Ph.D. in Economics and Finance
XXXI – year 2018/2019
Department of Social Science and Economics

How to disclose future performance of non-equity financial products. Evidence from a consumer testing exercise and a physiological experiment on retail investors.

Supervisor

Prof.ssa Rosella Castellano
Unitelma Sapienza,
Rome, Italy
rosella.castellano@unitelmasapienza.it

Ph.D. Candidate

Marco Mancinelli
Sapienza – University of Rome,
Rome, Italy
marco.mancinelli@uniroma1.it

KEY WORDS: behavioural finance, decision making, framing effect, risk disclosure, stochastic dominance, financial literacy, What-if, Probabilistic scenarios, consumer test and neuroscientific analysis.

EXECUTIVE SUMMARY

The empirical breakdown of rational preference hypothesis in the context of decision-making reveals the limits of classical economic theory. Behavioural finance, which draws on cognitive theory, provides an explanation of some of the biases that affect the decision-making process. The main pillars of behavioural finance include prospect theory and the framing effect, which have been analysed in detail in many studies of financial decisions. The influence of behavioural finance and neuro-economic theories on policy makers is increasing. Legislators and regulators have been trying to find ways of using policy to correct cognitive biases and thus reduce irrational behaviours. Given the various anomalies observed in data on preferences and portfolio choices, there is general interest in identifying a regulatory tool that may help to debiasing the behaviour of agents. Improvements in investor education, financial disclosure and investment advice services are considered the main tools for improving individuals' understanding of the tricks in the decision-making process. This research focused on disclosure of information about the likely future performance of non-equity financial products.

In recent decades, financial innovation has increased the complexity of financial products. In particular, non-equity products (such as mutual funds, structured bonds and some insurance products) are chosen also by retail investors and represents roughly 45% of total households' financial wealth. Their success might be increased by behavioural biases. A highly complex structure and a lack of standardisation often go along with poor disclosure of information and choosing between such products can be very stressful. Information on an investment instrument and the way it is represented (framing effect) are crucial in determining the outcome of decision-making. The representation of the characteristics of financial securities is particularly useful when the products are characterised by an intricate financial structure and a complex pattern of potential returns. In such cases the product embeds several risks that cannot be adequately perceived by the investor who relies on the usual metrics, such as price, internal rate of return and volatility.

This research looked at two schemes that can be employed to frame the random performances of non-equity financial products: the What-if and Probabilistic scenarios. As they have not been analysed before, applying these representation schemes on two Italian subordinated bonds we tested the methodological strengths and weakness of both approaches , as well as their influence on investment decisions.

Evidence shows that Probabilistic tables provide useful information whereas the What-if scenarios are affected by some methodological weaknesses. Firstly, the What-if approach considers three elementary outcomes out of an infinity of possible results of the investment, probably selected at the convenience of the issuer, whereas the Probabilistic method presents the entire probability distribution of the product's final performance, summarising it as a set of prescribed events that would be important to any investor. Secondly, since the What-if representation shows only projected future returns, it may be difficult for retail investors to link a probability of occurrence to each event. Not only is it impossible to identify the relevant probabilities, investors also cannot estimate the likelihood of general scenarios definable as negative, neutral and positive. Given the computational processes used to get the output tables, both approaches can be implemented quickly and can reflect all risks and market conditions. However, Probabilistic scenarios are based on pricing practices that have been already developed internally, by the issuer, for pricing and risk management purposes and widely shared among competitors. At the same time, allowing model selection to manufacturers leaves room for arbitrariness. Methodological weakness and strengths become more relevant when the manufacturer should periodically update the tables to reflect variation in market conditions and inform investors about the value of a security. The process and models underlying the Probabilistic scenarios reflect market expectations of the risk-reward profile of the products, whereas the What-if tables could not

be coherent with a forward looking and complete representation of the future performance and thus generate a lack of information for investors.

A consumer testing exercise and a laboratory experiment were conducted to assess the impact of the two schemes on investment decisions of potential retail investors as they are the most affected by behavioural biases and episodes of mis-selling.

A consumer testing exercise on a sample of 1130 potential retail investors stratified according to the Italian National Institute of Statistic (ISTAT) data aimed to explore: (i) how different representation formats are appraised in terms of complexity, usefulness and information content, (ii) how the different schemes influence risk perception, and (iii) how different templates affect investment decisions. Participants were divided into two homogeneous groups considering gender, age, geographic area and education as stratification variables and two different questionnaires were delivered. Both contained questions about sociodemographic characteristics, investment habits and a section measuring financial literacy. The only difference between the two questionnaires was the section about an investment choice: two subordinated bonds negotiated on the Italian retail bond market were presented using What-if or Probabilistic tables. It should be highlighted that the investment choice is between two financial products where a second-order stochastic dominance is verified.

Firstly, in order to investigate the effectiveness of the two schemes and how they are appraised by consumers, individuals were asked to evaluate the usefulness, clarity, adequacy and comparability of the information content of the risk-return representations. Although there was some heterogeneity related to sociodemographic characteristics, the information displayed in the What-if representation was perceived as more complex than the information displayed in the Probabilistic table. Probabilistic modelling was perceived to be more useful than the What-if format. Secondly, in order to assess the relation between information disclosure and risk perception, investors were asked to assess the products' riskiness. The percentage of respondents that ranked the riskiness of the products correctly was higher when the Probabilistic format was used. Frame and financial literacy were the main determinants of risk assessment. Finally, respondents were asked to invest in one product, given an initial endowment, a time horizon and an investment objective. This allowed observation of the impact of the frames on an investment decision. As expected, frame affected investment choice, with the Probabilistic format enabling risk averse investors to make less biased decisions. In this case, financial knowledge, personal traits and investment habits did not play a role.

In addition, a laboratory experiment was carried out to measure the impact of the two frames on potential retail investors' decision accuracy and peripheral nervous system activity. The goal was twofold: (i) to investigate the behavioural and physiological indices elicited in the decision-making process of investors, who had to make their choices on the basis of information represented in one of two formats; (ii) to investigate the effects of time pressure on the choice when Probabilistic scenarios or What-if frames are used. The first point encompasses a decision-making task (Task 1), while the second a perceptual one (Task 2). In particular, participants were asked to (i) complete a financial education questionnaire (the same as in the consumer testing) and (ii) to complete the financial investment tasks. Task 1 comprised twenty-two trials including Probabilistic and What-if scenarios. Each trial consisted of two financial products, one of which stochastically dominated the other. Participants were asked to make 22 financial choices during which behavioural and physiological indices were recorded. In particular, the behavioural measures included the accuracy of the financial decision and the reaction time associated with each decision. The physiological measures were galvanic skin response and heart rate, both considered reliable indices of stress level and cognitive load during the decision-making process. Task 2 aimed at investigating participants' preferences for What-if or Probabilistic frame. It consisted of the presentation of 22 trials in which the characteristics of a single financial products were displayed by means of the two different frames. No right or wrong responses were involved in this task.

During the financial decision task (Task 1), participants made more accurate and faster financial decisions when securities were presented via Probabilistic scenarios scheme. In addition, the integrated skin conductance response data suggested that Probabilistic scenarios induced less sympathetic nervous system activation during the decision-making process than the What-if representation, which induced more stress-related responding. A general conclusion from this analysis is that compared with What-if scenarios, Probabilistic scenarios schemes: (i) enhance financial decision accuracy, (ii) decrease financial decision reaction times and (iii) induce less stress-related responding. As regards the perceptual decision task (Task 2), the Probabilistic scheme was generally preferred by participants, but the What-if frame was preferred in trials in which the probability of a negative return was greater than 50% and the probability of a positive return was less than 50%. This observation could be explained by the fact that Probabilistic scenarios clearly alerted investors that those securities were too risky, so they ended up choosing the What-if scheme. This is an interesting observation which warrants future research.

Overall, both experiments show that risk perception and investment decisions of retail investors are context-dependent and mainly determined by the way financial information is disclosed. Evidence suggests that Probabilistic scenarios might help investors to make fair financial choices with respect to second-order stochastic dominance criterion. Indeed, What-if tables might foster biased beliefs in investors, since the partial representation provided by the three illustrative examples could be perceived as exhaustive of all possible future performance scenarios achievable by a product.

The research adds to the existing literature by providing new evidence about two representation schemes. Drawing on the appraisals made by a sample of Italian retail investors, the study provides insight into how people actually read and understand financial information, which may prove useful in the design of financial disclosure documents. Moreover, evidence about investors' heterogeneity and the behavioural biases affecting risk perception supports the idea that the 'one-size-fits-all' approach cannot be effective in ensuring a suitable level of investor protection, but there are representation schemes that help to reduce investment inconsistencies. This research is in line with the approach adopted by some regulators, who are increasingly engaged in the definition of evidence-based rules, and may offer useful insights for the design of more effective representation format.

CONTENTS

LIST OF TABLES

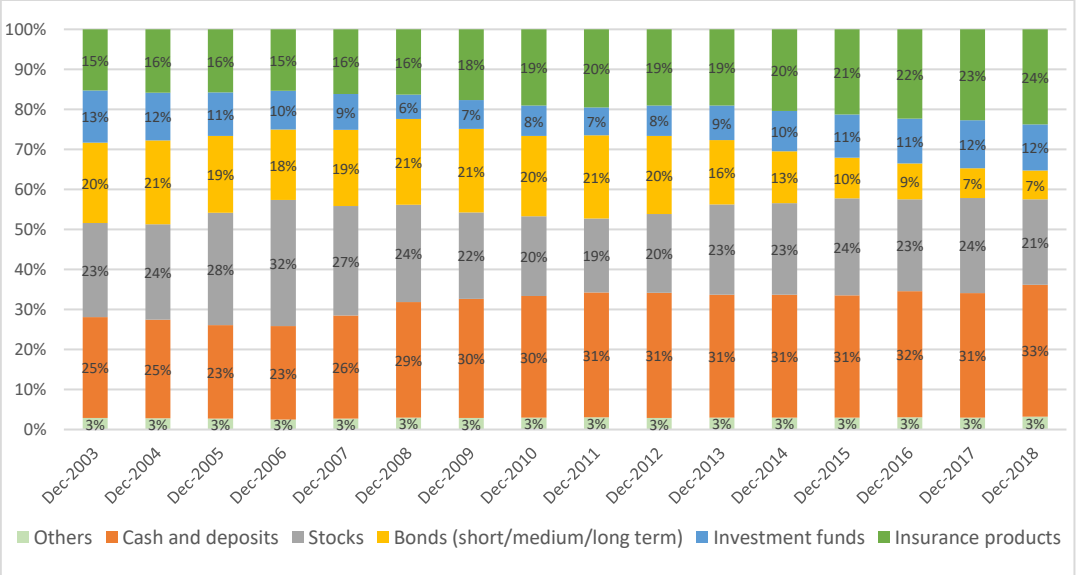
LIST OF FIGURES

1 How the decision-making process is affected by the irrationality of agents. A review of the literature.

1.1 Introduction

Financial markets are becoming more and more complex, to the detriment of disclosure. In general, retail investors suffer most from such complexity and, due to their lack of financial literacy, financial decisions are difficult even in case of non-structured securities (e.g. deposits and saving accounts). For this reason, investors are often forced to trust advice coming from intermediaries, who direct the investor toward the purchase of financial assets, which are often presented as simple and very profitable. However, these instruments are derived from a complex financial engineering process and are difficult even for expert investors to understand. In most cases, these securities are not presented clearly to potential buyers and are accompanied by prospectuses designed to steer investors in a particular direction, which is not always the appropriate one. As a consequence of poor information, shares of non-equity financial products which are more affected by this complexity are significant also among retail investors. In this regard, with specific reference to the Italian situations, Figure 1.1: *Breakdown of Italian households' portfolios of financial investment products*[1] shows the composition of households' wealth invested in financial products for the period December 2003 – December 2018. During this period non-equity financial products represent on average 45% of total households' financial wealth.

Figure 1.1: Breakdown of Italian households' portfolios of financial investment products.



Source: Bank of Italy - Conti finanziari.

Some anomalies in investors' behaviour highlight the weaknesses of the assumptions that classical financial theory makes about individuals' rationality and the homogeneity of expectation, which were

[1] Considering the reported categories, 'Bonds' includes different types of structure (plain vanilla, structured and) short- and medium/long-term securities issued by central government, local authorities, banks and corporates; 'Investment funds' are those issued by Italian and foreign institutions; 'Insurance products' is composed by life insurances, structured bonds issued by insurances and pension funds; the residual category 'Others' contains trade receivables and derivatives.

progressively overcome through introduction of a behavioural finance approach. A long series of behavioural and experimental studies shows that risk preference and financial decisions are extremely sensitive to framing, i.e. the way in which information is represented. Not only is there the possibility that the choice architecture may lead to biased and unconscious choices, but the negative effects could be reinforced by other factors such as heuristics, the level of financial literacy and the emotional components of the investors' decision-making process.

The extant evidence has generated an active debate on how financial information can be best presented to consumers in order to facilitate a mindful choice. Regulators are becoming increasingly aware of the need to change their way of defining measures and to deviate from the classical hypotheses. The standard regulatory approach is gradually making way for the so-called 'cognitive disclosure method' which refers to behaviours in the real-world.

This chapter is organised as follows. Section 2 analyses the transition from classical financial theory to behavioural finance. Over the years several studies have been conducted to identify framing effects in financial contexts and some of these results are described in Section 3. Then, Section 4 and 5 review some approaches to reducing cognitive biases and how policy makers have reacted to the emerging evidence. Section 6 deals with the inefficiencies in portfolio management of complex products and introduces the stochastic dominance criteria defined for a rational decision-making process under uncertainty. Finally, some general conclusions are stated in Section 7.

1.2 From classical financial theory to behavioural finance

Classical financial theory assumes that agents are totally rational, because they make their decisions using a complete and homogeneous set of information. Considering the literature on decisions under uncertainty, expected utility theory has a fundamental role to play in explaining the behaviour of investors (Von Neumann and Morgenstern, 1947).

However, during the last few decades, empirical researches have disproved assumptions about efficient markets and investors' rationality. It seems that individuals repeatedly commit reasoning and preference errors that are hard to accept if we assume that choices are rational. The empirical and experimental results reveal the limitations of the classical theory when it comes to describing the behaviour of people who are unable to use the available information and make errors in perceiving and processing the information itself.

The inconsistency between agents' behaviour and the hypotheses of classical theory has been demonstrated in several fields, including the financial context. The need for portfolio diversification is one of the main normative implications of the traditional theory, yet some studies have highlighted anomalies in investors' choices. Agents who own financial products are inclined to hold portfolios which are strongly biased towards a specific geographical area, issuer or type of security. Financial portfolios are subjected to the so-called 'familiarity bias' (or 'home bias') which refers to the situation where institutional or individuals investors prefer to buy domestic securities rather than foreign assets. Initially, French and Poterba (1991) and Tesar and Werner (1995) analysed equity familiarity bias in their studies where they found that possible reasons behind familiarity bias may be investment barriers, transaction costs, information asymmetry, inflation hedging and non-tradable assets. A study based on worldwide equity fund holdings data in 1999 and 2000 (Chan *et al.*, 2005) documented the existence of familiarity bias in every single country in the analysed sample of 48 countries. The authors linked investors' strong preference for domestic equities as a consequence of cross-border boundaries that give rise to exchange rate risk, variation in regulation, taxation, accounting standards, corporate governance, transaction costs, information asymmetries, and biased expectations. Coval and Moskowitz (1999) documented that US investment managers tend to invest in small, highly levered and locally situated firms that produce nontraded goods. Similarly, Strong and Xu (2003)

reported that fund managers are more optimistic towards investing in familiar securities and inclined to familiarity bias. More recently, Ke *et al.* (2010) examined the US equity holdings of more than 3000 non-US-based mutual funds from 22 developed countries for the period 2001-2002 and found that managers strongly prefer to invest in stocks of US firms that have presence in their home country. This familiarity bias is independent of the degree of global involvement and visibility of these US firms.

Considering the influence of familiarity bias on individual investors, Seasholes and Zhu (2010) found empirical evidence that individuals are more likely to hold local stocks in their portfolios and reported an information asymmetry among individuals which is considered as an indirect obstruction to foreign investment. Karlsson and Norden (2007) analysed individuals' demographic and socioeconomic characteristics and found that the likelihood of familiarity bias is caused by both rational and irrational factors. In particular, hedging against inflation and overconfidence are the reasons behind familiarity bias. Previously, Cooper and Kaplanis (1994) tested possible relations between familiarity bias in equity portfolios and intention of investors to hedge inflation risk. The authors explained that investors are risk-averse and that they hedge the inflation with domestic equity becoming more prone to familiarity bias.

These findings contradict the traditional finance framework suggesting that investors partially base their investment decisions on their familiarity and association with a stock rather than on the fundamentals of that particular stock. Thus they fail to achieve the greater efficiency of internationally diversified portfolios.

Moreover, the asset allocation puzzle is marked (Canner *et al.*, 1997). According to the mutual-fund separation theorem, given the three main classes of financial assets (cash, bonds and stocks), agent risk aversion should only affect the share of the portfolio invested in riskless assets without altering the composition of risky assets, which should be unique among individuals (i.e. the so-called 'market portfolio'). Nevertheless, as risk aversion and the reference time horizon increase, financial advisers tend to recommend a higher ratio of bonds to stocks.

Excessive trading represents another deviation from the rationality assumption: assets with positive performance are sold too quickly, whereas securities with negative performance are held for too long (Barber and Odean, 2000). This behaviour reflects the so-called 'disposition effect' introduced by Shefrin and Statman (1985).

Simple experiments conducted in the last few years do not fully support the notion that individuals behave rationally, even in simple investment scenarios. The incoherence of behaviour can be explained by many cognitive biases. In particular, individuals acquire and process information using a limited number of heuristic rules, which reduces the complexity of problems but causes systematic and significant errors (Tversky and Kahneman, 1975). Using simple choice problems the authors identify three main judgemental heuristics which affect assessment of the probability of uncertainty quantities: representativeness, availability and anchoring. The representativeness heuristic assumes that all probability estimates are formulated on the basis of stereotypes and familiar situations. The representativeness heuristic can hinder accurate judgments of probability by emphasising aspects of the considered event that are similar to a prototype, or by hiding other diagnostic information that demonstrates the event's dissimilarity to the prototype. Agents usually make a judgement based on what they can remember, rather than using a complete set of information. Thus, according to the availability heuristic, things more easily recalled from memory are assumed to be more probable. Finally, the heuristic of anchoring operates when individuals use only an initial piece of information to make subsequent judgments': once an anchor is set, other judgements are made by adjusting away from that anchor, and the interpretation of other information is biased by the anchor. Changing the starting point would change the evaluation.

Behavioural finance explains many anomalies by considering all these cognitive biases and removing the assumption of rationality of preferences. One of the main pillars of behavioural finance is prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). The authors identified two steps in the

decision-making process. Individuals analyse the prospectuses they are offered during an 'editing phase' and rank them after an 'evaluation process'. The main innovations with respect to the classical theory are introduced in the evaluation stage. Agents compute a value for each alternative using a 'value function', the equivalent of the utility function in the classical approach. The value function is S-shaped, reflecting the different treatment for gains and losses: the region of gains is concave, whilst the region of losses shows a more pronounced convexity. This formulation represents two characteristics of prospect theory, the reference point[2] and loss aversion[3].

Behavioural finance and prospect theory explain most of the anomalies in individual investment choices. For example, behavioural portfolio theory (Shefrin and Statman, 2000) can explain the already mentioned lack of diversification and the incoherence of some decisions. According to this theory, given a benchmark aimed at a global level, an investor builds his or her portfolios using a layer-by-layer approach that combines low-risk assets in the lower layers with very risky investments in the higher layers. This weak approach may lead to inefficient portfolio management, due to ignoring covariance among layers.

Furthermore, the behavioural assumptions help to explain familiarity bias. Experimental evidence (Kilka and Weber, 2000) highlights investors' use of the familiarity heuristic when choosing between domestic and foreign financial instruments: they consider themselves more competent to make judgements about domestic assets and exhibit optimism and overconfidence in this area.

Risk perception is also related to familiarity bias (Wang *et al.*, 2011). Survey respondents seem to be more confident about their evaluation when they are highly knowledgeable and experienced. In this case, high financial literacy generates overconfidence: people are overoptimistic and underestimate the risk of investments with which they are familiar.

These are only some examples of the link between behavioural finance theories and the inconsistency in investors' behaving with classical theory. There are also explanations of overconfidence, mental accounting, sunk costs and disposition effects (see Haigh and List, 2005; Thaler, 1985; Larrick *et al.*, 1990; Ferris *et al.*, 1988 respectively).

Classical theory suggests that agents' choices are frame-invariant, that if the same product is presented in two different ways people will be unaffected by the representation format and select the better prospect. Prospect theory takes a completely different view. One of prospect theory's important deviations from the notion of rational preferences is the so-called 'framing effect' (Tversky and Kahneman, 1981, 1986). Individuals' preferences and choices change according to the way in which the information is formulated: wording, presence of pictures, descriptions etc. In other words, an individual's choice might be strongly influenced by the way in which the options are presented. With respect to investment choices, framing results in a short-term bias, which, for example, may explain the inadequacy of retirement saving rates. It also affects attitude to risk because, for example, highlighting potential gains induces risk aversion, whereas emphasising potential losses determines risk appetite and loss aversion.

Later studies (Levin *et al.*, 1998) defined three categories of framing: risky choice framing, goal framing, and attribute framing. Risky choice framing occurs when agents should choose among options characterised by different levels of risk, but the presentation emphasises potential gains or potential losses. Goal framing focuses on the potential benefits associated with an option or highlights possible negative consequences of not choosing the proposed prospect. Finally, attribute framing influences valuations of a specific outcome through the choice of positive or negative words used to describe it.

Another aspect of framing explains how preferences change according to whether a decision is presented as a choice between alternative courses of action or an opportunity to pursue a specific course of action

[2] Agents think in terms of expected utility relative to a reference point (e.g. current wealth) rather than absolute outcomes.

[3] Loss aversion refers to people's tendency to prefer to avoid losses rather than make equivalent gains (losses loom larger than gains): in absolute terms, losing €10 generates a more intense negative feeling than the positive equivalent generated by gaining €10.

(Jones *et al.*, 1998). People focus their attention on those aspects of the decision problem that are explicitly represented and tend to pay less attention to those aspects that are only represented implicitly. In the real world many decisions are framed as opportunities rather than as choices, so people concentrate on the positive attributes of the options. On the contrary, offering the same product/action as one of the alternatives reduces the frequency of selecting that product/action rather than when it is simply proposed as an opportunity (i.e. only one product/action is offered).

When making portfolio selections individuals are implicitly influenced by whether the framing mode is narrow or broad (Kumar and Lim, 2008). Experimental evidence shows that when making stock investment decisions, investors are subject to narrow framing which isolates one option and ignores possible alternatives, multiple consequences, and combinations of outcomes. In particular, narrow framing increases the disposition effect, because investors' interest is focused on the possible cash flows of a single stock and not on the entire portfolio; furthermore, wider framing of choices would show a correlation among assets and thus increase the chance of portfolio diversification.

One of the first applications of framing effect to financial portfolio selection was provided by Shefrin and Statman (1993). Using covered calls and premium bonds, they showed that some investors prefer one financial product over another due to the way identical cash flows were framed. The behavioural model used by the authors to explain the phenomenon included prospect theory, hedonic framing[4], the behavioural life cycle[5] and cognitive errors. It represents the effect of framing on investors preferences with respect to covered call. These financial instruments are designed for several overlapping client types: prospect theory represents better investors who are highly risk-averse in gains and highly risk-seeking in losses and behavioural life cycle allows to describe behaviour of investors who are consuming from their portfolios.

1.3 Experiments on the framing effect

There are many examples of framing effect in many different research areas and lots of survey-based experiments have investigated the various aspects of this phenomenon.

People analyse several sources of information - newspapers, media and everyday situations. In particular, individuals are usually interested in weather forecasts. Ibrekk and Morgan (1987) analysed the impact of nine snow forecasting formats on non-technical people's mean estimates of the probability of snow fall. Providing the entire probability distribution (e.g. pie chart, density or cumulative function) without adding supplementary information leads some agents to mis-estimate the mean; more accurate estimates are obtained when the graphs explicitly mark the location of the mean.

Communication strategies and advertising try to take advantage of cognitive biases due to the framing effect. The likelihood of product purchase is driven by factors such as price, possible outcomes, attributes, valence, perspective and past experiences (Berger and Smith, 1998). Using a sample of qualified agents, the authors studied the simultaneous effects of these six factors on performance on a video camera purchasing task. Showing price and attributes increased propensity to purchase, because it made it easier to compare the alternatives with respect to these factors.

Advertising effects linked to frame have also been studied in the financial context. Bertrand *et al.* (2010) offered South African consumers different versions of information about a loan, in which features such as comparison to competitors, interest rates, images and suggestions on how use the loan were varied.

[4] Hedonic framing refers to how people try to maximise psychological pleasure and minimise pain (regret) when faced with decisions relating to gains and losses. This means that two individual gains are perceived to be more valuable than a single larger gain of the same value.

[5] The key assumption of the Behavioural Life Cycle theory (Shefrin and Thaler, 1988) is that households treat components of their wealth as non-fungible, even in the absence of credit rationing. Specifically, wealth is assumed to be divided into three mental accounts: current income, current assets, and future income. The temptation to spend is assumed to be greatest for current income and least for future income.

Estimating agents' sensitivity to all these factors, the authors found that advertising has a positive effect on demand for loans, but does not influence the amount borrowed. Therefore, simply by working on non-informative aspects without improving the quality of the loans (for instance interest rate and costs), financial intermediaries can significantly increase the demand for loans.

Moreover, the presence of distractors, such as advertisements and colourful texts (Hillenbrand and Schmelzer, 2015), in financial prospectuses reduces the amount of pertinent information absorbed by investors. The authors found that past experiences had no effect on assessments of product characteristics: individuals seem to miss the same essential information repeatedly, leading to consistent overvaluation of securities.

Each evaluation involves a stimulus which influences the response to another stimulus, generating an implicit process in our memory, the so-called 'priming effect'. Visual priming affects the evaluation process by focusing attention on more noticeable attributes and improving a person's awareness of them. The relationship between visual priming of investment information and financial knowledge was investigated by Wang and Dowding (2010). The analysis showed that financial literacy was not correlated with the impact of visual priming on understanding and evaluation of information. However, more knowledgeable investors handled data effortlessly and obtained a global picture of the financial instrument.

Concepts such as expected returns and risks come to mind when thinking about financial instruments. Therefore, it is crucial to understand how the presentation of information about risks and returns influences the decisions of retail investors. An empirical study of the impact of product risk presentation format on preferences provided interesting results (Cox *et al.*, 2006): presenting risk as a likelihood of loss rather than as variance in potential outcomes prompted respondents to avoid risk-seeking behaviours. Based on these results the authors proposed a 'valence congruity theory' and a 'risk discrimination theory'. The latter suggests that loss-framed messages induce risk aversion, whereas gain-framed messages increase consumers' tolerance of at least some types of product risks because they ignore minor or temporary dangers and focus on the more permanent ones.

More recently (Vlaev *et al.*, 2009) nonprofessional investors were asked to judge eleven types of risk framing of the same financial product and rate the usefulness, suitability and grade of information. The highest ranked format represented possible outcomes and risks using a probabilistic approach: the potential outcomes were linked to their likelihood of occurrence, showing both minimum and maximum predicted average amounts. On the contrary, formats that focused exclusively on gains or losses were not highly rated by agents because they did not show all possible scenarios.

Framing also influences investors' attitude to risk. Interesting empirical studies about the effect of risk representation on risk taking behaviour were carried out by Kaufman *et al.* (2013) and Kaufman and Weber (2013). Following the distinction between descriptive and experienced analysis[6] (Hertwig *et al.*, 2004), the authors analysed how people make a portfolio selection after interacting with specific risk data tools. Respondents buy more risky assets increasing their risk-taking behaviour when involved in an interactive risk tool exercise rather than when a simple numerical description of the products is delivered or when experienced with graph representation of risk. In general, agents are discouraged to invest because they do not understand information delivered in numerical description format or, distrustful of the financial sector, overestimate risk. Considering the evidence from such experiment, allowing investors to use an interactive portfolio selection tool might increase market participation and efficiency.

Differences in framing might affect the consistency of preferences. Changing the format of representations (graph vs. table, absolute vs. percentage values, size of confidence interval) generates a

[6] Descriptive analysis evaluates risk using disclosed statistical and stated probabilities whereas experience analysis uses probability estimates inferred from experience with the market and past decisions.

preference reversion in some agents. In particular the consistency of portfolio selections is positively correlated with level of financial knowledge (Glenzer *et al.*, 2014). Knowledge about financial markets increases understanding of presentation format and hence decision quality.

A lot of research on so-called 'presentational impression management' has shown that companies tend to influence the perception of a phenomenon through their choice of graphical techniques for representing accounting and financial information. Several findings on the attributes of the framing effect and its interactions with biases produced by impression management prompted analyses of 'visual framing' which investigated the effects of the graphical format used to represent a variable, data aggregation and order of presentation of options.

Desanctis and Jarvenpaa (1989) analysed the impact of disclosure of financial information about listed companies, showing that representing data with properly designed graphs improved the accuracy of professional agents' forecasts relative to other formats, such as tables or a combination of graphs and tables.

Because of the importance ascribed by investors to recent movements of share prices (Clark-Murphy and Soutar, 2004), it is crucial to analyse how past performance representation influences agents' evaluation processes. The results described by Ibrekk and Morgan (1987) for snow forecast do not allow one to evaluate the impact of the graphical format used to represent investment returns in financial prospectuses. Financial engineers generally prefer to use time series data representations and bar graphs to illustrate the past performance of an asset for retail investors. The impact of the format of historical returns on assessments of risk, returns and volatility is analysed in Weber *et al.* (2005). The authors provided two groups of individuals with the names of 16 domestic and foreign investment stocks and 10-years of historical returns data. Historical returns were presented as a bar graph or a continuous density distribution. When the latter format was used estimates of volatility were higher, probably because the density distribution format highlights extreme returns whereas the bar graph emphasises trends.

The impact of visual framing and graphical representation on investors' decisions was also examined by Mussweiler and Schneller (2003). Participants' evaluations of a time series stock price chart showed a strong anchoring effect: they focused on extreme data, using them as a reference point for the analysis. A historically salient peak or trough also affected decisions. In these experiments, individuals predicted that stocks with a salient peak in their past performance would perform well in the future, whereas they opted to sell stocks with a salient trough in their past performance, thinking that they would continue to follow a negative trend. This means that it is easy for product designers to steer investors towards choosing assets that display a specific and desired historical price series.

Turning to more recent research, Diacon and Hasseldine (2007) analysed a sample of English investors and identified two factors that influence risk perception of financial funds: how past performance is represented (histograms vs. price indices) and how long the time series is (12 vs. 45 months). They showed that the graphical format of past performance may be important and suggested that information should be delivered in several formats to avoid biasing investors, although this might generate information overload.

Appropriate graphical representations are a powerful way of debiasing the decision process. When provided with simple illustrations of risk-return indices alongside the usual information potential investors assessed risk more accurately and showed more awareness of the potential consequences of their decisions (De Goej *et al.*, 2014). By providing risk information in an understandable and standardised format it is possible to highlight similarities and differences between securities.

People also behave differently when confronted with absolute and percentage representations of the same quantity (Weathers *et al.*, 2012). Computing percentages and translating them into absolute amounts is not a simple task and so when percentages are displayed, agents resort to heuristics to simplify the evaluation process, whereas absolute value representation provides a better perception of the magnitudes and steers agents to a more appropriate evaluation.

A positive-negative imbalance in information may also alter the evaluation of a financial instrument (Cox and De Goeij, 2016): balanced information reduces the propensity to seek additional data about the financial instrument. Cox and De Goeij's experiment based on different advertisements for a hypothetical investment showed that explicit risk framing increased the perceived riskiness of the investment and decreased intention to search for additional information.

There is an ongoing and active debate about whether short or long information sets should be provided. Long documents offer more information, but quantity is not always related to quality. Moreover, providing additional information may move investors' attention away from the most important characteristics or generate overconfidence and optimism. Some survey-based experiments have been carried out to provide evidence on agents' reactions to short- or long-form information documents.

For instance, agents typically wish to minimise costs when they make investment choices. Showing an additional sheet containing detailed information about fees (such as front-end load, table of implicit costs of different funds, and expense ratio) should enable people to evaluate these data and make an appropriate choice, but despite - or because of - the information available, agents feel overwhelmed and make poor selections (Choi *et al.*, 2009). The results also suggest that suboptimal choices with respect to fees are usually linked to low financial literacy.

Because they did not understand why relatively expensive payday loans were so popular among investors, Bertrand and Morse (2011) decided to improve information quality by adding details of costs. They used three different treatments providing additional information about the products' costs (comparison between alternative products, cumulative fees and an example of a repayment profile). All the respondents benefited from the additional information, requesting lower amounts; the additional information decreased the probability that a payday loan would be requested by about 11%.

The US Department of Labor introduced the Model comparative chart to facilitate comparisons between types of retirement plan. This is a table containing different options, an appropriate benchmark and average annual total returns for different maturity dates. Its effectiveness was tested with a survey in which various additional types of information were provided to groups of respondents (Hung *et al.*, 2010) such as a risk-reward rating for each option or a bar chart representing the previous ten years of annual returns for each option. Showing a risk indicator affected the risk perception and portfolio allocation of both financially literate and financially illiterate people. This result shows that simplifying and increasing the saliency of information would mitigate problems related to limits on attention and information overload. In particular, there are frames, such as bar graphs or risk bars, that help people with lower financial literacy to make more appropriate, better informed decisions.

Because retail investors are not likely to read documents that are longer than 2/3 pages financial regulators require that prospectuses are supplemented by a short document that summarises the main characteristics of a security. The US Securities and Exchange Commission (SEC) introduced the so-called 'Summary Prospectus', a significantly simpler and shorter document which must be attached to the Full Prospectus. Beshears *et al.* (2011) used hypothetical documents to analyse how investors' portfolio choices are influenced by the new document. Participants spent less time researching options when the new document was available, and the summaries did not lead to better portfolio choices or greater sensitivity to fees. The availability of summaries reduced search costs, but it did not help investors to understand the impact of fees or the risk profile of the offered securities.

A similar analysis was conducted by Walther (2015) on key investor documents (KIDs), regulated by European law (UCITS 2009/65/EC). The author found that a shorter format has two effects on investor behaviour. Firstly, it reduces information overload because it seems more informative and helpful than the full prospectus. Secondly, it has a positive impact on diversification reducing the tendency to employ a naïve

strategy[7]. However, about 50% of survey respondents were still not able to understand all the information provided.

A general comparison of six world jurisdictions (Hong Kong, Singapore, European Union, Canada, Australia and New Zealand) that use short-form documentation for retail financial products highlighted the need for global standardisation and harmonisation to increase market efficiency (Godwin and Ramsay, 2015). The authors proposed some guidelines for ensuring that information was presented in a format that would increase the probability that investors would make informed decisions. They suggested that documentation should be short; use clear, succinct and comprehensible language; highlight risks using warnings and include appropriate graphs. In addition, a survey was conducted to assess the weaknesses and strengths of the six formats (Godwin and Ramsay, 2016). Respondents were asked to evaluate the short documentation in terms of the following variables (i) comprehensibility of the language; (ii) whether it offered a clear explanation of the key product features; (iii) appropriateness of the format; (iv) appropriateness of the length; (v) whether it provided a clear explanation of the risks of the investment. Respondent rated the Canadian format the easiest to understand and the Australian the most difficult. European summary documentation attracted intermediate ratings (only 8.2% of individuals considered the KID the easiest to understand, compared with 44.3% for the Canadian format, but only 7% disliked it, compared with 59.1% for the Australian documentation).

The existence of cognitive biases and the provided evidence might expose investors to opportunistic practices whenever the operators fulfil the mandatory disclosure obligation yet choose the content and format of the information they deliver. Providing understandable and standard risk indicators would make it easier to evaluate the similarities of products, whereas removing visual distractors, by standardising the visual frame, might help to improve disclosure efficacy. The introduction of Summary prospectuses did not generate the expected improvement suggesting that a further simplification of the formatting of information is required: a reduction in use of financial jargon, inclusion of some examples and more appropriate graphs.

The introduction of the behavioural approach to interpreting individuals' choices highlighted aspects not considered under rigid classical theory. Every day agents make decisions that may have positive or negative consequences. The latter can be easily avoided through appropriate measures by policy makers.

1.4 Debiasing through law

Observing the several anomalies in preferences and, above all, in portfolio choices, we can see that researchers share a common aim, namely to identify a tool that may help to debiasing the behaviour of agents. Investor education, financial disclosure and investment advice service are considered the main tools for improving understanding of the tricks used by manufacturers and sellers to influence the decision-making process. All these tools should be implemented collaboratively by regulators and investors. Taking a paternalistic approach, policy makers should find ways of correcting cognitive biases and use policy actions to reduce irrational behaviour. However, applying a paternalistic approach to the financial world, financial regulators should reduce the range of options in the markets and ban the riskier assets. Such an intrusive solution may affect market efficiency and competition. A libertarian paternalist or 'nudging' approach (Thaler and Sunstein, 2008) to regulation would not affect the range of available choices, but would involve modification of the choice architecture. Legal methods of debiasing agents' decisions try to reduce cognitive biases by operating on circumstances that promote decision-making errors (Jolls and Sunstein, 2006).

[7] Given *n* assets the naïve strategy is to *1/n* of one's capital in each product.

The role of context in determining investment choices can be investigated with the MINDSCAPE[8] framework (Dolan *et al.*, 2012): people act differently if simple features in the decision-making environment are changed. The authors designed a simple experiment that required investors to analyse the main features of six UK financial products. Subsequently the presentation of the information was redesigned to consider various aspects of the MINDISCAPE framework and subjects were required to make the same evaluation. The changes to the choice architecture had a positive effect, increasing the financial capability of subjects.

One of the main factors that generates financial decision-making mistakes is poor financial literacy. An in-depth survey of financial literacy around the world was conducted recently (Klapper *et al.*, 2015). People are not usually familiar with basic concepts such as portfolio diversification and inflation: on average about 30% and 40% of correct answer respectively were related to the question about diversification and inflation, with huge disparities between participants in emerging and advanced economies. Moreover, the relationship between financial literacy and age is a reverse U-shaped curve: 56% of young people (age lower than 35 years) have low financial literacy, but this percentage increases to 63% for adults (aged between 36 and 50 years) and falls for older people (45%). In this survey, only 37% of Italian respondents were financially literate, one of the lowest levels of financial literacy in Europe.

There is evidence that individuals also make mistakes with some basic financial choices. For example, a study in the US showed that people are not able to plan for retirement due to lack of financial education (Lusardi and Mitchell, 2011).

Furthermore, the link between financial literacy and financial advice seeking was analysed in Italian agents: higher financial literacy was negatively associated with propensity for self-assessment and positively associated with the seeking of professional advice (Gentile *et al.*, 2016).

A traditional way of increasing financial literacy consists of spreading financial information and educating agents to tackle their lack of financial knowledge. A lot of schools, associations and authorities are collaborating to develop a compulsory financial education curriculum. Some researchers have found that such projects have a positive impact on saving rates, wealth and pension investment (Lusardi and Mitchell, 2007). On the other hand, adopting a 'learning by doing' approach may reduce cognitive biases because it helps agents to remember errors, their consequences and ways of avoiding them (Weber and Welfens, 2008). However applying this approach to financial decisions would not be without costs.

Increasing financial literacy may improve both the financial judgement abilities of investors and awareness of the main behavioural biases. The financial literature also supports the idea that better financial advice would help potential investors to overcome some of their biases and obtain better performance from their investment portfolios. An Israeli study compared equity portfolios managed by professionals and independent investors, checking whether financial advice reduced biases such as the disposition effect, low diversification and excessive trading (Shapira and Venezia, 2001). The authors found that professionally managed portfolios were more consistent with the standards of classical theory because they were more diversified and less subject to disposition effect.

Use of advisory services seems to be particularly uncommon amongst Italian households. Gentile *et al.* (2006) showed that the probability that portfolio management will be delegated to professionals positively related to the size of the portfolio and to risk aversion, but it is not influenced by financial literacy.

In order to achieve the 'know your costumer' criterion[9] in the financial system, European regulators introduced MiFID (Markets in Financial Instruments Directive - 2004/39/EC) and some rules to protect investors. In particular, financial intermediaries are now required to define clients' risk tolerance by gathering

[8] MINDSCAPE is an abbreviation (messenger, incentives, norms, defaults, salience, priming, affect, commitment, and ego) to represent nine largely automatic behavioural effects.

[9] Know your customer is the process of a business verifying the identity of its clients and assessing their suitability, along with the potential risks of illegal intentions towards the business relationship.

some information via a questionnaire. Some basic data is common to all questionnaires, but sometimes there is a failure to collect information or a distortion of the results. Lucarelli (2011) asked a sample of 100 agents to complete several different questionnaires and compared the risk tolerance profiles they generated: in some cases an individual was classified both as risk averse and risk loving. This evidence simply shows that the information collected using these questionnaires is inadequate; thus, quantitative as well as qualitative assessments are necessary to characterise an investor's risk and time profile.

Kahneman and Riepe (1998) provided guidance on how financial advisers should behave with their customers to correct the cognitive biases caused by heuristics and emotions. Adoption of this guidance would lead to development of advisory services that would place more emphasis on consumers' welfare than is currently the case.

Although improving financial literacy and financial advisory services is undoubtedly an important goal, it cannot be the only strategy for optimising decision making. On the other hand, the means by which information is presented to consumers strongly influences their decisions. Indeed, how information about the characteristics of a financial instrument is presented affects people's understanding and evaluation of investment opportunities. Disclosure of relevant information is fundamental to rational decision-making: having a clear, simple and standardised representation of the characteristics of a financial product may increase awareness of the risk-rewards profile and its comparability with other products offered in the financial markets. Disclosure should promote the autonomy and good decision making, as well as increasing efficiency and avoiding market failures (for instance incomplete and asymmetric information).

However, the causes of mis-selling of financial services and products include the use of long, detailed information sheets to ostensibly remove information asymmetries, which may not be effective in protecting retail investors[10] (House of Commons Committee of Public Accounts, 2016; National Audit Office, 2016). Financial services are complex and difficult to understand even for the most knowledgeable consumers, which can mean that consumers in this market are particularly susceptible to mis-selling. Mis-selling occurs where, for example, an intermediary recommends unsuitable products to consumers or gives them misleading information. It is worth highlighting that since April 2011 over 12 million UK consumers have been mis-sold payment protection insurance and firms have paid over 22 billion pounds in compensation to them.

It is worth adding that regulators should consider the information overload phenomenon as relevant: having more data and indicators is not always useful and it does not mitigate the difficulties that investors have in understanding financial products. There are several factors that generate information overload: how much information is presented, the number of offered options and the similarity of the options (Agnew and Szykman, 2005). Through two survey-based experiments on retirement plan design Agnew and Szykman showed that even agents with a relatively high level of financial knowledge feel overwhelmed when confronted with a large information set and respond by selecting the proposed default option. Uninformed and less financially literate consumers tend to search less before making a choice. These results highlight the importance of developing and spreading financial knowledge.

Information overload may also generate confusion and hide the crucial elements (Kruschke and Johansen, 1999). Empirical research has highlighted that information is in competition with itself because the more salient items dominate and distract investors' attention from data that would be useful for making a decision.

According to some authors detailed disclosure is ineffective not only because of information overload, but also because it may encourage overconfidence and optimism (Willis, 2008). This effect has been reported, for example, by online traders: subjects who devote more time to the acquisition of information, trade more and hold less diversified portfolios, thus obtaining a lower return (Guiso and Jappelli, 2006).

[10] For example, the cultures of firms and the nature of their sales incentives have been identified as key factors behind mis-selling.

Loewenstein *et al.* (2014) highlighted the role of attention in decision-making based on information disclosures. The authors stated that psychological factors and mechanisms, such as limited attention span, biased probability judgments' and inattention to missing information might severely undermine the efficacy of disclosure as a tool for regulators.

Using data from a questionnaire completed by private and professional UK investors, Diacon (2004) highlighted group differences in risk perception. Private investors perceive risk in terms of four factors: scepticism about financial products, worry about significant losses, returns volatility and inappropriate regulation. In contrast professionals consider risk in terms of just two factors: the possibility of losing capital and insufficient knowledge of the products. Considering all these results, describing financial risk in the prospectuses of financial products using a one-dimensional measure, such as the returns volatility, might be not appropriate. When making rules to improve disclosure financial regulators should make use of all of the factors relevant to performance.

1.5 The increasing relevance of an evidence-based approach for policy actions

Behavioural finance and neuro-economic theories are increasingly influencing policy actions. Legislators and regulators have been trying to shape a more appropriate method of disclosure to improve its effectiveness. Following to the so-called 'information-based rules', investors and policy makers usually consider that increasing the quantity of information provided should be the main way of improving investor protection. However, the many cases of mis-selling of financial instruments and studies on information overload indicate that insisting on provision of detailed information may not be helpful.

Driven by the several experiments provided in the literature a new approach to regulation has been gradually spreading. The 'cognitive disclosure approach' deviates from the hypotheses of classical theory and considers how agents behave in the real world. Policy makers are following evidence-based rules: once the phenomenon to be regulated has been identified, consumer tests and technical discussions are conducted to find the most common behavioural strategies and heuristics for the problem and then all the evidence is analysed to identify the most appropriate policy measure.

Applying this approach first conducted to the UCITS Directive (2009/65/EU, so-called UCITS IV) and Regulation (EU No. 583/2010), and then to Guidelines issued by the Committee of European Securities Regulators (CESR, 2010a, 2010b), which introduced the Key Investor Information Document for open-ended mutual funds and SICAVs (KIID hereinafter). The structure and content of the KIID were defined on the basis of the evidence that emerged from a consumer test conducted across some countries belonging to the European Union (IFF Research and YouGov, 2009). The information needs of retail investors, the clearness of different presentation formats for the items to be included in the KIID (e.g. funds' strategy and objectives, past performance, risk, costs) and the real use of KIIDs in the decision-making process were investigated. As expected, the results showed that most consumers are unwilling to read lengthy documents or information hidden in the small print or in large blocks of text, whereas the use of more visual approaches, such as graphs, was felt to make the document more engaging. Moreover, consumers paid particular attention to the risk-return profile sections, sometimes missing more subtle messages delivered by the KIID. Finally, the use of synthetic indicators and narrative approaches to presenting risk-reward profiles was tested. The synthetic indicator scored better, being perceived by most investors as delivering information on the fund profile both more clearly than the narrative approach and making comparisons easier.

Due to the quality of the results obtained from this exercise, other authorities followed the same approach. In April 2013, the UK Parliament created the Financial Conduct Authority (FCA) and gave it a duty to promote effective competition. To achieve this the FCA is supposed to undertake integrated analysis of economic markets to explain how information problems, consumers' behavioural errors and firms'

competitive strategies combine to produce observed market outcomes (Erta *et al*., 2013). This involved some change from the existing practice of most conduct regulators, with one of the biggest changes relating to the emphasis on understanding consumer behaviour. Insights from behavioural economics, together with more traditional analysis of competition and market failures, are supposed to help the FCA to assess problems in financial markets, choose appropriate remedies and be a more effective regulator than in the past (Dambe *et al.*, 2013).

