The London School of Economics and Political Science

The Consequences of Behavioural Bias: Bandit Problems and Product Liability Law

Barbara Luppi

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, London, October 2009

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work, other than Chapter 4, which is a joint work with Francesco Parisi.

Abstract

The aim of the thesis is to explore how certain behavioural biases affect decision making. It focuses on two contexts, bandit problems and the case of legal decision making. In regard to bandit problems, the focus of interest is to examine the role of risk aversion and loss aversion, which are both excluded from the standard literature on bandit problems. We consider the standard bandit problem under alternative models of behaviour, from the standard expected utility model with risk neutral agents, to models with risk aversion, with loss aversion, and with risk and loss aversion. In Chapter 2 we show that new qualitative features of the prediction emerge in the presence of both risk aversion and loss aversion. Subjects who are moderately loss averse find it optimal to experiment more than highly loss averse or loss neutral individuals. In Chapter 3, an experimental study is conducted that involves a three-stage one-armed bandit problem. The main substantive finding in Chapter 3 is that there is a bias towards over-experimentation in stages 1 and 2, and a bias towards under-experimentation in stage 3, relative to the theoretical model, which is not explicable by reference to subjects' levels of risk or loss aversion. This remains a 'behavioural' puzzle. In regard to legal decision making, in Chapter 4 we examine the role of overconfidence in undermining the incentives provided by liability rules. Even factually informed, overconfident people tend to think that risks are less risky to materialize for themselves than for others and therefore inadequately react to legal threats and incentives such as liability rules. We examine the role of tort rules in "debiasing" overconfidence, showing paradoxically that the most effective way to correct overconfidence could be to forgive it, rather than to penalize it through liability. In Chapter 5 we examine applications to the field of product liability. Products regulations that impose liability on producers for not increasing the safety of their products in anticipation of consumer's overconfidence can be viewed as a legal strategy analogous to legal forgiveness of consumers' overconfidence.

Acknowledgements

I have received invaluable help and support from many people during my research: I could not have done it alone.

I would like to thank Professor John Sutton, my thesis advisor, for helping me during these years. His graduate lectures in Industrial Organization instilled in me the firm belief that economics can and should contribute of our understanding of reality. He has provided sound practical advice at all levels and he has done so with great kindness.

A special thank should also go to Professors Leonardo Felli and Kevin Roberts who provided valuable suggestions and criticisms as thesis examiners. Their effort was truly appreciated and led to a much better final product.

I must thank my family who must often have despaired of my ever getting to the end, but who nonetheless gave me emotional and financial support. I owe my gratitude to Francesco, who supported me through all the ups and downs of the last year.

All chapters in the thesis are my own work, except for Chapter 4 which builds on earlier joint work with Francesco Parisi.

I have received helpful comments on parts of this thesis from many people and audiences. For Chapter 2, I received comments from Robert Sugden and the participants to the University of East Anglia economics seminar. For Chapter 3, I have received comments from Graziella Bertocchi and the participants to the University of Modena and Reggio Emilia economics seminar. For Chapter 4 I have received comments from Aldo Rustichini and the participants to the University of Minnesota seminar, the 2009 American Law and Economics Association meeting and the 2009 European Law and Economics Association meeting.

Table of Contents

Chapter 1: An Introduction

1. A General Overview

Over the past three decades, behavioural and experimental economists have investigated the presence of systematic biases affecting human judgment in decision-making. Evidence both in the lab and in the field suggests the importance of behavioural biases in human decision-making. Investigating the role of behavioural biases and cognitive abilities in economic contexts has shed new light on how people make choices at the individual level. Even if we are far from "opening the black box of human reasoning", in the words of Rubinstein (2007), behavioural biases such as hyperbolic discounting, loss aversion, overconfidence, mental accounting, and biases involving framing effects have been shown to significantly affect various economic and financial decisions, such as buying and selling saving, investment and other financial choices. Such biases have been widely adduced to resolve anomalies in economic decision-making.

The present thesis aims to explore how certain behavioural biases affect decision-making. It focuses on two major behavioural biases: loss aversion and overconfidence, examined respectively in two alternative economic contexts: sequential choices under uncertainty and legal decision-making.

Loss aversion and reference dependence have been investigated at length both in economics and in neighbouring disciplines such as psychology and sociology, both at the empirical and experimental level. Evidence that agents treat losses differently from equal-sized gains has been found in experimental analysis, as shown by the presence of endowment effects (Kahneman, Knetsch, and Thaler 1990), selling behaviour in the housing market (where sellers are unwilling to sell below a reference point, represented usually by buying price (Genesove and Mayer 2001)) and more generally by the trading behaviour of individuals who are reluctant to realize losses. Additionally, Chen, Lakshminarayanan and Santos (2006) suggest the idea that loss aversion may be a universal bias arising regardless of experience and culture, by showing that even capuchin monkeys display loss aversion when faced with gambling situations.

We examine the role of loss aversion in the context of the bandit problems. The focus of the analysis is the role of risk aversion and loss aversion, which are both excluded from the standard literature on bandit problems. We consider the standard bandit problem under alternative models of behaviour, from the standard expected utility model with risk neutral agents, to models with risk aversion, with loss aversion, and with both risk and loss aversion. Chapter 2 investigates the predictions that emerge in the presence of both risk aversion and loss aversion. In Chapter 3, an experimental study is conducted that involves a three-stage one-armed bandit problem. The main experimental finding is that there is evidence of both over-experimentation and under-experimentation, and that this is true for subjects of all levels of risk aversion and loss aversion. At some stages in the process subjects of all types display over-experimentation, while at other stages there is a systematic bias towards under-experimentation.

In regard to legal decision-making, we turn our attention to other behavioural biases that may impact optimal choices in the field of tort law. Psychological research shows that there is systematic overconfidence in risk judgments. Data shows that overconfidence bias is pervasive in judgments that individuals make regularly in everyday life, not just in occasional activities (Jolls and Sunstein, 2006). Overconfidence creates a distinctive problem for legal policymakers: even factually informed people tend to think that bad outcomes are less likely to materialize for themselves than for others (Sunstein, 2000). Overconfident people therefore inadequately react to legal threats and incentives such as liability rules in various areas of law, and this carries some surprising implications. Chapter 4 examines the role of overconfidence in undermining the incentives provided by liability rules. We examine the role of tort rules in "debiasing" overconfidence, showing paradoxically that the most effective way to correct overconfidence could be to forgive it, rather than to penalize it through liability. Chapter 5 examines applications to the field of product liability. Product regulations that impose liability on producers for not increasing the safety of their products in anticipation of consumers' overconfidence can be viewed as a legal strategy analogous to legal forgiveness of consumers' overconfidence.

2. Experimentation and Loss Aversion

Rationality and emotions are intertwined and affect human decisions in a complex way. The well-known TV game show "Deal or No Deal" offers an interesting economic example of such situations. The game works as follows: a contestant picks one briefcase as his own from a set of 26 cases. Each case contains an unknown amount of money, varying from 1 cent to $1 million (in the US version of the game). The contestant is asked to open the remaining 25 cases. Each time, by process of elimination, information is partially revealed about what his own case might hold.

During the game, an anonymous player – called the banker – offers a deal to the contestant.[1] The proposal is: stop playing now and take the sure amount of money offered. Wall Street Journal reports that a player, Mr. Johnston, turned down an offer of $67,000 when he had only a chance of one third of winning $200,000.[2] Like Mr. Johnston, a number of players turn down offers of sure amounts of money that are higher than the expected value of the bet, in contrast with standard predictions of economic rationality. In situations like this, neoclassical economic models are neither descriptive nor predictive of human choice. At best they stand as prescriptive paradigms of human rationality. Choices like Mr. Johnston's pose a challenge to economic theory. How is it possible to explain the behaviour of such players through models of rationality?

Observing how people choose may help to shed light on the relative importance of rationality (in a neoclassical sense) and emotions when making a decision. One explanation is that the neoclassical models do not fully capture the entire decisional process: people may behave rationally and coherently with an objective function that differs from the one assumed in the standard expected utility framework. Among such explanations, Prospect Theory suggests that people may evaluate gains and losses from a psychological reference point shifting over time (Kahneman and Tversky, 1977).

The TV show game "Deal or No Deal" fits with the description of human activities in many areas: a person has to choose from a number of actions, each with a certain cost and an uncertain reward. Some of these actions are highly likely to produce a short-term gain, while others, such as gathering information to eliminate some of the uncertainty, may result in a long-term benefit only. The classic multi-armed bandit problem is a formalization of such a situation in economic and statistical theory: in each period the agent pays a cost to pull one of a fixed number of arms, different arms having different, unknown, and possibly interdependent payoff probabilities. The agent's problem is to maximize the expected discounted sum of the payoffs. Currently in the economic literature, bandit problems are equated with arms (Weitzman, 1981; Roberts and Weitzman, 1981).

In Chapter 2 we shall consider the bandit problem under alternative models of behaviour, from the standard expected utility model with risk neutral agents to models with risk aversion, with loss aversion and with both risk and loss aversion. We depart from the classical setting of bandit problems by endowing the agent with a disappointment-elation utility function (Sugden and Loomes, 1986), assuming that the player aims at maximizing the expected utility of profits, instead

[1] The "Deal or No Deal" game cannot be fully analogized with the one-armed bandit problem discussed in Chapter 1, due to the presence of the banker. In this example we study departures from standard predictions of expected utility theory, without stretching a fully analogy with the economic model developed in Chapter 2.

[2] Source: Wall Street Journal, Thursday 12 January 2006, "Game show is all risk, no skill –which is why economists love it".

of assuming the maximization of expected profits as an objective function. According to Sugden and Loomes (1986), the individual receives the utility derived directly from the actual consequence of an uncertain prospect, but he also feels some degree of disappointment and elation – measured as a deviation from the a priori expected payoff. If the actual consequence turns out to be worse than the expectation, he feels disappointment, while he experiences some degree of elation, if the actual consequence is better than the a priori expectation. We model the disappointment-elation utility function to capture the spirit of loss-aversion (Kahneman and Tversky, 1979): the disutility of a loss is greater than the elation associated with a same-sized gain.

What will be the optimal experimentation choice of a loss-averse player? Will this agent experiment more or less in the setting of a bandit problem? Chapter 2 addresses this question in a traditional one-armed bandit problem, where the player's decision problem has the same characteristics as the standard problem. The player faces the decision between an unknown arm and a safe one – in the TV show language the player should decide in each round whether to keep the case (risky arm) or accept the banker's offer (safe arm) and walk away with a sure prize.[3] Analytically the player faces an optimal stopping problem, i.e. he has to decide how many times to experiment with the unknown arm before switching to the safe one, so that once switched the agent won't return to the unknown arm again. Such experimentation is costly: in the short run, the agent bears the loss in case of a negative result. This loss must be traded off against the potential informational gain associated with experimentation, in terms of a more correct estimate of the payoffs obtainable when undertaking the risky project. In the standard framework, the agent's optimal strategy is found by weighing these two factors.

Chapter 2 shows that new qualitative features of the prediction emerge in the presence of both risk aversion and loss aversion. In the theoretical analysis we model the agent's preference to display loss aversion jointly with the presence of elation. Hence, we introduce an additional trade-off associated with experimentation. When playing "safe", the player feels neither disappointment nor elation, since he gets exactly the amount he expects to receive. When playing the risky arm, the player faces the additional cost of disappointment in case of failure and the possible gain of elation when the experimentation proves successful. If the player is loss-averse, he will give greater weight to disappointment than elation. This additional trade-off plays a key role in characterizing the optimal experimentation strategy of the player.

Common intuition would lead us to think that a loss-averse player would experiment less and choose to opt out the game before a loss-neutral player. In Chapter 2 we show that this standard

[3] Note that the analogy would fit the one-armed bandit problem only if the banker played no active role, keeping the outside offer to the player constant.

intuition holds true only for individuals who exhibit a high degree of loss-aversion. On the contrary, players who are moderately loss-averse will choose to experiment more. At every stage of the game, the player trades off the immediate cost of experimentation – due to the possible disappointment experienced in the short run when the failure occurs – with the long-run benefits from experimentation, which are measured by the presence of the additional benefit of elation in case of a success. The understanding of the long-run benefits of experimentation is important at this point. There are three main benefits that the player can obtain from experimentation. First, experimentation provides information on the payouts of the risky arm. Second, experimentation may provide elation when the attempt proves successful. Third, experimentation may "spare" disappointment in the long run due to the choice of a more profitable arm. For high levels of loss-aversion, the immediate short run cost of disappointment weighs more and may induce the player to opt for the safe arm, avoiding any disappointment cost even in the short run. A moderately loss-averse player, however, may be willing to do more experimentation, because the additional gain represented by the elation component may outweigh the higher cost of disappointment due to the loss aversion in the short run.

The departure point of the analysis is Rotschild's (1974) seminal paper on experimentation. He examines the pricing decision of a monopolist facing an unknown stochastic demand. The store can choose between two possible prices (arms), each with an unknown probability to make a sale; at each period the monopolist selects the price to charge and he updates his beliefs on the basis of the resulting sales to customers. The monopolist faces the trade-off between the short-term benefit, represented by charging the price that maximizes his payoff given his current information and the information gained on the demand at the other price, which has long-term value. Rotschild shows that with positive probability the monopolist will choose the inferior price, i.e. the price less likely to make the sale. Therefore, inefficiency may arise in the long run. Rothschild's analysis relates to the multi-armed bandit literature, on which Gittins (1979) and others have worked to characterize the index policy. In the economics literature Weitzman (1981) and Roberts and Weitzman (1981) provide two important applications of bandit problems with independent arms to R&D settings.

Another literature branch relevant in Chapter 2 examines alternative frameworks to the expected utility model of behaviour under uncertainty. Experimental research into choice under uncertainty has revealed that people behave in ways that systematically violate the set of basic axioms formulated by Von Neumann and Morgenstern (1947) and Savage (1954), upon which the conventional expected utility theory is built. In particular, the empirical evidence presented by Kahneman and Tversky (1979) and others show a number of patterns of choice that reveal behavioural regularities that systematically contradict the predictions of conventional expected utility. After the discovery of Allais paradox and Ellsberg paradox, many attempts have been done to develop alternative frameworks for the analysis of choice under uncertainty, consistent with the observed behavioural regularities. A subset of them starts from an attempt at a psychological explanation of the Allais Paradox phenomena. One of the earliest was prospect theory (Kahneman and Tversky, 1979), generalized later by cumulative prospect theory (1992). Two more intuitive and parsimonious psychologically based theories are regret theory (Sugden and Loomes, 1982, 1987) and disappointment theory (Sugden and Loomes, 1986). Both theories incorporate ex ante considerations of ex post psychological feelings: regret or rejoicing in the former, and disappointment or elation in the latter theory. "The fundamental idea behind regret theory is that the psychological experience of 'having x' can be influenced by comparison between x and y that one might have had, had one chosen differently. If, for example, I bet on a horse which fails to win, I may experience something more than a reduction in my wealth: I may also experience a painful sense of regret arising out of the comparison between my current state of wealth and the state that I would have enjoyed, had I not bet" (Sugden, 1991).

3. Experimentation and Loss Aversion: An Experimental Perspective

Chapter 2 investigated the optimal experimentation decision of a risk-averse and a loss-averse player in the setting of a one-armed bandit problem. Chapter 3 addresses the same question adopting an experimental perspective. We set up an experiment to investigate how risk aversion and loss aversion affect the optimal experimentation decisions in an experimentation environment modeled as a bandit problem.

In total 254 subjects – first year undergraduate students in Microeconomics - took part in the experiment. The experimental study involved for each subject the elicitation of individual preferences (loss aversion and risk aversion). Subjects participated in an experimentation environment involving a three-stage one-armed bandit problem.

In the first part of the study we measure loss aversion at the individual level in a within-subject design. Loss aversion can occur in riskless and in risky choices (Kahneman and Tversky, 1979, 1991). An example of loss aversion with riskless choice is the 'endowment effect' – the observation that experimental subjects who are randomly endowed with a commodity ask a selling price that substantially exceeds the buying price offered by subjects who can buy the commodity (Kahneman, Knetsch and Thaler, 1990). An example of loss aversion in risky choices is the observation that people reject small-scale gambles that have a positive expected value but may involve losses (e.g., Rabin, 2000; Fehr and Goette, 2007; Tom, Fox, Trepel and Poldrack, 2007).[4] We adopt an individual measure of loss aversion in a risky choice task, adapting a lottery choice similar to Fehr and Goette (2002). We complement this measure with one of risk aversion, in order to disentangle possible interlinked effects on the experimentation choices of subjects. Anderson (2001) shows that the presence of risk aversion induces agents to undertake a lower level of search.

In the second part of the study we investigate the experimentation decision of each subject in a three-stage one armed-bandit problem. The experimental design is based on the three-stage game example illustrated in Section 3 of Chapter 2 – where the optimal experimentation strategy is fully characterized for a risk-neutral player, a risk averse one, a loss averse one and for a player who displays both risk aversion and loss aversion. The experiment is designed as a traditional one-armed bandit problem, where the player faces the decision between an unknown arm and a safe one and the player should decide in each round whether to keep experimenting on the risky arm or to opt out of the game with a sure prize on the safe arm. The experiment involves an optimal stopping problem, i.e., each subject in the experiment has to decide how many times to experiment with the unknown arm before switching to the safe one, so that once he has switched the agent won't return to the unknown arm. In the experimental design we ask each subject whether he or she would accept to play on a slot machine at a pre-specified price and at which highest price they would accept to play. By doing so, we are able to fully characterize his or her experimentation decisions at each stage of the game for each participant in the experiment.

The experimental results confirm that individual experimentation decision is decreasingly monotonic in the degree of risk aversion, consistently with Anderson (2001), and in loss aversion. The main experimental finding is that there is evidence of both over-experimentation and under-experimentation, and this is true for subjects of all level of risk aversion and loss aversion. At some stages in the process subjects of all types display over-experimentation, while at other stages there is a systematic bias towards under-experimentation.

[4] Loss aversion has been invoked to explain many naturally occurring phenomena that are hard to understand under the assumption of reference-point independence. Camerer (2004) provides an overview of the field evidence, and Starmer (2000) a survey of theoretical explanations. See Sugden, (2003), Schmidt, Starmer and Sugden (2005) and Köszegi and Rabin (2006) for recent theoretical frameworks of reference-dependent preferences that can explain many of these phenomena. Loss aversion has been cited to explain behaviour in financial markets (Benartzi and Thaler, 1995; Gneezy and Potters, 1997; Odean, 1998; Haigh and List, 2005); selling patterns in housing markets (Genesove and Mayer, 2001); consumption behavior (Bowman, Minehart and Rabin 1999; Chua and Camerer 2004); marketing practices (Hardie, Johnson and Fader 1993; Carmon and Ariely 2000); trade policy (Tovar, 2006); labor supply (Camerer, Babcock, Loewenstein and Thaler, 1997;Goette, Huffman and Fehr, 2004; Fehr and Goette, 2007) and the importance of defaults and the status-quo bias in decision making (Samuelson and Zeckhauser, 1988; Johnson and Goldstein, 2003).

4. Overconfidence in Tort Law: General Principles

Behavioural biases have important effects on the optimality of economic choices. Chapters 2 and 3 investigated the role of risk aversion and loss aversion in an experimentation environment from a theoretical and experimental perspective. Chapter 4 and 5 consider the role of behavioural biases in regard to legal decision making, in the field of tort law by adopting a law and economics approach.

The economic analysis of tort law characterizes the care incentives of a potential tortfeasor, defined as an economic agent who undertakes a risky activity that may cause an injury to another agent, called the victim, and who chooses which level of care to undertake in order to reduce the likelihood of an injury. In many situations, not only potential injurers but also potential victims may undertake care in order to make the injury less likely. Liability regimes provide different care incentives to potential injurers and victims, since they shield parties from liability in different ways. The economic analysis of tort law studies how the liability regimes affect the optimal care incentives of tortfeasors (in a unilateral care model) and victims (in the bilateral one).

Loss aversion plays a role in the field of tort law. An implication of loss aversion is that people treat opportunity costs differently than "out-of-pocket" costs. Foregone gains are less painful than perceived losses. This perception is strongly manifested in people's judgments about fair behaviour. The presence of loss aversion raises questions about the policy objectives of the tort system. For example, Sunstein (2000) questions whether the liquidation of damages should aim at restoring an injured party to the position in which the victim was prior to the accident, or whether damages should instead reflect the amount the victim would have demanded prior to the accident to accept the injury. Sunstein suggests that "juries appear to believe that the amount that would be demanded pre-injury is far greater than the amount that would restore the status quo ex ante. The legal system generally appears to take the latter stance on the compensation question, though it does not seem to have made this choice in any systematic way" (Sunstein, 2000). Cohen and Knetsch (1990) showed that this principle, embodied in the old expression that "possession is nine tenths of the law," is reflected in many judicial opinions. For example, in tort law judges make the distinction between "loss by way of expenditure and failure to make gain." In one case, several bales fell from the defendant's truck and hit a utility pole, cutting off power to the plaintiff's plant. The plaintiff was able to recover wages paid to employees which were considered "positive outlays" but could not recover lost profits which were merely "negative losses consisting of a mere deprivation of an opportunity to earn an income"[5].

The definition of loss matters when potential tortfeasors display loss aversion. Bigus (2005) shows that the care incentives of a potential injurer remain optimal if only damage payments are considered in calculating a loss, thereby obtaining results similar to the Expected Utility Theory. Additional efficient care incentives are preserved even if both liability payments and costs of care are considered as a loss. This occurs even under the rule of strict liability, while it is not true in the case of risk aversion in an Expected Utility Theory framework. The presence of loss aversion does not undermine care incentives in either liability regime. On the contrary, it enhances them above the social optimum. Hence, loss aversion is not a source of inefficiency in a tort problem.

