and if high individual levels of overconfidence can be reduced by implementing both possible mechanisms: feedback availability and group decisions making.

The topic of Overconfidence and Decision-Making in Financial Markets has been important during the last years of the degree. However, little has been taught about overconfidence and decision-making in financial markets and the different possible mechanisms to reduce overconfidence. At the beginning of the third year, decision-making in financial markets was introduced by two different professors through didactic lectures, and in the middle of the year behavioural biases was introduced by two other professors through experimental papers conducted in different fields.

However, the fact that the relationship between different financial ratios is complex, that there is not always a clear decision rule, and that overconfidence was not studied and correlated with financial markets, raised a key question: would overconfidence influence the way in which we make decisions in financial markets?

This question, in turn, raised another two key questions: can high levels of individual overconfidence be reduced by making decisions in groups? And, could a constant provided feedback reduce overconfidence levels in individual judgements (decision-making)? To answer these questions, a suitable methodology conducted in a laboratory experiment would be needed because it will allow us to generate data in a controlled way, with a full control over the environment and we will have the advantage to evaluate the internal validity of a specific scientific relation. Additionally, it is an investigation in which the variance of all or almost all possible independent variables is kept to a minimum, isolating research in a separate physical situation and handling one or more independent variables under controlled conditions since the objective of the experiment is to study the relationships between variables under “pure”, “uncontaminated” conditions, check the predictions that derive from a theory and refine theories and hypotheses (construction of theoretical systems).

Following the above-mentioned context, the purpose of this study was established. The study aims to expand the knowledge of overconfidence and decision-making in financial markets. Particularly, on the relationship between overconfidence levels in general knowledge questions and financial decisions, as well as, the effects of overconfidence in asset markets and the different possible mechanisms, such as group decision-making and constant feedback in individual decision-making above any other, to possibly reduce the level of overconfidence.

From and academic point of view, the topic is directly related to the degree. Overconfidence and Decision-Making in Financial Markets is related to Business and Innovation Management Degree on how people have been undertaking measures to reduce overconfidence levels when making decisions in financial markets. This study became a challenge that motivated me to keep looking for information to the point that it became equally interesting. Additionally, the practical application in the business field made it more enjoyable.

The present study will be structured as follows. In Section 2 we will talk about the related literature. Section 2 will be dedicated to explain overconfidence in psychological research in which we will address our main attention to explain the foundations of overconfidence, the reasons for overconfidence, the different forms of overconfidence: miscalibration and positive illusions (better than the average effect, illusion of control and unrealistic optimism) since those definitions were the most used in different papers, and the two main possible mechanisms to reduce individual overconfidence when making decisions: feedback availability and group-decision making. Section 3 will be dedicated to explain the conclusions and, Section 4 the objectives of the paper. In Section 5 we will reflect the hypothesis, and in Section 6 we will describe different ways in which overconfidence can be measured and the methodology of the experiment.

2. RELATED LITERATURE

2.1 Overconfidence in psychological research

Overconfidence has been studied in psychology since 1960s and it is usually defined as the frequent tendency to overestimate the abilities, success perspectives, the probability of positive results, or the precision of someone's own knowledge (Michailova and Schmidt 2016). Overconfidence in psychology is closely related to the probability judgement research and calibration[5]. It may occur when our confidences are related to our judgements, predictions or inferences are too high in comparison to the corresponding precision (Hoffrage 2016). Although overconfidence is common, it is not universal and it is typically eliminated and even reversed for easy questions (Brenner et al. 1996). This is called the difficulty effect, or the hard-easy effect, identified already by Fischhoff, Slovic, and Lichtenstein 1977, and it is hardly surprising because, according to Brenner et al. 1996, overconfidence is defined as the difference between mean confidence and overall accuracy.

However, the most important extension of the definition of overconfidence, generally applied by economists, are the studies of overconfidence in the positive illusion context, i.e. illusion of control, better-than-the-average effect and unrealistic optimism (Skała, 2008) and, although many economists tend to consider miscalibration and positive illusions together, we will discuss psychological studies on overconfidence separately.

2.1.1 Overconfidence as miscalibration

In psychology, overconfidence can be measured in different ways, miscalibration is one of them. Miscalibration is the tendency to overestimate the precision of our knowledge[6] (Lucy F. Ackert et al. 2010, p. 106). Calibration is usually measured with general knowledge questions that researches develop and, participants are asked to give an upper and lower limit such that they are 90% sure that the actual value will fall within the range specified[7]. Appropriate calibration happens when the proportion that is true is equal to the probability assigned, therefore miscalibration is the difference between the accuracy rate and probability assigned, to the event that a given answer is correct (Skała 2008). However, studies of calibration have shown that people's

[5] See for some example (Brenner et al. 1996; Dawes and Mulford 1996; Koriat et al. 1980)
[6] One can also be underconfident and miscalibrated, but the norm is overconfidence and miscalibration.
[7] See for example (Hilton et al. 2011; Russo and Schoemaker 1992)

confidence often exceeds their accuracy[8] (Brenner et al., 1996). Oskamp (1965), who studied overconfidence as a simply excess of confidence over accuracy, concluded that accuracy did not increase when more information was provided, but confidence increased steadily and significantly. However, the discussion about whether overconfidence is a constant feature (steady) or a dynamic process liable to manipulation has not given conclusive answers to date (Skała, 2008).

There are some studies that noted that being well-calibrated (being neither overconfident, nor under-confident) is a learnable and teachable skill, and by repetitiveness and continuance of tasks, the accuracy of predictions might increase when feedback is provided.

There are two possible ways in which we might be able to achieve a better subject's calibration. This might be possible by motivating subjects through outcome feedback and reward for their assessment to be more precise (Arkes et al. 1987).These two concepts will be further explained in section 3.1. There might be a third way in which we can achieve a better subject's calibration and, at same time, reduce overconfidence which is by assessing judgements (decision-making) in groups rather than individually. This third way to reduce overconfidence will further explained in section 3.2.

However, one of the most important concepts in overconfidence research, as mentioned in section 2.1, is the "hard-easy effect". This finding demonstrates that overconfidence appears mostly when difficult, or very difficult, tasks are undertaken, while undertaking easy tasks may result in under-confidence which is when the proportion of correct answers surpasses the expressed probability judgement (Skała 2008). This effect has been present in most calibration researches since then, and Lichtenstein, Fischoff, and Phillips 1982 strongly confirmed this hard-easy effect.

The following sub-section will be dedicated to outline the reasons for overconfidence which is one of the most important concepts in the miscalibration area.

2.1.1.1 Reasons for overconfidence

There have been many research papers on overconfidence, however, the roots of overconfidence and the reason of existence, have not been clearly identified. Researchers take the concept of overconfidence for granted, or they analyse the degree of overconfidence during their research (Skała 2008).

There are some psychological reasons for overconfidence that arise from the literature. Some authors such as Russo and Schoemaker, 1992, divided these reasons

[8] See for example (Keren 1991; McClelland and Bolger 1994)

into three main areas: cognitive, psychological and motivational. The first one is the *cognitive area*. In this area Russo and Schoemaker considered four main causes: the availability bias (when people have difficulties in visualizing all the possible ways that an event can occur), the anchoring bias (the tendency to anchor on one idea or value and not being able to adjust it adequately), the confirmation bias (when we make a forecast, or prediction, we tend to have preferences toward one point of view and, we naturally seek support for our initial view rather than looking for evidences), and the hindsight bias (which makes us believe that everything is more predictable than it really is). Among all of them, the confirmation bias has been widely explored in the literature and it has been considered as a strong support in the emergence of overconfidence. Thus, this implies that miscalibration is not just a cognitive mistake when translating the confidence judgment into a numerical probability (Skała 2008).

The second area is the *psychological* one. Russo and Schoemaker, defined overconfidence as a distortion of judgment. It is often considered as a pure mental phenomena. They consider that euphoria may cause this overconfidence effect, however, they also consider that drugs like alcohol and cocaine can produce overconfidence. The third and last area is the *motivational* one, Russo and Schoemaker consider that another cause for overconfidence is related to our need to believe in our abilities. This need provokes in people a distorting reality, however their optimism has a huge motivational value.

Other authors such as Mishra and Metilda 2015, considered 'Self-attribution' bias has a cognitive and motivational component that is strongly associated with overconfidence, because it causes individuals to learn how to be overconfident rather than being more accurate with their self-assessment. This bias is when people relate success with aspects such as talent and prevision, and failure with situational factors. An example would be students. Students tend to attribute their own intelligence and hard work when having higher grades, and attribute 'unfair grading' when having low grades.

Going further, the *hard-easy effect* is also included as a reason for overconfidence. This effect is most of the times a standard finding in most research papers, however, some authors claim that difficult questions are the reason why overconfidence might arise and, without those difficult questions no miscalibration is observed (Skała 2008).

Some researches associate different overconfidence levels with *gender issues*. There is a common belief that men are more confident than women with the same level of knowledge. There is a small but growing literature showing evidences that men are more confident that women (Bengtsson, Persson, and Willenhag 2005). For example,

Barber and Odean 2001, used gender as a proxy for overconfidence in financial tasks, and they found that men were inclined to feel more competent than women in this field since they documented that men traded 45% more than women, but also performed worse than women, presumably due to a higher confidence in their own abilities[9] so, indeed, gender differences were confirmed. However, there are other studies such as Lundeberg, M. A., Fox, P. W., & Punćcohaŕ 1994, that did not find any difference in single item assessment (calibration) of confidence between both genders.

The last reason for overconfidence, presented by Gigerenzer, Hoffrage, and Kleinbolting 1991, is related to the *faulty procedures*, by researchers, when studying overconfidence. Gigerenzer et al. 1991, considered that the main reason for overconfidence was not a cognitive or motivational factor but, a biased structure of the task and its relation with the environment.

2.1.2 Other forms in which overconfidence can be manifested

Although psychological research mainly concentrates on measuring overconfidence through miscalibration, there are other ways in which overconfidence can be manifested such as the better-than-the-average effect, the illusion of control and the excessive optimism (Lucy F. Ackert, Richard Deaves 2010).

2.1.2.1 Better-than-the-average effect

This better than the average effect has been the main topic of research and discussion in the psychological area since 34 years ago. As it has been often the case, many researches have focused on two related task: (1) identifying possible variables that could moderate the better than the average effect and (2) generating explanations for why it occurs (Brown 2012). Regarding the first issue, researches have shown that the better than the average effect is reduced when the characteristics that are being compared are unambiguous or uncontrollable[10]. This is mainly observed when difficult skills are being evaluated (Kruger 1999), when people believe that they have to justify their affirmation to an audience (Sedikides et al. 2002), when comparisons are made directly (Otten and Van Der Pligt 1996), and when a comparison among people involves the comparison with a specific individual rather than an aggregated individual that represents the majority of people (Klar and Giladi 1997).

Regarding the explanations, the better than the average effect was originally thought to be motivated by self-improvement needs. People value themselves more

[9] See also (Correll 2001)
[10] See for example (Alicke 1985; Dunning, Meyerowitz, and Holzberg 1989)

positively than they value others because, it makes them feel good about themselves when believing that they are above the average (Brown 2012). However, other researches have described different cognitive mechanisms that could also produce the effect[11] without any motivated need. For example, it has been proposed that focalism (i.e. the tendency to focus on oneself when comparing to others), egocentrism (i.e. to overvalue one's own perspective), informational differences (i.e. the tendency to know more about ourselves than about others), and the naïve realism (i.e. the tendency to take for granted that one's own view of the world is the reflection of the world as it actually is) are different ways in which the BTA effect can be manifested (Brown 2012).

These studies of the BTA effect have helped other researches to understand the specific processes that form comparative judgements and, at the same time, it has helped readers to see that the better than the average effect is not just focused on motivational factors (Brown 2012). However, other researches consider that the motivational and cognitive mechanisms are behind the better-than-average effect. On the motivational side, thinking that you are better than average improves your self-esteem. On the cognitive side, the performance criteria that most easily comes to mind are often those that you are best at (Lucy F. Ackert, Richard Deaves 2010).

The better than the average effect, is measured by comparing a person's performance with others' performances. For example, a person's actual position in a test score distribution (percentile) is compared with the person's estimation of the position she/he estimated (Olsson 2014).

A very typical example to explain the better than the average effect is the driving example. In 2016, a survey conducted by the KPMG and ICADE Business School to 69 investors in Spain, 39% of the sample considered that they were driving above the average, and although in some cases it might be true, it was not the same for all of them. Therefore, this academic study showed that being considered better than the average was related to overconfidence and, therefore, these people were susceptible to this bias.

People always have in their minds the definition that will make them look best. However, based on the previous driving example some might see 'best' as being an expert at steering; others might see 'best' as the most competent at anticipating danger; while others might still see 'best' as the one who is most skilful at managing the steering wheel while speeding down the highway (Lucy F. Ackert, Richard Deaves 2010).

[11] See for example (Chambers and Windschitl 2004; Giladi and Klar 2002; Krizan and Suls 2008; Krizan and Windschitl 2007)

2.1.2.2 Illusion of Control

Another way in which overconfidence can be manifested is the so-called illusion of control. This happens when people think that they have more control over a situation than they actually have which causes an inaccurate estimation of risks (Langer 1975; Schwenk 1984). The higher the perception of control, the higher the chance to underestimate the risk (Schwenk 1986). In other words, a student's positive misconception of control will lead her/him to overestimate the ratio of success for an exam. Illusion of control is the tendency to underestimate positive information and overestimate negative information which normally leads into limited forecasts that will probably reduce the odds of overestimation[12] (Durand 2003).

Nonetheless, Illusion of control has been since 45 years ago related with two different concepts such as skills and luck. Langer 1975, claims that, in principle, the distinction between skill and luck seems clear. In skill situations there is a link between behaviour and outcome, thus success in skill situations is controllable. Luck, on the contrary, happens in a fortuitous way and the success of it is, apparently, uncontrollable (Langer 1975). However, the distinction between both concepts seems not to be generally recognized since many people behaves as if chance events were controllable, and a number of different researches have provided support for the position that people assume that there exists a relationship between skill orientations with chance situations (Langer 1975).

According to Skała 2008, a true fact is that in every kind of situation participants are more likely to express excessive confidence in their control over results that are chance-driven tasks. Presson and Benassi 1996, in a meta-analysis study, documented the dominance of illusion of control effects in different range of studies and experimental variations. The situational variations that increase the illusion of control included choice, outcome sequence, familiarity with the task, information regarding the outcome of the task, and the active participation in the task. However, some authors use these situational variations when doing researches but, most of the time they use proxies instead of situational variations. This proxies use the participant's willingness to trade lottery tickets, their judgement of contingency, their judgement on their prediction ability, or even their confidence on succeeding a task. This last proxy makes illusion of control studies closer to overconfidence (Skała 2008).

Illusion of control is measured by comparing a person's performance with respect to the person's belief of own performance. For example, the number of correct answers

[12] See for example (Langer 1975; Schweitzer and Cachon 2000)

a person does in a test is compared to the amount of correct answers that person though she/he would do (Olsson 2014).

To conclude, a very common example of this illusion of control effect is the dice example. People tend to insist on throwing a dice personally as if the outcome could then be more favourable. Additionally, if people expect a certain outcome and these outcome actually occurs, the participant will be more liable to assign this outcome as his/her skill rather than luck, and he/she will re-affirm his/her belief in control over a situation where the only factor is probability (Skała 2008).

2.1.2.3 Unrealistic optimism – Optimism bias

A third way in which overconfidence may manifest itself is the so-called unrealistic optimism, or optimistic bias, which is frequently analysed in the context of biased self-attribution or in the context of the BTA effect[13] (Skała 2008).

The optimism bias is the difference between an expectation and the real outcome. If the expectations that we have are better than the real outcome then we call this optimistic bias, the other way around, if the reality is better than the expectations we had, then we call this pessimistic bias. The scope of the optimism bias is thus empirically measured by recording the expectations of individuals before an event happens and then, contrasting those expectations with the resulted outcomes (Sharot 2011).

In general, optimistic bias can be seen as an error when evaluating future events, either in the sense of absolute terms or in the sense of the BTA effect. Actually, several research papers suggest that people tend to be unrealistically optimistic with future events. Researches concerning crime, disease (Harris and Guten 1979) and automobile accidents (Robertson 1977) found many people who were saying that their risk was less than the average but very few people who said that their risk was greater than the average. Also, when people are asked to predict future political and social events they tend to predict according to their preferences (McGuire 1960; Sedikides et al. 2002). Even for chance events such as picking up a card out of a deck, people show optimistic biases (Irwin 1953).

In past research, optimistic biases were generally related to wishful thinking or defensiveness. However, in recent years, explanations related to the capabilities of human being to handle information became a new perspective to explain why people might be unrealistically optimistic. The lack of certain information to make accurate risk assessments or to judge future probabilities that might introduce systematic errors arises

[13] BTA effect; Better than the average effect

unrealistic optimism (Weinstein 1980). Also, when comparing risk evaluations, the optimistic bias may arise because people have difficulties to adopt someone else's point of view (Ross and Sicoly 1979). However, many factors that make us feel that something is going to happen, or not, may also make other people feel the same way. If people just take into account their own circumstances and think that their chances are different compared to others, then they will conclude incorrectly. Any factor that influences people's own chances, could also affect comparative judgements and, one of the factors that will affect this belief is the perceived probability of the event (Weinstein 1980).

A last factor that provokes the optimistic bias is the past personal experience (Lichtenstein et al. 1978). Personal experience should make it easier to have a better perception of the probability through the mechanism of 'availability' (Tversky and Kahneman 1973, 1974).

However, it is important to mention that being unrealistically optimistic has huge consequences such as time and costs, disappointment to not meet established goals, the loss of self-esteem, the reduction of social regard, and the inability to strive for other future goals that are within his/her grasp (Lucy F. Ackert, Richard Deaves 2010).

Optimism bias is frequently measured through two risk determinants: absolute risk and comparative risk. Absolute risk (assessment of own) is when participants are asked to estimate their probability of experiencing a negative event in comparison to the real probability they have of experiencing that negative event. Comparative risk (self *versus* others), is when individuals estimate the probability they have of experiencing a negative event compared to others of the same age and sex (Lipkus et al. 2000).

To conclude, a very common example is that students expect to get higher grades than they actually do get, and they overestimate the number of jobs offers that they will receive after finishing their studies. Further, although many studies show high rates of divorce, almost all newlyweds expect their marriages to last forever. Last, the construction sector is also an example, for instance, the Sydney opera house was supposed to be finished in 1963 with a cost of $7 million. However, it was completed 10 years later at a cost of $102 million (Lucy F. Ackert, Richard Deaves 2010).

2.2 Overconfidence in Finance

Economists started to add psychological findings into economic models in the 70s, and during the 90s a rapid development of this practice began. From then until now, overconfidence has been a topic of interest for economists, with a main focus on the context of behaviour in financial markets. Overconfidence is generally defined as an overestimation of one's own knowledge, or the precision of the private information. However, it is also considered as an underestimation of the variation of signals or the volatility of asset values (Skała 2008).

Some issues found in financial markets such as miss-valuations, the disposition effect and the excessive trading volumes, were properly treated once the investor's overconfidence was assumed. Previously, with the standard economic theory, this issues could not be solved. The strong presence of overconfidence in markets and its long-term perseverance raised a continuous discussion about the idea of efficient markets and the rationality of economic agents. Overconfidence has been continuously demonstrated, in financial markets, through methods that go from experimental studies and questionnaires to financial market data and formal models, despite some doubtful thoughts among economists about the existence and effect of overconfidence (Skała 2008).

As seen in the previous psychological literature, overconfidence manifests itself in decision making under uncertainty, in both extremes, miscalibration and other forms of overconfidence (better than the average effect, positive illusions and unrealistic optimism). Actually, according to several studies (e.g. Hribar and Yang 2016) the research in finance also uses both extremes, implicitly or explicitly, to estimate predictions about the effect of individual overconfidence in economic decisions.

Additionally, the literature reviewed suggests that an overconfident person tends to under-estimate the risks that a financial decision may imply and, at the same time, they tend to over-estimate the expected earnings which, in general, leads people to purchase excessively, incurring high transaction costs that reduce profitability and lack of different investment portfolios (Michel M. Pompian 2012). Several studies claimed that the effects of overconfidence are well-known in the literature and that helps to explain anomalies such as under- and overreaction (Daniel, Hirshleifer, and Subrahmanyam 1998), excess price volatility, asset bubbles (Scheinkman and Xiong 2003), the forward premium puzzle (Burnside et al. 2011) and the excessive trading (Odean 1998). Many authors have addressed the study of overconfidence - in financial areas - in different ways. However, the majority of authors have related overconfidence to the profitability of

trading. Biais et al.(2005) and Grinblatt and Keloharju (2009) confirmed that overconfidence had negative effects on the profitability of trading, which is normally due to the fact that traders might overestimate their precision of private information.

As previously mentioned, overconfidence in financial areas is also measured through experiments like overconfidence in psychological researches. Some authors (e.g. Menkhoff, Schmidt, and Brozynski 2006) tried to perceive the impact of experience on overconfidence, risk taking and group behaviour by conducting surveys but, this method did not provide knowledge on whether certain levels of overconfidence led to a specific market outcome (e.g. average price or trade volume). Other authors (e.g.Kirchler and Maciejovsky 2002) conducted a study on development of overconfidence, using multi-period experimental market methods where the main result showed that subjects were well-calibrated in some periods and under/overconfident in other periods.

Last, a study conducted by Ackert et al. (2009) to test whether overconfidence was leading to an increase in trading activity and that such (trading activity) was gender influenced through differences in overconfidence, used a 12 single period method per experimental session. The main findings reported that the higher the overconfidence, the higher the trading volume and the lower the earnings. However, there was no evidence that overconfidence and trading activity was affected by gender.