One of the first examples of the use, by regulators, of experiments to diagnose problems rather than test specific remedies is provided by the Office of Fair Trading, which analysed the effects of different price formats on consumers' ability to shop around (OFT, 2010). A sample of 166 UK subjects made more mistakes and achieved worse outcomes under complex price frames, for instance, when the sellers promote a special price but only a limited number of products is available at that price or understanding the price requires some computation (e.g. '3 for the price of 2' instead of direct pricing, or 'product A costs X pounds'). This research led to a subsequent study on advertising (OFT, 2012a, 2012b) which identified risks associated with specific promotional practices. The results of the study have also been used to challenge the lack of transparency in airlines' payment card surcharges and to reach agreement with supermarkets on principles for fair advertising of discounts.

The Association of British Insurers investigated the relationship between risk presentation and investment choices in order to find a feasible way of presenting the risks associated with different investment funds that would help retail investors make decisions (Driver *et al.*, 2010). Participants assessed several graphic- and text-based formats for the synthetic risk-return indicator and numerous representation modes using the following criteria: the usability of the disclosure designs, the ability of people to rank different funds according to risk and return and people's ability to assess the suitability of funds when making decisions. The collected results highlighted that visual displays of information were more effective than text-based disclosures of risk in helping consumers to make sound financial decisions. Agents also benefit from standardisation of the disclosure format, whilst inclusion of complex charts may reduce their ability to understand the information being presented.

Research for the European Commission (Chater *et al.*, 2010) documented the factors that affect the investment decisions made with and without advice, and tested potential remedies (for example, standardisation of information formats; disclosure of advisers' conflicts of interest). The test, run on six thousand European purchasers of retail financial services, consisted of a combination of drastically simplified online investment choices (with and without advice condition) and a laboratory experiment involving direct computer-based interaction between non-specialist participants randomly allocated to the role of consumer or financial advisor. The results showed that over 40% of the participants were unaware that financial product providers, salespeople, brokers and financial advisers may receive payments for promoting particular products and were thus uninformed of advisers' potential conflict of interest. The European Commission explicitly referred to this research showing the beneficial effects of simplified and standardised information formats in its proposals on key information documents for investment products (EC, 2012).

The Australian Securities and Investments Commission (ASIC) examined the social and emotional impacts of financial losses due to misconduct by providers of financial services and assessed the effectiveness of the compensation system (ASIC, 2011). The Authority ran a multi-stage quantitative and qualitative study. The quantitative analysis involved an online research panel; data were collected about the demographic characteristics of investors who lost money and the impact of these losses on their lives. Twenty-nine representative investors were interviewed in order to assess the depth and breadth of their experiences with the various kinds of financial loss and the different compensation mechanisms used. The survey showed that failure to compensate investors who lost money because of the conduct of their managed investment scheme or financial planner in full may cause them severe emotional and financial distress. The research also

showed that investors found the compensation system difficult to use and that their experience of losses reduced their trust in the financial system.

Behavioural finance theories are also used as a tool by financial market regulators. An example is the method developed by the Netherlands Authority for the Financial Markets (AFM) for achieving a balanced relationship between consumers and financial institutions and simplifying the decision-making process of agents. An online survey of 800 consumers was conducted to find a solution to over-exposure to debt, which is probably due to behavioural biases (AFM, 2014).

The British authorities have increasingly been using evidence to drive policy development. The Financial Service Authority (FSA) ran a mystery shopping exercise[11] to get practical information that was then used to increase the fairness of firms' conduct when selling financial products to retail investors (FSA, 2013). A total of 231 mystery shoppers visited six major firms in the UK's retail banking sector and the researchers analysed the quality of advice given to customers looking to invest a lump sum. Although approximately 75% of the mystery shoppers received good advice, some concerns arose from the responses received by the rest of the sample, for example 15% of mystery shoppers received poor risk profiling, due to use of complex tools and a limited set of questions, and 42% were not given correct information (e.g. about the firm, service and remuneration). The Financial Conduit Authority has also used an open investor survey to investigate whether the selling of insurance as an add-on to a primary product impairs competition, by discouraging investors from searching for stand-alone products (FCA, 2014). Sixty-five responses were received from a cross-section of stakeholders, showing that the add-on mechanism has a clear effect on consumer behaviour and often affects consumers' decision-making, because they had little awareness that they were buying an additional product or knowledge of the price they had paid. In particular, it emerged that there was little pressure on firms to offer good value, because add-on buyers were less likely to shop around and less price sensitive: their attention was on the primary investment product rather than the add-on, leading to purchase of add-on products they did not need or understand. The FCA also conducted a survey which investigated how well consumers understand and value structured deposits and whether giving targeted information improves their understanding (FCA, 2015). Data from a sample of 384 retail investors who had already bought or would consider buying structured deposits or other structured products showed that they significantly overestimated the expected returns of all structured deposits, including the simplest. Behavioural biases, combined with the features of structured deposits, may give investors unrealistically high expectations of product returns and impair their ability to evaluate and compare structured products. Based on these findings, the FCA considered the issue relevant for its policy on structured deposits in order to understand which are product features and behavioural biases that drive investors' misperceptions of complex investments.

More recently the Italian authority for regulation of financial markets (CONSOB) conducted a consumer-test on 254 Italian retail investors to investigate understanding of different framings of financial information and perception of risk (Linciano *et al.*, 2018). The authors used different risk indicators and approaches to presenting performance scenarios and found a negative relationship between perceived complexity and willingness to invest. The authors concluded that, given the behavioural biases shown by participants and the way in which format influenced risk assessment, there is no single, optimal disclosure format and, therefore, regulators should not attempt to use one-size-fits all criteria.

Finally, European regulators have used open discussion between financial experts to define the format and content of the prospectuses of the so-called PRIIPs (Packaged Retail and Insurance-based Investment Products). All the proposed options for presenting the characteristics of a PRIIP, such as costs, risk-reward

[11] Mystery shopping involves the use of individuals trained to experience and measure customer service, by acting as potential customers and reporting back on their experiences in a detailed and objective way. Mystery shopping is used extensively in many industry sectors to measure the quality of services.

profile and risk indicators, are analysed in detail in a discussion paper published by the Joint Committee of the European Supervisory Authorities (JC/DP/2014/02).

Looking at the effort provided by competent authorities, there is a progressive improvement in disclosure which would reduce inefficient behaviours. Simplification of financial prospectuses so that they provide complete, clear, accurate yet comprehensible information is a good starting point. Moreover, combining pictures and text with statistical data may have more effect on decision making than a narrative representation. Finally, using standard and comparable information would alter the framing of investment decisions, so that they are presented as choices rather than opportunities and provide a more complete representation of a decision problem. Investors would then be able to assess possible trade-offs between financial products and make a more coherent and rational choice.

1.6 Inefficiency of complex financial products and stochastic dominance

In the last few decades financial innovation has increased the complexity of financial instruments. Celerier and Vallee (2013) are amongst those who have examined the drivers of financial complexity. They carried out a lexicographic analysis of the term sheets of 55,000 structured products targeted at retail investors and issued in 17 European countries between 2002 and 2012. They observed an increase in financial complexity, even after the recent financial crisis, and a positive linear relationship between the mark-up of a product and its complexity. In particular, ex-post analyses of the performance of retail structured products confirmed that these relatively high mark-ups translate into relatively low performance in the case of the more complex products. In addition, analysing the data at country level, Celerier and Vallee (2014) found that competition actually increases complexity, rather than acting as a brake on the trend: average complexity rises when a simple substitute product enters the market, or the number of competitors increases.

Several studies have focused on the diffusion of such assets among investors. It has been shown that investment in composite securities is not consistent with the assumption of rational behaviour and that such products are seldom convenient (Fischer, 2007). Although complex financial products are not conducive to portfolio optimisation (Hens and Rieger, 2008) they are very popular among retail investors, perhaps due to behavioural biases. A highly complex structure and lack of standardisation go hand-in-hand with low information disclosure causing wrong risk-reward assessment and portfolio choices by retail investors (Rieger, 2012).

Given that investors show limited rationality, the demand for some structured products might be explained better by behavioral theories rather than using expected utility theory. Breuer and Perst (2007) used expected utility theory and prospect theory to analyse the demand for reverse convertible and discounted reverse convertible bonds. The authors found that demand for structured products is determined by retail investors' subjective perception of their own competence, and reverse convertibles seem to be more attractive to individuals with lower financial competence. In particular, lack of investment experience reduces the desirability of uncertain prospects and differences among individuals become more relevant to subjective evaluations of certain and uncertain prospects, such that the hedonic framing rule becomes more important. Döbeli and Vanini (2010) described the demand of structured products using concepts from behavioural finance. They corroborated the results obtained in other studies: agent behaviour can be explained by the behavioural finance model.

Research in the decision-making field has led to development of various stochastic dominance (SD) criteria (Fishburn, 1969; Hanoch and Levy, 1969; Hadar and Russell, 1969; Scarsini, 1986). A decision under uncertainty is usually presented as a choice between several stochastic variables, which can be represented by their cumulative distribution functions. Using three different orders of stochastic dominance criteria an agent can rank different options, identify the least risky choice unambiguously and be capable of making a

sound decision. However, there are situations when stochastic dominance orders are not satisfied, and individuals would not be able to solve a decision problem using these criteria.

Levy and Levy (2002, 2003) extended stochastic dominance theory to risk averters and risk seekers by developing prospect SD (PSD) and Markowitz SD (MSD) theory for investors with S-shaped and reverse S-shaped (RS-shaped) utility functions, respectively. Davidson and Duclos (2000) developed SD tests for risk averters agents, whereas Sriboonchitta *et al.* (2009) modified precedent statistics to obtain SD tests for risk seekers and set out numerous applications of stochastic dominance criteria, including financial diversification, evaluation of hedge funds and income inequality. Their work was extended by Bai *et al.* (2011), who developed new statistics for both PSD and MSD of the three SD orders.

Devising an efficient portfolio strategy involves applying the diversification criterion. Kuosmanen (2004) updated stochastic dominance criteria to get diversified portfolio by considering the cross-sectional relationship between asset returns. The implemented test was used to assess whether the US stock market has diversified across industries.

Stochastic dominance is influenced by very unlikely outcomes. It is possible that stochastic dominance tests fail because they are not verified in the extreme values of the cumulative distribution functions. In these cases, Osuna (2012) proposes to cut the distribution and only test the central region, in order to define quasi-total stochastic dominance, so-called 'tail-restricted stochastic dominance'.

Other than securities with deterministic cash-flows, financial instruments can be considered as lotteries with a continuous cumulative distribution function. When evaluating the suitability of an asset, agents are generally interested in possible future performance. It is possible to obtain the distribution of potential outcomes of financial securities at maturity by considering all their characteristics and using appropriate simulation methods, deriving the relative cumulative distribution function and then applying stochastic dominance criteria.

It is possible to rank also structured products generated by the financial engineering process using stochastic dominance criteria and to analyse investors' irrational behaviour. By comparing locally and globally capped contracts Castellano and Cerqueti (2013, 2016) demonstrated that retail investors behave irrationally. Although simpler products dominate the set of available products, agents prefer locally capped products, which are more complex and less profitable than the others.

All this evidence support doubts about the efficiency of financial prospectuses, which are usually written in a long-winded, narrative format and fail to provide information that is required to enable an objective evaluation of the product. A legal approach to debiasing financial choices should, therefore, be applied immediately to prevent opportunistic behaviour by financial institutions and to protect consumers.

1.7 Considerations

The empirical breakdown of the rational preference hypothesis in the context of decision-making reveals the limits of classical theory. Cognitive theory provides an explanation of some of the biases and the approach based on these ideas is known as behavioural finance. The main pillars of behavioural finance include prospect theory and the framing effect, which has been analysed in detail in many financial studies. Given the several psychological and cognitive biases that affect retail investors' selection of financial products, policymakers need to identify ways of correcting these biases and use policy tools to reduce the irrationality of investors' behaviour.

In last few decades financial innovation has increased the complexity of financial products. A highly complex structure and lack of standardisation is often accompanied by poor disclosure of information. The success of complex products might be increased by behavioural biases. In particular, the recent financial crisis and the cases of securities mis-selling has prompted reflection on the financial engineering underlying these

structured products and analysis of the context in which financial decisions are made. Stronger and more direct supervision is the way to prevent extreme dangerous consequences of opportunistic behaviour by financial institutions and to restore trust in the financial and banking system that, in recent times, appears to have been severely compromised. Policy measures should also be based on the nudging theory where authorities help people to manage complex situation and to avoid legal tricks. The main tools are promoting activities to increase financial literacy, regulating the modality of financial advice servicing and improving the quality of financial information. Disclosure is an issue of fundamental importance, especially with respect to complex investment products, since without access to good information it is difficult for retail investors to distinguish between a sound, appropriate investment and one that should be avoided. Furthermore, investors might be discouraged from investing because of lack of understanding of relative risks and the elements that distinguish the proposed investment.

When investors choose between products, they should know their properties and characteristics. It is, therefore, necessary to provide information on financial cash-flows, which are generally not certain. The debate on disclosure is ongoing and focuses on three main factors: risk, costs and representation of possible performance. Retail investors will not feel informed and protect until more effective, direct methods of delivering financial information are introduced.

Regardless of education, nationality and social condition all potential investors ask the same questions: 'What are the risks?' and 'How much do I risk investing in this product and comparing with those of the alternatives?'. Conscious of the deficiencies in the information delivered to retail investors, both the academy and policy makers have been trying to find the best format in which to present information that helps to answer these questions. A lot of work has been done in the last decade, leading to the provision of several legal texts which attempt to standardise the format of information that has to be included in financial prospectuses. Several consumer tests have also been conducted among potential retail investors. However, it seems that it is only a starting point and further work is needed.

After a brief description of the legal texts, the second chapter of this work analyses two methods used to disclose the risk-reward profile of non-equity financial instruments[12]: the What-if approach and Probabilistic scenarios. They are applied to two subordinated bonds in order to assess their strengths and weaknesses under the methodological point of view.

[12] This expression is used, for instance, to refer to open-ended mutual funds, SICAVs, unit-linked and index-linked financial-insurance products, and financial products issued by banks such as covered warrants, certificates and structured bonds.

1.8 References Chapter 1

AFM (2014), Netherlands Authority for the Financial Markets, mimeo.

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

Australian Securities and Investments Commission (2011). Compensation for retail investors: the social impact of monetary loss, *Report 240*.

Bai, Z., Li, H., Liu, H., & Wong, W. K. (2011). Test statistics for prospect and Markowitz stochastic dominances with applications. *The Econometrics Journal*, *14*(2), pp. 278-303.

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

Berger, P. D., & Smith, G. E. (1998). The impact of prospect theory based framing tactics on advertising effectiveness. *Omega*, *26*(5), pp. 593-609.

Bertrand, M., & Morse, A. (2011). Information disclosure, cognitive biases, and payday borrowing. *The Journal of Finance*, *66*(6), pp. 1865-1893.

Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E., & Zinman, J. (2010). What's advertising content worth? Evidence from a consumer credit marketing field experiment. *The Quarterly Journal of Economics*, *125*(1), pp. 263-306.

Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2011). How Does Simplified Disclosure Affect Individuals' Mutual Fund Choices?' In Explorations in the Economics of Aging, edited by David A. *Chicago: University of Chicago*.

Breuer, W., & Perst, A. (2007). Retail banking and behavioral financial engineering: The case of structured products. *Journal of Banking & Finance*, *31*(3), pp. 827-844.

Canner, N., Mankiw, N. G., & Weil, D. N. (1994). *An asset allocation puzzle* (No. w4857). National Bureau of Economic Research.

Castellano, R., & Cerqueti, R. (2013). Roots and effects of financial misperception in a stochastic dominance framework. *Quality & Quantity*, *47*(6), pp. 3371-3389.

Castellano, R., & Cerqueti, R. (2016). A theory of misperception in a stochastic dominance framework and its application to structured financial products. *IMA Journal of Management Mathematics*, *29*, pp. 23-37.

Celerier, C., & Vallee, B. (2013). What drives financial complexity? A look into the European retail structured products market. HEC Paris Research Paper No. FIN-2013-1013, available at: http://ssrn. com/abstract, 2289890.

Celerier, C., & Vallee, B. (2014). The motives for financial complexity: An empirical investigation. HEC Paris Research Paper No. FIN-2013-1013, 2014. Available online at: http://papers. ssrn. com/sol3/papers. cfm.

CESR (2010a). CESR guidelines on the methodology for the calculation of the synthetic risk and reward indicator in Key Investor Information Document. CESR/10-673.

CESR (2010b). CESR guidelines on the selection and presentation of performance scenarios in the Key Investor Information document (KII) for structured UCITS. CESR/10-1318.

Chan, K., Covrig, V., & Ng, L. (2005). What determines the domestic bias and foreign bias? Evidence from mutual fund equity allocations worldwide. *The Journal of Finance*, *60*(3), pp.1495-1534.

Chater, N., Huck, S., & Inderst, R. (2010). Consumer decision-making in retail investment services: A behavioural economics perspective. *Report to the European Commission/SANCO*.

Choi, J. J., Laibson, D., & Madrian, B. C. (2009). Why does the law of one price fail? An experiment on index mutual funds. *The Review of Financial Studies*, *23*(4), pp. 1405-1432.

Clark-Murphy, M., & Soutar, G. N. (2004). What individual investors value: Some Australian evidence. *Journal of Economic Psychology*, *25*(4), pp. 539-555.

Cooper, I., & Kaplanis, E. (1994). Home bias in equity portfolios, inflation hedging, and international capital market equilibrium. *The Review of Financial Studies*, *7*(1), pp. 45-60.

Coval, J. D., & Moskowitz, T. J. (1999). Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance*, *54*(6), pp. 2045-2073.

Cox, R., & de Goeij, P. (2016). The Effects of Risk Framing, Balanced Information and a Regulatory Seal of Approval on Investor Perception and Behaviour.

Cox, A. D., Cox, D., & Zimet, G. (2006). Understanding consumer responses to product risk information. *Journal of Marketing*, *70*(1), pp. 79-91.

Dambe, K., Hunt, S., Iscenko, Z., & Brambley, W. (2013). Applying behavioural economics at the Financial Conduct Authority. *Occasional Paper No.1*, Financial Conduct Authority.

Davidson, R., & Duclos, J. Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, *68*(6), pp. 1435-1464.

de Goeij, P., Hogendoorn, T., & Campenhout, G. V. (2014). Pictures are worth a thousand words: Graphical information and investment decision making.

Desanctis, G., & Jarvenpaa, S. L. (1989). Graphical presentation of accounting data for financial forecasting: An experimental investigation. *Accounting, Organizations and Society*, *14*(5-6), pp. 509-525.

Diacon, S. (2004). Investment risk perceptions: Do consumers and advisers agree? *International Journal of Bank Marketing*, *22*(3), pp. 180-199.

Diacon, S., & Hasseldine, J. (2007). Framing effects and risk perception: The effect of prior performance presentation format on investment fund choice. *Journal of Economic Psychology*, *28*(1), pp. 31-52.

Dolan, P., Elliott, A., Metcalfe, R., & Vlaev, I. (2012). Influencing financial behaviour: from changing minds to changing contexts. *Journal of Behavioral Finance*, *13*(2), pp. 126-142.

Driver, R., Chater, N., Cheung, B., Latham, M., Lewis, R., & Stott, H. (2010). Helping consumers understand investment risk. Association of British Insurers (ABI). *Research Paper no.25.*

Döbeli, B., & Vanini, P. (2010). Stated and revealed investment decisions concerning retail structured products. *Journal of Banking & Finance*, *34*(6), pp. 1400-1411.

European Commission (2012). Proposal for a Regulation of the European Parliament and of the Council on key information documents for investment products. *Commission Staff Working Document: Impact Assessment.*

Erta, K., Iscenko, Z., Hunt, S. and Brambley, W. (2013). Appling behavioural economics at the Financial Conduct Authority, *FCA Occasional Paper 1.*

Ferris, S. P., Haugen, R. A., & Makhija, A. K. (1988). Predicting contemporary volume with historic volume at differential price levels: Evidence supporting the disposition effect. The Journal of Finance, 43(3), pp. 677-697.

Financial Conduct Authority (2014). General Insurance Add-Ons: Final Report – Confirmed Findings of the Market Study, *Market Study, MS14/1*.

Financial Conduct Authority (2015). Two plus two makes five? Survey evidence that investors overvalue structured deposits*, Occasional paper no. 9.*

Financial Services Authority (2013). Assessing the quality of investment advice in the retail banking sector. A mystery shopping review.

Fishburn, P.C. (1964). Decision and Value Theory. Wiley, New York.

Fischer, R. (2007). Do investors in structured products act rationally? Technical report, European Business School.

French, K. R., & Poterba, J. M. (1991). Investor Diversification and International Equity Markets. *American Economic Review*, *81*(2), pp. 222-226.

Gentile, M., Linciano, N., & Siciliano, G. (2006). Le scelte di portafoglio delle famiglie italiane e la diffusione del risparmio gestito. *Banca Impresa Società*, *25*(3), pp. 425-476.

Gentile, M., Linciano, N., & Soccorso, P. (2016). Financial Advice Seeking, Financial Knowledge and Overconfidence. Evidence from the Italian Market, *Quaderni di Finanza n.83*, CONSOB.

Glenzer, F., Gründl, H., & Wilde, C. (2014). 'And Lead Us Not into Temptation': Presentation Formats and the Choice of Risky Alternatives. ICIR Working Paper Series No. 16/14.

Godwin, A., & Ramsay, I. (2015). Financial products and short-form disclosure documents: a comparative analysis of six jurisdictions. *Capital Markets Law Journal*, *10*(2), pp. 212-238.

Godwin, A., & Ramsay, I. (2016). Short-form disclosure documents—an empirical survey of six jurisdictions. *Capital Markets Law Journal*, *11*(2), pp. 296-316.

Guiso, L., & Jappelli, T. (2006). Information acquisition and portfolio performance. CEPR Discussion Paper n. 5901, www.cepr.org.

Hadar, J., & Russell, W. R. (1969). Rules for ordering uncertain prospects. *The American economic review*, *59*(1), pp. 25-34.

Haigh, M. S., & List, J. A. (2005). Do professional traders exhibit myopic loss aversion? An experimental analysis. *The Journal of Finance*, *60*(1), pp. 523-534.

Hanoch, G., & Levy, H. (1969). The efficiency analysis of choices involving risk. *The Review of Economic Studies*, *36*(3), pp. 335-346.

Hens, T., & Rieger, M. O. (2008). *The dark side of the moon: structured products from the customer's perspective* (No. 459). Working paper.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological science*, *15*(8), pp. 534-539.

Hillenbrand, A., & Schmelzer, A. (2015). Beyond information: Disclosure, distracted attention, and investor behaviour. MPI Collective Goods Preprint, No. 2015/20.

House of Commons Committee of Public Accounts (2016). Financial Services Mis-selling: Regulation and Redress. Forty-first Report of Session 2015–16.

Hung, A., Heinberg, A., & Yoong, J. (2010). Do Risk Disclosures Affect Investment Choice? RAND Labour and Population Working Paper W-788.

Ibrekk, H., & Morgan, M. G. (1987). Graphical communication of uncertain quantities to nontechnical people. *Risk analysis*, *7*(4), pp. 519-529.

IFF Research and YouGov (2009). UCITS Disclosure Testing Research Report, prepared for European Commission.

Jolls, C., & Sunstein, C. R. (2006). Debiasing through law. *The Journal of Legal Studies*, *35*(1), pp. 199-242.

Jones, S. K., Frisch, D., Yurak, T. J., & Kim, E. (1998). Choices and Opportunities: Another Effect of Framing on Decisions. *Journal of Behavioral Decision Making*, *11*, pp. 211-226.

Kahneman, D., & Riepe, M. W. (1998). Aspects of investor psychology. *The Journal of Portfolio Management*, *24*(4), pp. 52-65.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, pp. 263-291.

Karlsson, A., & Nordén, L. (2007). Home sweet home: Home bias and international diversification among individual investors. *Journal of Banking & Finance*, *31*(2), pp. 317-333.

Kaufmann, C., & Weber, M. (2013). Sometimes less is more–The influence of information aggregation on investment decisions. *Journal of Economic Behavior & Organization*, *95*, pp. 20-33.

Kaufmann, C., Weber, M., & Haisley, E. (2013). The role of experience sampling and graphical displays on one's investment risk appetite. *Management Science*, *59*(2), pp. 323-340.

Ke, D., Ng, L., & Wang, Q. (2010). Home bias in foreign investment decisions. *Journal of International Business Studies*, *41*(6), pp. 960-979.

Kilka, M., & Weber, M. (2000). Home bias in international stock return expectations. *The Journal of Psychology and Financial Markets*, *1*(3-4), pp. 176-192.

Klapper, L., Lusardi, A., & Van Oudheusden, P. (2015). Financial Literacy Around the World. Standard & Poor's Ratings Services Global Financial Literacy Survey., Access mode: http://media. mhfi. com/documents/2015-Finlit_paper_17_ F3_SINGLES. pdf.

Kruschke, J. K., & Johansen, M. K. (1999). A model of probabilistic category learning. *Journal of Experimental Psychology-Learning Memory and Cognition*, *25*(5), pp. 1083-1119.

Kumar, A., & Lim, S. S. (2008). How do decision frames influence the stock investment choices of individual investors? *Management science*, *54*(6), pp. 1052-1064.

Kuosmanen, T. (2004). Efficient diversification according to stochastic dominance criteria. *Management Science*, *50*(10), pp. 1390-1406.

Larrick, R. P., Morgan, J. N., & Nisbett, R. E. (1990). Teaching the use of cost-benefit reasoning in everyday life. *Psychological Science*, *1*(6), pp. 362-370.

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational behavior and human decision processes*, *76*(2), pp. 149-188.

Levy, M., & Levy, H. (2002). Prospect theory: much ado about nothing? *Management Science*, *48*(10), pp. 1334-1349.

Levy, H., & Levy, M. (2003). Prospect theory and mean-variance analysis. *Review of Financial Studies*, *17*(4), pp. 1015-1041.

Linciano, N., Lucarelli, C., Gentile, M., & Soccorso, P. (2018). How financial information disclosure affects risk perception. Evidence from Italian investors' behaviour. *The European Journal of Finance*, *24*(15), pp. 1311-1332.

Loewenstein, G., Sunstein, C. R., & Golman, R. (2014). Disclosure: Psychology changes everything. *Annual Economic Review* 6.

Lucarelli, C. (2011). La rappresentazione del rischio. Alcune evidenze empiriche. *Quaderni di Finanza n.68*, CONSOB, pp. 25-48.

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

Lusardi, A., & Mitchell, O. S. (2011). Financial literacy around the world: an overview. *Journal of pension economics & finance*, *10*(4), pp. 497-508.

Mussweiler, T., & Schneller, K. (2003). ' What Goes Up Must Come Down'-How Charts Influence Decisions to Buy and Sell Stocks. *The Journal of behavioral finance*, *4*(3), pp. 121-130.

National Audit Office (2016). Financial Services Mis-selling: Regulation and Redress.

Office of Fair Trading (2010). The impact of price frames on consumer decision making. *A report by London Economics.*

Office of Fair Trading (2012a). Airlines/Payment Surcharges: Investigation into pricing practices by 14 airlines.

Office of Fair Trading (2012b). Eight supermarkets sign up to OFT principles on special offers and promotions. Press release.

Osuna, E. E. (2012). Tail-restricted stochastic dominance. *IMA Journal of Management Mathematics*, *24*(1), pp. 21-44.

Rieger, M. O. (2012). Why do investors buy bad financial products? Probability misestimation and preferences in financial investment decision. *Journal of Behavioral Finance*, *13*(2), pp. 108-118.

Scarsini, M. (1986). Comparison of random cash flows. *IMA Journal of Management Mathematics*, *1*(1), pp. 25-32.

Seasholes, M. S., & Zhu, N. (2010). Individual investors and local bias. *The Journal of Finance*, *65*(5), pp. 1987-2010.

Shapira, Z., & Venezia, I. (2001). Patterns of behavior of professionally managed and independent investors. *Journal of Banking & Finance*, *25*(8), pp. 1573-1587.

Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of finance*, *40*(3), pp. 777-790.

Shefrin, H., & Statman, M. (1993). Behavioral aspects of the design and marketing of financial products. *Financial Management*, pp. 123-134.

Shefrin, H., & Statman, M. (2000). Behavioral portfolio theory. *Journal of financial and quantitative analysis*, *35*(2), pp. 127-151.

Shefrin, H. M., & Thaler, R. H. (1988). The behavioral life-cycle hypothesis. *Economic inquiry*, *26*(4), pp. 609-643.

Sriboonchitta, S., Wong, W.K., Dhompongsa, S., & Nguyen, H.T. (2010). Stochastic Dominance and Applications to Finance, Risk and Economics. Chapman & Hall, New York.

Strong, N., & Xu, X. (2003). Understanding the equity home bias: Evidence from survey data. *Review of Economics and Statistics*, *85*(2), pp. 307-312.

Tesar, L. L., & Werner, I. M. (1995). Home bias and high turnover. *Journal of international Money and Finance*, *14*(4), pp. 467-492.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science*, *4*(3), pp. 199-214.

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

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141-162). Springer Netherlands.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, New Series, Vol. 211, No. 4481.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business*, S251-S278.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, *5*(4), pp. 297-323.

Vlaev, I., Chater, N., & Stewart, N. (2009). Dimensionality of risk perception: Factors affecting consumer understanding and evaluation of financial risk. *The Journal of Behavioral Finance*, *10*(3), pp. 158-181.

Von Neumann, J., & Morgenstern, O. (1947). Theory of games and economic behavior, 2nd rev.

Walther, T. (2015). Key investor documents and their consequences on investor behavior. *Journal of Business Economics*, *85*(2), pp. 129-156.

Wang, A., & Dowding, T. (2010). Effects of visual priming on improving web disclosure to investors. *The Journal of Behavioral Finance*, *11*(1), pp. 11-20.

Wang, M., Keller, C., & Siegrist, M. (2011). The less You know, the more You are afraid of - A survey on risk perceptions of investment products. *Journal of Behavioral Finance*, *12*(1), pp. 9-19.

Weathers, D., Swain, S. D., & Carlson, J. P. (2012). Why consumers respond differently to absolute versus percentage descriptions of quantities. *Marketing Letters*, *23*(4), pp. 943-957.

Weber, M., & Welfens, F. (2008). Splitting the Disposition Effect: Asymmetric Reactions Towards 'Selling Winners' and 'Holding Losers'. http://ssrn.com/abstract=1176422.

Weber, E. U., Siebenmorgen, N., & Weber, M. (2005). Communicating asset risk: how name recognition and the format of historic volatility information affect risk perception and investment decisions. *Risk Analysis*, *25*(3), pp. 597-609.

Willis, L. E. (2008). Against financial-literacy education. *Iowa L. Rev.*, *94*, 197.

2 Possible approaches to disclose risk-reward profile of non-equity financial instruments.

2.1 Introduction

An international framework characterised by increasing globalisation of financial markets and integration of banking, asset management and insurance activities has generated a rapid growth in the complexity of traded financial instruments. Disclosure of the essential characteristics of non-equity investment products became important for the protection of retail investors and is intended to boost their awareness of risk during the investment decisions and to safeguard confidence in the financial system. Approaches to regulation of disclosure are quite variable at both European and national level. Several regulatory products are issued and differentiated in relation to the category of the issuer. This regulatory fragmentation ignores the strong similarities in the financial engineering of products issued by subjects belonging to different categories (e.g. banks, investment firms and insurance companies), thus jeopardising the creation of a level playing field.

Following discussion amongst policy-makers, stakeholders and academics about the best format to adopt for disclosure of information, there is now a consensus that the traditional narrative descriptions of the various risks and costs is no longer effective and, therefore, it would be better to use synthetic indicators which are immediately comprehensible to investors and defined in relation to robust and objective quantitative metrics. However, it is not clear which approach is more appropriated to represent the single characteristics of a product (for instance indicators of risk, future performance representation and costs profile) based on to the criteria established by the CESR to assess the weakness and strengths of a presentation format (namely that it should be engaging, understandable, comparable, compatible, balanced and cover a range of types of product) and its underlying methodology (which should be reliable, robust, broadly applicable, stable, comparable, feasible/proportional and easy to reply by supervisor)[13].

This chapter is organised as follows. In Section 2 the evolution of the regulatory framework on disclosure of complex financial products is described. After a brief introduction to the main problems with the financial information set provided to investors (Section 3), two methods of representing the possible future performance of a financial product, namely What-if (Section 4) and Probabilistic scenarios (Section 5), are analysed. Section 6 applies these methods to two financial securities. In order to assess the weaknesses and strengths of the approaches two aspects are analysed: possible distortions of the represented amounts due to model risk (Section 7) and how to update the information set provided by each scheme following the variation of market conditions during the life of the investment (Section 8). Then, Section 9 summarises the results, comparing the What-if analysis and the Probabilistic scenarios. Finally, some general conclusions on the two methods are set out in Section 10.

2.2 Evolution of the normative framework on disclosure of non-equity products

Cases of mis-selling highlight the ineffectiveness of the documentation about financial products delivered to potential investors. Therefore, several people have recommended that the sale of particularly risky investment products to retail customers should be prohibited, rather than requiring further provisions for a

[13] For further details on the listed criteria see Table 6.1: *Assessment of the criteria for presentation format* and Table: 6.2: *Assessment of the criteria for underlying methodologies* in Annex III.

clearer information set, because these products are generally evaluated as 'toxic' finance. However, it is worth noting that not all structured financial products are regarded as 'toxic' and that they (derivatives for example) may contribute to the efficient functioning of the financial system. Eliminating these products would have a negative impact on the overall management of risks. Derivatives and other complex securities are in fact the basis of the modern financial system, making it possible to hedge against risks and providing operators with opportunities to diversify their portfolios and speculate on the basis of reference values. What is important, therefore, is to provide accurate information on the risks associated with specific contracts, rather than prohibiting trading in them, reducing opacity should help investors to identify securities that are too risky or 'toxic'.

Adequate regulation of financial complexity represents one of the greatest challenges of modern financial markets and regulation activities in order to reduce the potential for failures (Schwarcz, 2009). During the last twenty years the European Union has introduced several laws defining compulsory elements of a prospectus for the public offering of non-equity investment products. The variety of directives and regulations reflects the intention of European policy-makers to tailor the regulation of prospectuses[14] to the type of issuer of financial instruments, in other words a distinction is drawn between banks, asset management companies/SICAVs and insurance companies.

The UCITS (Undertakings for Collective Investment in Transferable Securities) directives[15] and subsequent recommendations of the European Commission (EC) regarding open-ended mutual funds and SICAVs grant Member States the option to require prior authorisation for the publication of the prospectus, which must include two separate documents: the Key Investor Information Document (KIID), which must be given to potential investors before they sign a contract, and the Full Prospectus, to be issued to potential investors upon request. The aim of the KIID is to illustrate, in a few pages, essential information about the risk-return profile and costs of the investment. The KIID contains also a concise representation of costs and expected returns of a mutual fund, based on synthetic risk indicators and on charts and tables. As regards the set of information provided in the prospectus, the UCITS directives give Member States the discretion to define specific templates for the prospectus and to specify the format and the minimum content, provided that it is consistent with the aforementioned directives and recommendations. When a Member State chooses to exercise this discretion it can only define templates for the prospectuses of mutual funds offered by issuers with registered offices in its territory; its competence does not extend to the prospectuses of mutual funds offered within its borders by issuers with a registered office in another Member State.

Directive 2003/71/EC[16] (the 'Prospectus directive'), which also refers to non-equity investment products issued by banks (i.e. ordinary and structured bonds, covered warrants and certificates), requires prior approval of the prospectus and specific authorisation by the competent authorities of the home Member State. In general, the information on the product is contained in two documents, which do not have to be delivered to investors, the Base Prospectus and the Final Terms. The subsequent Commission regulation 809/2004/EC, which is directly applicable in all Member States, implements the Prospectus directive, defining the format and the minimum content of the Base Prospectus, which follows a narrative logic and provides a detailed listing of all the costs and risks of the investment. This document is organised according to a product information sheet template and so potential investors should be able to base their assessment of the proposed investment on it. Basically, the Base Prospectus discloses the intention of the issuer to launch a public offering of an investment product and provides some general information about it, whereas the Final Terms provides a detailed description.

[14] This aims to standardise the format and minimum content of prospectus templates, align authorisation and publication regimes and explain how to transmit the prospectus to the competent Authority and deliver the documentation to investors.

[15] For the scope of this research, the more important directive is the so-called UCITS IV (i.e. Directive 2009/65/EC).

[16] Further modified by Directive 2010/73/EU.

On 30 June 2017 EU Regulation 1129/2017 regulation (hereafter, 'Prospectus regulation') was published in the Official Journal of the European Union. It deals with the structure of prospectuses and procedures for publishing public offerings and obtaining admission to the trading of securities in the regulated markets of the European Union. It brings about the progressive repeal of the provisions of the previously mentioned Prospectus directive. Within the scope of this revision, the European Commission evaluated changes to the format and to the contents of the issued documentations. In particular, documentation should provide retail investors with a clear understanding of the essential features of a financial product via a sheet including information on potential returns, the payoff structure, all direct and indirect costs, the financial risks, and terms and conditions of any capital guarantee. Increasing the effectiveness of pre-contractual information in protecting investors is the main driver of Commission policy since it found out that previous regulation had stimulated the drafting of long, complex and not easily comprehensible prospectuses. The European Commission is now considering various alternatives, including making a summary document[17] as an integral part of the prospectus or an approach similar to that underlying the definition of the KID[18].

A consequence of increased financial engineering is the introduction of the Packaged Retail and Insurance-based Investment Products (PRIIPs), which represent the core of the retail investment market[19]. These are investment products that banks typically offer to consumers who, for example, want to save for a specific objective such as a house purchase or a child's education. Despite their potential benefits for retail investors, PRIIPs are often complicated, lacking in transparency and the information which institutions make available to investors when selling these products is overly complex. The documentation often contains too much jargon and can make it difficult to compare investment products belonging to different categories. Since the institutions selling these products often act as advisors, conflicts of interest may arise, resulting in advice which may not be in the investor's best interests. After lengthy negotiation with stakeholders (ESAs JC[20], 2015) the EU has adopted regulations on PRIIPs[21] that address these problems but raise a number of important concerns, challenges, and criticisms (Colaert, 2016). The Regulation obliges those who produce or sell these investment products to provide investors with key information document (KID[22]). The KID should be no more than three pages long and present the investment product clearly: it should include the name of the product and the identity of the producer; the types of investor for whom it is intended; the risk and reward profile of the product (i.e. a summary risk indicator, the possible maximum loss of invested capital and appropriate performance scenarios); all the costs payable when investing in the product; information about how and to whom an investor can make a complaint if there is a problem with the product or the person producing, advising on or selling the product.

[17] The Prospectus regulation reiterates that the summary is preliminary to the prospectus. The contents of the summary document should be accurate, clear, correct and not misleading and provide the key information that the investor needs to know to make a well-informed judgement of the potential investment. In formal terms the requirement for minimum information to be included in the summary is eliminated and so issuers have greater discretion as to the informational content. It is expected that the document will not exceed seven pages and contain no more than fifteen risk factors. The summary is also expected to be divided into the following four sections: (i) introduction and warnings; (ii) key information on the issuer; (iii) key information on the securities; and (iv) key information relating to the public offering and/or admission to negotiation.

[18] In order to harmonise the regulations on prospectuses with the disclosure obligations of the EU Regulation 1286/2014 (so-called Regulation on PRIIPs), the Prospectus regulation gives the issuer the right to replace the section on the key information about the securities with the Key Information Document (KID) if the latter document is required for the type of product in question (see Regulation 1286/2014 in the following paragraphs).

[19] There are two categories of PRIIP: (i) non-insurance products, whose cash flows are not fixed because they are influenced by market indices or other predetermined baskets not directly bought by investor (e.g. structured bonds, securitised assets, derivatives) and (ii) insurance products, in the case of which maturities or repayments are fully or partially, directly or indirectly, linked to market fluctuations (e.g. unit- and index-linked securities).

[20] The ESAs JC is the Joint Committee of the European Supervisory Authorities, namely the European Banking Authority (EBA), the European Insurance and Occupational Pensions Authority (EIOPA) and the European Securities and Markets Authority (ESMA).

[21] In 2014 the EU introduced Regulation 1286/2014 (so-called Regulation on PRIIPs) and supplementary provisions (Regulation 2017/653) laying down regulatory technical standards for the presentation, content, review and revision of key information documents and the conditions for providing such documents. In 2017 the EU also published guidelines for the interpretation of the Regulation on PRIIPs (2017/C 2018/02).

[22] To clarify, the KID of PRIIPs are different from the KIID of UCITS. Although the form and structure of the documents are similar, there are profound differences in the methodology used to compute the risk indicators, the presentation of the performance scenarios and the various elements dealing with costs.

This regulation aims to deal with two important concerns in the field of investor protection: (i) investors do not read, understand or digest extensive or technical information and (ii) they do little to compare the products and services of different financial institutions. The first concern relates to the long-established problems of information overload and lack of rationality. The PRIIPs Regulation tackles those problems by introducing a short information document that uses clear, succinct and understandable language and supportive visual indicators. The second concern relates to the fact that information documents are often hard to compare because of their different formats. Standardisation seems an obvious answer to this problem. The use of standardisation, however, should be applied to all products even when they have different characteristics. The PRIIPs Regulation is innovative in this respect, being the first piece of 'horizontal' or 'cross-sectoral' legislation in the history of EU financial regulation; it covers banking, investment and insurance products, instead of applying to just one of these sectors.

Directive 2004/39/EC (also referred to as MiFID - Markets in Financial Instruments Directive) and the related second level directive (Directive 2006/73/EC) set out additional regulations on disclosure and conduct of business in relation to distribution of non-equity financial products. According to these directives the offer prospectus should be the primary source of information on the risk-return profile and on the costs of investments in open-ended mutual funds and SICAVs. Suppliers are required to supplement the prospectus information where it would be insufficient to comply with the principles of suitability or appropriateness to the investment in question. It is worth noting that, as each directive, the MiFID establishes principles in general terms and national regulators are required to implement them through the issuing of detailed rules.

Developed in response to the global financial crisis in 2008, Directive 2014/65/EU (MiFID II) affects all businesses involved in the manufacturing, distribution and trading of financial instruments in the European Union. Among other things[23], it requires that every portfolio is constructed with the client's objectives in mind and that portfolios are monitored for suitability on an ongoing basis. It also prohibits inducements for independent advice[24] and is designed to ensure that anyone providing investment advice puts the client's best interest first. Firms should review their entire advice process to make sure that the investments they offer are suited to investors' needs. The directive also emphasises transparency, particularly with respect to fees, making it much easier for investors to understand what services they are paying for throughout the investment process. Asset managers are required to provide investors and advisers with more information about the costs of their products, both ex-ante and ex-post underwriting. These costs must be broken down into four main components: the ongoing charge for the fund; one-off fees (such as entry and exit fees); incidental fees (such as performance fees paid by the fund); and transaction fees related to the investment product.

The standardisation of the contents in the prospectus is shared among regulators. However, the described regulatory framework shows that there is significant heterogeneity in regulatory provisions affecting the public offering of financial products which, however, may share similar financial engineering. A heterogeneous regulatory framework may generate three undesirable consequences: (i) regulatory arbitrage opportunities, which could be exploited by choosing the category of issuer; (ii) investors find it difficult to make a meaningful and fair comparison of the risk-return profile of products even when they share the same financial structure; and (iii) manufacturers and advisors' compliance costs increase.

The aim of simplifying offer documentations and introducing a wide regulatory framework should be to ensure a harmonised approach to the various product categories. Therefore, European policy-makers should intervene to align the disclosure requirements for products issued by subjects belonging to distinct categories

[23] The directive can be broken down into four main areas: internal and external controls, market structure, investor protection and fee transparency. Only the last two are relevant here.

[24] MiFID II prohibits an EU investment firm which carries out portfolio management or provides investment advice from accepting and retaining third party inducements (fees, commissions or monetary and non-monetary benefits) in relation to the provision of services to clients.

and to ensure effective harmonisation of regulations across Member States, in order to protect investors and to guarantee competition between the various national financial systems. Implementation of a level playing field requires the adoption of common standards for risk measurement methods, which enables investors to compare the numerous financial products offered in the market fairly.

2.3 Representation of future performance

In general, the price is the first aspect that agents look at when deciding which product to buy, but in the financial context the price of a security is closely related to the mean of the probability distribution of its future performances (first moment of the distribution) and, as is well known, it does not provide any indication of the risks associated with a financial product so that the variability of the possible value should be identified (second moment of the distribution). Although the first two moments are provided, the information set could sometimes be insufficient. The pattern of expected returns may be complex and irregular, such that the probability distribution is characterised by features such as multimodality, asymmetry or kurtosis, making the first two moments of the distribution insufficient for a complete and useful representation of the security. For these reasons, the main goal of regulation is to ensure that retail investors are provided with adequate information about the key characteristics of financial products (not just price) and the associated risks and costs, so that they can select investments that suit their needs.

There are three main pieces of information that an investor needs to make an investment choice: information about risks, all costs and expected future performance. Regulators and academics have been debating the best ways to represent the future performance of non-equity products[25]. The traditional approach used by regulators to outline disclosures focused solely on the set of information to present without taking into consideration its format. For instance, it is now recognised that framing of information so that is both salient and easy for consumers to understand is crucial to reducing the impact of cognitive biases on investment decisions.

As for other investments' features, different representation formats of future performance scenarios could be addressed in three main categories: narrative, single visual element and multiple visual elements. If a narrative format is adopted no visual presentation of performance is given and the possible outcomes are explained only in narrative terms. There is a consensus on the weaknesses of this approach: a textual representation is less engaging than a visual representation and may be difficult to standardise, which has implications for comparability. However, comprehensive and simple wording may well be understood and may be a useful supplement to visual representations of possible outcomes such as tables, scales or graphs. The single representation differs from the multiple representation because in the first case only one format is used to represent one characteristic of the security while in the other case several visual elements could be displayed to investors.

Although visual representations are preferred to textual representations there is still an active debate about what format is most appropriate and how detailed the documentation should be. The set of axioms provided by Gilboa and Schmeidler (1989) included the notion of 'uncertainty aversion' which is defined as a preference for knowing probabilities rather than having to infer them subjectively from limited information[26]. Following this axiom, some studies are conducted to find its empirical evidence in the financial context. A more recent study (Glenzer *et al.*, 2014) showed that presenting the 75th and 25th percentiles of the distribution of returns, instead of the 95th and 5th percentiles, increased participants' appetite for risk as

[25] The following sections of the thesis are focused only on the future performance issue without analysing the other aspects.