We turn our attention to other behavioural biases that may impact optimal choices in the field of tort law and create inefficiencies. Psychological research shows that there is systematic overconfidence in risk judgments – which is one of the most widespread psychologically generated biases in human judgment. Data shows that overconfidence bias is pervasive in the judgments that individuals make regularly in everyday life, rather than in activities that are seldom carried out (Jolls and Sunstein, 2006). Overconfident individuals overestimate their own ability and underestimate the risk that they face. Overconfidence is one of the two biases that result from what psychologists know as optimism bias.[6] Optimism bias affects people's subjective estimates of the likelihood of future events, and causes them to overestimate the likelihood of positive or desirable events and to underestimate the likelihood of negative or undesirable events (Colman, 2001). Optimism bias was reported by psychologists starting from the early twentieth-century (Lund, 1925; and Cantril 1938) and was rigorously studied and documented by the US psychologist Neil David Weinstein, who named it "unrealistic optimism" in a 1980 article published in the *Journal of Personality and Social Psychology*.[7]

Several subsequent psychological studies confirm the stylized fact that people exhibit an unrealistic optimism bias (see the survey of the literature in Wenglert and Rosén, 2000).[8] An interesting example of overconfidence for the purpose of this analysis is the finding by Svenson

[5] See Cohen and Knetsch (1990), p. 18.

[6] There are two forms of judgment bias that follow from unrealistic optimism: overconfidence bias which implies an overestimation of one's own ability, and self-serving bias which is a tendency to evaluate evidence or make judgments in a way that benefits oneself (Muren, 2004).

[7] Weinstein (1980) asked students to estimate the likelihoods of various events happening to them and showed that they rated their chances of experiencing positive (negative) events significantly above (below) the average for their peers.

[8] Economic experiments confirm the existence of optimism and related overconfidence bias. Forsythe, Rietz and Ross (1999) find experimental evidence of the human tendency to overestimate the probability of desirable events. Babcock, Loewenstein, Issacharoff and Camerer (1995) find evidence of a self-serving bias in an experiment where subjects were given roles as plaintiffs and defendants in a legal dispute over a claim for damages. Kaplan and Ruffle (2001) argue that strategic behavior may affect the measure of these biases.

(1981) that most survey respondents rated themselves as better and more competent drivers than average. Overconfidence appears to be robust with respect to a variety of accident risks (see, among others, Sunstein, 1997 and Jolls, 1998).[9] Studies on traffic accidents show that people's assessment of accident risks faced by others is fairly accurate (Lichtenstein et al. 1978) and at the same time people tend to be unrealistically optimistic about themselves, underestimating the likelihood that they will cause an accident (see Svenson 1981; Svenson, Frischhoff, and MacGregor 1985; Finn and Bragg 1986; Matthews and Moran 1986; DeJoy 1989; McKenna, Stanier and Lewis 1991; Guppy 1992; Jolls, 1998).[10]

Overconfidence creates a distinctive problem for legal policymakers: even factually informed people tend to think that risks are less likely to materialize for themselves than for others (Sunstein, 2000).[11] Overconfident people therefore inadequately react to legal threats and incentives such as liability rules in a number of areas of law, showing some surprising implications.

The traditional law and economics view on issues of overconfidence is that the problem is caused by imperfect information and can be appropriately corrected through the provision of additional information (see, e.g., Stiglitz, 1986 on consumer optimism). However, as the extensive evidence suggests, overconfidence leads many individuals to underestimate their personal risks even if they receive accurate information about average risks. Evidence indeed suggests that debiasing strategies through risk education and information are only partially effective. As pointed out by Jolls and Sunstein (2006), educated people could hold accurate statistical information about the probability of contracting cancer, and yet believe that they are less likely than most people to contract it. As Viscusi (2002) puts it, people may accurately learn the aggregate statistical risk, but their information only has a limited effect on the optimistic "above average" illusion that leads them to underestimate the actual risk that they face.

---

[9] Empirical evidence suggests that optimism bias affects not only the perception of risky events, but it affects other choices, such as litigation choices. It has been shown that lawyers and litigants are systematically optimistic with respect to the outcome at trial. Bar-Gill (2006) studies the persistence of optimism bias in litigation using a setting of evolutionary game theory. The adaptive force of optimism is linked to its instrumental value in the pre-trial bargaining: optimistic lawyers can credibly threaten to resort to costly litigation, and are therefore more successful in extracting bargaining surplus from settlements.

[10] Hanson and Kysar (1999a and 1999b) report evidence of the resilience of consumer optimism, also in the face of risk disclosure and product warnings. Jolls (1998) further reports that people underestimate their own likelihood of being victims of natural disasters, as evidenced by their frequent failure to buy insurance for floods and earthquakes. Armour and Taylor (2002) discuss the case of professional financial experts, who consistently overestimate expected earnings, and business school students who overestimate starting salaries and probabilities of employment at graduation. According to Jolls and Sunstein (2006), this evidence suggests that people sometimes exhibit optimism bias not only with respect to relative risk, but also in the estimation of actual probabilities in their category of profession.

[11] Sunstein (2000) considers overconfidence and the role of law in debiasing this judgment error, considering the special challenges posed by the fact that the vast majority of people believe that they are less likely than other people to be subject to automobile accidents, infection from AIDS, heart attacks, asthma, and many other health risks, even though they do not lack statistical information about these risks in general.

Legal scholars have introduced the concept of debiasing through law, both with reference to procedural rules governing adjudication (Babcock, Loewenstein, Issacharoff and Camerer, 1995; and Babcock, Loewenstein and Issacharoff, 1997) and substantive law (Jolls and Sunstein, 2006). The idea of debiasing through law, as most notably presented by Jolls and Sunstein (2006), is that the design of substantive law should consider the need to correct the systematic judgment errors of individuals.[12]

Behavioural law and economics scholars have addressed the issue of overconfidence, considering the possible role of law in restraining or correcting this judgment error.[13] Chapter 4 considers the role of tort rules in debiasing overconfidence, unveiling another surprising and counterintuitive implication. We consider alternative legal strategies to correct overconfidence problems in tort law. We illustrate the possible use of threat strategies (threatening liability when overconfidence leads to an accident) and forgiveness strategies (foregoing liability when the accident is solely caused by a biased perception of risk) as alternative ways in which law could be used to debias overconfidence. We compare the alternative threat and forgiveness strategies in reducing the cost of accidents due to overconfidence, allowing for the possibility that government investment in information fails to guarantee full debiasing of agents. Under each liability rule we characterize the care and activity levels chosen by injurer and victim in the presence of overconfidence and we rank each combination of rule/strategy according to the efficiency level that it will induce. The model highlights the role of tort law and the optimal design of liability rules for correcting overconfidence biases. We compare the effectiveness of alternative legal strategies under alternative liability rules, unveiling an interesting paradox: legal forgiveness of overconfidence may be a valuable second-best solution, when debiasing through information and threat strategies proves ineffective. The most effective way to correct overconfidence in tort law may well be to forgive it, rather than to penalize it through liability. Product regulations that impose liability on producers for not increasing the safety of their products in anticipation of consumer's overconfidence can be viewed as a legal strategy germane to legal forgiveness of overconfidence (consumer are effectively shielded against the consequences of their overconfident errors).

---

[12] An important concern in the legal literature is to avoid the adoption of legal solutions that restrict choices or impose significant costs on unbiased individuals in order to improve the outcome for those who exhibit cognitive biases (Camerer, Issacharoff, Loewenstein, O'Donoghue and Rabin 2003; Mitchell 2002).

[13] A growing body of law and economics literature focuses on the departures of human behavior from full rationality and attempts to explain the positive and normative implications of bounded rationality in the formulation of legal policy. Sunstein (1997) and Jolls, Sunstein, and Thaler (1998) point out the need of a more accurate understanding of behavior and individual choice in legal context, in order to take into account the shortcomings in human behavior when structuring the law. Jolls and Sunstein (2006) discuss the idea of "debiasing through law", instead of "debiasing law", i.e., to insulate legal outcomes from the effects of boundedly rational behavior. Debiasing through law is instead aimed at developing legal strategies attempting to reduce or eliminate boundedly rational behavior. Jolls and Sunstein provide a general description of debiasing through law with application to many areas, such as consumer safety law, corporate law and property law.

5 Overconfidence in Tort Law: Applications in Product Liability Law

In this section, we begin with a survey of older cases that construed errors due to overconfidence and unrealistic optimism as negligence. In these cases, "threat strategies" were utilized, imposing liability for accidents caused by parties' overconfidence. In these cases, courts utilized the threat of liability to induce the parties to correct their biases. We subsequently discuss the gradual trend of case law towards "forgiveness strategies" with an analysis of cases in which forgiveness strategies are fully implemented and extended to areas outside of product liability, including premises liability, wrongful death, and negligence. We conclude by looking at the different treatment of overconfidence in other areas of tort law involving unilateral care, such as medical malpractice law.

Chapter 2: Loss Aversion and Experimentation

"Footfalls echo in the memory
Down the passage we did not take"

T.S. Eliot

1. Introduction

Rationality and emotions are intertwined and affect human decisions in a complex way. The well known TV game show "Deal or No Deal" offers an interesting economic example of such situations. The game works as follows: a contestant picks one briefcase as his own from a set of 26 cases. Each case contains an unknown amount of money, varying from 1 cent to $1 million (in the US version of the game). The contestant is asked to open the remaining 25 cases. Each time, by process of elimination, information is partially revealed about what his own case might hold. During the game, an anonymous player – called the banker – offers a deal to the contestant. The proposal is: stop playing now and take the sure amount of money offered. Wall Street Journal reports that a player, Mr. Johnston, turned down an offer of $67,000 when he had only a chance of one third of winning $200,000.[14] Like Mr. Johnston, a number of players turn down offers of sure amounts of money that are higher than the expected value of the bet, in contrast with standard predictions of economic rationality. In situations like this, neoclassical economic models are neither descriptive nor predictive of human choice. At best they stand as prescriptive paradigms of human rationality. Choices like Mr. Johnston's pose a challenge to economic theory. How is it possible to explain the behavior of such players through models of rationality?

Observing how people choose may help to shed light on the relative importance of rationality (in a neoclassical sense) and emotions when making a decision. One explanation is that the neoclassical models do not fully capture the entire decisional process: people may behave rationally and coherently with an objective function which differs from the one assumed in the standard

[14] Source: Wall Street Journal, Thursday 12 January, 2006, "Game show is all risk, no skill –which is why economists love it".

expected utility framework. Among such explanations, Prospect Theory suggests that people may evaluate gains and losses from a psychological reference point shifting over time (Kahneman and Tversky, 1977).

The TV show game "Deal or No Deal" fits with the description of human activities in many areas: a person has to choose from a number of actions, each with a certain cost and an uncertain reward. Some of these actions are highly likely to produce a short-term gain, while others, such as gathering information to eliminate some of the uncertainty, may result in a long-term benefit only. The classic multi-armed bandit problem is a formalization of such a situation in economic and statistical theory: in each period the agent pays a cost to pull one of a fixed number of arms, different arms having different, unknown, and possibly interdependent payoff probabilities. The agent's problem is to maximize the expected discounted sum of the payoffs. Currently in the economic literature, bandit problems are equated with arms (Weitzman, 1981; Roberts and Weitzman, 1981).

In Chapter 2 we shall consider the bandit problem under alternative models of behaviour, from the standard expected utility model with risk neutral agents to models with risk aversion, with loss aversion and with both risk and loss aversion. We depart from the classical setting of bandit problems by endowing the agent with a disappointment-elation utility function (Sugden and Loomes, 1986), assuming that the player aims at maximizing the expected utility of profits, instead of assuming the maximization of expected profits as an objective function. According to Sugden and Loomes (1986), the individual receives the utility derived directly from the actual consequence of an uncertain prospect, but he also feels some degree of disappointment and elation – measured as a deviation from the a priori expected payoff. If the actual consequence turns out to be worse than the expectation, he feels disappointment, while he experiences some degree of elation, if the actual consequence is better than the a priori expectation. We model the disappointment-elation utility function to capture the spirit of loss-aversion (Kahneman and Tversky, 1979): the disutility of a loss is greater than the elation associated with a same-sized gain.

What will be the optimal experimentation choice of a loss-averse player? Will this agent experiment more or less in the setting of a bandit problem? Chapter 2 addresses this question in a traditional one-armed bandit problem, where the player's decision problem has the same characteristics as the standard problem. The player faces the decision between an unknown arm and a safe one – in the TV show language the player should decide in each round whether to keep the case (risky arm) or accept the banker's offer (safe arm) and walk away with a sure prize.3 Analytically the player faces an optimal stopping problem, i.e. he has to decide how many times to experiment with the unknown arm before switching to the safe one, so that once switched the agent won't return to the unknown arm again. Such experimentation is costly: in the short run, the agent bears the loss in case of a negative result. This loss must be traded off against the potential informational gain associated with experimentation, in terms of a more correct estimate of the payoffs obtainable when undertaking the risky project. In the standard framework, the agent's optimal strategy is found by weighing these two factors.

Chapter 2 shows that new qualitative features of the prediction emerge in the presence of both risk aversion and loss aversion. Chapter 2 shows the non-equivalence of experimentation strategies in the presence of loss aversion and risk aversion. In the theoretical analysis we model the agent's preference to display loss aversion jointly with the presence of elation. Hence, we introduce an additional trade-off associated with experimentation. When playing "safe", the player feels neither disappointment nor elation, since he gets exactly the amount he expects to receive. When playing the risky arm, the player faces the additional cost of disappointment in case of failure and the possible gain of elation when the experimentation proves successful. If the player is loss-averse, he will give greater weight to disappointment than elation. This additional trade-off plays a key role in characterizing the optimal experimentation strategy of the player.

Common intuition would lead us to think that a loss-averse player would experiment less and choose to opt out the game before a loss-neutral player. In Chapter 2 we show that this standard intuition holds true only for individuals who exhibit a high degree of loss-aversion. On the contrary, players who are moderately loss-averse will choose to experiment more. At every stage of the game, the player trades off the immediate cost of experimentation – due to the possible disappointment experienced in the short run when the failure occurs – with the long-run benefits from experimentation, which are measured by the presence of the additional benefit of elation in case of a success. The understanding of the long-run benefits of experimentation is important at this point. There are three main benefits that the player can obtain from experimentation. First, experimentation provides information on the payouts of the risky arm. Second, experimentation may provide elation when the attempt proves successful. Third, experimentation may "spare" disappointment in the long run due to the choice of a more profitable arm. For high levels of loss-aversion, the immediate short run cost of disappointment weighs more and may induce the player to opt for the safe arm, avoiding any disappointment cost even in the short run. A moderately loss-averse player, however, may be willing to do more experimentation, because the additional gain represented by the elation component may outweigh the higher cost of disappointment due to the loss aversion in the short run.

The chapter is organized as follows. In Section 2 we provide a brief overview of the literature, both in two-armed bandit problems and in the alternative frameworks to standard expected utility.

In Section 3 we develop an illustrative example to explain the nature of the experimentation problem faced by the economic agent and to illustrate the analytical set-up of the decision procedure in a three stage game. In Section 4 an existence theorem of the optimal experimentation strategy of a loss-averse player is derived in an infinite horizon, two-armed bandit problem. In Section 5 the optimal experimentation strategy of a risk-averse and of a loss-averse player is characterized in a one-armed bandit problem with finite horizon. Section 6 concludes.

2. Related literature

The departure point of our analysis is Rotschild's (1974) seminal paper on experimentation. He examines the pricing decision of a monopolist facing an unknown stochastic demand. The store can choose between two possible prices (arms), each with an unknown probability to make a sale; at each period the monopolist selects the price to charge and he updates his beliefs on the basis of the resulting sales to customers. The monopolist faces the trade-off between the short-term benefit, represented by charging the price that maximizes his payoff given his current information and the information gained on the demand at the other price, which has long-term value. Rotschild shows that with positive probability the monopolist will choose the inferior price, i.e. the price less likely to make the sale. Therefore, inefficiency may arise in the long run. Rothschild's analysis relates to the multi-armed bandit literature, on which Gittins (1979) and others have worked to characterize the index policy. In the economics literature Weitzman (1981) and Roberts and Weitzman (1981) provide two important applications of bandit problems with independent arms to R&D settings.

Another literature branch relevant to the present chapter examines alternative frameworks to the expected utility model of behaviour under uncertainty. Experimental research into choice under uncertainty has revealed that people behave in ways that systematically violate the set of basic axioms formulated by Von Neumann and Morgenstern (1947) and Savage (1954), upon which the conventional expected utility theory is built. In particular, the empirical evidence presented by Kahneman and Tversky (1979) and others show a number of patterns of choice that reveal behavioural regularities that systematically contradict the predictions of conventional expected utility. After the discovery of Allais paradox and Ellsberg paradox, many attempts have been done to develop alternative frameworks for the analysis of choice under uncertainty, consistent with the observed behavioural regularities. A subset of them starts from an attempt at a psychological explanation of the Allais Paradox phenomena. One of the earliest was prospect theory (Kahneman and Tversky, 1979), generalized later by cumulative prospect theory (1992). Two more intuitive and procedure of an agent characterized simultaneously by risk aversion and loss aversion is discussed in section 3.3.

3.1 The Bandit Problem in Expected Utility Theory Framework

In the standard framework of expected utility theory, we consider at first a risk neutral agent, whose preferences are represented by a linear utility function. With no loss of generality, the agent is assumed to have a utility level equal to the payoff of the experimentation decision x, i.e. $U(x)=x$.

We will take an advantage of the basic results in 'bandit' theory which establishes that the optimal strategy for this class of games can be obtained by backward induction and we calculate this for the present example. An alternative method to backward induction is the method of Bradt, Johnson and Karlin (1956). Calculations of the optimal strategy are performed by using both methods, i.e. the backward induction and the method of Bradt, Johnson and Karlin.

3.1.1 Backward induction

According to Gittins (1979), the optimal strategy of the experimenter can be characterized by backward induction. Denote with $h=\{h_1,h_2,...\}$ the set of histories that leads to continued experimentation, where h_i , the history in period i, is a binary variable that takes value 1 when the experimentation along the risky project lead to a success in period i or 0 otherwise.

In the initial period the experimenter knows the prior probability ϕ of a success of the experimentation strategy along the risky arm. In each subsequent period, the experimenter updates the prior ϕ through the calculation of a posterior on the basis of the history of play. The information set regarding the experimentation history can be summarized through the help of the following two statistics:

$$\rho_{h_i} \equiv \frac{1}{1+n_{h_i}}$$

and

$$\mu_{h_i} \equiv \frac{s_{h_i}}{1+n_{h_i}}$$

parsimonious psychologically based theories are regret theory (Sugden and Loomes, 1982, 1987) and disappointment theory (Sugden and Loomes, 1986). Both incorporate ex ante considerations of ex post psychological feelings: of regret, or rejoicing, in the former and of disappointment, or elation, in the latter. "The fundamental idea behind regret theory is that the psychological experience of 'having x' can be influenced by comparison between x and y that one might have had, had one chosen differently. If, for example, I bet on a horse which fails to win, I may experience something more than a reduction in my wealth: I may also experience a painful sense of regret arising out of the comparison between my current state of wealth and the state that I would have enjoyed, had I not bet" (Sugden, 1991).

3. Setting the Stage: An Illustrative Example

We present a simple example to explain the nature of the experimentation problem faced by the economic agent and to illustrate the analytical set-up of the decision procedure of the experimenter.

In each period, the individual faces the decision of whether to undertake a risky project or a safe one. The risky project pays 1 with unknown probability p and 0 otherwise, while the safe project pays a known sum equal to q with certainty, where $0<q<1$. The agent has a prior on p denoted by ϕ. Each time the agent makes a decision, he faces the trade-off between a sure payoff yielded by the safe project and the cost of experimenting by investing in the risky project, represented by the loss in case of a failure. The agent faces the decision problem in each period along a time horizon of N periods. We indicate with n the number of remaining periods of experimentation. In this simple example for illustrative purposes, without loss of generality, we limit our attention to a finite horizon of three periods of experimentation, i.e. $N=3$.

The example is organized as follows. In section 3.1we provide a full characterization of the experimental decision procedure of a risk neutral agent in the standard framework of expected utility. The strategy characterization is performed by using two alternative methods: backward induction (presented in subsection 3.1.1) and the general method of Bradt, Johnston and Karlin (in subsection 3.1.2). The extension to the case of risk aversion is discussed in subsection 3.1.3. In section 3.2 we reanalyze the bandit problem discussed in section 3.1 in a non-expected utility framework, where the experimenter exhibits loss aversion,n and we characterize the decision procedure of experimentation. We compare the experimentation decision procedures by reference to the loss aversion parameter and we discuss alternative hypotheses on the updating of the reference point in the course of experimentation (in subsections 3.2.1 and 3.2.2). The optimal experimentation procedure of an agent characterized simultaneously by risk aversion and loss aversion is discussed in section 3.3.

3.1 The Bandit Problem in Expected Utility Theory Framework

In the standard framework of expected utility theory, we consider at first a risk neutral agent, whose preferences are represented by a linear utility function. With no loss of generality, the agent is assumed to have a utility level equal to the payoff of the experimentation decision x, i.e. $U(x)=x$.

We will take an advantage of the basic results in 'bandit' theory which establishes that the optimal strategy for this class of games can be obtained by backward induction and we calculate this for the present example. An alternative method to backward induction is the method of Bradt, Johnson and Karlin (1956). Calculations of the optimal strategy are performed by using both methods, i.e. the backward induction and the method of Bradt, Johnson and Karlin.

3.1.1 Backward induction

According to Gittins (1979), the optimal strategy of the experimenter can be characterized by backward induction. Denote with $h=\{h_1,h_2,...\}$ the set of histories that leads to continued experimentation, where h_i , the history in period i, is a binary variable that takes value 1 when the experimentation along the risky project lead to a success in period i or 0 otherwise.