2.3 Possible Mechanisms to reduce overconfidence

Overconfidence can have some serious consequences. Researches have used overconfidence as an explanation for, strikes, litigation, entrepreneurial failures, wars, and stock market bubbles (Moore and Healy 2008). Some consequences raised from the literature are the following.

On the one side, when talking about overconfidence in psychological research, we saw that considering yourself as being better than the average, thinking that you have control over some situations that in reality you do not have - which ends up in an inaccurate estimation of risks - and to be unrealistically optimistic - which can result in a loss of time and high costs, disappointment because of not meeting established goals, the loss of self-esteem, reduction of social regard and the inability to strive for other future goals that are within your grasp – leads you to be overconfident and miscalibrate.

On the other side, when applying the psychological concept of overconfidence into financial decision making, being overconfident resulted in situations in which we under or overreact to decisions such as underestimating the risks that financial decisions may imply, or overestimating the expected earnings which leads to purchase excessively, incurring high transaction costs that reduce profitability and lack of different investment portfolios. Additionally, being overconfident in the finance sector is related to the excess price volatility, asset bubbles, the forward premium puzzle and excessive trading.

In order to reduce overconfidence in decision making the literature suggests different mechanisms to reduce overconfidence such as: including self-evaluations, training, choice prompt inclusion, foil plausibility, lower-familiarity options, providing monetary incentives, feedback availability, and decision-making in groups among others. However, we will mainly focus on feedback availability and group decision making as two possible mechanisms to reduce overconfidence. Both concepts will be further explained in section 2.3.1 and 2.3.2 respectively.

2.3.1 Overconfidence and feedback availability

According to Ayton and McClelland 1997, the most favoured explanation for overconfidence is that it occurs due to a mapping error, when mapping true feelings of trust according to a corresponding response scale. Thus, this response error and overconfidence could be reduced with feedback, just if the error is correctable, rather than not correctable[14] (Bolger and Önkal-Atay 2004).

[14] For instance, ‘cognitive noise’ which is not correctable.

A number of theorists have identified different types of feedback. The first one is the *outcome feedback*, which is the so-called 'knowledge of results' such as the true value of the variable that was forecasted. A second type of feedback is the *cognitive feedback*, which is the relationship between the outcome and your own prediction, or the relationship between the outcome and the characteristics of the series (i.e. variability or trend) (Todd and Hammond 1965).

Balzer, Doherty, and O'Connor 1989, went on to differentiate three components of *cognitive feedback* in probability learning*:* cognitive information (CI) which is the link between someone's judgement and the cues (i.e. level and variability of judgement), functional validity information (FVI), which details the relationship between a judge's cognitive strategy and the task (i.e. the correlation between judged and actual values) and task information (TI) which shows the relationship between the cues and the criteria.

Benson and Önkal 1992, on the contrary, investigated two types of *performance feedback*. The first one is the *scoring-rule feedback*, which penalizes or prizes the forecaster depending on the predicted forecast and the real outcome/result. The scoring-rule feedback is basically the score a forecaster for a set of probability forecasts. The second one is the *calibration feedback* which provides information to forecaster about the ability they have to assign appropriate probabilities to outcomes. Contrary to scoring-rule feedback, calibration feedback has not been standardized. It can consist of graphical displays or numerical summaries of the forecasted probabilities, the correct proportion associated to each forecasted probability value, and the number of evaluation of each value. However, calibration has been proved to improve the performance of forecasters when forecasting probabilities are provided.

Nonetheless, *performance feedback* (which at the end is the information about how accurately forecasters predict and the type of feedback that we will use), is one of the four types of feedback that exist in judgemental forecasting tasks, the three other types of feedback are: *outcome feedback* (information about the realization of a previously predicted event), *the process feedback* (is the information received about the cognitive process of the forecaster), and the *environmental feedback,* also called: 'task feedback' (which is the information that researches provide about the event that is going to be predicted) (Benson and Önkal 1992)

Several authors that studied the effects of feedback on overconfident people found that people who were initially overconfident could *learn to be better calibrated* (reduce overconfidence) after making 200 decisions and receiving constant performance feedback (Slovic et al. 1982). Other authors found that *overconfidence could be directly eliminated* by proving performance feedback to participants during the experiment (Plous

1995), and that overconfidence could improve performance - e.g. by maximizing payoffs - (Berlin and Dargnies 2016).

Practical studies have also been conducted to study whether feedback reduces overconfidence levels. In one particular study performed by Koriat, Lichtenstein, and Fischhoff (1980), participants were asked to answer two sets of general knowledge questions. First, the question was conducted under controlled instructions and then under reason instructions[15]. The results suggested that participants were overconfident when given control instructions but became well calibrated after analysing the pros and cons of the reason instructions.

Other empirical studies suggested that feedback had a significant effect on overconfidence, when questions provided were 'consistently hard' (the hard-easy effect), and that there was no significant effect on confidence, accuracy or overconfidence when questions were categorized as 'easy questions' (Pulford and Colman 1997). The reason why this may happen is because there is more social pressure to reduce confidence when hard tasks have to be done, rather than increasing confidence for medium, or easy, tasks where people are not so confident and social pressure is lower.

However, it is important to take into account that feedback is not useful to reduce overconfidence in every situation. For example, in studies of probability where they use general-knowledge questions, researches have failed when trying to reduce overconfidence with feedback. These tests usually require judgments for unrelated events (see, for example, Keren, 1991) and responses that are deliberately selected to be counter-intuitive (see, for example, Gigerenzer, Hoffrage and Kleinbolting, 1991); none of these conditions is ideal for learning. On the contrary, a forecasting task offers better opportunities to learn from feedback and also from training, as probability judgements have to be continuously done for events that were not selected to be misleading (Bolger and Önkal-Atay 2004).

For that reason, it seems possible and relevant to investigate whether providing performance feedback (one of the four types of feedback that exist in judgemental forecasting tasks) during a forecasting activity is a useful mechanism to reduce individual overconfidence.

[15] Reason intructions; why the chosen answer has been done? Analysis of the pros and cons of each step / decision - Koriat, Lichtenstein, and Fischhoff (1980)

2.3.2. Group decision-making and overconfidence

There is a limited economic literature on group decision-making and the existing one is always divided into studies of decision-making, or preferences in non-strategic settings and the involvement of strategic games (Cheung and Palan 2012).

However, we all know that generally speaking, when making decisions in group, people tend to follow a set of six general processes: set-up objectives, search for alternatives, compare and evaluate alternatives, choose, implement the decisions, and follow-up and control (Harrison 1996).

Group decision-making starts when objectives are settled down, and the aim of the process is to achieve those established objectives together. The second step, is to generate alternatives, which implies to look for external and internal relevant information and then design different alternatives that can help to achieve the final objective. The third step, is to compare and evaluate, among the group members, all chosen alternatives based on the perceived uncertainty and the preferences of the decision makers to achieve the objective (Harrison 1996). The fourth step, is the act of choice. It is the moment when decision makers choose a set of alternatives that will be followed to achieve the objective.

The two lasts steps of the group decision-making process are: implementing the decision, and the follow-up and control. The implementation of the decision is when the abstract decision turns into an operational reality and, finally, we have the follow-up control which is ensuring that the implemented decisions have the estimated outcome and that objectives are achieved (Harrison 1996).

Although few papers have been dedicated to study group decision-making, there are some authors that compared small groups and individual behaviours in different aspects of our daily life, and studied in which situations making decisions in groups can reduce risks. However, to the question *'if groups do better than individuals when sizing up uncertainty in financial decision-making'*, the answer is mixed. Group judgements can be better than individual ones, because in groups we are forced to acknowledge that other people see the world differently than we do, we take lower risks, learn faster, play more strategically, and we provide more normative and informational effects on the other person's judgements. However, group decision-making can also increase levels of confidence and decrease accuracy on decision-making.

Some evidences of the above mentioned effects of making decisions in groups are the following.

Taking lower risks. A very famous example to explain the difference between groups and individuals' behaviour is the lottery-choice experiment by Baker, Laury, and

Williams (2011). The aim of the experiment was to compare risk preferences from three people's behaviour (group) versus isolated individuals. The main results pointed out that groups tend to choose lower risk lotteries than the average choice of individual group members. Other studies (e.g. Rockenbach, Sadrieh, and Mathauschek 2007) investigated the difference between team decisions and individual decisions following different models such as *the Expected Utility Theory (EUT) and the Portfolio Selection Theory (PST).* The main results suggested that they did not find any evidence pointing out that there was a difference between team decisions and individual decisions with the principles of EUT. However, they found a substantial evidence that teams were taking lower risks than individuals in their decisions by using the PST.

Learn faster and play more strategically. Team decision making can influence other factors such as the ability to learn faster and to play in a more strategical way. A beauty contest experiment conducted by Kocher and Sutter (2006) shows that, as the game was repeated, teams were learning faster than individuals, and payoffs were higher under low time pressure rather than under high time pressure. Another experiment, based on a signalling game by Cooper and Kagel (2005), also compared individuals with two-person teams and found out that teams tend to play more strategically than individuals and generate positive synergies in more complicated games.

Perspective thinking. Russo and Schoemaker 1992, tested group overconfidence through a simple experiment with eighty-three managers. On average groups were doing better than individuals. In the worst scenario, they forced a compromise, and in the best scenario, they encourage open-mindedness (perspective thinking). However, individually, people were still anchoring too strong to their initial view and returned to it when they had a chance. This persistence can be reduced with the so called 'Delphi techniques' and other procedures, in which you share opinions, when the analysis / experiment is conducted in a network PC environment.

Normative and informational effects. Further, other authors studied the influence of group interaction on individual members' behavioural judgements. The results suggested that groups *provided' normative social influence'* (a process of social comparison of some kind) *and 'informational effects'* (it is when relevant arguments are exchanged) on the individual members' judgements in several ways (T. M. Ostrom and H. S. Upshaw 1968).

Confidence increases. According to Heath and Gonzalez (1995) asking other people's opinion is not likely to reduce very high levels of overconfidence. In fact, they hypothesized that interacting in groups just invites people to create explanations for their

own beliefs, which at the end just provokes that confidence levels in someone's own predications increase. They performed several experiments which proved this effect. These experiments were based on making predictions and stating their confidence level of the selected choices in the sports area. The results of the experiments, ended up as predicted, confidence was increasing when interacting in groups. Therefore, this results supports the fact that interacting in groups can be dangerous since it ends up increasing confidence. Other authors such as Bang et al. (2014) also described and evaluated the process through groups by combining their individual judgement and their determined confidence level into a group judgement with different levels of confidence, and the results suggested that groups had the tendency to be more confident when making decisions than individuals.

Decreases accuracy. In the same experiment were Heath and Gonzalez (1995), found that group decision making increases confidence levels, they also found out that 98% of participants though that the accuracy of their group was above the median. This result is to be expected if the confidence level of the group improves more than the accuracy. This results was also proved by Sniezek (1992). Sniezek (1992) suggests that single person forecasts should be less accurate than group's forecasts, because group members believe that their forecast discussions are more accurate. The theory that the participant implicitly thinks may be that a single person perform less work than a group and, therefore, the result of this difference in performance should result into a better product. Additionally, Boje and Murnighan (1982) proved that the confidence of the group's members increased when they did many trials, but their accuracy decreased.

To end up, we can summarize the literature in this following way: (1) in some cases the increase in confidence can be justified, because groups may learn faster, take lower risk and play more strategically than individuals which improves performance; (2) groups generally tend to have higher confidence levels than individuals; (3) group decision making can influence overconfidence but also the accuracy of answers; (4) group decision making can also open mind-sets and have different perspective thinking; (5) group decision making can provide normative social influence and informational effects.

For that reason, since overconfidence was considered as a negative / risky impact on decision-making in general but, specifically, in financial matters, and group judgment (team decision-making) is proved to have effects on overconfidence levels, it seems relevant to study whether group judgement could be a possible mechanism to reduce overconfidence in asset markets by following a laboratory experiment.

2.4 Conclusions

The literature reviewed has shown different perspectives and points of views of overconfidence. Nevertheless, those different approaches were not considered as opposite conclusions, but just different ways in which overconfidence could be studied. Additionally, overconfidence can be seen in many areas, it can be manifested in different ways, and there are also many reasons why overconfidence may appear. However, generally speaking, it can be concluded that overconfidence has almost never a positive effect on individuals' decision making, and in the financial sector this negative effect is reflected in negative, or risky, financial decision-making, which led us to the first hypothesis: *'there is a correlation between overconfidence in general knowledge questions and overconfidence in financial decisions'.*

Following the literature, the forms in which overconfidence can be manifested, in the psychological field, are also ways in which overconfidence can be manifested in the financial area. Additionally, the literature suggests that overconfident people tend to under-estimate and over-estimate their risks and earnings when making financial decisions. This under/over estimations end up in excessive trading, asset bubbles, excessive price volatility and premium puzzles, which led us to the second hypothesis: '*There is a correlation between overconfidence and riskier financial decisions in asset markets'.*

However, it seems that overconfidence bias can be reduced by constant feedback on individual judgment. According to the literature, subjects can be trained to be better calibrated and this would include receiving constant feedback, as well as, repeating tasks several times. Additionally, other ways in which overconfidence may decrease are when providing reason instructions. However, according to the literature, overconfidence can just be reduced when tasks are considered to be 'hard tasks', and since financial decision making is considered to be a hard task, we can apply these mechanisms to our experiment to identify whether providing feedback decreases, or not, overconfidence levels. This part of the literature led us to the third hypothesis: '*Overconfidence is reduced when individual feedback is provided'.*

On the contrary, group judgement has not been useful when trying to reduce confidence levels since it rather reduced accuracy and increased confidence levels. This suggestion was only proved in non-financial areas. Therefore, these mechanisms can be tested to try to reduce overconfidence when making financial decisions. It is also possible to apply this mechanism because, again, financial decisions are considered to be 'hard tasks'.

In summary, it has been possible to determine whether there is a correlation between overconfidence and decision making. However a deep analysis of how overconfidence influences decision-making in asset markets, how overconfidence levels in general knowledge questions and financial decisions are correlated and, at the same time, studying the possibility to reduce overconfidence by using two different possible mechanisms such as group judgements and feedback availability in financial markets has not been studied yet, which led us to identify the main objectives of the project, explained in the following section.

3. OBJECTIVES

Following the conclusions, our proposal is to use a calibration test and a laboratory experiment to demonstrate that there is a correlation between being overconfident in decision making at a psychological and financial level, and a correlation between being overconfident and making riskier financial decisions, as well as, to provide evidence on the possibility to reduce overconfidence through the two previously mentioned possible mechanisms: feedback availability and group judgment. Therefore, **this study aims to fill in the gap in the literature by investigating whether overconfidence biases are influential in financial decision making, and whether overconfidence can be reduced using different possible mechanisms: feedback availability and group judgement.**

This research is relevant to distinguish judgemental overconfidence from positive illusions, review relevant work in psychology and behavioural finance, as well as, to measure the effect of overconfidence at group level compared to overconfidence at individual level, and present evidence of the effectiveness of two possible alternative mechanisms to reduce overconfidence.

Additionally, this research can contribute in two other ways. First, our research can be relevant to understand whether overconfidence may appear. Second, it can be relevant as well to understand *which are the negative effects of being overconfident*. Overconfident people are more likely to make erroneous financial decisions which, at the end, can end up in infinite debts, business to fail, bankruptcy and the loss of assets among others.

To sum up, the specific research objectives are **to determine whether subjects who are overconfident in psychological areas are also overconfident in financial decision making; to determine whether overconfident subjects make riskier financial decisions; and to test whether individual overconfidence can be reduced by providing feedback and group judgement.**

4. HYPOTHESIS

As we have seen, overconfidence can appear in many areas such as the personal and financial area, and it can be manifested in different ways such as: miscalibration, better than the average effect, positive illusions and unrealistic optimism. There are also many reasons why overconfidence may appear, we mentioned some reasons such as: cognitive, motivational and psychological reasons, also the hard-easy effect, the faulty procedures and gender. Thus, a first hypothesis was postulated. This hypothesis tries to answer whether there is a correlation of overconfidence levels when subjects make decisions at a general knowledge level and a financial level. The first hypothesis is as follows:

Hypothesis 1: There is a correlation between overconfidence in general knowledge questions and overconfidence in financial decisions.

According to the psychological literature, miscalibration and the other forms of overconfidence are ways in which overconfidence is manifested, and so happens in the financial area. Finance researchers also use miscalibration and other forms of overconfidence to estimate predictions about the effect of individual overconfidence in economic decisions. Additionally, the literature reviewed suggests that an overconfident person tends to under-estimate the risks that a financial decision may imply and, at the same time, they tend to over-estimate the expected earnings which, in general, leads people to purchase excessively, incurring high transaction costs that reduce profitability and lack of different investment portfolios (Michel M. Pompian 2012). This under- and over-estimations can end up in excessive trading (Odean 1998), asset bubbles (Scheinkman and Xiong 2003), excess price volatility and the forward premium puzzle (Burnside et al. 2011). Thus, a second hypothesis was postulated. This hypothesis tries to answer whether there exists a correlation between over-confidence and riskier financial decisions in asset markets. The second hypothesis is as follows:

Hypothesis 2 (H_2): There is a correlation between overconfidence and riskier financial decisions in asset markets.

The third hypothesis is based on the literature reviewed of overconfidence and individual feedback availability. As previously mentioned subjects can be trained to be better calibrated by constantly receiving feedback and repeating tasks several times. Other studies suggested that providing controlled instructions increased overconfidence on participants while providing reason instructions was calibrating overconfidence.

Additionally, according to the literature, feedback can only be used as a mechanism to reduce overconfidence when tasks are considered has 'hard or difficult' ones. Thus, a third hypothesis was postulated. This hypothesis tries to answer whether individual feedback can be useful to reduce overconfidence and guide participants to make correct decisions. The third hypothesis is as follows:

Hypothesis 3 (H_3): Overconfidence is reduced when individual feedback is provided.

The fourth hypothesis is based on the group-decision making and overconfidence literature. As previously mentioned, several studies have explored mechanisms to reduce overconfidence, however, almost all the results were related to overconfidence in individual judgements rather than studying individual overconfidence when making decisions in groups. Nevertheless, there were some studies considering group judgements rather than individual judgements, and found that individuals taking decisions in groups had the tendency to be more confident and reduce their accuracy when making decisions than those taking decisions individually (Plous 1995). However, such evidence has not been proven in the financial area. Thus, a fourth hypothesis was postulated. This fourth hypothesis tries to answer whether individual confidence can be reduced if decisions are taken in groups rather than individually. The fourth hypothesis is as follows:

Hypothesis 4 (H_4): When decisions are taken in groups, there is a decrease in confidence and an increase of accuracy.

5. METHODOLOGY

As concluded from the literature review, the methodologies most often used to study overconfidence are calibration tests to measure the levels of overconfidence on individuals, and laboratory experiments to show evidence of the influence of overconfidence in financial decision making, or in other situations. For that reasons, to test our first hypothesis, we will use the calibration test methodology, and to test the second, third and fourth hypotheses, we will use a laboratory experiment.

A laboratory experiment is justified in this context because it has helped economists to resolve important empirical challenges such as going beyond correlational analysis to come up with a better understanding on causation (Peón, Antelo, and Calvo 2016), and laboratory experiments were also used to test a variety of issues, including the assimilation of information (Levitt and List 2009).

There are empirical studies using experiments that are closely related to the topic of our research. In particular, Hilton et al. (2011) used an experiment to study whether judgmental overconfidence, evaluated by probability miscalibration, is related to positive illusions about the self, where they demonstrate that hard-easy effect works differently on self-placement of someone's own performance relative to others' performance and over precision. At the same time, they demonstrated that subjects had a tendency to consider themselves 'better-than-the-average' (Hilton et al. 2011).

In addition, in the literature related to experimental asset markets the use of experiments is common, following the pioneering work of Smith, Suchanek, and Williams (1988). This literature shows that there is a tendency for such markets to follow a common pattern: bubble and then crash even when the dividend process was of common knowledge. An important finding is that in experiments with subjects without experience, prices were following a consistent pattern of starting below intrinsic value before abruptly rising above it and then crashing (Cheung and Palan 2012).

Nevertheless, a classic issue with controlled experiments is their external validity. Specially, the fact that people are in an environment where they are aware that their way of acting is being recorded, monitored, and then examined. This could compromise generalization (Levitt and List 2007). Market incorporation, replication, and monetary incentives would improve the validity of the experiment, but may not completely solve this problem (Peón et al. 2016). However, monetary incentives will be introduced to provide validity to the experiment and increase participants' motivation to respond accurately. Therefore, the external validity may be questionable. Market incorporation,

replication, and monetary incentives could improve the validity of the experiment, although may not completely solve this problem (Peón et al. 2016).

However, there is an extensive literature that demonstrates that the results with inexperienced students are not statistically different from the experiments that have been done with people who are familiar with financial concepts. There are three studies that consider the effect of trader characteristics about the bubble-and-crash fact. Van Boening, Williams, and LaMaster (1993) realized a call market trading experiment with business and professional people where they concluded from this that their results were not a consequence of using student subjects. Alike, Ackert and Church (2001) reported an experiment in which they compared markets in which students were business students with markets in which students were non-business students, and they observed that there was the same bubble and crash pattern with both types of students. Finally, King et al. (1993) carried out an experiment with two markets in which the subjects were stock market dealers and corporate executives. They conclude, as well, that this did not affect the general pattern of trade commonly observed with student subjects without experience.

Being aware of this limitation and with the caveat that the results may not be generalizable, we introduce monetary incentives to increase participants' motivation to respond accurately.

5.1 Experiment Structure

Participants

A total of 140 undergraduate TecnoCampus students completed one questionnaire and participated in a laboratory experiment in one occasion, during classes. These students were enrolled in three different academic degrees: 71 business management and marketing (in Spanish) students, 41 business and innovation management (in English) students, and 28 business management (in Spanish) students.