[26] The distinction between ambiguity aversion and risk aversion is important but subtle. Risk aversion is defined as the preference for less risky options in contexts where probabilities can be assigned to possible outcomes. An ambiguity-averse individual would rather choose an alternative where the probability distribution of the outcomes is known over one where the probabilities are unknown (Epstein, 1999).

reflected in preference order: reducing the emphasis on downside risk and upside potential returns increases risk taking substantially. More relevantly, the authors found that showing the full probability distribution caused a further decrease in risk taking, indicating that subjects react to having the full loss potential of riskier investments. Another study (Donker *et al.*, 2013) showed that presenting the probability distribution as a hundred little dots distributed over the possible range of outcomes (i.e. as frequencies) increased consumers understanding of risk significantly.

These evidences are the basis of some of the methodologies implemented for performance representation which disclose the likelihood of occurrence of specific events computed from historical data or modelled. Indeed, possible performance scenarios could be selected on the basis of the probability distribution of expected returns, which could be derived from historical data (backward-looking) or by modelling the market instruments underlying the financial product (forward-looking). Recently European regulators (see Delegated Regulation n.2017/653) considered the results from empirical studies and introduced a requirement to show possible returns of PRIIPs by partitioning the probability distribution of future performance and showing four scenarios: (i) a favourable scenario (10th percentile of the distribution of the value of the product at maturity); (ii) a moderate scenario (50th percentile); (iii) an unfavourable scenario (90th percentile); (iv) a stress scenario which is computed on the basis of specific hypotheses set out in Annex IV of the Regulation.

Despite the progress made by regulators the debate about presentation of information continues. The remaining sections of this chapter analyse two methods that are the subject of some controversy in the academic world and amongst policy-makers, namely the use of the What-if approach and the Probabilistic scenarios.

2.4 What-if approach

2.4.1 General aspects

In general, What-if analysis is an intensive simulation of data whose objective is to capture the behaviour of complex systems, such as a corporate business or a part of such a business, via specific hypotheses called scenarios (Rizzi, 2016). The analysis estimates how variations in a set of input variables, considering possible impacts on intermediate variables, affect certain output variables through reference to a simplified model of the phenomenon of interest. As immediately understandable, formulating scenarios allows the determination of a hypothetical world and, thus, it remains a partial reproduction of reality (Hull, 2010).

What-if analysis is based on a simulation model that defines a set of complex relationships between variables corresponding to significant entities in the domain of interest. A simulation model supports one or more scenarios, each of which describes one or more alternative future paths. A scenario is characterised by a set of input variables and underlying scenario parameters that the analyst must evaluate to run the model and obtain a forecast (Golfarelli *et al.*, 2006). The scenario parameters provide technical information that is relevant to the evaluation and specify the relationships assumed by the evaluation. Each scenario should give rise to different simulations, one for each set of the input variables and scenario parameters. Therefore, given the general framework, an appropriate procedure would involve: (i) determination of which factors should be used to construct the scenarios; (ii) determination of the number of scenarios to be analysed for each factor, given that analysing a lot of scenarios (i.e. a fine-grained analysis) would give a more realistic view of the phenomenon, but make it more difficult to collect the required information and differentiate between scenarios; (iii) estimation of performance under the different scenarios. It is also important not to confuse the What-if representation with ‘sensitivity analysis’, which is an assessment of how sensitive the behaviour of a system is to a slight change in one or more parameters.

Although the What-if approach are commonly used in several fields[27], there have been very few attempts to define a clear and precise methodology and modelling procedure. As far as the economic field is concerned, What-if analysis is mainly used in marketing, production planning and business planning. It is important to note, however, that there are no significant applications of this methodology in the financial field.

2.4.2 Methodological procedure for financial products

Using the What-if methodology for representing future performance of financial instruments, the scenario should be computed by using the structure of the cash-flows[28] as well as making assumptions about certain market conditions that could cause a positive, neutral or negative result for the investor. This method consists, therefore, of calculating the expected return of the product under particular scenarios. The purpose is to indicate to a potential investor how the security operates by giving a series of answers to the question 'how much could I get if event X happened?'.

The What-if method is considered the clearest, simplest and least misleading way of representing potential future returns for structured UCITS (see article. 36 of Regulation 583/2010/EU). Guidelines on the What-if procedure are set out by the Committee of European Securities Regulators (CESR, 2010a, 2010b), although detailed instructions are not provided. According to the Guidelines, manufacturers should represent performance in at least three scenarios to illustrate how pay-out is affected by market conditions. The underlying hypotheses should be based on reasonable assumptions about future market conditions and price movements. In particular, the scenarios should capture:

- the formula underlying the performance of the financial instrument under market conditions leading to an unfavourable, medium and favourable result (labelled as negative, neutral and positive scenario, respectively);
- the specific characteristics of the formula;
- situations in which some mechanisms of the formula have a positive or negative impact on the final performance;
- positive and negative aspects of the security and its future performance.

In order to avoid misunderstanding the scenarios should be accompanied by a statement that they are illustrative examples, included only to describe how cash-flows are determined, and do not constitute forecasts of future performance. Moreover, it should be made clear that the scenarios may not be equally probable. The scenarios may be presented to investors via tables or graphs, but the management company must select the clearest way to represent the possible performance of the product.

Although defined for structured funds, this approach can also be applied to other non-equity products such as subordinated bonds, securities with a floating interest rate structures and products linked to market indices.

[27] The method is used in many fields, from physics to chemistry, biology, medicine, economics and psychology.
[28] For example, the interest rate of return, caps and floors and indexing of returns.

2.5 A risk-based method

2.5.1 General aspects

A risk-based method for non-equity products could be used to disclose information (Minenna *et al.*, 2009; Minenna, 2011 and Movement of Risk Transparency[29]). It is based on possible returns of a financial instrument at maturity. This method is built on three pillars and uses probabilistic tools. It is an objective method of determining and representing the value of three synthetic risk indicators that are useful for comparing the various non-equity products available in the global financial market: (i) the price unbundling and the Probabilistic scenarios (first pillar); (ii) the degree of ongoing risk (second pillar); and (iii) the recommended time horizon for the investment (third pillar).

The first pillar relies on two complementary tables, the financial investment table and the Probabilistic scenarios table. Core information is extracted from the risk-neutral density distribution of non-equity products considered at two specific points in time: the issue date and the end of the recommended investment time horizon. The second pillar highlights the degree of ongoing global risk: a synthetic indicator summarises the overall riskiness of the product over the entire recommended investment period. By working on the simulated trajectories of the product's value process used in the first pillar, it is possible to analyse their variability through the volatility measure. The third pillar shows a recommended investment period for a security, based on its financial structure and cost regime.

This thesis focuses on the representation of risk-reward profiles using Probabilistic scenarios tables as part of the first pillar. Using the implied probability distribution of financial products and selecting appropriate reference thresholds, manufacturer can identify events that are crucial from the perspective of retail investors. The number of events to be disclosed should allow an effective reading of the main statistical features of the distribution (e.g. multimodality and asymmetry) but the maximum number of partitions of the distribution is three. The events should be accompanied by their relative probabilities of occurrence and absolute indicator of performance, namely the mean of the product value in each event.

All the data are collected in a table, which provides a synthetic representation of the probability distribution of the investment's possible payoffs at the end of the recommended investment or holding period.

2.5.2 Methodological procedure for Probabilistic scenarios

In the case of non-equity products with simple structures, the expected value at maturity and hence the price can be calculated using closed-form formulae (following the reasoning by Black and Sholes, 1973). However in the case of most financial products, calculation of the expected value at maturity requires the preliminary determination of the distribution of 'pricing at maturity' using a simulation procedure. The probabilistic comparison underlying the first pillar of the risk-based approach requires the numerical simulation of both the final value of capital invested in the security (C) and the notional capital invested in a reference financial asset[30] (hereinafter RFA) over the recommended investment period (T). Hence the value of any non-equity product can be described by a specific stochastic process denoted by $\{S_t\}_{t\in[0,T]}$. For $t = T$, the final value of the product (i.e. S_T) is a random variable whose risk-neutral density is the main data which

[29] A group of academics, consumers associations, unions and other representatives of investors' interests who want to express a common view about the issues related to disclosure of the risk financial products placed with retail customers or with public institutions such as municipalities or other local authorities. For more details visit the website: movement for risk transparency .

[30] In the risk-based approach, a reference financial asset is a security for which the distribution of possible values at maturity is given by the variance of the risk rate at which its payoffs are invested up to maturity.

needs to be analysed in order to construct performance scenarios. In addition, partial preliminary knowledge of the functional form assumed by the probability density function at maturity (f_{S_T}) is required.

Most of the stochastic models used to describe the above processes could be chosen by the manufacturer of a financial product for the purposes of calculating a risk measure and determining potential future performance. Indeed, many of them are defined in continuous time and then discretised, where necessary, to perform simulations. In general, the preference for continuous-time models stems from their greater flexibility, since they allow the temporal dynamics of the variables of interest to be described, capturing how they affect the value of a financial product. Such models are useful for representing quite complex securities, whose payoffs depend on specific quantitative algorithms and which are exposed to a multiplicity of risk factors.

Some non-equity securities are easy to model, as their performance is usually directly related to the performance of a portfolio of underlying assets. In fact, they can be accurately represented by means of slightly revised common stochastic differential equations, such as that for geometric Brownian motion, which reflect features connected to the stochastic term structure of the benchmark's volatility. On the other hand, other non-equity products are contingent claims, namely securities whose pay-off structure is determined over a specific time period and is linked to underlying assets or reference values, according to specific formulae and subject to the fulfilment of precise conditions. This implies that stochastic models used to describe the temporal variation in their performance must consider all relevant risk factors and the specific way in which, due to financial engineering choices, they may affect the future cash-flows of the investment until its maturity date.

Parameters and variables associated with different risk factors have to be properly calibrated by the manufacturer on the basis of current market data as well as consistently with the features of any single financial product and the actual state of the reference market. Moreover, since most non-equity products are intended to be held for more than a year, variables such as interest rates, credit spreads, volatility and correlations cannot be assumed to be constant and so the models used to perform simulations must include a suitable set of stochastic differential equations to tackle this element of complexity.

The criterion to identify the scenarios aims at disclosing (i) the performance risks of a non-equity product and (ii) potential results obtainable by investing in an alternative security. Hence, an alternative investment should be determined in order to minimise any arbitrary assumptions about investors' preferences and, at the same time, to represent in a clear, immediate and significant manner how the specific risk factors and financial structure of the non-equity asset will affect the payoffs that could be obtained. In the Probabilistic scenarios table the risk-neutral density of the product is compared with the density associated with investment of the same amount of money in an RFA. The latter should be an investment which has same time horizon of the non-equity product and pays a return following the risk-free rate structure.

Moreover, to conduct the partition of the probability distribution of the returns, the percentiles are fixed and identified by exploiting the information contained in the risk-neutral probability distribution and in the RFA (described by the process $\{B_t\}_{t\in[0,T]}$). This allows identification of three main scenarios: negative return (loss of the invested capital), neutral return (in line with the return on the RFA) and positive return (exceeds return on the RFA). Each category is linked to a probability of occurrence and an absolute value (i.e. the conditional expected return of each partition), which together capture the returns achievable in that scenario. This method is also known as the 'superimposition technique'. From a technical point of view, since the probability density of RFA returns and the non-equity product are calculated under the risk-neutral measure $\mathbb{Q}$, it is possible to use this measure to compare their final value and to define events to be quantified.

The starting point of the simulations of non-equity product and the RFA are supposed to be determined according to specific rules. In the case of the non-equity product, it should be equal to the difference between

the notional capital and the up-front costs, if any. The simulations should also consider the size and the time schedule of any deferred costs that might apply over the recommended investment period[31], as well as any periodic or one-off amounts paid out to the investor or invested in other financial assets[32]. Because investment in the RFA has usually a minimal or even negligible mark-up, it is assumed that the initial size of the alternative investment in this asset is exactly equal to the notional capital. In this way, the probabilities attached to the scenarios implicitly consider the higher costs that non-equity investment products usually charge.

The two simulation processes described above provide the probability distributions of the final values of the capital effectively invested in the securities and of the notional capital ideally invested in the RFA. Both simulations must be performed according to models developed internally, by intermediaries to serve pricing and risk management activities. The numerical simulations of the final value of the invested capital must be compliant with the risk neutrality principle to ensure methodological consistency.

The risk-neutral probability density at time T by definition associates a predetermined quantile of probability with a possible final value of the asset. Thus it is possible to connect a value of α with the event 'the final value of the RFA is lower than α', in terms of probability quantiles on the final distributions. This is formally stated as follows $\mathbb{Q}[CB_T \leq \alpha] = P$ where T denotes the recommended investment period and C the amount invested in the security. The choice of a specific quantile P on the density of the RFA implicitly assumes the exclusion of a fixed percentage of the trajectories of the process of the RFA that are not considered representative of the potential behaviour of the process itself at time T. Hence, the cutting procedure can be considered a sort of correction aimed at excluding extreme events from the risk-neutral distribution of the RFA and, to this end, the choice of the threshold α is related to a very high quantile. In the implementation of the risk-based approach, the value of α is set at the 97.5th percentile. In formal terms this means that: $\mathbb{Q}[CB_T \leq \alpha] = 97.5$. It is interesting to observe that, when this identification criterion is used, the reference thresholds are automatically anchored to variations in the position and dispersion of the density of the risk-free asset distribution, and consequently these thresholds objectively reflect changes in the volatility of interest rates and overall market conditions. Considering the obtained α value and picking it from the distribution at maturity of the non-equity product, the neutral and positive performance scenarios could be set as all the trajectories below α but up to the amount C and above the threshold α, respectively.

Because information about possible poor performances is crucial to making an investment decision, the notional capital is used as the threshold for identifying losses, as its value represents the liquidity which the investor actually renounces during the recommended investment period. Finally, the mean values at the end of the recommended investment period are calculated using the formulae reported in Table 2.1: *The superimposition technique*.

[31] If there are deferred fees for some non-financial service, the starting point of the simulation must be equal to the difference between the price paid for the product and the total costs.

[32] Similarly, in products with periodic payments, the simulations should reflect the discontinuities in the value of the invested capital at the various payment dates.

Table 2.1: The superimposition technique.

EVENTS	PROBABILITY	MINIMUM VALUE IN €
Negative performance	$\mathbb{Q}(S_T < C)$	$E^{\mathbb{Q}}(S_T \mid S_T < C) = \frac{1}{\mathbb{Q}(S_T < C)} \int_{-\infty}^{C} x f_{S_T(x)dx}$
Neutral performance	$\mathbb{Q}(C \leq S_T < \alpha)$	$E^{\mathbb{Q}}(S_T \mid C \leq S_T < \alpha) = \frac{1}{\mathbb{Q}(C \leq S_T < \alpha)} \int_{C}^{\alpha} x f_{S_T(x)dx}$
Positive performance	$\mathbb{Q}(S_T \geq \alpha)$	$E^{\mathbb{Q}}(S_T \mid S_T \geq \alpha) = \frac{1}{\mathbb{Q}(S_T \geq \alpha)} \int_{\alpha}^{\infty} x f_{S_T(x)dx}$

2.6 Application of the analysed techniques

In the following sections, using evidence from two case studies, the approaches described above are analysed to compare their methodology and technical properties. Based on the recent Italian cases of miss-selling, two subordinated bonds , which characteristics are obtained from the relative financial prospectuses and resumed in Table 2.2: *Summary characteristics of the analysed products,* drives such an analysis. In 2018, about 1% of the Italian households' financial wealth were invested in this type of non-equity products, representing roughly 15% of the bonds issued by banks hold by households and even 37% of the overall Italian subordinated bonds (Bank of Italy, 2018).

Table 2.2: Summary characteristics of the analysed products.

	Product I	Product II
Rating issuer	BB-	BBB+
Debt seniority	lower Tier II	lower Tier II
Costs	No	No
Issue date	January 2012	January 2012
Maturity	7 years	7 years
Coupons payment	yearly	yearly
Gross interest rate	5%	4,5%
Net interest rate	4%	3,6%
Issue price	100% of nominal value	100% of nominal value
Redemption price	• at par • from July 2015 yearly cash flow of 20% of nominal value	• at par • unique solution

The selected securities were issued by two Italian banks in January 2012; they are two seven-year subordinated bonds (hereafter Product I and Product II) with fixed rate coupons of 5% and 4.5% per annum respectively. Because they are subordinated bonds (lower Tier II securities) investors were not guaranteed to get back all the capital they had invested: if the issuer goes into liquidation before the maturity of the product, investors could lose all the money they had invested. In these circumstances subordinated bonds

would only be redeemed after all other preferential creditors, unsecured or with a lesser degree of subordination, had been satisfied, and the liquidity of the issuer might not be sufficient to make even partial repayment of subordinated bonds. It should be noted that the two issuers had different ratings, implying a difference in the assumed default risk. Other than the solvency level, the main difference between Product I and Product II is the modality of repayment of the invested capital: from 2015 to maturity Product I provided a yearly refund of 20% of the invested capital whereas Product II was to be repaid in full at maturity via a one-off payment.

2.6.1 What-if analysis

Based on the described indications, the What-if analysis requires the specification of a set of possible scenarios (namely negative, neutral and positive, indicated respectively with $I = N,E,P$), which in turn depend on a set of variables. Indexing Product I and Product II respectively as $j = I$ and $j = II$, the main drivers to consider were the default risk and related factors (time to default, $t_{D,i,j}$; and recovery rate, $RR_{i,j}$); the expected returns were computed using the IRR formula. The assumptions used to obtain the illustrative examples for both products were the following:

- Negative scenario: issuer defaults before maturity ($t_{D,N,I} = t_{D,N,II} = 5$) and only part of the invested amount is paid back ($RR_{N,I} = RR_{N,II} = 0.5$).
- Neutral scenario: coupons are paid regularly but the issuer defaults at maturity ($t_{D,E,I} = t_{D,E,II} = 7$ and $RR_{E,I} = RR_{E,II} = 0.8$).
- Positive scenario: all coupon and capital payments are made on schedule (no default occurs).

The results obtained from the What-if analysis are summarised below in Figure 2.1: *What-if tables*.

Figure 2.1: What-if tables.

Product I

ILLUSTRATIVE EXAMPLES*	EXPECTED RETURN PER YEAR**
Negative scenario - Issuer defaults before the maturity and only part of the invested amount is paid back.	-3,02%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity.	4,15%
Positive scenario – All coupons and invested capital are paid regularly.	5%

* They are illustrative examples and do not represent a forecast. The shown scenarios may not have an equal probability of occurrence

** If held to maturity

Product II

ILLUSTRATIVE EXAMPLES*	EXPECTED RETURN PER YEAR**
Negative scenario - Issuer defaults before the maturity and only part of the invested amount is paid back.	-8,39%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity.	1,12%
Positive scenario – All coupons and invested capital are paid regularly.	4,5%

* They are illustrative examples and do not represent a forecast. The shown scenarios may not have an equal probability of occurrence

** If held to maturity

2.6.2 Building Probabilistic scenarios

Considering the risk-neutral world, the Monte Carlo technique was used to simulate the scenarios in the Probabilistic tables. A sufficient number of possible combinations of the risk drivers was generated and the values used to calculate the prices of the two products at maturity. The end of the recommended investment period is the maturity date of the bonds and so the investment period was set at seven years[33]. This simulation method requires identification of the exogenous variables and the parameters, definition of reference models (mathematical relationships between the input variables and the parameters), attribution of the distributions of probability and the setting of the simulations. In this specific case, calculation of the *i*-th simulated paths (with *i = 1, …, N* and *N* = 100,000) required consideration of two risk drivers: the default risk of the issuers and interest rate risk[34], which were modelled separately.

The default risk was modelled in discrete time using the binomial tree representation proposed by Jarrow and Turnbull (1995), with the assumption that default may only occur instantaneously before the dates of coupon payments. Under this approach, the binomial tree is described by a Bernoulli process calibrated using the issuers' probability of default for the relevant time period. Using the CDS spreads[35] for various maturities, the risk-neutral probability of default for the connected time horizons was estimated. Remember that Product I and Product II are indexed respectively with *j* = I and *j* = II. Let $p_{j,t}$ indicate the cumulative probability of default for product *j* for a period of *t* years, i.e. the probability that the issuer will default between the observation date and the end of the *t*-th year (with *t = 1, …, T* where in our cases *T* = 7). Assuming risk neutrality, the cumulative probability of default is given by:

$$p_{j,t} = \frac{1 - e^{-CDS_{j,t} * t}}{1 - RR}$$

Let the probability that the debtor does not default between the issue date (*t* = 0) and the end of year *t* be $s_{j,t} \equiv 1 - p_{j,t}$ and let $s'_{j,t}$ be the marginal survival probability during year *t*, i.e. the probability that the debtor will not default during year *t* conditional on there being no default before the end of year *t-1*. Accordingly, for any *t*, the following is given:

$$s_{j,t} = s_{j,t-1} s'_{j,t}$$

In other words, the probability of survival from 0 to *t* is given by the product of the probability of survival from 0 to *t-1* and the marginal probability of survival for year *t*. It follows, then, that the marginal probability of survival can be expressed as follows:

$$s'_{j,t} = \frac{s_{j,t}}{s_{j,t-1}}$$

The marginal probability of default during year *t* ($p_{j,t}$) will then be one minus the related marginal probabilities of survival:

$$p'_{j,t} = 1 - \frac{1 - p_{j,t}}{1 - p_{j,t-1}}$$

Thus, using the CDS market quotations the default probability curves were bootstrapped. At this stage, for each issuer, a Bernoulli distribution was considered to generate *N*-th default paths:

[33] As there are no events that would trigger early redemption. The investor could only sell the products in a secondary market.

[34] It is necessary to compound any bond coupon payment, at the risk-free rate, forward to the end of the recommended investment time horizon.

[35] For each issuer the linked Junior CDS spread curve in January 2012 is considered. Given that Product I issuer is a non-listed company, the calibration of the parameters (see also models in Section 2.7 and Section 2.8) is done on using market data for similar listed companies.

$$I_{ijt} = \begin{cases} 0, & \text{with probability } p'_{j,t} \\ 1, & \text{with probability } 1 - p'_{j,t} \end{cases}$$

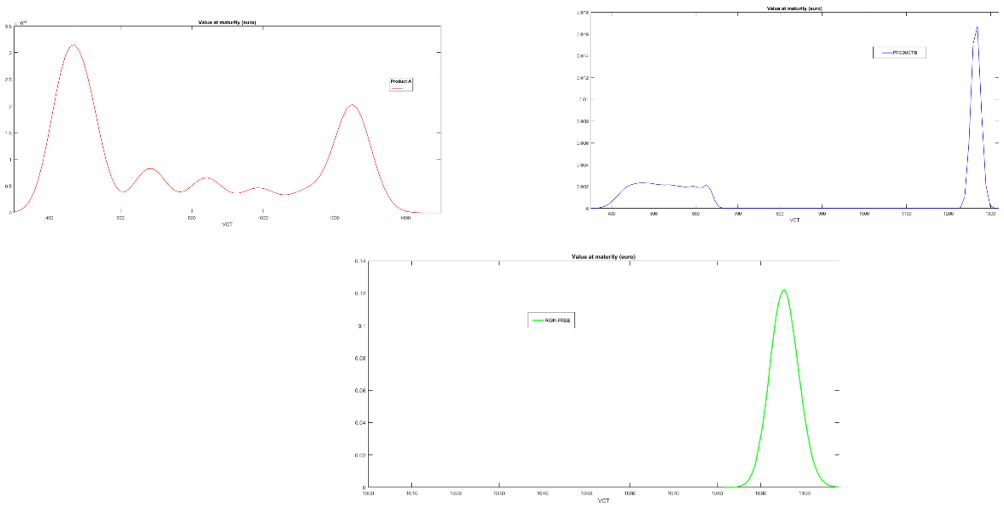
Moreover, it should be noted that the simulation of the *i*-th paths stopped whenever default occurred (i.e. I_{ijt} = 0). After combining the simulated default events and the linked fixed cash-flows ($FC_{i,j,t}$[36]), the distributions of price at maturity represent the capitalisation of the simulated paths (capital at maturity - $CM_{i,j,T}$).

In order to represent the interest rate risk, *N*-th simulations of the OIS term structure (Hull and White, 2012) were performed according to the Hull and White one-factor model[37] (Hull and White, 1990). In this model, the calibration of the parameters α and σ was done explicitly, by minimising the square distance between the option prices quoted in the financial market[38] and the theoretical prices given by the Hull-White trinomial tree approach (Hull and White 1994, 1996).

A 7-year German Bund issued in 2012 (1.48% gross yearly coupon) was selected as RFA and its price at maturity computed as above, but without taking into consideration the default risk. It should be noted that the already-simulated interest rate term structures were also used to capitalise at maturity the fixed cash-flows generated by the Bund.

The distributions of price at maturity for Product I (red line), Product II (blue line) and the RFA (green line) are shown in Figure 2.2: *Price at maturity of the analysed products*.

Figure 2.2: Price at maturity of the analysed products.



The bimodal shape of the probability densities of Product I and Product II reflects the seven-year default risk for the issuers, under the hypothesis of both recovery rates equal to 0.4, due to the seniority of the debts. In particular, the left side of the distributions incorporates all the trajectories in which the default events occur and, therefore, this part of the distributions falls in the region corresponding to negative returns; whereas the other mode represents the trajectories not affected by a credit event and consequently lies in the area of positive returns. This bimodality is more evident in the case of Product II due to the difference in the capital repayment structures of the products.

Applying the superimposition technique, the probability of occurrence of negative/neutral/positive events for Product I and Product II were calculated. In order to compute the medium value at maturity in

[36] The coupons are computed using net interest rates in order to consider the reduction in the coupons' value due to taxes.

[37] $dr = [\theta(t) - \alpha r]dt + \sigma dz$

[38] Firstly, the strike price is computed from the quoted ATM caps volatility for different maturities. Then, using market volatilities and the ATM strike prices, the prices of the different options are obtained from the Black formula.

euros, for each scenario, the mean value of the simulated partition was measured following the criteria described above in Table 2.1: *The superimposition technique*. The final Probabilistic scenarios tables for Product I and Product II are shown in the Figure below.

Figure 2.3: Probabilistic scenarios tables.

Product I

EVENTS	PROBABILITY	MEDIUM VALUE IN €*
NEGATIVE performance	66,38%	576
NEUTRAL performance	3,76%	1045
POSITIVE performance	29,86%	1249

* at maturity for an initial investment of 1000 euros

Product II

EVENTS	PROBABILITY	MEDIUM VALUE IN €*
NEGATIVE performance	50,21%	529
NEUTRAL performance	-	-
POSITIVE performance	49,79%	1268

* at maturity for an initial investment of 1000 euros

2.7 Model risk

Given that all models are a simplified representation of reality, there is a risk that something will not be accounted for (i.e. model risk). The assumptions made to develop a model and the corresponding input parameters may vary widely. Risk managers and traders determine the possible returns for a particular product from market or historical data, after specifying the model, the calibration procedure and a sufficient number of assumptions. This process and, thus, model risk also affect manufacturers when applying the What-if and Probabilistic scenarios methods.

2.7.1 Changing the assumptions underlying the illustrative examples

The procedure used to produce the What-if tables can be used to generate different outputs if the hypotheses underpinning the illustrative examples are changed without altering the formula used to determine the potential returns, which is strongly linked to the characteristics of the financial products and, therefore, not manipulable. For these reasons, in the case of Product I and Product II, model risk is determined by the several sets of assumptions relating to the risk parameters used in the IRR formula for computing the expected annual return and not by the IRR formula itself. For example, variations in expected returns can be obtained by changing the time at which the issuer would default (i.e. modifying $t_{D,i,j}$) or the

relative recovery rate (i.e. $RR_{i,j}$), which may be different for the two products[39]. Thus, stressing the hypotheses underlying the illustrative examples, it is possible to get a wide range of output tables for each product (for some examples see Figure 2.4: *Stressing the assumptions in the What-if tables*).

Figure 2.4: Stressing the assumptions in the What-if tables.

Product I

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - All the coupons are paid regularly but the issuer defaults at maturity, paying back a small part of the invested capital.	1,4%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity, paying back most of the invested capital.	4,15%
Positive scenario – All coupons and capital payments are made on schedule.	5%

Product II

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - All the coupons are paid regularly but the issuer defaults at maturity, paying back a small part of the invested capital.	-8,9%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity, paying back most of the invested capital.	1,58%
Positive scenario – All coupons and capital payments are made on schedule.	4,5%

Product I

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - All the coupons are paid regularly but the issuer defaults at maturity, paying back a small portion of the invested capital.	-6,95%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity, paying back part of the invested capital.	1,4%
Positive scenario – All coupons and capital payments are made on schedule.	5%

Product II

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - All the coupons are paid regularly but the issuer defaults at maturity, paying back a small portion of the invested capital.	-11,84%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity, paying back most of the invested capital.	1,12%
Positive scenario – All coupons and capital payments are made on schedule.	4,5%

Product I

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - All the coupons are paid regularly but the issuer defaults at maturity, paying back a small portion of the invested capital.	-4,89%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity, paying back a small portion of the invested capital.	2,65%
Positive scenario – All coupons and capital payments are made on schedule.	5%

Product II

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - All the coupons are paid regularly but the issuer defaults at maturity, paying back a small portion of the invested capital.	-11,84%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity, paying back a small portion of the invested capital.	-6,67%
Positive scenario – All coupons and capital payments are made on schedule.	4,5%

First of all it should be noted that changing the assumptions linked to the different risk drivers would also affect the wording of the illustrative examples, thus reducing the comparability of the products. Moreover, focusing on the negative scenarios (the positive ones are unchanged since they always represent the full repayment of the debt), the risk of losses is not homogeneous among representations of the same product. Although the investor buying either Product I or Product II might not get back all the invested capital, the first table for Product I in Figure 2.4: *Stressing the assumptions in the What-if tables* does not indicate possible losses, whereas the investor in Product II would incur losses, despite Product II has a lower risk of default. One can see that this representation may affect investors' perception of the relative riskiness of these two securities. Finally, the expected return values reported in the tables do not indicate the underlying risks which are strongly related to the differing structures of the products: the amount invested in Product I is refunded following an amortisation programme whereas the amount invested in Product II is paid back via a one-off payment at maturity.

[39] Remember that the illustrative examples (namely negative, neutral and positive) are indicated with i = N, E, P; while j = I, II index Product I and Product II respectively.

2.7.2 The pricing models and the variations in probabilities of occurrence

Let us now consider how the Probabilistic scenarios tables are compiled. For any given product a suitable family of models for dealing with the associated pricing problem and the underlying risks must be identified, but the various models in this family may produce dissimilar risk-neutral densities for returns at maturity. Choosing a model and corresponding calibration data input and process is a challenging task, because both the model and its calibration should be selected to give a distribution of outcomes that coincides with the expected distribution of future outcomes. Although there is no market consensus on the type of model to be used or the calibration procedure, intermediaries consider and measure the parameters and the variables corresponding to all the risk factors relevant to the product offered, taking care that they are consistent with the reality and complexity of financial markets, and also using, where necessary, suitable stochastic processes to model the various variables involved.

In order to stress the model risk for the Probabilistic scenarios, the distributions of price at maturity for the subordinated bonds analysed here were obtained using different models for the relative risk drivers: the interest rate was simulated with the stochastic term structure model developed by Cox *et al.* (1985) (the so-called CIR[40]) whereas the risk-neutral probabilities of default were calculated with the Merton structural credit model[41] (Merton, 1974). Moreover, a combination of the models used above [42] is performed in order to stress the effects of model choice on the distributions at maturity.

The risk-neutral density functions of Product I and Product II are not influenced by the methods used for interest rate term structure, because both the chosen models belong to the class of one-factor affine short-rate models and, therefore, base the term structure of interest rates on the same assumptions. In particular, using two different stochastic processes belonging to the class of one-factor affine short-rate models (i.e. HW and CIR), the variation in the risk-neutral densities of the subordinated bonds is not significant because shape of the interest rate curve is not the main risk driver for the analysed instruments, which have fixed coupons rates. In contrast, greater variation in probabilities - although it is not significant since most of the trajectories associated with negative and positive events do not vary - is obtained changing the approach used to model the default risk, which is the main risk driver for a subordinated bond with fixed coupons.

Analysing Figure 2.5: *Effect on probabilities changing pricing models*, the variation in the probabilities of the scenarios due to the choice of pricing model is attenuated, because the distributions are partitioned into only three segments, and thus, many elementary events are aggregated and not overestimated. Hence, applying the superimposition technique, the information embedded in the risk-neutral density functions is exploited efficiently and the differences between the densities generated by different models are smoothed and should have negligible influence on the decision-making process of retail investors.

[40] This process describes the evolution of the interest rate in the risk-neutral world.

[41] The Merton model, like any model, simplifies the reality to make things tractable. Important assumptions are: no transactions cost, no bankruptcy cost, no taxes, unrestricted borrowing and lending at the risk-free interest rate, no short selling restrictions, no uncertainty about liabilities, log-normally distributed assets.

[42] The Hull and White One-factor model (HW) and the approach based on corporate CDS market quotes (CDS) already used in Section 2.6.2, respectively for interest rate risk and credit risk.

Figure 2.5: Effect on probabilities changing pricing models.

Product I: CDS-CIR

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	69,14%	549
NEUTRAL performance	4,85%	1073
POSITIVE performance	26,01%	1201

Product II: CDS-CIR

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	49,6%	495
NEUTRAL performance	-	-
POSITIVE performance	50,4%	1252

Product I: MERTON-HW

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	64,23%	644
NEUTRAL performance	5,85%	1043
POSITIVE performance	29,92%	1243

Product II: MERTON-HW

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	41,33%	557
NEUTRAL performance	-	-
POSITIVE performance	58,68%	1268

Product I: MERTON-CIR

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	68,67%	621
NEUTRAL performance	6,87%	1073
POSITIVE performance	24,47%	1201

Product II: MERTON-CIR

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	41,41%	524
NEUTRAL performance	-	-
POSITIVE performance	58,59%	1252

2.8 Update the tables

Often people find it difficult to understand the financial implications of a purchase. In the case of simple purchases, such as buying a coffee, the financial transaction is trivial, but it becomes tricky when a decision requires analysis of several factors, as in the case of financial products. For example, there are aspects/events that may occur during the investment period which for a retail investor are not easy to consider before investing, for instance annual fees, changes to market conditions, variations in interest rates and taxes. In the case of financial investments and other contracts that have a complex pricing structure (for example mortgages) that is neither transparent nor comprehensible to consumers, the approach termed RECAP (Record, Evaluate, and Compare Alternative Prices) proposed by Thaler and Sunstein (2008) is appropriate. To produce a RECAP for a product the seller must collect all the relevant information and deliver it to the consumer in the form of a periodical report; this would be extremely useful to individuals to be informed about past and updated expected future performance of a financial product.

Investors should be guaranteed delivery of dynamic information so that they are kept informed about market conditions, the indicators used should capture variation in risk and update investors about expected returns. The data used for computing the indicators, such as the risk or costs indicators, may change over time, leading to variations in the information to be included in financial prospectuses. Therefore, in order to represent the performance risk associated with a financial product, the tables should also be used to detect

cases where changes in the risk factors affecting a given investment over the recommended investment period require updates to the information previously provided to investors. For these reasons, manufacturers should periodically review the information contained in financial prospectuses, starting with an assessment of whether changes in the data necessitate revision and republication of the documentation. Where a periodic review of a prospectus identifies changes to the information set that is required to be included in the document or concludes that information contained in it is no longer accurate, fair, clear and not misleading, the manufacturer should revise the key information document to take the changes into account. Moreover, retail investors should also be informed if conditions remain unchanged, so that they have a better understanding about the characteristics of the security.

Figure 2.6: Updating the What-if tables.

Product I – 2013

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - Issuer defaults before the maturity and only part of the invested amount is paid back.	- 4,89%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity.	3,42%
Positive scenario – All coupon and capital payments are made on schedule.	5%

Product II – 2013

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - Issuer defaults before the maturity and only a small part of the invested amount is paid back.	- 8,90%
Neutral scenario – Issuer defaults before the maturity and most of the invested amount is paid back.	0,46%
Positive scenario – All coupon and capital payments are made on schedule.	4,5%

Product I – 2014

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - Issuer defaults in the short term and only a small part of the invested amount is paid back.	- 11,43%
Neutral scenario – Issuer defaults in the long term and the invested amount is paid back almost fully.	1,40%
Positive scenario – All coupon and capital payments are made on schedule.	5%

Product II – 2014

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - All the coupons are paid regularly but the issuer defaults at maturity, paying back a small part of the invested capital.	- 6,75%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity, paying back most of the invested capital.	1,12%
Positive scenario – All coupon and capital payments are made on schedule.	4,5%

Product I – 2015

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - Issuer defaults in the short term and only a small part of the invested amount is paid back.	- 15,13%
Neutral scenario – Issuer defaults in the medium term and the invested amount is paid back almost fully.	0,25%
Positive scenario – All coupon and capital payments are made on schedule.	5%

Product II – 2015

ILLUSTRATIVE EXAMPLES*	EXPECTED RETURN PER YEAR
Negative scenario - All the coupons are paid regularly but the issuer defaults at maturity, paying back a small part of the invested capital.	- 6,75%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity, paying back most of the invested capital.	1,12%
Positive scenario – All coupon and capital payments are made on schedule.	4,5%

Product I – 2016

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - Issuer defaults in the short term and only a small part of the invested amount is paid back.	- 4,89%
Neutral scenario – Issuer defaults in the medium term and the invested amount is paid back almost fully.	1,4%
Positive scenario – All coupon and capital payments are made on schedule.	5%

Product II – 2016

ILLUSTRATIVE EXAMPLES	EXPECTED RETURN PER YEAR
Negative scenario - All the coupons are paid regularly but the issuer defaults at maturity, paying back a small part of the invested capital.	- 6,75%
Neutral scenario – All the coupons are paid regularly but the issuer defaults at maturity, paying back most of the invested capital.	1,12%
Positive scenario – All coupon and capital payments are made on schedule.	4,5%

If the What-if approach is being used, the tables could be updated simply by changing the hypotheses underlying the three illustrative examples. For Product I and Product II, in particular, the time of default and the linked recovery rate should reflect the status of the issuers and market conditions. The tables covering 2013 to 2016 are shown above in Figure 2.6: *Updating the What-if tables*.

Figure 2.7: Updating the Probabilistic scenarios tables.

Product I – 2013

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	44,46%	688
NEUTRAL performance	3,87%	1047
POSITIVE performance	51,67%	1231

Product II – 2013

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	38,77%	556
NEUTRAL performance	-	-
POSITIVE performance	61,23%	1262

Product I – 2014

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	41,2%	685
NEUTRAL performance	3,56%	1063
POSITIVE performance	55,24%	1220

Product II – 2014

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	16,9%	555
NEUTRAL performance	-	-
POSITIVE performance	83,1%	1258

Product I – 2015

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	45,49%	793
NEUTRAL performance	8,24%	1058
POSITIVE performance	46,27%	1205

Product II – 2015

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	13,1%	568
NEUTRAL performance	-	-
POSITIVE performance	86,9%	1256

Product I – 2016

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	89,62%	830
NEUTRAL performance	1,01%	1008
POSITIVE performance	9,37%	1207

Product II – 2016

EVENTS	PROBABILITY	MEDIUM VALUE IN €
NEGATIVE performance	16,61%	582
NEUTRAL performance	-	-
POSITIVE performance	83,39%	1254

The three pillars of the risk-based method consider the implied probability distribution and its valuable information content. In particular, Probabilistic scenarios are strictly linked to market conditions because of its direct connection with the quoted prices of liquid assets. For these reasons, the information included in the Probabilistic tables could be refreshed simply, by using input data reflecting the new market conditions to re-calibrate the underlying pricing models. Examples of the Probabilistic tables for Product I and Product

II that could be produced and annually delivered to update investors from 2013 to 2016 are reported above in Figure 2.7: *Updating the Probabilistic scenarios tables*.

Analysis of the output tables obtained leads to some observations about the efficacy of updating the information in the tables and how effectively the approaches capture variations in market conditions. Firstly, as shown in both What-if and Probabilistic tables, over the years, the expected performance of Product I is decreasing but it is characterised by an increasing risk, whereas the expected performance of Product II shows the opposite pattern. Although the temporal variation in risk is captured by the What-if approach, it does not show how high the risk of losing the invested capital is, whereas this information is easily gleaned from the display of probabilities provided by the Probabilistic scenarios. Indeed, the simulated trajectories in the two modes of the risk-neutral distributions vary consistently, reflecting the variation in the issuers' default risk. It is worth noting that the variation in risk is also made more salient for investors by the choice of colour scheme. The negative scenario is shown in red, which is a sensitive colour for individuals who, in general, link it to negative and dangerous situations (in this case loss of invested capital). The What-if representations do not permit the use of visual distractors such as colours[43], so this visual salience is lost. Secondly, it should be noted that the repayment of the invested capital structures, which strongly differentiates the two products, is reflected in the medium values displayed in the Probabilistic scenarios[44] whereas it is not directly available from the presentation of expected returns in the What-if tables. Moreover, the choice of underlying hypotheses for the two approaches affects the output tables (model risk came out here too). Although the principle underlying the information updating process is common to the What-if and the Probabilistic scenarios (i.e. change the input data in order to reflect the market conditions), the former approach uses manufacturers expectation of future market condition, whereas the latter considers real market data. Finally, changing the market conditions implies a variation in the wording used to describe the illustrative examples in the What-if tables, which makes it less easy to compare the two products and less easy to compare the new estimates of expected performance with previous tables. So, in order to standardise both the criteria underlying the two approaches and the framing of information, it seems more appropriate to use Probabilistic tables, the format of which is not affected by the change of market conditions.

2.9 Comparisons between methodologies

Starting from the analyses provided in the previous sections and given that some criticism of the two methods emerged during the debate on this issue, the assessment of the two approaches might be useful to understand positive and negative aspects under the methodological point of view[45].

Manufacturers that use the What-if approach select events and construct hypotheses on a somewhat arbitrary basis. The presentation table shows only three single state of the world in the future out of an infinity of possible states (the wording used is 'illustrative examples' which implies the lack of a complete representation or not based on real empirical data), and as such has a zero likelihood of occurrence. The manufacturers could cherry-pick, choosing the scenarios that are likely to steer investors towards buying the product. Moreover, the same product could be presented in diverse ways under the What-if format, simply by changing the hypotheses and, thus, the wording included in the output table. This lack of standardisation makes comparison of financial products difficult, even impossible across asset classes. In contrast, with the

[43] It should be highlighted that the use of colours is a key distinguishing characteristic of the two methodologies and, thus, the linked conclusions are valid for their overall assessment and not just for the results in this Section.

[44] The medium value linked to the negative scenarios is progressively increasing for Product I, because over the years part of the invested capital is reimbursed, reducing possible losses in the event of default; it remains constant for Product II, which pays back the entire capital at maturity and, thus, the amount exposed to default risk does not vary.

[45] Their effect on the decision-making process will be analysed in Chapter 3.

Probabilistic scenarios not only are the events standardised (as a consequence of the superimposition technique), a complete representation of possible future performance is presented to investors (all possible scenarios are collected and measured). It may, therefore, be easier to compare products when information about future performance is presented in terms of probability. However, it should be noted that the Probabilistic scenarios methodology does not prescribe the use of a specific pricing model for the quantitative determinations needed to calculate the values in the table. Therefore, arbitrariness in the probabilistic approach is not completely eliminated since the choice of pricing model is still left to intermediaries.

Turning to the computational processes used to obtain the tables, both the approaches can be implemented quickly and to reflect all risks and market conditions. However, Probabilistic scenarios are based on real and forward-looking market data shared among agents, whereas What-if tables simply use assumptions based on market scenarios to formulate the hypotheses underlying the illustrative examples.

As stated by regulators and some academics without providing concrete evidence, a disadvantage of using Probabilistic scenarios refers to the increased error linked to the models used in the calculation of the probability of the performance scenarios (i.e. model risk). Manufacturers model future performance of a security for pricing and risk management activities. Indeed, asset managers and manufacturers are interested in the same basic question as retail investors: 'How much better am I likely to do with this product than with other products?', but unlike retail investors they have the technical tools and skills to provide a reasonable answer. Given the availability of such knowledge, providing consumers with this key information does not impose any additional burden on top of their usual pricing and risk management activities[46]. In their business market practitioners make use of pricing and hedging models defined under the risk-neutral measure, which is the only one consistent with the no-arbitrage principle providing the connection between the fair value of any contingent claim with a time horizon *T* and the risk-neutral probability density function of the possible final values of the contingent claim at time *T*. Only under this measure can any arbitrary assumption about the future evolution of market variables be discarded and an effective comparability across different securities is meaningful. The use of a risk-neutral measure is a basic methodological requirement if the information conveyed is to be objective, meaningful and consistent both intrinsically and across the various indicators[47]. However, when financial markets are incomplete, risk-neutral measure does not provide a real representation of the future performance of a financial product but it allows to get a comparable representation of a security by computing the probability of occurrence of each scenario using a unique risk measure. In addition, each indicator[48] built for disclosing information requires the use of specific models under stringent hypotheses. Even though the calculation of the probabilities of the performance scenarios implies the use of assumptions and models, the model risk is mitigated (see Section 2.7: *Model risk* and Minenna, 2011[49]) because of the reduction in granularity (given by a limited number of scenarios) which ignores slight shifts in the probability distributions of the final values of the invested capital. The above reasoning does not apply to the What-if representation, since pricing models are not used and the procedure for computing expected returns is linked to the formula describing the cash-flows mechanism of the product. In addition, as shown in Section 2.7: *Model risk*, the What-if representation implies a high model risk due to

[46] This approach also avoids the costly and useless implementation of 'parallel models', one for the internal activities of the intermediaries and the other for compliance with disclosure obligations.

[47] It is worth recalling that, by the Fundamental Theorem of Asset Pricing, the risk-neutral probability measure is - under the complete markets hypothesis - the only one under which the stochastic process of the discounted final payoff of the product is a martingale; and thus, only this measure allows a meaningful comparison of different quantities. More specifically, the risk neutrality ensures that the probabilistic comparison with the reference financial asset is valid. In fact, it is only under the risk-neutral measure that the expected returns of the product and the reference financial asset do coincide and, therefore, the probability table can highlight the role played by the volatility of the elementary components of the financial investment, by the costs attached to it and by the specific structure of the amounts paid out to the investor over the recommended investment period.

[48] For example, see the synthetic risk index introduced in the KID.

[49] The author shows the slight variation in the risk-neutral density for a subordinated floating-coupon bond obtained from two different stochastic term structure models belonging to the class of one-factor affine short-rate models, namely Hull and White and CIR.

the discretion given to the manufacturers who have to choose the events that are represented in the output table.

Following the results in Section 2.8: *Update the tables*, methodological weakness and strengths become more relevant when the manufacturers should periodically calculate the tables reflecting the variation of market conditions in order to update investors about the value of the product. Indeed, the process and the models underlying the Probabilistic scenarios reflect market expectations of the risk-reward profile of the products, whereas the What-if tables would not be coherent with a forward-looking and complete representation of the future performance, which means they do not provide investors with as much information.