In the initial period the experimenter knows the prior probability ϕ of a success of the experimentation strategy along the risky arm. In each subsequent period, the experimenter updates the prior ϕ through the calculation of a posterior on the basis of the history of play. The information set regarding the experimentation history can be summarized through the help of the following two statistics:

$$\rho_{h_i} \equiv \frac{1}{1+n_{h_i}}$$

and

$$\mu_{h_i} \equiv \frac{s_{h_i}}{1+n_{h_i}}$$

where s_{h_i} denotes the number of recorded successes along the risky project in n_{h_i} trials given a history of play $h=\{h_1,h_2,...,h_i\}$ recorded up to period i. For ease of notation we indicate h_i as subscript of the statistics ρ and μ , to remind us that the statistics are calculated according to the history of play $h=\{h_1,h_2,...,h_i\}$ recorded up to period i.

We denote with ψ_{h_i} the posterior in period $i+1$, which represents the updated probability of winning along the risk arm after the history $h=\{h_1,h_2,...,h_i\}$ is realized. Given the history $h=\{h_1,h_2,...,h_i\}$ realized up to the period i, the updated posterior ψ_{i+1,h_i} of a success on the risky arm in period $i+1$ becomes:

$$\psi_{i+1,h_i}=\frac{\mu_{h_i}+\rho_{h_i}}{1+\rho_{h_i}}$$

We characterize the experimentation strategy of the agent by backward induction. In period 3, the experimenter finds it optimal to play along the risky arm only if the payoff q on the safe arm is lower than the expected payoff of the risky project. The optimal strategy is defined conditional on the state of the world, represented by the history that leads to continued experimentation, which is given by $h=\{h_1,h_2\}$ in the period $N=3$. For each history the optimal experimentation strategy is defined in terms of a threshold value of q below which it is optimal to experiment. In this illustrative example, this condition becomes:

$$q\le\psi_{3,\{h_1,h_2\}}$$

In the last period of the game four states of the world are possible: the experimentation has led to two successes in the first two periods, i.e. $\{h_1,h_2\}=\{1,1\}$, one success in the first period and one failure in the second period, i.e. $\{h_1,h_2\}=\{1,0\}$, one failure in the first period and one success in the second period, i.e. $\{h_1,h_2\}=\{0,1\}$, or two failures, i.e. $\{h_1,h_2\}=\{0,0\}$.

The agent's optimal strategy in period 3 will be to experiment if:

$q\le\psi_{3,\{0,0\}}$, where $\psi_{3,\{0,0\}}=\dfrac{\mu_{\{0,0\}}+\rho_{\{0,0\}}}{1+\rho_{\{0,0\}}}$ if $\{h_1,h_2\}=\{0,0\}$;

where s_{h_i} denotes the number of recorded successes along the risky project in n_{h_i} trials given a history of play $h=\{h_1,h_2,...,h_i\}$ recorded up to period i. For ease of notation we indicate h_i as subscript of the statistics ρ and μ, to remind us that the statistics are calculated according to the history of play $h=\{h_1,h_2,...,h_i\}$ recorded up to period i.

We denote with ψ_{h_i} the posterior in period $i+1$, which represents the updated probability of winning along the risk arm after the history $h=\{h_1,h_2,...,h_i\}$ is realized. Given the history $h=\{h_1,h_2,...,h_i\}$ realized up to the period i, the updated posterior ψ_{i+1,h_i} of a success on the risky arm in period $i+1$ becomes:

$$\psi_{i+1,h_i}=\frac{\mu_{h_i}+\rho_{h_i}}{1+\rho_{h_i}}$$

We characterize the experimentation strategy of the agent by backward induction. In period 3, the experimenter finds it optimal to play along the risky arm only if the payoff q on the safe arm is lower than the expected payoff of the risky project. The optimal strategy is defined conditional on the state of the world, represented by the history that leads to continued experimentation, which is given by $h=\{h_1,h_2\}$ in the period $N=3$. For each history the optimal experimentation strategy is defined in terms of a threshold value of q below which it is optimal to experiment. In this illustrative example, this condition becomes:

$$q\le\psi_{3,\{h_1,h_2\}}$$

In the last period of the game four states of the world are possible: the experimentation has led to two successes in the first two periods, i.e. $\{h_1,h_2\}=\{1,1\}$, one success in the first period and one failure in the second period, i.e. $\{h_1,h_2\}=\{1,0\}$, one failure in the first period and one success in the second period, i.e. $\{h_1,h_2\}=\{0,1\}$, or two failures, i.e. $\{h_1,h_2\}=\{0,0\}$.

The agent's optimal strategy in period 3 will be to experiment if:

$q\le\psi_{3,\{0,0\}}$, where $\psi_{3,\{0,0\}}=\dfrac{\mu_{\{0,0\}}+\rho_{\{0,0\}}}{1+\rho_{\{0,0\}}}$ if $\{h_1,h_2\}=\{0,0\}$;

$q \le \psi_{3,\{0,1\}}$, where $\psi_{3,\{0,1\}} = \dfrac{\mu_{\{0,1\}} + \rho_{\{0,1\}}}{1+\rho_{\{0,1\}}}$ if $\{h_1, h_2\} = \{0,1\}$;[15]

$q \le \psi_{3,\{1,1\}}$, where $\psi_{3,\{1,1\}} = \dfrac{\mu_{\{1,1\}} + \rho_{\{1,1\}}}{1+\rho_{\{1,1\}}}$ if $\{h_1, h_2\} = \{1,1\}$.

Proceeding by backward induction, we calculate the optimal strategy of the experimenter in stage 2 of the game ($N = 2$), through the characterization of the threshold value of q below which the economic agent will continue to experiment with risky project in period 2, given the history of play in period $n = 1$. The experimentation strategy in period 2 is therefore defined contingently on a success or a failure in period 1.

Suppose $\{h_1\} = \{0\}$. If the agent switches to the safe project in $N = 2$, by Lemma in Gittins (1979), the agent sticks with the safe arm in period 3, with an expected payoff equal to $2q$. If the agent plays the risky arm in period 2, two states of the world after period 2 are possible: $\{h_1, h_2\} = \{0,0\}$ and $\{h_1, h_2\} = \{0,1\}$.

We need therefore to distinguish two cases. If $q \le \psi_{3,\{0,0\}}$, the agent will always experiment and has a combined expected payoff from periods 2 and 3 equal to $\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)$.

If $\psi_{3,\{0,0\}} < q < \psi_{3,\{0,1\}}$, the agent will play the risky arm in $n = 3$ only if a success occurs in $N = 2$ and switches to the safe project in $N = 3$ otherwise (i.e., if a failure occurs in $N = 2$). In this case, the combined expected payoff from periods 2 and 3 is equal to $\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)q$.

Henceforth the agent will find it optimal to choose to experiment in period 2 if the expected combined payoff $\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)q$ is higher than the payoff from playing twice on the safe arm, equal to $2q$. Hence, the critical value of q below which it is optimal to experiment in period 2 following a failure in period 1 will be:

$$q \le \frac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}\left(1+\psi_{3,\{0,1\}}\right)$$

Suppose $\{h_1\} = \{1\}$. The optimal experimentation strategy is characterized in an analogous way as a threshold of the parameter q. From a similar calculation, the expected payoff combined from

[15] Note that $\psi_{\{0,1\}} = \psi_{\{1,0\}}$, i.e. the posterior calculated after history $\{h_1, h_2\} = \{1,0\}$ coincides with the posterior after history $\{h_1, h_2\} = \{0,1\}$, where $\psi_{3,\{1,0\}} = \dfrac{\mu_{\{1,0\}} + \rho_{\{1,0\}}}{1+\rho_{\{1,0\}}}$

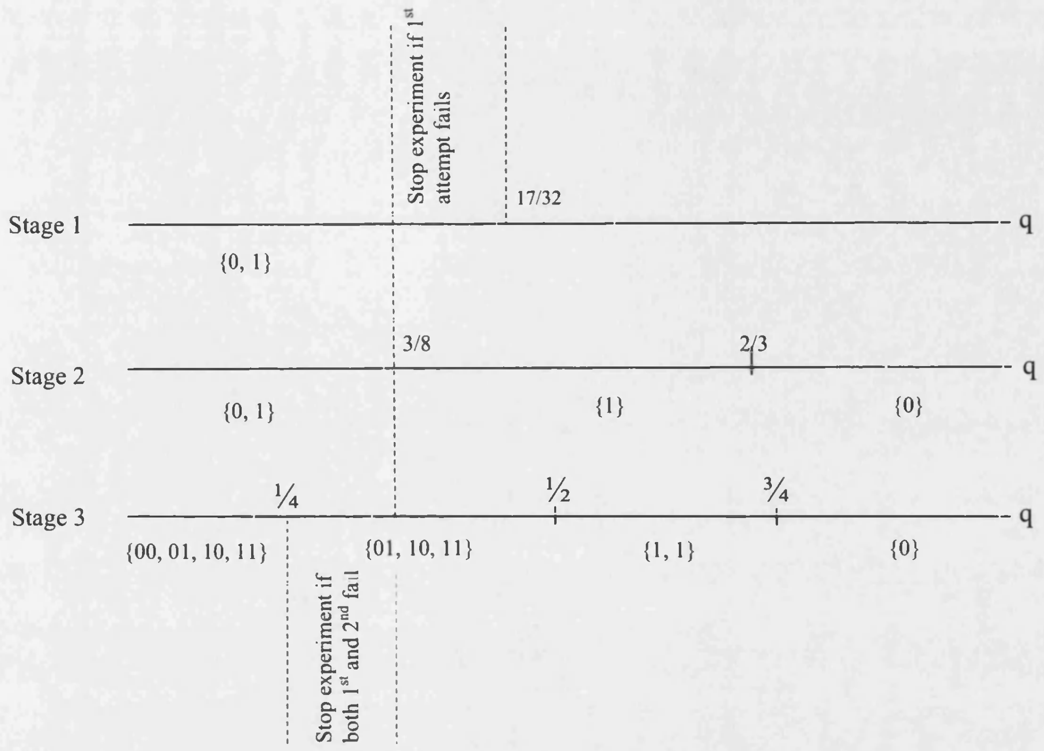


Figure 1: An example: Optimal experimentation strategy if $\phi = \dfrac{1}{2}$

3.1.2 A general method

According to Karlin, Bradt and Johnston (1956), the optimal experimentation strategy can be specified in terms of a sequence $\{k_1, k_2, ..., k_N\}$, where k_i is defined as the number of successes associated with i-th failure, such that the player will stay with the risky arm after the i-th failure only if at least a number of positive results equal to $\sum_{j=1}^{i} k_j$ has been realized. Each sequence $\{k_1, k_2, ..., k_n\}$ attaches to the risky arm an index, called the *dynamic allocation index*, representing the expected payoff of the risky arm given the experimentation strategy $\{k_1, k_2, ..., k_n\}$. We denote the dynamic allocation index with $Q(n, \phi, h)$, which depends on the number of experimentation periods n, the prior ϕ and the history of play h. According to Karlin et al., the experimenter will act optimally by selecting the experimentation strategy in terms of the sequence $\{k_1, k_2, ..., k_n\}$ that maximizes the dynamic allocation index. The characterization of an optimal experimentation period 2 and 3 is equal to $\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}$. The agent will find it optimal to experiment in period 2 after a success in period 1 if the expected combined payoff is higher than the payoff from playing twice along the safe arm, equal to $2q$. This implies that the critical value of q below which it is optimal to experiment in period 2 following a success in period 1 becomes:

$$q \leq \frac{1}{2}\left[\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]$$

Summarizing, the agent's optimal strategy in period 2 will be to experiment if:

$q \leq \dfrac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}\left(1+\psi_{3,\{0,1\}}\right)$ if $\{h_1\}=\{0\}$;

$q \leq \dfrac{1}{2}\left[\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]$ if $\{h_1\}=\{1\}$

Proceeding by backward induction, we need to characterize the player's optimal experimentation strategy in period 1 ($N=1$). Given the prior ϕ, the expected payoff will be $\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)q\right]$. In the first period, the agent will optimally choose to experiment if the payoff from optimal experimentation is higher than the payoff from playing three times on the safe arms, equal to $3q$, i.e. if:

$$q \leq \frac{\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\right\}}{2+\psi_{2,\{0\}}+\phi\left(1+\psi_{2,\{0\}}\right)}$$

Figure 1 summarizes the optimal experimentation strategy of the economic agent under the assumption of an initial prior $\phi=\dfrac{1}{2}$.

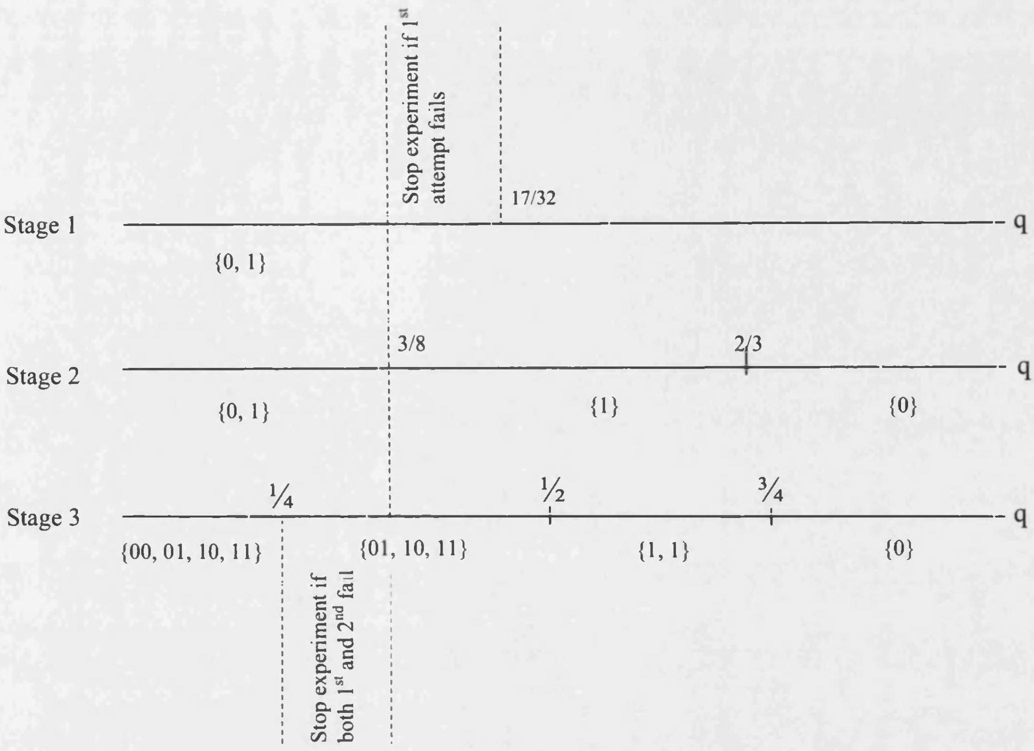


Figure 1: An example: Optimal experimentation strategy if $\phi = \frac{1}{2}$

3.1.2 A general method

According to Karlin, Bradt and Johnston (1956), the optimal experimentation strategy can be specified in terms of a sequence $\{k_1, k_2, ..., k_N\}$, where k_i is defined as the number of successes associated with i-th failure, such that the player will stay with the risky arm after the i-th failure only if at least a number of positive results equal to $\sum_{j=1}^{i} k_j$ has been realized. Each sequence $\{k_1, k_2, ..., k_n\}$ attaches to the risky arm an index, called the *dynamic allocation index*, representing the expected payoff of the risky arm given the experimentation strategy $\{k_1, k_2, ..., k_n\}$. We denote the dynamic allocation index with $Q(n, \phi, h)$, which depends on the number of experimentation periods n, the prior ϕ and the history of play h. According to Karlin et al., the experimenter will act optimally by selecting the experimentation strategy in terms of the sequence $\{k_1, k_2, ..., k_n\}$ that maximizes the dynamic allocation index. The characterization of an optimal experimentation strategy according to Karlin et al. (1956) is based on a "stay-with-the-winner" rule[16], according to which the experimenter keeps experimenting after every success is realized on the risky arm and updates the dynamic allocation index attached to the risky arm after every failure, thereby updating the posterior probability of success given the history of play.

In the first stage of the game, the time horizon is $N = 3$. The experimenter chooses a strategy in terms of a couple $\{k_1, k_2\}$ and calculates the dynamic allocation index $Q(3, \phi, h = \{\varnothing\})$. The dynamic allocation index is determined by applying the following procedure. In order to be indifferent between experimenting and not, the agent calculates an "indifference" boundary by equating the expected payoffs under two alternative strategies: a non-experimentation strategy (play "safe" for all three periods), which yields a sure payoff equal to $3q$ and an experimentation strategy defined according to the couple $\{k_1, k_2\}$. Note that there exist six possible combinations of $\{k_1, k_2\}$. k_1 denotes the number of successes realized before the first failure occurs and required by the agent to stay with the risky arm and it can either take the value 0, 1 or 2 when the time horizon is $N = 3$. The experimenter will stick with the risky arm after a first failure if at least k_1 successes have been realized prior to the first failure. Similarly k_2 - the number of successes realized after the first and before the second failure occurs and required by the agent to stay with the risky arm - can either take the value 0, 1 or 2. The player will not switch to the safe arm after a second failure if at least $k_1 + k_2$ successes have been realized prior to the second failure. Note that any combination of $\{k_1, k_2\}$ should fulfill the requirement $k_1 + k_2 \leq n - 1$.

We characterize, for example, the index attached to the strategy $\{k_1 = 0; k_2 = 1\}$.[17] This strategy means that the agent plays the risky arm in the first stage of the game and sticks with the unsafe arm in the second period either in the case of a success or a failure in the first stage, i.e. $\{h_1\} = \{0\}$ or $\{h_1\} = \{1\}$. In the third stage of the game, the agent keeps playing the risky arm only if at least one success is realized either in first or second stage of the game, i.e., in the case of a history $\{h_1, h_2\} = \{0,1\}$, $\{h_1, h_2\} = \{1,0\}$ or $\{h_1, h_2\} = \{1,1\}$, while he will switch to the safe arm in the case of two successive failures, i.e. after a history $\{h_1, h_2\} = \{0,0\}$. Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

[16] The "stay-with-the-winner" rule is proved to be optimal in Lemma 1 in Karlin et al. (1956).

[17] We refer to the appendix of Chapter 2 for a detailed analytical characterization of the experimenter's optimal strategy.

strategy according to Karlin et al. (1956) is based on a "stay-with-the-winner" rule[16], according to which the experimenter keeps experimenting after every success is realized on the risky arm and updates the dynamic allocation index attached to the risky arm after every failure, thereby updating the posterior probability of success given the history of play.

In the first stage of the game, the time horizon is $N=3$. The experimenter chooses a strategy in terms of a couple $\{k_1,k_2\}$ and calculates the dynamic allocation index $Q(3,\phi,h=\{\varnothing\})$. The dynamic allocation index is determined by applying the following procedure. In order to be indifferent between experimenting and not, the agent calculates an "indifference" boundary by equating the expected payoffs under two alternative strategies: a non-experimentation strategy (play "safe" for all three periods), which yields a sure payoff equal to $3q$ and an experimentation strategy defined according to the couple $\{k_1,k_2\}$. Note that there exist six possible combinations of $\{k_1,k_2\}$. k_1 denotes the number of successes realized before the first failure occurs and required by the agent to stay with the risky arm and it can either take the value 0, 1 or 2 when the time horizon is $N=3$. The experimenter will stick with the risky arm after a first failure if at least k_1 successes have been realized prior to the first failure. Similarly k_2 - the number of successes realized after the first and before the second failure occurs and required by the agent to stay with the risky arm - can either take the value 0, 1 or 2. The player will not switch to the safe arm after a second failure if at least k_1+k_2 successes have been realized prior to the second failure. Note that any combination of $\{k_1,k_2\}$ should fulfill the requirement $k_1+k_2\le n-1$.

We characterize, for example, the index attached to the strategy $\{k_1=0;k_2=1\}$.[17] This strategy means that the agent plays the risky arm in the first stage of the game and sticks with the unsafe arm in the second period either in the case of a success or a failure in the first stage, i.e. $\{h_1\}=\{0\}$ or $\{h_1\}=\{1\}$. In the third stage of the game, the agent keeps playing the risky arm only if at least one success is realized either in first or second stage of the game, i.e., in the case of a history $\{h_1,h_2\}=\{0,1\}$, $\{h_1,h_2\}=\{1,0\}$ or $\{h_1,h_2\}=\{1,1\}$, while he will switch to the safe arm in the case of two successive failures, i.e. after a history $\{h_1,h_2\}=\{0,0\}$. Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

[16] The "stay-with-the-winner" rule is proved to be optimal in Lemma 1 in Karlin et al. (1956).

[17] We refer to the appendix of Chapter 2 for a detailed analytical characterization of the experimenter's optimal strategy.

$\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)q\right]$ and he prefers to play along the risky arm only if:

$$3q\leq\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)q\right]$$

In analytical terms, the agent will optimally choose to experiment if:

$$q\leq\frac{1}{2+\psi_{2,\{0\}}+\phi\left(1-\psi_{2,\{0\}}\right)}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\right\}\equiv\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}}$$

The individual will choose the strategy as a pair of $\{k_1,k_2\}$ in order to maximize the index attached to the risky arm (representing the expected payoff from experimenting). Analytically, the experimenter will choose:

$$Q(3,\phi,h=\{\varnothing\})=\max_{\{k_1,k_2\}}\left\{\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}\right\}$$

The individual will decide to experiment if and only if[18]:

$$q\leq Q(3,\phi,h=\{\varnothing\})$$

In the case in which the prior $\phi=\frac{1}{2}$, $Q(3,\phi,h=\{\varnothing\})=\frac{17}{32}$ associated with strategy $\{k_1=0,k_2=1\}$.

Hence the agent will choose to experiment if $q\leq\frac{17}{32}$, which coincides with the optimal experimentation strategy calculated by using backward induction.

In the case of a history $\{h_1\}=\{1\}$, the agent will keep playing on the risky arm. In the case of a history $\{h_1\}=\{0\}$, the individual will calculate the updated dynamic allocation index to attach to the risky arm and compare q with the expected payoff along the risky arm, given that he experienced a failure in the first run. The time horizon of the decisional problem reduces to $n=2$ and the optimal strategy can be defined in terms of k_1, which can take only two possible values, 0 or 1.