Procedure

The questionnaire consisted of a calibration scale in a confidence range format (as Hilton et al., 2011). The scale was followed by self-placement questions concerning participants' performance on this scale, which allowed us to identify participants' levels of overconfidence. The level of miscalibration based on those general knowledge questions was the main variable to be measured, which along with the levels of overconfidence measured in the second part of the experiment, were analysed through a correlation matrix to test the first hypothesis.

In the second part of the experiment we followed the pioneering work of Smith, Suchanek, and Williams (1988) which consisted of a continuous anonymous double auction experiment, where we studied participants' behaviour in three different scenarios with the same market situation: in the *first scenario*, participants had to make individual financial decisions without receiving any information about their position in the game with respect to others. This group was the control group and was used to test the third and fourth hypothesis. In the *second scenario*, participants had to make individual financial decisions while receiving feedback in periods 3, 6, 9 and 12, to know their real position with respect to others. This second scenario was Treatment 1, and this scenario was used to test hypothesis three. Finally, in the *third scenario*, participants had to make financial decisions in groups of two people without receiving any feedback, about their position, with respect to others. This third scenario was Treatment 2, and was used to test hypothesis four.

Both treatments were tested in an OLS regression in which the dependent variable was the level of risk in asset trading, which was measured by taking into account the variability of prices (overpricing of the shares) at the end of period 15, and the increase in trading by taking into account the variability of volume in shares at the end of period 15. Finally, the main independent variable was a dummy variable of having been treated or not.

5.2 Calibration Test

As seen in the literature, we defined at least four ways in each overconfidence can appear: Miscalibration, Better-than-the-average effect (BTA effect), Illusion of control, and Unrealistic optimism (Optimism bias).

In this study we will mainly focus on overconfidence as miscalibration. Studies in calibration usually use one of the following response formats to measure overconfidence: half-range format, full-range format, or interval estimation format. The half-range format is when participants select one of two options and assess the probability that the selected answer is correct within a specified scale which is normally from .5 to 1, which is usually expressed in percentages (50% to 100%) and the scores are calculated as the difference between the mean probability of an individual's judgement of a correct answer and the proportion of correct answers. If the difference is positive, it indicates overconfidence. If the difference is negative, then it indicates under confidence (Olsson 2014).

In a full-range format the probability that something is true or not, is assessed between 0 and 1, again usually expressed as percentages (Olsson 2014). In this case, overconfidence is assumed to appear when participants give probabilities larger than .5, and that the statement is false when they give ratings lower than .5. The probabilities of the 50% are randomly assigned to favour both options, the truth and the falsity of the statement (Olsson 2014).

In an interval estimation format participants are faced with statements such as:

	Lower	*Upper*
Give an upper and lower limit such that you are 90% sure that the mean birth rate of the population of Nigeria lies between _____ years and _____ years.		

In this case, overconfidence scores are calculate by taking into account the difference between the stated probability interval and the true value (Olsson 2014). This last format has proved the existence of overconfidence when difficult questions were assessed. For that reason, we decided to mainly focus on the *'interval estimation format'* to develop the calibration test.

In order to determine the confidence level of each individual before proceeding with the experiment, a calibration test was conducted at the TecnoCampus University in Mataró, Spain. During lectures, participants were announced that they had the opportunity to participate on a short experiment in which a general knowledge test had to be filled in, followed by a laboratory experiment which will be defined later. The test was done with *Mach Form* and it consisted of 11 general knowledge questions which

were not related to financial markets, previous experiments or economics to avoid any kind of advantage, and two different calibration tests were designed to avoid that participants of the same university degree could anticipate results. See *Annex 1 - Test 1 – IP*[16] *scale* and *Annex 2-Test 2 – IP scale* for the calibration tests.

As we already mentioned, we used the interval estimation format defined by Olsson (2014) to design the calibration test. This method consists of a typical setting, where participants are asked to reveal a lower and upper bound for the n-percent confidence interval – which in our case will be 90-percent – of a correct answer to a general knowledge question.

The eleven general knowledge questions followed the Hilton, D., Régner, I., Cabantous, L., Charalambides, L. and Vautier, S. (2011) model, since it is a model that proved the existence of overconfidence when difficult questions are to be answered and, at the same time, proved that difficult questions increased the worse than the average effect on the self-placement task. For that reason, a comparative analysis would be possible.

Each calibration scale was followed by two questions designed to capture overconfidence when self-placing to measure the WTA effect with the calibration self-placement (cSP) index. This two questions also followed the Hilton et al. (2011) model since, as previously mentioned, they have shown evidences of the appearance of the worse than the average effect on self-placement tasks when overconfidence increases.

The first question asked participants to indicate to what extent they thought that they had succeeded at the previous calibration task. The second question asked participants to indicate to what extent they thought that most of the other students had succeeded at the same task. For both questions, participants had to answer on an 11-point scale rating base where 0 is '*completely failed'* and 10 is '*completely succeed'.*

Experimental papers in psychology do not provide monetary incentives for reporting confidence interval, they usually offer a small reward to participants if the value they provided falls within the true value (Blavatskyy 2009), which is the system that we used. For that reason, to motivate participants to provide accurate answers, incentives were provided (which will be described later).

This pre-experimental test allowed us to gather participants' data in our database, and their overconfidence levels which were useful to test whether hypothesis 1 was rejected, or not.

[16] IP = Interval production task

Finally, the duration of the activity lasted, approximately, 10 minutes and the test was conducted by each individual. The test was done before starting the laboratory experiment and feedback was not provided. Communication among subjects was not allowed.

5.2.1 Measurement

In order to identify overconfidence levels we followed the Hilton et al. (2011) method, which consists of subtracting the others' evaluation score from the self-evaluation score where the key variable will be the miscalibration indicator. We identified the following:

If the result was *positive*, it indicated that participants evaluated themselves as being better than the average (BTA) and, as mentioned in *Section 2.1.2.1*, the better than the average effect is a form of overconfidence.

On the contrary, if the result was *negative,* it indicated that participants evaluated themselves worse than the average (WTA), which means that they are under-confident.

Therefore, the following table will summarize the indicators used to test the first hypothesis, which will be measured through a correlation matrix.

Table 1. Summary of Calibration Test indicators to measure Overconfidence

Variable	***Values***	***Interpretation***	***Calculation***	***Literature***
cSP	+ → BTA - → WTA	If cSP + → overconfident. If cSP - → underconfident.	Self-calibration – Others' calibration (C-D)	Calibration self-placement (cSP) index: Hilton et al. (2011)
C_T	1 = Yes 0 = No	Calibration Task. If 1, the true answer is in the interval. If 0, the true answer is not in the interval.	Range = U_B-L_B C_T → =IF(AND(TA<=U_B ;TA>=L_B);1;0)	Calibration self-placement (cSP) index: Hilton et al. (2011)

However we wanted to go further and identify if the true value was NOT in the suggested range, and how far it was from the range. We wanted to do so because

believing that you have done better than the average is just a way to identify, or measure overconfidence. However, from a financial point of view, it is much more interesting to know if the participants' estimation is well above, or far below, the real or fundamental values of the asset, which it is precisely the reason why there are market bubbles. To identify if the true value was inside the range and how far it was from the range, we used the Hilton et al. (2011) method.

Table 2. Summary of Calibration Test indicators – Range and Distance

Variable	***Values***	***Interpretation***	***Calculation***	***Literature***
U_C	1 = Yes 0 = No	Under-confidence. If U_B is less than true value, 1= Yes, 0= No	=IF(AND(C_T=0;U_B< TA);1;0)	Calibration tasks: Hilton et al. (2011)
		If under-confidence, distance to the true value.	=IF(AND(TA=0;U_B< TA);U_B-TA;0)	Calibration tasks: Hilton et al. (2011)
O_C	1 = Yes 0 = No	Overconfidence. If L_B is greater than the true value, 1= Yes, 0= No	=IF(AND(C_T=0;U_B> TA);1;0)	Calibration tasks: Hilton et al. (2011)
		If over-confidence, distance to the true value	=IF(AND(TA=0;U_B> TA);L_B-TA;0)	Calibration tasks: Hilton et al. (2011)

5.2.2 Incentives

The individual, or group, of each respective session, who was, in general, closer to the correct answer, was the one getting the incentive. Therefore, the monetary payoff was based on the final performance.

At the end of the calibration test, the participant, or group, (of each session) that provided the most accurate answers received a total amount of *3 EUR.*

The main indicators to identify the individual, or group, with the most accurate answers were the following:

Table 3. Summary of test indicators - Incentives

Variable	***Values***	***Interpretation***	***Calculation***
MP		Middle point of the L_B and U_B	= (L_B+U_B) / 2
A_{CC}		Accuracy. Distance to the true value	= (TA-MP)^2
SS		Sum of Squares from question 1 to question 11	= (Q_1 + Q_2 + Q_3 + … + Q_{11})
MIN		Minimum of all sum of squares	= MIN (SS_1:SS_{11})
R	True = Most accurate answers False = Not most accurate answers	Raking. Individual, or group, with the most accurate answers	(= SS = MIN)

5.3 Design of the laboratory experiment

5.3.1 General characteristics of the experiment

The experiment consisted of a set of 20 experimental sessions that were conducted at the TecnoCampus – UPF University, between January and March 2020, during the teaching period. For each session, 7 to 8 students were participating, having a total of 140 undergraduate TecnoCampus students who completed the laboratory experiment. These students were enrolled in three different undergraduate degrees: business management and marketing (in Spanish, 71 students), business and innovation management (in English, 41 students), and business management (in Spanish, 28 students).

Each session lasted, approximately, 1 hour and 30 minutes. Instructions were provided at the beginning of the experiment and participants were familiarized with the rules of the experimental market before starting. As in any laboratory experiment, different incentives were established, which will be described later.

All experimental sessions were conducted in a computer lab and 7-8 players were participating in each of the experimental asset markets. The asset market situation was the same for all participants, and the experiment was programmed and conducted with the software z-Tree (Fischbacher 2007).

At the beginning of each session participants were able to read in detail the instructions and ask questions and doubts, if any. Two trial periods were provided to familiarize participants with the software and, questions and doubts were allowed. However, this trial periods were not used when interpreting results (participants were informed before starting the trial). Each subject started the first period with the same number of shares and amount of money in their cash balance. Communication among subjects was prohibited during the experiment.

As previously mentioned, the experimental design followed the pioneering work of Smith, Suchanek, and Williams (1988) which was performed as a continuous anonymous double auction. Within this framework, we could study whether overconfident people were taking riskier financial decisions, whether feedback availability reduces participants' overconfidence levels, and whether making decisions in group might decrease, or not, overconfidence levels and accuracy, through their changes of strategy with respect to prices and volume of shares.

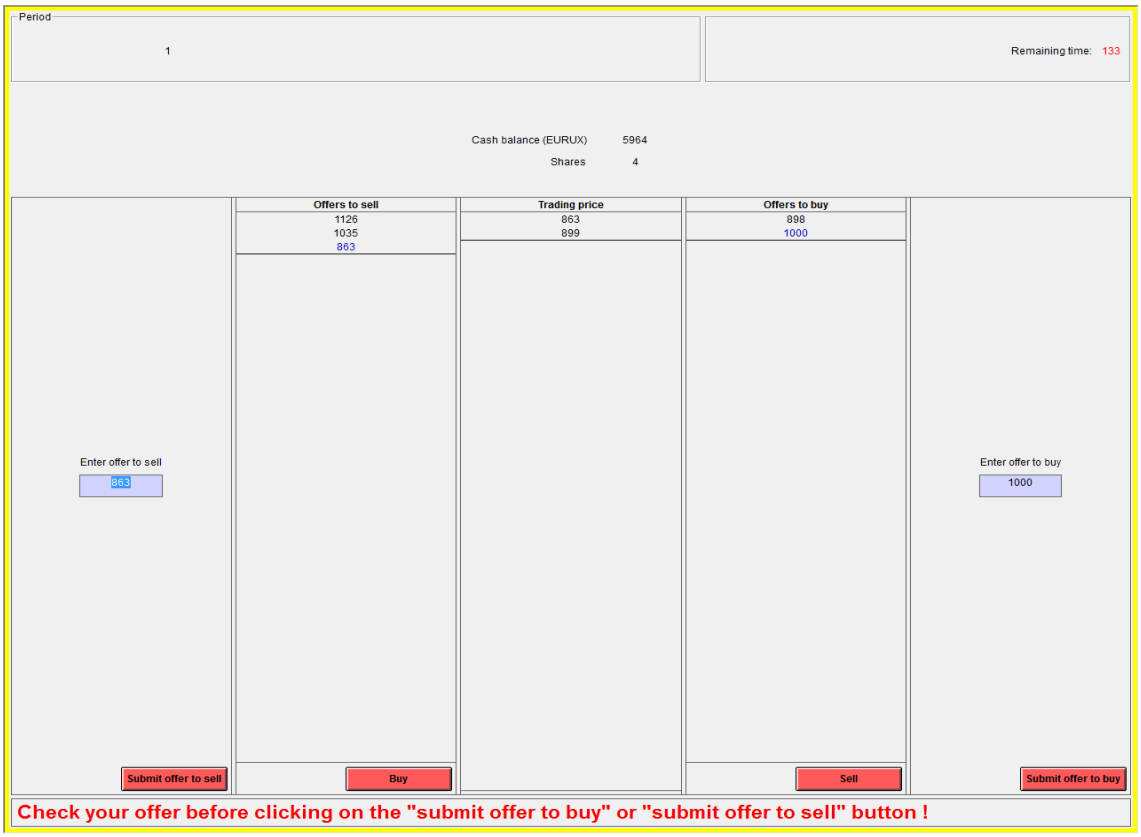
The game consisted of a trading mechanism in a continuous double auction (Figure 1) with open order books. Participants initiated a transaction by posting offers to buy (bids) and offers to sell (offers). Subjects could put multiple offers, however, each

offer was just for the transaction of 1 share. The prices at which shares were bought or sold, were classified in two columns on real time where the best offer was appearing at the bottom of the order book, but this did not mean that such offer was a good one. Participants executed an operation when they had select the offer that suits them best and pressed the "buy" or "sell" button at the bottom of the order books (Giusti, Jiang, and Xu 2016).

Shares had a life of 15 periods. Each share was paid with a random dividend at the end of each period. The distribution of the dividend was over time. In each period, cash earned dividends and was maintained in the cash-balance. Subjects could use money (EURUX) from the cash balance account to purchase shares. Revenues, or expenses, from buying or selling shares were automatically deposited into, or deducted, from the cash-balance account, and it was maintained for the following period (Giusti et al. 2016).

In period 3 – 6 – 9 – 12 participants had to insert their prediction with respect to the position they were in the market, according to the benefits (EURUX) obtained, compared to other players, and the percentage of how confident they felt with the answer previously provided.

Figure 1. Z-Tree program continuous double auction market



Source: Z-tree screenshot

5.3.2 Incentives

As in any laboratory experiment, a set of monetary and non-monetary payoff, which are experimental incentives that participants earn based on participants' performance, were established. The final performance obtained was calculated based on two elements. First, the cash held at the end of the last period, the final result. Second, the estimation of their position with respect to others during the experiment. At the end of the period 3 – 6 – 9 – 12, two questions appeared on the screen:

- *Question 1: In which ranking position do you think you are?*
- *Question 2: On a scale of 1 to 100, what is your degree of confidence about the position you just entered?*

In question 1, participants had to indicate their current position with respect to other players, in their group, and with respect to the money they had in their savings account and, in question 2 participants had to indicate how confident they felt with the answer provided in question one. If for example they put 80, it means that they were 80% confident of the answer provided in question 1.

The final result of the experiment depended on their choices and considered 2 elements:

First, their performance in the financial market: *How much money have you earned at the end of period 15?*

- The *1st performing* subject obtained a grade of 10 in the corresponding percentage of the final grade of the course.
- The *2nd* obtained a grade of 9 in the corresponding % of the final grade of the course.
- The *3rd* obtained an 8 in the corresponding percentage of the final grade of the course.
- The *4th* obtained a 7 in the corresponding percentage of the final grade of the course.
- The *5th* obtained a 6 in the corresponding percentage of the final grade of the course.
- And, the *6th and 7th* obtained a grade of 5 in the corresponding percentage of the final grade of the course.

Second, the estimation with respect to others.

At the end of the last period, the two subjects from each group, who best estimated their position with respect to the others, received money:

- The *1st performing* subject obtained a monetary incentive of 7 EUR.
- The *2nd performing* subject obtained a monetary incentive of 5 EUR.

5.3.3 Data analysis

5.3.3.1 Collecting data

With the data that we collected from the laboratory experiment, explained in the previous section, we used two measures to calculate and describe the trading behaviour of each session. The two variables that we used were: the Relative Absolute Deviation (RAD) and the Relative Deviation (RD); both have been used in previous works such as: Giusti et al. (2016) and Stöckl, Huber, and Kirchler (2010).

The RAD measures the average level of mispricing compared to the average fundamental value of the period, and the RD measures the extent of over or under-valuation (Giusti et al. 2016).

- Relative Absolute Deviation. $RAD = \frac{1}{N}\sum_{p=1}^{N}\frac{|\overline{P}p-FV|}{|\overline{FV}|}$
- Relative Deviation. $RD = \frac{1}{N}\sum_{p=1}^{N}\frac{(\overline{P}p-FV)}{|\overline{FV}|}$

Being:

- *p* = indexes period
- *N* = total number of periods
- $\overline{P}_p$ = average price in period p
- FV = fundamental value
- $\overline{FV}$ = average fundamental value of the market (equals 192 EUR).

Additionally, for a better understanding of the trading behaviour of each individual, we used the 'Bid-Ask Spread' variable. The Bid-Ask spread indicates the difference in prices between the offer, or sale, and the demand, or buy. The Bid is the buying price, the price at which a share is offered, and the Ask is the sale price, the price they ask you. Normally the buying price is lower than the selling price. Typically, buyers want to buy low and sellers want to sell high.

To end up, we used two other measures to calculate and describe the confidence behaviour of each observation. The two variables used were: the Confidence Level (CL) and the Ranking Distance (cSP). The CL measures the confidence level about self-reported position in percentage, and the cSP measures the difference between the self-reported position and the true position.

5.3.3.2 Statistical measures

Using the above framework, we investigated one possible factor that may contribute to overconfidence studies in financial decision making which is the so called 'risk' factor. We studied whether there exists a correlation between overconfidence and riskier financial decisions in asset markets. To measure risk, we took into account the variability of prices which was the main indicator. Furthermore, a linear regression was used to explain the causality of the results obtained. Therefore, to construct a model used to approximate the dependency relation between a dependent variable Y, in this case 'risk' in form of the variation of prices to sell and buy, and the independent variables X, in this case 'overconfidence' in form of: BTA, calibration self-placement index (cSP), and average confidence. This regression provided us with enough evidences of the double auction market behaviour to reject, or not, hypothesis 2.

Taking advantage of the flexibility of the double auction market, we ran two treatments with the same market situation but different scenarios. The first treatment features feedback availability in period 3-6-9 and 12. This treatment was related to overconfidence, which means that overconfidence was the dependent variable and feedback the independent variable. Additionally, we studied whether overconfidence was reduced when individual performance feedback was provided. This treatment was used to answer hypothesis three, and was measured with four main indicators: their established ranking position, the real raking position, the distance of the estimated value, the real value, and the average confidence in self-reporting their positions in period 3-6-9 and 12.

Before testing the hypotheses we investigated whether the differences among treatments were significant. To do so, we elaborated a t-test analysis which compares the means among both treatments. Furthermore, the hypothesis was tested through a '*Linear regression'* in which the key explanatory variable was a dummy variable equals 1 if the individual was treated (received feedback) and zero if the individual belonged to the control group. However, we ran two linear regression estimators; OLS for panel data on the level value of the dependent variable and on its first differences (using the command xtreg in Stata) to reaffirm the results obtained with the aforementioned regression.

The second treatment features group judgement without feedback availability. In this case we also carried out a comparison among means, to investigate whether the differences among treatments were significant. To do so, we elaborated the same t-test (T-student distribution) analysis (previously mentioned) which compares the means among both treatments.

This treatment followed the same structure and market situation as the first study of risk measurement. However, decisions were taken in groups of two. We studied whether there was an increase, or decrease, of overconfidence levels and accuracy when making decisions in groups of two, to reject, or not our fourth hypothesis. Again, this second treatment was related to overconfidence which means that overconfidence keeps being the dependent variable and group decision making an independent variable.

The main indicators were the same as in treatment two, and the hypothesis was tested through a *'Linear regression'* in which the key variable was a dummy variable equals 1 if the individual was treated (took decisions in groups of two) and zero if the individual belonged to the control group.

6. RESULTS

6.1 Test results

Characteristics of the participants

In the first part of the experiment, a total of 137 students conducted a calibration task test. It should be mentioned that, 3 participants did not conduct the calibration test, but participated in the laboratory experiment, thus becoming 140 participants in total.

Table 4. Summary of the descriptive statistics of participants

	Gender				**Age**							
	Individually		**In pairs**		**Individually**				**In pairs**			
	N	**%**	**N**	**%**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Female	47	34.31	-	-	20.49	1.89	18	27	-	-	-	-
Male	58	42.34	-	-	21.69	2.67	18	30	-	-	-	-
FF[17]	-	-	10	7.29	-	-	-	-	19.70	.949	18	21
MM[18]	-	-	13	9.49	-	-	-	-	20.77	1.59	19	25
FM[19]	-	-	9	6.57	-	-	-	-	21.33	2	19	25
Total	105	76.65	32	23.35	21.09	2.28	18	30	20.6	1.51	18	25

Table 4 summarizes descriptive statistics of participants for the calibration test. All subjects participating individually were between 18 and 30 years old, having an average of 21.09 years old and, those participating in groups of two were between 18 and 25 years old, having an average of 20.6 years old. Due to the fact that in our sample the age is very similar, the following results are going to be presented differentiating results by gender (female and male) and pairs (individually and in pairs).