Finally, although the introduction of specific and technical rules steer a correct implementation and monitoring of the outcome, regulators and supervisors can easily monitor the processes underlying the What-if and Probability tables.

The weaknesses and strengths of the analysed approaches are summarised in Table 2.3: *Comparison between the approaches*, where red crosses are used to indicate a weakness and green ticks indicate a strength.

Table 2.3: Comparison between the approaches.

	WHAT-IF	RISK-BASED
Standard/arbitrary	x	x
Fully representation	x	✓
Easy to compute	✓	✓
All risks	✓	✓
Market conditions	✓	✓
Model risk	x	x
Updating	x	✓
Supervision	✓	✓

2.10 Considerations

Disclosure of the risks of structured securities should result in elimination of excessively risky products from the market and improve the quality of the offer as well as providing investors with greater protection. Without a comprehensive and understandable set of information, particularly in the case of non-equity products, retail investors' decisions are heavily influenced by their confidence in the financial institution supplying a particular investment product. This can mean that investors fail to comprehend important information about products and hence make uninformed financial decisions. Representations of the characteristics of financial securities are particularly useful when the products are characterised by structured financial engineering and by a complex pattern of potential returns. In those cases, in fact, the product embeds several risks that cannot be adequately perceived by retail investors through the usual metrics such as price, the internal rate of return and the volatility.

Practical solutions for representation of expected future performance are provided by the What-if and Probabilistic scenarios approaches. The former involves using illustrative examples that describe how the formula underlying the products works in specific circumstances, whereas the latter uses three thresholds connected with events of high financial relevance: loss of capital and performance in line with or above that of a reference financial asset and getting back the invested amount.

Some conclusions about the two methods can be drawn from the analyses presented in this chapter. Firstly, given the complex and fragmentary regulatory framework, it seems reasonable that all the authorities should work on a legislative revision of the EU provisions with the aim of producing a single and shared regulatory framework for disclosure of all non-equity securities, focusing on the set of information necessary to informed investment decisions.

Moreover, a deep analysis of What-if and Probabilistic scenarios method highlights some positive and negative aspects of the two approaches. In particular, robust evidence shows the quality of the information provided by the Probabilistic tables whereas there are some weaknesses in the procedure used to produce What-if scenarios. The What-if approach considers just three elementary outcomes out of an infinity of possible outcomes, and these are probably selected for the convenience of the issuer, who may wish to present the product optimistically or ambiguously, overemphasising unlikely positive performance or omitting negative results. Such a representation may bias investors' opinion of the prospects of the investment, since the three illustrative examples could be perceived as exhaustive of all possible outcomes. Furthermore, it may be difficult for retail investors to link a probability of occurrence to each event: not only are these likelihoods impossible to be identified for the events shown, but it is also not possible for investors to estimate the likelihood of generally defined negative, neutral and positive scenarios. All these flaws are avoided by using the probabilistic approach, which presents the entire probability distribution of a product's final performance and summarises it in a set of prescribed events - calculated with reference to a product-specific time frame that corresponds to the recommended holding period - which are important to any investor. It is also highlighted that the probabilistic approach leaves room for arbitrariness and model risk.

All these aspects deserve further analyses stressing the methodological features of these representation schemes applied to different types of securities (e.g. structured bonds, unit-link, index-link and derivatives).

Recognising the quality of the risk-based approach, the public authority responsible for regulating the Italian financial markets (CONSOB) had successfully implemented measures for mutual funds and class III and V financial-insurance products offered in Italy requiring the application of the three pillars of the risk-based approach to obtain the information for the prospectuses. However, in 2011, the Authority removed these requirements because they were not considered to be in line with those of the European regulators.

Although the positive and negative aspects of the two approaches have long been debated without any consensus being reached, their relative usefulness to potential retail investors and the impact of the way the information is framed have not yet been tested to determine which method would do most to reduce the impact of cognitive biases on financial decisions and increase investor awareness. After comparing Product I and Product II using stochastic dominance criteria, Chapter 3 examines how different framings of the analysed products affect subjective understanding and perception of the financial information provided and investment decisions via a consumer-test exercise and a physiological experiment. Consumer testing was used to explore: (i) how the different framings influence risk perception; (ii) how the different framings affect investment decisions; (iii) how the different representation formats are appraised in terms of complexity, usefulness and information content. The physiological laboratory experiment measured the impact of the two frames on investors' peripheral nervous system activity. The objective was twofold: (i) to investigate behavioural and physiological variables during the decision-making process of investors who were required to choose on the basis of the two different schemes of representing information; (ii) to compare the effects of time pressure on the choice when Probabilistic scenarios and What-if frames are used.

2.11 References Chapter 2

Bank of Italy (2018), Financial stability report, n.1., April 2018.

Black, F. & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy,* vol. 81(3), pp. 637-654.

CESR (2010a). Selection and presentation of performance scenarios in the Key Investor Information Document (KID) for structured UCITS. Guidelines, CESR/10-1318, 20/12/2010.

CESR (2010b). CESR's level 3 guidelines on the selection and presentation of performance scenarios in the Key Information Document (KID) for structured UCITS. Feedback Statement, CESR/10-1396, 20/12/2010.

Colaert, V. (2016). The Regulation of PRIIPs: Great Ambitions, Insurmountable Challenges?. *Journal of Financial Regulation*, vol. *2*(2), pp. 203-224.

Cox, J. C., Ingersoll Jr, J. E., & Ross, S. A. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica: Journal of the Econometric Society*, pp. 363-384.

Donkers, B., Lourenço, C., Goldstein, D. & Dellaert, B., (2013). Building a distribution builder. Design considerations for financial investment and pension decisions. *Design paper Network for Studies on Pensions*, Paper n. 20.

Epstein, L. G. (1999). A definition of uncertainty aversion. *The Review of Economic Studies*, vol.66(3), pp. 579-608.

ESAs JC (2015). Technical Discussion Paper on Risk, Performance Scenarios and Cost Disclosures for Key Information Documents (KIDs) for Packaged Retail and Insurance-based Investment Products (PRIIPs).

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of mathematical economics*, vol.18(2), pp. 141-153.

Glenzer, F., Gründl, H., & Wilde, C. (2014). 'And Lead Us Not into Temptation': Presentation Formats and the Choice of Risky Alternatives. ICIR Working Paper Series No. 16/14

Golfarelli, M., Rizzi, S. & Proli, A. (2006). Designing What-if Analysis: Towards a Methodology. *Proceedings 9th International Workshop on Data Warehousing and OLAP* (DOLAP 2006), Arlington, USA, pp. 51-58.

Hull, J. C. & White, A. (1990). Pricing interest-rate derivative securities. *Review of Financial Studies*, vol. 3, pp. 573-592.

Hull, J. & White, A., (1994). Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models. *Journal of Derivatives*, vol. 2, pp.7-16.

Hull, J. & White, A., (1996). Using Hull–White Interest Rate Trees. *Journal of Derivatives*, pp. 26-36.

Hull, J., & White, A. (2012). LIBOR vs. OIS: The Derivatives Discounting Dilemma. *Journal of Investment Management*, vol. *11*(3), pp. 14-27.

Hull, T. (2010). A Deterministic Scenario Approach to Risk Management. *Presented at 2010 Enterprise Risk Management Symposium*, Society of Actuaries.

Jarrow, R. A. & Turnbull, S. M., (1995). Pricing derivatives on financial securities subject to credit risk. *The Journal of Finance*, March, L(1), pp. 53-85.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, vol. *29*(2), pp. 449-470.

Minenna, M. (2011). A Quantitative Framework to Assess the Risk-Reward Profile of Non-Equity Products. London, Risk Books.

Minenna, M., Boi, G. M., Russo A., Verzella P. & Oliva, A. (2009). A quantitative risk-based approach to the transparency on non-equity investment products. *Quaderno di Finanza n.63*, CONSOB.

Rizzi, S. (2016). What-if Analysis. *Encyclopedia of Database Systems*, 2nd Edition, L. Liu, T. Özsu (Eds.), Springer.

Schwarcz, S. L. (2009). Regulating complexity in financial markets. *Washington University Law Review*, vol. 87 (2), pp. 211-268.

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

3 How frame affects investment decisions. Empirical evidence.

3.1 Introduction

Performing experiments in the financial field is a way of systematically studying how people's behaviour varies across different conditions. In essence, this methodology involves assigning participants to various different circumstances and observing how the choices they make change. The controlled nature of this process allows researchers to identify causal links between the circumstances and associated differences in behaviour with far greater confidence than could normally be achieved by simply observing consumer choices in real life. Randomized studies are classified according to the settings in which they take place: laboratory experiments involve an intervention in a setting created and controlled by the researcher; survey experiments are conducted face-to-face, over the telephone, or via a web opinion survey; and field experiments are carried out in a naturally occurring setting (Druckman *et al.*, 2011).

Field trials, which explore the effects of particular changes in a controlled way but in real-world situations without manipulating the context, are more likely to be descriptive, developmental and correlational than in a controlled setting. The principal advantage of field research is its generalizability to real-life contexts because they represent a greater variety of situations and environments that subjects experience in their natural habitat (Fisher and Wood, 2007). However, this advantage can be misleading: the lack of control and the impossibility of precisely characterizing the field environment may make it very difficult to judge the generalizability of the study.

On the contrary, controlled experiments normally study decision-making in more stylized environments to capture the main aspects of the decision process that are being investigated. This level of abstraction has important practical advantages because it allows researchers to identify causes and effects more precisely by stripping out irrelevant features that might influence consumer behaviour as well as to investigate whether there are general problems that affect consumers across multiple specific contexts. Due to the greater control that researchers are generally able to exert in a laboratory setting than in a field experiment, controlled studies are more likely to represent a true experimental design. For example, laboratory studies have the advantage of greater control of irrelevant variables that might otherwise influence the results and thus of clearer clues of the behaviour being observed (Bland and Bland, 1994). If controlling all extraneous influences is successfully accomplished, any change observed in the subjects is presumed to be caused by the variable that has been manipulated. This approach comes close to establishing a cause-and-effect relationship. Nevertheless, caution should be taken when considering such a relationship: there is always the risk that an uncontrolled variable influenced the behaviour of the participants biasing the results that were obtained by chance. Another advantage of laboratory research is its reproducibility (Braga *et al.*, 2016). The environmental conditions can be neatly controlled and documented. Like any other research method, disadvantages can be encountered in laboratory research: it may represent an artificial environment that may influence the manner in which subjects behave and therefore alter the results that may be different in the real life.

Reliable causal identification usually gives a high internal validity of findings. In order to be useful with regard to answering questions with policy or academic relevance, it is important for a study to have external validity; the effects established in the experimental setting should still hold true in the real world. A common

perception is that experiments cannot be used to make reliable predictions for the real world because they are conducted in abstract and simplified settings. This idea is questionable: a well-designed experiment can be structured in a way that uses simplicity to its advantage. Indeed, the high degree of control over the environment, which allows the causal effects of specific changes to be isolated, is a defining feature of experimental methodologies. In the real world, although certain things happen together (for example, a complex pricing tariff being used by sellers and consumers failing to get the best deal), it is very difficult to establish that one causes the other. When treatments have an effect on the experiment, despite features of experimental design that should weaken this effect relative to real life, it is reasonable to expect that the identified treatment effects would be present beyond the experiment. However, the experimental setting may be deliberately made simpler than the choices in the real world by making the products consumers wish to compare differ only in price. For example, if consumers struggle to make good comparisons in a very simple setting where the only difference between products is price, we can be confident that their ability to make good decisions would be no better in similar real-life settings which involve comparisons of a much more diverse range of products.

The external validity of results from very abstract laboratory experiments is a topic of ongoing academic debate[50]. An earlier analysis of the use of experiments for policy carried out for Ofcom by Duke *et al*. (2010) also observed that even the more abstract laboratory experiments provide reliable predictions about the real-world differences between treatments. There are surprisingly few studies that explore the same hypotheses, in experimental and in real-life settings, to directly test generalizability of behavioural experiments. However, a systematic review of such studies by Camerer (2011) shows that across a wide range of settings, including collectable card markets, fishing, donations, and exam grading tasks, experimental results generalize well to the real world.

To summarize, experiments can often be particularly useful in exploring consumers' ability to assess offers effectively and make the right decisions about product value. Because the experimenter controls the content and the format of the information that is presented to participants, it is possible to isolate the effects of both the information format and the inherent product complexity on consumers' decisions. With representative samples of participants, it is also possible to explore which types of consumers are most at risk of making errors. Experimental settings are simplified compared with real-life equivalents; thus, if participants have difficulties in an experiment, it is generally easy to predict that problems would also arise in the real world.

A consumer testing exercise and a laboratory experiment were conducted to assess the impact of the two schemes described in Chapter 2 on investment decisions. The analyses are focused on potential retail investors considered as the most affected by behavioral biases as well as the investor category which needs more attention by policy regulators.

Following the guidelines for conducting an experiment with a qualitative external validity and obtaining more realistic quantitative predictions (Iscenko *et al.*, 2014), a consumer test was set to assess how potential investors' choices might be affected by two different disclosure framings of possible future performance obtained by the application of the What-if analysis (WI) and Probabilistic scenarios (PS) approach. Specifically, we aimed to investigate whether the way disclosures of financial information are delivered may influence the perception of riskiness of financial products comparing a score attributed to the analysed subordinated bonds. The research design was a stylized setup, as behavioural experiments typically are, and therefore did not have the purpose of fully reflecting the real market environment. Moreover, the methodology used aimed to explore the effect of the format for future performance representation, while leaving open the

[50] For further discussion on the application of results obtained in laboratory experiments to the real world, the work by Harrison and List (2004) is fundamental. Experimental research is criticized by several studies such as that by Levitt and List (2007a and 2007b), whereas others defend it (for example, Camerer, 2011, or Falk and Heckman, 2009).

question about which costs and risk indicators should be adopted. How investors arrive at their judgements could be studied more deeply in a wider, more detailed investigation.

In addition, a laboratory experiment was conducted. It aimed to objectively measure the impact of Probabilistic and What-if scenarios related to financial decisions through a close look inside consumers' nervous systems. In particular, the main purpose was to investigate behavioural and physiological indexes elicited in the decision-making process of consumers facing the two above-mentioned types of scenarios.

The rest of this chapter is organized as follows. In Section 2, all the aspects linked to the consumer testing exercise are discussed (such as experimental setting, descriptive and multivariate analyses). Section 3 reports all the characteristics and results from the laboratory experiment. Section 4 presents a conclusion.

3.2 A consumer testing exercise[51]

It is valuable to observe how consumer behaviour changes in response to different framings of the same financial products obtained by the What-if and the Probabilistic scenarios schemes analysed in the previous chapter. Not considering details and other aspects (such as picture, name of the issuers, or table of costs) increases control in the experiment because removing factors that are not relevant for the purpose of the research reduces the risk that participants' behaviour would be influenced by contextual features in ways that would make it difficult to isolate the effects of genuine interest. Indeed, introducing more details may threaten internal validity because it may become more difficult to establish that changes in consumer behaviour are caused only by the treatments being tested, rather than other specific details of the designed experiment. Therefore, a simple investment decision was stylized in order to capture the key elements of the consumer decision.

Before describing how the experiment was structured, a further analysis of Product I and Product II introduced in Chapter 2 is necessary: stochastic dominance criteria were tested in order to assess which financial product would represent the best investment solution for a retail investor.

3.2.1 Testing stochastic dominance

When risky prospects are compared, assessments for an optimal choice can be carried out following the literature related to decision under uncertainty. Indeed, by looking at the distribution functions of random variables representing risky prospects and, in particular, comparing the respective cumulative distribution functions, several alternatives could be ordered for investors belonging to a given class of decision-makers, without specifying a particular expression of their utility function.

This section is related to the comparison of the random outcomes of the two considered subordinated bonds, via stochastic dominance criteria (Levy, 2015), which was useful for the design of the consumer-testing exercise. After verifying the necessary conditions concerning the expected value[52] and the so-called 'left tail condition'[53], considering the common domain (*D*) between the two distributions of Product I (PI) and Product II (PII), the sufficient conditions for the three stochastic orders (Table 3.1: *Stochastic dominance tests*) were tested, taking into consideration the cumulative distribution functions ($F_1(x)$ and $F_2(x)$).

[51] Special thanks go to Prof. Fiorenza Deriu from Sapienza - University of Rome, for valuable advice in framing the questionnaires and to Prof. Pasquale Sarnacchiaro from Unitelma Sapienza for supporting the multivariate analysis. I would also thank all the participants that agreed to contribute the research project.

[52] In order to have first order stochastic dominance of F over G, EF(x) > EG(x) should be verified (the relation becomes ≥ when the second and third order are tested).

[53] The G distribution starts to accumulate area (or probability) before distribution F. This is called the 'left tail condition' because the cumulative distributions imply that G has a thicker (in the weak sense) left tail. Note that if the necessary condition does not hold and Min F(x) < Min G(x) = k, there will be a value x_o such that $F(x_o) > G(x_o) = 0$ and, therefore, F cannot dominate G. Thus, if FDG (F Dominates G), it is necessary that Min F(x) > Min G(x).

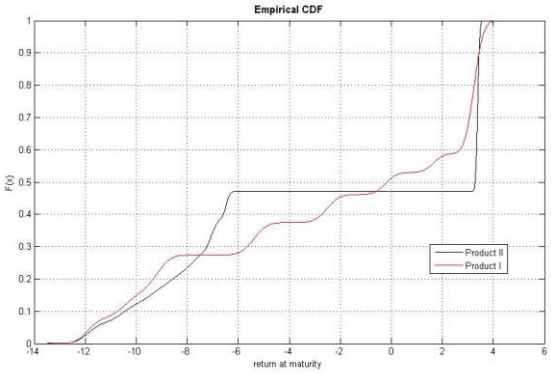
Although for first order stochastic dominance the set of all utility functions representing non-satiable investors (i.e., the set containing all non-decreasing utility functions) was considered, for second order criterion only the subset of all utility functions which were non-decreasing and concave (risk-averse investors) were used. Lastly, the third stochastic order required the set of all utility functions which were non-decreasing, concave, and with a non-negative third derivative ($u'''(x) \geq 0$)[54].

Table 3.1: Stochastic dominance tests.

ORDER	TEST	PII over PI	PI over PII
I	$F_1(x) \leq F_2(x) \quad \forall x \in D$	x	x
II	$\int_{-\infty}^{z} F_1(x)\,dx \leq \int_{-\infty}^{z} F_2(x)\,dx \quad \forall z \in D$	✓	x
III	$\iint_{-\infty}^{z} F_1(x)\,dx \leq \iint_{-\infty}^{z} F_2(x)\,dx \quad \forall z \in D$	✓	x

The first-order stochastic dominance criterion, which is related to investors characterized by strictly increasing utility function, does not allow us to draw any conclusions about the best option, because there is no evidence of dominance between Product I and Product II. On the contrary, results related to second-order stochastic dominance and, consequently, for third type, demonstrate that investors characterized by risk-averse utility function and who are expected utility maximisers should prefer Product II to Product I. The results of the tests are confirmed by observing the cumulative probability distributions of yearly returns obtained when applying the pricing at maturity of the analysed securities (Figure 3.1: *Cumulative distributions of expected yearly return at maturity*).

Figure 3.1: Cumulative distributions of expected yearly return at maturity.



For a better reading of the following sections, a clarification is needed. So far, Product I and Product II were introduced such that Product II stochastically dominated Product I. In the experiments, the analysed financial instruments were randomly renamed Product A and Product B so that: (i) if Product I is renamed Product A, then Product B represents Product II; (ii) if Product II is renamed Product A, then Product B would be Product I. Following this reclassification, the relationship B > A means that Product B stochastically

[54] U^3 represents the class of non-satiable, risk-averse investors who prefer positive over negative skewness. Utility functions exhibiting a decreasing absolute risk aversion (DARA) are important because the investors they represent favour positive over negative skewness. This is a consequence of decreasing risk aversion — at higher payoff levels such investors are less inclined to avoid risk in comparison with lower payoff levels at which they are much more sensitive to risk-taking. Technically, a utility function with a decreasing absolute risk aversion has a non-negative third derivative, $u'''(x) \geq 0$, because this means that the second derivative is non-decreasing.

dominates Product A, whereas A > B means that Product A stochastically dominates Product B, becoming case (i) and case (ii) respectively.

3.2.2 Structure of the experiment and the questionnaires

The experimental setting followed the case-control methodology where the two treatments were represented by the What-if and Probabilistic scenarios schemes. Two different questionnaires (hereinafter Q1 and Q2) were assigned to participants who were divided into two groups (namely G1 and G2). Both the questionnaires were composed by four sections: Section A investigated investment habits and experience, Section B measured the basic level of financial literacy, Section C provided an investment decision between Product A and Product B, and Section D collected sociodemographic information. The only difference between Q1 and Q2 was the format used to represent future performance of Product A and Product B in Section C: in Q1 What-if tables were used whereas the Probabilistic scenarios scheme was used in Q2. The entire questionnaires are presented in Section 4.1: *The questionnaires* of Annex I.

In order to prevent the investment decision from being driven by familiarity, anchoring, and representative heuristics, the issuer names of Product A and Product B were substituted respectively by Bank Alpha and Bank Beta. Moreover, a randomization of the order in which the financial products were presented to the participants was set in order to exclude biases in participants' investment choice caused by the so-called 'first impression effect'. In this way, the experimental setting-controlled cases (recorded by the variable *R*) when A > B (i.e., case (ii) above) and B > A (i.e., case (i) above) following the second order stochastic dominance criterion.

The questionnaires were delivered to participants using a web computer-based task which was the least time-consuming way to engage with a wider range of potential retail investors.

The survey aimed to investigate the impact of financial information disclosure on risk perception and investment decisions generated by the two frames. In particular, it was possible to assess which methodology would reduce potential biases and allowed interviewees to a correct risk perception and an investment decision coherent with the stochastic dominance criteria. Thanks to the set of questions used, additional analyses could be performed on the appraisal of the representation schemes in terms of complexity, usefulness, and sufficiency of the information content.

3.2.3 The sample

In general, experimental data are obtained from studies that involve a random allocation of subjects to different treatments of one sort or another; however, in sample survey studies subjects are randomly chosen from a larger study population. In a lot of experimental research, university students are used as participants because their generally high cognitive skills and adaptability to abstract setups reduces the errors made in the real world by the general public: if students make mistakes in an experiment, then the errors among a more diverse population of consumers are likely to be at least as large (Druckman and Kam, 2011). Although these arguments hold in several cases, there are situations where a different target population should be considered. For example, considering the purpose of this thesis, it may be particularly important to understand whether more vulnerable consumer groups make certain mistakes or benefit from the two schemes being tested, without forgetting that errors are not universal across all consumers. Indeed, the consumer test was also designed to identify which categories of potential retail investors are particularly at risk of errors and to examine whether the effects are similar across consumers, and, if not, what the relative magnitude of the effects might be for different groups. Moreover, to explore the effectiveness of the analysed treatments and use the obtained results as a starting point for practical solutions, a wide range of

potential beneficiaries of that treatments should be considered as the target population. Thus, considering a wider demographic range of participants than students (or other sub-groups) is more appropriate.

Considering the available budget, the sample selection process and sample size were aimed at reducing the possible risk of distorted representation of particular categories in the sample, and obtaining evidence as close as possible to the target population. Due to the variability of characteristics among individuals in the population, the used sampling design is a combination of a sample survey and experimental data processes: after selecting a study population, treatments were randomly assigned to the resulting study subjects. Therefore, Italian people aged between 20 and 64 years old was considered to be the target population. In particular, using data from the most recent survey produced by ISTAT (2017), two stratified samples (i.e., G1 and G2) were obtained, dividing the target population into subpopulations using gender, age, geographical area[55], and education level as stratification variables (for more details on the sampling procedure see Section 4.2: *Sampling procedure* of Annex I)[56]. In the following subsections, the sample is described in terms of sociodemographic characteristics, investment habits and experience, and financial knowledge.

3.2.3.1 Characteristics of the sample: sociodemographic, investment habits and experience

The web survey methodology, which had a CAWI approach (Computer Assisted Web Interviewing), was used to collect data. The final dataset was composed of 1130 valid interviews[57] (G1 = 544 units under WI treatment and G2 = 586 units for PS). The strata of the sample design was obtained using gender, age, geographical area, and education as stratification variable. The weight of each stratum was proportional to the ones in the target population represented by potential Italian retail investors aged between 20 and 64 years. The obtained sample represents the target population with a marginal error of ±3%. In particular, about 49% were male. The North-West was represented by 25% of the sample, the North-East by about 17%, 25% were from the Centre, and the remaining 33% were from the South of Italy. Considering the level of education, 29% of the respondents had obtained at most the degree of 'Licenza Scuola Media' or lower, whereas the participants with at least a University degree made up 28% of the sample. The remaining 43% of the sample had not achieved more than the degree of 'Diploma Scuola Superiore'. The age distribution was homogeneous between classes: 32% were aged between 20-34, 35% in the range 35-49, and 33% were 50-64 years old.

Moreover, about 36% were single, while about 58% were married or cohabiting, and less than 6% belonged to residual categories (5% were divorced or separated and 1% were widowed). On average, a family accounted for three units of which at least 1 had a stable income. About 46% of the respondents were household head[58] (of which 64% male and 81% aged between 35 and 64 years old). Analysing their employment status, 63% were employed (of which 79% were employees, 6% did occasional work or were freelance, and about 15% were self-employed); 18% were unemployed; and inactive people were approximately equal to 19% (of which roughly 36% were students, 44% were housewives, 6% were unemployed and not looking for a job, and 13% were retired). Considering the field of specialization, there was heterogeneity among the participants: 19% were experts in biology, chemistry, or medicine; 17% had

[55] As specified by the NIELSEN Company: North-East – Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia and Emilia-Romagna; North-West – Valle d'Aosta, Liguria, Piemonte and Lombardia; Centre – Toscana, Umbria, Marche, Lazio and Sardegna; South – Abruzzo, Molise, Campania, Basilicata Puglia, Calabria and Sicilia.

[56] Each stratum was obtained by crossing the weights of each variable. Therefore, the sample was composed of 72 strata.

[57] A total of 328 suspicious observations were deleted from the dataset: i.e., very short time to complete the questionnaire, too many 'I do not know' answers, or it appeared that answers were selected randomly without evaluating the options.

[58] Household head is identified as the person who is responsible for the financial and economic choices of the household. The Questionnaires did not contain a specific question to identify them. Therefore the answer to question D.10 ("Consider who has the higher income in your family. Please, select your relationship with this person") was used as a proxy: a household head is identified when the respondent answered 'myself' to the question.

skills in humanities, philosophy, or psychology; 31% were economists; 13% were specialized in law; and 14% in physics, engineering, maths, or statistics. The remaining 6% of the participants belonged to the residual category.

Considering the described characteristics of the sample, not only was the typical household financial decision-maker considered in the experiment, but also marginal categories of potential retail investors. Indeed, according to GfK Eurisko, which collects retail market survey data relating to a representative sample of approximately 2500 Italian households, the typical household financial decision-maker is aged within the 35-39 years age bracket on average, and holds a bachelor's degree in only 15% of the cases (Linciano *et al.*, 2018).

As for investment habits among the participants, saving was a widespread behaviour: 42% of people declared that they are able to save part of their income. A large proportion of respondents (about 64%) have invested in a financial instrument at least once in their life, but only 58% of them made their investment decision after having consulted a financial expert (71%), a relative (26%), or the media (1%). About 42% of candidates made the decision on their own. Considering the respondents with past experience in financial investment, the vast majority know and use deposits (48%), bonds (33%), and pension funds (32%). Only a small percentage of individuals declared to have invested in derivatives (7%).

More detailed information about the distribution of the sample, such as split between the two treatments, is reported in Table 4.3: *Summary of sociodemographic characteristic of the sample* and Table 4.4: *Investment habits and experience* in Annex I.

3.2.3.2 Definition and measure of financial literacy

Financial literacy has been strongly increasing over time and how to define it has become an important aspect to consider. Some financial literacy measures reflect only the ability to use information to make correct judgments in order to manage personal financial conditions (Noctor *et al.*, 1992; Anderson *et al.*, 2000; Danes and Habermann, 2007). Empirical studies build financial literacy indicators exclusively on the basis of numeracy – that is, the skill to use numbers and mathematical approaches – and investors' knowledge of basic financial concepts (Hilgert *et al.*, 2003; Lusardi, 2008a; Lusardi,2008b; Lusardi *et al.*, 2012). Other experts evaluate financial literacy levels from the basic knowledge necessary to manage simple financial tasks (see, for example, Kim, 2001). More sophisticated studies provide a clear and comprehensive definition of financial literacy which should consider both the knowledge of basic concepts and how to apply them in financial decisions (US FLEC[59], 2006; Lusardi and Tufano, 2015; Mandell, 2008) or take into consideration decision-making competence defined as the ability to avoid investment mistakes which is driven also by behavioural biases (see Bachman and Hens, 2015; Bruine de Bruin *et al.*, 2007; Finucane *et al.*, 2005; Levin *et al.*, 2007; Parker and Fischhoff, 2005). Considering these approaches, the need to separately measure knowledge and ability has become fundamental (Huston, 2010; Remund, 2010). In a recent survey aimed at measuring the level of financial literacy around the world, the Organisation for Economic Co-operation and Development (OECD) computed a global indicator which takes into consideration the basic components of financial knowledge, individuals' behaviour when taking financial decisions, and the financial attitude to medium-/long-term investment (Atkinson and Messy, 2012).

As demonstrated by this short overview, the definition and measurement of financial literacy used in empirical literature is not univocal (Hung *et al.*, 2009). Therefore, when research involves an assessment of financial literacy, it is necessary to understand which aspects are interesting for that particular study. A possible solution is to consider the investigation of general financial principles, such as inflation and the

[59] US-FLEC is the Financial Literacy and Education Commission in the United States.

money illusion, the interest effect, and the diversification of risk in order to assess possible relationships between financial and other variables (among others, see Chen and Volpe, 2002; NEFE, 2006; Kempson *et al.*, 2006; and ANZ, 2008).

Because this research does not aim to infer the level of financial literacy in the target population, introducing simple questions to explore the level of basic financial knowledge could be useful to analyse financial behaviours. Following a multiple-choice design, both the questionnaires used in the consumer test contained five questions to test numerical skills (Question B.1) and to investigate whether participants were familiar with financial concepts such as inflation, funds, and diversification (see Questions B.2, B.3, and B.4 respectively). All these questions are similar to the ones used by the OECD (2005). In addition, a question was introduced in order to test participants' ability to evaluate two different financial products when their possible future returns are shown (see Question B.5). It should be noted that the 'I do not know' option (abbreviated as 'dnk') was recommended to participants rather than selecting a random choice when they did not know the correct answer. This option was also used to assess the self-confidence of the participants: besides being an admission of financial illiteracy, by selecting the 'I do not know' option, individuals demonstrate a feeling of under-confidence in a financial context, which is more useful than the alternative of not even trying to answer. Table 3.2: *Questions to test financial knowledge* shows the distribution of the answers to the test on financial literacy.

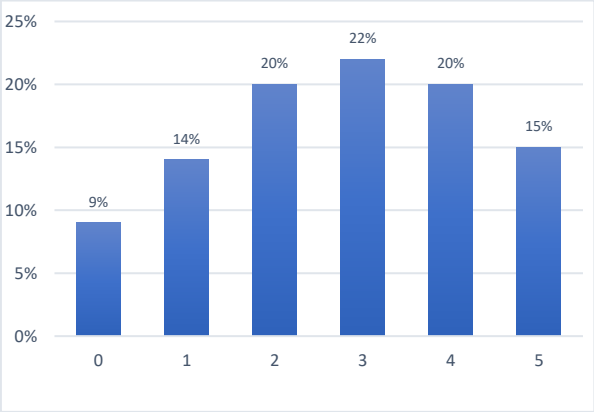
Table 3.2: Questions to test financial knowledge.

		Frequency			**Percentage**		
		dnk	wrong	correct	dnk	wrong	correct
Question on:							
B.1	Interest rate	107	235	788	9.47	20.8	69.73
B.2	Inflation	196	242	692	17.34	21.42	61.24
B.3	Fund	362	118	650	32.04	10.44	57.52
B.4	Diversification	342	346	542	21.42	30.62	47.96
B.5	Prospect	217	449	464	19.2	39.73	41.07

Despite showing familiarity with investment decisions and concerns for economics and financial news, respondents who knew the meaning of portfolio diversification accounted for only 48% and the percentage of participants who were able to compute simple calculations of future returns was about 70%. Inflation was correctly understood by about 61% of the participants. Although investments in funds was spread among candidates, most of the subjects (42%) were not able to correctly identify the main characteristics of these financial products. Only 41% of the sample were able to identify the correct investments when possible returns and the linked likelihood of occurrence were shown. These results are in line with the more accurate survey conducted on Italian people in 2013 (Bongini *et al.*, 2014). Only 57% answered correctly to at least three questions (Figure 3.2: *Number of correct answers about financial knowledge*). In order to obtain a financial literacy indicator, correct answers should be aggregated into a single measure. Although the financial literacy level could be measured by different approaches such as, for example, discriminating individuals that answered all questions correctly compared with the others (as in Lusardi and Mitchell, 2011; Almenberg and Save-Söderberg, 2011) or by summing the correct and wrong answers after being weighted by degree of relative complexity (Lusardi *et al.*, 2012), in the present work, a financial literacy score was

obtained by summing the correct answers (see Moore, 2003; Agnew and Szykman, 2005; Robb and Sharpe, 2009; Gerardi *et al.*, 2010)[60].

Figure 3.2: Number of correct answers about financial knowledge.



More specifically, firstly, answers to Section B of the questionnaires were dichotomized as being equal to 1 when they were correct and 0 otherwise; then, for each individual all correct answers were summed in order to obtain a score from 0 to 5. Finally, the distribution was divided into three categories: low, medium, and high, which respectively collected all the individuals with a score between 0–1, 2–3, and 4–5. Moreover, it is worth noting that because the questions were on very basic concepts, individuals having 'high' financial literacy does not imply that they are financially skilled: indeed, correct answers to easy questions do not grant that respondents are financially literate, while failing some questions is enough to identify individuals as having low or no literacy.

The average score in the sample was 2.7. As in most of previous works on financial literacy (such as Bongini et al., 2014; Hung et al., 2009; Lusardi et al., 2012; Van Rooij et al., 2007;), men tended to be more literate, with an average score equal to 3 (2.5 for women). In particular, as in Bucher-Koenen et al. (2017), data showed the so-called 'gender gap': women are less financially literate than men, independent of age, educational level, area of residence, marital status, employment status, and their experience in financial investments. Moreover, the degree of financial knowledge tended to be positively correlated with education. More details about the distribution average scores and the number of correct answers by sociodemographic attributes are reported in Table 3.3: *Average financial literacy score* and in Annex I (Figure 4.2: *Number of correct answers about financial knowledge by attributes*).

Because definition and measurement of financial literacy are not univocal and some indicators may perform better than others, the described values were not be in line with more accurate surveys aimed at analysing the level of financial literacy. Keeping this in mind, the financial literacy score helped to identify which sub-groups in the sample were more vulnerable to financial mistakes, due to their poor knowledge. Finally, an alternative indicator of financial knowledge was used as a robustness check in the multivariate analysis (i.e., a dummy variable equal to 1 when all the five answers were correct).

[60] Other approaches used to build a financial literacy score are: principal component analysis applied to the answers rescaled on the basis of the easiness of questions (i.e., questions recording a lower rate of correct answers have a greater weight) in order to use information about respondents' financial sophistication level (Lusardi *et al.*, 2012), and factor analysis and the distinction between correct, incorrect, and 'do not know' answers (Kramer, 2014; van Rooij *et al.*, 2011).

Table 3.3: Average financial literacy score.

		Average score					Average score		
		F	M	Tot			F	M	Tot
Age					Employment status				
	20-34	2.6	3.2	2.9		employed	2.7	3.1	2.9
	35-49	2.5	2.8	2.6		not employed	2.3	2.3	2.3
	50-64	2.5	3	2.7		inactive	2.3	3.3	2.7
Education					Marital status				
	low	2.1	2.3	2.2		single	2.6	3	2.8
	medium	2.5	2.9	2.7		cohabitant/married	2.5	3	2.7
	high	3.2	3.7	3.5		divorced/widower	2.4	2.6	2.5
Geographical Area					Financial experience				
	North-West	2.7	2.9	2.8		yes	2.8	3.3	3.1
	North-East	2.4	2.9	2.7		no	2.1	2.3	2.2
	Centre	3	3.4	3.2	Household head				
	South	2.1	2.8	2.4		yes	2.6	3.1	2.9
						no	2.6	2.9	2.7

3.2.4 A descriptive analysis of investment choices, risk assessment and frames appraisal

In Section C of the questionnaires, the effect of the two treatments on decision-making and risk assessment of the presented products were analysed. Moreover, interviewees were asked to rate the usefulness, clearness, sufficiency, and comparability of the information provided by the two frames. These questions were used to evaluate frames appraisal.

3.2.4.1 Investment choice and risk perception

In order to investigate whether framing affects risk assessment of the presented products, it should be remembered that Product I is riskier than Product II. This assessment is based on returns distributions at maturity obtained in Chapter 2 which consider market and credit risks. Participants were asked to give a score from 1 to 7 (where a 7 means 'very risky') according to their perceived risk levels. If the presentation format does not affect risk perception, on average respondents should be able to assess the same level of risk for the same product across the different treatments.

The number of respondents that assigned the same risk level to Product I and Product II was consistent between treatments: 38% in WI and 36% for PS. However, when comparing WI and PS, marked differences could be noted. In particular, Product I was correctly considered to be the risker product only in 23% of the cases, when it was framed using the What-if scheme; this proportion was almost double (43%) when the Probabilistic tables were used.

Figure 3.3: Risk assessment by treatment.

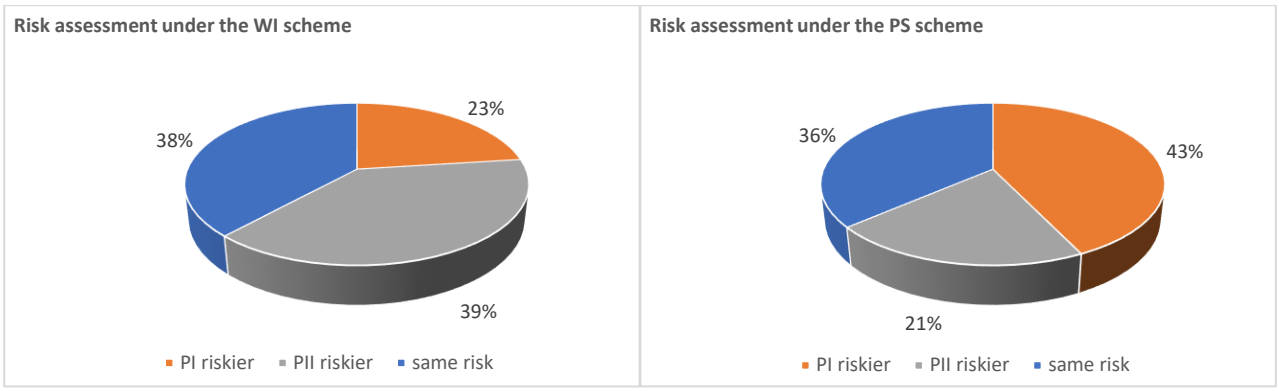
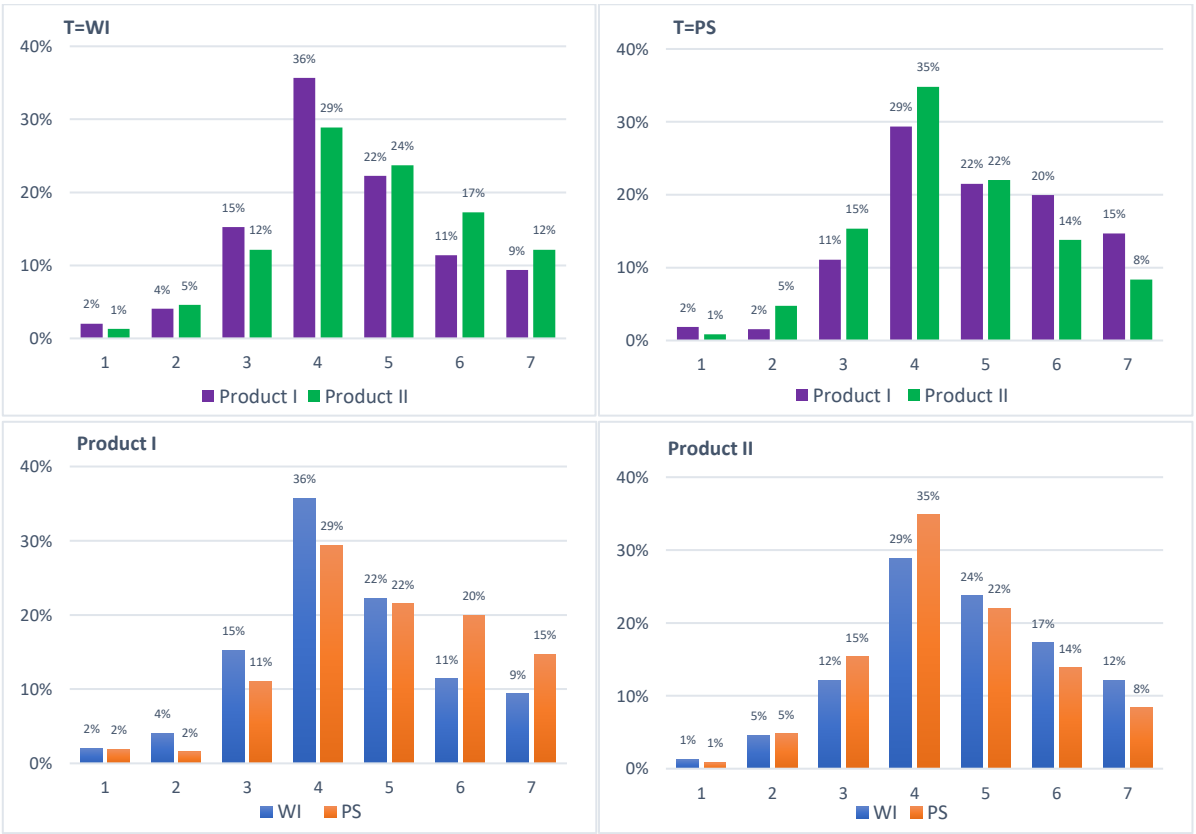


Figure 3.4: *Risk score of Product I and Product II by treatment* shows the single score (i.e., risk assessment) assigned to Product I and Product II splitting the answers both by treatments and by products. Higher levels of risk (i.e., 6 and 7) were assigned to Product I in 35% of the cases under PS, while this percentage was only 20% under the WI scheme. A wrong evaluation of Product II (i.e., level of risk between 6 and 7) was roughly equal to 29% in WI representation. On the contrary, respondents deeming Product II as having a high-risk level was reduced to 22% when Probabilistic tables were used. Under both representation schemes, the proportions of participants who perceived the two products as low-risk (i.e., score between 1 and 3) were equal to 39% and 36% for WI and PS, respectively. In particular, looking at lower score of riskiness, WI representation communicates a wrong level of risk of Product I for about 21%. These percentage is lowered to 15% in case of PS representation. On the contrary, only 18% and 21% of respondents correctly evaluated Product II as having a low-risk level under WI and PS, respectively.

Figure 3.4: Risk score of Product I and Product II by treatment.



Considering the overall impact of the different frames on risk assessment, individuals showing similar sociodemographic characteristic may be affected differently by the two risk-return representations. Group differences were tested along some variables, accounting for individuals' characteristics with respect to participants' products risk assessment (see Table 3.4: *High risk perception and sociodemographic characteristics*). For this purpose, seven dichotomized variables were considered, accounting for sociodemographic characteristics and personal features of the participants (e.g., financial knowledge). As for the sociodemographic aspects, no regular patterns were detected, excluding geographical area: the perception of high risk associated with Product I was more frequent for individuals from the South of Italy to whom What-if tables were displayed. As for financial knowledge, highly financially literate individuals seemed to be more frequently able to rank Product I and Product II correctly. In particular, interviewees classifying Product I as the riskiest predominantly had a high level of financial knowledge when Probabilistic tables were used.

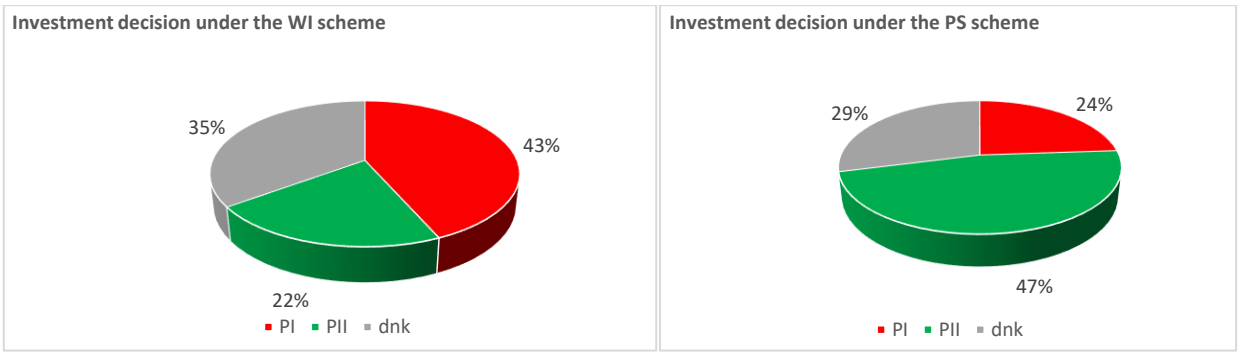
Table 3.4: High risk perception and sociodemographic characteristics.

	Sample		**WI**		**PS**	
	PI	PII	PI	PII	PI	PII
Male vs Female						
50-64 years old vs other age classes						
High vs other levels of education						
South vs other geographic areas			+	-		
Employed vs other						
Household head vs other						
High vs other financial literacy levels	+	-			+	-
Experienced vs no experience						

Note: The tests were performed on 709 observations after excluding individuals who considered Product I to be as risky as Product II. The variables were dichotomized: for example the variable 'age' was equal to 1 if the participant was between 50 and 64 years old, the variable 'financial literacy score' was equal to 1 if the participant scored 4 or 5 for the linked questions, and so on. The notation +/- indicates the sign of an at least 5% statistically significant difference between the risk assessment of the two products (PI = Product I riskier, PII = Product II riskier) by dividing the sample along the lines of the attributes considered to be dichotomous independent variables. Moreover, the analysis was performed according to a Pearson's chi-squared test and considering the treatments separately (i.e., sample, WI and PS). As an example, individuals with x = 1 classified PI as riskier (y = 1) more frequently than individuals with x = 0 under the WI scheme. Blank cells indicate that the difference was not statistically significant.