[18] By convention, the individual will experiment when indifferent between the safe and the uncertain project, i.e. when at the boundary.

In case of a failure, the agent will keep experimenting in the last period if $k_1 = 0$, while he will switch to the safe arm in case $k_1 = 1$. The agent will choose the strategy k_1 attached to the highest value of the dynamic allocation index.

For the sake of clarity, let us compute the dynamic allocation index under the assumption that $k_1 = 1$. The combined expected payoff in the remaining periods ($N = 2$ and $N = 3$) is equal to $\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)q$, since he will play the safe arm in the case of a failure in period $N = 2$. This payoff should be compared with the alternative payoff, which is to play the safe arm in both periods, which yields $2q$. An analogous calculation is performed when we assume $k_1 = 0$.

Hence, the agent will experiment on the risky arm at $n = 2$, if:

$$q \leq Q\left(n=2, \phi, h_1 = 0\right)$$

where the dynamic allocation index takes the following form:

$$Q(2, \phi, h_1 = 0) = \max_{k_1=\{0,1\}}\left\{\tilde{q}\Big|_{k_1=0}; \tilde{q}\Big|_{k_1=1}\right\}$$

where:

$$\tilde{q}\Big|_{k_1=0} = \frac{1}{2}\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]$$

$$\tilde{q}\Big|_{k_1=1} = \frac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}\left(1+\psi_{3,\{0,1\}}\right)$$

In the case of the prior $\phi = \frac{1}{2}$, $Q\left(2, \phi = \frac{1}{2}, h_1 = 0\right) = \frac{3}{8}$ associated with strategy $k_1 = 1$. Hence the agent will choose to experiment if $q \leq \frac{3}{8}$, which coincides with the optimal experimentation strategy calculated by backward induction.

In the case of a history $\{h_1, h_2\} = \{1,1\}$ or $\{h_1, h_2\} = \{0,1\}$, the agent will keep playing on the risky arm. In the case of a history $\{h_1, h_2\} = \{0,0\}$, the individual will calculate the updated dynamic allocation index to attach to the risky arm and compare q with the expected payoff along the risky arm, given that he experienced a failure in the first two runs of trial. The decision problem simplifies to a one-stage decision problem. Hence, the agent will play the risky arm if:

$$q \leq Q\left(n=1, \phi, h_1 = h_2 = 0\right)$$

where the dynamic allocation index simplifies to:

$$Q\left(n=1, \phi, h_1 = h_2 = 0\right) = \psi_{3,\{0,0\}}$$

and

$$\psi_{3,\{0,0\}} = \frac{\mu_{\{0,0\}} + \rho_{\{0,0\}}}{1+\rho_{\{0,0\}}} \text{ if } \{h_1,h_2\} = \{0,0\}$$

As shown before, the experimentation strategy coincides with the one calculated by backward induction. In the case in which the prior $\phi = \frac{1}{2}$, $Q\left(n=1, \phi=\frac{1}{2}, h_1 = h_2 = 0\right) = \frac{1}{4}$.

3.1.3 Introducing risk aversion

Up to this point we have followed the standard literature in working with a risk neutral agent who maximizes expected utility. We now extend the standard treatment first to the case of a risk averse agent. We characterize the sequential strategy of the agent facing the same decisional problem described above and using the general method developed in section 3.1.2. In what follows, we set out the main steps in the analysis. Full details of the calculations are given in the Appendix 1 to Chapter 2.

Agents may be characterized by the presence of risk aversion. We model risk aversion by using a standard quadratic utility function (with increasing constant absolute risk aversion) in the form:

$$U(x) = x - rx^2$$

We characterize the optimal experimentation strategy by following the same analytical steps in the general formulation described in section 3.1.2[19].

In the first stage of the game, $N=1$, the individual will choose the strategy as a pair of $\{k_1,k_2\}$ in order to maximize the index attached to the risky arm (representing the expected payoff from experimenting). We denote with Q^{Ra} the dynamic allocation index of a risk averse agent. Analytically, the experimenter will choose:

$$Q^{Ra}(3,\phi,h=\{\varnothing\}) = \max_{\{k_1,k_2\}}\left\{ \tilde{q}\Big|_{\substack{k_1=0\\k_2=0}}, \tilde{q}\Big|_{\substack{k_1=0\\k_2=1}}, \tilde{q}\Big|_{\substack{k_1=0\\k_2=2}}, \tilde{q}\Big|_{\substack{k_1=1\\k_2=0}}, \tilde{q}\Big|_{\substack{k_1=1\\k_2=1}}, \tilde{q}\Big|_{\substack{k_1=2\\k_2=0}} \right\}$$

The individual will decide to experiment if and only if[20]:

$$q \leq Q^{Ra}(3,\phi,h=\{\varnothing\})$$

[19] Calculations are presented in the appendix.

[20] By convention, the individual will experiment when he or she is indifferent between the safe and the uncertain project, i.e. when at the boundary.

for any time horizon of the game n (number of periods of experimentation), prior probability of a success ϕ and game history h.

In the presence of a positive level of risk aversion $r > 0$, the agent will always experiment less for any value of the risk aversion parameter r[21].

3.2 Reanalyzing the Bandit Problem Using Loss Aversion

In this section, we characterize the sequential strategy of the agent facing the same decisional problem described in the previous section and using the general method developed above. Additionally here we assume the player experiences the psychological feelings of disappointment and elation and that he or she takes them into account when evaluating the optimal strategy. The agent's preferences are described by the disappointment-elation utility function a la Sugden and Loomes (1986)[22], according to which the agent gets utility directly from the payoff (denoted with x) and additionally he experiences disappointment if the payoff x is lower than the expected one (denoted with $\bar{x}$) and elation on the contrary. Analytically:

$$U(x) = x + D(x - \bar{x})$$

where $D(x - \bar{x})$ is the disappointment-elation function, which takes the following form:

$$D(x) = \begin{cases} x - \bar{x} \; for \; x \geq \bar{x} \\ \lambda(x - \bar{x}) \; for \; x < \bar{x}, \lambda > 1 \end{cases}$$

According to the specified disappointment-elation function, the parameter λ measures loss aversion, which can be interpreted as the ratio at which the experimenter values a loss compared to a gain, i.e. gaining \$1 more than expected provides an additional utility of \$1, while losing \$1 leads to a higher disutility, equal to λ. Loss aversion arises when $\lambda > 1$.

In the example discussed here, the payoff x can either take the value 1, with an associated utility equal to $U(1) = 2 - \bar{x}$ in the case of a success or 0 in the case of a failure, with utility $U(0) = -\lambda \bar{x}$.

The expected payoff $\bar{x}$ of the risky arm plays the role of a reference point in the experimentation process.

We now introduce a key distinction, which relates to the following question: in the course of experimentation, the agent may choose to use new information to reconsider or update the original

[21] See Andersom (2001) for an analogous result.

[22] The choice of modeling preferences according to the disappointment-elation utility function a la Sugden and Loomes (1986)among others is motivated by the need of working in a dynamic framework.

In the case of the prior $\phi = \frac{1}{2}$, $Q^{Ra}\left(3,\phi,h=\{\varnothing\}\right)=\frac{17}{32}(1-r)$ associated with strategy $\{k_1=0,k_2=1\}$. In stage 2 of the game ($N=2$), following the history $\{h_1\}=\{1\}$, the agent will keep playing on the risky arm. In the case of a history $\{h_1\}=\{0\}$, the player will experiment on the risky arm at $n=2$, if

$$q \le Q^{Ra}\left(n=2,\phi,h_1=0\right)$$

where the dynamic allocation index takes the following form (the derivation is given in the Appendix):

$$Q^{Ra}(2,\phi,h_1=0)=\max_{k_1=\{0,1\}}\left\{\left.\tilde{q}^{Ra}\right|_{k_1=0};\left.\tilde{q}^{Ra}\right|_{k_1=1}\right\}$$

where

$$\left.\tilde{q}^{Ra}\right|_{k_1=0}=\frac{1}{2}(1-r)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]$$

$$\left.\tilde{q}^{Ra}\right|_{k_1=0}=\frac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}(1-r)\left(1+\psi_{3,\{0,1\}}\right)$$

In the case of the prior $\phi=\frac{1}{2}$, $Q^{Ra}\left(2,\phi=\frac{1}{2},h_1=0\right)=\frac{3}{8}(1-r)$ associated with strategy $k_1=1$.

In the case of a history $\{h_1,h_2\}=\{1,1\}$ or $\{h_1,h_2\}=\{0,1\}$, the agent will keep playing on the risky arm. In the case of a history $\{h_1,h_2\}=\{0,0\}$, the agent will play the risky arm if:

$$q \le Q^{Ra}\left(n=1,\phi,h_1=h_2=0\right)$$

where the dynamic allocation index simplifies to:

$$Q^{Ra}\left(n=1,\phi,h_1=h_2=0\right)=\left(1-r\right)\psi_{3,\{0,0\}}$$

and

$$\psi_{3,\{0,0\}}=\frac{\mu_{\{0,0\}}+\rho_{\{0,0\}}}{1+\rho_{\{0\}}} \text{ if } \{h_1,h_2\}=\{0,0\}$$

In case the prior $\phi=\frac{1}{2}$, $Q\left(n=1,\phi=\frac{1}{2},h_1=h_2=0\right)=\frac{1}{4}(1-r)$.

To study whether a risk averse player will experiment more than a risk neutral one, we compare the dynamic allocation index in the presence of risk aversion with the index in the presence of risk neutrality. In order to induce more experimentation, the following condition should be satisfied:

$$Q^{Ra}\left(n,\phi,h\right)\ge Q^{EUT}\left(n,\phi,h\right)$$

for any time horizon of the game n (number of periods of experimentation), prior probability of a success ϕ and game history h.

In the presence of a positive level of risk aversion $r > 0$, the agent will always experiment less for any value of the risk aversion parameter r[21].

3.2 Reanalyzing the Bandit Problem Using Loss Aversion

In this section, we characterize the sequential strategy of the agent facing the same decisional problem described in the previous section and using the general method developed above. Additionally here we assume the player experiences the psychological feelings of disappointment and elation and that he or she takes them into account when evaluating the optimal strategy. The agent's preferences are described by the disappointment-elation utility function a la Sugden and Loomes (1986)[22], according to which the agent gets utility directly from the payoff (denoted with x) and additionally he experiences disappointment if the payoff x is lower than the expected one (denoted with $\bar{x}$) and elation on the contrary. Analytically:

$$U(x) = x + D(x - \bar{x})$$

where $D(x - \bar{x})$ is the disappointment-elation function, which takes the following form:

$$D(x) = \begin{cases} x - \bar{x} \ for\ x \geq \bar{x} \\ \lambda(x - \bar{x})\ for\ x < \bar{x},\ \lambda > 1 \end{cases}$$

According to the specified disappointment-elation function, the parameter λ measures loss aversion, which can be interpreted as the ratio at which the experimenter values a loss compared to a gain, i.e. gaining \$1 more than expected provides an additional utility of \$1, while losing \$1 leads to a higher disutility, equal to λ. Loss aversion arises when $\lambda > 1$.

In the example discussed here, the payoff x can either take the value 1, with an associated utility equal to $U(1) = 2 - \bar{x}$ in the case of a success or 0 in the case of a failure, with utility $U(0) = -\lambda\bar{x}$.

The expected payoff $\bar{x}$ of the risky arm plays the role of a reference point in the experimentation process.

We now introduce a key distinction, which relates to the following question: in the course of experimentation, the agent may choose to use new information to reconsider or update the original

[21] See Andersom (2001) for an analogous result.

[22] The choice of modeling preferences according to the disappointment-elation utility function a la Sugden and Loomes (1986)among others is motivated by the need of working in a dynamic framework.

reference point. There is no persuasive a priori argument that favors one or other of these two cases. So, in what follows, we explore both cases. In the next subsection we assume "no updating" of the reference point. In section 3.2.2 we assume it is updated. The effect of this updating will be to modify the quantitative prediction, while leaving the qualitative features of the solution unchanged.

3.2.1 The "No Updating" Case

We now proceed with the analysis under "no updating". In the calculation performed here $\bar{x}$ is assumed to be the expected payoff given the prior ϕ.

We proceed following the exact analytical steps shown in the previous section[23] to calculate the dynamic allocation index in each stage of the game and to provide a full characterization of the experimentation strategy of the player[24] – summarized in Table 2. The critical values in the absence and in the presence of loss-aversion are provided in Table 2 for each stage of the game.

We characterize the optimal experimentation strategy of a loss averse agent by following the same analytical steps in the general formulation described in section 3.1.2. Full details of calculation are provided in the Appendix to Chapter 2.

In the first stage of the game, $N=1$, the individual will choose the strategy as a pair of $\{k_1,k_2\}$ in order to maximize the index attached to the risky arm (representing the expected payoff from experimenting). We denote with Q^{La} the dynamic allocation index of a loss averse agent. Analytically, the experimenter will choose:

$$Q^{La}\left(3,\phi,h=\{\varnothing\}\right)=\max_{\{k_1,k_2\}}\left\{\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}\right\}$$

The individual will decide to experiment if and only if:

$$q\le Q^{La}\left(3,\phi,h=\{\varnothing\}\right)$$

In the case of a history $\{h_1\}=\{1\}$, the agent will keep playing on the risky arm. In the case of a history $\{h_1\}=\{0\}$, the player will experiment on the risky arm at $n=2$, if :

$$q\le Q^{La}\left(n=2,\phi,h_1=0\right)$$

where the dynamic allocation index takes the following form (the derivation is given in the Appendix):

[23] Full characterization of the optimal strategies and relative calculations are presented in the appendix.

[24] There exists no general proof that backward induction works independently of the context exists. The equivalence theorem works only for expected utility. Therefore we check directly in the case of present example that the calculation coincides. For an example in which the backward induction does not work see Berry, Fristedt (1985).

$$Q^{La}(2,\phi,h_1=0)=\max_{k_1=\{0,1\}}\left\{\tilde{q}\Big|_{k_1=0};\tilde{q}\Big|_{k_1=1}\right\}$$

where

$$\tilde{q}^{La}\Big|_{k_1=0}=\frac{1}{2}\left[\begin{array}{l}\psi_{2,\{0\}}\left[\left(1+\psi_{3,\{0,1\}}\right)\left(2-\psi_{2,\{0\}}\right)-\left(1-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)\right]+\\+\left(1-\psi_{2,\{0\}}\right)\left[-\left(2-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)+\psi_{3,\{0,1\}}\left(2-\psi_{2,\{0\}}\right)\right]\end{array}\right]$$

$$\tilde{q}^{La}\Big|_{k_1=1}=\frac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}\left[\left(2-\psi_{2,\{0\}}\right)\left(1+\psi_{3,\{0,1\}}\right)-\lambda\left(1-\psi_{2,\{0\}}\psi_{3,\{0,1\}}\right)\right]$$

In the case of a history $\{h_1,h_2\}=\{1,1\}$ or $\{h_1,h_2\}=\{0,1\}$, the agent will keep playing on the risky arm. In the case of a history $\{h_1,h_2\}=\{0,0\}$, the agent will play the risky arm if:

$$q\le Q^{La}\left(n=1,\phi,h_1=h_2=0\right)$$

where the dynamic allocation index simplifies to:

$$Q^{La}\left(n=1,\phi,h_1=h_2=0\right)=\left(2-\psi_{3,\{0,0\}}-\lambda\left(1-\psi_{3,\{0,0\}}\right)\right)\psi_{3,\{0,0\}}$$

and

$$\psi_{3,\{0,0\}}=\frac{\mu_{\{0,0\}}+\rho_{\{0,0\}}}{1+\rho_{\{0,0\}}}\text{ if }\{h_1,h_2\}=\{0,0\}$$

In the case in which the prior $\phi=\frac{1}{2}$, $Q^{La}\left(n=1,\phi=\frac{1}{2},h_1=h_2=0\right)=\frac{1}{16}(7-3\lambda)$.

Figure 2 summarizes the change in the critical values of dynamic allocation index in the presence of loss aversion for a prior $\phi=\frac{1}{2}$.

$$Q^{La}(2,\phi,h_1=0)=\max_{k_1=\{0,1\}}\left\{\tilde{q}\Big|_{k_1=0};\tilde{q}\Big|_{k_1=1}\right\}$$

where

$$\tilde{q}^{La}\Big|_{k_1=0}=\frac{1}{2}\left[\begin{array}{l}\psi_{2,\{0\}}\left[\left(1+\psi_{3,\{0,1\}}\right)\left(2-\psi_{2,\{0\}}\right)-\left(1-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)\right]+\\+\left(1-\psi_{2,\{0\}}\right)\left[-\left(2-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)+\psi_{3,\{0,1\}}\left(2-\psi_{2,\{0\}}\right)\right]\end{array}\right]$$

$$\tilde{q}^{La}\Big|_{k_1=1}=\frac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}\left[\left(2-\psi_{2,\{0\}}\right)\left(1+\psi_{3,\{0,1\}}\right)-\lambda\left(1-\psi_{2,\{0\}}\psi_{3,\{0,1\}}\right)\right]$$

In the case of a history $\{h_1,h_2\}=\{1,1\}$ or $\{h_1,h_2\}=\{0,1\}$, the agent will keep playing on the risky arm. In the case of a history $\{h_1,h_2\}=\{0,0\}$, the agent will play the risky arm if:

$$q\leq Q^{La}\left(n=1,\phi,h_1=h_2=0\right)$$

where the dynamic allocation index simplifies to:

$$Q^{La}\left(n=1,\phi,h_1=h_2=0\right)=\left(2-\psi_{3,\{0,0\}}-\lambda\left(1-\psi_{3,\{0,0\}}\right)\right)\psi_{3,\{0,0\}}$$

and

$$\psi_{3,\{0,0\}}=\frac{\mu_{\{0,0\}}+\rho_{\{0,0\}}}{1+\rho_{\{0,0\}}}\text{ if }\{h_1,h_2\}=\{0,0\}$$

In the case in which the prior $\phi=\frac{1}{2}$, $Q^{La}\left(n=1,\phi=\frac{1}{2},h_1=h_2=0\right)=\frac{1}{16}(7-3\lambda)$.

Figure 2 summarizes the change in the critical values of dynamic allocation index in the presence of loss aversion for a prior $\phi=\frac{1}{2}$.

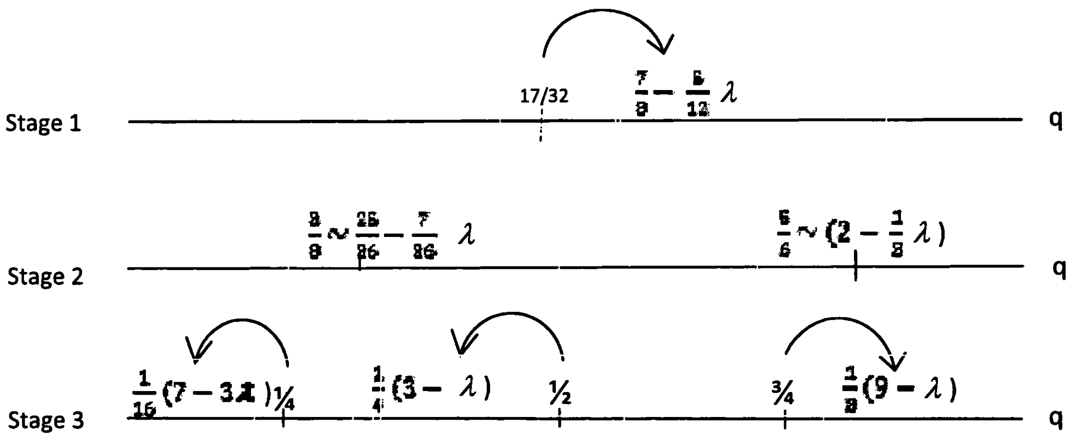


Figure 2: *Introducing loss aversion:*

The change in critical values relative to Figure 1, when $\lambda \neq 0$

To study whether a loss averse player will experiment more than a risk neutral one, we compare the dynamic allocation index in the presence of loss aversion with the index in the presence of risk neutrality. In order to induce more experimentation, the following condition should be satisfied:

$$Q^{La}(n,\phi,h) \geq Q^{EUT}(n,\phi,h)$$

for any time horizon of the game n, a prior probability of a success ϕ and a game history h.

Figure 2a and 2b summarize the range of values for loss-aversion, measured by the parameter λ, that allow for more experimentation in the game presented in this section. The figure provides a straight comparison between the degree of experimentation with and without loss-aversion, providing the upper bound value on the parameter of loss-aversion that allows for more experimentation in the presence of loss-aversion. This upper bound is calculated as the highest value of λ, such that the dynamic allocation index in the presence of loss-aversion is higher than the index in absence of loss-aversion, i.e.:

$$Q^{La}(n,\phi,h) \geq Q^{EUT}(n,\phi,h)$$

3.2.1.1 A clarification: The Non-Equivalence of the $\lambda = 1$ case and the Risk Neutrality case

A subject who does not display loss aversion[25] is characterized by $\lambda = 1$. It is worth to note that the case of a loss neutral subject does not collapse to the case of risk-neutrality. When $\lambda = 1$, the utility function of the agent simplifies:

$$U(x) = \begin{cases} 2 - \bar{x} & if\ x = 1 \\ -\bar{x} & if\ x = 0 \end{cases}$$

The utility function of a risk neutral agent is $U(x) = x$.

Hence, the cutoff of optimal experimentation of a loss-neutral agent, represented by the Gittins index $Q(n, \phi, h)$ calculated setting $\lambda = 1$, does not coincide with the cutoff of optimal experimentation of a risk-neutral agent, represented by the Gittins index $Q(n, \phi, h)$ calculated setting the disappointment-elation function $D(\bullet) = 0$

In the case of loss neutrality – when the parameter of loss aversion is $\lambda = 1$ - there are both an elation and a disappointment effect. In the case of risk neutrality, the disappointment-elation function is set equal to zero and there are neither elation nor disappointment and the utility function simplifies to the linear case, $U(x) = x$. The cases of loss neutrality and risk neutrality are related by an affine transformation. However, they are not equivalent because of the role of the status quo point.