[17] Female / Female
[18] Male / Male
[19] Female / Male

Miscalibration

Table 5 summarizes overconfidence rates for the calibration tasks used in both questionnaires. The average proportion of surprises[20] for the test (M = 72.06%, SD = 18.18) were high enough to assume that the scale of the test was difficult since the usual average proportion of surprises obtained, with the same scale, in previous studies was between 40 and 60% which was already considered as a high percentage (Hilton et al., 2011). Fact that can also be seen through the better than the average indicator. Participants described themselves with a 44.5%, better than the average (M = 0.445, SD = 0.50) which means that 55.5% of participants considered themselves worse than the average.

Self-Placement Scores on the Calibration Tasks

The perceived difficulty of these scales was supported by the findings on the calibration self-placement scores *(Table 5)*. Participants reported that they had performed worse than most other students (WTA effect), as indicated by the negative value of the cSP index (M = -.002, SD = 1.22) and the BTA (.445) indicator. This result shows the expected dissociation mentioned by Hilton et al. (2011), Larrick et al. (2007), and Kruger (1999): where they mention that task difficulty increases overconfidence on the calibration task but increases the worse than the average (WTA) effect on the self-placement task.

Table 5. Better than the average and Self-placement Scores - Test

	Test – IP Scale			
Variable	**Mean**	**SD**	**Min**	**Max**
Percentage of surprises	72.06	17.43	18.18	100
Better than the average (BTA)	0.445	0.50	0	1
Calibration self-placement (cSP)	-.002	1.22	-4.82	4

[20] Additionally to surprises' rates, accuracy (to what extend midpoints of intervals were close to the true value) and informativeness scores (which is indicative of the intervals' width) were also calculated for the miscalibration scale following the procedure of Hilton et al. (2011) and Yaniv and Foster (1995). Findings on this scores indicated that lower rates of surprises were because of midpoints closer to the true answer, and/or wider intervals.

Better than the average effect and accuracy by gender and pairs

Table 6 summarizes the better than the average effect and accuracy intervals by gender for the calibration tasks used in both questionnaires. The average of the BTA indicator for female (M = .234, SD = .424) was lower than the BTA indicator for male (M = .517, SD = .500) which indicates that men were considering themselves, in a higher proportion, better than the average, and women were considering themselves more worse than the average. The average of intervals of accuracy is higher male than for female (Female; M = .246, SD = .180 – Male; M = .288, SD = .153) which indicates that men were providing more accurate answers than women.

Table 6. Better than the average and Accuracy Intervals by Gender

	Test – IP Scale			
Variable	**Mean**	**SD**	**Min**	**Max**
Better than the average – Female	.234	.424	0	1
Intervals of Accuracy – Female	.246	.180	0	.727
Better than the average – Male	.517	.500	0	1
Intervals of Accuracy – Male	.288	.153	0	.636

Table 7 summarizes the better than the average effect and accuracy intervals by pairs for the calibration tasks used in both questionnaires. The average of the BTA indicator was higher for participants making decisions in pairs (M = .625, SD = .485), rather than for those making decisions individually (M = .390, SD = .488), which it also indicates that those who were making decisions individually considered themselves, in a higher proportion, worse than the average. Finally, the average of intervals of accuracy was higher for those making decisions in pairs (M = .313, SD = .193) than for those making decisions individually (M = .269, SD = .167) which indicates that participants in groups of two were providing more accurate answers.

Table 7. Better than the average and Accuracy Intervals by Pairs

	Test – IP Scale			
Variable	**Mean**	**SD**	**Min**	**Max**
Better than the average – Individual	.390	.488	0	1
Intervals of Accuracy – Individual	.269	.167	0	.727
Better than the average – In pairs	.625	.485	0	1
Intervals of Accuracy – In pairs	.313	.193	0	.818

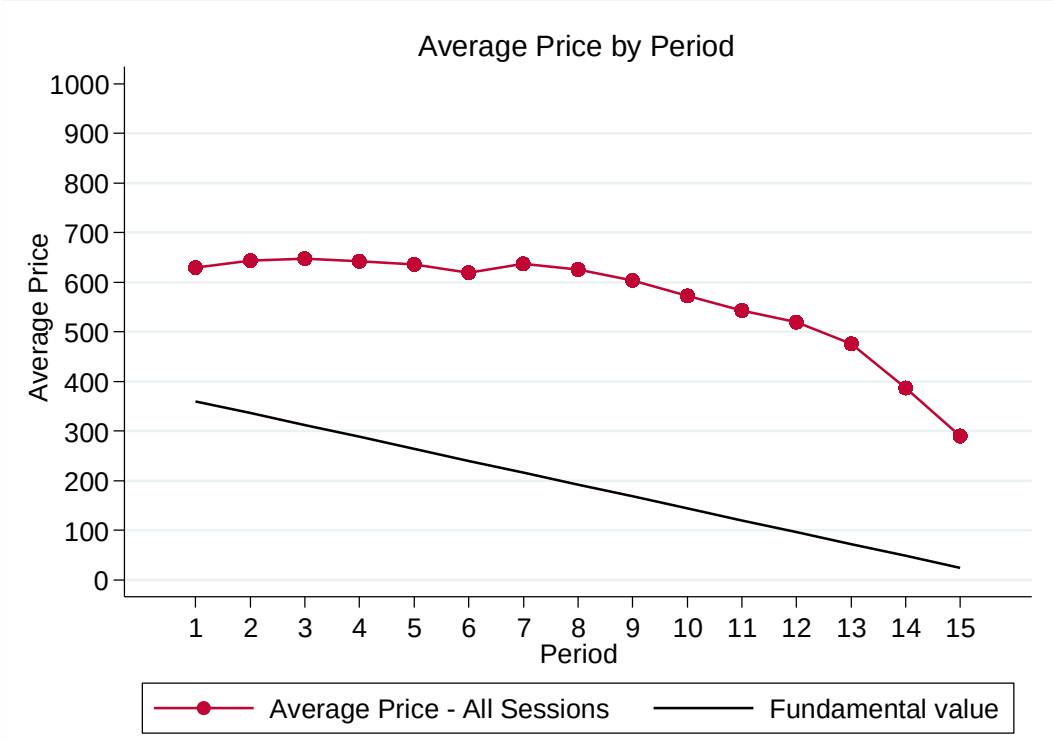
6.2 Market results

6.2.1 Trading behaviour

Figure 2 presents and initial impression of the trajectory of market prices in the double auction market, by showing the average transaction price over the 15 periods. The red line represents the average transaction price per period of all the sessions while the black line shows the fundamental value of the market. The data from all the markets were integrated in the same graph for a better understanding of the trend regarding average prices.

Some patterns are evident in Figure 2. First, prices deviate significantly from the fundamental value (FV) which is represented with a black line. The second pattern distinctly identifiable is found in the average price per period, the figure shows a coherent dynamic with what was expected; a decrease in prices as periods go by.

Figure 2. Average price evolution for the 15 periods

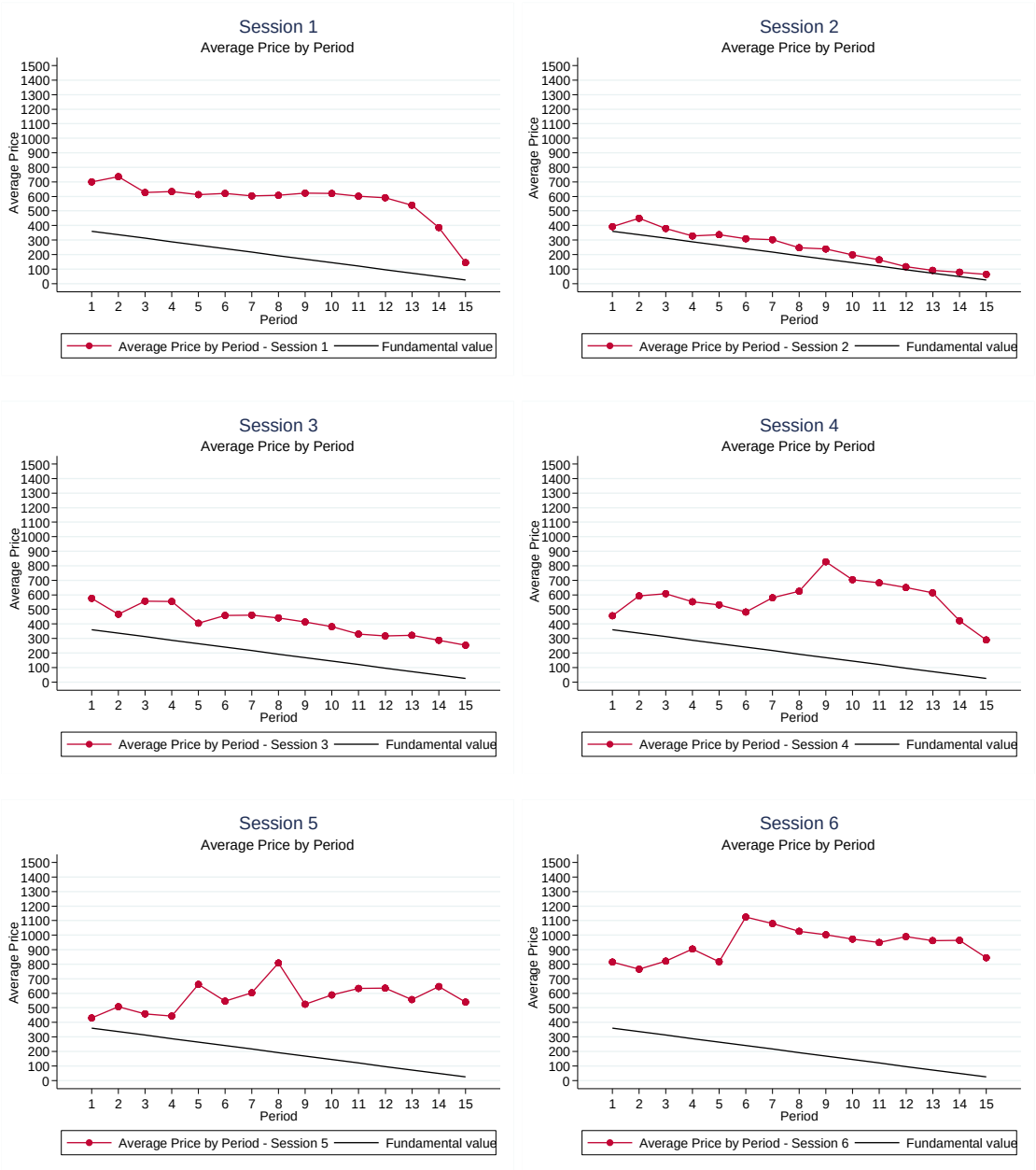


Source: Self-elaboration with Stata

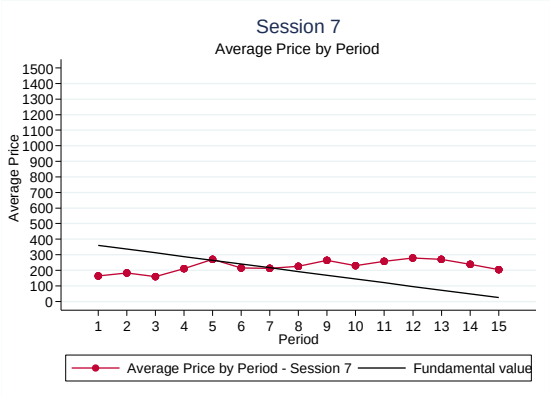
In order to have a better understanding of where the results come from, we followed a deep analysis of the results obtained from each individual observation which will explain the pattern seen in Figure 2. Figure 3 shows the time series of transaction prices of each session, as well as the fundamental value (FV).

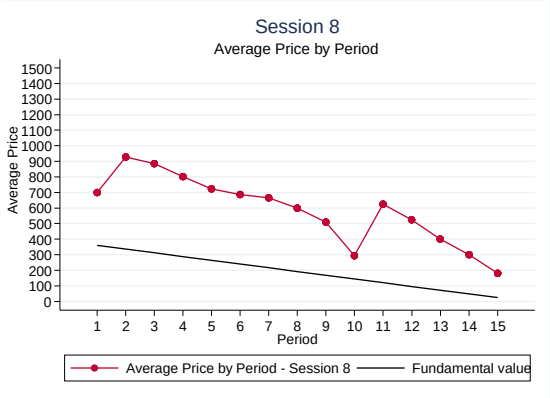
There is some heterogeneity between sessions within each treatment, they follow the basic overall pattern shown in Figure 2. Prices tend to be above the fundamental value and closer to it at the end of the market. We can also observe, in its majority, that there is a tendency for prices to fall as periods go by, but still being above the FV. However, it is important to mention that only in four sessions out of ten (session 7, 10, 16, and 20) had the average share price below the FV during several periods.

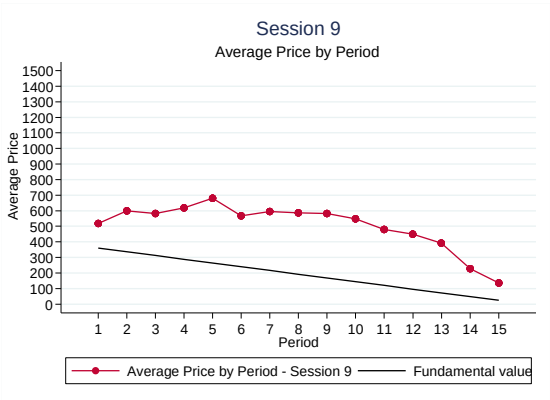
Figure 3. Average price evolution per period by Session

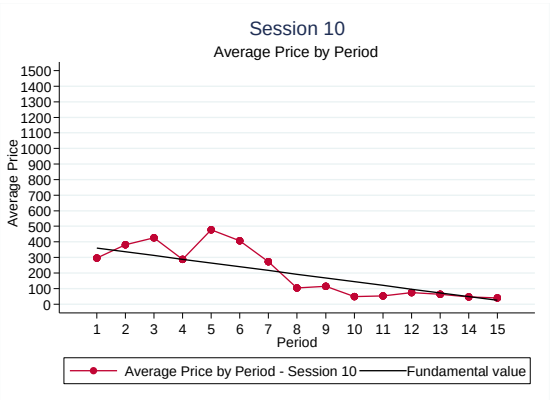


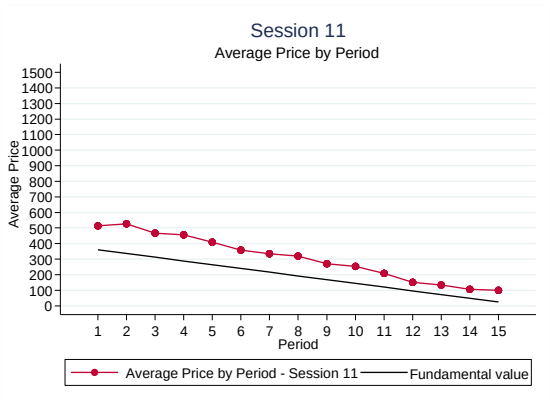
Source: Self-elaboration with Stata

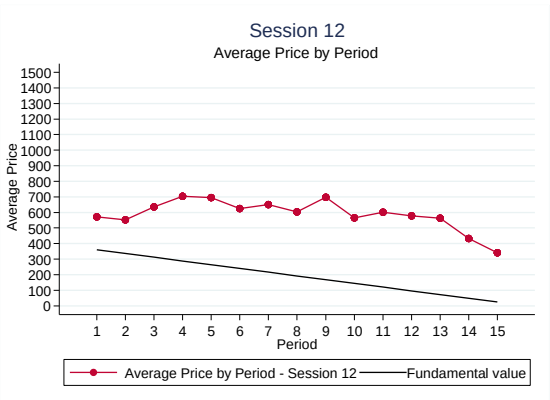


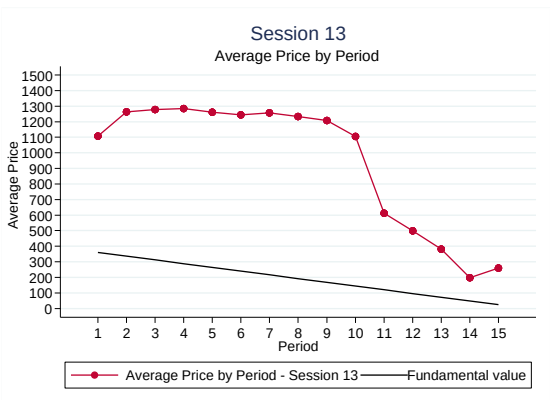


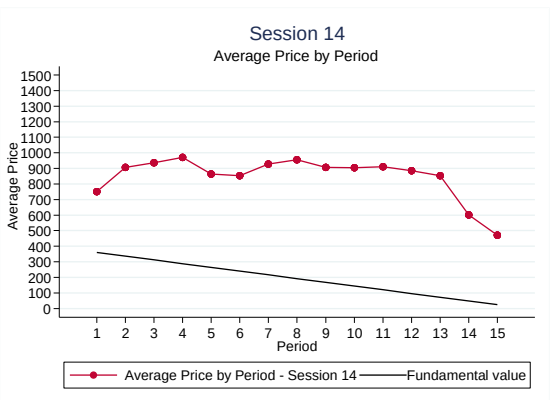




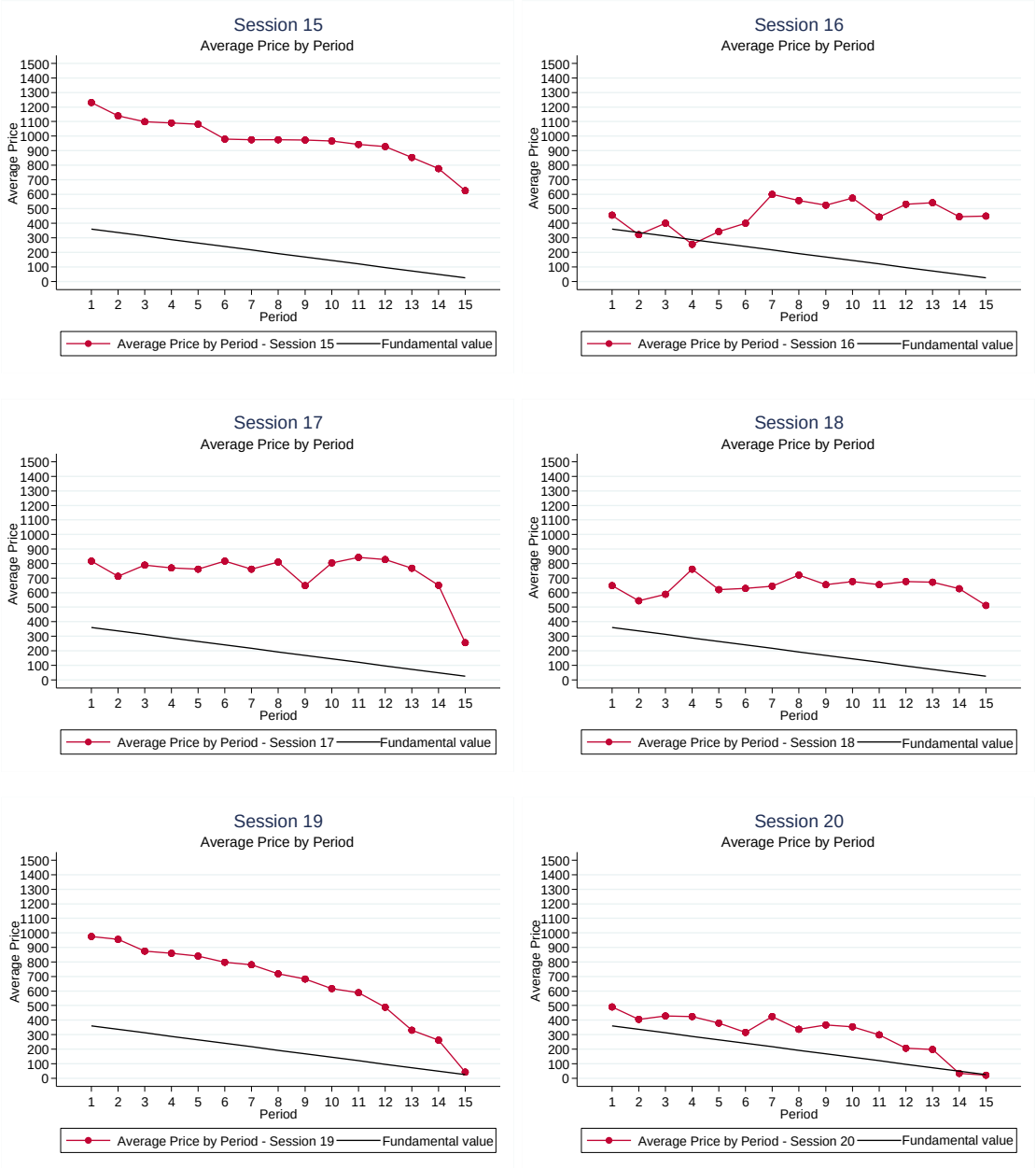








Source: Self-elaboration with Stata



Source: Self-elaboration with Stata

Going further with the analysis and, in order to measure the market behaviour per session, we used the RAD (Relative Absolute Deviation) and the RD (Relative Deviation) variables, as previously used by Giusti et al. (2016) and Stöckl et al. (2010) in past studies. Table 8 outlines the number of observations (N), the mean, the standard deviations (SD) and, the minimum (Min) and maximum (Max) for each variable; RAD and RD, for the sessions.