Before making the risk assessment and eliciting an appraisal of the frame, respondents were asked to state their willingness to invest in each of the offered products, taking into consideration only the available information. In order to control for individual time horizon, investment objectives, wealth, and mental accounting effects, all subjects were given the same hypothetical investment conditions: they were invited to assume that 1,000 euros should be invested over the next 7 years. It should be remembered that by choosing Product II a decision coherent with the stochastic dominance criteria is made. Therefore, the higher the probability of choosing Product II, the higher is the effectiveness of the representation scheme in reducing potential biases for a risk-averse investor due to framing effect. As result of this task (Figure 3.5: *Investment decision by framing*), we see that Product I was preferable under the What-if representation (43%, only 24% in case of PS representation) while Product II was mostly chosen when Probabilistic tables were shown (47%, only 22% in case of WI representation). Also, in this case, the frames influenced investment decision. Hesitant individuals (i.e., who answered 'do not know') were not significantly different between treatments.

Figure 3.5: Investment decision by framing.



Considering the investment decision and the risk assessment, there was a significant relationship between the variables (p-value < 0.001 when Pearson's chi-squared test was performed). This evidence seems reasonable because each investor makes their investment decision only after evaluating the risk-return profile of the proposed investments. Therefore, the frame firstly affects the risk assessment process and only indirectly the investment decision. Overall, this preliminary evidence shows that the representation format matters. A deeper analysis of these aspects is provided in paragraph 3.2.5: *Multivariate analysis of risk perception and investment choice*.

3.2.4.2 Appraisal of the two representation schemes

Probabilistic tables are favoured when interviewees were asked to choose which representation format is preferred (see Question C.5 in the questionnaires). Indeed, about 53% of the participants selected the PS frame whereas only 18% opted for WI scheme. The residual 29% did not express any preference (for more details see Table 4.15: *Framing preference* in Annex I).

In order to investigate more the effectiveness of the two schemes, individuals were asked to evaluate the risk-return representations in terms of usefulness, clearness, sufficiency, and comparability of the information content. Specifically, each respondent rated the presented framing considering the four attributes on a 0-5 Likert scale. In addition, for each participant a final score (labelled 'score') was calculated as the average of votes given to each attribute. The average score for each attribute rises moving from the What-if representation to the Probabilistic scenarios (see Table 3.5: *Frame evaluation (frequencies)* and, for more details, Annex I). The evidence supports that, overall, the Probabilistic scenarios scheme is more appreciated than the What-if information sheet (final average scores are equal to 2.7 and 2.5 respectively).

Table 3.5: Frame evaluation (frequencies).

	Useful		**Clear**		**Sufficient**		**Confrontable**	
Score	*WI*	*PS*	*WI*	*PS*	*WI*	*PS*	*WI*	*PS*
0	15	6	45	13	49	36	19	6
1	71	70	83	76	101	108	46	57
2	144	150	161	155	181	168	134	138
3	206	210	166	189	149	174	193	191
4	86	110	71	121	56	84	110	140
5	22	40	18	32	8	16	42	54
Avg. score	**2.63**	**3**	**2.35**	**2.93**	**2.16**	**2.54**	**2.84**	**3.19**

Using the non-parametric Mann-Whitney-Wilcoxon test[61], the difference in means between the two representation schemes was determined. The mean of the distribution of the overall score is significantly different between frames. The difference is mainly driven by clearness and sufficiency of the information provided. Considering the results reported in Table 3.5: *Frame evaluation (frequencies)*, a higher average score is expected for the Probabilistic tables. Therefore, the one-side test[62] was also used for the average. Excluding the comparability attribute, the mean score given to the Probabilistic representation is higher than the What-if one. The analysis was repeated using the non-parametric test for the median. The results confirm that clearness is the attribute which mainly differentiates the score distribution between the schemes (two-sided test) and that the median score for clearness and sufficiency are significantly higher for the Probabilistic representation. In addition, the non-parametric Kolmogorov-Smirnov test for two groups[63] was performed to verify that the distribution of *y* (i.e., score, usefulness, clearness, sufficiency, and comparability) was the same between the treatments. The analysis shows a significant difference in distribution for the overall score and clearness between representation schemes. More details are in Table 3.6: *Differences in scoring between treatments* below.

Table 3.6: Differences in scoring between treatments.

	Mann-Whitney Wilcoxon Pr>|Z| (two-sided)					**Mann-Whitney Wilcoxon Pr<Z (one-sided)**				
	Score	Useful	Clear	Suff.	Conf.	Score	Useful	Clear	Suff.	Conf.
WI vs PS	***	*	***	**		***	*	***	**	

	Median Pr>|Z|(two-sided)					**Median Pr<Z (one-sided)**				
	Score	Useful	Clear	Suff.	Conf.	Score	Useful	Clear	Suff.	Conf.
WI vs PS			***	*		*	*	***	**	

	Kolmogorov-Smirnov two-sample test (Pr>KSa)				
	Score	Useful	Clear	Suff.	Conf.
WI vs PS	*		**		

Note: The tests were performed on the entire sample. The notations *, **, and *** indicate the sign of a 5% , 2.5%, and 1% statistically significant difference between the two groups given by the treatment's variable (i.e., p-value < 0.05, p-value < 0.025, and p-value < 0.01). Moreover, the analysis was performed according to the non-parametric Mann-Whitney Wilcoxon test for the mean, the median test (both the two-sided and one-sided), and the Kolmogorov-Smirnov two-sample test. Blanks cells indicate that the difference is not statistically significant.

Sociodemographic characteristics of respondents affected frame appraisal. In order to gain an insight into the relationship between the appraisal of the representation of financial products and the respondents' sociodemographic characteristics, statistical tests were conducted for group differences in the perceived complexity, sufficiency, comparability, and usefulness, by dividing the sample into mutually exclusive groups along the lines of some independent variables. In particular, seven variables were selected: gender, age, employment status, area of residence, education, previous investment experience, and financial literacy. Given the results of the non-parametric Kruskal-Wallis two-sided test for the median, the overall median

[61] H0: no difference in mean of *y* between groups given by WI and PS schemes. The two-sided test verifies differences in mean: if Pr > |Z| < alpha, then H0 is rejected and the mean of *y* (score, usefulness, clarity, sufficiency, and comparability) is significantly different between treatments.

[62] H0: mean of *y* in Group 0 (i.e., WI) is higher than mean of *y* in Group 1 (i.e., PS). If Pr < Z < alpha, then H0 is rejected and the mean of *y* in WI is lower than the one in PS.

[63] With (Pr > KSa) < alpha, the null hypothesis (H0: no difference in distribution) is rejected; the distributions of *y* were not the same for the two sub-populations (treatment).

score was significantly different under the Probabilistic scheme considering age and geographical area, whereas the financial literacy score influenced both the What-if and Probabilistic representations. Stratifying the sample among categories[64], the above-described non-parametric tests for the mean and the median were performed considering WI and PS as Group 1 and Group 2, respectively. The gender seems to be relevant in the assessment of clearness: both men and women consider PS to be clearer than WI. With regard to age stratification, almost all categories prefer PS to WI in terms of clearness, usefulness, and sufficiency of the information. Also, interviewees with low and high financial literacy assigned higher scores to the Probabilistic representation. On average, people from the South, the Centre and North-East of Italy appreciate the PS schemes more.

More details on these statistical tests are reported in Table 4.16: *Differences in scoring between sociodemographic characteristics by treatments*, Table 4.17: *Differences in mean scoring between treatments by levels of sociodemographic variables*, Table 4.18: *Differences in median scoring between treatments by levels of sociodemographic variables* and Table 4.19: *Differences in distribution between treatments by levels of sociodemographic variables* in Annex I.

3.2.5 Multivariate analysis of risk perception and investment choice

The main assumption is that the two different risk-reward representations might impact directly on the risk assessment of investment products. In this respect, each participant in the experiment was asked to assess the risk of the selected two subordinated bonds which were characterized by different level of risk. Hence, the analysis is focused on the impact of the two frames (described by the variable $T = 0$ for WI and $T = 1$ for PS) on the binomial variable 'Risk', Y_i, for the individual i evaluating the risk level of Product I and Product II (obtained from Question C.3), labelled as Risk I and Risk II respectively:

$$Y_i = \begin{cases} 1, & \text{if Risk I} > \text{Risk II} \\ 0, & \text{otherwise} \end{cases}$$

Although the analysis could be performed by a logistic regression (see Table 4.21: *Estimation results from logistic regression* in Annex I), given the structure of the data there are reasons why it might be worth moving to a multilevel analysis for purposes of causal inference. In particular, individual- and group-level variations could be taken into consideration when estimating group-level regression coefficients and to model variation among individual-level regression coefficients[65]. Therefore, the following research questions should be answered:

Q1: What is the extent of between-treatment variation in assessing the risk of the two products?

Q2: Do individual-level variables, such as age, gender and financial literacy, have different effects in different treatments?

3.2.5.1 Model selection

In order to consider the heterogeneity of the individuals' data (level 1) grouped by the two treatment (level 2), the multilevel random effect models were used with the purpose of identifying the factors

[64] For example, for the variable 'education' the analysis was performed considering three univocal sub-samples composed respectively of participants with low, medium, and high levels of education.

[65] In classical regression, one can do this using indicator variables, but multilevel modeling is convenient when we want to model the variation of these coefficients across groups, make predictions for new groups, or account for group-level variation in the uncertainty for individual-level coefficients.

influencing the risk assessment of the financial products. In particular, Y_{ij} was specified for each *i*-th individual and the *j*-th group was represented by the treatment such that:

$$Y_{ij} = \begin{cases} 1, & \text{if Risk I} > \text{Risk II} \\ 0, & \text{otherwise} \end{cases}$$

Because the variable is binomial, the multilevel logistic regression model specification was used. Firstly, the random-intercept specification was identified: a frame-specific (level 2) random intercept was included in the linear predictor, thus enabling us to explicitly model the two clusters of the data and their potential unobserved heterogeneity (Goldstein, 2011; Raudenbush and Bryk, 2002). Secondly, a random-slope specification was estimated in order to assess whether and to what extent the effect of individuals' characteristics varied across treatments. In this way, the assumption of the random-intercept model could be relaxed (i.e., the slope was fixed across treatments), thus analysing the between-treatment heterogeneity. For the *i*-th individual in the *j*-th group, the multilevel odds model was based on the logit link and fitted according to the following equations (Hox, 2017; O'Connell *et al.*, 2008; Raudenbush and Bryk, 2002):

$$\text{level 1: } logit(Y_{ij}) = logit\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \sum_{q=1}^{Q} \beta_q X_{qij} \quad (1)$$

$$\text{level 2: } \beta_{0j} = \gamma_{00} + u_{0j} \quad (2)$$

In these expressions, the level 1 equation represents the individual-level model and the level 2 equation defines the treatment-level models. For the *i*-th individual in the *j*-th group, Y_{ij} represents the odds for the success category (Y_{ij} = 1) based on q = 1, …, Q individual-level explanatory variables. The expression on the left-hand side of equation (1) is the natural log of the odds and it is referred to as the logit distribution. The expression in the middle of equation (1) – $\sum_{q=1}^{Q} \beta_q X_{qij}$ – can be interpreted in the same way as the regression coefficients for the multilevel logistic model. The intercept for the *j*-th group is represented by β_{0j}. The level 2 equation describes how the within-group effects may vary according to group-level characteristics. In particular, the intercept is the level 2 variable which is composed of the estimated value (γ_{00}) and the term u_{0j} which is the group-specific random effect and which also represents the level 2 error[66]. It should be noted that specific group variables were not considered in the experiment. For random slope models, β_q was replaced with β_{qj} which represented the collection of treatment-specific slopes for the group in terms of expected change in the logit for each change in the respective within-group explanatory variables X_{qij}.

Several models were estimated using R software and implementing the lmer4 package, which allowed us to fit multilevel mixed-effects logistic models (Bates *et al.*, 2015). Because the response variable was binary, a generalized linear mixed model with a binomial distribution was needed. The parameters estimation of these models involved log-likelihood calculations which required integrating out the random effects. Because there are fewer than five random effects, the Laplace approximation was used, which is a special case of a more accurate parameter estimation method called Gauss–Hermite quadrature, with one iteration. Finally, because the log-likelihood itself was estimated, log-likelihood ratio tests were performed to choose the model that best fit the data.

3.2.5.2 Model specification

Bearing in mind the experimental set, the selection of the variables to be included in the models was carried out by computing association tests according to the nature of each variable, as reported in Table 4.20:

[66] $u_{0j} \sim N(0, \sigma_u^2)$

Mantel-Haenszel Statistical test for independence of Annex I (Hosmer and Lemeshow, 2000). Following the common process for developing multilevel analysis (Peugh, 2010; Aguinis *et al.* 2013; Hox, 2017; Sommet and Morselli, 2017), the sociodemographic characteristics of the respondents were introduced in the models as individual-level covariates, namely: gender; age expressed in three classes (20-34, 35-49, and 50-64); area of residence (coded 'GEO') considering four categories North-West, North-East, Centre, and South of Italy; level of education ('EDUC') in three classes[67], marital status, job status and a dummy variable to identify household head status ('HH'). Then, the dimensions related to the investment habits of the respondents, such as being updated about financial news (dichotomous variable labelled as 'NEWS') and a dummy variable 'EXP' to capture their past experience in financial investment, as well as a categorical measure of financial literacy level ('FL') were considered. Finally, in order to control for possible effects of the order in which Product I and Product II were displayed to the participants (see Section 3.2.2), the dummy variable *R* was included[68]. It should be highlighted that because most of the variables were multi-categorical, one level should be considered as a reference value and then omitted from the estimation results.

As a first step, an empty two-level model was fitted (also referred to as an 'unconditional mean model') – i.e., a model containing only an intercept defined by the treatment effects (*T*) and without the fixed part of explanatory variables – so as to assess the variation of the log-odds from one cluster to another. Then, the intraclass correlation coefficient (ICC) was calculated to measure the degree of homogeneity of the outcome within clusters. It transpired that a variation between clusters existed and that 10% of the chance of making the correct risk assessment could be explained by the between-strata difference and, conversely, the 90% was represented by the within strata differences.

The analysis was continued by building an intermediate model adding explanatory variables (the so-called CIM, constrained intermediate model) so as to assess the variation of the lower-level effects from one cluster to another. Finally, the augmented intermediate model (AIM) was obtained, starting from the best CIM: the model is similar to the constrained intermediate model with the exception that it includes the random slope term of FL or EDUC. Then, the goal was to determine whether the AIM (random intercept–random slope) achieves a better fit to the data than the CIM (random intercept). In other words, our goal was to determine whether considering the cluster-based variation of the effect of the lower-level variable improves the model. To do so, after gathering and storing the deviance of the CIM and AIM, a likelihood-ratio test (LRT) was performed. Comparing the random slope model, which allows FL to vary across strata and the CIM, the test was statistically significant (p-value < 0.05); hence, there is evidence that the effect of financial literacy varies across treatments.

3.2.5.3 Results

Before considering possible effects of covariates, the multilevel model with only the intercept *T* was compared to the empty logit model. The LRT statistics had a minimal p-value, so the null hypothesis was rejected[69]. Therefore, there is evidence of unobserved heterogeneity at treatment level: as expected, individuals evaluate the risk of Product I and Product II differently under the two frames.

The analysis continued by adding covariates at the individual level (level 1), representing respondents' sociodemographic characteristics (such as age, gender, and education), their investment habits, household head status and financial literacy score, and a parameter related to the experimental setting (*R*). Some interactions among them were also performed. Estimates of the fixed effects (β_{qj}) are reported in Table 4.22:

[67] The higher degree of education was obtained by three classes: 'low' when the respondent obtained the degree of Licenza Scuola Media or lower; 'medium' when the respondent earned at most the degree of Licenza Scuola Superiore ; and 'high' when the respondent at least gained a university degree.

[68] When R = 0, Product I appeared as the first investment proposal, whereas when R = 1 Product II was first.

[69] H0: variance between groups was 0.

Estimation of fixed effects from multilevel logistic regression with random intercept in Annex I. Because most of the sociodemographic individual characteristics, as well as attitudes to financial investment, seemed not related to the respondents' propensity to make a correct risk assessment, the only variables which were statistically significant in the model were the education level and the financial literacy score. The CIM was therefore identified and coefficient estimates (β_{qj}) are summarised below (standard errors are reported in parentheses).

Table 3.7: Fixed effects estimate (beta - random intercept).

T	(Intercept)	EDUC_medium	EDUC_high	FL_medium	FL_high
0	-2.195900 *** (0.377)	0.2421075 (0.169)	0.6622724 *** (0.186)	0.5568624 *** (0.197)	1.189597 *** (0.202)
1	-1.296494 *** (0.377)	0.2421075 (0.169)	0.6622724 *** (0.186)	0.5568624 *** (0.197)	1.189597 *** (0.202)

Note: *p-value < 0.1; **p-value < 0.05; ***p- value < 0.01

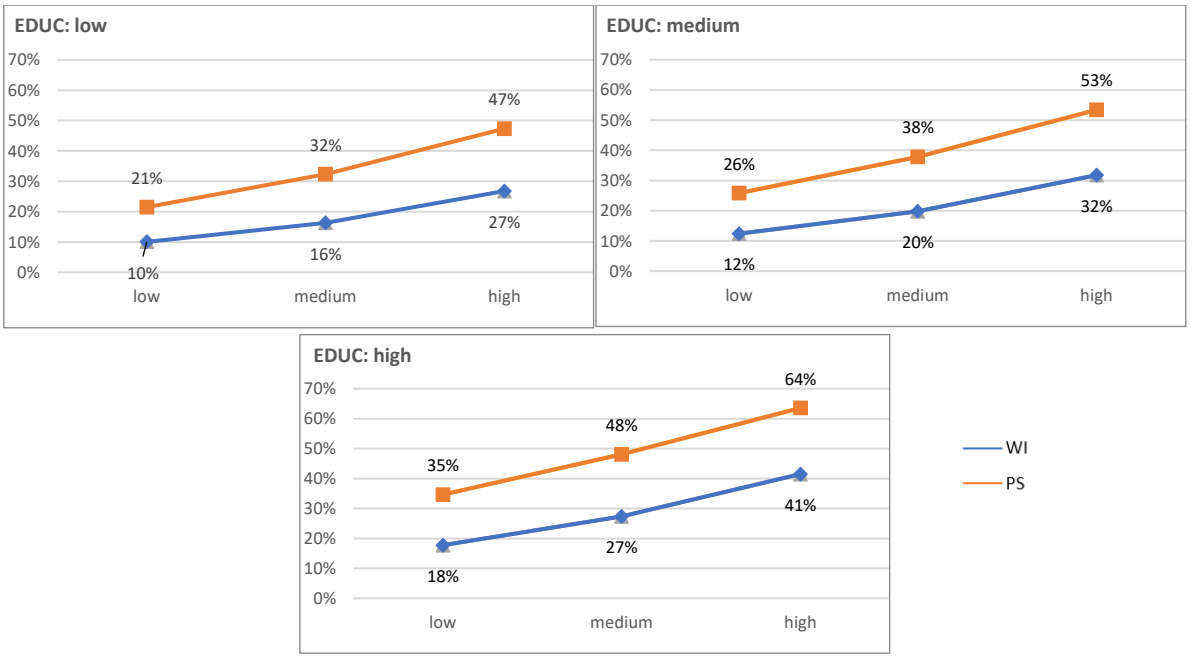
After transforming β_{qj} into probabilities, the results were as follows:

Table 3.8:Fixed effects estimate (likelihood - random intercept).

T	(Intercept)	EDUC_medium	EDUC_high	FL_medium	FL_high
0	10.01%	56.02%	65.98%	63.57%	76.67%
1	21.48%	56.02%	65.98%	63.57%	76.67%

From Table 3.8: *Fixed effects estimate (likelihood - random intercept)*, we see that the higher the level of education or the financial literacy score, the higher is the probability of making a correct risk assessment. For example, participants with a high education level assessed riskiness correctly in about 66% more cases than individuals with a low education level. This percentage was equal to 77% when comparing individuals with low and high financial literacy. Figure 3.6: *CIM-predicted probabilities by FL and stratified by education level* reports the predicted probabilities of making the correct risk assessment using the parameter estimates obtained in the CIM. In particular, the results show the relation between FL and the dependent variable by using EDUC as a stratification variable. Firstly, the Probabilistic representation allows potential investors to make the correct risk assessment of the two products more often than the What-if scheme. In particular, participants with low skills (i.e., FL_low, EDUC_low, or both) seem more advantaged by the Probabilistic scenarios tables, which constantly encourage participants towards more appropriate answers. Considering the What-if treatment, the increase in correct risk assessment moves smoothly from individuals with a low financial literacy score to ones with a high financial literacy score. This is not the case under the Probabilistic representation: a higher level of financial literacy significantly increases the chance of making the correct risk assessment.

Figure 3.6: CIM-predicted probabilities by FL and stratified by education level.



Because financial literacy score and education level were expected to impact the dependent variable differently, random slope models were performed in order to consider the effect of possible variation between treatments of individuals' characteristics. Starting from the CIM, only the education level and the financial literacy score were considered as level 1 variables (detailed estimates are in Table 4.23: *Fixed effects estimates of random intercept and random slope models* in Annex I). M14 uses the financial literacy score as a random slope variable while M15 accounts for possible variation between education level groups. The LRT comparing the CIM and the random slope model M14 was statistically significant (p-value < 0.05) which confirms that the level of financial literacy affects the risk assessment differently between two treatments. Therefore, M14 was relabelled as the AIM. The values below report the estimated marginal likelihood of making the correct risk assessment (the standard errors are reported in parentheses).

Table 3.9:Fixed effects estimate (likelihood - FL random slopes).

T	(Intercept)	EDUC_medium	EDUC_high	FL_medium	FL_high
0	11.38% *** (0.310)	55.88% (0.170)	65.77% *** (0.187)	65.46% *** (0.207)	67.77% *** (0.341)
1	20.10% *** (0.310)	55.88% (0.170)	65.77% *** (0.187)	61.63% *** (0.207)	81.66% *** (0.341)

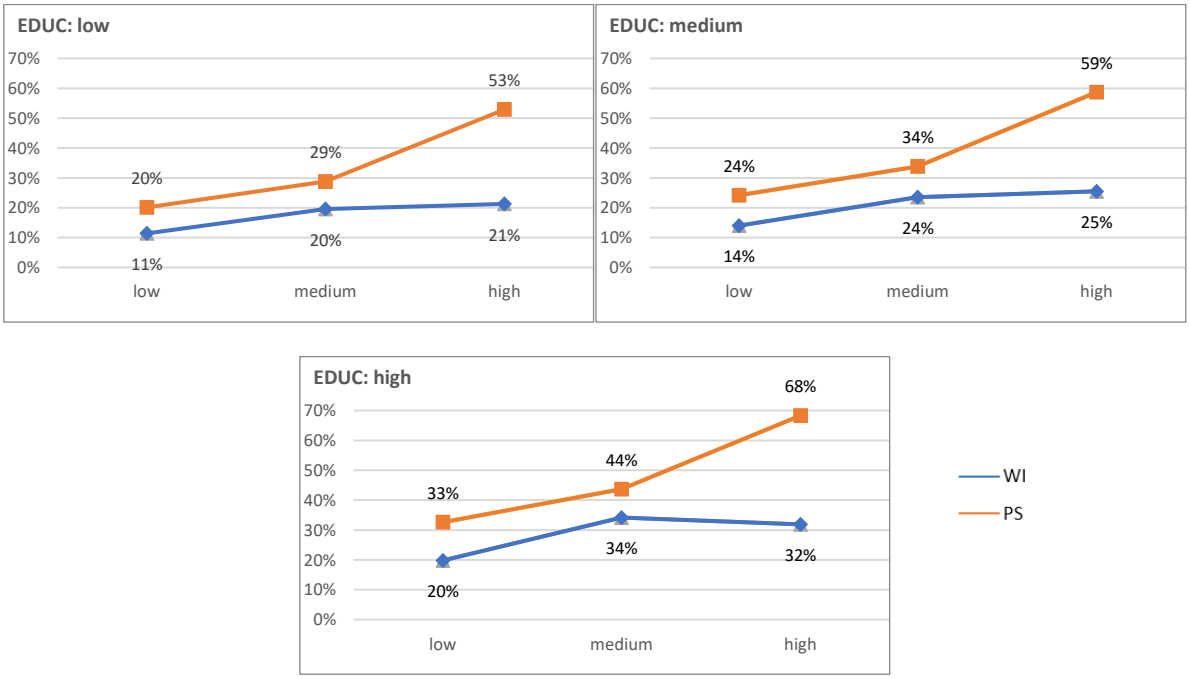
Note: *p-value < 0.1; **p-value < 0.05; ***p- value < 0.01

Financial literacy is a fundamental factor which acts differently in relation to the two treatments. In general, individuals with higher financial literacy have more chance of making the correct risk assessment. However, respondents with a higher level of financial literacy evaluated Product I and Product II riskiness correctly more frequently under the PS scheme than the WI one: individuals with high financial literacy have about 81% more chance than individuals with low financial literacy when Probabilistic representation is used; this relation becomes only 67% under the What-if treatment. This phenomenon suggests that: (i) the marginal contribution of higher financial literacy level is strictly positive under the Probabilistic scheme; and (ii) because the chance of getting the correct answer is not significantly different between FL_medium and FL_high with respect to FL_low, increasing financial literacy level is not the correct tool to use to potentially

debias investment decisions when What-if representations are used. As for the last point, indeed, even participants with high financial literacy score do not always make the correct risk assessment of financial products. This highlights the complexity of the What-if representation and its unreadable format. All these statements are confirmed by Figure 3.7: *AIM-predicted probabilities by FL and stratified by education level* below which shows the marginal contribution of higher level of financial literacy to the chance of achieving the correct risk assessment by using education as a stratification variable.

The chance of making the correct risk assessment always increases moving from individuals with low financial literacy to ones with a medium score. The benefit of a further increase of financial literacy score (i.e., moving from medium to high) is not consistent between groups. Indeed, the relationship between financial literacy score and the likelihood of getting the correct risk assessment seems to be always positive under the Probabilistic scheme; whereas under the What-if representation this relationship becomes almost constant at the higher level of financial literacy scores.

Figure 3.7: AIM-predicted probabilities by FL and stratified by education level.



3.2.6 *A further analysis to investigate relationship with stochastic dominance*

The consumer testing exercise was built considering two financial products where one of them always stochastically dominates (second order) the other, therefore representing the best investment choice for risk-averse expected-utility-maximiser individuals. As shown in the previous paragraphs, the treatment directly impacts the risk assessment which influences the final investment decision. Therefore, a multilevel logistic model was used to investigate the following question:

Q3: Are there any individual characteristics that affect the investment choice?

In order to answer to this research question, an analysis was performed using a reduced sample of 474 individuals after excluding: (i) all the participants who wrongly evaluated Product I as being as risky as Product II; (ii) all the participants who did not express any preference between Product I and Product II (i.e., answered 'I do not know' to Question C.1); (iii) risk lovers (i.e., after a correct risk assessment, the individual chose Product I); and (iv) potential random investment choice. Table 3.10: *Sub-sample distribution* reports the

distribution of the sample in terms of frequencies and clearly shows a path between treatment, risk assessment, and investment decision. In particular, 219 respondents (i.e., 147 + 72) represent the negative effect of the frames on the investment decision, while the remaining 255 participants (i.e., 57 + 198) show the positive influence of the treatment on risk averse investment choice.

Table 3.10: Sub-sample distribution.

Frequency		**Choice**			
		wrong		correct	
		WI	PS	WI	PS
Riskier					
	wrong	147	72	-	-
	correct	-	-	57	198

The output variable of the multilevel model is represented by the answer given to Question C.1 and it was labelled as *choice.* In particular, *Y* is described for each individual *i* as follows:

$$Y_i = \begin{cases} 1, & \text{if Product II is selected} \\ 0, & \text{otherwise} \end{cases}$$

The multilevel logit model with non-nested factors was built where β_{0j} represents the treatment variable and β_{0k} is for the variable 'risker' specified above.

$$\text{level 1: } logit(Y_{ijk}) = logit\left(\frac{\pi_{ijk}}{1-\pi_{ijk}}\right) = \beta_{0j} + \beta_{0k} + \sum_{q=1}^{Q} \beta_q X_{qijk} \quad (3)$$

$$\text{with } \beta_{0j} = \gamma_{00} + u_{0j} \quad (4)$$

$$\text{and } \beta_{0k} = \gamma_{10} + u_{0k} \quad (5)$$

For the sake of clarity in these expressions, the level 1 equation (3) represents the individual-level model while equations (4) and (5) show the non-nested group factors. For the *i*-th individual in the *j*-th and *k*-th groups, Y_{ijk} represents the odds for the success category (y_{ijk} = 1) based on q = 1, …, Q individual-level explanatory variables. Similar to Section 3.2.5, $\sum_{q=1}^{Q} \beta_q X_{qijk}$ are the regression coefficients of the multilevel logistic model. The intercept for the *j*-th group is represented by β_{0j} while β_{0k} is the intercept for the *k*-th group. The level 2 equations describe how the within-group effects may vary according to group-level characteristics. In particular, the intercept is the level 2 variable which is composed of the estimated values (γ_{00} and γ_{10}) and the terms u_{0j} and u_{0k} which are the group-specific random effects that also represent the level 2 errors. Finally, it should be noted that specific group variables were not considered in the experiment.

The procedure used to obtain the models is the same as above. Although personal traits in general affect investment behaviour, the estimation results of multilevel logistic models showed no relationship between individual characteristics (such as sociodemographic, investment habits, and financial literacy level) and investment decision coherent with the stochastic dominance criteria (detailed estimates are reported in Table 4.24: *Analysis of the relation with stochastic dominance: fixed effects estimation* in Annex I). However, the unconditional mean model – i.e., an empty model containing only the intercept – was statistically significant. This evidence implies that how information is framed represents the main factor that influences the decision-making process.

3.3 Analysis of the decision-making process from a physiological perspective[70]

The excessive complexity of financial instruments and the format in which they are represented often leads investors to make their decisions in a stressful environment and under time pressure since stakes are high and the consequences of wrong choices can be particularly critical. Recent behavioral evidence suggests that stress potentiates decision biases because heuristics may dominate over deliberate processes in decision making under uncertainty (Yu, 2016). Stress is thought to trigger riskier decisions (Pabst *et al.*, 2013) but the way the object of the decision is presented is pivotal. In fact, stress enhances a preference for safer decisions in gain-domain frames and enhances a preference for risky decisions in loss-domain ones (Porcelli and Delgado, 2009; for a review see Porcelli and Delgado, 2017). Generally, stress is associated to the employment of dysfunctional strategies (Baradell and Klein, 1993), lowered self-control (Maier *et al.*, 2015), altered feedback processing and enhanced reward sensitivity (Gray, 1999; for a review see Starcke and Brand, 2012).

Scientific evidence suggests that bodily responses (i.e., electrodermal activity[71], heart rate, respiration rate, and face temperature) may be employed to reveal implicit features of decision processes and to monitor decisors' emotional states during decision making (Tieri *et al.*, 2015; Ponsi *et al.*, 2019). In particular, electrodermal activity is an index widely employed in decision theory (Dawson *et al.*, 2011; Palomaki *et al.*, 2013) and economic sciences such as consumer and organizational research (Christopoulos *et al.*, 2016). In scientific literature a large set of empirical facts about the influence of cognitive processes on financial decision making is found (for an exhaustive survey see Frydman and Camerer, 2016). These findings mainly relate with departures from rationality occurring at all levels in the economy, from households to CEOs. What is less investigated from a cognitive science perspective is the psychology that might generate the observed patterns of investing, particularly in relation to framing effects. One of the goals of this experiment is precisely to fill this gap in the literature.

Given the general framework described above, a laboratory experiment was conducted to investigate the effects of the What-if and the Probabilistic scenarios frames. In particular, the impact of the two methodologies on investors' sympathetic nervous system[72] (SNS) activity was measured. The experiment aimed to answer the following questions:

Q4: How do behavioural and physiological indexes react in the investors' decision-making process when the investor must choose on the basis of two different schemes of representing information?

Q5: Are the effects of time pressure on the choice relevant when the two schemes of representing information are presented?

[70] This Section is an abstract from the article 'What if versus Probabilistic scenarios: a neuroscientific analysis' written with Rosella Castellano, Georgia Ponsi and Gaetano Tieri. The manuscript was submitted to the Journal 'Annals of Operations Research' on 15th November 2018. In order to be published, the reviewers asked to clarify some parts of the article. It was resubmitted on 3rd April 2019 and published on 23th May 2019 (se Castellano *et al.*, 2019).

[71] Electrodermal activity is the property of the human body that causes continuous variation in the electrical characteristics of the skin. The traditional theory of EDA holds that skin resistance varies with the state of sweat glands in the skin. Sweating is controlled by the sympathetic nervous system and skin conductance is an indication of psychological or physiological stimulation. If the sympathetic branch of the autonomic nervous system is highly stimulated, then sweat gland activity also increases, which, in turn, increases skin conductance. In this way, skin conductance can be a measure of emotional and sympathetic responses.

[72] Sympathetic nervous system is the part of the autonomic nervous system that is concerned especially with preparing the body to react to situations of stress or emergency. Specifically, when the entire SNS is activated, there is a cascade of reactions from all the organ systems of the body, which prepare the individual to deal with an emergency. This includes an increase in heart rate, bronchial dilation, increase in cardiac output, and dilation of pupils, all of which are directed towards heightened awareness and preparation to combat danger. Blood circulation is preferentially targeted towards skeletal muscle, with a reduction in blood flow towards non-essential organs. Therefore, there is vasoconstriction in the gastrointestinal tract and skin, and compensatory piloerection to allow the body to remain warm. While the physical danger that you have to either fight or escape from may be short-lived, the SNS can also be activated in response to long-term psychological or emotional stress.

As far as we know, this is the first scientific work in which the two above mentioned types of investment framings are subjected to an analysis of this type.

3.3.1 The experiment: tools, methods and stimuli

Forty-one volunteers took part in the study (19 male; age mean ± SD, 31.805 ± 9.73, socio-demographic characteristics are resumed in Table 5.1: *Attributes distribution* in Annex II) conducted at the Social and Cognitive Neuroscience Laboratory, Department of Psychology, Sapienza University of Rome. All the participants were naive to the purposes of the experiment and gave their written informed consent to take part in the study. Participants were comfortably seated in front of a rectangular table (120 × 75 cm) and watched a full HD monitor placed on the table at a distance of ≈ 60 cm from their eyes. The experiment was implemented using homemade software developed in Unity 5. It consisted of two different two-choice tasks where the participants were required to make financial decisions by using a mouse with their right hand. The first task (Task 1) consisted of 22 trials (11 WI frames and 11 PS ones, respectively) presented on the PC monitor in a pseudo-randomized order across participants. For each trial, two different financial products were displayed using the same representation scheme (WI or PS frame). It is important to highlight here that pairs of investment products were chosen such that one out of two financial products always stochastically dominated (second order) the other, thus representing the financial asset that should be selected by risk-averse expected-utility-maximiser individuals – those characterized by increasing and concave utility functions. Participants were asked to make their choices by taking all the time they needed (Figure 5.1: *Example of Task 1* in Annex II shows an example of a trial of Task 1). After each choice, the trial disappeared and the next trial started automatically after a pause of 8 sec during which a white screen was displayed. The second task (Task 2) was a perceptual one. It consisted of the presentation of 22 trials in which the characteristics of a single financial product were displayed by means of the two different frames – i.e., WI and PS schemes. It is worth remarking here that participants were asked to choose the preferred scheme as quickly as possible (Figure 5.2: *Example of Task 2* in Annex II shows an example of a trial of Task 2). No right or wrong responses were involved in this task. This strategy allowed us to investigate participants' preferences for What-if or Probabilistic scenarios frames.

In both tasks and in each trial, participants made their decision by using a PC mouse and clicking on one of three grey buttons placed in the lower part of the screen (left and right circle button meaning left or right choice, rectangle button meaning 'I do not know'). For each decision, the software recorded: (i) the reaction time (RT) – i.e. the time occurring between the presentation of the stimulus (the trial) and the decision; (ii) the accuracy (ACC), only in Task 1 – i.e., the correctness of the decision; and (iii) the frequency of choice, only in Task 2 – i.e., the number of times in which WI or PS were selected. During Task 1, participants' electrodermal activity (EDA) was acquired (see Section 3.3.2: *Electrodermal activity* for further details). Before starting the experiment, participants were asked to complete a questionnaire about financial education (the same as in the aforementioned consumer testing) in order to understand whether individuals were accustomed to basic financial concepts such as compound interest, inflation, and diversification. A financial literacy score was calculated as described. The questionnaire aimed also to collect sociodemographic characteristics and investment habits of the candidates. Participants were asked to read the instructions of the experimental protocol, while the experimenter mounted the EDA electrodes on their hands.

To implement the 11 trials for each of the two analysed frames, 22 financial instruments were identified. In particular, subordinated bonds with seven years' maturity and annual fixed coupons were selected. All the securities were characterized by different contractual characteristics: i.e., issuer rating, coupon rate, and repayment of principal. Each bond was priced using standard financial models. The What-if and Probabilistic

tables were obtained by using the same approach as described in Chapter 2[73]. Finally, the pairs of securities to be included in the first trial were identified by applying second order stochastic dominance criteria, to obtain an ordering of preferences for a rational risk-averse investor (Castellano and Cerqueti, 2013, 2016).

3.3.2 Electrodermal activity.

To measure the skin's electrodermal activity during Task 1, AD-Instruments PowerLab 8/35 and ML116 GSR (galvanic skin response[74]) amplifier (providing a 75Hz AC excitation and automatic zeroing) devices were used to amplify the signals. Specific GSR sensors consisting of two bipolar finger electrodes were placed on the right hand to record EDA. The signal was sampled at 1KHz and recorded using LabChart 7 (AD-Instruments Inc.) software. The trigger signals were sent to the AD-Instruments amplifier at stimulus onset, participant response, and stimulus offset by means of a TriggerStation (BrainTrends Ltd., Italy). One participant was excluded from the sample because of a technical problem that occurred during the electrodermal activity recording. The EDA analysis was thus performed on a sample of 40 participants. EDA data pre-processing and analysis were carried out using the MATLAB-based toolbox Ledalab V3.4.9 (Leipzig, Germany – www.ledalab.de). Data were down-sampled at 10Hz (Lajante *et al.*, 2012) and then analysed by the method of continuous decomposition analysis (CDA; Benedek and Kaernbach, 2010), which allows the extraction of stimulus-related sympathetic activity by decomposing the EDA signal into separate phasic (fast) and tonic (slow) components. To determine skin conductance responses[75] (SCR), a threshold criterion of 0.01 μS was chosen. Among many indexes, CDA allows us to compute a stimulus-locked integrated SCR (ISCR, SCR multiplied by the size of response window; mean μS x second). In particular, ISCR represents the time integral of the phasic driver within the response window. Because ISCR allows the integration of both spatial and temporal dimensions of SCR, it is considered an unbiased and time-sensitive measure of sympathetic activity in response to each stimulus. Individual ISCR data were normalized through *z* scores. In order to derive information about changes from the baseline, ISCRs were extracted for both the 5 sec. window starting 1 sec. after the stimulus onset (post-stimulus ISCR) and the 0.3 sec window (as in Tieri *et al.*, 2015) before stimulus onset (pre-stimulus ISCR). For each trial, ISCR delta scores (ΔISCR) were computed by subtracting pre-stimulus ISCR (our baseline) from post-stimulus ISCR.

3.3.3 Results

In this section, the results of the laboratory experiment composed of the two-choice tasks are presented.

[73] The parameters were calibrated using market data. In particular, the WI tables were obtained by changing the main risk factors characterizing the selected financial instruments – that is, the issuer's default date and the recovery rate. Under the WI frame, the expected return of the bond was calculated using the deterministic internal rate of return. The construction of the tables grounded on the PS frame considered two risk factors: the interest rate and issuer's default risk. To model the interest rate dynamics, the Hull and White single factor model was implemented. The main parameters of the model were estimated by minimizing the distance between observed market prices and theoretical ones. The latter were computed on a trinomial tree. The issuer's default risk was estimated by using the binomial tree model proposed by Jarrow and Turnbull while the default probabilities were extracted from CDS market quotes for different maturities. To apply the superimposition technique, a 7-year sovereign bond issued by Germany was selected as a RFA security.

[74] The galvanic skin response (GSR, which falls under the umbrella term of electrodermal activity, or EDA) refers to changes in sweat gland activity that are reflective of the intensity of individual emotional state, otherwise known as emotional arousal. Individual level of emotional arousal changes in response to the environment individuals are in – if something is scary, threatening, joyful, or otherwise emotionally relevant, then the subsequent change in emotional response that is experienced also increases eccrine sweat gland activity. Research has shown how this is linked to emotional arousal. It is noteworthy that both positive ('happy' or 'joyful') and negative ('threatening' or 'saddening') stimuli can result in an increase in arousal – and in an increase in skin conductance. The GSR signal is therefore not representative of the type of emotion, but the intensity of it.

[75] The skin conductance response, also known as the electrodermal response (and in older terminology as the 'galvanic skin response'), is the phenomenon that occurs when the skin momentarily becomes a better conductor of electricity when either external or internal stimuli occur that are physiologically arousing. Arousal is a broad term referring to overall activation and is widely considered to be one of the two main dimensions of an emotional response. Measuring arousal is therefore not the same as measuring emotion, but is an important component of it. Arousal has been found to be a strong predictor of attention and memory.

3.3.3.1 Task 1: behavioural responses - reaction times and accuracy

During Task 1 – the one in which participants were asked to make financial decisions under the PS and WI frames – mean reaction times (RT) and accuracy (ACC) were calculated for each trial. The collected data were analysed by using a two-tailed paired sample t-test. Results show that participants are faster at making investment decisions under the PS scheme (Mean ± SE, 23.72 ± 2.07 sec.) compared with WI frames (28.87 ± 2.3 sec.): $t(40) = 5.53$, p-value < 0.001.

With regard to ACC, it can be observed that participants took more accurate financial decisions in PS trials (67 ± 0.02%) compared with WI ones (33 ± 0.02%): $t(40) = -10.09$, p-value < 0.001. These results suggest that the adoption of a PS frame resulted in participants making more accurate and faster decisions with respect to the WI representation.

In order to provide more robust results, a multilevel mixed log-linear regression analysis was performed (Panasiti *et al.*, 2018; Ponsi *et al.*, 2017a and 2017b). The applied models belong to the family of linear mixed models (LMM) which were estimated by using the lme4 package (Bates *et al.*, 2015). Indeed, LMM are suitable for: (i) analysing hierarchical data structures (i.e., in which not all levels of a categorical factor co-occur in all levels of another categorical factor); (ii) analysing the entire data set (not just the mean observations for each participant and condition) to better evaluate the variations of data usually left out in analysis-of-variance-style (ANOVA) analyses; (iii) accounting for the non-independence of observations with correlated error; (iv) separately treating the effects caused by the experimental manipulation (fixed effects) and the ones that were not (random effects; Pinheiro and Bates, 2000). The log-likelihood ratio statistics were asymptotically approximated to a χ^2 distribution to compute the p-value (Pinheiro and Bates, 2000).

Because the financial decision Reaction Times (RT) represent a continuous variable, a linear mixed model was implemented (as in Liuzza *et al.*, 2017) where RT was the dependent variable (y_{ij}). In particular, RT_{ij} was specified for each *i*-th observation and the *j*-th grouping variable was represented by the individual, meaning that participants were entered into the model as a random factor, while the fixed effects of frame (*T*) were modelled as random slope over participant (Barr *et al.*, 2013; Ponsi *et al.*, 2016). Using a random intercept–random slopes approach, the following model was estimated:

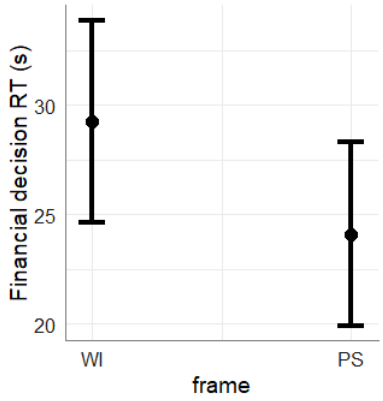
$$\text{level 1: } y_{ij} = \beta_{0j} + \beta_{1j}T_{ij} + \sum_{q=2}^{Q}\beta_q X_{qij} + e_{ij} \quad (6)$$

$$\text{with } \beta_{0j} = \gamma_{00} + u_{0j} \quad (7)$$

$$\text{and } \beta_{1j} = \gamma_{01} + u_{1j} \quad (8)$$

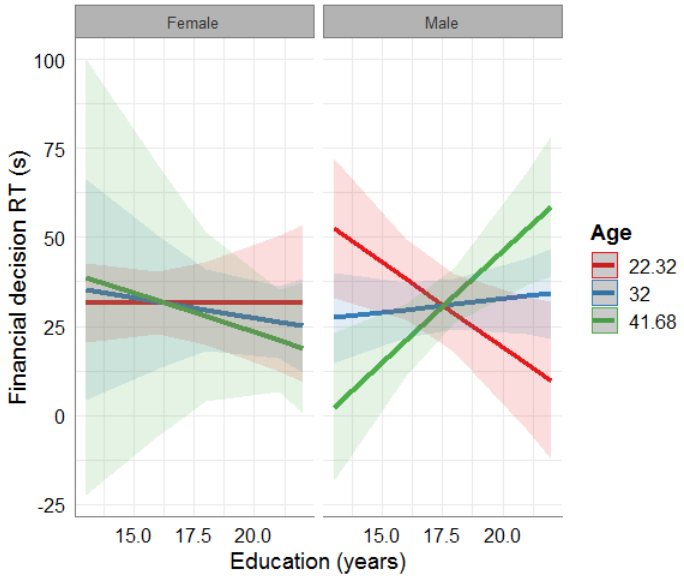
For the sake of clarity, equation (6) represents the level 1 model while equations (7) and (8) show the random intercept and the random slope, respectively. The random slope is only for the framing ($\beta_{1j}T_{ij}$) while the aggregate $\sum_{q=2}^{Q}\beta_q X_{qij}$ shows the regression coefficients of the multilevel model where $q = 2, \ldots, Q$ are the individual-level explanatory variables, excluding *T*. The intercept for the *j*-th group is represented by β_{0j}, while β_{0j} is the random slope. Finally, it should be noted that specific group variables were not considered in the experiment. The two frames were firstly considered as level 1 categorical predictors and then all other variables (e.g., age, gender, and financial literacy score) were added. According to this model, the PS scheme allows participants to make faster choices with respect to WI ($\chi^2 = 20.11$; p-value < 0.001). This result is reported in Figure 3.8: *Predicted values of reaction time* which shows the predicted values of reaction time (measured in seconds) when considering the scheme as main effect and the error bars indicate 95% confidence intervals.

Figure 3.8: Predicted values of reaction time.