The equivalence of the two cases holds when $N = 1$. The proof is straightforward. Let p be the probability to win on the risky arm and $\bar{x}$ the status quo.

The optimal cutoff of the risk-neutral agent is $q_{Risk-neutral} = p$.

The optimal cutoff of the loss-neutral agent is: $q_{Loss-neutral} = p(2 - \bar{x}) - (1 - p)\bar{x} = 2p - \bar{x}$

Note that $\bar{x}$ - the expected payoff of the risky arm – plays the role of the status quo point in the example presented here. In the last stage of the experimentation game, $N = 1$, the expected payoff of the risky arm $\bar{x}$ coincides with p, the probability to win on the risky arm. It follows immediately that the optimal cutoff of the loss-neutral agent reduces to p and coincides with the optimal cutoff of the risk-neutral agent.

[25] We refer to the case $\lambda = 1$ as loss neutrality.

for any time horizon of the game n, a prior probability of a success ϕ and a game history h. Figure 2a illustrates the upper bound on λ in the first stage of the game. Figure 2b summarizes the upper bound on λ in the first and second period of experimentation. Note that in the last period of experimentation the presence of loss-aversion induces less experimentation for any value of $\lambda > 1$.

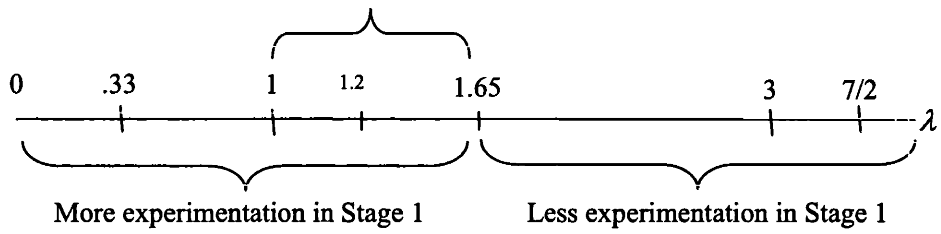


Figure 2a: *Loss-aversion parameter and First Stage Experimentation*

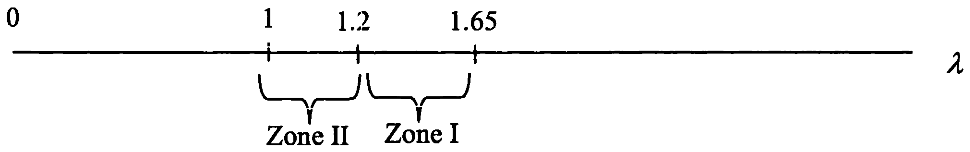


Figure 2b: *Loss-aversion parameter and Optimal Experimentation in the 3-Stage Game*

3.2.1.1 A clarification: The Non-Equivalence of the $\lambda = 1$ case and the Risk Neutrality case

A subject who does not display loss aversion[25] is characterized by $\lambda = 1$. It is worth to note that the case of a loss neutral subject does not collapse to the case of risk-neutrality. When $\lambda = 1$, the utility function of the agent simplifies:

$$U(x) = \begin{cases} 2 - \bar{x} & \text{if } x = 1 \\ -\bar{x} & \text{if } x = 0 \end{cases}$$

The utility function of a risk neutral agent is $U(x) = x$.

Hence, the cutoff of optimal experimentation of a loss-neutral agent, represented by the Gittins index $Q(n,\phi,h)$ calculated setting $\lambda = 1$, does not coincide with the cutoff of optimal experimentation of a risk-neutral agent, represented by the Gittins index $Q(n,\phi,h)$ calculated setting the disappointment-elation function $D(\bullet) = 0$

In the case of loss neutrality – when the parameter of loss aversion is $\lambda = 1$ - there are both an elation and a disappointment effect. In the case of risk neutrality, the disappointment-elation function is set equal to zero and there are neither elation nor disappointment and the utility function simplifies to the linear case, $U(x) = x$. The cases of loss neutrality and risk neutrality are related by an affine transformation. However, they are not equivalent because of the role of the status quo point.

The equivalence of the two cases holds when $N = 1$. The proof is straightforward. Let p be the probability to win on the risky arm and $\bar{x}$ the status quo.

The optimal cutoff of the risk-neutral agent is $q_{Risk-neutral} = p$.

The optimal cutoff of the loss-neutral agent is: $q_{Loss-neutral} = p(2-\bar{x}) - (1-p)\bar{x} = 2p - \bar{x}$

Note that $\bar{x}$ - the expected payoff of the risky arm – plays the role of the status quo point in the example presented here. In the last stage of the experimentation game, $N = 1$, the expected payoff of the risky arm $\bar{x}$ coincides with p, the probability to win on the risky arm. It follows immediately that the optimal cutoff of the loss-neutral agent reduces to p and coincides with the optimal cutoff of the risk-neutral agent.

[25] We refer to the case $\lambda = 1$ as loss neutrality.

However, the equivalence does not hold in case the agent chooses a different status quo point $\bar{x}$.

Additionally, the equivalence does not hold even when the agent chooses $\bar{x}$ as the expected payoff of the risky arm, if the time horizon of the game is equal or bigger than two periods, $N \geq 2$. In the following we illustrate the non-equivalence of the experimentation strategy of the loss-neutral agent and of the risk-neutral agent when $N = 2$.

In stage *N=2* (with two periods of experimentation ahead), after a history of one failures: $h = \{0\}$, the updated probability of a success is equal to $\frac{1}{3}$ and the expected payoff is equal to $\bar{x}_{\{0\}} = \frac{1}{3}$.

The utility function of a loss neutral agent is equal to

$$U(x) = \begin{cases} 2 - \bar{x}_{\{0\}} = \frac{5}{3} & \text{if} \quad \text{success} \\ -\bar{x}_{\{0\}} = -\frac{1}{3} & \text{if} \quad \text{fail} \end{cases}$$

We should distinguish two cases along the risky arm:

- if $q \leq \frac{1}{4}$, the agent chooses to experiment in both periods
- if $\frac{1}{4} < q \leq \frac{1}{2}$, the agent chooses to experiment in period N=2 and switches to the safe arm in period N=3.

If $q \leq \frac{1}{4}$, the expected payoff of a loss-neutral agent who experiments in both periods becomes equal to $\frac{2}{3}$.
In alternative, the agent can play twice on safe arm, with a payoff equal to 2q. Hence, the cutoff value of the agent becomes at *N=2*:

$$q \leq \frac{1}{3}$$

If $\frac{1}{4} < q \leq \frac{1}{2}$, the expected payoff of a loss-neutral agent is $\frac{5}{9} + \frac{2}{3}q$

In alternative, the agent can play twice on safe arm, with a payoff equal to 2q. Hence:

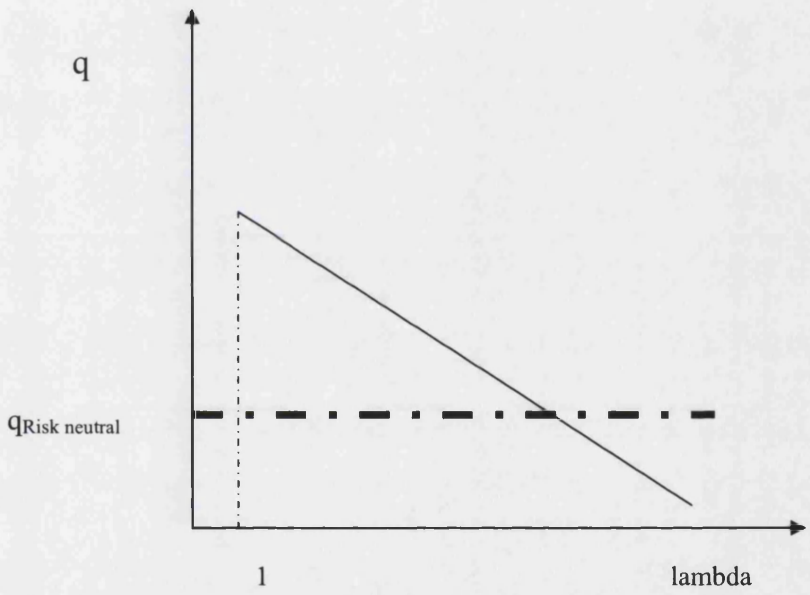
$$q \leq \frac{5}{12}$$

The optimal cutoff value of a a loss-neutral agent takes the following form:

$$Q(2, h = \{0\}) = \max\left\{\frac{1}{3}; \frac{5}{12}\right\} = \frac{5}{12}$$

While the risk-neutral agent finds it optimal to experiment when the payoff of the sure arm is lower or equal to $\frac{1}{3}$, the loss-neutral agent finds it optimal to experiment when the payoff of the sure arm is lower or equal to $\frac{5}{12}$. We provide the full characterization of the loss-neutrality case in the Appendix to Chapter 2.

The following graph illustrates how the cutoff of optimal experimentation behaves in the loss-neutrality case, i.e., when $\lambda = 1$. The cutoff of optimal experimentation of the agent with loss aversion is shown to be decreasing in the level of loss aversion λ. In the special case of loss-neutrality, $\lambda = 1$, the cutoff of the subject q does not coincide with the cutoff of the risk-neutral agent, $q_{\text{Risk neutral}}$, illustrated as the horizontal dotted line in the Figure below.[26]



[26] Calculations for each stage in the game are in the Appendix to Chapter 2.

However, the equivalence does not hold in case the agent chooses a different status quo point $\bar{x}$.

Additionally, the equivalence does not hold even when the agent chooses $\bar{x}$ as the expected payoff of the risky arm, if the time horizon of the game is equal or bigger than two periods, $N \geq 2$. In the following we illustrate the non-equivalence of the experimentation strategy of the loss-neutral agent and of the risk-neutral agent when $N = 2$.

In stage *N=2* (with two periods of experimentation ahead), after a history of one failures: $h = \{0\}$, the updated probability of a success is equal to $\frac{1}{3}$ and the expected payoff is equal to $\bar{x}_{\{0\}} = \frac{1}{3}$.

The utility function of a loss neutral agent is equal to

$$U(x)=\begin{cases} 2-\bar{x}_{\{0\}} = \frac{5}{3} & if \quad success \\ -\bar{x}_{\{0\}} = -\frac{1}{3} & if \quad fail \end{cases}$$

We should distinguish two cases along the risky arm:

- if $q \leq \frac{1}{4}$, the agent chooses to experiment in both periods
- if $\frac{1}{4} < q \leq \frac{1}{2}$, the agent chooses to experiment in period N=2 and switches to the safe arm in period N=3.

If $q \leq \frac{1}{4}$, the expected payoff of a loss-neutral agent who experiments in both periods becomes equal to $\frac{2}{3}$.
In alternative, the agent can play twice on safe arm, with a payoff equal to 2q. Hence, the cutoff value of the agent becomes at *N=2*:

$$q \leq \frac{1}{3}$$

If $\frac{1}{4} < q \leq \frac{1}{2}$, the expected payoff of a loss-neutral agent is $\frac{5}{9} + \frac{2}{3}q$

In alternative, the agent can play twice on safe arm, with a payoff equal to 2q. Hence:

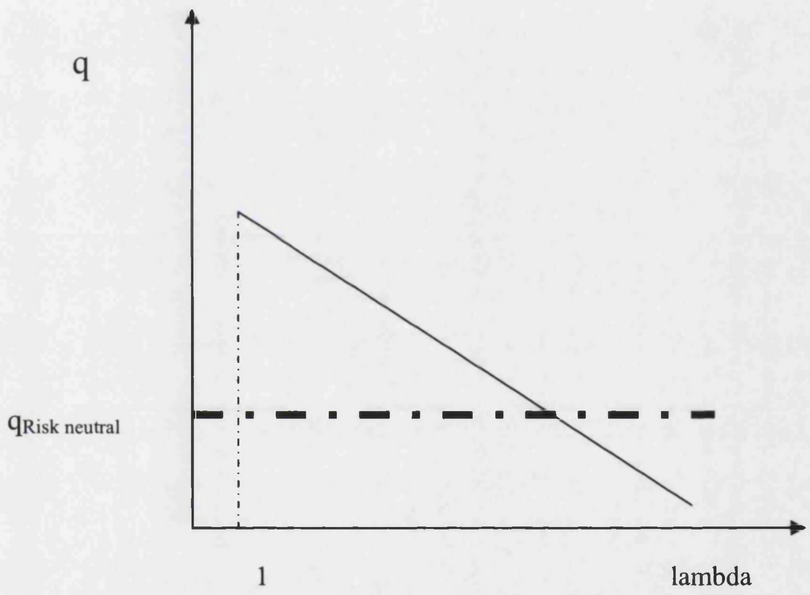
$$q \leq \frac{5}{12}$$

The optimal cutoff value of a a loss-neutral agent takes the following form:

$$Q(2, h = \{0\}) = \max\left\{\frac{1}{3}; \frac{5}{12}\right\} = \frac{5}{12}$$

While the risk-neutral agent finds it optimal to experiment when the payoff of the sure arm is lower or equal to $\frac{1}{3}$, the loss-neutral agent finds it optimal to experiment when the payoff of the sure arm is lower or equal to $\frac{5}{12}$. We provide the full characterization of the loss-neutrality case in the Appendix to Chapter 2.

The following graph illustrates how the cutoff of optimal experimentation behaves in the loss-neutrality case, i.e., when $\lambda = 1$. The cutoff of optimal experimentation of the agent with loss aversion is shown to be decreasing in the level of loss aversion λ. In the special case of loss-neutrality, $\lambda = 1$, the cutoff of the subject q does not coincide with the cutoff of the risk-neutral agent, $q_{\text{Risk neutral}}$, illustrated as the horizontal dotted line in the Figure below.[26]



[26] Calculations for each stage in the game are in the Appendix to Chapter 2.

3.2.2 The "Updating" case

In this section we allow the player to update the expected payoff $\bar{x}$ when calculating $Q(n,\phi,h)$ on the basis of the implied evolution of the game. We characterize the optimal experimentation strategy of a loss averse agent by following the same analytical steps in the general formulation described in section 3.1.2. Full details of calculation are provided in the Appendix to Chapter 2. Note that in each stage of the game the analytical derivation of the updated $\bar{x}$ is required in order to characterize fully the equilibrium strategy of the player.

In the first stage of the game, $N=1$, the individual will choose the strategy as a pair of $\{k_1,k_2\}$ in order to maximize the index attached to the risky arm (representing the expected payoff from experimenting). Analytically, the experimenter will choose:

$$Q^{La,Updating}\left(3,\phi,h=\{\varnothing\}\right)=\max_{\{k_1,k_2\}}\left\{\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}\right\}$$

The individual will decide to experiment if and only if:

$$q\le Q^{La,Updating}\left(3,\phi,h=\{\varnothing\}\right)$$

In the second stage of the game, in the case of a history $\{h_1\}=\{1\}$, the agent will keep playing on the risky arm. In the case of a history $\{h_1\}=\{0\}$, the player will experiment on the risky arm at $N=2$, if :

$$q\le Q^{La,Updating}\left(n=2,\phi,h_1=0\right)$$

where the dynamic allocation index takes the following form (the derivation is given in the Appendix):

$$Q^{La,Updating}(2,\phi,h_1=0)=\max_{k_1=\{0,1\}}\left\{\tilde{q}\Big|_{k_1=0};\tilde{q}\Big|_{k_1=1}\right\}$$

and

$$\tilde{q}^{La}\Big|_{k_1=0}=\frac{\psi_{2,\{0\}}\begin{bmatrix}2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\\-\left(1-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{3,\{0,1\}}\right)\end{bmatrix}+\left(1-\psi_{2,\{0\}}\right)\begin{bmatrix}-\lambda\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\\-\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{3,\{0,1\}}\end{bmatrix}}{2}$$

$$\tilde{q}\Big|_{h_1=1} = \frac{\psi_{2,\{0\}}\begin{bmatrix} 2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)- \\ -\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{3,\{0,1\}} \end{bmatrix} - \left(1-\psi_{2,\{0\}}\right)\lambda\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}$$

In the case in which the prior $\phi=\frac{1}{2}$, $Q^{La,Updating}\left(2,\phi=\frac{1}{2},h_1=0\right)=\frac{2}{3}-\frac{11}{48}\lambda$.

In the case of a history $\{h_1,h_2\}=\{1,1\}$ or $\{h_1,h_2\}=\{0,1\}$, the agent will keep playing the risky arm.

In the case of a history $\{h_1,h_2\}=\{0,0\}$, the agent will play the risky arm if:

$$q \leq Q^{La,Updating}\left(n=1,\phi,h_1=h_2=0\right)$$

where the dynamic allocation index simplifies to:

$$Q^{La}\left(n=1,\phi,h_1=h_2=0\right)=\psi_{3,\{0,0\}}\left(2-\psi_{3,\{0,0\}}\right)-\left(1-\psi_{3,\{0,0\}}\right)\left(\lambda\psi_{3,\{0,0\}}\right)$$

and

$$\psi_{3,\{0,0\}}=\frac{\mu_{\{0,0\}}+\rho_{\{0,0\}}}{1+\rho_{\{0,0\}}} \text{ if } \{h_1,h_2\}=\{0,0\}$$

In the case in which the prior $\phi=\frac{1}{2}$, $Q^{La,Updating}\left(n=1,\phi=\frac{1}{2},h_1=h_2=0\right)=\frac{1}{16}(7-3\lambda)$. Note that in the last stage of experimentation the updating does not change the prior.

Figure 3 summarizes the change in the critical values of dynamic allocation index in the presence of loss aversion and updating of the reference point for a prior $\phi=\frac{1}{2}$

$$\tilde{q}\Big|_{k_1=1} = \frac{\psi_{2,\{0\}}\begin{bmatrix} 2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)- \\ -\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{3,\{0,1\}} \end{bmatrix} - \left(1-\psi_{2,\{0\}}\right)\lambda\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}$$

In the case in which the prior $\phi=\frac{1}{2}$, $Q^{La,Updating}\left(2,\phi=\frac{1}{2},h_1=0\right)=\frac{2}{3}-\frac{11}{48}\lambda$.

In the case of a history $\{h_1,h_2\}=\{1,1\}$ or $\{h_1,h_2\}=\{0,1\}$, the agent will keep playing the risky arm.

In the case of a history $\{h_1,h_2\}=\{0,0\}$, the agent will play the risky arm if:

$$q \le Q^{La,Updating}\left(n=1,\phi,h_1=h_2=0\right)$$

where the dynamic allocation index simplifies to:

$$Q^{La}\left(n=1,\phi,h_1=h_2=0\right)=\psi_{3,\{0,0\}}\left(2-\psi_{3,\{0,0\}}\right)-\left(1-\psi_{3,\{0,0\}}\right)\left(\lambda\psi_{3,\{0,0\}}\right)$$

and

$$\psi_{3,\{0,0\}}=\frac{\mu_{\{0,0\}}+\rho_{\{0,0\}}}{1+\rho_{\{0,0\}}} \text{ if } \{h_1,h_2\}=\{0,0\}$$

In the case in which the prior $\phi=\frac{1}{2}$, $Q^{La,Updating}\left(n=1,\phi=\frac{1}{2},h_1=h_2=0\right)=\frac{1}{16}(7-3\lambda)$. Note that in the last stage of experimentation the updating does not change the prior.

Figure 3 summarizes the change in the critical values of dynamic allocation index in the presence of loss aversion and updating of the reference point for a prior $\phi=\frac{1}{2}$

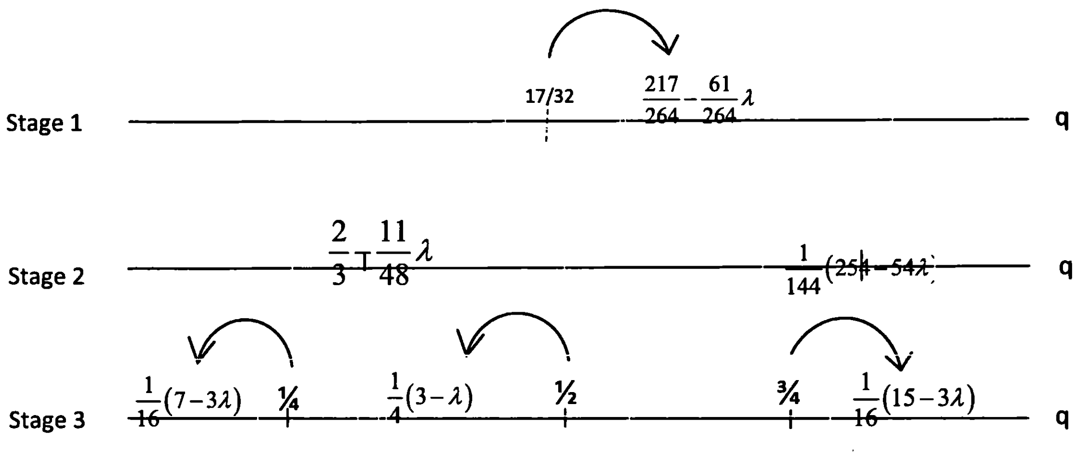


Figure 3: *The change in critical values relative to Figure 1, when* $\lambda \neq 0$ *and the experimenter updates his reference point*

Figures 3a and 3b summarize the range of values for loss-aversion, measured by the parameter λ, that allows for more experimentation in the game. The figure provides a straight comparison on the degree of experimentation with and without loss-aversion, providing the upper bound value on the parameter of loss-aversion that allows for more experimentation in the presence of loss-aversion. As before, this upper bound is calculated as the highest value of λ, such that the dynamic allocation index in the presence of loss-aversion is higher than the index in the absence of loss-aversion, i.e.:

$$Q^{La,Updating}(n,\phi,h) \geq Q^{EUT}(n,\phi,h)$$

for any time horizon of the game n, a prior probability of a success ϕ and a game history h. Figure 3a summarizes the upper bound on λ in the first and second period of experimentation.