Table 8. Summary of descriptive statistics for the RAD and RD variables

	N	Mean	SD	Min	Max
RAD	140	1.975	1.355	.0023	5.427
RD	140	1.939	1.405	-1.026	5.427

It is observable that the results of the RAD variable (M = 1.975, SD = 1.335) and the RD variable (M = 1.939, SD = 1.405) indicate that there is a clear tendency to overprice the shares with respect to the fundamental value. In Table 9, you will be able to identify the RAD and RD variables obtained for each session:

Table 9. RAD and RD values obtained for each session

	RAD		RD	
	2,00	3,88	2,00	3,88
	0,28	0,60	0,28	0,17
	1,16	2,06	1,16	2,06
	1,99	1,63	1,99	1,63
	1,98	0,35	1,98	0,07
	0,60	1,41	0,60	1,38
	2,06	2,83	2,06	2,83
	3,93	2,34	3,93	2,34
	3,41	2,41	3,41	2,41
	4,08	0,64	4,08	0,63

According to the general results obtained, we can identify that the general trend in average prices of shares, in all the sessions, is to be above the fundamental value (FV). In other words; shares are being overvalued.

However, we went further and we studied the spread, which is a very important factor when dealing with assets, as it will significantly influence the profitability of an operation. Table 10 represents the Average Spread variable obtained for each period of all the sessions.

Spread is generated when in a balanced market, supply and demand meet and make an immediate transaction; however, as long as the sale price and the purchase price are not found, a time difference is generated which is the so called ‘spread’. For example, if we take into account Table 10 the spread for period one is of -149, the difference between the sell price (bid price) and the price to buy (ask price). Since the purchase price is higher than the sale price, if a speculator bought and sold immediately, he would lose € 149, which is the amount the market maker earns to give liquidity. The same happens for each period except for period 15, where the average spread is 19, which indicates that the sale price is higher than the purchase price, which means that if

a speculator bought and sold immediately, he would earn €19, which is the amount the market maker loses.

Table 10. Average Spread by Period for All the Sessions

Period	Mean	SD	Min	Max
1	-149	361	-1039	900
2	-166	299	-800	600
3	-143	300	-1020	750
4	-178	335	-950	700
5	-161	315	-925	480
6	-110	194	-600	300
7	-120,8	200	-600	400
8	-80	174	-629	300
9	-93	182	-555	300
10	-82	174	-690	150
11	-26	349	-530	1900
12	-50	181	-585	288
13	-37	146	-400	350
14	-30	147	-350	580
15	19	106	-349	216

Going further, we compared the average Spread by period and treatment to further analyse the trading behaviour. Table 11 represents the average Spread by treatment for each period. Thus, as we can observe, both treatments (Feedback and Group treatment) have very large means, which indicates that the difference between supply and demand has been very large, and therefore, the probability that a transaction has not been carried out is greater.

Taking into account the results, it can also be observed that the stakes have been very high for both treatments, however, we can see that for the Feedback treatment the spread (M = -108) is larger than for the Group treatment (M = -53), which it can also indicate that the transactions of the Feedback treatment have been more risky.

Table 11. Average Spread by Period and Treatment

Period	Feedback Treatment	Group Treatment
1	-189	-233
2	-267	-183
3	-176	72
4	-196	-76
5	-297	-191
6	-150	-90
7	-179	70
8	-111	-5
9	-44	-102
10	-67	-13
11	95	126
12	-12	-115
13	15	-40
14	-41	-37
15	-7	16
Total Spread Average	-108	-53

In this context, market makers -market creators- should have been the responsible for keeping spreads low, for example € 0.02, so that for 1 share there is a bid of € 99.98 and an ask of € 100.00, which means that the market maker will buy at (he will offer) € 99.98 for the share to whoever wants to sell them, and he will sell them at (he will ask for) € 100.00 to whoever wants to buy them. Thus, the market maker is buying cheap and selling expensive because the spread is their benefit.

This data is especially crucial for a speculator in an over the counter market as the market maker may incur high spreads to earn more profits at the expense of the speculator. Thus, those results are indicating us that the spread has been very high which minored the liquidity and increased the risk. However, in hypothesis 3 and 4 we will analyse whether the variable spread and the confidence variables have a statistically significant correlation, to investigate whether the confidence variables increase, or decrease, risk levels when each treatment is conducted.

6.2.2 Confidence levels and performance feedback behaviour

Table 12 shows overconfidence average rates over period 3, 6, 9, and 12 (it is when overconfidence was measured) when feedback was provided (treated group) and participants were making decisions individually, while Table 13 shows overconfidence average rates over period 3, 6, 9, and 12, when feedback was not provided and participants were making decisions individually (Control group). Out of the results, we can see that 107 out of 140 participants were taking decisions individually from which; 51 of them received feedback and 56 did not receive feedback.

A priori, we can observe that the overconfidence average rate of the treated group (Feedback treatment) was slightly higher (M = 66.97, SD = 22.51) than the overconfidence average rate of the Control group (M = 66.34, SD = 19.77). It is also observable that there is a clear tendency, in both treatments, to increase confidence levels when periods go by. These results show an independence between overconfidence and performance feedback as a method to directly eliminate overconfidence levels mentioned by Plous (1995), as well as with the study of Slovic et al. (1982) where he found out that overconfident people could be better calibrated when feedback was provided. However, later on, we will perform different statistical measures to confirm, or not, a disassociation among both variables.

Table 12. Confidence average rates by period – Feedback treatment

	Period	Obs.	Mean	SD	Min	Max
	3	12	62.08	21.32	8	100
	6	13	66.94	23.03	5	100
	9	13	69.60	21.72	6	100
	12	13	69.24	23.97	6	100
Total	[3,6,9,12]	51	66.97	22.51	6,25	100,00

Table 13. Confidence average rates by period – Control Group

	Period	Obs.	Mean	SD	Min	Max
	3	14	65.52	19.47	6	100
	6	14	65.88	20.64	5	100
	9	14	65.95	18.24	6	100
	12	14	68.00	20.71	6	100
Total	[3,6,9,12]	56	66.34	19.77	5,75	100,00

For a better understanding of the results, to know where results come from, we went through the obtained outcomes from each individual, of both: Feedback treatment and Control group, which will explain the results obtained in the previous tables: *Table*

12 and *Table 13*. Table 14 (Annex 1) shows the confidence average rates of each session for period 3, 6, 9, and 12 of the Feedback, and Table 15 (Annex 2) shows the same but for the Control Group.

There is some heterogeneity among sessions within each treatment. In both treatments, the confidence average rates oscillate between 60 and 80%, which are considered to be quite high taking into account the fact that they had no previous experience in double auction markets.

Once again, it is observable that there is almost no difference among treatments which excise the idea of providing feedback to reduce confidence levels. However, as previously mentioned, later on we will use different regressions to reaffirm, or not, these results and reject, or not, our hypothesis 3.

Additionally, in Table 16 (Feedback treatment) and Table 17 (Control group) we measured confidence levels through another variable which is the so called: *cSP* which measures accuracy (difference self-reported position minus the true position, also called: *'rank distance'*). In this case, we can observe that the overall average of the cSP index was smaller for the Control group (M = 0.101, SD = 1.861) than for the Feedback treatment (M = 0.145, SD = 1.499). These results indicate that subjects of the Feedback treatment were being less accurate when providing the estimation of their positions in the market with respect to others, than the Control Group. Fact that is correlated to the study of Oskamp (1995) where he concludes that accuracy does not increase when more information is provided. Nevertheless, the relation among variables will be studied in the following sections.

Table 16. Calibration Self-placement averages by period – Feedback treatment

	Period	**Obs.**	**Mean**	**SD**	**Min**	**Max**
	3	12	0.240	2.264	-5,00	5,00
	6	13	0.020	1.647	-3,00	6,00
	9	13	0.260	1.242	-2,00	3,00
	12	13	0.060	0.843	-2,00	2,00
Total	[3,6,9,12]	51	0.145	1.499	-3,00	4,00

Table 17. Calibration Self-placement averages by period – Control group

	Period	**Obs.**	**Mean**	**SD**	**Min**	**Max**
	3	14	0.178	1.830	-3,00	5,00
	6	14	-0.035	1.964	-4,00	5,00
	9	14	0.105	1.839	-4,00	5,00
	12	14	0.158	1.811	-4,00	5,00
Total	[3,6,9,12]	56	0.101	1.861	-4,00	5,00

For a better understanding of the results we went through the obtained outcomes from each individual, of both: Feedback treatment and the Control group when decisions were done individually, which will explain the results obtained in the previous tables. Table 18 (Annex 3) shows the calibration self-placement of each session for the Feedback treatment, and Table 19 (Annex 4) shows the same but for the Control Group. It is observable that there is some heterogeneity among sessions within each treatment. In the Feedback treatment just 4 out of 10 sessions had a negative cSP average which means that most of the participants of those sessions were placing themselves in a better position than they actually were *(i.e., self-reported position: 2nd, true position: 6th, cSP = 2 – 6 = -4).* However, the overall average of the cSP index was still positive because the 6 remaining sessions had positive averages. The same happened with the Control Group sessions (Table 19), 6 out of 10 sessions had a positive cSP average and the 4 remaining sessions had a negative average, however, those negative averages were not high enough to impact the overall average which continued being positive. Finally, comparing both treatments, we can observe that the Control group (M = 0.05, SD = 1.65) continues being more accurate than the Feedback treatment (M = 0.073, SD = 1.53).

6.2.3 Confidence levels and group-decision making behaviour

Table 20 shows the average confidence and calibration self-placement results by Treatment (Group treatment and Control Group) and, without the availability of performance feedback. Out of the results, we can observe that the confidence average rate of the treated group (Group treatment) was slightly higher (M = 67.54, SD = 22.39) than the confidence average rate of the Control group (M = 66.34, SD = 19.77). Additionally, another important factor to take into account is the calibration self-placement (cSP) used to measure accuracy. We can clearly observe that participants making decisions in groups of two have the tendency to be less accurate when indicating their positions in the market (M = -0.34, SD = 1.83) than participants making decisions individually (M = 0.1, SD = 1.86).

At this point, we can observe that the obtained results agree with the study of Heath and Gonzalez (1995) where they proved that interacting in groups of two makes that confidence levels in someone's own predictions increases and accuracy decreases. However, we will later on study the relationship among both variables, through different statistical measures, to investigate if there is a significant correlation.

Table 20. Confidence levels and Calibration Self-Placement by Treatments

	Confidence level		cSP	
	Group Treatment	**Control Group**	**Group Treatment**	**Control Group**
Obs.	14	57	14	57
Mean	67,54	66,34	-0,34	0,10
SD	22,39	19,77	1,83	1,86
Min	7	5	-5	-4
Max	100	100	3	5

In order to explain the results obtained in Table 20 we went through the obtained outcomes from each individual by period, for both treatments.

Table 21 shows the average confidence levels and calibration self-placement of each period for the Group Treatment, and Table 22 shows the average confidence levels and the calibration self-placement of each period for the Control group.

If we take a look to the average confidence levels, it is observable that the Control group follows a more stable pattern than the Group treatment. However, it is visible that that participants making decisions in group of two were a little bit more confident than those making decisions individually. In the case of the cSP index, we can observe that the Group Treatment had the tendency to be less accurate when providing their positions in the market than the Control Group. However, later on, we will use different statistical measures to compare both means and to investigate whether we can affirm, or not, these results and reject, or not, our hypothesis 4.

Table 21. Average confidence levels and cSP index by period – Group Treatment

	Period	**Obs**	**Mean**	**SD**	**Min**	**Max**
Confidence levels	3	14	74,64	15,25	40	100
	6	14	60,50	27,53	8	95
	9	14	63,43	21,82	8	90
	12	14	71,57	22,63	7	100
Total	**[3,6,9,12]**	**14**	**67,54**	**21,81**	**16**	**96**
cSP Index	3	14	0,07	2	-3	3
	6	14	0,07	2	-3	3
	9	14	-0,71	2	-3	3
	12	14	-0,79	2	-5	2
Total	**[3,6,9,12]**	**14**	**-0,34**	**1,83**	**-3,50**	**2,75**

Table 22. Average confidence levels and cSP index by period – Control Group

	Period	Obs	Mean	SD	Min	Max
Confidence levels	3	56	65,52	19.47	6	100
	6	57	65,88	20.64	5	100
	9	57	65,95	18.24	6	100
	12	57	68,00	20.71	6	100
Total	**[3,6,9,12]**	**57**	**66,34**	**19.77**	**5,75**	**100**
cSP Index	3	56	0,18	1,83	-3	5
	6	57	-0,04	1,96	-4	5
	9	57	0,11	1,84	-4	5
	12	57	0,16	1,81	-4	5
Total	**[3,6,9,12]**	**57**	**0,10**	**1,86**	**-3,75**	**5**

In the following sections, to verify whether the differences encountered in the previous sections are statistically significant, we will carry out different statistical measures.

7. TESTING HYPOTHESIS

7.1 Validating Hypothesis 1

In previous sections we mentioned that overconfidence appears in many areas of our daily lives, it is manifested in different ways and for many reasons. However, we identified that there was a huge gap in the literature about relating overconfidence levels in general knowledge test and overconfidence levels in financial markets. Therefore, with the objective of filling in the gap in the literature, we postulated the first hypothesis, which is as follows:

Hypothesis 1: There is a correlation between overconfidence in general knowledge questions and overconfidence in financial decisions.

> H_0: Overconfidence in general knowledge questions is associated to overconfidence in financial decisions.
> H_A: Overconfidence in general knowledge questions is independent from overconfidence in financial decisions

We expect to have a high and positive correlation among variables which would indicate that the variables measure the same characteristic. In other words: participants being overconfident in general knowledge questions are also overconfident when making financial decisions. To do the correlation, we used the Pairwise Correlation Matrix to evaluate the strength and direction of the relationship between confidence indicators: better than the average effect (BTA) and the calibration self-placement (cSP) index of the test, with the average difference in position (cSP) by ID and the average confidence in self-positioning by ID of the laboratory experiment.

Table 23 shows the Pairwise correlation coefficients between confidence indicators. The results are as follows:

As we can see in Table 23 there are four statistically significant values at 5%. The first significant value at .05 level exists between the variable: Better than the average effect – Test (BTA of the test), and the Calibration self-placement (cSP) index – test. Both variables have 0.745 of correlation. This value is indicating us that there is a high correlation among the variables, as the closer the correlation coefficient is to +1, the more patent this covariation will be. In addition to being a high correlation, it is also positive, which it is evident since both variables are measured in the same way. This positive relationship indicates that both variables are directly correlated.

The second statistically significant correlation at 5% is the correlation between the BTA of the test and the Average difference in position cSP by ID of the Lab. experiment, with a significant linear correlation of -0.1695. This value is indicating that there is a significant negative correlation between both variables, although it is very low.

The same pattern happens between the average difference in position cSP by ID of the Lab. experiment and the Calibration self-placement (cSP) index of the Test. There is a significant negative correlation at 5%, however the correlation is very low and far from -1 (correlation at 5%: -0.3253), which means that there is no correlation between both variables in reverse.

The fourth significant correlation at 5% occurs between the average confidence in self-position by ID of the Lab. experiment variable, and the Calibration self-placement (cSP) index of the Test. Both variables have a positive correlation with each other, however the correlation is very low (0.2116). This low positive correlation is indicating that both variables are directly correlated, but not all the high values of one variable correspond to the high values of the other variable, and the same with the low values.

Therefore, regarding the postulated hypothesis and in accordance with the results obtained from the Pairwise Correlation of Coefficients, we can conclude that we have enough evidence to reject, at p-value < 0.05, the null hypothesis.

Table 23. Pairwise correlation coefficients of confidence indicators

	(1) Better than the average effect - Test	(2) Calibration self-placement (cSP) index - Test	(3) Avg. difference in position cSP by ID - Lab. Experiment	(4) Avg. confidence in self-position by ID - Lab. Experiment
Better than the average effect - Test	1.0000			
Calibration self-placement (cSP) index - Test	0.7455* 0.0000	1.0000		
Avg. difference in position cSP by ID - Lab. Experiment	-0.1695* 0.0477	-0.3253* 0.0001	1.0000	
Avg. confidence in self-position by ID - Lab. Experiment	0.0425 0.6221	0.2116* 0.0131	-0.1119 0.1881	1.0000

p < 0.05 *

7.2 Validating Hypothesis 2

The literature reviewed suggests that an overconfident person tends to under-estimate the risks that a financial decision may imply and, at the same time, they tend to over-estimate the expected earnings which, in general, leads people to purchase excessively, incurring high transaction costs that reduce profitability and lack of different investment portfolios (Michel M. Pompian 2012). This under- and over-estimations can end up in excessive trading (Odean 1998), asset bubbles (Scheinkman and Xiong 2003), excess price volatility and the forward premium puzzle (Burnside et al. 2011). Thus, from this information we postulated hypothesis 2, which tries to answer if there exists a correlation between over-confidence and riskier financial decisions in the asset market. The second hypothesis is as follows:

Hypothesis 2 (H_2): There is a correlation between overconfidence and riskier financial decisions in asset markets.

H_0: There is a relation between being overconfident and make riskier financial decisions.

H_A: There is no relation between being overconfident and make riskier financial decisions.

To understand whether there is a relation between being overconfident and take riskier financial decisions, we run two regressions in which we use as dependent variables the standard deviation of the selling price by ID and the standard deviation of the buying price by ID, and as independent variables: the average confidence in self-positioning by ID, the average difference in position (cSP) by ID, the calibration self-placement index (cSP), and the better than the average effect (BTA). We reported the results in two different tables for a better understanding of them.

Table 24 shows the relationship between confidence indicators and risk indicators for the standard deviations of selling prices. It is observable that none of the independent variables has a statistically significant coefficient, at any level (* $p<0.10$, ** $p<0.05$, *** $p<0.01$), which indicates a disassociation between the confidence and the risk variables.

Table 24. Relationship between Confidence and Risk indicators - Selling Price

	(1) Std. dev. sell price by ID	(2) Std. dev. sell price by ID	(3) Std. dev. sell price by ID	(4) Std. dev. sell price by ID
Avg. confidence in self-position by ID	-0.515			
	(-0.98)			
Avg. difference in position CsP by ID		4.554		
		(0.61)		
Calibration self-placement (cSP) index			6.995	
			(0.91)	
BTA effect				5.069
				(0.27)
Constant	202.6***	166.8***	167.7***	165.3***
	(5.46)	(18.40)	(18.27)	(13.55)
R-Squared	0.007	0.003	0.006	0.001
Obs.	133	133	130	130
AIC	1614.680	1615.273	1579.929	1580.692
Pseudo-Log Likelihood	-805.340	-805.637	-787.964	-788.346

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Furthermore, Table 25 reports the results of the second regression where the dependent variable is the standard deviation of the buying price by ID, and the independent variables are the same confidence indicators previously mentioned. These results show that none of the independent variables has a statistically significant coefficient at any level (* p<0.10, ** p<0.05, *** p<0.01) which, again, indicates that there is no association between the confidence and risk variables.

At this point, we can see that the results are consistent and robust, there is no correlation between overconfidence and risk. The conclusion we obtain from these regressions is that, overconfidence indicators do not influence if a financial decision is riskier, or not. This conclusion, was correlated with the data collected in section 6.2.2, in which results showed a dissociation between overconfidence and performance feedback as a method to directly eliminate overconfidence levels mentioned by Plous (1995), as well as with the study of Slovic et al. (1982) where he found out that overconfident people could be better calibrated when feedback was provided. Thus, according to the information obtained from the regressions and the data previously collected, we have enough evidences to reject the null hypothesis (H_0).

Table 25. Relationship between Confidence and Risk indicators - Buying Price

	(1) Std. dev. buy price by ID	(2) Std. dev. buy price by ID	(3) Std. dev. buy price by ID	(4) Std. dev. buy price by ID
Avg. confidence in self-position by ID	-0.119			
	(-0.25)			
Avg. difference in position CsP by ID		-0.0152		
		(-0.00)		
Calibration self-placement (cSP) index			2.106	
			(0.29)	
BTA effect				-1.576
				(-0.09)
Constant	159.4***	151.3***	152.3***	153.1***
	(4.78)	(17.87)	(17.82)	(13.17)
R-Squared	0.001	0.000	0.001	0.000
Obs.	127	127	124	124
AIC	1518.786	1518.850	1483.605	1483.684
Pseudo-Log Likelihood	-757.393	-757.425	-739.802	-739.842

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

However, we wanted to go further and investigate whether the fact of being female, or male, could have an impact on the trading behaviour. To do so, we ran a linear regression where the dependent variable was the standard deviation of the selling and buying price by ID. Regarding the independent variables, we considered the dummy variable gender (0 = female, 1 = male), and the overconfidence variables: average confidence in self-positioning by ID and average difference in position (cSP).

Table 26 summarizes the information obtained by the linear regression. We can clearly observe that there is a statistically significant correlation at 1% (55.85***) between the dependent variable ‘standard deviation selling price’ and the gender variable ‘male’.

Additionally, the result can be interpreted in the following way; if the individual is male, the average difference in position increases and, in consequence, the risk increases as well. Thus, we can observe that the factor gender has an impact on participants’ trading behavior, which makes it an interesting application for further research.

Table 26. Relationship Overconfidence and Risk - Gender

	(1) Std. dev. sell price by ID	(2) Std. dev. buy price by ID
Avg. confidence in self-position by ID	-0.688	-0.175
	(-1.20)	(-0.32)
Avg. difference in position cSP by ID	3.423	0.503
	(0.44)	(0.06)
Female	0	0
	(.)	(.)
Male	55.85***	9.583
	(2.82)	(0.49)
Constant	185.9***	158.0***
	(4.56)	(4.06)
R-Squared	0.087	0.003
Obs.	102	97
AIC	1230.715	1162.145
Pseudo-Log Likelihood	-611.358	-577.072

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

7.2 Validating Hypothesis 3

The third hypothesis is based on the literature reviewed of overconfidence and individual feedback availability. As previously mentioned, subjects can be trained to be better calibrated by constantly receiving feedback and repeating tasks several times. However, feedback could only be used as a mechanism to reduce overconfidence when tasks were 'hard or difficult'. Thus, since financial decisions are considered to be a 'hard/difficult' task we postulated the third hypothesis which tries to answer whether individual feedback could reduce overconfidence levels when financial decisions were done. The third hypothesis is as follows:

Hypothesis 3 (H_3): Overconfidence is reduced when individual feedback is provided

H_0: Individual Feedback reduces overconfidence
H_A: Individual Feedback does not reduce overconfidence

In this section, we are going to run different statistical tests to reject or not the hypothesis above mentioned. The section will be structured in the following way. First, we will elaborate a t-test with a T-student distribution. Second, we will study whether there is a correlation between spread and overconfidence to identify if the ones who are more overconfident have also higher spreads when feedback is provided and, to further support hypothesis 3, we will analyse the relationship between the feedback treatment and the average confidence in self-position by ID, with the purpose to clarify the results obtained in the double auction market.