Then, additional variables were added as continuous (age, education) or categorical (gender, financial literacy) predictors to the multilevel mixed linear regression analysis (see Table 5.2: *Fixed effects estimates on Reaction time (Task 1)* in Annex II). We found a main effect of gender (χ^2 = 5.68, p-value < 0.05), a two-way interaction between age and gender (χ^2 = 4.84, p-value < 0.05), and a two-way interaction between age and gender (χ^2 = 4.84, p-value < 0.05), and gender and education (χ^2 = 6.27, p-value < 0.05). The main effect and the two-way interactions were qualified by a three-way interaction between age, gender, and education (χ^2 = 5.87, p-value < 0.05). Then, we analysed the triple interaction and tested its simple effects in a regression model (including also the random effects; as in Ponsi *et al.*, 2017a). The simple slopes for the association between education and reaction times were tested for low (-1 SD below the mean, < 22.32 years), mean (32 years), and high (+1 SD above the mean, > 41.68 years) levels of the continuous variable (age) and for the two levels of the categorical variable (gender). The analysis shows that in the case of older (+1 SD above the mean, > 41.68 years) male participants, there is a significant relationship between education and reaction times (*b* = 12.675, SE = 4.59, *t* = 2.76, upper CI = 21.68, lower CI = 3.675, see Figure 3.9). The same relationship is also significant in the case of younger (-1 SD below the mean, < 22.32 years) male participants (*b* = -13.90, SE = 4.47, *t* = -3.09, upper CI = -5.035, lower CI = -22.56). The other combinations of age mean/male (*b* = -0.56, SE = 2.68, *t* = -0.21, upper CI = 4.70, lower CI = -5.82), age -1 SD /female (*b* = -2.44, SE = 3.39, *t* = -0.72, upper CI = 4.19, lower CI = -9.08), age mean/female (*b* = -3.83, SE = 4.68, *t* = -0.82, upper CI = 5.34, lower CI = -12.99), and age +1 SD /female (*b* = -5.21, SE = 8.75, *t* = -0.595, upper CI = 11.94, lower CI = -22.36) do not entail significant effects. In particular, the three-way interaction shows that the more older male participants are educated, the more they need time in order to take a financial decision (see Figure 3.9: *Predicted values of financial decision reaction time*, right-sided panel, green slope). The exact opposite behavioural pattern is true for younger male participants: the more they are educated, the more they make fast financial decisions (see Figure 3.9: *Predicted values of financial decision reaction time,* right-sided panel, red slope).

Figure 3.9: Predicted values of financial decision reaction time.



Note: Shaded bands indicate 95% confidence intervals. RT(s) = reaction time measured in seconds.

In reference to the financial decisions' accuracy, the same model was run. However, it is worth highlighting here that, in the experiment, ACC represented a binomial variable (Y_{ij}) specified for each *i*-th observation and *j*-th participant:

$$Y_{ij} = \begin{cases} 1, & \text{if choice is coherent with SD criteria} \\ 0, & \text{otherwise} \end{cases}$$

Hence, ACC was treated as the dependent variable of a random intercept random slope multilevel logistic model. As well as the RT, participants were entered as random factor and fixed effect of frame was modelled as random slope over participants. In addition, individual variables were considered as fixed effects. The model was specified as:

$$\text{level 1: } logit(Y_{ij}) = logit\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \beta_{1j}T_{ij} + \sum_{q=1}^{Q}\beta_q X_{qij} \quad (9)$$

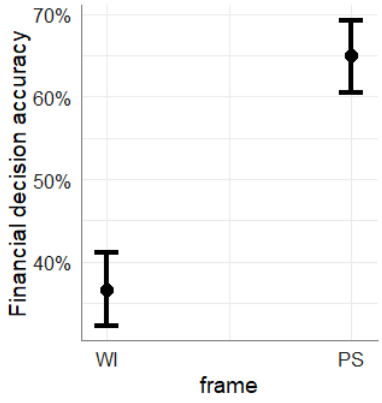
$$\text{with } \beta_{0j} = \gamma_{00} + u_{0j} \quad (10)$$

$$\text{and } \beta_{1j} = \gamma_{01} + u_{1j} \quad (11)$$

The only difference in the model specified for RT (equations (6)–(8)) is equation (9), which represents the relation between the natural log of the odds for the logit distribution and the independent variables. Results show that a main effect on ACC is observable under the PS frame (χ^2 = 70.64; p-value < 0.001), implying that the PS information scheme allows participants to be more accurate with respect to the WI one (see Figure 3.10: *Predicted values of financial decision accuracy* which shows the predicted values of financial decision accuracy when considering the main effect of scenario and the error bars indicate 95% confidence intervals).

Additional variables were added as continuous (age, education) or categorical (gender, financial literacy) predictors to the multilevel mixed log-linear regression analysis (see Table 5.3: *Fixed effects estimates on Accuracy (Task 1)* in Annex II). We can see that males have a higher propensity to make more accurate decisions and having a higher financial literacy score improves the accuracy level.

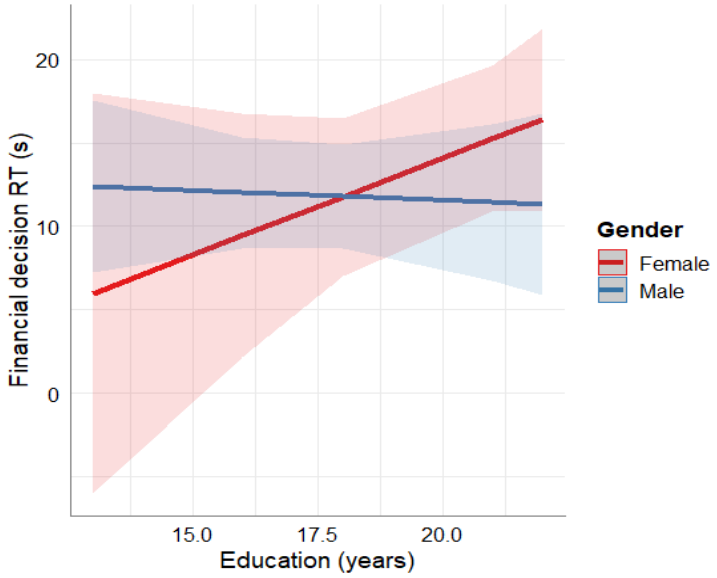
Figure 3.10: Predicted values of financial decision accuracy.



3.3.3.2 Task 2: behavioural responses - reaction times and frequencies of choices

During Task 2, the mean frequency of choices (FC) – i.e., the number of times participants chose WI or PS schemes – and the associated RT, were calculated for each trial and analysed using a two-tailed paired sample t-test. The analysis showed that participants chose the PS representation significantly more often (67 ± 0.02%) compared with the WI one (33 ± 0.02%): $t(40) = -10.09$, p-value < 0.001. The one-sample proportions (one-sided) test confirmed this result ($\chi^2 = 2.91$, p-value < 0.05). Looking at the RT, no significant difference was found when comparing the RT related to WI and PS schemes, 11.25 sec and 10.24 sec respectively: $t(31) = -1.39$; p-value = 0.17. Following the linear mixed model described by equations (6)–(8) (as in Liuzza et al., 2017) in which the continuous variable RT is the dependent variable and the choice between WI and PS is the categorical predictor, no main effect was found ($\chi^2 = 2.78$, p-value > 0.05), showing that participants take the same amount of time in choosing the preferred scheme between PS and WI. Then, the additional variables were added as continuous (age, education) or categorical (gender, financial literacy) predictors to the multilevel mixed linear regression (see Table 5.4: Fixed effects estimates on Reaction time (Task 2) in Annex II). We found a main effect of gender ($\chi^2 = 4.91$, p-value = 0.03) which was qualified by a two-way interaction between gender and education ($\chi^2 = 5.39$, p-value = 0.02). We focused on the double interaction and we also tested its simple effects in a regression model (including also the random effects). The simple slopes for the association between education and reaction times were tested for the two levels of the categorical variable (gender). The analysis shows that in the case of female participants there is a significant relationship between education and reaction times (b = 0.77, SE = 0.37, t = 33.27, upper CI = 1.50, lower CI = 0.037, see Figure 3.11: *Predicted values of choice reaction time when considering the two-way interaction*). The same relationship is not significant in the case of male participants (b = -0.32, SE = 0.46, t = 33.23, upper CI = 0.59, lower CI = -1.23). In particular, the two-way interaction shows that the more female participants are educated, the more they need time in order to make a choice between What-if and Probabilistic scenarios (see Figure 3.11: *Predicted values of choice reaction time when considering the two-way interaction*, red slope). This effect of education on reaction times is not present for male participants (see Figure 3.11: *Predicted values of choice reaction time when considering the two-way interaction*, blue slope).

Figure 3.11: Predicted values of choice reaction time when considering the two-way interaction.



Note: Shaded bands indicate 95% confidence intervals. RT(s) = reaction time measured in seconds.

A multilevel analysis was performed in order to check the role of age, gender, education and financial literacy in explaining the choice of financial frame (see Table 5.5: *Fixed effects estimates on Decision (Task 2)* in Annex II). It is worth highlighting here that, in the experiment, the variable 'decision' represented a binomial variable (y_{ij}) specified for each *i*-th observation and *j*-th participant:

$$Y_{ij} = \begin{cases} 1, & \text{if participant prefers PS} \\ 0, & \text{otherwise} \end{cases}$$

Hence, decision was treated as the dependent variable of a random intercept random slope multilevel logistic model where participants were entered as random factor and individual variables were considered as fixed effects. The model was specified as:

$$\text{level 1: } logit(Y_{ij}) = logit\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \sum_{q=1}^{Q} \beta_q X_{qij} \quad (12)$$

$$\text{with } \beta_{0j} = \gamma_{00} + u_{0j} \quad (13)$$

The analysis shows that more educated participants (p-value < 0.05) or with a higher level of financial literacy (p-value < 0.01) tend to prefer the Probabilistic scenarios scheme than the What-if ones.

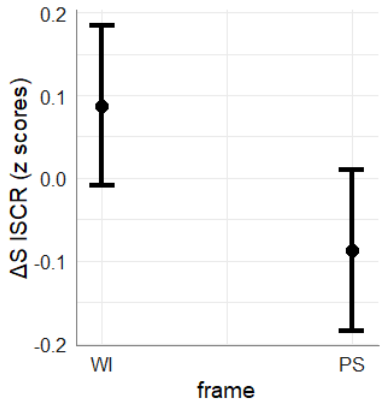
4.3 Electrodermal activity

During Task 1, the ISCR delta scores (ΔISCR) were calculated for each trial under the influence of both frames: WI and PS. The collected data were analysed using a two-tailed paired sample t-test. The analysis of ΔISCR shows that the electrodermal activity elicited by the PS scheme was significantly lower (Mean ± SE, -0.09 ± 0.21 sec.) than the one elicited by the WI scheme (0.09 ± 0.21 sec.): $t(39) = 2.65$, p-value < 0.01. This finding suggests that financial decisions made under the PS representations induce lower sympathetic activity in the participants than under WI.

To confirm this result, because the ISCR delta score (ΔISCR) represents a continuous variable, a linear mixed model was implemented (same structure as equations (6)–(8) above). In the model, the ISCR delta scores was the dependent variable and the scheme (WI and PS) were the categorical predictor. As before, participants were entered as random factor and the fixed effects of the two schemes were modelled as random slope over participant.

A main effect of the scheme was found ($\chi^2 = 6.27$; p-value <0.05), confirming that financial decisions under the PS scheme induce lower sympathetic activity (see Figure 3.12: *Predicted values of integrated skin conductance response delta scores*, which shows the predicted values of integrated skin conductance response delta scores when considering the main effect of the frame and the error bars indicate 95% confidence intervals). No main or interaction effects involving age, gender, as well as education and financial literacy level, were found (more details are in Table 5.6: *Fixed effects estimates on electrodermal activity (Task 1)* in Annex II).

Figure 3.12: Predicted values of integrated skin conductance response delta scores.



3.4 Considerations

Many experiments have been conducted to test the effectiveness of representation format used to disclose financial information. Based on some simplifications of the investment decision context, this research analysed the impact of the What-if and Probabilistic scenarios schemes on the investment decision process by a consumer test exercise and a laboratory experiment.

The consumer testing exercise shows that Probabilistic tables obtained a higher grade of preference (57%) than the What-if representation (18%). Indeed, participants' appraisal of the information provided by the Probabilistic scheme was on average higher for all the considered attributes (i.e., usefulness, clearness, sufficiency, and comparability) with some differences among sociodemographic characteristics and investment habits. Moreover, the Probabilistic format increases the chance of a correct risk assessment (43% of the participants) and risk-averse efficient investment decision than the What-if frame (23% of the participants). In particular, financial literacy and education level of investors may reduce the impact of framing effects further, leading to unbiased risk perception. This evidence provides useful insights into how financial knowledge could be strengthened in order to improve the decision-making process. In particular, a high level of financial literacy allows investors to correctly evaluate the risk of financial investment when Probabilistic tables are used. This is might not be the case for the What-if representation.

Most of these results are confirmed by the physiological laboratory test. During the financial decision task (Task 1) participants took more accurate financial decisions when securities were presented via the

Probabilistic scheme, which also allowed participants to make more rapid decisions. This finding suggests that Probabilistic scenarios might help risk-averse investors to make fair financial choices with respect to second order stochastic dominance. In addition, the integrated skin conductance response data suggest that Probabilistic representation induced less sympathetic nervous system activation during the decision-making process. Importantly, this differential sympathetic nervous system activation took place in a rapid way, because it was already evident 6 seconds after the presentation of visual stimuli. These data corroborate the idea that the What-if scheme induces a greater degree of stress response compared with the Probabilistic one. As a general conclusion of the analysis, it is possible to state that the Probabilistic scenarios scheme with respect to What-if: (i) enhances financial decision accuracy, (ii) decreases financial decision reaction times, and (iii) induces less stress response. In regard to the perceptual decision task (Task 2), Probabilistic representation was largely preferred by participants (67%). However, it should be noted that the trials in which participants preferred the What-if frame were trials in which the probability of a negative return was greater than 50% (signalled by the red pie chart) and the probability of a positive return was less than 50% (signalled by the green pie chart). These results could be explained by the fact that Probabilistic scenarios clearly alerted investors on too risky securities, so they ended up choosing the WI scheme. This is an interesting observation which deserves future research.

In line with existing findings of behavioural studies, this research shows that risk evaluation and financial decisions are sensitive to the way that financial information is disclosed. As for financial disclosure, Probabilistic scenarios seem to represent a possible solution to reducing potential biases that occur during the decision-making process. Knowing the probability associated with the future scenarios of an investment would not eliminate uncertainty, but provides customers with a further element on the basis of which they can decide if and how to invest their savings. However, the analysis highlights that frame simplification (i.e. removing narrative description) of the Probabilistic approach may not be sufficient to ensure correct risk perception and unbiased investment choices. Indeed, besides the visual elements, more explanation can be given by, for example, a proper narrative: a disclaimer and warning may be added to Probabilistic representation in order to avoid the wrong interpretation of the probabilities.

To the authors' knowledge, no previous studies have examined the What-if and Probabilistic scenarios schemes to date. Therefore, this research represents a starting point for future analyses which involve different experimental setting to guarantee the external validity of the results.A more complex consumer test could be run in order to analyse the impact of these schemes on risk-return representation of more complex non-equity products than a subordinated bond or measure the impact on individuals with different risk tolerances and impulsivity levels. In addition, What-if and Probabilistic schemes could be modified (e.g., add/remove colours for both the frames or add a risk indicator) and the best balance of information could be found. Future researches focused on more complex products are necessary to investigate the role of frame of different type of risk embedded in a security (e.g. credit risk and market risk). In addition, further laboratory experiments could be run. In particular, future studies are needed in order to better understand the hierarchy of the perceptual factors affecting the investor's behaviour. It is important to note that, due to the explorative nature of the laboratory experiment, we tested only a small sample of participants for validating the experimental method and to obtain a sensitive measure of behavioral and physiological differences between representation schemes. The preliminary results obtained in relation to gender and education as reported in the previous section need to be further investigate by analysing a larger and more representative sample of participants. Thus, an important future direction of research will be to test a larger and stratified sample (according to ISTAT ranks), in order to better understand the influence of individuals, socio-demographic factors. Another promising future direction is to combine the present experimental paradigm with new technologies, such as immersive virtual reality (Vecchiato *et al.*, 2015a, 2015b; Jelić *et al.*, 2016; Tieri *et al.*, 2018) which would allow the stimulation of a realistic external environment (e.g., a bank)

and to investigate how the environmental factors, including the gender of the proposer, affects the investor's behaviour and physiological reactivity. Finally, the methodologies used aimed to explore the effect of the format for future performance representation, while leaving open the question about which costs and risk indicators should be adopted.

3.5 References Chapter 3

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

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modelling. *Journal of Management*, Vol. 39(6), pp.1490-1528.

Almenberg, J., & Säve-Söderbergh, J. (2011). Financial literacy and retirement planning in Sweden. *Journal of Pension Economics & Finance*, *10*(4), pp. 585-598.

Anderson, C., Kent, J., & Deanna, M. (2000). Personal finance and the rush to competence: Financial Literacy Education in the US. *Middleburg, Virginia–2011*.

ANZ (2008). ANZ Survey of Adult Financial Literacy in Australia. *The Social Research Centre*.

Atkinson, A. & F. Messy (2012). Measuring Financial Literacy: Results of the OECD / International Network on Financial Education (INFE) Pilot Study, *OECD Working Papers on Finance, Insurance and Private Pensions*, No. 15, OECD Publishing, Paris.

Bachmann, K., & Hens, T. (2015). Investment competence and advice seeking. *Journal of Behavioral and Experimental Finance*, *6*, pp. 27-41.

Baradell, J. G., & Klein, K. (1993). Relationship of life stress and body consciousness to hypervigilant decision making. *Journal of Personality and Social Psychology*, *64*(2), pp.267-273.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), pp. 255-278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, Vol. 67, Issue 1, pp. 1-48.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. Journal of neuroscience methods, 190(1), pp. 80-91.

Bland, J. M., & Bland, D. G. (1994). Statistics Notes: One and two sided tests of significance. *Bmj*, *309*(6949), 248.

Bongini, P., Trivellato, P., Zenga, M., Di Chiacchio, C. & S. Greco (2014). Gli indici di financial literacy, in *Le competenze economico-finanziarie degli italiani*. Bancaria editrice.

Braga, L. H., Easterbrook, B., Jegatheeswaran, K., & Lorenzo, A. J. (2016). From research question to conducting a randomized controlled trial on continuous antibiotic prophylaxis in prenatal hydronephrosis: a rational stepwise process. *Frontiers in pediatrics*, *4*, 27.

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of personality and social psychology*, *92*(5), 938.

Bucher-Koenen, T., Lusardi, A., Alessie, R., & Van Rooij, M. (2017). How financially literate are women? An overview and new insights. *Journal of Consumer Affairs*, *51*(2), pp. 255-283.

Camerer, C. (2011). The promise and success of lab-field generalizability in experimental economics: a critical reply to Levitt and List. *Working Paper*.

Castellano, R., & Cerqueti, R. (2013). Roots and effects of financial misperception in a stochastic dominance framework. *Quality & Quantity*, *47*(6), pp. 3371-3389.

Castellano, R., & Cerqueti, R. (2016). A theory of misperception in a stochastic dominance framework and its application to structured financial products. *IMA Journal of Management Mathematics*, *29*, pp. 23-37.

Castellano, R., Mancinelli, M., Ponsi, G., & Tieri G. (2019). What if versus probabilistic scenarios: a neuroscientific analysis, *Annals of Operational Research*, https://doi.org/10.1007/s10479-019-03272-5.

Chen, H., & Volpe, R. P. (2002). Gender differences in personal financial literacy among college students. *Financial services review*, *11*(3), 289.

Christopoulos, G. I., Uy, M. A., & Yap, W. J. (2016). The body and the brain: measuring skin conductance responses to understand the emotional experience. *Organizational Research Methods*, 127.

Danes, S., & Haberman, H. (2007). Teen Financial Knowledge, Self-Efficacy, and Behavior: A Gender View. *Journal of Financial Counseling and Planning*, *18(2)*, pp. 48-60.

Dawson, M. E., Schell, A. M., & Courtney, C. G. (2011). The skin conductance response, anticipation, and decision-making. *Journal of Neuroscience, Psychology, and Economics*, *4*(2), pp. 111-116.

Druckman, J. N., Green, D. P., Kuklinski, J. H., & Lupia, A. (2011). Experiments: an introduction to core concepts. *The Cambridge Handbook of Experimental Political Science*.

Druckman, J. N. & Kam, C. D. (2011). Students as experimental participants: A defence of the 'narrow data base'. *Cambridge handbook of experimental political science*.

Duke, C., Huck, S. & Wallace, B. (2010). Using experiments in consumer research. Experimental Economic Research. *A London Economics report for Ofcom*.

Falk, A. & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, *326*(5952), pp. 535-538.

Finucane, M. L., Mertz, C. K., Slovic, P., & Schmidt, E. S. (2005). Task complexity and older adults' decision-making competence. *Psychology and aging*, *20*(1), 71.

Fisher, C. G., & Wood, K. B. (2007). Introduction to and techniques of evidence-based medicine. *Spine*, *32*(19), S66-S72.

Frydman, C., & Camerer, C. F. (2016). The psychology and neuroscience of financial decision making. *Trends in cognitive sciences*, 20(9), pp. 661-675.

Gerardi K., Goette L. & Meier S. (2010). Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data. *Federal Reserve Bank of Atlanta Working Paper, Working paper 2010-10*.

Goldstein, H., (2011). Multilevel Statistical Models, vol. 922. John Wiley & Sons.

Gray, J.R., (1999). A Bias Toward Short-Term Thinking in Threat-Related Negative Emotional States. *Personality and Social Psychology Bulletin*, *25*, pp. 65-75.

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic literature*, *42*(4), pp. 1009-1055.

Hilgert, M. A., Hogarth, J. M., & Beverly, S. G. (2003). Household financial management: The connection between knowledge and behavior. *Fed. Res. Bull.*, *89*, 309.

Hosmer Jr., D.W., & Lemeshow, S. (2000). Applied Logistic Regression. John Wiley & Sons.

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.

Hung, A., Meijer, E., Mihaly, K., & Yoong, J. (2009). Building Up, spending down: Financial literacy, retirement savings management, and decumulation, Working paper.

Huston S. (2010). Measuring Financial Literacy. *Journal of Consumer Affairs*,.*44(2)*, pp. 296-316.

Iscenko, Z., Duke, C., Huck, S. & Wallace, B. (2014). How does selling insurance as an addon affect consumer decisions? A practical application of behavioural experiments in financial regulation. *FCA Occasional Paper 3.*

ISTAT (2017). Annuario statistico italiano 2017.

Jelić, A., Tieri, G., De Matteis, F., Babiloni, F., & Vecchiato, G. (2016). The Enactive Approach to Architectural Experience: A Neurophysiological Perspective on Embodiment, Motivation, and Affordances. *Frontiers in Psychology*, 7(March), pp. 1-20. https://doi.org/10.3389/fpsyg.2016.00481

Kempson, E., Collard, S., Turtle, J., & Worley, A. (2006). Financial capability baseline survey: Questionnaire. *Financial Services Authority Consumer Research Paper: 47b*.

Kim, J. (2001). Financial knowledge and subjective and objective financial well-being. *Consumer Interests Annual*, *47*, pp. 1-3.

Kramer, M. M. (2014). *Financial literacy, overconfidence and financial advice seeking*. EFMA Working paper.

Lajante, M., Droulers, O., Dondaine, T., & Amarantini, D. (2012). Opening the 'black box' of electrodermal activity in consumer neuroscience research. *Journal of Neuroscience, Psychology, and Economics*, 5(4), 238.

Levin, I., Weller, J., Pederson, A., & Harshman, L. (2007). Age-related differences in adaptive decision making: Sensitivity to expected value in risky choice. *Judgement and decision making*, *2*, pp. 225-233.

Levitt, S. D. & List, J. A. (2007a). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic perspectives*, *21*(2), pp. 153-174.

Levitt, S.D. & List, J. A. (2007b). Viewpoint: On the generalizability of lab behaviour to the field. *Canadian Journal of Economics, 40*, pp. 347–370.

Levy, H. (2015). Stochastic dominance: Investment decision making under uncertainty. 3rd edition, Springer.

Linciano, N., Caivano, V., Gentile, M., & Soccorso, P. (2018). Report on financial investments of Italian households. Behavioural attitudes and approach. *Statistics and analyses*. CONSOB.

Liuzza, M. T., Olofsson, J. K., Sabiniewicz, A., & Sorokowska, A. (2017). Body odor trait disgust sensitivity predicts perception of sweat biosamples. *Chemical senses*, *42*(6), pp. 479-485.

Lusardi, A. (2008a). Household saving behavior: The role of financial literacy, information, and financial education programs (No. w13824). National Bureau of Economic Research.

Lusardi, A. (2008b). Financial literacy: an essential tool for informed consumer choice? (No. w14084). National Bureau of Economic Research.

Lusardi, A., & Mitchell, O. S. (2011). Financial literacy and retirement planning in the United States. *Journal of Pension Economics & Finance*, *10*(4), pp. 509-525.

Lusardi, A., Mitchell, O. S., & Curto, V. (2012). Financial sophistication in the older population. *National Bureau of Economic Research, WP17863*.

Lusardi, A., & Tufano, P. (2015). Debt literacy, financial experiences, and overindebtedness. *Journal of Pension Economics & Finance*, *14*(4), pp. 332-368.

Maier, S. U., Makwana, A. B., & Hare, T. A. (2015). Acute stress impairs self-control in goal-directed choice by altering multiple functional connections within the brain's decision circuits. *Neuron*, *87*(3), pp. 621-631.

Mandell, L. (2008). Financial education in high school. *Overcoming the saving slump: How to increase the effectiveness of financial education and saving programs*, pp. 257-279.

Moore, D. L. (2003). Survey of financial literacy in Washington State: Knowledge, behavior, attitudes, and experiences. *Washington State University Social and Economic Sciences Research Center*, Technical Report 03-39.

NEFE - National Endowment for Financial Education (2006). Closing the Gap Between Knowledge and Behavior: Turning Education into Action. *Journal of Financial Counseling and Planning*, *17(1)*, pp. 73-90.

Noctor, M., Stoney, S., & Stradling, R. (1992). Financial literacy: a discussion of concepts and competences of financial literacy and opportunities for its introduction into young people's learning. *National Foundation for Educational Research*.

O'Connell, A. A., Goldstein, J., Rogers, H. J. and Peng, C. J. (2008). Multilevel logistic models for dichotomous and ordinal data. Multilevel modelling of educational data, pp.199-242.

OECD (2005), Recommendation on principles and good practices for financial education and awareness, Recommendation of the Council, June 2005.

Pabst, S., Brand, M., & Wolf, O. T. (2013). Stress and decision making: a few minutes make all the difference. *Behavioural brain research*, *250*, pp. 39-45.

Palomäki, J., Kosunen, I., Kuikkaniemi, K., Yamabe, T., & Ravaja, N. (2013). Anticipatory electrodermal activity and decision making in a computer poker-game. *Journal of Neuroscience, Psychology, and Economics*, *6*(1), pp. 55-70.

Panasiti, M. S., Ponsi, G., Monachesi, B., Lorenzini, L., Panasiti, V., & Aglioti, S. M. (2018). Cognitive load and emotional processing in psoriasis: a thermal imaging study. *Experimental brain research*, pp. 1-12.

Parker, A. M., & Fischhoff, B. (2005). Decision-making competence: External validation through an individual-differences approach. *Journal of Behavioral Decision Making*, *18*(1), pp. 1-27.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of school psychology*, *48*(1), pp. 85-112.

Pinheiro, J. C., & Bates, D. M., (2000), Mixed-Effects Models in S and S-PLUS. New York: Springer.

Ponsi, G., Panasiti, M. S., Scandola, M., & Aglioti, S. M. (2016). Influence of warmth and competence on the promotion of safe in-group selection: Stereotype content model and social categorization of faces. *The Quarterly Journal of Experimental Psychology*, *69*(8), pp. 1464-1479.

Ponsi, G., Panasiti, M. S., Aglioti, S. M., & Liuzza, M. T. (2017a). Right-wing authoritarianism and stereotype-driven expectations interact in shaping intergroup trust in one-shot vs multiple-round social interactions. *PloS one*, *12*(12).

Ponsi, G., Panasiti, M. S., Rizza, G., & Aglioti, S. M. (2017b). Thermal facial reactivity patterns predict social categorization bias triggered by unconscious and conscious emotional stimuli. *Proc. R. Soc. B*, *284*(1861).

Ponsi, G., Monachesi, B., Panasiti, V., Aglioti, S. M., & Panasiti, M. S. (2019). Physiological and behavioral reactivity to social exclusion: a functional infrared thermal imaging study in patients with psoriasis. *Journal of neurophysiology*.

Porcelli, A. J., & Delgado, M. R. (2009). Acute stress modulates risk taking in financial decision making. *Psychological Science*, *20*(3), pp. 278-283.

Porcelli, A. J., & Delgado, M. R. (2017). Stress and decision making: effects on valuation, learning, and risk-taking. *Current opinion in behavioral sciences*, *14*, pp. 33-39.

Raudenbush, S.W., and Bryk, A.S. (2002). Hierarchical linear models: Applications and data analysis methods (Second Edition). Thousand Oaks, CA: Sage Publications.

Remund, D. L. (2010). Financial literacy explicated: The case for a clearer definition in an increasingly complex economy. *Journal of consumer affairs*, *44*(2), pp. 276-295.

Robb, C., & Sharpe, D. (2009). Effect of Personal Financial Knowledge on College Students' Credit Card Behavior. *Journal of Financial Counseling and Planning*, *20(1)*, pp. 25-43.

Sommet, N., & Morselli, D. (2017). Keep calm and learn multilevel logistic modeling: A simplified three-step procedure using Stata, R, Mplus, and SPSS. *International Review of Social Psychology*, *30*(1), pp.203-2018.

Starcke, K., & Brand, M. (2012). Decision making under stress: a selective review. *Neuroscience & Biobehavioral Reviews*, *36*(4), pp. 1228-1248.

Tieri, G., Tidoni, E., Pavone, E. F., & Aglioti, S. M. (2015). Body visual discontinuity affects feeling of ownership and skin conductance responses. *Scientific reports*, *5*, pp. 18.

Tieri, G., Morone, G., Paolucci, S., & Iosa, M. (2018). Virtual reality in cognitive and motor rehabilitation: facts, fiction and fallacies. *Expert Review of Medical Devices*, 15(2), pp. 107-117. https://doi.org/10.1080/17434440.2018.1425613

US FLEC - U.S. Department of the Treasury, Financial Literacy and Education Commission (2006), Taking ownership of the future*: The National Strategy for Financial Literacy*.

Van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, *101*(2), pp. 449-472.

Vecchiato, G., Jelic, A., Tieri, G., Maglione, A. G., De Matteis, F., & Babiloni, F. (2015a). Neurophysiological correlates of embodiment and motivational factors during the perception of virtual architectural environments. *Cognitive Processing*, pp. 3-7.

Vecchiato, G., Tieri, G., Jelic, A., Maglione, A. G., De Matteis, F., & Babiloni, F. (2015b). Electroencephalographic Correlates of Sensorimotor Integration and Embodiment during the Appreciation of Virtual Architectural Environments. *Frontiers in Neuroscience*, 16(S1), pp. 425-429. https://doi.org/10.3389/fpsyg.2015.01944

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

4 Annex I: Web survey.

4.1 The Questionnaires

Questionnaire A and Questionnaire B have the same structure, but they are different only in Section C where the tables of future performance of the selected products are shown according to the two representation schemes. After a brief description of the purpose underlying the research, the participants answered mandatory to each question. For a faster reading and completion of the survey, filtered questions were used. Moreover, at the beginning of Section B respondents were encouraged to answer 'I do not know' when they do not have any idea about the correct option with the purpose to capture actual levels of knowledge rather than lucky guesses. Both the questionnaires are reported in the following subsections 4.1.1 : *The Questionnaire 1 (Q1)* and 4.1.2: : *The Questionnaire 2 (Q2)*. The questions are divided as were displayed to participants in the screen of their device (for example 'S2' collects all the questions showed in the second screen, 'S3' collects all the questions showed in the third screen and so on).

4.1.1 The Questionnaire 1 (Q1)

S1 .....

Disclosure in prospectuses of financial instruments.

Sapienza University of Rome and Unitelma Sapienza started a statistical analysis about 'Disclosure in prospectuses of financial instruments'.

Thanks to you, we test both the efficiency and the level of understanding of some types of information provided in prospectuses of financial products.

We aim at realising a standardized and simple documentations in order to offer clear and meaningful information about the features of financial instruments. Investors will be able to assess the risk-return profile of a product and make an informed investment choice.

The questionnaire is composed by four sections on:

- investment habits and experience;
- the level of financial literacy;
- a choice of a financial product;
- sociodemographic aspects of the respondent.

In filling the questionnaire, please remember the following instructions:

- all questions with the symbol * are mandatory;
- answer independently without be helped by other people.

NEXT

S2 .....

Section A. Investment habits and experience.

A.1. Are you updated about economics and financial news? *

- ○ Yes, every day
- ○ Yes, at least two times per week
- ○ Yes, rarely
- ○ No

A.2.[76] Have you ever bought financial instruments (such as cash deposit, life insurance, pension fund, bond, stock etc.)? *

- ○ Yes
- ○ No

BACK NEXT

S3 .....

A.3. What did you buy? (more answers are possible) *

- ☐ Deposit
- ☐ Bond (e.g. sovereign bond, corporate bond)
- ☐ Stocks
- ☐ Insurance product (e.g. life insurance)
- ☐ Investment fund
- ☐ Pension fund
- ☐ Derivative, covered warrant, ETF
- ☐ I do not know
- ☐ Other (please indicate what):

A.4. Consider last investment, did you gain or loss? *

- ○ Gain
- ○ Loss
- ○ Neither gain nor loss

A.5.[77] Consider last investment, did you choose it by yourself or after an advice? *

- ○ By myself
- ○ After an advice

BACK NEXT

S4 .....

[76] Filter question: if the participant answers 'Yes' S3 would be shown, otherwise it skips to S5.
[77] Filter question: if the participant answers 'After an advice' S4 would be shown, otherwise it skips to S5.

A.6. Please, select who advised you. *

- Parents or friends
- Bank or financial advisor as old client
- Bank or financial advisor as new client
- Newspaper, internet or other media

BACK NEXT

S5 .....

Section B. Financial literacy.

ATTENTION

In answering to next questions, please choose 'I do not know' only after evaluating all the alternatives.

Moreover, 'I do not know' option is more preferable than a random selection.

BACK NEXT

S6 .....

B.1. Suppose you had € 100 in a saving account and the interest rate was 2% per year. Without considering costs and fees, after 5 years, how much do you think you would have in the account if you left the money to grow? *

- More than € 102
- Exactly € 102
- Less than € 102
- I do not know

B.2. Suppose that you had € 100 in a saving account. Imagine that the interest rate on your saving account was 1% per year and inflation rate was 2% per year. After 1 year, how much would you be able to buy with the money in this account: *

- more than today
- the same as today
- less than today
- I do not know

BACK NEXT

S7 .....

B.3. Please, read the following sentence:

Buying stocks of one firm is a safer investment than buying shares of an equity investment fund.

Is the above sentence true or false? *

- o True
- o False
- o I do not know

B.4. Diversifying an investment means investing …

- o ... in financial products with different interest rates
- o ... in financial products of corporates belonging to different sectors
- o ... in several financial products linked by the same level of risk
- o I do not know

BACK NEXT

S8 ……

B.5. Consider the following prospectuses which describe two financial investment. Please evaluate the possible pros and cons of each package. *

Investment 1 Package which pays		
euro 7	with probability	33%
euro 5	with probability	33%
euro -2	with probability	33%
euro 0	with probability	1%

Investment 2 Package which pays		
euro 5	with probability	33%
euro 1	with probability	33%
euro 9	with probability	33%
euro 0	with probability	1%

Which one do you prefer? *

- o Investment 1
- o Investment 2
- o It is not possible to answer basing on the provided information
- o I do not know

Random number generator [78]

BACK NEXT

S9 ……

[78] At this stage, the software drawn in background (i.e. not shown to the participant) an integer number *x* from a uniform distribution in the interval [0;9] in order to randomize the order of presenting the choice between Product I and Product II. When $x \leq 4$, S9 is shown; otherwise the questionnaire skips to S10.

Section C. Choice of a financial product.

You have €1000 and you want to invest them on a financial product for 7 years. The financial advisor offers you Product A and Product B displaying the following information on their main characteristics and expected returns at maturity.

Product A

A bond issued by Bank Alpha with rating BBB+. With the maturity of 7 years, the security pays yearly a fixed interest rate of 4,5%. The invested capital will be paid back entirely (€1000) at maturity.

In case of issuer default, the investor may incur in a loss even of the invested capital.

Expected returns at maturity are summarized in the following table.

SCENARIO	EXPECTED RETURN PER YEAR*
Negative scenario - Issuer defaults before the maturity and only part of the invested amount is paid back.	-8,38%
Neutral scenario – All the coupons are regularly paid but the issuer defaults at maturity.	1,12%
Positive scenario – All coupons and invested capital are regularly paid.	4,5%

*if held to maturity

Product B

A bond issued by Bank Beta with rating BB-. With the maturity of 7 years, the security pays yearly a fixed interest rate of 5%. From the third year and for each year up to maturity, the invested capital will be paid back at 20% calculated on nominal value (€1000).

In case of issuer default, the investor may incur in a loss even of the invested capital.

Expected returns at maturity are summarized in the following table.

SCENARIO	EXPECTED RETURN PER YEAR*
Negative scenario - Issuer defaults before the maturity and only part of the invested amount is paid back.	-3,02%
Neutral scenario – All the coupons are regularly paid but the issuer defaults at maturity.	4,15%
Positive scenario – All coupons and invested capital are regularly paid.	5%

*if held to maturity

C.1. Basing your evaluation only on the provided information, which product do you prefer? *

- Product A
- Product B
- I do not know

C.2. How much confident are you about your choice? *

	1	2	3	4	5	
Little confident	O	O	O	O	O	Very confident

C.3. Consider an evaluation scale from 1 to 7 (where 1 = few, 7 = a lot). Please, select the risk level of the two products. *

	1	2	3	4	5	6	7
Titolo A	O	O	O	O	O	O	O
Titolo B	O	O	O	O	O	O	O

C.4.[79] Evaluating only the information provided in the tables, in order to make an investment decision, you think that the information is (consider a scale of number from 0 to 5, where 0 = at all, 5 = a lot). *

	0	1	2	3	4	5
Useful	O	O	O	O	O	O
Clear	O	O	O	O	O	O
Sufficient	O	O	O	O	O	O
Comparable	O	O	O	O	O	O

BACK NEXT

S10 [80]……

Section C. Choice of a financial product.

You have €1000 and you want to invest them on a financial product for 7 years. The financial advisor offers you Product A and Product B displaying the following information on their main characteristics and expected returns at maturity.

Product A

A bond issued by Bank Alpha with rating BB-. With the maturity of 7 years, the security pays yearly a fixed interest rate of 5%. From the third year and for each year up to maturity, the invested capital will be paid back at 20% calculated on nominal value (€1000).

In case of issuer default, the investor may incur in a loss even of the invested capital.

Expected returns at maturity are summarized in the following table.

[79] After answering to C.4. and clicking on 'next', S11 is shown.
[80] This Screen is shown only when the random number is x ≥ 5.

SCENARIO	EXPECTED RETURN PER YEAR*
Negative scenario - Issuer defaults before the maturity and only part of the invested amount is paid back.	-3,02%
Neutral scenario – All the coupons are regularly paid but the issuer defaults at maturity.	4,15%
Positive scenario – All coupons and invested capital are regularly paid.	5%

*if held to maturity

Product B

A bond issued by Bank Beta with rating BBB+. With the maturity of 7 years, the security pays yearly a fixed interest rate of 4,5%. The invested capital will be paid back entirely (€1000) at maturity.

In case of issuer default, the investor may incur in a loss even of the invested capital.

Expected returns at maturity are summarized in the following table.

SCENARIO	EXPECTED RETURN PER YEAR*
Negative scenario - Issuer defaults before the maturity and only part of the invested amount is paid back.	-8,38%
Neutral scenario – All the coupons are regularly paid but the issuer defaults at maturity.	1,12%
Positive scenario – All coupons and invested capital are regularly paid.	4,5%

*if held to maturity

C.1. Basing your evaluation only on the provided information, which product do you prefer? *

- Product A
- Product B
- I do not know

C.2. How much confident are you about your choice? *

	1	2	3	4	5	
Little confident	O	O	O	O	O	Very confident

C.3. Consider an evaluation scale from 1 to 7 (where 1 = few, 7 = a lot). Please, select the risk level of the two products. *

	1	2	3	4	5	6	7
Titolo A	O	O	O	O	O	O	O
Titolo B	O	O	O	O	O	O	O

C.4. Evaluating only the information provided in the tables, in order to make an investment decision, you think that the information is (consider a scale of number from 0 to 5, where 0 = at all, 5 = a lot). *

	0	1	2	3	4	5
Useful	O	O	O	O	O	O
Clear	O	O	O	O	O	O
Sufficient	O	O	O	O	O	O
Comparable	O	O	O	O	O	O

BACK NEXT

S11 .....

Comparison between modalities.

You have €1000 and you want to invest them on a financial product for 7 years. The financial advisor offers you Product C displaying the following information on its main characteristics and two different tables of expected returns at maturity.

Product C

A bond issued by Bank Gamma with rating BB-. With the maturity of 7 years, the security pays yearly a fixed interest rate of 5%. From the third year and for each year up to maturity, the invested capital will be paid back at 20% calculated on nominal value (€1000).

In case of issuer default, the investor may incur in a loss even of the invested capital.

Expected returns at maturity are summarized using two different tables.

Method 1

SCENARIO	EXPECTED RETURN PER YEAR*
Negative scenario - Issuer defaults before the maturity and only part of the invested amount is paid back.	-3,02%
Neutral scenario – All the coupons are regularly paid but the issuer defaults at maturity.	4,15%
Positive scenario – All coupons and invested capital are regularly paid.	5%

* if held to maturity

Method 2

EVENTS	PROBABILITY	MEDIUM VALUE IN €*
NEGATIVE performance	66,38% 	576
NEUTRAL performance	3,76% 	1045
POSITIVE performance	29,86% 	1249

* at maturity for an initial investment of 1000 euro

C.5. What type of representation of expected returns do you prefer? *

- Method 1
- Method 2
- I do not know

NEXT

S12 .....

Section D. Sociodemographic aspects.

D.1. Sex *

- Male
- Female

D.2.[81] Year of birth. *

Example: 1948

D.3. Where is your permanent location? *

(All Italian region are listed in a dropdown menu)

D.4. Select your marital status. *

- single
- cohabitant
- married
- divorced
- widower

D.5. Please, including yourself, write how many people your family is composed with.*

Example: 3

BACK NEXT

S13 .....

[81] In order to get only participants aged between 20 and 64 years old, the minimum value in of this question is 1998 while the maximum is 1954.

D.6. What is your higher level of education? *

- ○ Primary school
- ○ Secondary school
- ○ Tertiary school
- ○ Bachelor's degree
- ○ Master's degree
- ○ Specialization
- ○ PhD

D.6.1.[82] Please, select the field in which you are specialized. *

- ○ Biology, Chemistry, Medicine
- ○ Law
- ○ Economics
- ○ Humanities, Philosophy, Psychology
- ○ Physics, Engineering, Maths, Statistics
- ○ Other (please indicate):

D.7.[83] Which is your actual work condition? *

- ○ Employed
- ○ Unemployed and looking for a job
- ○ Inactive (student, unemployed and not looking for a job, housewife, retired)

○

BACK NEXT

S14……

D.8.[84] You are: *

- ○ employee
- ○ occasional jobs/freelance
- ○ self employed

BACK NEXT

S15 ……

[82] Filter question: if the participant answers 'Bachelor's degree', 'Master's degree', 'Specialization' or 'PhD' to Question D.6, the D.6.1 is automatically displayed; otherwise D.6.1 would be not shown.

[83] Filter question: if 'employee', S16 is shown; if 'Unemployed and looking for a job', S17 is shown; if 'Inactive (student, unemployed and not looking for a job, housewife, retired)', S18 is shown.

[84] This question is visualised only if the participant would answer 'employee' to the Question D.7. After answering to D.8. and clicking the button 'next', S17 is displayed.

D.8.[85] You are: *

- a short-term unemployed (less than 3 months)
- a medium-term unemployed (from 3 to 12 months)
- a long-term unemployed (more than 1 year)
- looking for first job

BACK NEXT

S16 .....

D.8.[86] You are: *

- student
- housewife
- retired
- unemployed and not looking for a job

BACK NEXT

S17.....

D.9.[87] Including yourself, how many people in your family have a 'stable income'? Define 'stable income' as all the revenues obtained constantly on monthly basis (for instance: salary, pension, financial returns etc.). *

Example: 3

D.10. Consider who has the higher income in your family. Please, select your relationship with this person. *

- Myself
- Partner
- Parent
- Son/daughter

D.11.[88] Consider your monthly income. Do you save part of it? *

- Yes
- No
- I do not have a monthly income

[85] This question is visualised only if the participant would answer 'Unemployed and looking for a job' to the Question D.7. After answering to D.8. and clicking the button 'next', S17 is displayed.
[86] This question is visualised only if the participant would answer 'Inactive' to the Question D.7. After answering to D.8. and clicking the button 'next', S17 is displayed.
[87] In order to avoid incoherent answers, the number provided in this question should be lower or equal to the answer of Question D.5.
[88] Filter question: when the participant answers 'yes', S18 is shown; otherwise, it skips to S19.

BACK NEXT

S18 ……

D.12. How much do you save?

- Less than 10% of your monthly income
- Between 10% and 30% of your monthly income
- More than 30% of your monthly income

BACK NEXT

S19 ……

SEND THE QUESTIONNAIRE

4.1.2 The Questionnaire 2 (Q2)

S1 ……

Disclosure in prospectuses of financial instruments.

Sapienza University of Rome and Unitelma Sapienza started a statistical analysis about 'Disclosure in prospectuses of financial instruments'.

Thanks to you, we test both the efficiency and the level of understanding of some types of information provided in prospectuses of financial products.

We aim at realising a standardized and simple documentations in order to offer clear and meaningful information about the features of financial instruments. Investors will be able to assess the risk-return profile of a product and make an informed investment choice.

The questionnaire is composed by four sections on:

- investment habits and experience;
- the level of financial literacy;
- a choice of a financial product;
- sociodemographic aspects of the respondent.

In filling the questionnaire, please remember the following instructions:

- all questions with the symbol * are mandatory;
- answer independently without be helped by other people.

NEXT

S2 ……

Section A. Investment habits and experience.

A.1. Are you updated about economics and financial news? *

- ○ Yes, every day
- ○ Yes, at least two times per week
- ○ Yes, rarely
- ○ No

A.2.[89] Have you ever bought financial instruments (such as cash deposit, life insurance, pension fund, bond, stock etc.)? *

- ○ Yes
- ○ No

BACK NEXT

S3 .....

A.3. What did you buy? (more answers are possible) *

- ☐ Deposit
- ☐ Bond (e.g. sovereign bond, corporate bond)
- ☐ Stocks
- ☐ Insurance product (e.g. life insurance)
- ☐ Investment fund
- ☐ Pension fund
- ☐ Derivative, covered warrant, ETF
- ☐ I do not know
- ☐ Other (please indicate what):

A.4. Consider last investment, did you gain or loss? *

- ○ Gain
- ○ Loss
- ○ Neither gain nor loss

A.5.[90] Consider last investment, did you choose it by yourself or after an advice? *

- ○ By myself
- ○ After an advice

BACK NEXT

S4 .....