The choice of the reference point confirms the general result, according to which a loss averse player may choose to experiment more than a loss neutral one which chooses to experiment. The updating of $\bar{x}$ reduces the upper bound on λ, the highest degree of loss aversion that can be born in order to experiment more with respect to the case of no-updating. Again, as before, in the last period of experimentation the presence of loss-aversion induces less experimentation for any value of $\lambda > 1$.

Figure 3a summarizes the range of values for loss-aversion, measured by the parameter λ, that allow for more experimentation in the first and second stage of the game.

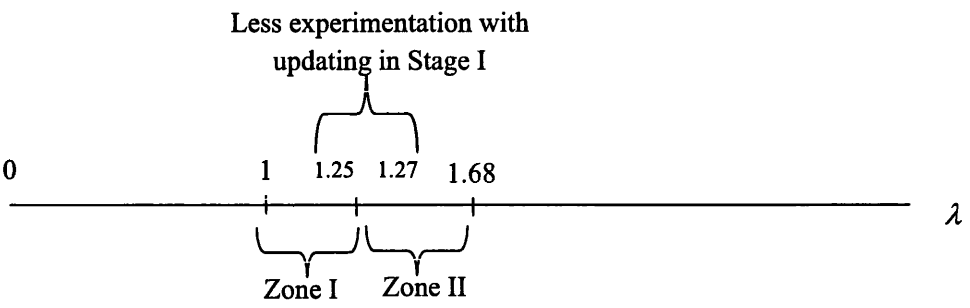


Figure 3a: *Loss-aversion parameter when the player updates his reference point and Optimal Experimentation in the 3-Stage Game*

3.3 Risk-Aversion and Loss-Aversion Together

Agents may be characterized by the simultaneous presence of risk aversion and loss aversion. We model a player's preference in order to tie together the presence of these two attitudes in the following form:

$$U(x) = x - rx^2 + D\left(x - \bar{x}\right)$$

where $D\left(x-\bar{x}\right)$ is the disappointment-elation function, which takes the following form:

$$D(x) = \begin{cases} x - \bar{x} \; for \, x \geq \bar{x} \\ \lambda\left(x - \bar{x}\right) for \, x < \bar{x}, \lambda > 1 \end{cases}$$

We proceed by calculating the optimal experimentation strategy of an agent with loss aversion measured by λ and risk aversion measured by r following the same steps described in the previous section.

In the first stage of the game, the individual will choose the strategy as a pair of $\{k_1, k_2\}$ in order to maximize the index attached to the risky arm (representing the expected payoff from experimenting). We denote with $Q^{Ra,La}$ the dynamic allocation index of a risk and loss averse player. Analytically, the experimenter will choose:

In the last stage of the game, following history $\{h_1, h_2\} = \{1,1\}$ or $\{h_1, h_2\} = \{0,1\}$, the agent will keep playing on the risky arm. In the case of a history $\{h_1, h_2\} = \{0,0\}$, the agent will play the risky arm if:

$$q \leq Q^{Ra,La}\left(n=1, \phi, h_1 = h_2 = 0\right)$$

where the dynamic allocation index simplifies to:

$$Q^{Ra,La}\left(n=1, \phi, h_1 = h_2 = 0\right) = \psi_{3,\{0,0\}}\left(2 - r - \psi_{3,\{0,0\}}\right) - \left(1 - \psi_{3,\{0,0\}}\right)\left(\lambda \psi_{3,\{0,0\}}\right)$$

The formulation discussed in this section allows us to determine, for any payoff value q of the safe arm, the indifference curve along which a loss averse player (with zero risk aversion) would undertake the same experimentation of a risk averse player (with zero loss aversion). The iso-curve is determined by imposing the equivalence between $Q^{La,Ra}\big|_{\lambda=0} = Q^{La,Ra}\big|_{r=0}$. From inspection of $Q^{Ra,La}(n,\phi,h)$ we immediately determine a negative linear relationship between the loss aversion and risk aversion parameter[27] in order to produce the same optimal experimentation strategy for the agent.

3.4 The testable predictions of the theory

We will use the set-up discussed in this Section as the basis for the design of the experimental analysis developed in Chapter 3. The analysis of the three-stage game allows us to state testable predictions, that we will summarize as follows.[28]

Proposition 1: Monotonicity

1a: Monotonicity of optimal experimentation in risk aversion r

The cutoff of optimal experimentation is monotonically decreasing in r

1b. Monotonicity of optimal experimentation in loss aversion λ

The cutoff of optimal experimentation is monotonically decreasing in λ

---

[27] Details are in the appendix.

[28] Proofs are in the Appendix to Chapter 2

$$Q^{Ra,La}\left(3,\phi,h=\{\varnothing\}\right)=\max_{\{k_1,k_2\}}\left\{\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}\right\}$$

The individual will decide to experiment if and only if:

$$q\le Q^{Ra,La}\left(3,\phi,h=\{\varnothing\}\right)$$

In the second stage of the game ($N=2$), the agent will keep playing on the risky arm following a history $\{h_1\}=\{1\}$. In case of a history $\{h_1\}=\{0\}$, the player will experiment on the risky arm at $n=2$, if:

$$q\le Q^{Ra,La}\left(n=2,\phi,h_1=0\right)$$

The dynamic allocation index under the strategy $k_1=0$ is calculated as follows. The combined expected payoff in the two remaining periods of experimentation is equal to $\psi_{2,\{0\}}\begin{bmatrix}\left(1+\psi_{3,\{0,1\}}\right)\left(2-r-\psi_{2,\{0\}}\right)-\\-\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{2,\{0\}}\end{bmatrix}+\left(1-\psi_{2,\{0\}}\right)\begin{bmatrix}-\left(2-\psi_{3,\{0,1\}}\right)\lambda\psi_{2,\{0\}}+\\+\psi_{3,\{0,1\}}\left(2-r-\psi_{2,\{0\}}\right)\end{bmatrix}$ since he will play the risky arm in both periods. This payoff should be compared with the alternative payoff, which is to play the safe arm in both periods, which yields $2q$.

The dynamic allocation index under the strategy $k_1=1$ is calculated as follows. The combined expected payoff in the remaining periods ($N=2$ and $N=3$) is equal to $\psi_{2,\{0\}}\left[\left(1+\psi_{3,\{0,1\}}\right)\left(2-r-\psi_{2,\{0\}}\right)-\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{2,\{0\}}\right]+\left(1-\psi_{2,\{0\}}\right)\left(-\lambda\psi_{2,\{0\}}+q\right)$, since he will play the safe arm in the case of a failure in period $n=2$. This payoff should be compared with the alternative payoff, which is to play the safe arm in both periods, which yields $2q$.

Hence the agent will not experiment at $n=2$, if:

$$q\le Q^{Ra,La}\left(n=2,\phi,h_1=0\right)$$

where the dynamic allocation index takes the following form:

$$Q^{Ra,La}(2,\phi,h_1=0)=\max_{k_1=\{0,1\}}\left\{\tilde{q}\Big|_{k_1=0};\tilde{q}\Big|_{k_1=1}\right\}$$

and

$$\tilde{q}\Big|_{k_1=0}=\frac{1}{2}\begin{bmatrix}\psi_{2,\{0\}}\left[\left(1+\psi_{3,\{0,1\}}\right)\left(2-r-\psi_{2,\{0\}}\right)-\left(1-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)\right]+\\+\left(1-\psi_{2,\{0\}}\right)\left[-\left(2-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)+\psi_{3,\{0,1\}}\left(2-r-\psi_{2,\{0\}}\right)\right]\end{bmatrix}$$

$$\tilde{q}\Big|_{k_1=1}=\frac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}\left[\left(2-r-\psi_{2,\{0\}}\right)\left(1+\psi_{3,\{0,1\}}\right)-\lambda\left(1-\psi_{2,\{0\}}\psi_{3,\{0,1\}}\right)\right]$$

In the last stage of the game, following history $\{h_1, h_2\} = \{1,1\}$ or $\{h_1, h_2\} = \{0,1\}$, the agent will keep playing on the risky arm. In the case of a history $\{h_1, h_2\} = \{0,0\}$, the agent will play the risky arm if:

$$q \leq Q^{Ra,La}(n=1, \phi, h_1 = h_2 = 0)$$

where the dynamic allocation index simplifies to:

$$Q^{Ra,La}(n=1, \phi, h_1 = h_2 = 0) = \psi_{3,\{0,0\}}(2 - r - \psi_{3,\{0,0\}}) - (1 - \psi_{3,\{0,0\}})(\lambda \psi_{3,\{0,0\}})$$

The formulation discussed in this section allows us to determine, for any payoff value q of the safe arm, the indifference curve along which a loss averse player (with zero risk aversion) would undertake the same experimentation of a risk averse player (with zero loss aversion). The iso-curve is determined by imposing the equivalence between $Q^{La,Ra}\big|_{\lambda=0} = Q^{La,Ra}\big|_{r=0}$. From inspection of $Q^{Ra,La}(n, \phi, h)$ we immediately determine a negative linear relationship between the loss aversion and risk aversion parameter[27] in order to produce the same optimal experimentation strategy for the agent.

3.4 The testable predictions of the theory

We will use the set-up discussed in this Section as the basis for the design of the experimental analysis developed in Chapter 3. The analysis of the three-stage game allows us to state testable predictions, that we will summarize as follows.[28]

Proposition 1: Monotonicity

1a: Monotonicity of optimal experimentation in risk aversion r

The cutoff of optimal experimentation is monotonically decreasing in r

1b. Monotonicity of optimal experimentation in loss aversion λ

The cutoff of optimal experimentation is monotonically decreasing in λ

---

[27] Details are in the appendix.

[28] Proofs are in the Appendix to Chapter 2

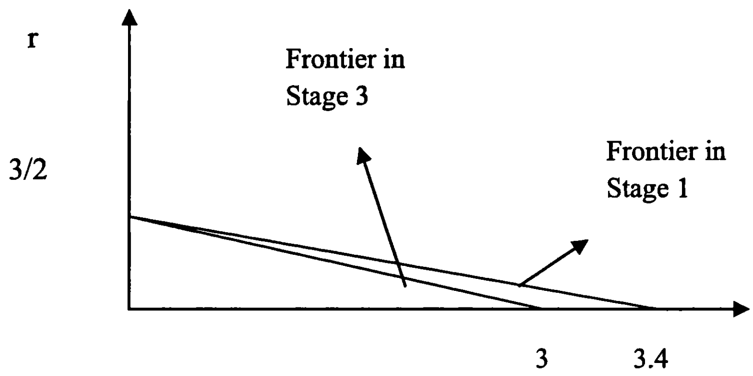
Proposition 2: Symmetry

Experimentation cutoffs in stage 3 are symmetrical following history $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$ for any pair λ, r. The optimal experimentation strategies in stage 3 will coincide following history $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$ for any pair λ, r

Proposition 3: Final stage of experimentation

Any pair λ, r that leads to experimentation in stage 1 will lead to experimentation in stage 3 following history $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$.

Proposition 3 states that subjects who choose to experiment in Stage 1 should find optimal to experiment in period 3 in the presence of the same posterior following history $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$. We illustrate the statement of Proposition 3 in the following graph. The cutoff for optimal experimentation in Stage 1 will always lie above the optimal cutoff of experimentation in Stage 3.



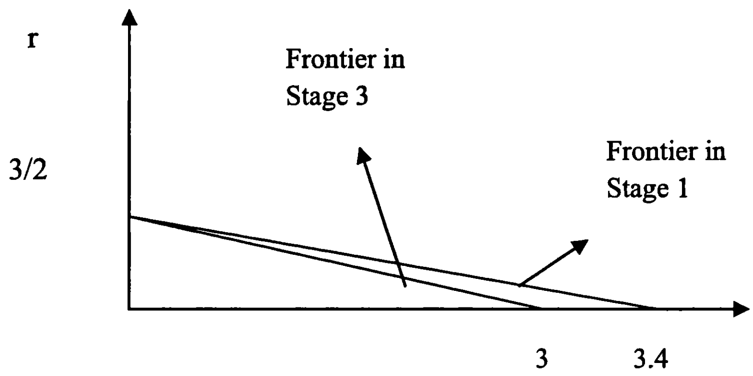
Proposition 2: Symmetry

Experimentation cutoffs in stage 3 are symmetrical following history $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$ for any pair λ, r. The optimal experimentation strategies in stage 3 will coincide following history $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$ for any pair λ, r

Proposition 3: Final stage of experimentation

Any pair λ, r that leads to experimentation in stage 1 will lead to experimentation in stage 3 following history $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$.

Proposition 3 states that subjects who choose to experiment in Stage 1 should find optimal to experiment in period 3 in the presence of the same posterior following history $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$. We illustrate the statement of Proposition 3 in the following graph. The cutoff for optimal experimentation in Stage 1 will always lie above the optimal cutoff of experimentation in Stage 3.



4. Towards a Specific Model

So far, we have considered the effect of risk aversion and loss aversion in a fairly general setting. In this section and in the next, we apply the method of analysis set out above to two specific models. This section considers a preliminary model involving loss aversion (but not risk aversion) that serves to illustrate some ideas that will be of interest in what follows. In the next section (Section 5) we set out the model involving both loss aversion and risk aversion that will be used in the experimental study in Chapter 2 below.

Consider an experimenter who is facing the following decision problem. The agent has to decide which project (chosen from a fixed set) to undertake in each period. Here we assume only two projects are available to the agent. Each project pays out a fixed reward at an unknown probability. In the case in which he knows the true probability of the reward of each project i, Π_i he would simply choose the project with the highest expected reward. The agent does not know Π_i but he can learn the probability of the reward through experimentation.

The generality of the description lends itself to a number of possible interpretations: we can think of the agent as a researcher employed in an R&D department, which has been assigned the task of finding a more efficient way to produce some commodity. He may have to choose among two substitute technologies, each offering uncertain benefits, until the development work is completed. Or consider the research work of a PhD student, who has to select the topic of his thesis from a pool of possible interesting areas, each with an uncertain probability of yielding a publishable paper. We can interpret the reward yielded by a project as a preliminary positive result or as a significative t-test yielded by regressions with the new data at the basis of his research. On the other hand, the project can hold no result, in the sense that it creates no significative advances toward the goal of the research. The framework also fits the case of a monopolistic store trying to price its commodity and learning the demand function through experimentation. This was studied by Rotschild (1974), where a positive result indicates that a customer buys the commodity, associated with a return equal to the price of the good minus the cost of production. Symmetrically, the negative result represents the customer leaving the store without buying the good[29].

[29] Note that since we endow the agent with a utility function, not all economists would be satisfied with this interpretation.

losses generally decreases with their magnitude[33]. This "loss aversion' is measured by the parameter λ , which captures the intensity of loss aversion in case of disappointment: The choice of this utility function allows us to investigate the effect of a psychological attitude, such as loss aversion, on the experimentation strategy optimally chosen by the experimenter.

4.1 The Analytical Set-Up

We consider two projects. The generic project i pays off 1 with probability Π_i and 0 with probability $(1-\Pi_i)$. The experimenter does not know the true probability Π_i. He decides which project to undertake according to his prior beliefs about the parameter Π_i and selects the one with the highest probability of a reward. We assume the agent has an infinite horizon.

Let N_i be the number of trials along project i and s_i be the number of positive results along i with $i = 1,2$.

The information in the sample can be represented by using the statistics defined as follows:

$$\rho_i = \frac{1}{1+N_i}$$

$$\mu_i = \frac{s_i}{1+N_i}$$ [34]

When the agent chooses project i, ρ_i becomes $\frac{\rho_i}{1+\rho_i}$. In case of a positive result on i, the statistics μ_i becomes:

$$s(\mu_i) = \frac{\mu_i + \rho_i}{1+\rho_i}$$

while in the case of a negative result, μ_i becomes:

$$f(\mu_i) = \frac{\mu_i}{1+\rho_i}$$

---

[33] The parameter λ is defined in an analogous way to the coefficient of loss-aversion in Kahneman and Tversky (1979, 1992). Empirical studies conducted by Kahneman and Tversky indicate that the best estimate of the coefficient of loss-aversion is 2,25.

[34] As the number of trials increases, μ_i approaches the sample mean $\overline{\mu_i} = \frac{s_i}{N_i}$ i.e. $\lim \mu = \lim \bar{\mu}$

In mathematical and statistical literature this problem is known as the two armed bandit problem and it has been widely analyzed[30]. Here, we assume that the agent's preferences are described by the disappointment-elation utility function (Sugden and Loomes, 1986):

$$U(x_{is}) = x_{is} + D(x_{is} - \bar{x}_i)$$

where x_{is} denotes the utility of project i under state of the world s and $\bar{x}_i$ denotes the expected utility of project i.

According to the utility function, the economic agent receives not only the utility derived directly from the actual consequence of an uncertain prospect, but in addition he feels some degree of disappointment and elation. When the agent evaluates a prospect, he forms an a priori expectation about any uncertain prospect and after uncertainty is resolved he compares the actual consequence of the prospect with the a priori expectation. If the actual consequence turns out to be worse than the expectation, he feels disappointment. On the other hand, the individual experiences some degree of elation if the actual consequence is better than the a priori expectation.

We assume x_{is} to be linear and exactly equal to the payoff in state s of project i and the disappointment-elation utility function $D(\bullet)$ to take the following functional form:

$$D(\xi) = \begin{cases} \xi \; if \; \xi \geq 0 \\ \lambda\xi \; if \; \xi < 0, \lambda > 1 \end{cases}$$

The utility function is defined on deviations from the reference point[31]. Additionally, we require the utility function to be steeper for losses than for gains, according to Kahneman and Tversky (1979), to reflect a salient feature of attitudes to changes in welfare: the disappointment experienced in losing a sum of money appears to be greater than the pleasure associated with gaining the same amount. This specification of the $D(\bullet)$ function captures the spirit of loss aversion in the Kahneman-Tversky descriptive theory of judgment under uncertainty (Kahneman and Tversky, 1979)[32]. The economic meaning of this assumption is that the marginal value of both gains and

[30] See Gittins (1979), Berry and Fristedt (1985)

[31] The reference point is represented by the origin, since we do not endow the agent with any legacy or other forms of wealth inherited from the past, i.e. the utility function is kinked at the origin.

[32] Here, for simplicity, we do not introduce Kahneman and Tversky's assumption according to which the utility function for changes of wealth is normally concave above the reference point ($D^{''}(x) < 0$ for $x > 0$) and often convex below it ($D^{''}(x) > 0$ for $x < 0$).

losses generally decreases with their magnitude[33]. This "loss aversion' is measured by the parameter λ , which captures the intensity of loss aversion in case of disappointment: The choice of this utility function allows us to investigate the effect of a psychological attitude, such as loss aversion, on the experimentation strategy optimally chosen by the experimenter.

4.1 The Analytical Set-Up

We consider two projects. The generic project i pays off 1 with probability Π_i and 0 with probability $(1-\Pi_i)$. The experimenter does not know the true probability Π_i. He decides which project to undertake according to his prior beliefs about the parameter Π_i and selects the one with the highest probability of a reward. We assume the agent has an infinite horizon.

Let N_i be the number of trials along project i and s_i be the number of positive results along i with $i=1,2$.

The information in the sample can be represented by using the statistics defined as follows:

$$\rho_i = \frac{1}{1+N_i}$$

$$\mu_i = \frac{s_i}{1+N_i}$$ [34]

When the agent chooses project i, ρ_i becomes $\frac{\rho_i}{1+\rho_i}$. In case of a positive result on i, the statistics μ_i becomes:

$$s(\mu_i) = \frac{\mu_i + \rho_i}{1+\rho_i}$$

while in the case of a negative result, μ_i becomes:

$$f(\mu_i) = \frac{\mu_i}{1+\rho_i}$$

---

[33] The parameter λ is defined in an analogous way to the coefficient of loss-aversion in Kahneman and Tversky (1979, 1992). Empirical studies conducted by Kahneman and Tversky indicate that the best estimate of the coefficient of loss-aversion is 2,25.

[34] As the number of trials increases, μ_i approaches the sample mean $\overline{\mu_i} = \frac{s_i}{N_i}$ i.e. $\lim \mu = \lim \bar{\mu}$

Therefore, the information contained in the sample is given by $(\mu_1, \mu_2, \rho_1, \rho_2)$, which is a subset of R^4, constituted by a fourfold copy of the closed unit interval $[0,1]$. Denote the domain of (μ, ρ) as Δ.

The experimenter's belief about the parameters Π_i with $i = 1,2$ are given by the prior density function $g(\pi_1, \pi_2)$. We do not assume the independence of the probabilities of success along the two projects, but we assume:

$$g(\pi_1, \pi_2) > 0 \quad for all \quad (\pi_1, \pi_2) \in (0,1) \times (0,1)$$

excluding all non extreme combinations of Π_1 and Π_2.

The experimenter with experience (μ, ρ) updates his prior beliefs from $g(\pi_1, \pi_2)$ to $h(\pi_1, \pi_2, \mu, \rho)$. The probability density is proportional to

$$\pi_1^{\frac{\mu_1}{\rho_1}} (1-\pi_1)^{\frac{[1-(\mu_1+\rho_1)]}{\rho_1}} \pi_2^{\frac{\mu_2}{\rho_2}} (1-\pi_2)^{\frac{[1-(\mu_2+\rho_2)]}{\rho_2}} g(\pi_1, \pi_2)$$

The posterior mean of the experimenter's belief about the parameter Π_i given the sample information (μ, ρ) and the prior density function g is defined as follows[35]:

$$\lambda_i(\mu, \rho) = \int_0^1 \int_0^1 \pi_i h(\pi_1, \pi_2, \mu, \rho) d\pi_1 d\pi_2$$

4.2 Dynamic Programming Equations and Properties

The experimenter maximizes the expected discounted utility of his rewards over the infinite horizon. The problem can be written in terms of dynamic programming equations, satisfying the following functional form[36]:

$$V(\mu, \rho) = \max_{\{i\}} W_i(\mu, \rho)$$

Where

$$W_i(\mu, \rho) = \lambda_i(\mu, \rho)\left[1 + D(1 - \lambda_i(\mu, \rho))\right] + (1 - \lambda_i(\mu, \rho)) D(-\lambda_i(\mu, \rho)) +$$
$$+ \delta\left[\lambda_i(\mu, \rho) V(\sigma_i(\mu, \rho)) + (1 - \lambda_i(\mu, \rho)) V(\psi_i(\mu, \rho))\right]$$

[35] $\lambda_i(\mu, \rho)$ is defined and continuous $\forall(\mu, \rho)$ such that $\rho_i > 0$, $i = 1,2$. It is possible to show that $\lambda_i(\mu, \rho)$ can be extended by continuity to $[0,1]$, since $\lim_{\rho_i \to 0} \lambda_i(\mu, \rho) = \mu_i$ by the law of large numbers.