As previously mentioned, the first statistical test will be the t-test for both variables: confidence about self-reported position (in %) and the difference self-reported minus the true position. The results are shown in Table 27, which summarizes the number of observations for each treatment, the mean for each treatment, and the t statistics and the p-value for each variable. The *diff* is defined as mean(0) - mean(1), being 0 if they belong to the Control group, and 1 if they belong to the Feedback treatment. Thus, the alternative hypothesis which is diff < 0 is also the hypothesis that the mean of confidence (in %) and cSP index, for the Feedback treatment is greater than the mean of confidence (in %) and cSP index for the Control group. All the probabilities are well above 0.05 which means that, no matter which alternative hypothesis we select to test, we will not reject the null hypothesis which says that the mean level of confidence (in %) and the cSP index for the Feedback treatment and the Control group is the same.

Table 27. Summary of the t-test (Confidence (in %) and cSP index)

	N	Confidence - Self-reported position (in %)	cSP index
Feedback treatment	69	66.96	0.145
Control Group	71	66.34	0.101
t statistics		-0.306	-0.260
p-values			
Ha: diff < 0		0.379	0.397
Ha: diff != 0		0.759	0.795
Ha: diff > 0		0.620	0.602

Furthermore, in order to identify whether overconfidence was influencing the spread, when feedback was provided, we conducted a linear regression where the dependent variable is represented by the Spread (difference among selling and buying price, in euros). The independent variables are; the Feedback treatment variable, and the confidence variables; confidence about self-reported position (in %) and difference self-reported minus true position (cSP).

Table 28 is reporting the linear regression's results. The dependent category is represented by the Spread, basically meaning that, the reference is the difference among prices, where the constant (36.38) is the average variance. Each coefficient is telling us the magnitude of the variation in variance in each independent variable compared to the reference variable (spread).

We can clearly observe that there is a negative statistically significant correlation at 0.05 level (-2.326**) between the dependent variable 'Spread' and the confidence variable 'Confidence about self-reported position (in %)'.

This model can be understood as, for each overconfident individual the spread decreases by 2.326. In other words; when feedback is provided and individuals are overconfident, their spread is lower which means that the risk is minimized.

Table 28. Relationship Spread and Overconfidence when Feedback is provided

	(1) Difference among selling and buying price, in euros (Spread)
Confidence about self-reported position (in %)	-2.326** (-2.14)
Difference self-reported minus true position CsP	-9.198 (-0.73)
Feedback treatment	12.79 (0.29)
Constant	36.38 (0.48)
R-Squared	0.038
Obs.	131
AIC	1819.066
Pseudo-Log Likelihood	-905.533

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Going further with the analysis and in order to identify whether the null hypothesis can be rejected, or not, we conducted two linear regressions where the dependent variables are represented by the confidence about self-reported position (in %) and the difference self-reported minus the true answer, and the independent variable is going to be a dummy variable represented by feedback. Both regressions are represented in the same table (Table 29) for a better understanding of the results. Therefore, after using Stata to fit a regression model, and verify the fit by checking the results, we interpreted the p-values and the coefficients that appeared in the output of the linear regression analysis.

Table 29. Relationship between Confidence, cSP and Individual Feedback

	(1) Confidence about self-reported position (in %)	(2) Difference self-reported minus true position (cSP index)
Feedback treatment	0.626 (0.31)	0.0437 (0.26)
Constant	66.34*** (47.41)	0.101 (0.88)
R-Squared	0.000	0.000
Obs.	427	427
AIC	3817.192	1681.691
Pseudo-Log Likelihood	-1906.596	-838.845

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 29 shows the results of the linear regression between overconfidence and individual feedback. The dependent category is represented by the 'Confidence about self-reported position (in %), basically meaning that, the reference is the confidence about the position reported in percentage, where the constant (66.34) is the average variance. Therefore, the feedback treatment coefficient is telling us the magnitude of the variation compared to the reference variable. Additionally, the other dependent category is represented by the 'Difference self-reported minus the true position' meaning that, the reference is the cSP index, where the constant (0.101) is the average variance.

We can clearly see that, in both cases, the coefficient is higher than alfa (* p<0.10, ** p<0.05, *** p<0.01) which means that it is not statistically significant at any level. Therefore, we cannot identify any correlation between performance feedback and these two confidence indicators. However, before rejecting our hypothesis 3, we decided to check other types of regressions to have a more robust conclusion.

Table 30. Relationship between Confidence, cSP and individual Feedback - xtreg

	(1) Confidence about self-reported position (in %)	(2) Difference self-reported minus true position (cSP)
Feedback treatment	0.578 (0.17)	0.0401 (0.16)
Constant	66.39*** (28.68)	0.105 (0.62)
R-Squared		
Obs.	427	427
AIC	.	.
Pseudo-Log Likelihood		

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 30 shows the results of an xtreg regression between overconfidence and performance feedback, and another xtreg regression between cSP and individual feedback. The data from the separated regressions were integrated to have a better understanding of the results.

In the first xtreg regression, the dependent category is represented by the 'confidence about self-reported position (in %)', meaning that the reference is the confidence self-reported position in percentage, where the constant (66.39) is the average variance. Therefore, the feedback treatment coefficient is, again, telling us the magnitude of the variation compared to the reference variable. We can clearly appreciate

that the coefficient (0.578) is almost the same as in Table 29, which means that the coefficient is not statistically significant at any level.

On the second xtreg regression, the dependent category is represented by the 'Difference self-reported minus the true position, cSP', meaning that the reference is the cSP, where the constant (0.105) is the average variance. Therefore, the feedback treatment coefficient is telling us the magnitude of the variation compared to the cSP variable. Again, we can clearly observe that the coefficient (0.0401) is not statistically significant at any level, which means that there is no relation between feedback and cSP.

Finally, we decided to run a first-difference estimator regression, which avoids biases that appear because of omitted, and time-invariant, variables by using repeated observations over time, to finally conclude if we can reject, or not, our third hypothesis.

Table 31. Relationship Overconfidence, cSP and Feedback - First differences

	(1) D.Confidence about self-reported position (in %)	(2) D.Difference self-reported minus true position (cSP)
Feedback treatment	1.698 (0.90)	-0.0482 (-0.24)
Constant	0.688 (0.53)	-0.0118 (-0.09)
R-Squared		
Obs.	427	427
AIC	.	.
Pseudo-Log Likelihood		

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 31 shows the relationship between differences on confidence about self-reported positions (in %) with performance feedback, and the relationship between differences in the cSP and performance feedback. We can clearly observe that, in both regressions, the coefficient is not statistically significant, being 1.698 for regression (1), and -0.0482 for regression (2), which means that there is no relationship between the dependent and independent variable.

For that reason, with the regressions reported in Tables 29, 30, and 31, which have robustly proved that there is no relationship between confidence indicators and individual performance feedback, plus the results obtained from the data collected from the laboratory experiment, we have enough evidences to reject the null hypothesis (H_0).

7.2 Validating Hypothesis 4

The fourth hypothesis is based on the group-decision making and overconfidence literature. As previously mentioned, several studies have explored mechanisms to reduce overconfidence, but few were focused on overconfidence levels when decisions were done in groups of two. Even so, those studies that focused on group judgements found that individuals making decisions in groups had the tendency to be more confident and reduce their accuracy when making decisions than those taking decisions individually (Plous 1995). However, such evidence has not been proven in the financial area. Thus, we postulated a fourth hypothesis where we try to answer if confidence levels can be reduced when decisions are done in groups of two and, at the same time, if accuracy can increase when decisions are taken in groups of two as well. The hypothesis was as follows:

Hypothesis 4 (H_4): When decisions are taken in groups, there is a decrease in confidence and an increase of accuracy.

H_0: Decisions taken in groups of two decrease confidence levels
H_A: Decisions taken in groups of two does not decrease confidence levels

H_0: Decisions taken in groups of two increases accuracy
H_A: Decisions taken in groups of two does not increase accuracy

In this section, we are also going to run different statistical tests to reject, or not, the hypothesis above mentioned. The section will be structured in four parts. First, we will elaborate a t-test with a T-student distribution. Second, we will analyse whether the spread and the overconfidence variables are correlated, to identify if the ones who are overconfident have also higher spreads when making decisions in groups of two. Third, we will analyse the relationship between the Group treatment and the confidence about self-reported position (in %), with the purpose to clarify the results obtained in the double auction market and to investigate whether the first part of the hypothesis can be rejected, or not. Fourth, to further support hypothesis 4, we will run a second regression between Group treatment and the accuracy variable 'difference self-reported minus true position (cSP)'.

As previously mentioned, the first statistical test will be the t-tes for both variables: average confidence in self-positioning by ID and the average difference in position by ID (cSP). The results are show in Table 32, which summarizes the number of observations and the mean for each treatment, the t statistics and the p-value for each variable. The

diff is defined as mean(0) - mean(1), being 0 if they belong to the Control group, and 1 if they belong to the Group treatment.

So, the alternative hypothesis which is diff < 0 is also the hypothesis that the mean of confidence (in %) and cSP index, for the Group treatment is greater than the mean of confidence (in %) and cSP index for the Control group. All the probabilities are above 0.05 which indicates that, no matter which alternative hypothesis we select to test, we will not reject the null hypothesis which says that the mean level of confidence (in %) and the cSP index for the Group treatment and the Control group is the same.

Table 32. Summary of the t-test (Avg. confidence and cSP index)

	N	*Confidence about self-reported position (in %)*	*Difference self-reported minus true position (cSP)*
Group treatment	14	67.54	-0.339
Control Group	58	66.34	0.101
t statistics		-0.396	1.598
p-values			
Ha: diff < 0		0.346	0.944
Ha: diff != 0		0.692	0.111
Ha: diff > 0		0.654	0.055

Furthermore, once having compared the means, we analysed whether overconfidence was influencing the spread when making decisions in groups of two. To do so, we conducted a linear regression where the dependent variable is represented by the Spread (difference among selling and buying price, in euros), and the independent variables which are the confidence variables are represented by the confidence about self-reported position (in %) and difference self-reported minus true position (cSP), and the Group treatment variable.

Table 33 is reporting the linear regression's results. The dependent category is represented by the Spread, basically meaning that, the reference is the difference among prices, where the constant (45.21) is the average variance. Each coefficient is telling us the magnitude of the variation in variance in each independent variable compared to the reference variable (spread).

We can clearly observe that there is a negative statistically significant correlation at 10% (-2.461*) between the dependent variable 'spread' and the confidence variable 'confidence about self-reported position (in %)'. This model can be understood as, for each overconfident individual the spread decreases by 2.461. In other words; when decisions are done in groups of two and individuals are overconfident, their spread is

lower which means that the risk is minimized.

Table 33. Relationship Spread and Overconfidence in Groups

	(1) Difference among selling and buying price, in euros (Spread)
Confidence about self-reported position (in %)	-2.461* (-1.83)
Difference self-reported minus true position CsP	-9.285 (-0.67)
Group treatment (pairs)	104.0 (1.43)
Constant	45.21 (0.49)
R-Squared	0.055
Obs.	92
AIC	1273.516
Pseudo-Log Likelihood	-632.758

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Going with the third part of the analysis and, in order to identify whether the null hypothesis can be rejected, or not, we conducted two linear regressions. In the first regression the dependent variable is represented by the confidence about self-reported position (in %), and the independent variable is going to be a dummy variable represented by pairs (group treatment). After using Stata to fit a regression model, and verify the fit by checking the results, we interpreted the p-values and the coefficients that appeared in the output of the linear regression analysis.

Therefore, Table 34 reports the linear regression results. The dependent category is represented as the average confidence in self-positioning by ID, which means that the reference is confidence average where the constant (66.34) is the average variance. We can clearly observe that the coefficient is not statistically significant at any level, telling us that the variance does not increase, or decrease, when decisions are done in groups of two (there is no relationship among variables). Thus, with the regression results and the previous analysis of the laboratory experiment, we have enough evidences to reject the null hypothesis *(H_0: Decisions taken in groups of two decrease confidence levels).*

Table 34. Relationship Average Confidence and Group decision making

	(1) Confidence about self-reported position (in %)ID
Group treatment (pairs)	1.197
	(0.40)
Constant	66.34***
	(49.37)
R-Squared	0.001
Obs.	72
AIC	2507.598
Pseudo-Log Likelihood	-1251.799

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Going further, in Table 35 we reported the linear regression results for the second part of the hypothesis. The dependent variable is represented as the Difference self-reported minus true position (cSP), which means that the reference is the cSP index where the constant (0.101) is the average variance of the variable. So, again, the coefficient is telling us the variation magnitude of the variance for the Group treatment.

We can see that the coefficient is not statistically significant (-0.441) at any level, which means that there is no relationship among both variables.

Table 35. Relationship Accuracy and Groups

	(1) Difference self-reported minus true position (cSP)
Group treatment (pairs)	-0.441
	(-1.60)
Constant	0.101
	(0.77)
R-Squared	0.009
Obs.	72
AIC	1152.571
Pseudo-Log Likelihood	-574.285

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Thus, concluding with hypothesis four, we have enough evidences to reject the null hypothesis *(H_0: Decisions taken in groups of two increases accuracy).*

8. DISCUSSION AND CONCLUSION

Our research covers four main purposes. The first purpose is to determine if people who are confident in general knowledge questions are also confident when making financial decisions. Previous studies attempt to understand whether participants who were overconfident on interval production tasks were most likely to behave in an experimental financial market. Bias et al. (2005) states that participants who showed high overconfidence on interval production tasks had more possibilities to lose money in an experimental financial market when a situation was highly ambiguous, where the fundamental value was low and was not being reflected in the traded price of the assets.

Additionally, Bruno Biais, Denis Hilton, Karine Mazurier and Sébastien Pouget (2004) examined the degree of overconfidence in judgement (as miscalibration) and also observed the behaviour of participants in an experimental financial market with identical information where they found that miscalibration was reducing trading performance.

Nevertheless, the main issue regarding these papers was the fact that none of them attempted to study whether being overconfident in an interval production task was related to being overconfident in an experimental financial market. Despite the evidence that the research in finance also uses both extremes; miscalibration and other forms of overconfidence (better than the average effect, positive illusions and unrealistic optimism), implicitly or explicitly, to estimate predictions about the effect of individual overconfidence in economic decisions, no previous studies have managed to achieve its causality.

Our contribution was to address the question with an interval production scale test and a laboratory experiment. To be more precise, we designed two calibration tests of 11 general knowledge questions, where participants were asked to reveal a lower and upper bound for the 90% confidence interval, and each calibration scale was followed by two questions designed to capture overconfidence when self-placing to measure the BTA effect with the calibration self-placement (cSP) index. Regarding the laboratory experiment, we designed an artificial double auction market in which different shares were traded during 15 periods and, in period 3, 6, 9, and 12 participants had to insert their prediction with respect to the position they were compared to other players, and the percentage of how confident they felt with the answer previously provided. Thanks to our designs, we are able to identify if there is a link, and to contribute to this part of the literature.

From the analysis, three statistically significant results were obtained. The first one is that the average difference in position by ID of the laboratory experiment and the better than the average effect of the test are inversely correlated, indicating that the high values

of one of the variables were usually corresponding to the low values of the other variable, and vice versa. The second result obtained is that the average difference in position by ID of the laboratory experiment and the calibration self-placement of the test, which are also inversely correlated, meaning that there is no correlation between both variables. The third result, is related the average confidence in self-position by ID of the laboratory experiment and the calibration self-placement of the test, which are positively correlated but with a very low correlation, indicating that not all the high values of one variable are corresponding to the high values of the other variable, and the same with the low values.

Generally speaking, the majority of students' answers about general knowledge questions were falling outside the range for more than 10% of the time. The usual proportion of surprises is between 40 and 60%, in our case it was 72%, it is a very huge percentage in comparison to the usual proportion of surprises and, even more, if we take into account that an incentive was provided. It has also to be said that it was a small monetary incentive (3 EUR), compared to the incentive provided in the lab experiment. It is possible, that if the incentive would have been more significant, (i.e., being related to the grade of the course) it could have happened that students would have tried to provide more accurate answers, however it is something that we cannot assume, or affirm, because we do not have evidences of it, it is just a possible explanation for such huge percentage. Thus, after having analysed the results, we are able to contribute to this part of the literature by showing that overconfidence in general knowledge questions is independent from overconfidence in financial decisions.

The second main purpose of this study is to determine if there is a correlation between overconfidence and making riskier decisions in financial markets. The literature reviewed suggests that overconfidence is generally defined as an overestimation of one's own knowledge, or the precision of the private information. However, it is also considered as an underestimation of the variation of signals or the volatility of asset values (Skała 2008). Additionally, previous studies have encompassed the idea that an overconfident person tends to under-estimate the risks that a financial decision may imply and, at the same time, they tend to over-estimate the expected earnings which, in general, lead people to purchase excessively, incurring high transaction costs that reduce profitability and lack of different investment portfolios (Michel M. Pompian 2012). This under- and over-estimations can end up in excessive trading (Odean 1998), asset bubbles (Scheinkman and Xiong 2003), excess price volatility and the forward premium puzzle (Burnside et al. 2011). Furthermore, a study conducted by Ackert et al. (2009) found that the higher the overconfidence, the higher the trading volume and the lower the earnings.

Thus, we have evidences suggesting that overconfidence and riskier decisions are linked. To identify the existence of this link, we used the same laboratory experiment (as previously mentioned), with the same artificial double auction market design in which different shares were traded during 15 periods. Selling and buying prices were all the time exposed in the market interface and, in period 3, 6, 9, and 12 participants had to insert their prediction with respect to the position they were compared to other players, and the percentage of how confident they felt with the answer previously provided. Thanks to this design, we are again able to identify if there is a link between overconfidence and risk.

From this study, we did not obtain any statistically significant correlation among variables. The results obtained just indicated that there is no relationship between overconfidence and making riskier decisions. Thus, the results obtained are not in line with previous papers that studied overconfidence and riskier decisions in financial markets (Skala, 2008; Michel M. Pompian, 2012; Odean, 1998; Scheinkman and Xiong, 2003; Burnside et al. 2011; Ackert et al., 2009).

The third main purpose of this study is to determine if overconfidence is reduced when feedback is provided. Several authors that studied the effects of feedback on overconfident people found that people who were initially overconfident could learn to be better calibrated (reduce overconfidence) after making many decisions and receiving constant performance feedback (Slovic et al., 1982). Other authors found that overconfidence could be directly eliminated by proving performance feedback to participants during the experiment (Plous, 1995), and that overconfidence could improve performance - e.g. by maximizing payoffs (Berlin and Dargnies, 2016). Unlike the previously mentioned authors, Oskamp (1965), studied overconfidence as a simply excess of confidence over accuracy, and concluded that accuracy did not increase when feedback was provided, but confidence increased steadily and significantly. However, the discussion about whether overconfidence is a constant feature (steady) or a dynamic process liable to manipulation has not given conclusive answers to date (Skała, 2008).

Thus, since there are contradictory evidences we hypothesized that overconfidence is reduced when performance feedback is provided, since we consider performance feedback as one of the possible mechanisms to reduce overconfidence.

To identify if there is a link between overconfidence and feedback we used the same laboratory experiment, previously mentioned. However, this time, in period 3 – 6 – 9 and 12 participants had to insert their prediction with respect to the position they were compared to other players, and the percentage of how confident they felt with the answer previously provided, and after everyone had provided their answers, the software was

automatically providing them their real position in the market. This method allowed us to examine whether there is a link, or not, between both variables.

To have a robust result we ran different regressions however, none of them gave us statistically significant results. Therefore, the only result obtained from this study is that overconfidence is not reduced when performance feedback is provided. Thus, our null hypothesis is rejected and these results are not in line with previous studies that found that performance feedback was reducing overconfidence (Slovic et al., 1982; Plous, 1995; Berlin and Dargnies, 2016).

Our last main purpose was to determine if overconfidence decreases and accuracy increases when decisions are done in groups of two. Previous studies attempt to understand whether overconfidence increases when decisions are done in groups of two. They suggest that asking other people's opinion is not likely to reduce very high levels of overconfidence. In fact, they consider that interacting in groups just invites people to create explanations for their own beliefs, which at the end provokes that confidence levels in someone's own predications increase (Heath and Gonzalez, 1995). Other authors such as Bang et al. (2014) also described and evaluated the process through groups by combining their individual judgement and their determined confidence level into a group judgement with different levels of confidence, and the results suggested that groups had the tendency to be more confident than individuals when making decisions.

In the same experiment were Heath and Gonzalez (1995) found that group decision making was increasing overconfidence levels, they also found out that 98% of participants though that the accuracy of their group was above the median. This result is to be expected if the confidence level of the group improves more than the accuracy. This results was also proved by Sniezek (1992). Sniezek (1992) suggests that single person forecasts should be less accurate than group's forecasts, because group members believe that their forecast discussions are more accurate. Additionally, Boje and Murnighan (1982) proved that the confidence of the group's members increased when they did many trials, but their accuracy decreased.

Nevertheless, the main issue regarding these papers is that such evidences have not been proved in the financial area, thus we still considered group decision making as a possible mechanism to reduce overconfidence. For that reason, we hypothesized that overconfidence decreases and accuracy increases when decisions are done in groups of two.