[89] Filter question: if the participant answers 'Yes' S3 would be shown, otherwise it skips to S5.
[90] Filter question: if the participant answers 'After an advice' S4 would be shown, otherwise it skips to S5.

A.6. Please, select who advised you. *

- Parents or friends
- Bank or financial advisor as old client
- Bank or financial advisor as new client
- Newspaper, internet or other media

BACK NEXT

S5 .....

Section B. Financial literacy.

ATTENTION

In answering to next questions, please choose 'I do not know' only after evaluating all the alternatives.

Moreover, 'I do not know' option is more preferable than a random selection.

BACK NEXT

S6 .....

B.1. Suppose you had € 100 in a saving account and the interest rate was 2% per year. Without considering costs and fees, after 5 years, how much do you think you would have in the account if you left the money to grow? *

- More than € 102
- Exactly € 102
- Less than € 102
- I do not know

B.2. Suppose that you had € 100 in a saving account. Imagine that the interest rate on your saving account was 1% per year and inflation rate was 2% per year. After 1 year, how much would you be able to buy with the money in this account: *

- more than today
- the same as today
- less than today
- I do not know

BACK NEXT

S7 .....

B.3. Please, read the following sentence:

Buying stocks of one firm is a safer investment than buying shares of an equity investment fund.

Is the above sentence true or false? *

- o True
- o False
- o I do not know

B.4. Diversifying an investment means investing …

- o ... in financial products with different interest rates
- o ... in financial products of corporates belonging to different sectors
- o ... in several financial products linked by the same level of risk
- o I do not know

BACK NEXT

S8 ……

B.5. Consider the following prospectuses which describe two financial investment. Please evaluate the possible pros and cons of each package. *

Investment 1 Package which pays		
euro 7	with probability	33%
euro 5	with probability	33%
euro -2	with probability	33%
euro 0	with probability	1%

Investment 2 Package which pays		
euro 5	with probability	33%
euro 1	with probability	33%
euro 9	with probability	33%
euro 0	with probability	1%

Which one do you prefer? *

- o Investment 1
- o Investment 2
- o It is not possible to answer basing on the provided information
- o I do not know

Random number generator [91]

BACK NEXT

S9 ……

[91] At this stage, the software drawn in background (i.e. not shown to the participant) an integer number *x* from a uniform distribution in the interval [0;9] in order to randomize the order of presenting the choice between Product I and Product II. When x ≤ 4, S9 is shown; otherwise the questionnaire skips to S10.

Section C. Choice of a financial product.

You have €1000 and you want to invest them on a financial product for 7 years. The financial advisor offers you Product A and Product B displaying the following information on their main characteristics and expected returns at maturity.

Product A

A bond issued by Bank Alpha with rating BBB+. With the maturity of 7 years, the security pays yearly a fixed interest rate of 4,5%. The invested capital will be paid back entirely (€1000) at maturity.

In case of issuer default, the investor may incur in a loss even of the invested capital.

Expected returns at maturity are summarized in the following table.

EVENTS	PROBABILITY	MEDIUM VALUE IN €*
NEGATIVE performance	50,21% 	529
NEUTRAL performance	-	-
POSITIVE performance	49,79% 	1268

* at maturity for an initial investment of 1000 euros

Product B

A bond issued by Bank Beta with rating BB-. With the maturity of 7 years, the security pays yearly a fixed interest rate of 5%. From the third year and for each year up to maturity, the invested capital will be paid back at 20% calculated on nominal value (€1000).

In case of issuer default, the investor may incur in a loss even of the invested capital.

Expected returns at maturity are summarized in the following table.

EVENTS	PROBABILITY	MEDIUM VALUE IN €*
NEGATIVE performance	66,38% 	576
NEUTRAL performance	3,76% 	1045
POSITIVE performance	29,86% 	1249

* at maturity for an initial investment of 1000 euros

C.1. Basing your evaluation only on the provided information, which product do you prefer? *

- ○ Product A
- ○ Product B
- ○ I do not know

C.2. How much confident are you about your choice? *

	1	2	3	4	5	
Little confident	O	O	O	O	O	Very confident

C.3. Consider an evaluation scale from 1 to 7 (where 1 = few, 7 = a lot). Please, select the risk level of the two products. *

	1	2	3	4	5	6	7
Titolo A	O	O	O	O	O	O	O
Titolo B	O	O	O	O	O	O	O

C.4.[92] Evaluating only the information provided in the tables, in order to make an investment decision, you think that the information is (consider a scale of number from 0 to 5, where 0 = at all, 5 = a lot). *

	0	1	2	3	4	5
Useful	O	O	O	O	O	O
Clear	O	O	O	O	O	O
Sufficient	O	O	O	O	O	O
Comparable	O	O	O	O	O	O

BACK NEXT

S10 [93]……

Section C. Choice of a financial product.

You have €1000 and you want to invest them on a financial product for 7 years. The financial advisor offers you Product A and Product B displaying the following information on their main characteristics and expected returns at maturity.

Product A

A bond issued by Bank Alpha with rating BB-. With the maturity of 7 years, the security pays yearly a fixed interest rate of 5%. From the third year and for each year up to maturity, the invested capital will be paid back at 20% calculated on nominal value (€1000).

In case of issuer default, the investor may incur in a loss even of the invested capital.

Expected returns at maturity are summarized in the following table.

[92] After answering to C.4. and clicking on 'next', S11 is shown.
[93] This Screen is shown only when the random number is x ≥ 5.

EVENTS	PROBABILITY	MEDIUM VALUE IN €*
NEGATIVE performance	66,38% 	576
NEUTRAL performance	3,76% 	1045
POSITIVE performance	29,86% 	1249

* at maturity for an initial investment of 1000 euros

Product B

A bond issued by Bank Beta with rating BBB+. With the maturity of 7 years, the security pays yearly a fixed interest rate of 4,5%. The invested capital will be paid back entirely (€1000) at maturity.

In case of issuer default, the investor may incur in a loss even of the invested capital.

Expected returns at maturity are summarized in the following table.

EVENTS	PROBABILITY	MEDIUM VALUE IN €*
NEGATIVE performance	50,21% 	529
NEUTRAL performance	-	-
POSITIVE performance	49,79% 	1268

* at maturity for an initial investment of 1000 euros

C.1. Basing your evaluation only on the provided information, which product do you prefer? *

- ○ Product A
- ○ Product B
- ○ I do not know

C.2. How much confident are you about your choice? *

	1	2	3	4	5	
Little confident	O	O	O	O	O	Very confident

C.3. Consider an evaluation scale from 1 to 7 (where 1 = few, 7 = a lot). Please, select the risk level of the two products. *

	1	2	3	4	5	6	7
Titolo A	O	O	O	O	O	O	O
Titolo B	O	O	O	O	O	O	O

C.4. Evaluating only the information provided in the tables, in order to make an investment decision, you think that the information is (consider a scale of number from 0 to 5, where 0 = at all, 5 = a lot). *

	0	1	2	3	4	5
Useful	O	O	O	O	O	O
Clear	O	O	O	O	O	O
Sufficient	O	O	O	O	O	O
Comparable	O	O	O	O	O	O

BACK NEXT

S11 .....

Comparison between modalities.

You have €1000 and you want to invest them on a financial product for 7 years. The financial advisor offers you Product C displaying the following information on its main characteristics and two different tables of expected returns at maturity.

Product C

A bond issued by Bank Gamma with rating BB-. With the maturity of 7 years, the security pays yearly a fixed interest rate of 5%. From the third year and for each year up to maturity, the invested capital will be paid back at 20% calculated on nominal value (€1000).

In case of issuer default, the investor may incur in a loss even of the invested capital.

Expected returns at maturity are summarized using two different tables.

Method 1

SCENARIO	EXPECTED RETURN PER YEAR*
Negative scenario - Issuer defaults before the maturity and only part of the invested amount is paid back.	-3,02%
Neutral scenario – All the coupons are regularly paid but the issuer defaults at maturity.	4,15%
Positive scenario – All coupons and invested capital are regularly paid.	5%

* if held to maturity

Method 2

EVENTS	PROBABILITY	MEDIUM VALUE IN €*
NEGATIVE performance	66,38% 	576
NEUTRAL performance	3,76% 	1045
POSITIVE performance	29,86% 	1249

* at maturity for an initial investment of 1000 euro

C.5. What type of representation of expected returns do you prefer? *

- Method 1
- Method 2
- I do not know

NEXT

S12 .....

Section D. Sociodemographic aspects.

D.1. Sex *

- Male
- Female

D.2.[94] Year of birth. *

Example: 1948

D.3. Where is your permanent location? *

(All Italian regions are listed in a dropdown menu)

D.4. Select your marital status. *

- single
- cohabitant
- married
- divorced
- widower

D.5. Please, including yourself, write how many people your family is composed with.*

Example: 3

BACK NEXT

S13 .....

[94] In order to get participants aged between 20 and 64 years old, the minimum value in of this question is 1998 while the maximum is 1954.

D.6. What is your higher level of education? *

- ○ Primary school
- ○ Secondary school
- ○ Tertiary school
- ○ Bachelor's degree
- ○ Master's degree
- ○ Specialization
- ○ PhD

D.6.1.[95] Please, select the field in which you are specialized. *

- ○ Biology, Chemistry, Medicine
- ○ Law
- ○ Economics
- ○ Humanities, Philosophy, Psychology
- ○ Physics, Engineering, Maths, Statistics
- ○ Other (please indicate):

D.7.[96] Which is your actual work condition? *

- ○ Employed
- ○ Unemployed and looking for a job
- ○ Inactive (student, unemployed and not looking for a job, housewife, retired)

○

BACK NEXT

S14 .....

D.8.[97] You are: *

- ○ employee
- ○ occasional jobs/freelance
- ○ self employed

BACK NEXT

S15 .....

[95] Filter question: if the participant answers 'Bachelor's degree', 'Master's degree', 'Specialization' or 'PhD' to Question D.6, the D.6.1 is automatically displayed; otherwise, D.6.1 would be not shown.

[96] Filter question: if 'employee', S16 is shown; if 'Unemployed and looking for a job', S17 is shown; if 'Inactive (student, unemployed and not looking for a job, housewife, retired)', S18 is shown.

[97] This question is visualised only if the participant would answer 'employee' to the Question D.7. After answering to D.8. and clicking the button 'next', S17 is displayed.

D.8.[98] You are: *

- a short-term unemployed (less than 3 months)
- a medium-term unemployed (from 3 to 12 months)
- a long-term unemployed (more than 1 year)
- looking for first job

BACK NEXT

S16 .....

D.8.[99] You are: *

- student
- housewife
- retired
- unemployed and not looking for a job

BACK NEXT

S17 .....

D.9.[100] Including yourself, how many people in your family have a 'stable income'? Define 'stable income' as all the revenues obtained constantly on monthly basis (for instance: salary, pension, financial returns etc.). *

Example: 3

D.10. Consider who has the higher income in your family. Please, select your relationship with this person. *

- Myself
- Partner
- Parent
- Son/daughter

D.11.[101] Consider your monthly income. Do you save part of it? *

- Yes
- No
- I do not have a monthly income

BACK NEXT

S18 .....

[98] This question is visualised only if the participant would answer 'Unemployed and looking for a job' to the Question D.7. After answering to D.8. and clicking the button 'next', S17 is displayed.
[99] This question is visualised only if the participant would answer 'Inactive' to the Question D.7. After answering to D.8. and clicking the button 'next', S17 is displayed.
[100] In order to avoid incoherent answers, the number provided in this question should be lower or equal to the answer of D.5.
[101] Filter question: when the participant answers 'yes', S18 is shown; otherwise it skips to S19.

D.12. How much do you save?

- Less than 10% of your monthly income
- Between 10% and 30% of your monthly income
- More than 30% of your monthly income

BACK NEXT

S19 .....

SEND THE QUESTIONNAIRE

4.2 Sampling procedure

Considering Italians people from 20 to 64 years old as target population (N) as well as time and budget constraints, the target sample size (n^*) was 600 interviews for each treatment T (n_T^*; with n_1^* for Questionnaire 1 and n_2^* for Questionnaire 2; then $n^*= n_1^* + n_2^*$). Both the questionnaires were pre-tested on a sample of 52 students. Following a non-probability sampling design, a link to fulfil the questionnaire was shared on different web sites (Facebook[102] and Unitelma Sapienza) and an invitation text was sent to all available contacts. After a three months period during which a reminder was sent to all members every three weeks, 352 completed interviews were collected (n_1 was equal to 159 whereas n_2 was 193).

Given the low sample size and the meaningless representation of some sub-populations (for example adult people or with a low level of education), a complementary survey design was implemented in order to get a representative sample of the target population. A stratified sample was built considering the following stratification variables (x):

- geographical area as defined by the Nielsen Institute: North-East, North-West, Centre and South (indexed by NE, NO, CE, and SO) [103];
- gender (i.e. Female and Male);
- age according to three ranges: 20-34; 35-49; 50-64.
- the higher degree of education obtained by the respondent which is categorized in 'Low' when the respondent got the degree of 'Licenza Scuola Media' or lower; 'Medium' when at most the degree of 'Licenza Scuola Superiore' was obtained; and 'High' when at least gained a university degree.

The corresponding target size of each stratum (n_g^* where g =1, …, G represents the strata) was proportional to the weights of each stratum in the target population ($W_g = N_g/N$ where N_g represents the size of the stratum in the target population and thus $n_g^* = W_g n^*$). The weights were obtained from the annual survey on Italian population run by the ISTAT (2017). The table below reports the values of W_g where, for a better reading, the strata for female and male are coloured in pink and blue, respectively.

[102] The recruitment was characterized by two approaches. Firstly, we sent a direct invitation to all the contacts and asking for sharing the link with their contacts. At the same time, after contacting the administrators of several Communities which are involved in the aims of the research (for example the key words were 'investor protection', 'financial literacy' and 'behavioral economics'), the ones who accepted to participate in the survey sent an invitation to all registered members and posted the link in their homepage.

[103] North-East: Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia and Emilia-Romagna; North-West: Valle d'Aosta, Liguria, Piemonte and Lombardia; Centre: Toscana, Umbria, Marche, Lazio and Sardegna; South: Abruzzo, Molise, Campania, Basilicata Puglia, Calabria and Sicilia.

Table 4.1: Weights in the target population.

W_g	low				medium				high			
	NO	NE	CE	SO	NO	NE	CE	SO	NO	NE	CE	SO
20-34	0.7%	0.5%	0.6%	1.0%	1.9%	1.4%	1.4%	2.5%	0.9%	0.6%	0.7%	1.2%
	1.0%	0.7%	0.8%	1.3%	2.1%	1.5%	1.6%	2.7%	0.6%	0.4%	0.4%	0.7%
35-49	1.8%	1.3%	1.3%	2.3%	2.2%	1.6%	1.7%	2.9%	1.1%	0.8%	0.8%	1.4%
	2.1%	1.5%	1.6%	2.7%	2.2%	1.6%	1.6%	2.8%	0.8%	0.6%	0.6%	1.0%
50-64	2.4%	1.7%	1.8%	3.1%	1.8%	1.3%	1.3%	2.3%	0.6%	0.4%	0.4%	0.8%
	2.2%	1.6%	1.7%	2.9%	1.7%	1.2%	1.3%	2.2%	0.6%	0.4%	0.4%	0.7%

Successively, the realized size of each stratum (n_g) was computed considering the data in the pre-test and the ones collected in the first step. The necessary units for each stratum to get the target population were computed as $\Delta_g = n_g^* - n_g$. This process was repeated separately for n_1 and n_2 getting two different values of Δ_{gT} (i.e., $\Delta_{g1} = n_g^* - n_{g1}$ and $\Delta_{g2} = n_g^* - n_{g2}$). At this stage, using the available budget, an institute that operates in the field of social and economic studies (hereinafter provider) collected all the questionnaires following the provided quotas (Δ_{g1} and Δ_{g2})[104]. The provider built two base sampling lists following a systematic random sampling technique without repetition and contacted the selected potential participants according to the quotas. Moreover, two separate reserve lists (each with a sample size four time higher than the base lists) were provided in order to deal with the redemption rate (equals to 25% according to provider's experience). The sampling frame was an available list of subscribers to telephone directories who are not signed in the register of opposition and thus could be contacted in line with new regulation on privacy (so-called GDPR). At the end of the survey process, the provider collected 1054 interviews of which 537 for Questionnaire 1 and 517 for Questionnaire 2.

Merging the interviewers obtained during the three phases, the overall realized samples size was 1458 but 328 observations were delated from the dataset because suspicious: very short time spent to fulfil the questionnaire; to many 'I do not know' answers; or seems that answers are selected randomly without evaluate the options (e.g. Likert scale). Therefore, the sample size is 1130 observations with n_1 = 544 and n_2= 586 distributed as in the following table.

Table 4.2: Sample distribution in each stratum (weights in parentheses).

n_g	low				medium				high			
	NO	NE	C	S	NO	NE	C	S	NO	NE	C	S
20-34	8 (0.4%)	5 (0.2%)	2 (0.9%)	10 (0.4%)	19 (1.7%)	12 (1.1%)	20 (1.8%)	29 (2.6%)	14 (1.2%)	9 (0.8%)	27 (2.4%)	17 (1.5%)
	8 (0.3%)	3 (0.3%)	3 (0.9%)	10 (0.3%)	24 (2.1%)	17 (1.5%)	21 (1.9%)	26 (2.3%)	6 (0.5%)	5 (0.4%)	43 (3.8%)	27 (2.4%)
35-49	19 (0.9%)	10 (0.9%)	10 (2.0%)	23 0.9%)	23 (2.0%)	16 (1.4%)	15 (1.3%)	31 (2.7%)	12 (1.1%)	11 (1.0%)	22 (1.9%)	16 (1.4%)
	19 (0.8%)	9 (0.8%)	9 (1.9%)	21 (0.8%)	22 (1.9%)	17 (1.5%)	19 (1.7%)	31 (2.7%)	8 (0.7%)	7 (0.6%)	14 (1.2%)	15 (1.3%)
50-64	27 (1.3%)	15 (1.6%)	18 (2.5%)	28 (1.3%)	21 (1.9%)	13 (1.2%)	12 (1.1%)	25 (2.2%)	10 (0.9%)	5 (0.4%)	9 (0.8%)	9 (0.8%)
	22 (1.2%)	14 (1.1%)	12 (1.9%)	22 (1.2%)	19 (1.7%)	15 (1.3%)	15 (1.3%)	26 (2.3%)	8 (0.7%)	4 (0.4%)	8 (0.7%)	9 (0.8%)

[104] The data were collected by Format Research S.r.l..

4.3 Descriptive analyses[105]

Table 4.3: Summary of sociodemographic characteristic of the sample.

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Gender (D.1.)				
	Male	268 23.72 48.03	290 25.66 51.97	558 49.38
	Female	276 24.42 48.25	296 26.19 51.75	572 50.62
Age (D.2.)				
	20-34	173 15.31 47.40	192 16.99 52.60	365 32.30
	35-49	192 16.99 48.12	207 18.32 51.88	399 35.31
	50-64	179 15.84 48.91	187 16.55 51.09	366 32.39
Geographical area (D.3.)				
	NW	147 13.01 50.87	142 12.57 49.13	289 25.58
	NE	87 7.70 46.52	100 8.85 53.48	187 16.55
	CE	121 10.71 43.37	158 13.98 56.63	279 24.69
	SO	189 16.73 50.40	186 16.46 49.60	375 33.19

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Marital status (D.4.)				
	single	199 17.61 49.38	204 18.05 50.62	403 35.66
	cohabitant	71 6.28 43.03	94 8.32 56.97	165 14.60
	married	234 20.71 47.95	254 22.48 52.05	488 43.19
	divorced	35 3.10 56.45	27 2.39 43.55	62 5.49
	widower	5 0.44 41.67	7 0.62 58.33	12 1.06
Education (D.6.)				
	low	149 13.19 45.57	178 15.75 54.43	327 28.94
	medium	242 21.42 49.59	246 21.77 50.41	488 43.19
	high	153 13.54 48.57	162 14.34 51.43	315 27.88

(table continues in the next page)

[105] In the following tables and figures, the value in parentheses (e.g. D.1.) represents the question number in the Questionnaires to which the linked variable is obtained.

Frequency Percent Row Pct	T		
	WI	PS	Total
Field of education (D.6.1.)			
biology, chemistry or medicine	32 10.16 53.33	28 8.89 46.67	60 19.05
law	16 5.08 39.02	25 7.94 60.98	41 13.02
economics	32 10.16 32.65	66 20.95 67.35	98 31.11
humanities, philosophy, or psychology	30 9.52 55.56	24 7.62 44.44	54 17.14
physic, engineering, math, or statistics	32 10.16 72.73	12 3.81 27.27	44 13.97
other	11 3.49 61.11	7 2.22 38.89	18 5.71
Work condition (D.7.)			
employed	341 30.18 47.83	372 32.92 52.17	713 63.10
of which (D.8.)			
employee	269 37.73 47.53	297 41.65 52.47	566 79.38
occasional work	19 2.66 43.18	25 3.51 56.82	44 6.17
self-employed	53 7.43 51.46	50 7.01 48.54	103 14.45
not employed	106 9.38 50.96	102 9.03 49.04	208 18.41
of which (D.8.)			
short term	13 6.25 52.00	12 5.77 48.00	25 12.02
medium term	21 10.10 51.22	20 9.62 48.78	41 19.71
long term	62 29.81 52.54	56 26.92 47.46	118 56.73
first job	10 4.81 41.67	14 6.73 58.33	24 11.54
inactive	97 8.58 46.41	112 9.91 53.59	209 18.50
of which (D.8.)			
student	34 16.27 44.74	42 20.10 55.26	76 36.36
housewife	45 21.53 48.91	47 22.49 51.09	92 44.02
retired	13 6.22 46.43	15 7.18 53.57	28 13.40
not looking for a job	5 2.39 38.46	8 3.83 61.54	13 6.22

(table continues in the next page)

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Number of people with stable income in the family (D.9.)				
	0	1 0.09 16.67	5 0.44 83.33	6 0.53
	1	279 24.69 49.64	283 25.04 50.36	562 49.73
	2	210 18.58 47.73	230 20.35 52.27	440 38.94
	3	40 3.54 43.01	53 4.69 56.99	93 8.23
	4	11 0.97 45.83	13 1.15 54.17	24 2.12
	5	3 0.27 60.00	2 0.18 40.00	5 0.44
Who has the higher income in the family (D.10.)				
	myself[106]	247 21.86 47.50	273 24.16 52.50	520 46.02
	partner	164 14.51 48.24	176 15.58 51.76	340 30.09
	parent	119 10.53 49.17	123 10.88 50.83	242 21.42
	son/daughter	8 0.71 53.33	7 0.62 46.67	15 1.33
	other	6 0.53 46.15	7 0.62 53.85	13 1.15

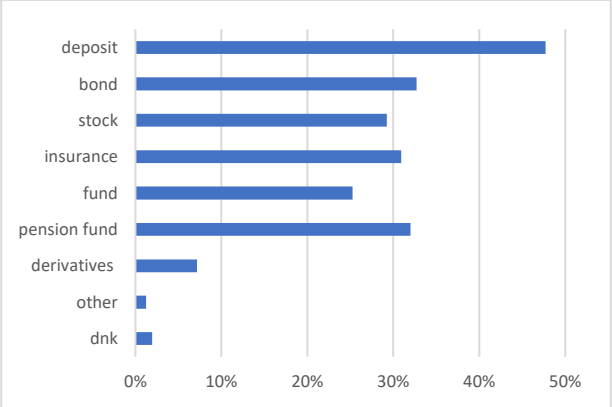
Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Savers (D.11.)				
	no income	73 6.46 48.99	76 6.73 51.01	149 13.19
	no	241 21.33 48.10	260 23.01 51.90	501 44.34
	yes	230 20.35 47.92	250 22.12 52.08	480 42.48
of which (D.12.)				
	Save less than 10%	85 17.71 45.95	100 20.83 54.05	185 38.54
	Save between 10% and 30%	121 25.21 50.21	120 25.00 49.79	241 50.21
	Save more than 30%	24 5.00 44.44	30 6.25 55.56	54 11.25

[106] These respondents were considered as household heads.

Table 4.4: Investment habits and experience.

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Updating about economics and financial news (A.1.)				
	never	104 9.20 53.33	91 8.05 46.67	195 17.26
	rarely	200 17.70 45.45	240 21.24 54.55	440 38.94
	sometimes	121 10.71 45.49	145 12.83 54.51	266 23.54
	daily	119 10.53 51.97	110 9.73 48.03	229 20.27
Past investment experience (A.2.)				
	No	198 17.52 48.89	207 18.32 51.11	405 35.84
	Yes	346 30.62 47.72	379 33.54 52.28	725 64.16

Figure 4.1: Types of investment (A.3.).



Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Considering who had past experience				
Gain or losses? (A.4.)				
	loss	51 7.03 48.11	55 7.59 51.89	106 14.62
	nor gain nor losses	149 20.55 48.22	160 22.07 51.78	309 42.62
	gain	146 20.14 47.10	164 22.62 52.90	310 42.76
Advice seeking (A.5.)				
	not advised	136 18.76 44.88	167 23.03 55.12	303 41.79
	advised	210 28.97 49.76	212 29.24 50.24	422 58.21
of which were advised by (A.6.)				
	family/friends	60 14.22 55.56	48 11.37 44.44	108 25.59
	old advisor	132 31.28 47.48	146 34.60 52.52	278 65.88
	new advisor	13 3.08 52.00	12 2.84 48.00	25 5.92
	media	2 0.47 40.00	3 0.71 60.00	5 1.18
	other	3 0.71 50.00	3 0.71 50.00	6 1.42

Table 4.5:Answers to the questions about financial knowledge[107].

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Interest rate (B.1.)				
	Do not know	60 5.31 56.07	47 4.16 43.93	107 9.47
	less than 102	54 4.78 54.00	46 4.07 46.00	100 8.85
	exactly 102	54 4.78 40.00	81 7.17 60.00	135 11.95
	more than 102	376 33.27 47.72	412 36.46 52.28	788 69.73
Inflation (B.2.)				
	Do not know	103 9.12 52.55	93 8.23 47.45	196 17.35
	equal	68 6.02 49.28	70 6.19 50.72	138 12.21
	more	54 4.78 51.92	50 4.42 48.08	104 9.20
	less	319 28.23 46.10	373 33.01 53.90	692 61.24
Investment fund (B.3.)				
	Do not know	179 15.84 49.45	183 16.19 50.55	362 32.04
	true	66 5.84 55.93	52 4.60 44.07	118 10.44
	false	299 26.46 46.00	351 31.06 54.00	650 57.52

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Diversification (B.4.)				
	Do not know	119 10.53 49.17	123 10.88 50.83	242 21.42
	different rate	54 4.78 45.38	65 5.75 54.62	119 10.53
	same risk	118 10.44 51.98	109 9.65 48.02	227 20.09
	different sector	253 22.39 46.68	289 25.58 53.32	542 47.96
Prospects (B.5.)				
	Do not know	107 9.47 49.31	110 9.73 50.69	217 19.20
	no sufficient info	154 13.63 51.33	146 12.92 48.67	300 26.55
	Investment 1	74 6.55 49.66	75 6.64 50.34	149 13.19
	Investment 2	209 18.50 45.04	255 22.57 54.96	464 41.06

[107] The underlined option represents the correct answer.

Figure 4.2: Number of correct answers about financial knowledge by attributes.

Table 4.6: Measures of financial literacy.

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Financial literacy (sum of correct answers)				
	0	51 4.51 52.04	47 4.16 47.96	98 8.67
	1	77 6.81 50.00	77 6.81 50.00	154 13.63
	2	114 10.09 50.22	113 10.00 49.78	227 20.09
	3	127 11.24 50.80	123 10.88 49.20	250 22.12
	4	105 9.29 46.26	122 10.80 53.74	227 20.09
	5	70 6.19 40.23	104 9.20 59.77	174 15.40
Financial literacy (dummy)				
	No	474 41.95 49.58	482 42.65 50.42	956 84.60
	Yes	70 6.19 40.23	104 9.20 59.77	174 15.40

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Financial literacy (by class)				
	Low	128 11.33 50.79	124 10.97 49.21	252 22.30
	Medium	241 21.33 50.52	236 20.88 49.48	477 42.21
	High	175 15.49 43.64	226 20.00 56.36	401 35.49

Table 4.7: Investment choice between Product I and Product II.

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
	Do not know	191 16.90 52.91	170 15.04 47.09	361 31.95
	Product I	236 20.88 62.77	140 12.39 37.23	376 33.27
	Product II	117 10.35 29.77	276 24.42 70.23	393 34.78
	Total	544 48.14	586 51.86	1130 100

Table 4.8: Randomization of the order of products presentation.

Frequency Row Pct		T		
		WI	**PS**	**Total**
Number of cases				
	A > B	222 49.66	225 50.34	447
	B > A	322 47.14	361 52.86	683
	Total	544 48.14	586 51.86	1130

Table 4.9: Investment decision controlling for order randomization.

Frequency Percent Row Pct	T		
	WI	**PS**	**Total**
A > B			
Investment choice (C.1.)			
Do not know	87 19.46 51.79	81 18.12 48.21	168 37.58
A (i.e. Product II)	35 7.83 31.25	77 17.23 68.75	112 25.06
B (i.e. Product I)	100 22.37 59.88	67 14.99 40.12	167 37.36
Total	222 49.66	225 50.34	447 100

Frequency Percent Row Pct	T		
	WI	**PS**	**Total**
B > A			
Investment choice (C.1.)			
Do not know	104 15.23 53.89	89 13.03 46.11	193 28.26
A (i.e. Product I)	136 19.91 65.07	73 10.69 34.93	209 30.60
B (i.e. Product II)	82 12.01 29.18	199 29.14 70.82	281 41.14
Total	322 47.14	361 52.86	683 100

Table 4.10: Coherence with the Stochastic dominance.

Frequency		T		
		WI	PS	Total
Investment decision				
	wrong	236	140	376
	correct	117	276	393
	Total	353	416	769
Frequency Missing = 361 participants answering 'do not know'				

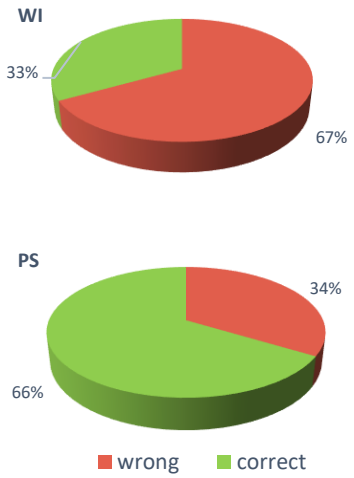
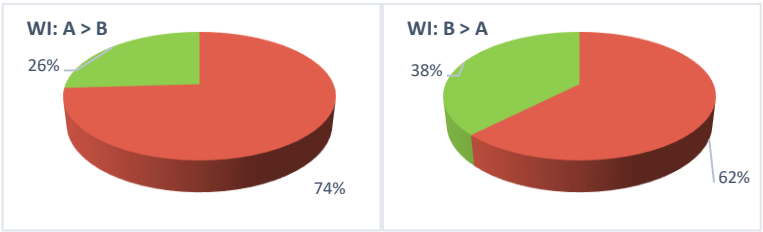


Table 4.11: Coherence with the Stochastic dominance controlling for order randomization.

Frequency		T		
A > B		WI	PS	Total
Investment decision				
	wrong	100	67	167
	correct	35	77	112
	Total	135	144	279
Frequency Missing = 168 participants answering 'do not know'				

Frequency		T		
B > A		WI	PS	Total
Investment decision				
	wrong	136	73	209
	correct	82	199	281
	Total	218	272	490
Frequency Missing = 193 participants answering 'do not know'				



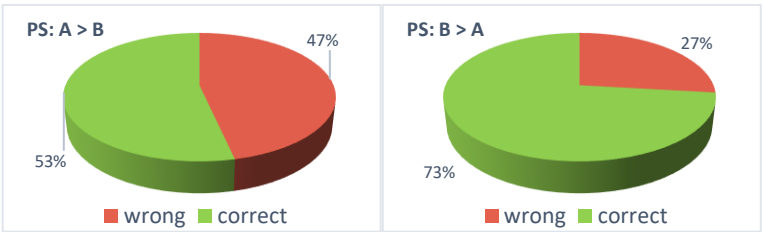


Figure 4.3: Risk assessment considering product labels linked to stochastic dominance (C.3.).

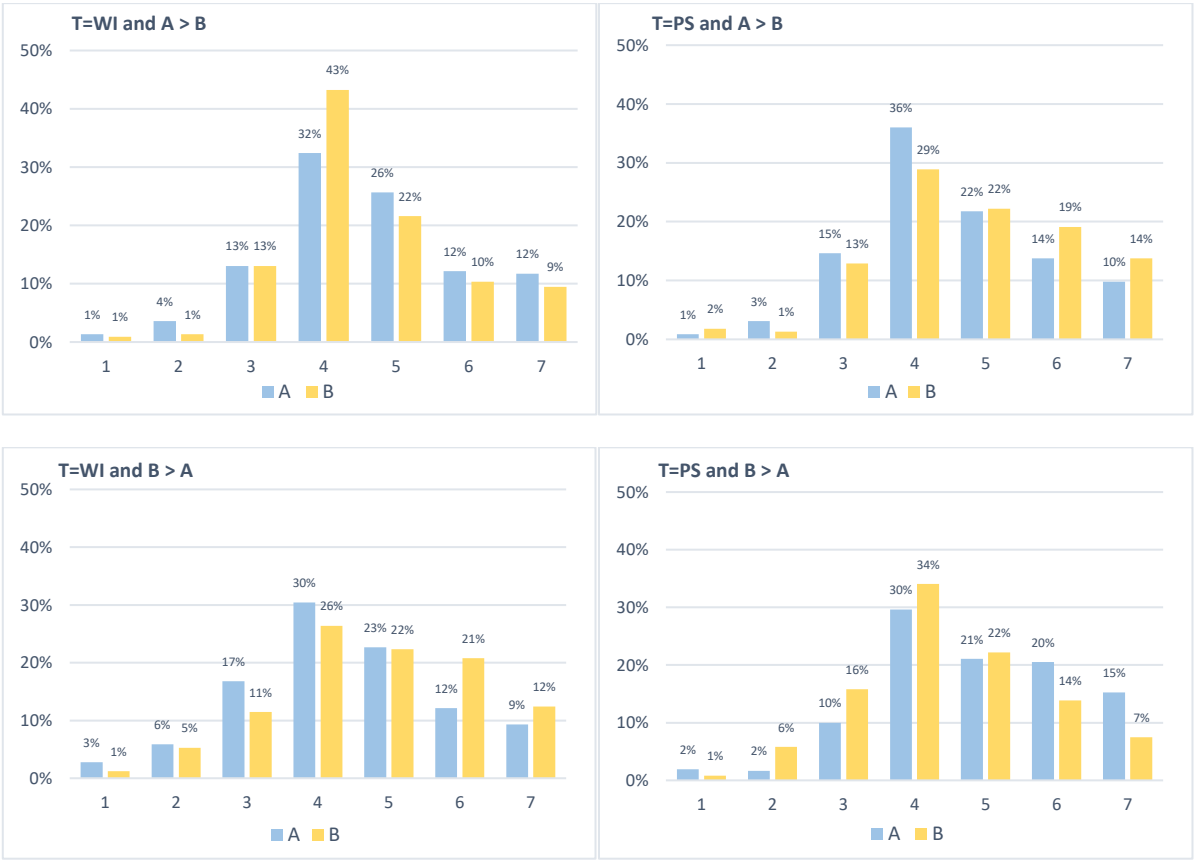
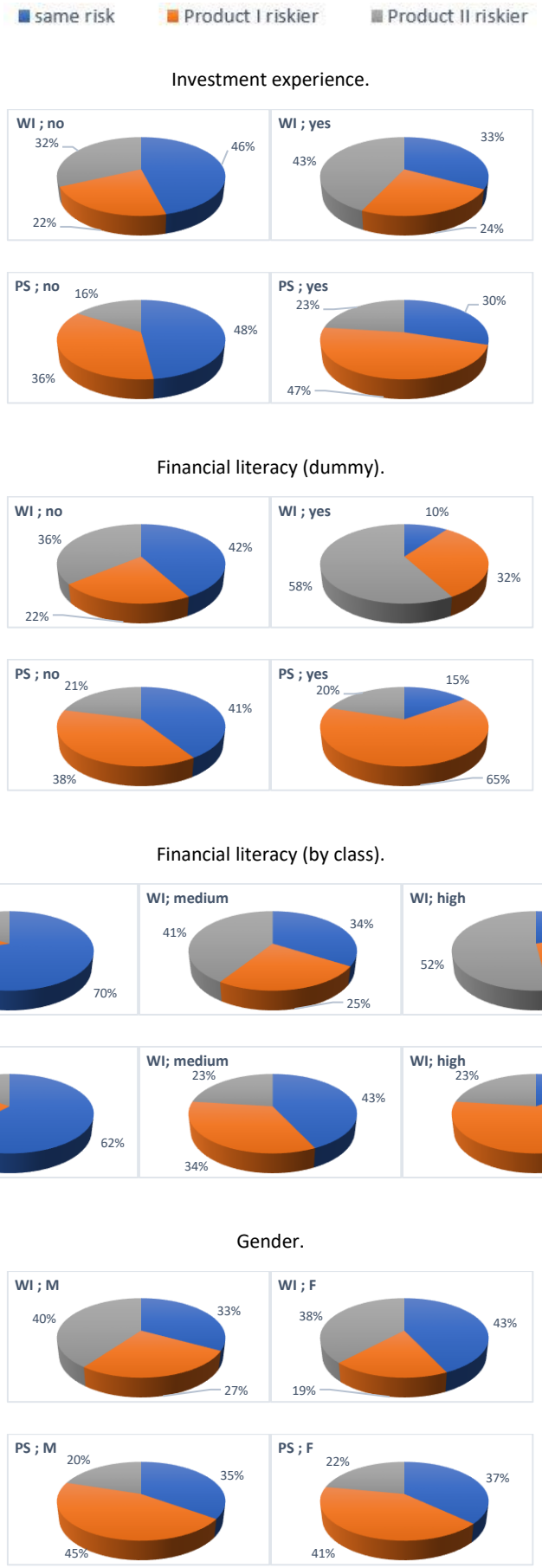


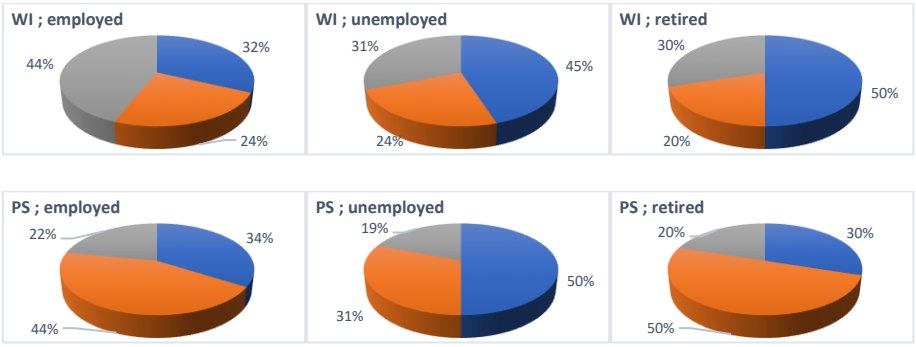
Table 4.12: Evaluation of difference in risk between Product I and Product II.

Frequency Percent Row Pct		T		
		WI	PS	Total
Evaluation of difference in risk				
	Wrong	419 37.08 55.57	335 29.65 44.43	754 66.73
	Correct	125 11.06 33.24	251 22.21 66.76	376 33.27
	Total	544 48.14	586 51.86	1130 100

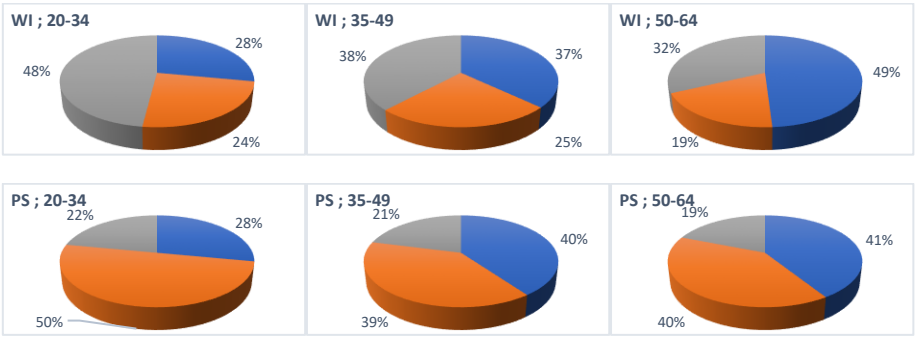
Figure 4.4: Risk assessment by treatments and attributes.



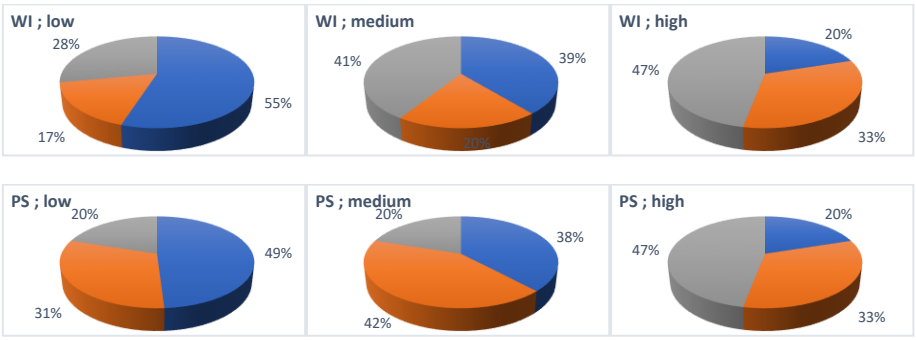
Job status.



Age.



Education level.



Geographical area.

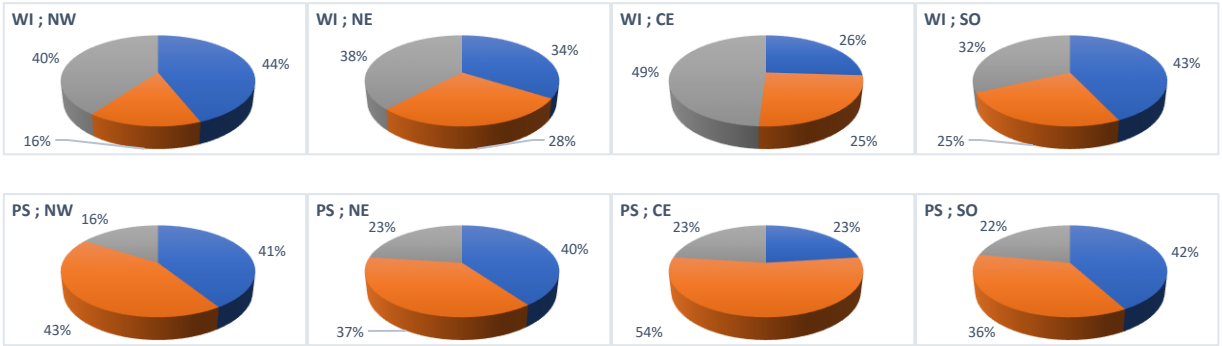


Table 4.13: Product choices controlling for risk assessment.

Frequency Percent Row Pct		Product I riskier			Product II riskier			same risk		
		T			T			T		
		WI	PS	Total	WI	PS	Total	WI	PS	Total
Investment decision										
	Do not know	24 6.38 48.98	25 6.65 51.02	49 13.03	30 9.01 66.67	15 4.50 33.33	45 13.51	137 32.54 51.31	130 30.88 48.69	267 63.42
	Product I	44 11.70 61.11	28 7.45 38.89	72 19.15	147 44.14 67.12	72 21.62 32.88	219 65.77	45 10.69 52.94	40 9.50 47.06	85 20.19
	Product II	57 15.16 22.35	198 52.66 77.65	255 67.82	34 10.21 49.28	35 10.51 50.72	69 20.72	26 6.18 37.68	43 10.21 62.32	69 16.39
	Total	125 33.24	251 66.76	376 100	211 63.36	122 36.64	333 100	208 49.41	213 50.59	421 100

Figure 4.5: Distributions of score attribute by treatment (C.4.).

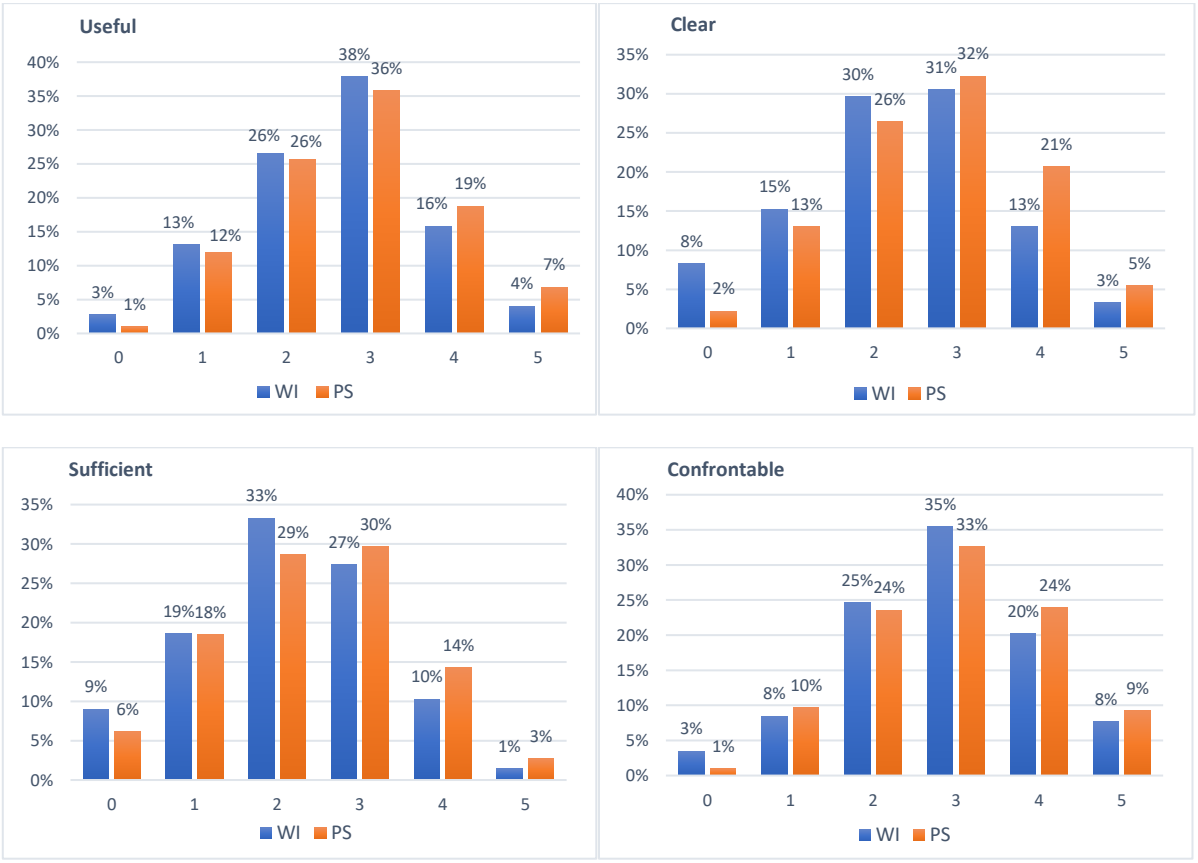


Figure 4.6: Average framing evaluation.