[36] From now on we use the general specification of the disappointment-elation function $D(\bullet)$ and we consider the specific functional form assumed only when necessary to derive the result.

Therefore, the information contained in the sample is given by $(\mu_1,\mu_2,\rho_1,\rho_2)$, which is a subset of R^4, constituted by a fourfold copy of the closed unit interval $[0,1]$. Denote the domain of (μ,ρ) as Δ.

The experimenter's belief about the parameters Π_i with $i=1,2$ are given by the prior density function $g(\pi_1,\pi_2)$. We do not assume the independence of the probabilities of success along the two projects, but we assume:

$$g(\pi_1,\pi_2)>0 \quad for all \quad (\pi_1,\pi_2)\in(0,1)\times(0,1)$$

excluding all non extreme combinations of Π_1 and Π_2.

The experimenter with experience (μ,ρ) updates his prior beliefs from $g(\pi_1,\pi_2)$ to $h(\pi_1,\pi_2,\mu,\rho)$. The probability density is proportional to

$$\pi_1^{\frac{\mu_1}{\rho_1}}(1-\pi_1)^{\frac{[1-(\mu_1+\rho_1)]}{\rho_1}}\pi_2^{\frac{\mu_2}{\rho_2}}(1-\pi_2)^{\frac{[1-(\mu_2+\rho_2)]}{\rho_2}}g(\pi_1,\pi_2)$$

The posterior mean of the experimenter's belief about the parameter Π_i given the sample information (μ,ρ) and the prior density function g is defined as follows[35]:

$$\lambda_i(\mu,\rho)=\int_0^1\int_0^1\pi_i h(\pi_1,\pi_2,\mu,\rho)\,d\pi_1 d\pi_2$$

4.2 Dynamic Programming Equations and Properties

The experimenter maximizes the expected discounted utility of his rewards over the infinite horizon. The problem can be written in terms of dynamic programming equations, satisfying the following functional form[36]:

$$V(\mu,\rho)=\max_{\{i\}}W_i(\mu,\rho)$$

Where

$$\begin{aligned}W_i(\mu,\rho)&=\lambda_i(\mu,\rho)\left[1+D(1-\lambda_i(\mu,\rho))\right]+(1-\lambda_i(\mu,\rho))D(-\lambda_i(\mu,\rho))+\\&+\delta\left[\lambda_i(\mu,\rho)V(\sigma_i(\mu,\rho))+(1-\lambda_i(\mu,\rho))V(\psi_i(\mu,\rho))\right]\end{aligned}$$

---

[35] $\lambda_i(\mu,\rho)$ is defined and continuous $\forall(\mu,\rho)$ such that $\rho_i>0$, $i=1,2$. It is possible to show that $\lambda_i(\mu,\rho)$ can be extended by continuity to $[0,1]$, since $\lim_{\rho_i\to 0}\lambda_i(\mu,\rho)=\mu_i$ by the law of large numbers.

[36] From now on we use the general specification of the disappointment-elation function $D(\bullet)$ and we consider the specific functional form assumed only when necessary to derive the result.

where $0<\delta<1$ and $\sigma_i(\mu,\rho)$ and $\psi_i(\mu,\rho)$ are vectors indicating the state of information after, respectively, a positive or a negative result on i.

Define the functions $V^t(\mu,\rho)$ and $W_i^t(\mu,\rho)$ as follows:

$$V^0(\mu,\rho)=0$$

$$V^t(\mu,\rho)=\max_{\{i\}} W_i^t(\mu,\rho)$$

where

$$\begin{aligned}W_i^t(\mu,\rho)=&\lambda_i(\mu,\rho)\left[1+D(1-\lambda_i(\mu,\rho))\right]+(1-\lambda_i(\mu,\rho))D(-\lambda_i(\mu,\rho))+\\&+\delta\left[\lambda_i(\mu,\rho)V^{t-1}(\sigma_i(\mu,\rho))+(1-\lambda_i(\mu,\rho))V^{t-1}(\psi_i(\mu,\rho))\right]\end{aligned}$$

Lemma 1

The functions $V^t(\mu,\rho)$ and $W_i^t(\mu,\rho)$ are continuous.

Proof of Lemma 1

The proof is given by induction. We know that:

$$V^0(\mu,\rho)=0$$

$$V^1(\mu,\rho)=\max_{\{i\}} W_i^1(\mu,\rho)$$

where

$$W_i^1(\mu,\rho)=\lambda_i(\mu,\rho)\left[1+D(1-\lambda_i(\mu,\rho))\right]+(1-\lambda_i(\mu,\rho))D(-\lambda_i(\mu,\rho))$$

since $V^0(\mu,\rho)=0$.

Given the continuity of $\lambda_i(\mu,\rho)$ and of the disappointment utility function $D(\cdot)$, $W_i^1(\mu,\rho)$ is continuous. Let us suppose that $V_i^{t-1}(\mu,\rho)$ and $W_i^{t-1}(\mu,\rho)$ are continuous. By the definition of $W_i^t(\mu,\rho)$, we have:

$$\begin{aligned}W_i^t(\mu,\rho)=&\lambda_i(\mu,\rho)\left[1+D(1-\lambda_i(\mu,\rho))\right]+(1-\lambda_i(\mu,\rho))D(-\lambda_i(\mu,\rho))+\\&+\delta\left[\lambda_i(\mu,\rho)V^{t-1}(\sigma_i(\mu,\rho))+(1-\lambda_i(\mu,\rho))V^{t-1}(\psi_i(\mu,\rho))\right]\end{aligned}$$

which is continuous since it is the sum of continuous functions. The same argument applies to $V_i^t(\mu,\rho)$, since:

$$V^1(\mu,\rho)=\max_{\{i\}} W_i^1(\mu,\rho)$$

Lemma 2

The functions $V^t(\mu,\rho)$ and $W_i^t(\mu,\rho)$ are monotonic.

Proof of Lemma 2

The proof is given by induction.

First, we show that $V^1 \geq V^0$ and $W_i^1 \geq W_i^0 = 0$. We know that $V^0(\mu,\rho)=0$. Therefore, we need to show that $V^1 \geq 0$. By simply applying the formula,

$$W_i^1(\mu,\rho)=\lambda_i(\mu,\rho)\left[1+D(1-\lambda_i)\right]+\left(1-\lambda_i(\mu,\rho)\right)D\left(-\lambda_i(\mu,\rho)\right)\geq 0$$

that is,

$$W_i^1(\mu,\rho)=\lambda_i(\mu,\rho)\left[2-\lambda_i(\mu,\rho)\right]-\left(1-\lambda_i(\mu,\rho)\right)b\lambda_i(\mu,\rho)\geq 0$$

if and only if $b \leq \dfrac{2-\lambda_i(\mu,\rho)}{1-\lambda_i(\mu,\rho)}$ since $\lambda_i(\mu,\rho)\in[0,1]$.

Let us assume that $V_i^{t-1}(\mu,\rho)$ is monotonically increasing and show the monotonicity of $V_i^t(\mu,\rho)$

By the definition of $W_i^t(\mu,\rho)$ we have:

$$\begin{aligned}W_i^t(\mu,\rho)&=\lambda_i(\mu,\rho)\left[1+D(1-\lambda_i)\right]+\left(1-\lambda_i(\mu,\rho)\right)D\left(-\lambda_i(\mu,\rho)\right)+\\&+\delta\left[\lambda_i(\mu,\rho)V^{t-1}\left(s_i(\mu,\rho)\right)+\left(1-\lambda_i(\mu,\rho)\right)V^{t-1}\left(f_i(\mu,\rho)\right)\right]\end{aligned}$$

which is monotonically increasing, under the condition stated on the parameter of loss aversion λ, since it is the sum of monotonically increasing functions. The same argument applies to $V_i^t(\mu,\rho)$, since $V^t(\mu,\rho)=\max_{\{i\}} W_i^t(\mu,\rho)$

Lemma 3

$V^t(\mu,\rho)$ and $W_i^t(\mu,\rho)$ converge uniformly to $V(\mu,\rho)$ and $W_i(\mu,\rho)$ respectively

Proof of Lemma 3

In order to show the uniform convergence of $V^t(\mu,\rho)$ and $W_i^t(\mu,\rho)$, we need to show that there exists a majorant M, greater than the maximum of $\lambda_i(\mu,\rho)\left[1+D(1-\lambda_i)\right]+\left(1-\lambda_i(\mu,\rho)\right)D\left(-\lambda_i(\mu,\rho)\right)$. Note that M exists, since $\lambda_i(\mu,\rho)\in[0,1]$,

$D(1-\lambda_i(\mu,\rho)) \in (0,1)$, $D(-\lambda_i(\mu,\rho)) \in R^- - \{-\infty\}$ [37] and $\lim_{\rho_i \to 0} \lambda_i(\mu,\rho) = \mu_i$. Then, it has to be the case that $V_i^t(\mu,\rho) \leq M\sum_{\tau=0}^{t}\delta^\tau \leq M\sum_{\tau=0}^{\infty}\delta^\tau = \frac{M}{1-\delta}$.

The sequences $V^t(\mu,\rho)$ and $W_i^t(\mu,\rho)$ are bounded above and converge respectively to $V(\mu,\rho)$ and $W_i(\mu,\rho)$, which are defined as follows:

$$V(\mu,\rho) = \lim_{t\to\infty} V^t(\mu,\rho)$$

$$W(\mu,\rho) = \lim_{t\to\infty} W_i^t(\mu,\rho)$$

Indicate with $\overline{V}^t(\mu,\rho)$ the present discounted value of the expected utility of the sum of rewards from the first t periods when following a policy described by the system of dynamic programming equations.

Then, $V^t(\mu,\rho) \geq \overline{V}^t(\mu,\rho)$. Then, $V(\mu,\rho) \leq \overline{V}^t(\mu,\rho) + \delta^t \sum_{\tau=t}^{\infty}\delta^\tau \leq V^t(\mu,\rho) + \delta^t \frac{M}{1-\delta}$, which implies $\left|V(\mu,\rho) - V^t(\mu,\rho)\right| \leq \delta^t \frac{M}{1-\delta}$. Since the previous inequality is independent of (μ,ρ), $V^t(\mu,\rho)$ converges uniformly to $V(\mu,\rho)$. A similar argument for uniform convergence applies to $W_i^t(\mu,\rho)$.

Proposition 1

$V(\mu,\rho)$ and $W_i(\mu,\rho)$ are continuous on Δ.

Proof of Proposition 1

This follows from Lemma 1 and 3.

Define $A_i = \left\{(\mu,\rho) \in [0,1]^4 : W_i(\mu,\rho) > W_j(\mu,\rho)\right\}$ the information set such that project i is optimally chosen.

Lemma 4

A_i is an open set.

---

[37] This is true under the condition on the parameter λ stated in Lemma 2

Proof of Lemma 4

This follows from Lemma 1.

4.3 Main results

Theorem 1

If the true parameters satisfy $0 < \Pi_1 < \Pi_2 < 1$, and the loss aversion parameter satisfies $\lambda \leq \frac{2-\lambda_i(\mu,\rho)}{1-\lambda_i(\mu,\rho)}$, then an agent who follows an optimal strategy will with positive probability choose project 1 infinitely often and project 2 only a finite number of times.

The proof of the theorem is given in an analogous way to the one stated in Rothschild (1974). The following Lemma 5 shows that while the agent is experimenting, he will find it optimal to undertake project 1 if he has a very bad run of trials along project 2, even if project 1 is the inferior of the two. The main difference in the proof lies in the additional condition on the loss aversion parameter λ characterizing the agent. This condition tells us that the agent will implement the same optimal experimentation strategy of a standard expected utility maximizing agent under the condition that the intensity of loss aversion he feels is not "too strong". Note that the threshold value under which the agent is going to follow this experimentation strategy is endogenously determined by his beliefs on the likelihood of the generic project i to pay out the reward, $\lambda_i(\mu,\rho)$, updated in each period on the basis of the state of information (μ,ρ). This surprising finding relies in the fact that, for every value of $\lambda_i(\mu,\rho)$, the condition satisfies the asymmetry on the weights given to losses and gains, i.e. for every value of $\lambda_i(\mu,\rho)$, the agent is allowed to consistently weigh more losses relative to gains but will nonetheless follow the same experimentation strategy as a standard EUT maximizer.

Lemma 5

Under the condition $\lambda \leq \frac{2-\lambda_i(\mu,\rho)}{1-\lambda_i(\mu,\rho)}$ and for every δ_1, δ_2 such that $\Pi_1 > \delta_1 > 0$ and $0 < \delta_2 < 1$, there exists $\varepsilon > 0$ such that $W_1(\mu,\rho) > W_2(\mu,\rho)$ whenever $\mu_2 + \rho_2 < \varepsilon$ and either $\mu_1 \geq \Pi_1 - \delta_1$ or $\rho_1 \geq \delta_2$.

Proof of Lemma 5

Consider the compact set:

$$K=\{(\mu_1,\mu_2,\delta_1,\delta_2)\in\Delta|\mu_1\geq\Pi_1-\delta_1>0\ or\ \rho_1\geq\delta_2\ and\ \mu_2=\rho_2=0\}$$

We want to show the conditions under which $K\subset A_1$. Consider:

$$W_1(\mu,0,\rho,0)\geq\lambda_1(\mu,0,\rho,0)\left[1+D(1-\lambda_1(\mu,0,\rho,0))\right]+(1-\lambda_1(\mu,0,\rho,0))D(-\lambda_1(\mu,0,\rho,0))>0$$

From Lemma 2, the second inequality holds if and only if:

$$\lambda\leq\frac{2-\lambda_i(\mu,\rho)}{1-\lambda_i(\mu,\rho)}$$

while if

$$W_2(\mu,\rho)\geq W_1(\mu,\rho)$$

then

$$V(\mu,0,\rho,0)=W_2(\mu,0,\rho,0)=0\left[1+D(1)\right]+1\left[0+D(0)\right]+\delta\left[0*V(\mu,0,\rho,0)+1*V(\mu,0,\rho,0)\right]$$

$$\Rightarrow V(\mu,0,\rho,0)=\delta V(\mu,0,\rho,0)$$

This equality holds only if $V(\mu,0,\rho,0)=0<W_1(\mu,0,\rho,0)$, a contradiction. From Lemma 4, A_1 is an open set, centered at any (μ,ρ). Therefore, K is an open ball contained in A_1. Let us indicate with $B_r(x)$ the generic ball centered at point x with radius r. Then, there is a finite number J of balls covering the set A_1, such that $K\subset\bigcup_{j\in J}B_{r_j}(x_j)$.

Let us denote a minimum radius r such that $\varepsilon=\min_{\{j\in J\}}r_j$. It has to be the case that any point $(\mu_1,\mu_2,\rho_1,\rho_2)$ such that $\mu_2+\rho_2<\varepsilon$ and either $\mu_1\geq\Pi_1-\delta_1$ or $\rho_1\geq\delta_2$ belongs to a open ball $B_{r_j}(x_j)\in A_1$.

For the second lemma needed to establish the theorem, refer to Rothschild (1974).

As in Rothschild, the proof of Theorem 1 is obtained without using the condition $\Pi_1<\Pi_2$. This means that there exists a support in the probability distribution such that the experimenter will find it optimal to choose to undertake project 1 only a finite number of times and project 2 infinitely often. This means that inefficiency may arise in the long run, since nothing guarantees even in the long term that the experimenter will end up choosing the arm more likely to pay out the reward.

the reservation value that the agent assigns to each project, and in turns to make explicit quantitative comparisons in terms of how much experimentation the agent will undertake when characterized by a disappointment-elation utility function, as defined in Section 3.3. We allow for the presence of risk aversion and loss aversion.

Within this framework, we characterize the condition on the parameter of loss aversion λ under which the agent, characterized by the disappointment-elation utility model of Loomes and Sugden, will choose to experiment more intensively than the player characterized by the standard expected utility model.

5.1 The Analytical Set-Up

Consider an agent who faces the decision between a project that pays off a reward with an uncertain probability and a "safe" project. Assume the first project X_1 has a binomial distribution with $p = \Pr(X_1 = 1)$ unknown but selected by a known priori distribution, F, while the second X_2 has a binomial distribution with a known parameter, q. In other terms, the experimenter knows that the second project pays off q with certainty, while with project 1 he gets a reward equal to 1 with probability $p = \Pr(X_1 = 1)$, which is drawn from a known (general) distribution $F(p)$. In statistical terms, the problem is reduced to a one armed bandit problem.

The experimenter is allowed to play n times and his objective is to determine the sequential strategy which will maximize the expected utility of rewards of the unsafe project (out of n independent observations). We consider n-horizon uniform discounting. The problem is therefore reduced to an optimal stopping one, in which the experimenter has to decide in each period whether to go on with experimentation along the unknown project or to undertake the safe project.

The agent's preferences are described by the disappointment-elation utility function (Sugden and Loomes, 1986), as defined in Section 3.3. We allow for the presence of risk aversion and loss aversion.

We characterize the optimal experimentation strategy in terms of the dynamic allocation index or Gittins index, indicated as $Q(n,F)$, which is a function of n trials remaining and F, which is the a priori distribution of p at that time. According to the optimal strategy, the experimenter will choose the safe project if $q > Q(n,F)$[38] and he will undertake project 1, yielding an uncertain reward, in the opposite case. The optimal strategy followed by the experimenter takes the form of an optimal stopping rule.

[38] Note that along the safe project the agent feels neither elation nor disappointment.

This result holds even when we characterize the preferences of the agent in terms of the disappointment-elation utility function in the presence of loss aversion.

In the following proposition we establish the limiting conditions on the parameter of loss aversion, λ.

Proposition 2

As the number of trials on project i increases, the condition on the parameter λ becomes:

$$\lambda \leq \frac{2-\lambda_i(\mu,\rho)}{1-\lambda_i(\mu,\rho)} \text{ as } \rho_i \rightarrow 0$$

$$\lambda \leq 2 \text{ as } \mu_i \rightarrow 0 \text{ and } \rho_i \rightarrow 0$$

$$\lambda \leq \infty \text{ as } \mu_i \rightarrow 1 \text{ and } \rho_i \rightarrow 0$$

Proof of Proposition 2

The proof is straightforward.

From Proposition 2 there are two main observations. First, λ is allowed to assume values greater than 1, consistent with the Kahneman and Tversky psychological findings of a higher disutility associated with same size losses relative to gains. Second, and more surprisingly, even if the experience on project i is very poor, the condition on the loss aversion parameter λ never violates the requirement on the utility function.

5. The Model with Dynamic Allocation Index and a Finite Horizon

In this section the problem of the optimal experimentation of the agent is handled by using the analytical tool of the dynamic allocation indexes. Gittins (1979) showed that the optimal policy in the framework of multi-armed bandit problems can be described in terms of the so-called dynamic allocation index: if the arms are independent (that is pulling one arm is uninformative about other arms) then it is possible to attach to each arm an index, which depends only on the current state of information on that arm. According to the optimal strategy, the experimenter will find it optimal to choose the arm with the highest index. This index acts therefore as a reservation value. The tractability of the problem with dynamic allocation indexes allows us to determine a closed form for the reservation value that the agent assigns to each project, and in turns to make explicit quantitative comparisons in terms of how much experimentation the agent will undertake when characterized by a disappointment-elation utility function, as defined in Section 3.3. We allow for the presence of risk aversion and loss aversion.

Within this framework, we characterize the condition on the parameter of loss aversion λ under which the agent, characterized by the disappointment-elation utility model of Loomes and Sugden, will choose to experiment more intensively than the player characterized by the standard expected utility model.

5.1 The Analytical Set-Up

Consider an agent who faces the decision between a project that pays off a reward with an uncertain probability and a "safe" project. Assume the first project X_1 has a binomial distribution with $p = \Pr(X_1 = 1)$ unknown but selected by a known priori distribution, F, while the second X_2 has a binomial distribution with a known parameter, q. In other terms, the experimenter knows that the second project pays off q with certainty, while with project 1 he gets a reward equal to 1 with probability $p = \Pr(X_1 = 1)$, which is drawn from a known (general) distribution $F(p)$. In statistical terms, the problem is reduced to a one armed bandit problem.

The experimenter is allowed to play n times and his objective is to determine the sequential strategy which will maximize the expected utility of rewards of the unsafe project (out of n independent observations). We consider n-horizon uniform discounting. The problem is therefore reduced to an optimal stopping one, in which the experimenter has to decide in each period whether to go on with experimentation along the unknown project or to undertake the safe project.

The agent's preferences are described by the disappointment-elation utility function (Sugden and Loomes, 1986), as defined in Section 3.3. We allow for the presence of risk aversion and loss aversion.

We characterize the optimal experimentation strategy in terms of the dynamic allocation index or Gittins index, indicated as $Q(n,F)$, which is a function of n trials remaining and F, which is the a priori distribution of p at that time. According to the optimal strategy, the experimenter will choose the safe project if $q > Q(n,F)$[38] and he will undertake project 1, yielding an uncertain reward, in the opposite case. The optimal strategy followed by the experimenter takes the form of an optimal stopping rule.

[38] Note that along the safe project the agent feels neither elation nor disappointment.