Our contribution was to address these questions by using the same laboratory experiment, previously mentioned. To be more precise, the design of the artificial double

auction market and the game system was exactly the same but, this time, participants had to make decisions in groups of two. This design allowed us to identify if there is a link among variables. Two main results were obtained from this study. The first result is related to overconfidence and group decision making. The results indicate that there is no relationship among variables, which means that group decision making and overconfidence are independent variables. The second result, which is related to the accuracy, indicates that there is, again, no relationship among variables, which means that group decision making and accuracy are independent variables. Thus, after having analysed the results, we are able to contribute to this part of the literature by showing that group decision making does not decrease overconfidence levels, nor increase accuracy, in financial markets. In other words; overconfidence and accuracy are independent from group decision making.

These findings provided us more insights about participants than about any other factor. As previously mentioned, the majority of students' answers about general knowledge questions were falling outside the range for more than 10% of the time, fact that could have been driven by the lack of motivation, or knowledge. However, we cannot assume, or affirm, such claim since we did not investigate the reasons of such event.

As follows, regarding the lab experiment and from a strategical perspective, we may consider that subjects had the tendency to decide between two main strategies during the double auction market; 1) going for the money, focusing on positioning themselves on very low positions with respect to others, 2) going for the grade, by completely focusing on the trading of shares. However, it cannot be assumed that subjects were following one, or the other, strategy. Nevertheless, we could observe a tendency to decline for one strategy, or the other. The most significant impact of this behaviour was that those going for the money were prone to barely participate in the double auction market.

Another observation gathered from the behaviour during sessions was that those subjects making decisions in groups of two were prone to convince themselves that they were doing well and, there was always one subject who was having more initiative than the other one. However, this behaviour is very common when working in groups of two. Finally, a last observation gathered from sessions was that those subjects who were not receiving feedback had the tendency to follow a more stable pattern, with less ups and downs in shares' prices.

The key point is that people have the tendency to consider themselves as worse than the average when answering questions about general knowledge, but have the

tendency to overvalue prices and over-estimate their positions in financial markets. Thus, understanding the profile of each subject could be an interesting practical application to, at least, have a better understanding of their behaviour.

This study, however, has some limitations that must be recognized. On the one hand, the lab experiment, as a research methodology, has the advantage of promoting the internal validity of the results. Additionally, as mentioned in previous sections the external validity may be questionable but there is an extensive literature that demonstrates that the results with inexperienced students are not statistically different from the experiments that have been done with people who are familiar with financial concepts.

On the other hand, the use of a sample with young students of the same age and without any knowledge in financial markets can affect the trading price evolution since they have the tendency to overvalue shares' prices in the trading market. It could be interesting to see whether overconfidence levels remained the same with a sample composed by professionals who are dedicated to the finance sector.

9. FURTHER RESEARCH

This study may be a turning point for future studies. The literature, suggested many interesting points such as; the effect of feedback on confidence levels, the tendency to over-value and under-value shares' prices due to overconfidence, about the effect of group work on your levels of overconfidence and accuracy, among many.

However, some authors have found that there could be a gender issue regarding trader's behaviour in asset markets. Thus, it could be an interesting point to take into account if we re-do the experiment with a bigger sample in which we can distribute sessions by gender. Also, another interesting observation could be studying the differences between women and men who have knowledge and experience in the finance sector, to see if the same effect would apply. This question could be answered by implementing a double auction market in which there is a distinction by gender, for example, comparing the results obtained by experienced women in a double auction market compound just by women, with another double auction market were only men are trading.

Furthermore, other authors have found that the feedback effect has an asymmetric effect on trading behaviour by increasing the profitability of buying, and decreases the profitability of selling, giving rise to an endogenous limit to arbitrage, whereby participants may abstain from trading on negative information. Thus, a feedback that provides negative information (a reduction in the profitability of selling) is incorporated more slowly into prices than positive information, leading to over-investment. Therefore, it could be another interesting point to take into account for future research, the study of the effect of performance feedback on trading behaviour.

In general, this research has shed very little light on the gap found in the literature, however it may be a point of reflection for future research since we have already been able to experience the fact of working with students without prior knowledge in financial markets.

REFERENCES

Ackert, Lucy F., Narat Charupat, Richard Deaves, and Brian D. Kluger. 2009. “Probability Judgment Error and Speculation in Laboratory Asset Market Bubbles.” *Journal of Financial and Quantitative Analysis*.

Ackert, Lucy F. and Bryan K. Church. 2001. “The Effects of Subject Pool and Design Experience on Rationality in Experimental Asset Markets.” *Journal of Psychology and Financial Markets*.

Alicke, Mark D. 1985. “Global Self-Evaluation as Determined by the Desirability and Controllability of Trait Adjectives.” *Journal of Personality and Social Psychology*.

Arkes, Hal R., Caryn Christensen, Cheryl Lai, and Catherine Blumer. 1987. “Two Methods of Reducing Overconfidence.” *Organizational Behavior and Human Decision Processes*.

AYTON, PETER and ALASTAIR G. R. McCLELLAND. 1997. “How Real Is Overconfidence?” *Journal of Behavioral Decision Making*.

Baker, Ronald J., Susan Laury, and Arlington W. Williams. 2011. “Comparing Small-Group and Individual Behavior in Lottery-Choice Experiments.” *SSRN Electronic Journal*.

Balzer, William K., Michael E. Doherty, and Raymond O’Connor. 1989. “Effects of Cognitive Feedback on Performance.” *Psychological Bulletin*.

Bang, Dan, Riccardo Fusaroli, Kristian Tylén, Karsten Olsen, Peter E. Latham, Jennifer Y. F. Lau, Andreas Roepstorff, Geraint Rees, Chris D. Frith, and Bahador Bahrami. 2014. “Does Interaction Matter? Testing Whether a Confidence Heuristic Can Replace Interaction in Collective Decision-Making.” *Consciousness and Cognition*.

Barber, Brad M. and Terrance Odean. 1999. “The Courage of Misguided Convictions.” *Financial Analysts Journal*.

Barber, Brad M. and Terrance Odean. 2001. “Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment.” *Quarterly Journal of Economics*.

Bengtsson, Claes, Mats Persson, and Peter Willenhag. 2005. “Gender and Overconfidence.” *Economics Letters*.

Benson, P. George and Dilek Önkal. 1992. “The Effects of Feedback and Training on the Performance of Probability Forecasters.” *International Journal of Forecasting*.

Berlin, Noémi and Marie Pierre Dargnies. 2016. “Gender Differences in Reactions to Feedback and Willingness to Compete.” *Journal of Economic Behavior and Organization*.

Biais, Bruno, Denis Hilton, Karine Mazurier, and Sébastien Pouget. 2005. “Judgemental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental

Financial Market.” *Review of Economic Studies*.

Blavatskyy, Pavlo R. 2009. “Betting on Own Knowledge: Experimental Test of Overconfidence.” *Journal of Risk and Uncertainty*.

Van Boening, Mark V., Arlington W. Williams, and Shawn LaMaster. 1993. “Price Bubbles and Crashes in Experimental Call Markets.” *Economics Letters*.

BOJE, DAVID M. and J. KEITH MURNIGHAN. 1982. “GROUP CONFIDENCE PRESSURES IN ITERATIVE DECISIONS.” *MANAGE SCI*.

Bolger, Fergus and Dilek Önkal-Atay. 2004. “The Effects of Feedback on Judgmental Interval Predictions.” *International Journal of Forecasting*.

Brenner, Lyle A., Derek J. Koehler, Varda Liberman, and Amos Tversky. 1996. “Overconfidence in Probability and Frequency Judgments: A Critical Examination.” *Organizational Behavior and Human Decision Processes*.

Brown, Jonathon D. 2012. “Understanding the Better than Average Effect: Motives (Still) Matter.” *Personality and Social Psychology Bulletin*.

Burnside, Craig, Bing Han, David Hirshleifer, and Tracy Yue Wang. 2011. “Investor Overconfidence and the Forward Premium Puzzle.” *Review of Economic Studies*.

Camerer, Colin and Dan Lovallo. 1999. “Overconfidence and Excess Entry: An Experimental Approach.” *American Economic Review*.

Chambers, John R. and Paul D. Windschitl. 2004. “Biases in Social Comparative Judgments: The Role of Nonmotivated Factors in above-Average and Comparative-Optimism Effects.” *Psychological Bulletin*.

Cheung, Stephen L. and Stefan Palan. 2012. “Two Heads Are Less Bubbly than One: Team Decision-Making in an Experimental Asset Market.” *Experimental Economics*.

Cooper, David J. and John H. Kagel. 2005. “Are Two Heads Better than One? Team versus Individual Play in Signaling Games.” *American Economic Review*.

Correll, Shelley J. 2001. “Gender and the Career Choice Process: The Role of Biased Self-Assessments.” *American Journal of Sociology*.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. 1998. “Investor Psychology and Security Market Under- and Overreactions.” *Journal of Finance*.

Dawes, Robyn M. and Matthew Mulford. 1996. “The False Consensus Effect and Overconfidence: Flaws in Judgment or Flaws in How We Study Judgment?” *Organizational Behavior and Human Decision Processes*.

Dunning, David, Judith A. Meyerowitz, and Amy D. Holzberg. 1989. “Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability.” *Journal of Personality and Social Psychology*.

Durand, Rodolphe. 2003. “Predicting a Firm’s Forecasting Ability: The Roles of Organizational Illusion of Control and Organizational Attention.” *Strategic Management Journal*.

Fischbacher, Urs. 2007. “Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments.” *Experimental Economics*.

Fischhoff, Baruch, Paul Slovic, and Sarah Lichtenstein. 1977. “Knowing with Certainty: The Appropriateness of Extreme Confidence.” *Journal of Experimental Psychology: Human Perception and Performance*.

Gigerenzer, Gerd, Ulrich Hoffrage, and Heinz Kleinbolting. 1991. “Probabilistic Mental Models: A Brunswikian Theory of Confidence.” *Psychological Review*.

Giladi, Eilath E. and Yechiel Klar. 2002. “When Standards Are Wide of the Mark: Nonselective Superiority and Inferiority Biases in Comparative Judgments of Objects and Concepts.” *Journal of Experimental Psychology: General*.

Giusti, Giovanni, Janet Hua Jiang, and Yiping Xu. 2016. “Interest on Cash, Fundamental Value Process and Bubble Formation: An Experimental Study.” *Journal of Behavioral and Experimental Finance*.

Glaser, Markus, Markus Nöth, and Martin Weber. 2008. “Behavioral Finance.” in *Blackwell Handbook of Judgment and Decision Making*.

Glaser, Markus and Martin Weber. 2007. “Overconfidence and Trading Volume.” *GENEVA Risk and Insurance Review*.

Griffin, Dale and Lyle Brenner. 2008. “Perspectives on Probability Judgment Calibration.” in *Blackwell Handbook of Judgment and Decision Making*.

Grinblatt, Mark and Matti Keloharju. 2009. “Sensation Seeking, Overconfidence, and Trading Activity.” *Journal of Finance*.

Harris, D. M. and S. Guten. 1979. “Health-Protective Behavior: An Exploratory Study.” *Journal of Health and Social Behavior*.

Harrison, E. Frank. 1996. “A Process Perspective on Strategic Decision Making.” *Management Decision*.

Heath, Chip and Rich Gonzalez. 1995. “Interaction with Others Increases Decision Confidence but Not Decision Quality: Evidence against Information Collection Views of Interactive Decision Making.” *Organizational Behavior and Human Decision Processes*.

Hilton, Denis, Isabelle Régner, Laure Cabantous, Laetitia Charalambides, and Stéphane Vautier. 2011. “Do Positive Illusions Predict Overconfidence in Judgment? A Test Using Interval Production and Probability Evaluation Measures of Miscalibration.” *Journal of Behavioral Decision Making*.

Hoffrage, Ulrich. 2016. “Overconfidence.” in *Cognitive Illusions: Intriguing Phenomena in Judgement, Thinking and Memory*.

Hribar, Paul and Holly Yang. 2016. “CEO Overconfidence and Management Forecasting.” *Contemporary Accounting Research*.

IRWIN, FRANCIS W. 1953. “Stated Expectations as Functions of Probability and Desirability of Outcomes.” *Journal of Personality*.

Keren, Gideon. 1991. “Calibration and Probability Judgements: Conceptual and Methodological Issues.” *Acta Psychologica*.

Kirchler, Erich and Boris Maciejovsky. 2002. “Simultaneous Over- and Underconfidence: Evidence from Experimental Asset Markets.” *Journal of Risk and Uncertainty*.

Klar, Yechiel and Eilath E. Giladi. 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*.

Kocher, Martin G. and Matthias Sutter. 2006. “Time Is Money-Time Pressure, Incentives, and the Quality of Decision-Making.” *Journal of Economic Behavior and Organization*.

Koriat, Asher, Sarah Lichtenstein, and Baruch Fischhoff. 1980. “Reasons for Confidence.” *Journal of Experimental Psychology: Human Learning and Memory*.

Krizan, Zlatan and Jerry Suls. 2008. “Losing Sight of Oneself in the Above-Average Effect: When Egocentrism, Focalism, and Group Diffuseness Collide.” *Journal of Experimental Social Psychology*.

Krizan, Zlatan and Paul D. Windschitl. 2007. “The Influence of Outcome Desirability on Optimism.” *Psychological Bulletin*.

Kruger, Justin. 1999. “Lake Wobegon Be Gone! The ‘below-Average Effect’ and the Egocentric Nature of Comparative Ability Judgments.” *Journal of Personality and Social Psychology*.

Langer, Ellen J. 1975. “The Illusion of Control.” *Journal of Personality and Social Psychology*.

Levitt, Steven D. and John A. List. 2007. “Viewpoint: On the Generalizability of Lab Behaviour to the Field.” *Canadian Journal of Economics*.

Lichtenstein, S., B. Fischoff, and L. D. Phillips. 1982. “Calibration of Subjective Probabilities: The State of the Art to 1980.” in *Judgement under uncertainty: Heuristics and biases*.

Lichtenstein, Sarah, Paul Slovic, Baruch Fischhoff, Mark Layman, and Barbara Combs. 1978. “Judged Frequency of Lethal Events.” *Journal of Experimental Psychology: Human Learning and Memory*.

Lipkus, I. M., M. Kuchibhatla, C. M. McBride, H. B. Bosworth, K. I. Pollak, I. C. Siegler, and B. K. Rimer. 2000. “Relationships among Breast Cancer Perceived Absolute Risk, Comparative Risk, and Worries.” *Cancer Epidemiology Biomarkers and Prevention*.

Lucy F. Ackert, Richard Deaves, 2010. 2010. “Behavioral Finance: Psychology, Decision-Making, and Markets - Lucy Ackert, Richard Deaves - Google Libros.” Retrieved February 8, 2020 (https://books.google.es/books?hl=es&lr=&id=aSXZGwFPXJMC&oi=fnd&pg=PR7&dq=Behavioral+Finance:+Psychology,+Decision-Making,+and+Markets&ots=sSwGt67PXz&sig=PrvdvVyNlPITSXUL9Wn6gHTloQs#v=onepage&q=Behavioral Finance%3A Psychology%2C Decision-Making%2C and Mar).

Lundeberg, M. A., Fox, P. W., & Punćcohaŕ, J. (1994). 1994. “Highly Confident but Wrong: Gender Differences and Similarities in Confidence Judgments.” *Journal of Educational Psychology* 86(1), 114–121.

Malmendier, Ulrike and Geoffrey Tate. 2005. “Does Overconfidence Affect Corporate Investment? CEO Overconfidence Measures Revisited.” *European Financial Management*.

McClelland, Alastair G. R. and Fergus Bolger. 1994. “The Calibration of Subjective Probability: Theories and Models 1980-94.” *Subjective Probability.*

McGuire, W. J. 1960. “A Syllogistic Analysis of Cognitive Relationships.” in *Attitude organization and change*.

Menkhoff, Lukas, Ulrich Schmidt, and Torsten Brozynski. 2006. “The Impact of Experience on Risk Taking, Overconfidence, and Herding of Fund Managers: Complementary Survey Evidence.” *European Economic Review*.

Michailova, Julija and Ulrich Schmidt. 2016. “Overconfidence and Bubbles in Experimental Asset Markets.” *Journal of Behavioral Finance*.

Michel M. Pompian. 2012. *Behavioral Finance and Wealth Management*.

Mishra, K. C. and Mary J. Metilda. 2015. “A Study on the Impact of Investment Experience, Gender, and Level of Education on Overconfidence and Self-Attribution Bias.” *IIMB Management Review*.

Moore, Don A. and Paul J. Healy. 2008. “The Trouble With Overconfidence.” *Psychological Review*.

Natalia Cassinello Plaza, 2016. n.d. “Sesgo de Exceso de Confianza.” Retrieved December 29, 2019 (https://es.fundspeople.com/news/sesgo-de-exceso-de-confianza).

Odean, Terrance. 1998. “Odean, Terrance (1998): Volume, Volatility, Price and Profit When All Traders Are Above Average, Journal of Finance 53, 1887-1934.” *The Journal of Finance*.

Olsson, Henrik. 2014. “Measuring Overconfidence: Methodological Problems and Statistical Artifacts.” *Journal of Business Research*.

Oskamp, Stuart. 1965. “Overconfidence in Case-Study Judgments.” *Journal of Consulting Psychology*.

Otten, Wilma and Joop Van Der Pligt. 1996. “Context Effects in the Measurement of Comparative Optimism in Probability Judgments.” *Journal of Social and Clinical Psychology*.

Peón, David, Manel Antelo, and Anxo Calvo. 2016. “Overconfidence and Risk Seeking in Credit Markets: An Experimental Game.” *Review of Managerial Science*.

Plous, S. 1995. “A Comparison of Strategies for Reducing Interval Overconfidence in Group Judgments.” *Journal of Applied Psychology*.

Presson, Paul K. and Victor A. Benassi. 1996. “Illusion of Control: A Meta-Analytic Review.” *Journal of Social Behavior and Personality*.

Pulford, Briony D. and Andrew M. Colman. 1997. “Overconfidence: Feedback and Item Difficulty Effects.” *Personality and Individual Differences*.

Robertson, Leon S. 1977. “Car Crashes - Perceived Vulnerability and Willingness to Pay for Crash Protection.” *Journal of Community Health*.

Rockenbach, Bettina, Abdolkarim Sadrieh, and Barbara Mathauschek. 2007. “Teams Take the Better Risks.” *Journal of Economic Behavior and Organization*.

Ross, Michael and Fiore Sicoly. 1979. “Egocentric Biases in Availability and Attribution.” *Journal of Personality and Social Psychology*.

Russo, J. and P. Schoemaker. 1992. “Managing Overconfidence.” *Sloan Management Review*.

Scheinkman, José A. and Wei Xiong. 2003. “Overconfidence and Speculative Bubbles.” *Journal of Political Economy*.

Schweitzer, Maurice E. and Gérard P. Cachon. 2000. “Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence.” *Management Science*.

Schwenk, C. H. 1986. “Information, Cognitive Biases, and Commitment to a Course of Action.” *Academy of Management Review*.

Schwenk, Charles R. 1984. “Cognitive Simplification Processes in Strategic Decision-making.” *Strategic Management Journal*.

Sedikides, Constantine, Deletha P. Hardin, Kenneth C. Herbst, and Gregory J. Dardis.

2002. “Accountability as a Deterrent to Self-Enhancement: The Search for Mechanisms.” *Journal of Personality and Social Psychology*.

Sharot, Tali. 2011. “The Optimism Bias.” *Current Biology*.

Skała, Dorota. 2008. “Overconfidence in Psychology and Finance - an Interdisciplinary Literature Review.” *Bank i Kredyt*.

Slovic, Paul, Baruch Fischhoff, and Sarah Lichtenstein. 1982. “Why Study Risk Perception?” *Risk Analysis*.

Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams. 1988. “Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets.” *Econometrica*.

Sniezek, Janet A. 1992. “Groups under Uncertainty: An Examination of Confidence in Group Decision Making.” *Organizational Behavior and Human Decision Processes*.

Stöckl, Thomas, Jürgen Huber, and Michael Kirchler. 2010. “Bubble Measures in Experimental Asset Markets.” *Experimental Economics*.

T. M. Ostrom and H. S. Upshaw. 1968. *Psychological Foundations of Attitudes*.

Todd, F. J. and K. R. Hammond. 1965. “Differential Feedback in Two Multiple-Cue Probability Learning Tasks.” *Behavioral Science*.

Tversky, Amos and Daniel Kahneman. 1973. “Availability: A Heuristic for Judging Frequency and Probability.” *Cognitive Psychology*.

Tversky, Amos and Daniel Kahneman. 1974. “Judgment under Uncertainty: Heuristics and Biases.” *Science*.

Weinstein, Neil D. 1980. “Unrealistic Optimism about Future Life Events.” *Journal of Personality and Social Psychology*.