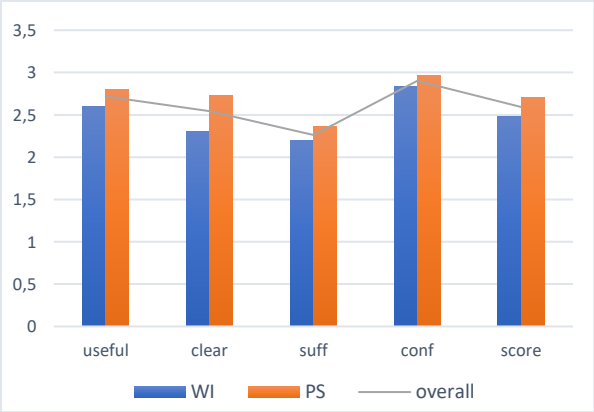


Table 4.14: Total score.

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Score by class				
	low	101 8.94 56.42	78 6.90 43.58	179 15.84
	medium	344 30.44 48.25	369 32.65 51.75	713 63.10
	high	99 8.76 41.60	139 12.30 58.40	238 21.06
	Total	544 48.14	586 51.86	1130 100

Table 4.15: Framing preference.

Frequency Percent Row Pct		T		
		WI	**PS**	**Total**
Which frame do you prefer? (C.5.)				
	Do not know	154 13.63 47.09	173 15.31 52.91	327 28.94
	WI	74 6.55 35.75	133 11.77 64.25	207 18.32
	PS	316 27.96 53.02	280 24.78 46.98	596 52.74
	Total	544 48.14	586 51.86	1130 100

Table 4.16: Differences in scoring between sociodemographic characteristics by treatments.

	Mann-Whitney Wilcoxon/Kruskal Wallis Pr > Z (two-sided)														
	Score			Useful			Clear			Sufficient			Confrontable		
X	Sample	WI	PS	Sample	WI	PS	Sample	WI	PS	Sample	WI	PS	Sample	WI	PS
gender	*			*						**	*				
geo	**		**	**		*	*		*	*					
age	***		***	**		**	***		***	*		*	***	*	**
edu															
job													**		
exp				*		*									
HH															
FL_01	***		***	***		***	***		***	**		**	***		***
FL_class	***	**	***	***	**	***	**	*	**			*	***	**	***

Note: The non-parametric Kruskal-Wallis two-sided test for the median was performed on the entire sample and two univocal sub-samples (i.e. Wi and PS). The test aimed to analyse the impact of sociodemographic variables (i.e. each raw in the table) on the distribution of *y*. The output variables were represented by the overall score and the rank given to usefulness, clearness, sufficiency and comparability. The notations *, ** and *** indicate the sign of a 5% , 2.5% and 1% statistically significant difference between the two groups given by the treatment's variable (i.e. p-value < 0.05, p-value < 0.025 and p-value < 0.01). Blank cells indicate that the difference is not statistically significant. Only two-sided test was performed with the following null hypothesis: 'considering *x*, the median of *y* across categories of *x* is homogeneous'. For example, considering the sub-sample of PS, the median of the distribution of the overall score, usefulness and clearness are statistically different among geographical areas or age whereas they are not when the gender is considered.

Table 4.17: Differences in mean scoring between treatments by levels of sociodemographic variables.

	Mann-Whitney Wilcoxon Pr > |Z| (two-sided)					**Mann-Whitney Wilcoxon Pr < Z (one-sided)**				
	Score	Useful	Clear	Sufficient	Confrontable	Score	Useful	Clear	Sufficient	Confrontable
Gender										
M	*	*	***			**	*	***		
F	*		***	*		**		***	**	
GEO										
NW										
NE			*			*		*		
CE	*		*	**		**		**	**	
SO	*		***			**		***		*
Age										
20-34	**		***	*		**	*	***	*	
35-49	**	*	**			**	*	***	*	*
50-64								*		
Edu										
Low	*		***			*		***		
Medium			**					**		
High	**	*	*	**		**	*	*	**	
job										
Employed	*	*	**			**	*	***	*	
Unemployed								*		
Inactive	*		***			*		***	*	
exp										
No			**			*		***		
Yes	**	*	***	**		**	*	***	**	
FL_class										
Low	***	*	***	**	*	***	**	***	**	**
Medium										
High	**		**	*		**	*	**	*	
HH										
No	**		***	*		**		***	**	
Yes	**	*	***			**	*	***	*	

Note: The non-parametric Mann-Whitney Wilcoxon tests was performed on subpopulations represented by each level of the selected sociodemographic variables (i.e. each raw in the table). The notations *, ** and *** indicate the sign of a 5% , 2.5% and 1% statistically significant difference between the two groups given by the treatment variable (i.e. p-value < 0.05, p-value < 0.025 and p-value < 0.01). Blank cells indicate that the difference is not statistically significant. Both two-sided and one-sided tests were performed. For the one-sided test the null hypothesis is the following: 'mean of y in Group 0 (i.e. WI) is higher than mean of y in Group 1 (i.e. PS).' The output variable y is represented by the overall score and the rank given to usefulness, clearness, sufficiency and comparability. For example, considering the sub-sample of male, the mean of the distribution of the overall score, usefulness and clearness are statistically different between What-if and Probabilistic scenarios (two-sided) and, in particular, they are higher in Group 1 (one-sided).

Table 4.18: Differences in median scoring between treatments by levels of sociodemographic variables.

	Median Pr>|Z|(two-sided)					**Median Pr<Z (one-sided)**				
	Score	Useful	Clear	Sufficient	Confrontable	Score	Useful	Clear	Sufficient	Confrontable
Gender										
M			**			*		**		
F			**	*				**	*	
GEO										
NW										
NE						*		*		
CE				*				*	*	
SO	*		***			*		***	*	*
Age										
20-34	*		**			*		***	*	
35-49			*					**	*	
50-64										
Edu										
Low			**					**		
Medium			*					**		
High		*		**			*		**	
job										
Employed			*					**	*	
Unemployed								*		
Inactive			**			*	*	***		
exp										
No			***					***		
Yes			***	*				***	*	
FL_class										
Low	*	*	**	*		**	**	***	**	*
Medium										
High	**		***	**		***		***	***	*
HH										
No		*	***					***	*	
Yes			**					**		

Note: The non-parametric Median tests were performed on subpopulations represented by each level of the selected sociodemographic variables (i.e. each raw in the table). The notations *, ** and *** indicate the sign of a 5% , 2.5% and 1% statistically significant difference between the two groups given by the treatment variable (i.e. p-value < 0.05, p-value < 0.025 and p-value < 0.01). Blank cells indicate that the difference is not statistically significant. Both two-sided and one-sided tests were performed. For the one-sided test the null hypothesis is the following: 'median of *y* in Group 0 (i.e. WI) is higher than median of *y* in Group 1 (i.e. PS).' The output variable *y* is represented by the overall score and the rank given to usefulness, clearness, sufficiency and comparability. For example, considering the sub-sample of male, the median of the distribution of clearness is statistically different between What-if and Probabilistic scenarios (two-sided) and, in particular, they are higher in Group 1 (one-sided).

Table 4.19: Differences in distribution between treatments by levels of sociodemographic variables.

	Kolmogorov-Smirnov two-sample test (Pr>KSa)				
	Score	Useful	Clear	Sufficient	Confrontable
Gender					
M					
F			*		
GEO					
NW					
NE					
CE					
SO			***		
Age					
20-34			**		
35-49					
50-64					
Edu					
Low			*		
Medium					
High	**			*	
job					
Employed			*		
Unemployed					
Inactive			**		
exp					
No					
Yes	**		**		
FL_class					
Low	***		***		
Medium					
High	**				
HH					
No			**		
Yes	*				

Note: The non-parametric Kolmogorov-Smirnov two-sample test was performed on subpopulations represented by each level of the selected sociodemographic variables (i.e. each raw in the table). The notations *, ** and *** indicate the sign of a 5% , 2.5% and 1% statistically significant difference in distribution between two groups given by the treatment variable (i.e. p-value < 0.05, p-value < 0.025 and p-value < 0.01). Blank cells indicate that the difference is not statistically significant. Only two-sided test was performed with the following null hypothesis: 'considering the subpopulation, the distribution of *y* is homogeneous between treatments'. The output variables *y* is represented by the overall score and the rank given to usefulness, clearness, sufficiency and comparability. For example, considering the sub-sample of female, the distribution of clearness is statistically different between What-if and Probabilistic scenarios.

4.4 Estimation results

Table 4.20: Mantel-Haenszel Statistical test for independence.

X	Y	Strata	p-value	Sign.
Gender	riskier	T	0.0264	*
Geo	riskier	T	0.0205	**
Age	riskier	T	0.0675	
Edu	riskier	T	<0.0001	***
Marital status	riskier	T	0.0375	*
job	riskier	T	0.1693	
Household Head	riskier	T	0.6682	
Exp	riskier	T	0.0209	**
News	riskier	T	0.0006	***
FL (dico)	riskier	T	<0.0001	***
FL_	riskier	T	<0.0001	***
R	riskier	T	0.0216	**

Note: In general, for categorical variables statistical analysts are interested in whether there are overall differences in proportions with *Y* (Pearson's Chi-squared test). However, they might be concerns that the populations at the different level of *X* are sufficiently different. By a stratified analysis, it is possible to examine the association between two variables while adjusting for the effects of others. In general, the strata may represent explanatory variables and each contingency table (not reported in this Section) corresponds to one stratum such as the strata are determined by the levels of the explanatory variables (one for each unique combination of the levels of the explanatory variables). The idea is to evaluate the association between the row variable and the response variable, while adjusting, or controlling, for the effects of the stratification variables. In this analysis, the treatment (*T*) is used as stratification variable. In some cases, sociodemographic characteristics resulting from the study design, such as age or gender, are used as independent variables; whereas in other cases, they may arise from poststudy variables, such as the financial literacy score or the past experience in financial investment, that are thought to be related to the response variable (e.g. the risk assessment). The Mantel-Haenszel strategy potentially removes the confounding influence of the explanatory variables that comprise the stratification. In addition, it can provide increased power for detecting association in a randomized study by comparing subjects. It can also remove the bias that can occur in an observational study from imbalances in confounding factors, but possibly at the cost of decreased power. In some sense, the strategy is similar to adjustment for blocks in a two-way analysis of variance for randomized blocks; it is also like covariance adjustment for a categorical explanatory variable. In the table above, the notations *, ** and *** indicate the sign of a 5% , 2.5% and 1% statistically significant difference in distribution between two groups given by the treatment variable (i.e. p-value < 0.05, p-value < 0.025 and p-value < 0.01). Blank cells indicate that the difference is not statistically significant.

For sake of clarity, in order to simplify the multivariate analyses, some variables were reclassified as follows:

- the variable 'News' came from Question A.1. and it was dichotomised as follows: 0 if the answer was 'no' or 'yes, rarely'; 1 otherwise.
- the variable 'marital status' was obtained from Question D.4. and the recorded answers were aggregated in the following categories: 'single', if the answer was 'single'; 'married', if the answer was 'cohabitant' or 'married'; and 'separated', if the answer was 'divorced' or 'widow'.

Table 4.21: Estimation results from logistic regression.

	Dependent variable: riskier						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
T	0.921***	0.953***	0.930***	0.930***	0.677**	0.733***	1.134***
	(0.132)	(0.135)	(0.137)	(0.137)	(0.338)	(0.188)	(0.238)
gender		-0.279**	-0.166	-0.165	-0.163	-0.411*	-0.163
		(0.132)	(0.136)	(0.142)	(0.137)	(0.213)	(0.143)
GEO_NO		-0.017	-0.142	-0.142	-0.133	-0.142	-0.141
		(0.178)	(0.182)	(0.182)	(0.183)	(0.182)	(0.183)
GEO_NE		0.088	0.034	0.033	0.058	0.034	0.024
		(0.199)	(0.203)	(0.204)	(0.204)	(0.203)	(0.204)
GEO_CE		0.278	0.140	0.140	0.139	0.136	0.129
		(0.175)	(0.179)	(0.179)	(0.181)	(0.179)	(0.180)
EDUC_medium		0.325*	0.198	0.197	0.190	0.200	0.199
		(0.169)	(0.173)	(0.173)	(0.174)	(0.173)	(0.174)
EDUC_high		0.911***	0.583***	0.582***	0.575***	0.582***	0.579***
		(0.186)	(0.197)	(0.198)	(0.198)	(0.197)	(0.198)
age_medium		-0.072	-0.044	-0.045	-0.047	-0.052	0.243
		(0.161)	(0.164)	(0.170)	(0.165)	(0.165)	(0.255)
age_high		-0.084	-0.131	-0.132	-0.139	-0.140	-0.088
		(0.171)	(0.175)	(0.182)	(0.176)	(0.175)	(0.276)
FL_medium			0.556***	0.556***	0.628**	0.568***	0.615**
			(0.199)	(0.199)	(0.304)	(0.200)	(0.263)
FL_high			1.165***	1.165***	0.695**	1.179***	1.249***
			(0.207)	(0.207)	(0.319)	(0.208)	(0.269)
HH				0.002			0.155
				(0.149)			(0.344)
T:FL_medium					-0.151		
					(0.396)		
T:FL_high					0.789*		
					(0.403)		
T:gender						0.415	
						(0.275)	
FL_medium:HH							-0.155
							(0.398)
FL_high:HH							-0.200
							(0.400)
T:age_medium							-0.493
							(0.328)
T:age_high							-0.083
							(0.344)
Constant	-1.210***	-1.533***	-2.028***	-2.029***	-1.871***	-1.921***	-2.204***
	(0.102)	(0.228)	(0.266)	(0.269)	(0.322)	(0.275)	(0.323)
Observations	1,130	1,130	1,130	1,130	1,130	1,130	1,130
Log Likelihood	-693.364	-671.574	-653.420	-653.420	-648.156	-652.275	-651.988
Akaike Inf. Crit.	1,390.728	1,363.148	1,330.840	1,332.840	1,324.312	1,330.551	1,337.975
Note:			*p-value < 0.1; **p-value < 0.05; ***p- value < 0.01				

(table continues in the next page)

	Dependent variable: riskier						
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
T	1.126***	1.130***	0.705**	0.931***	0.949***	0.933***	0.931***
	(0.238)	(0.238)	(0.275)	(0.137)	(0.138)	(0.137)	(0.136)
gender	-0.415**	-0.162	-0.163	-0.168	-0.182	-0.144	-0.220
	(0.193)	(0.136)	(0.136)	(0.136)	(0.140)	(0.139)	(0.135)
GEO_NO	-0.127	-0.140	-0.144	-0.142	-0.114	-0.140	-0.068
	(0.183)	(0.182)	(0.182)	(0.182)	(0.185)	(0.183)	(0.179)
GEO_NE	0.028	0.027	0.038	0.033	0.056	0.026	0.091
	(0.205)	(0.203)	(0.203)	(0.203)	(0.207)	(0.203)	(0.200)
GEO_CE	0.133	0.132	0.135	0.147	0.143	0.153	0.234
	(0.180)	(0.180)	(0.180)	(0.183)	(0.186)	(0.183)	(0.180)
EDUC_medium	0.186	0.198	-0.018	0.197	0.216	0.188	0.311*
	(0.174)	(0.173)	(0.278)	(0.173)	(0.176)	(0.174)	(0.169)
EDUC_high	0.559***	0.581***	0.423	0.587***	0.626***	0.566***	0.775***
	(0.198)	(0.197)	(0.292)	(0.198)	(0.207)	(0.199)	(0.193)
age_medium	0.268	0.243	-0.048	-0.047	0.063	-0.052	-0.065
	(0.256)	(0.252)	(0.165)	(0.165)	(0.182)	(0.165)	(0.163)
age_high	-0.068	-0.089	-0.139	-0.134	-0.076	-0.139	-0.077
	(0.276)	(0.272)	(0.175)	(0.175)	(0.197)	(0.177)	(0.173)
job_unemployed					-0.003		
					(0.196)		
job_inactive					0.231		
					(0.191)		
maritalstatus_married					-0.222		
					(0.166)		
maritalstatus_separated					0.401		
					(0.300)		
FL_medium	0.561***	0.553***	0.562***	0.557***	0.567***	0.543***	
	(0.200)	(0.199)	(0.199)	(0.199)	(0.200)	(0.201)	
FL_high	1.192***	1.165***	1.171***	1.168***	1.191***	1.140***	
	(0.208)	(0.207)	(0.207)	(0.208)	(0.209)	(0.214)	
HH	-0.260						
	(0.201)						
gender:HH	0.550						
	(0.281)						
T:age_medium	-0.485	-0.492					
	(0.329)	(0.328)					
T:age_high	-0.084	-0.077					
	(0.345)	(0.344)					
T:EDUC_medium			0.338				
			(0.347)				
T:EDUC_high			0.245				
			(0.365)				
news						0.118	
						(0.142)	
exp						0.024	
						(0.152)	
FL_dummy							0.614***
							(0.184)
R				-0.030	-0.040	-0.028	-0.006
				(0.149)	(0.151)	(0.149)	(0.148)
Constant	-2.008***	-2.143***	-1.881***	-1.982***	-2.013***	-2.038***	-1.576***
	(0.303)	(0.293)	(0.305)	(0.353)	(0.376)	(0.360)	(0.329)
Observations	1,130	1,130	1,130	1,130	1,130	1,130	1,130
Log Likelihood	-650.200	-652.114	-652.942	-653.400	-649.369	-653.014	-665.982
Akaike Inf. Crit.	1,332.401	1,332.228	1,333.884	1,332.801	1,332.738	1,336.029	1,355.963
Note:				*p-value < 0.1; **p-value < 0.05; ***p- value < 0.01			

Table 4.22: Estimation of fixed effects from multilevel logistic regression with random intercept.

	Dependent variable: riskier						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
gender		-0.279**	-0.165	-0.168	-0.631*	-0.159	-0.168
		(0.132)	(0.136)	(0.136)	(0.336)	(0.136)	(0.136)
GEO_NO		-0.017	-0.141	-0.143	-0.156	-0.147	-0.142
		(0.177)	(0.182)	(0.182)	(0.182)	(0.182)	(0.182)
GEO_NE		0.089	0.035	0.022	0.034	0.050	0.034
		(0.199)	(0.202)	(0.203)	(0.203)	(0.203)	(0.202)
GEO_CE		0.280	0.142	0.137	0.138	0.160	0.149
		(0.174)	(0.179)	(0.180)	(0.179)	(0.180)	(0.183)
EDUC_medium		0.323*	0.195	0.148	0.196	0.185	0.194
		(0.168)	(0.173)	(0.371)	(0.173)	(0.173)	(0.173)
EDUC_high		0.908***	0.578***	1.043**	0.569***	0.566***	0.583***
		(0.186)	(0.196)	(0.483)	(0.197)	(0.197)	(0.198)
age_medium		-0.072	-0.045	-0.046	-0.044	-0.444	-0.047
		(0.161)	(0.164)	(0.164)	(0.164)	(0.392)	(0.165)
age_high		-0.084	-0.132	-0.132	-0.134	-0.629	-0.135
		(0.171)	(0.174)	(0.175)	(0.175)	(0.436)	(0.175)
FL_medium			0.555***	0.570*	0.212	0.086	0.557***
			(0.199)	(0.330)	(0.287)	(0.340)	(0.199)
FL_high			1.166***	1.337***	0.904***	0.919***	1.169***
			(0.207)	(0.361)	(0.287)	(0.336)	(0.208)
EDUC_medium:FL_medium				0.127			
				(0.446)			
EDUC_high:FL_medium				-0.456			
				(0.560)			
EDUC_medium:FL_high				-0.076			
				(0.476)			
EDUC_high:FL_high				-0.617			
				(0.561)			
gender:FL_medium					0.629		
					(0.395)		
gender:FL_high					0.480		
					(0.397)		
age_medium:FL_medium						0.731	
						(0.466)	
age_high:FL_medium						0.653	
						(0.507)	
age_medium:FL_high						0.211	
						(0.467)	
age_high:FL_high						0.541	
						(0.501)	
R							-0.029
							(0.149)
Constant	-0.747**	-1.053***	-1.560***	-1.610***	-1.294***	-1.268***	-1.514***
	(0.326)	(0.392)	(0.409)	(0.445)	(0.441)	(0.458)	(0.470)
Observations	1,130	1,130	1,130	1,130	1,130	1,130	1,130
Log Likelihood	-697.585	-675.822	-657.588	-656.787	-656.318	-655.471	-657.569
Akaike Inf. Cri	1,399.169	1,371.644	1,339.176	1,345.573	1,340.636	1,342.942	1,341.138
Bayesian Inf. Crit.	1,409.229	1,421.944	1,399.536	1,426.053	1,411.056	1,423.422	1,406.528
Note:			*p-value < 0.1; **p-value < 0.05; ***p- value < 0.01				

(table continues in the next page)

	Dependent variable: riskier						
	(8)	(9)	(10)	(11)	(12)	(13)	(CIM)
gender	-0.165	-0.168	-0.423**	-0.181	-0.144	-0.219	
	(0.142)	(0.142)	(0.192)	(0.140)	(0.139)	(0.134)	
GEO_NO	-0.142	-0.142	-0.129	-0.114	-0.139	-0.068	
	(0.182)	(0.182)	(0.183)	(0.185)	(0.183)	(0.179)	
GEO_NE	0.034	0.033	0.036	0.057	0.027	0.092	
	(0.204)	(0.204)	(0.205)	(0.207)	(0.203)	(0.199)	
GEO_CE	0.142	0.140	0.143	0.145	0.155	0.236	
	(0.179)	(0.179)	(0.179)	(0.186)	(0.183)	(0.180)	
EDUC_medium	0.195	0.196	0.183	0.214	0.186	0.308*	0.242
	(0.173)	(0.173)	(0.174)	(0.176)	(0.174)	(0.169)	(0.169)
EDUC_high	0.578***	0.577***	0.557***	0.621***	0.562***	0.771***	0.662***
	(0.197)	(0.198)	(0.198)	(0.207)	(0.199)	(0.193)	(0.186)
age_medium	-0.045	-0.044	-0.014	0.062	-0.052	-0.065	
	(0.170)	(0.170)	(0.171)	(0.181)	(0.165)	(0.163)	
age_high	-0.133	-0.133	-0.113	-0.077	-0.140	-0.078	
	(0.181)	(0.182)	(0.182)	(0.197)	(0.177)	(0.173)	
job_unemployed				-0.005			
				(0.195)			
job_inactive				0.231			
				(0.191)			
maritalstatus_married				-0.220			
				(0.165)			
maritalstatus_separated				0.399			
				(0.300)			
FL_medium	0.555***	0.621**	0.562***	0.567***	0.543***		0.557***
	(0.199)	(0.262)	(0.199)	(0.200)	(0.201)		(0.197)
FL_high	1.166***	1.250***	1.192***	1.191***	1.141***		1.190***
	(0.207)	(0.268)	(0.208)	(0.209)	(0.213)		(0.202)
HH	0.003	0.157	-0.266				
	(0.148)	(0.343)	(0.200)				
FL_medium:HH		-0.164					
		(0.397)					
FL_high:HH		-0.204					
		(0.399)					
gender:HH			0.558				
			(0.280)				
news					0.117		
					(0.142)		
exp					0.024		
					(0.151)		
FL_dummy						0.617***	
						(0.183)	
R				-0.039	-0.028	-0.006	
				(0.150)	(0.149)	(0.147)	
Constant	-1.560***	-1.620***	-1.424***	-1.535***	-1.568***	-1.108**	-1.748***
	(0.410)	(0.430)	(0.414)	(0.492)	(0.475)	(0.453)	(0.377)
Observations	1,130	1,130	1,130	1,130	1,130	1,130	1,130
Log Likelihood	-657.588	-657.456	-655.606	-653.570	-657.187	-670.172	-659.926
Akaike Inf. Cr	1,341.176	1,344.912	1,339.211	1,341.139	1,344.375	1,364.345	1,331.852
Bayesian Inf. Crit.	1,406.565	1,420.361	1,409.631	1,426.649	1,419.824	1,424.704	1,362.032
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 4.23: Fixed effects estimates of random intercept and random slope models.

	Dependent variable: riskier	
	(14)	(15)
FL_medium	0.557***	0.562***
	(0.207)	(0.197)
FL_high	1.118***	1.193***
	(0.341)	(0.202)
EDUC_medium	0.236	0.199
	(0.170)	(0.209)
EDUC_high	0.653***	0.627***
	(0.187)	(0.209)
Constant	-1.717***	-1.718***
	(0.310)	(0.312)
Observations	1,130	1,130
Log Likelihood	-654.763	-659.465
Akaike Inf. Crit.	1,331.526	1,340.931
Bayesian Inf. Crit.	1,386.856	1,396.260

Note:*p-value < 0.1; **p-value < 0.05; ***p- value < 0.01
Model 14 and Model 15 have FL and EDUC as random slope, respectively.

Table 4.24: Analysis of the relation with stochastic dominance: fixed effects estimation.

	Dependent variable: choice						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
gender		-0.130	-0.173	-0.189	-0.319	-0.221	-0.156
		(7.292)	(7.265)	(7.353)	(28.376)	(7.191)	(7.395)
GEO_NO		-0.109	-0.004	0.063	-0.126	-0.147	-0.063
		(9.798)	(9.798)	(9.805)	(10.090)	(9.616)	(10.030)
GEO_NE		-0.066	-0.017	0.151	0.022	0.025	0.016
		(11.457)	(11.075)	(11.337)	(11.750)	(11.102)	(11.540)
GEO_CE		-0.062	0.009	0.114	0.082	-0.001	-0.092
		(9.427)	(9.348)	(9.429)	(9.469)	(9.175)	(9.766)
EDUC_medium		0.049	0.189	0.225	0.010	0.179	0.023
		(10.172)	(10.074)	(20.508)	(10.293)	(9.795)	(10.398)
EDUC_high		0.200	0.131	0.839	0.124	0.227	0.185
		(10.683)	(10.584)	(23.500)	(11.329)	(10.332)	(11.197)
age_medium		0.007	-0.129	-0.036	0.035	-0.863	0.044
		(8.882)	(8.627)	(8.714)	(8.837)	(18.428)	(9.048)
age_high		0.136	0.130	0.093	0.159	-1.422	0.142
		(9.420)	(9.243)	(9.516)	(9.639)	(19.520)	(9.791)
FL_medium			-0.343	0.645	-0.303	-1.458	-0.257
			(12.956)	(17.733)	(22.333)	(15.956)	(13.988)
FL_high			-0.062	0.642	-0.147	-1.025	-0.125
			(12.791)	(17.143)	(20.968)	(15.739)	(14.149)
EDUC_medium:FL_medium				-0.439			
				(23.625)			
EDUC_high:FL_medium				-0.908			
				(26.449)			
EDUC_medium:FL_high				-0.271			
				(23.224)			
EDUC_high:FL_high				-0.831			
				(24.740)			
gender:FL_medium					0.175		
					(31.297)		
gender:FL_high					0.127		
					(30.340)		
age_medium:FL_medium						1.175	
						(20.945)	
age_high:FL_medium						1.935	
						(22.193)	
age_medium:FL_high						0.851	
						(20.152)	
age_high:FL_high						1.581	
						(21.440)	
R							0.170
							(8.621)
Constant	18.211***	15.025	15.205	14.534	15.230	16.057	14.903
	(5.474)	(13.723)	(16.343)	(18.301)	(24.527)	(16.517)	(23.006)
Observations	474	474	474	474	474	474	474
Log Likelihood	-1.767	-1.907	-1.905	-1.907	-1.906	-1.906	-1.906
Akaike Inf. Crit.	9.535	25.814	29.809	37.814	33.811	37.812	31.812
Bayesian Inf. Crit.	22.018	71.587	83.905	108.554	96.230	108.552	90.069
Note:		*p-value < 0.1; **p-value < 0.05; ***p- value < 0.01					

(table continues in the next page)

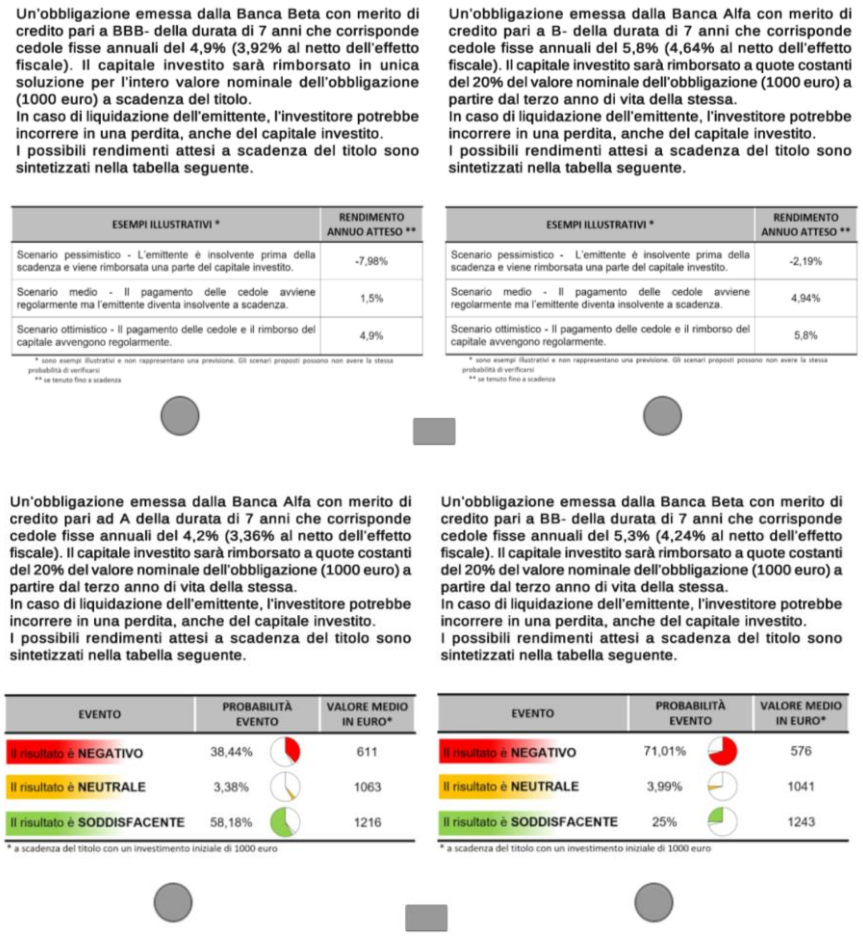
	Dependent variable: choice					
	(8)	(9)	(10)	(11)	(12)	(13)
gender	-0.207	-0.261	-0.395	-0.176	-0.169	-0.372
	(7.554)	(7.654)	(9.866)	(7.425)	(7.658)	(7.379)
GEO_NO	-0.158	-0.066	-0.061	-0.141	-0.132	-0.026
	(9.941)	(10.066)	(10.018)	(10.051)	(10.265)	(9.917)
GEO_NE	-0.038	-0.114	-0.066	-0.118	-0.201	0.092
	(11.812)	(11.782)	(11.689)	(11.357)	(11.962)	(11.815)
GEO_CE	-0.096	-0.059	-0.018	-0.210	-0.234	-0.015
	(9.343)	(9.498)	(9.353)	(9.717)	(10.119)	(9.687)
EDUC_medium	-0.067	-0.153	0.091	-0.101	-0.118	0.009
	(10.197)	(10.193)	(9.975)	(10.358)	(10.574)	(9.987)
EDUC_high	0.177	0.071	0.148	0.004	0.094	0.134
	(11.011)	(10.854)	(10.741)	(11.340)	(11.514)	(10.794)
age_medium	0.175	0.027	0.019	0.247	-0.122	0.097
	(9.044)	(9.024)	(9.059)	(9.980)	(9.291)	(8.785)
age_high	0.365	0.082	0.142	0.377	-0.002	0.280
	(9.801)	(9.836)	(9.723)	(11.052)	(10.219)	(9.637)
job_unemployed				-0.169		
				(10.910)		
job_inactive				0.247		
				(10.378)		
maritalstatus_married				-0.205		
				(9.279)		
maritalstatus_separated				0.046		
				(17.621)		
FL_medium	-0.080	0.024	-0.016	-0.311	-0.395	
	(12.898)	(15.495)	(13.107)	(13.059)	(14.266)	
FL_high	0.079	0.512	0.141	-0.042	-0.339	
	(12.823)	(15.186)	(13.100)	(13.051)	(14.396)	
HH	-0.036	-0.363	-0.218			
	(8.050)	(22.729)	(10.656)			
FL_medium:HH		0.604				
		(25.445)				
FL_high:HH		0.081				
		(24.377)				
gender:HH			0.537			
			(15.305)			
news					0.079	
					(7.625)	
exp					0.052	
					(8.522)	
FL_dummy						0.134
						(8.804)
R				0.213	0.309	0.175
				(8.790)	(8.961)	(8.519)
Constant	15.021	15.041	15.066	14.957	15.080	14.726
	(16.316)	(18.078)	(16.533)	(24.109)	(25.704)	(19.183)
Observations	474	474	474	474	474	474
Log Likelihood	-1.907	-1.904	-1.907	-1.907	-1.906	-1.907
Akaike Inf. Crit.	31.814	35.807	33.813	39.814	35.811	29.814
Bayesian Inf. Crit.	90.071	102.387	96.232	114.716	102.390	83.910

Note: *p-value < 0.1; **p-value < 0.05; ***p- value < 0.01

5 Annex II: Neuroscientific analysis.

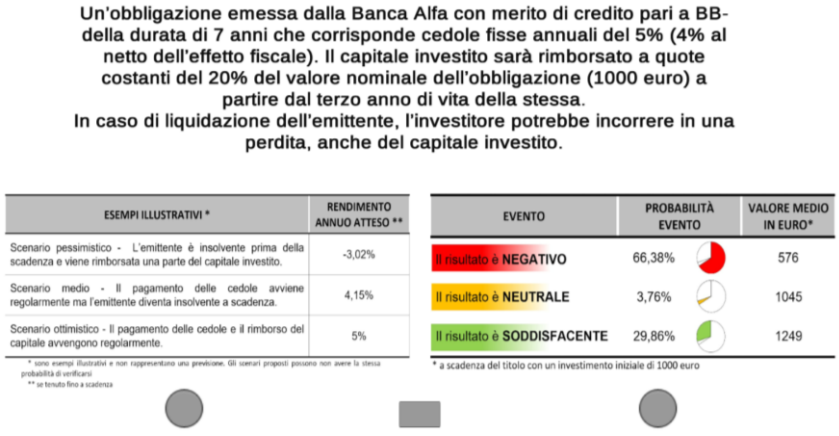
5.1 Tasks

Figure 5.1: Example of Task 1.



Two trials of Task 1 using WI and PS schemes, respectively.

Figure 5.2: Example of Task 2.



Single trial of Task 2 which included a single security and the two representation formats.

5.2 Characteristics of the sample

Table 5.1: Attributes distribution.

		Frequency	Percent
Gender			
	Male	18	49
	Female	19	51
Frequency missing = 5			
Age			
	20-34	26	72
	35-49	8	22
	50-64	2	6
Frequency missing = 5			
Employment status			
	Employed	24	65
	Unemployed	3	8
	Inactive	10	27
Frequency missing = 4			
Education			
	High school	10	27
	Bachelor's degree	2	5
	Master's degree	15	41
	Specialization	1	3
	PhD	9	24
Frequency missing = 4			

		Frequency	Percent
Past experience in financial investments			
	No	20	56
	Yes	16	35
Frequency missing = 5			
Correct answers about financial literacy test			
	Interest rate	30	73
	Inflation	25	61
	Fund	23	56
	Diversification	20	49
	Prospect	29	71
Financial literacy score (sum correct answer)			
	0	5	12
	1	4	10
	2	2	5
	3	10	24
	4	11	27
	5	9	22
Financial literacy score (class)			
	low	9	22
	medium	12	29
	high	20	49

5.3 Estimation results

Table 5.2: Fixed effects estimates on Reaction time (Task 1).

	Dependent variable: reaction time			
	(1)	(2)	(3)	(4)
T_PS	-5.150***	-5.428***	-5.428***	27.335
	(1.148)	(1.280)	(1.267)	(58.788)
age		-0.080	-0.151	1.830
		(0.275)	(0.279)	(4.395)
Gender_Male		5.641	6.488	345.291**
		(4.979)	(5.569)	(144.831)
education		-0.823	-1.025	2.489
		(0.821)	(0.847)	(5.530)
FL_medium			-10.468	
			(7.893)	
FL_high			-4.918	
			(7.789)	
T_PS:age				-0.818
				(2.447)
T_PS:gender_Male				-36.971
				(80.646)
age:gender_Male				-11.794**
				(5.359)
T_PS:education				-2.041
				(3.079)
age:education				-0.113
				(0.219)
gender_Male:education				-19.899**
				(7.949)
T_PS:age:gender_Male				1.411
				(2.984)
T_PS:age:education				0.049
				(0.122)
T_PS:gender_Male:education				2.024
				(4.426)
age:gender_Male:education				0.680**
				(0.281)
T_PS:age:gender_Male:education				-0.074
				(0.156)
Constant	29.250***	44.490***	55.859***	-8.887
	(2.349)	(14.028)	(16.511)	(105.582)
Observations	879	769	769	769
Log Likelihood	-3,771.315	-3,320.185	-3,313.645	-3,312.475
Akaike Inf. Crit.	7,554.630	6,658.369	6,649.290	6,664.950
Bayesian Inf. Crit.	7,583.302	6,700.175	6,700.386	6,757.851
Note:	*p-value < 0.1; **p-value < 0.05; ***p- value < 0.01			

Table 5.3: Fixed effects estimates on Accuracy (Task 1).

	Dependent variable: accuracy			
	(1)	(2)	(3)	(4)
T_PS	1.168***	1.220***	1.230***	1.092
	(0.139)	(0.150)	(0.150)	(6.878)
age		-0.008	-0.005	0.006
		(0.009)	(0.009)	(0.206)
gender_Male		0.429***	0.363**	2.543
		(0.162)	(0.181)	(6.723)
education		0.028	0.045	0.036
		(0.026)	(0.027)	(0.264)
FL_medium			0.616**	
			(0.260)	
FL_high			0.351	
			(0.258)	
T_PS:age				-0.078
				(0.283)
T_PS:gender_Male				-3.970
				(9.447)
age:gender_Male				-0.079
				(0.249)
T_PS:education				0.108
				(0.363)
age:education				-0.001
				(0.010)
gender_Male:education				-0.134
				(0.369)
T_PS:age:gender_Male				0.216
				(0.347)
T_PS:age:education				0.001
				(0.014)
T_PS:gender_Male:education				0.120
				(0.520)
age:gender_Male:education				0.005
				(0.013)
T_PS:age:gender_Male:education				-0.009
				(0.018)
Constant	-0.550***	-1.014**	-1.769***	-1.131
	(0.098)	(0.448)	(0.548)	(5.008)
Observations	902	792	792	792
Log Likelihood	-588.268	-510.488	-507.476	-507.667
Akaike Inf. Crit.	1,186.537	1,036.977	1,034.951	1,053.334
Bayesian Inf. Crit.	1,210.560	1,074.373	1,081.697	1,142.151

Note: *p-value < 0.1; **p-value < 0.05; ***p- value < 0.01

Table 5.4: Fixed effects estimates on Reaction time (Task 2).

	Dependent variable: reaction time			
	(1)	(2)	(3)	(4)
decision_PS	-1.277*	-1.640**	-1.542*	25.236
	(0.766)	(0.801)	(0.806)	(37.630)
age		-0.103	-0.125	0.661
		(0.111)	(0.114)	(1.797)
gender_Male		3.216	3.403	132.026**
		(1.998)	(2.245)	(58.655)
education		0.511	0.387	2.555
		(0.325)	(0.340)	(2.290)
FL_medium			-3.838	
			(3.200)	
FL_high			-1.589	
			(3.181)	
decision_PS:age				-0.509
				(1.622)
decision_PS:gender_Male				-19.589
				(47.560)
age:gender_Male				-3.432
				(2.165)
decision_PS:education				-1.986
				(1.914)
age:education				-0.044
				(0.090)
gender_Male:education				-7.583**
				(3.217)
decision_PS:age:gender_Male				0.436
				(1.855)
decision_PS:age:education				0.041
				(0.079)
decision_PS:gender_Male:education				1.955
				(2.553)
age:gender_Male:education				0.198*
				(0.113)
decision_PS:age:gender_Male:education				-0.047
				(0.095)
Constant	10.518***	3.872	8.682	-29.951
	(1.013)	(5.468)	(6.641)	(43.741)
Observations	902	792	792	792
Log Likelihood	-3,304.226	-2,917.895	-2,913.248	-2,916.128
Akaike Inf. Crit.	6,620.452	5,849.791	5,844.496	5,868.255
Bayesian Inf. Crit.	6,649.279	5,882.512	5,886.567	5,952.397
Note:	*p-value < 0.1; **p-value < 0.05; ***p- value < 0.01			

Table 5.5: Fixed effects estimates on Decision (Task 2).

	Dependent variable: **decision**			
	(1)	(2)	(3)	(4)
age	-0.020	-0.013	-0.290	
	(0.035)	(0.030)	(0.543)	
gender_Male	0.415	-0.239	-6.847	
	(0.635)	(0.613)	(17.693)	
education	0.163	0.231**	0.054	0.504*
	(0.101)	(0.093)	(0.684)	(0.259)
FL_medium		2.980***		5.699
		(0.884)		(5.700)
FL_high		2.413***		9.940*
		(0.860)		(5.406)
age:gender_Male			0.428	
			(0.656)	
age:education			0.011	
			(0.027)	
Gender_Male:education			0.281	
			(0.970)	
age:gender_Male:education			-0.020	
			(0.034)	
education:FL_medium				-0.118
				(0.312)
education:FL_high				-0.423
				(0.282)
Constant	-1.262	-4.693**	2.859	-10.250**
	(1.699)	(1.846)	(13.071)	(5.014)
Observations	792	792	792	814
Log Likelihood	-426.977	-421.603	-426.052	-435.759
Akaike Inf. Crit.	863.955	857.207	870.104	885.518
Bayesian Inf. Crit.	887.328	889.929	912.175	918.432

Note: *p-value < 0.1; **p-value < 0.05; ***p- value < 0.01

Table 5.6: Fixed effects estimates on electrodermal activity (Task 1).

	Dependent variable: **ISCR**			
	(1)	(2)	(3)	(4)
T_PS	-0.175**	-0.182**	-0.182**	4.991
	(0.070)	(0.074)	(0.074)	(3.397)
age		-0.000	-0.000	0.097
		(0.004)	(0.004)	(0.100)
Gender_Male		-0.000	-0.000	0.252
		(0.080)	(0.089)	(3.295)
education		0.000	0.000	0.141
		(0.013)	(0.014)	(0.126)
FL_medium			0.000	
			(0.127)	
FL_high			-0.000	
			(0.125)	
TPS:age				-0.195
				(0.141)
T_PS:gender_Male				-0.504
				(4.659)
age:gender_Male				-0.016
				(0.122)
T_PS:education				-0.282
				(0.178)
age:education				-0.005
				(0.005)
gender_Male:education				-0.012
				(0.181)
T_PS:age:gender_Male				0.032
				(0.172)
T_PS:age:education				0.010
				(0.007)
T_PS:gender_Male:education				0.024
				(0.256)
age:gender_Male:education				0.001
				(0.006)
T_PS:age:gender_Male:education				-0.001
				(0.009)
Constant	0.087*	0.091	0.091	-2.495
	(0.049)	(0.227)	(0.267)	(2.402)
Observations	880	770	770	770
Log Likelihood	-1,282.679	-1,126.869	-1,129.554	-1,160.221
Akaike Inf. Crit.	2,577.357	2,271.737	2,281.108	2,360.443
Bayesian Inf. Crit.	2,606.037	2,313.555	2,332.219	2,453.371
Note:	*p-value < 0.1; **p-value < 0.05; ***p- value < 0.01			

6 Annex III: criteria to assess representation format.

The following tables show some criteria used by the CESR to assess weakness and strengths of presentation format and methodology used for disclosing information of financial instruments (e.g. cost and risk-reward profile).

Table 6.1: Assessment of the criteria for presentation format.

CRITERIA	DESCRTIPTION
Engaging	The presentation of the elements in the prospectus should be as simple as possible, to be easily understood by retail investors and to engage them more, motivate them to use the prospectus and increase attention for the prospectus
Understandable	This criterion refers to the level of complexity of the information as whether consumers can interpret the information. The information in the prospectus should be understood by the retail investors with the assumption that the consumer may not have an adviser, distributor or seller on hand to explain the information.
Comparable	One of the main purposes of the prospectus is to enable investors to compare investment opportunities with each other. This can be done by an investor by comparing several prospectuses, but it is also possible to include reference points within the prospectus that make the information more comparable, for example through the use of benchmarks.
Compatible	This criterion refers to the compatibility of the proposals for presentation of information with the requirements formulated in the several regulations. All options will need to be strictly compliant with the regulations.
Balanced presentation	This criterion holds that we should look for a balanced presentation of the different aspects within sections of the prospectus but also a balanced presentation of the sections of the prospectus. By this criterion we try to make sure that the upsides and downsides of certain products are balanced in an objective fashion
Coverage of types of products	The goal is to develop presentational forms that are suitable and applicable to all different types of products that fall into the scope of the regulation. Given the heterogeneity of products in scope, this will be an important challenge.

Table 6.2: Assessment of the criteria for underlying methodologies.

CRITERIA	DESCRTIPTION
Reliable	The information within the prospectus is reliable where it provides a fair estimate of the actual risks and costs involved.
Robust	The measurement of risk, reward and costs should not be easy to manipulate. It should be an objective representation of the risk, reward and costs that are being measured.
Applicable	The measurements should be applicable to all types of products. Where a measurement (e.g. of historic volatility) is available for some products but not available for those without a track record, or might be a misleading measure for some, effective methodologies for combining or synthesizing different measures in an objective way may be necessary.
Stable	The output of the measurements needs to be relatively stable. It is important that risk or cost indicators are reliable forecasts, and that they are not overly sensitive to relatively minor changes in conditions.
Comparable	The measurements should lead to values that are comparable amongst different types of products.
Feasible/ Proportional	An indicator or measure that is overly sophisticated in relation to the granularity and accuracy of the information included in the prospectus could be seen as disproportionate. This does not imply that the simplest or least costly option should always be selected.
Supervision	Will it be possible for regulators to assess whether product manufacturers are complying with any proposed prescribed methodology?