ANNEX

Annex 1. Table 14. Confidence average rates by session - Feedback treatment

Table 14. Confidence average rates by session - Feedback treatment

	Session	Mean	SD	Min	Max
	1	65.39	20.16	6	90
	2	70.18	13.64	50	95
	3	44.33	23.23	6	90
	4	65.00	17.32	50	80
	5	60.54	26.31	5	100
	6	77.46	23.24	15	100
	7	67.08	8.91	50	80
	8	75.83	19.29	50	100
	9	76.89	12.41	50	100
	10	54.79	30.22	5	100
Total	[1 to 10]	65.75	20.26	28.70	93.5

Annex 2. Table 15. Confidence average rates by session – Control Group

Table 15. Confidence average rates by session – Control Group

	Session	Mean	SD	Min	Max
	1	73.04	11.81	50	100
	2	67.19	15.96	40	100
	3	63.44	16.80	40	100
	4	52.50	5.00	50	60
	5	60.71	23.16	15	100
	6	68.36	18.84	30	90
	7	66.04	22.76	15	90
	8	66.52	21.29	10	100
	9	60.04	26.36	5	90
	10	75.31	12.04	50	100
Total	[1 to 10]	65.32	17.40	30.50	93

Annex 3. Table 18. Calibration Self-placement by session – Feedback treatment

Table 18. Calibration Self-placement by session – Feedback treatment

	Session	Mean	SD	Min	Max
	1	0.36	1.22	-2	4
	2	-0.36	1.37	-4	3
	3	-0.17	1.80	-4	2
	4	0.50	1.29	-1	2
	5	0.54	1.97	-3	5
	6	0.32	1.18	-2	3
	7	-0.75	1.21	-3	1
	8	-0.67	1.97	-5	2
	9	0.28	1.46	-2	3
	10	0.67	1.88	-2	6
Total	[1 to 10]	0.073	1.53	-2,8	3,1

Annex 4. Table 19. Calibration Self-placement by session – Control Group

Table 19. Calibration Self-placement by session – Control Group

	Session	Mean	SD	Min	Max
	1	-0.14	1.65	-3	3
	2	-0.62	1.93	-4	4
	3	0.19	1.28	-2	2
	4	-0.75	1.26	-2	1
	5	0.64	1.81	-2	5
	6	0.57	1.89	-4	5
	7	0.54	2.64	-4	5
	8	0.39	1.47	-2	3
	9	-0.67	1.66	-4	2
	10	0.31	0.95	-2	2
Total	[1 to 10]	0.05	1.65	-2.90	3.20

Annex 5. TEST 1 - IP Scale

TEST 1 - IP Scale						
Question number	**Questions**	**90% Confidence Range**		**Indicate to what extent you think that you have succeeded at the previous calibration task *(0 = entirely failed, 10 = entirely succeeded).* (C)**	**Indicate to what extent you think that most of the other students have succeeded at the same task *(0 = entirely failed, 10 = entirely succeeded).* (D)**	**True answers (TA)**
		LB (A)	**UB (B)**			
1	**What is Donald Trump's age?**					**73**
2	**The inflation rate in Spain in January 2020 was? (in percentage)**					**1,1**
3	**Population of China in 2019?**					**1433783686**
4	**Number of countries in the world?**					**194**

TEST 1 - IP Scale *(Continuation…)*						
Question number	**Questions**	**90% Confidence Range**		**Indicate to what extent you think that you have succeeded at the previous calibration task *(0 = entirely failed, 10 = entirely succeeded).* (C)**	**Indicate to what extent you think that most of the other students have succeeded at the same task *(0 = entirely failed, 10 = entirely succeeded).* (D)**	**True answers (TA)**
		Lower bound (A)	**Upper bound (B)**			
5	**Number of provinces in Spain?**					**50**
6	**Number of countries which have designated French as their official language**					**29**
7	**How many hectares were burnt during the Amazonas fire in 2019? (in millions)**					**2,5**
8	**Number of languages spoken in the world**					**6500**

TEST 1 - IP Scale *(Continuation…)*						
Question number	**Questions**	**90% Confidence Range**		**Indicate to what extent you think that you have succeeded at the previous calibration task *(0 = entirely failed, 10 = entirely succeeded).* (C)**	**Indicate to what extent you think that most of the other students have succeeded at the same task *(0 = entirely failed, 10 = entirely succeeded).* (D)**	**True answers (TA)**
		Lower bound (A)	**Upper bound (B)**			
9	**Height of the Eiffel Tower (in meters)**					**300**
10	**Number of countries that are members of the EU**					**27**
11	**Altitude of Mont Blanc (in meters)**					**4810**

Annex 6. TEST 2 - IP scale

TEST 2 - IP scale						
Question number	**Questions**	**90% Confidence Range**		**Indicate to what extent you think that you have succeeded at the previous calibration task *(0 = entirely failed, 10 = entirely succeeded).* (C)**	**Indicate to what extent you think that most of the other students have succeeded at the same task *(0 = entirely failed, 10 = entirely succeeded).* (D)**	**True answers (TA)**
		Lower bound (A)	**Upper bound (B)**			
1	**What is Will Smith's age?**					**51**
2	**The annual average of the inflation rate in 2019 in Spain was? (In percentage).**					**0,7**
3	**Population of United States in 2019? (in millions)**					**329,97**

TEST 2 - IP Scale *(Continuation…)*						
Question number	**Questions**	**90% Confidence Range**		**Indicate to what extent you think that you have succeeded at the previous calibration task** ***(0 = entirely failed, 10 = entirely succeeded).*** **(C)**	**Indicate to what extent you think that most of the other students have succeeded at the same task** ***(0 = entirely failed, 10 = entirely succeeded).*** ***(D)***	**True answers (TA)**
		Lower bound (A)	**Upper bound (B)**			
4	**Number of countries in the world?**					**194**
5	**Number of regions in France?**					**18**
6	**Number of countries which have designated Spanish as their official language**					**21**
7	**How many hectares were burnt during the Amazonas fire in 2019? (in millions)**					**2,5**

TEST 2 - IP Scale *(Continuation…)*						
Question number	**Questions**	**90% Confidence Range**		**Indicate to what extent you think that you have succeeded at the previous calibration task *(0 = entirely failed, 10 = entirely succeeded).* (C)**	**Indicate to what extent you think that most of the other students have succeeded at the same task *(0 = entirely failed, 10 = entirely succeeded).* (D)**	**True answers (TA)**
		Lower bound (A)	**Upper bound (B)**			
8	**Number of languages spoken in EU**					**24**
9	**Height of the Statue of Liberty (in meters)**					**93**
10	**Number of countries that are members of the UN 2020?**					**193**
11	**Altitude of Mount Everest (in meters)**					**8848**

Annex 7 - Experiment Instructions

General Instructions

Participants will take part in a behavioural investment experiment. If they follow carefully the instructions, and make the right decisions, they will earn money which will be paid at the end of the experiment (Giusti et al. 2016).

The experiment consists of a sequence of 15 trading periods, each period will last for 120 seconds. During each period, participants will make decisions to invest their money in shares of stocks of an imaginary Company. The currency that we will use in this market is the so-called EURUX (Giusti et al. 2016).

The experiment will be done in the following sequence. First, we will explain participants the trading interface they will use, and then, before starting the real game, they will have time to practice during two periods. After this first step, we will give them further instructions and information to help them to make trading decisions (Giusti et al. 2016).

Additionally, if anything is unclear, during the instructions or practice, participants will be able to raise their hands and ask the administrator all the questions they may have (Giusti et al. 2016).

Trading Interface

In each trading period, they will start with some cash (in EURUX) and a certain number of shares. Shares earn dividends, which will be described later. During each trading period, participants will make investment decisions to buy or sell shares. Here is a sample trading screen (Giusti et al. 2016).

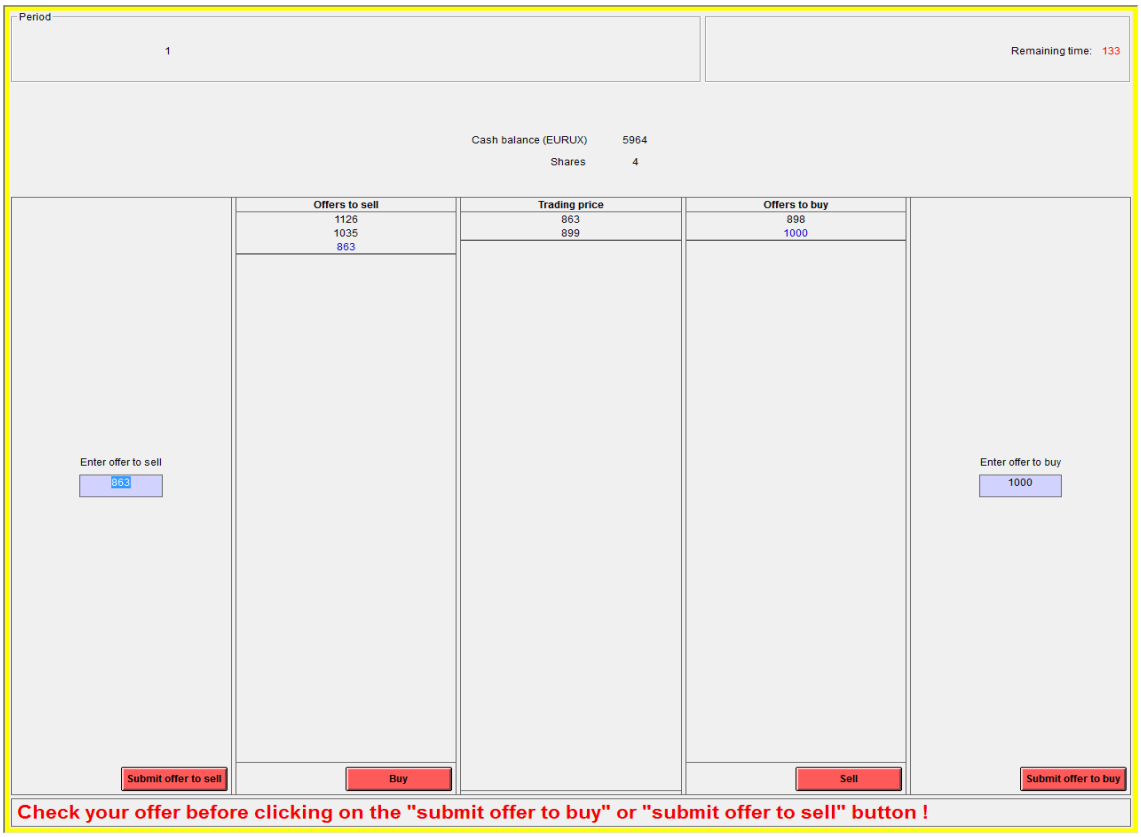


Figure 4. Interface Laboratory Experiment

On the left corner, at the top, participants will be able to see the current trading period in which they are, and on the top right corner participants will be able to see how much time (in seconds) is left for the current period. Their cash balance (in EURUX) and the number of shares they own are shown in the middle of the screen, right above the five columns. On this screen they will be able to buy, or sell shares, in *four* ways (Giusti et al. 2016).

First, they can initiate a sale of shares by submitting an offer to sell.
If they have shares, they may choose to sell them. They can initiate a sale in the text area below “Enter offer to sell” in the first column. Here they can enter the price at which they are offering to sell a share. To send the offer, they have to click the “Submit offer to sell” button. After that, their offer to sell will appear in the second column named “Offers to sell”. Each offer introduced corresponds to one single share. If they want to sell more shares, they have to repeat this process (Giusti et al. 2016).

Important: by submitting an offer to sell, they initiate a sale, but the sale will not be executed until someone accepts it (Giusti et al. 2016).

At this moment they should try to offer to sell a share. They could do so by writing a number (integer) in the text area named “Enter offer to sell” and then clicking on the button “Submit offer to sell”. They will be able to see that a set of numbers will appear in

the column named “Offers to sell”. Each number corresponds to an offer from one of the participants. Their own offers are shown in blue; others’ offers are shown in black. The offers to sell are ranked from high to low, so that the cheapest (best) price is displayed at the bottom of the list (Giusti et al. 2016).

Second, they can realize a purchase of shares by accepting an offer to sell.
If they have enough money in their savings account, they can buy a share at one of the prices in the “Offers to sell” column (which also contains their previously submitted offer to sell). They can buy a share by selecting one of the others’ offers and then clicking on the red button “Buy”. They are not allowed to accept their own offers (shown in blue) (Giusti et al. 2016).

If they click on the “Buy” button without selecting any offer, the program will automatically buy for them at the cheapest (best) price among the offers to sell posted by others. The best offer is highlighted in deep blue (Giusti et al. 2016).

It may also happen that when they select the best price and press the “Buy” button, someone else is doing the same action but acting slightly faster than them. In that case, a message “someone has been faster than you” will show up (Giusti et al. 2016).

They now should try to buy a share now. They can do so by choosing a price in the column “Offer to sell” and then clicking on the “Buy” button; or directly clicking on the “Buy” button and buy at the cheapest price listed in the column “Offers to sell” (Giusti et al. 2016).

Every time an offer is accepted, a transaction is done. Immediately, when they accept an offer to sell, they can immediately see that a purchase and the number of EURUX in their cash balance decreases by the trading price; at the same time, their trading partner realizes a sale and the balance in her/his cash balance increases by the trading price. On the contrary, when their offer to sell is accepted, they realize a sale, their trading partner does a purchase, and cash is transferred from their trading partner to them by the amount of the trading price (Giusti et al. 2016).

Given that they all submitted one offer to sell and accepted one offer to sell, they all realized one purchase and one sale so they have the same number of shares as they started out with (Giusti et al. 2016).

Third, they can initiate a purchase of a share by submitting an offer to buy.

If they have cash and would like to buy a share, they can initiate the purchase by submitting an offer to buy. They just have to enter a number in the text box under the words “Enter offer to buy” located on the right side of the screen and then click the button of “Submit offer to buy” (Giusti et al. 2016).

At this moment, participants should try to submit an offer to buy a share. They have to write a number in the text box “Enter offer to buy” and then they have to press the red button named “Submit offer to buy”. Once they pressed the red button “Submit offer to buy”, they will immediately see, in the column named “Offers to buy”, a list of numbers ranked from low to high , so that the highest (best) price is displayed at the bottom of the list. If they want to sell more shares, they have to repeat this process. Again, their own offers are shown in blue; others’ offers are shown in black (Giusti et al. 2016).

Fourth, by accepting an offer to buy, they can realize a sale of a share.
They can sell a share at one of the prices that appear in the ‘Offers to buy’ column – which also contains their previously submitted offer. They have to select one of the offers and then click on the red button “Sell”. If they click on the “Sell” button without selecting any offer, the program will automatically sell one share for them at the highest (best) price listed in the column of “Offers to buy”. Again, they are not allowed to accept their own offers (shown in blue). They have to remember that the highest (best) price is displayed at the bottom of the list (Giusti et al. 2016).

At this moment they have to try to sell a share. They have to choose a price from the column “Offer to buy” and then they have to click on the “Sell” button; or directly click on the “Sell” button and sell at the highest price listed in the column “Offers to buy” (Giusti et al. 2016).

Again, a transaction is done when an offer to buy is accepted. If they accept an offer to buy posted by others, they realize a sale and as a result, their cash balance increases by the trading price. On the contrary, when their offer to buy is accepted by someone else, they realize a purchase and the number of EURUX in their savings account decreases by the trading price. The contrary happens to their trading partner (Giusti et al. 2016).

These four trading methods are complementary: it is possible to initiate a trade by offering a price to sell, or buy and wait for others to accept the offer; or it is also possible to realize or execute a trade by accepting an offer, submitted by others, to buy or sell (Giusti et al. 2016).

In the column situated in the middle of the screen named “Trading price”, participants will be able to see all prices at which shares have been traded during the trading period by other participants present in the market (Giusti et al. 2016).

The above explanation was the trading interface that participants will use during the experiment. A continuation, we will provide more instructions about dividend payments. After the explanation, participants will have time to practice a full trading

period. However, they have to take into account that they cannot press the 'continue' button until the instructor tells so (Giusti et al. 2016).

DIVIDEND

Once the trading period ends, participants will receive dividends for the shares they hold. Dividends are paid in cash (in units of EURUX) (Giusti et al. 2016).

The amount of dividend per share will be determined by a random device (the Company's business may go well or bad, which will affect how much dividend you get) and will take one of four values with the same probability (Giusti et al. 2016):

1/4 probability you get 0 EURUX per share,

1/4 probability you get 8 EURUX per share,

1/4 probability you get 28 EURUX per share, and

1/4 probability you get 60 EURUX per share

Each participant will get the same dividend per share, and for each new trading period, there is a new random dividend draw. Since the four results have the same probability to happen, we can calculate the average dividing as: (0 + 8 + 28 + 60) / 4 = 24 EURUX (Giusti et al. 2016).

Here an example to show how dividends will be paid. If after trading, participants have 4 shares and 2000 EURUX in their savings account. The random device shows that each share receives a dividend of 8 EURUX. At the end of the period, participants will receive 8 x 4 = 32 EURUX of dividend. As a result, the balance in their savings account, at the end of the period, will be 2000 + 32 = 2032 EURUX (Giusti et al. 2016).

END-OF-PERIOD: Information Screen

At the end of the trading period, after dividends and those are paid, they will be shown an information screen which will include the information about dividend payment, information about their end-of-period inventory of shares and their cash balance (Giusti et al. 2016).

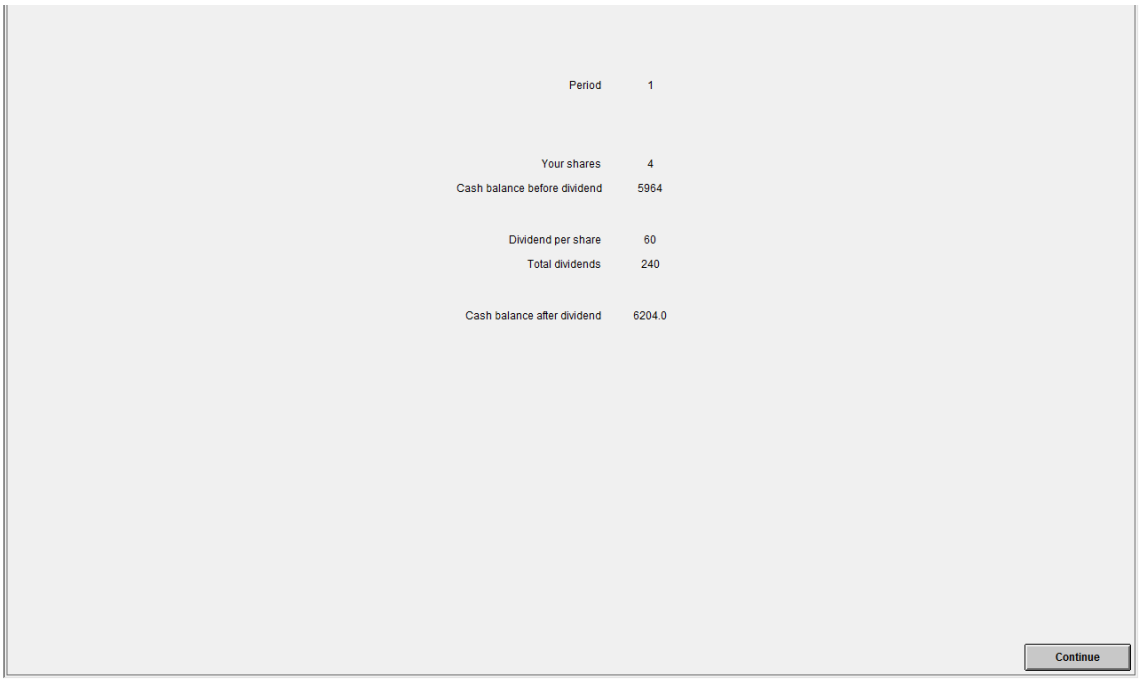


Figure 5. End of period interface

The information screen will contain the following information (Giusti et al. 2016):

1. *Period:* the period just finished.
2. *Your shares:* the numbers of shares participants will have after trading in the period.
3. *Cash balance before dividend:* participants' cash balance (in EURUX) right after trading and before dividend payment.
4. *Dividend per share:* the amount of dividends in EURUX that participants will receive for each share they own.
5. *Total dividend*: is the amount that results from the following calculation: *the amount of shares you own* x *dividend per share.*
6. *Cash balance after dividend:* participants' cash balance after dividend payment, which is calculated as *Cash balance before dividend* + *Total dividend.*

TRIAL PERIOD

Before starting the experiment participants will be able to practice trading in this market, during 3-4 minutes, for one period. Their actions in this period will not influence their position and earnings in the real experiment (Giusti et al. 2016).

To start the trial participants will have to click on the 'Continue' button on their screen. After they have practiced with the presented interface and they understand how to trade shares and how dividends are paid, we will go through some instructions to help them maximizing their earnings (explained later) before the experiment starts (Giusti et al. 2016).

The experiments consists of 15 consecutive trading periods. Each period will last for 120 seconds. They will start in period 1 with a certain amount of money in their savings account and a certain investment portfolio of shares. In each 15 trading periods, they will trade among themselves using the interface they just practiced with. At the end of each trading period they will be able to see the 'information screen' which will show them their end-of-period portfolio position after dividends (Giusti et al. 2016).

Their shares inventory and savings account balance will be transferred from one period to the other. For example, if at the end of period 4 they have 4 shares and 2000 EURUX, then they will start period 5 with the same portfolio of 4 shares and 2000 EURUX before trading (Giusti et al. 2016).

AVERAGE VALUE TABLE

How to maximize earnings?

The objective of their investment decisions is to maximize your earnings at the end of the experiment. In each exchange period participants decide how many shares to buy and sell and at what price (Giusti et al. 2016).

To facilitate their decisions we will provide them with a table called "Average value table" that can be used throughout the experiment. The table calculates the average amount of money they would earn if they buy a share in the current period and keep it in their wallet until the end of the game. Clearly they can choose to sell that stock at any time, for example if someone offers them a good price to buy it. The average value table only serves as a reference (Giusti et al. 2016).

The table has 6 columns, which will be explained one by one (Giusti et al. 2016):

1. *Current period:* The exchange period in which you are.
2. *Average dividend:* The average dividend per share and per period. This, as explained above, is equal to 24 EURUX.
3. *Average remaining dividends:* If they keep an action in your portfolio from this period until the end of the game, they will receive a dividend at the end of each of the remaining periods. The remaining dividend is calculated as the number of remaining periods of the experiment multiplied by the average dividend. For example, for every action they have in period 14, there will be two dividend payments: one at the end of period 14, and one at the end of period 15. On average, dividends will increase their earnings at the end of the game by 24 x 2 = 48 EURUX.

Table 36. Average Value Table

1	***2***	***3***
Actual period	*Average dividend*	*Remaining average dividends*
1	24	360
2	24	336
3	24	312
4	24	288
5	24	264
6	24	240
7	24	216
8	24	192
9	24	168
10	24	144
11	24	120
12	24	96
13	24	72
14	24	48
15	24	24