According to Bradt, Johnson and Karlin (1956), within this analytical set-up, the optimal strategy has the following form for the appropriate k_i:

1. Observe the result on the unknown project X_1 until a failure occurs.
2. There exists an integer $k_1 \geq 0$ such that if at least k_1 positive results preceded the first negative result, continue with X_1; otherwise switch to the safe project X_2 for the remaining trials.
3. There is an integer $k_2 \geq 0$ attached to the second negative result such that if at least $k_1 + k_2$ positive results with X_1 precede the second negative one of X_1, continue with X_1; otherwise switch to X_2 for the remaining trials.
4. In general, let S_r be the number of positive results that precede the r-th negative one of X_1. If $S_r \geq k_1 + k_2 + ... + k_r$, continue with X_1; otherwise switch to X_2 for the remaining trials.

Thus, any sequence $k = (k_1, k_2, ..., k_n)$ of integers, $0 \leq k_i \leq n$, corresponds to a strategy of the same form as the optimal strategy.

Let E_k denote the expectation given k and let $W_n(F, q)$ denote the expected value of utility of the n observations against a priori distribution function F on project 1 and a given parameter q on project 2, pursuing an optimal strategy. In using any strategy for n trials, X_1 will be used a certain number of times, N_x , and there will be used a certain number, S_x of positive results with X_1; similarly for X_2.

In the following Proposition, a closed formula for the dynamic allocation index $Q(n,F)$ is determined. We use a specification of the disappointment-elation utility function, where $\bar{x}$ denotes the a priori expected reward yielded by project 1 calculated when n periods of experimentation are available to the agent and against a priori distribution function F. In what follows we consider an agent who chooses a fixed reference point $\bar{x}$ represented by the a priori expected reward of the risky project. In section 3.2 we discussed the role of the experimenter's decision to update the reference point according to the new beliefs formed after each trial. We have showed that the effect of this updating is to change the quantitative prediction, leaving unchanged the qualitative features of the solution. The assumption of "no updating" allows us to obtain a well-defined closed formula

According to Bradt, Johnson and Karlin (1956), within this analytical set-up, the optimal strategy has the following form for the appropriate k_i:

1. Observe the result on the unknown project X_1 until a failure occurs.
2. There exists an integer $k_1 \geq 0$ such that if at least k_1 positive results preceded the first negative result, continue with X_1; otherwise switch to the safe project X_2 for the remaining trials.
3. There is an integer $k_2 \geq 0$ attached to the second negative result such that if at least $k_1 + k_2$ positive results with X_1 precede the second negative one of X_1, continue with X_1; otherwise switch to X_2 for the remaining trials.
4. In general, let S_r be the number of positive results that precede the r-th negative one of X_1. If $S_r \geq k_1 + k_2 + \ldots + k_r$, continue with X_1; otherwise switch to X_2 for the remaining trials.

Thus, any sequence $k = (k_1, k_2, \ldots, k_n)$ of integers, $0 \leq k_i \leq n$, corresponds to a strategy of the same form as the optimal strategy.
Let E_k denote the expectation given k and let $W_n(F,q)$ denote the expected value of utility of the n observations against a priori distribution function F on project 1 and a given parameter q on project 2, pursuing an optimal strategy. In using any strategy for n trials, X_1 will be used a certain number of times, N_x , and there will be used a certain number, S_x of positive results with X_1; similarly for X_2.
In the following Proposition, a closed formula for the dynamic allocation index $Q(n,F)$ is determined. We use a specification of the disappointment-elation utility function, where $\bar{x}$ denotes the a priori expected reward yielded by project 1 calculated when n periods of experimentation are available to the agent and against a priori distribution function F. In what follows we consider an agent who chooses a fixed reference point $\bar{x}$ represented by the a priori expected reward of the risky project. In section 3.2 we discussed the role of the experimenter's decision to update the reference point according to the new beliefs formed after each trial. We have showed that the effect of this updating is to change the quantitative prediction, leaving unchanged the qualitative features of the solution. The assumption of "no updating" allows us to obtain a well-defined closed formula

5.2 Main Results

In this section we investigate the role of loss aversion in the experimentation choice of an agent. According to the literature on the armed bandit problem, the natural benchmark is provided by the framework in which the agent is characterized by the standard expected utility paradigm. In the standard optimizing literature on bandit problem, it is assumed that the agent is risk neutral and maximizes expected profits; we consider, therefore, the case where the utility function of the agent is linear and exactly equal to the payoff[39].

Proposition 4

A loss averse agent experiments more than a risk neutral agent when the following condition on the loss aversion parameter λ is satisfied:

$$\lambda \leq \min\left[\frac{E_k[S_x](1-r-\bar{x})}{\bar{c}[E_k[N_x]-E_k[S_x]]}, k+\frac{E_k[S_x](1-r-\bar{x})}{\bar{c}[E_k[N_x]-E_k[S_x]]}\right]$$

Proof

The agent will decide to experiment more when characterized by disappointment when the following inequality is satisfied:

$$Q^D(n,F) \geq Q^{EUT}(n,F)$$

where:

$Q^D(n,F)$ is calculated according to Proposition 3.

$$Q^{EUT}(n,F) = \max_k\left\{\frac{E_k[S_x]}{E_k[N_x]}\right\}^{40}$$

Under the specific assumption of a piecewise linear disappointment-elation utility function, this is equivalent to checking under which conditions the following inequality holds:

$$\max_k\left\{\frac{E_k[S_x]}{E_k[N_x]}(2-r-\bar{x})-\left(1-\frac{E_k[S_x]}{E_k[N_x]}\right)\lambda\bar{x}\right\} \geq \max_k\left\{\frac{E_k[S_x]}{E_k[N_x]}\right\}$$

[39] This choice is motivated for homogeneity with the standard economic literature on bandit problems.

[40] See Karlin, Bradt and Johnson (1956)

for $Q(n,F)$. Proposition 3 provides the explicit formula of the dynamic allocation index associated with the risky project.

Proposition 3

$$Q(n,F) = \max_k \left\{ \frac{E_k[S_x]}{E_k[N_x]}\left(1-r+D\left(1-\bar{x}\right)\right) + \left(1-\frac{E_k[S_x]}{E_k[N_x]}\right)D\left(-\bar{x}\right) \right\}$$

Proof

At the boundary, $q = Q(n,F)$. This implies that $nq = W_n(F,q)$. Since the optimal strategy is defined in terms of a sequence of k, it has to be the case that:

$$nq = \max_k \left\{ E_k\left[S_{x_1}\right]\left(1-r+D\left(1-\bar{x}\right)\right) + \left(E_k\left[N_{x_1}\right]-E_k\left[S_{x_1}\right]\right)D\left(-\bar{x}\right) + E_k\left[N_{x_2}\right]q \right\}$$

where $\bar{x} = \int_0^1 p dF$. Note that neither $E_k\left[S_{x_1}\right]$ nor $E_k\left[N_{x_1}\right]$ depend on q. Moreover, neither $E_k\left[S_{x_1}\right]$ nor $E_k\left[N_{x_2}\right]$ depend on the way the agent evaluates his payoffs. Note $E_k\left[S_{x_1}\right] = n - E_k\left[N_{x_2}\right]$. Hence, $q = Q(n,F)$ implies

$$nq = \max_k \left\{ E_k\left[S_{x_1}\right]\left(1-r+D\left(1-\bar{x}\right)\right) + \left(E_k\left[N_{x_1}\right]-E_k\left[S_{x_1}\right]\right)D\left(-\bar{x}\right) + E_k\left[N_{x_2}\right]q \right\}$$

$$\Leftrightarrow q \geq \frac{E_k[S_x]}{E_k[N_x]}\left(1-r+D\left(1-\bar{x}\right)\right) + \left(1-\frac{E_k[S_x]}{E_k[N_x]}\right)D\left(-\bar{x}\right) \text{ for all } k$$

For some k, it holds with equality:

$$Q(n,F) = \max_k \left\{ \frac{E_k[S_x]}{E_k[N_x]}\left(1-r+D\left(1-\bar{x}\right)\right) + \left(1-\frac{E_k[S_x]}{E_k[N_x]}\right)D\left(-\bar{x}\right) \right\}$$

According to Proposition 3, the agent attaches to the unsafe project an index $Q(n,F)$, which can be interpreted as the expected payoff yielded by the unsure project. Note that $Q(n,F)$ depends on the expected success to trial ratio and on the evaluation of project payoffs in case of a positive and a negative result.

5.2 Main Results

In this section we investigate the role of loss aversion in the experimentation choice of an agent. According to the literature on the armed bandit problem, the natural benchmark is provided by the framework in which the agent is characterized by the standard expected utility paradigm. In the standard optimizing literature on bandit problem, it is assumed that the agent is risk neutral and maximizes expected profits; we consider, therefore, the case where the utility function of the agent is linear and exactly equal to the payoff[39].

Proposition 4

A loss averse agent experiments more than a risk neutral agent when the following condition on the loss aversion parameter λ is satisfied:

$$\lambda \leq \min\left[\frac{E_k[S_x](1-r-\bar{x})}{\bar{c}[E_k[N_x]-E_k[S_x]]}, k+\frac{E_k[S_x](1-r-\bar{x})}{\bar{c}[E_k[N_x]-E_k[S_x]]}\right]$$

Proof

The agent will decide to experiment more when characterized by disappointment when the following inequality is satisfied:

$$Q^D(n,F) \geq Q^{EUT}(n,F)$$

where:

$Q^D(n,F)$ is calculated according to Proposition 3.

$$Q^{EUT}(n,F) = \max_k \left\{\frac{E_k[S_x]}{E_k[N_x]}\right\}^{40}$$

Under the specific assumption of a piecewise linear disappointment-elation utility function, this is equivalent to checking under which conditions the following inequality holds:

$$\max_k \left\{\frac{E_k[S_x]}{E_k[N_x]}(2-r-\bar{x}) - \left(1-\frac{E_k[S_x]}{E_k[N_x]}\right)\lambda\bar{x}\right\} \geq \max_k \left\{\frac{E_k[S_x]}{E_k[N_x]}\right\}$$

[39] This choice is motivated for homogeneity with the standard economic literature on bandit problems.

[40] See Karlin, Bradt and Johnson (1956)

Two possible cases arises.

1. $\max_k \left\{ \frac{E_k[S_x]}{E_k[N_x]} \right\}$ coincide in both expressions. In this case, the previous inequality simplifies to

$$-\lambda\bar{x} + \frac{E_k[S_x]}{E_k[N_x]}\left\{1-r-\bar{x}-\left(-\lambda\bar{x}\right)\right\} \geq 0 \Rightarrow \lambda \leq \frac{E_k[S_x]\left(1-r-\bar{x}\right)}{\bar{x}\left[E_k[N_x]-E_k[S_x]\right]}$$

2. $\max_k \left\{ \frac{E_k[S_x]}{E_k[N_x]} \right\}$ do not coincide in the expressions of $Q^D(n,F)$ and $Q^{EUT}(n,F)$.

Let us indicate $\overline{\frac{E_k[S_x]}{E_k[N_x]}} = \max_k \left\{ \frac{E_k[S_x]}{E_k[N_x]} \right\}$ under EUT. In this case, the previous inequality becomes

$$\left[\frac{E_k[S_x]}{E_k[N_x]} - \overline{\frac{E_k[S_x]}{E_k[N_x]}}\right] - \lambda\bar{x} + \frac{E_k[S_x]}{E_k[N_x]}\left\{1-r-\bar{x}-\left(-\lambda\bar{x}\right)\right\} \geq 0 \Rightarrow \lambda \leq k + \frac{E_k[S_x]\left(1-r-\bar{x}\right)}{\bar{x}\left[E_k[N_x]-E_k[S_x]\right]}$$

Where $k = \frac{\left[\frac{E_k[S_x]}{E_k[N_x]} - \overline{\frac{E_k[S_x]}{E_k[N_x]}}\right]}{\bar{x}\left[E_k[N_x]-E_k[S_x]\right]}$.

In the standard framework of expected utility theory, the agent faces a well-known trade-off, according to which, in the short run, he bears the loss in case of a negative result. This loss must be traded off against the potential informational gain associated with experimentation, in terms of a more correct estimate of the probability of having a positive result when undertaking a certain project. The agent's optimal strategy consists of finding the optimal way of weighing these two opposite effects associated with the experimental procedure.

The presence of risk aversion reduces the incentives of the player to collect information (as in Anderson, 2001). The experimentation decision is monotonically decreasing in r (for any value of λ), leading the agent to experiment less. The presence of loss-aversion may instead lead the agent to experiment more. The reason may be explained as follows. Loss aversion introduces an additional trade-off associated with experimentation. When playing "safe", the player feels neither disappointment nor elation, since he gets exactly the amount he expects to receive. When playing the risky arm, the player faces the additional cost of disappointment in case of failure and the possible gain of elation when the experimentation proves successful. If the player is loss-averse, he

Two possible cases arises.

1. $\max_k \left\{ \frac{E_k[S_x]}{E_k[N_x]} \right\}$ coincide in both expressions. In this case, the previous inequality simplifies to

$$-\lambda\bar{x} + \frac{E_k[S_x]}{E_k[N_x]}\left\{1 - r - \bar{x} - \left(-\lambda\bar{x}\right)\right\} \geq 0 \Rightarrow \lambda \leq \frac{E_k[S_x]\left(1 - r - \bar{x}\right)}{\bar{x}\left[E_k[N_x] - E_k[S_x]\right]}$$

2. $\max_k \left\{ \frac{E_k[S_x]}{E_k[N_x]} \right\}$ do not coincide in the expressions of $Q^D(n,F)$ and $Q^{EUT}(n,F)$.

Let us indicate $\overline{\frac{E_k[S_x]}{E_k[N_x]}} = \max_k \left\{ \frac{E_k[S_x]}{E_k[N_x]} \right\}$ under EUT. In this case, the previous inequality becomes

$$\left[\frac{E_k[S_x]}{E_k[N_x]} - \overline{\frac{E_k[S_x]}{E_k[N_x]}}\right] - \lambda\bar{x} + \frac{E_k[S_x]}{E_k[N_x]}\left\{1 - r - \bar{x} - \left(-\lambda\bar{x}\right)\right\} \geq 0 \Rightarrow \lambda \leq k + \frac{E_k[S_x]\left(1 - r - \bar{x}\right)}{\bar{x}\left[E_k[N_x] - E_k[S_x]\right]}$$

Where $k = \frac{\left[\frac{E_k[S_x]}{E_k[N_x]} - \overline{\frac{E_k[S_x]}{E_k[N_x]}}\right]}{\bar{x}\left[E_k[N_x] - E_k[S_x]\right]}$.

In the standard framework of expected utility theory, the agent faces a well-known trade-off, according to which, in the short run, he bears the loss in case of a negative result. This loss must be traded off against the potential informational gain associated with experimentation, in terms of a more correct estimate of the probability of having a positive result when undertaking a certain project. The agent's optimal strategy consists of finding the optimal way of weighing these two opposite effects associated with the experimental procedure.

The presence of risk aversion reduces the incentives of the player to collect information (as in Anderson, 2001). The experimentation decision is monotonically decreasing in r (for any value of λ), leading the agent to experiment less. The presence of loss-aversion may instead lead the agent to experiment more. The reason may be explained as follows. Loss aversion introduces an additional trade-off associated with experimentation. When playing "safe", the player feels neither disappointment nor elation, since he gets exactly the amount he expects to receive. When playing the risky arm, the player faces the additional cost of disappointment in case of failure and the possible gain of elation when the experimentation proves successful. If the player is loss-averse, he will give greater weight to disappointment than elation. This additional trade-off plays a key role in characterizing the optimal experimentation strategy of the player.

The condition stated on the parameter λ can be written as follows[41] to highlight the existence of an additional trade-off for the agent:

$$-\lambda\bar{x}+\frac{E_k[S_x]}{E_k[N_x]}\left\{1-r-\bar{x}-\left(-\lambda\bar{x}\right)\right\}\geq 0$$

where the first term represents the loss associated with the disappointment sensation in the short run caused by a negative result when undertaking the risky project. The second term represents the gain associated with the elation sensation in the long run given by the "real" elation sensation of the agent because of a positive result and the elation sensation due to not feeling disappointment on that occasion. This overall elation sensation is weighted by the success to failure ratio (in expected terms).

Common intuition would lead us to think that a loss-averse player would experiment less and choose to opt out the game before a loss-neutral player. We have shown that this standard intuition holds true only for individuals who exhibit a high degree of loss-aversion. On the contrary, players who are moderately loss-averse will choose to experiment more. The reason for this counter-intuitive result can be explained as follows. At every stage of the game, the player trades off the immediate cost of experimentation – due to the possible disappointment experienced in the short run when the failure occurs – with the long-run benefits from experimentation. The understanding of the long-run benefits of experimentation is important at this point. There are three main benefits that the player can obtain from experimentation. First, experimentation provides information on the payouts of the risky arm. Second, experimentation may provide elation when the attempt proves successful. Third, experimentation may "spare" disappointment in the long-run due to the choice of a more profitable arm. For high levels of loss-aversion, the immediate short run disappointment weighs more and may induce the player to opt for the safe arm, avoiding any disappointment cost even in the short run. A moderately loss-averse player, however, may be willing to do more experimentation, facing the risk of a bad run of luck in the short run for the choice of more profitable arm in the remainder of the game.

In other terms, the agent will experiment more under disappointment if the intensity of loss aversion captured by the parameter λ is lower than the ratio of the expected elation to the expected disappointment.

[41] The condition is stated referring to case 1, according to the Proof in Proposition 4, which is the more interesting from an economic point of view.

6. Conclusions

In the present chapter we consider the bandit problem under alternative models of rational behavior, bridging the methodological gap between the standard optimizing analysis of the bandit problem and the modern literature on alternatives to the expected utility model of behavior under uncertainty, pioneered by Loomes and Sugden, among others. We depart from the classical setting of bandit problems by endowing the agent with a disappointment-elation utility function (Sugden and Loomes, 1986), in presence of loss aversion.

Common intuition would lead us to think that in an experimental setting a loss-averse player would experiment less and choose to opt out the game before a loss-neutral player. We have shown that this standard intuition holds true only for individuals who exhibit a high degree of loss-aversion. On the contrary, players who are moderately loss-averse will choose to experiment more. A moderately loss-averse player, in fact, may be willing to do more experimentation, facing the risk of a bad run of luck in the short run for the choice of more profitable arm in the remainder of the game.

Appendix

The derivation of the numerical example

3.1.1 Backward induction

In the three-stage example under the assumption of a prior $\phi = \frac{1}{2}$, the probabilistic structure of the game will be as follows.

In the first stage of the game ($N = 1$), the history of the game is the empty set ($h = \{\varnothing\}$) and the player has three periods of experimentation ($n = 3$). The experimenter will evaluate the likelihood of a success along the risky arm at the prior $\phi = \frac{1}{2}$.

In the second stage of the game ($N = 2$), the history in stage 1 of the game $h = \{h_1\}$ can either be a failure, i.e. $h_1 = \{0\}$ or a success, i.e. $h_1 = \{1\}$. The player has two periods of experimentation ($n = 2$). In case $h_1 = \{0\}$, the experimenter will evaluate the likelihood of a success along the risky arm at the prior $\psi_{2,\{0\}} = \frac{1}{3}$. In case $h_1 = \{1\}$, the experimenter will evaluate the likelihood of a success along the risky arm at the prior $\psi_{2,\{1\}} = \frac{2}{3}$.

In the third stage of the game ($N = 3$), the history in stage 1 and 2 of the game $h = \{h_1, h_2\}$ can take four possible values: either two failures in both periods, i.e. $\{h_1, h_2\} = \{0,0\}$, one success and one failure, i.e. $\{h_1, h_2\} = \{0,1\}$, one failure and one success, i.e. $\{h_1, h_2\} = \{1,0\}$ or two successes, i.e. $\{h_1, h_2\} = \{1,1\}$. The player has one period of experimentation ($n = 1$). In case $\{h_1, h_2\} = \{0,0\}$, the experimenter will evaluate the likelihood of a success along the risky arm at the prior $\psi_{3,\{0,0\}} = \frac{1}{4}$. In case of histories $\{h_1, h_2\} = \{0,1\}$ or $\{h_1, h_2\} = \{1,0\}$ the experimenter will evaluate the likelihood of a success along the risky arm at the prior $\psi_{3,\{0,1\}} = \psi_{3,\{1,0\}} = \frac{1}{2}$. In case $\{h_1, h_2\} = \{1,1\}$, the experimenter will evaluate the likelihood of a success along the risky arm at the prior $\psi_{3,\{1,1\}} = \frac{3}{4}$.

In this section we provide a full characterization of experimentation strategy by backward induction. The optimal strategy is defined conditionally on the state of world, represented by the history that leads to continued experimentation.

In the last stage of the game $N = 3$ the history of play is given by $h = \{h_1, h_2\}$. For each history the optimal experimentation strategy is defined in terms of a threshold value of q below which it is optimal to experiment. In this illustrative example, this condition becomes:

$$q \leq \psi_{3,\{h_1,h_2\}}$$

In the last period of the game four states of the world are possible: the experimentation has lead to two successes in the first two periods, i.e. $\{h_1, h_2\} = \{1,1\}$, one success in first period and one failure in second period, i.e. $\{h_1, h_2\} = \{1,0\}$, one failure in first period and one success in second period, i.e. $\{h_1, h_2\} = \{0,1\}$, or two failures, i.e. $\{h_1, h_2\} = \{0,0\}$.

The agent's optimal strategy in period 3 will be to experiment if:

$q \leq \frac{1}{4}$ if $\{h_1, h_2\} = \{0,0\}$;

$q \leq \frac{1}{2}$ if$\{h_1, h_2\} = \{0,1\}$;

$q \leq \frac{3}{4}$ if $\{h_1, h_2\} = \{1,1\}$.

The optimal strategy of the experimenter in stage 2 of the game ($N = 2$) is defined contingently on a success or a failure in period 1.

Suppose $\{h_1\} = \{0\}$. If the agent switches to the safe project in $N = 2$, by Lemma in Gittins (1979), the agent sticks with the safe one in period 3, with an expected payoff equal to $2q$. If the agent plays risky in period 2, two states of the world after period 2 are possible: $\{h_1, h_2\} = \{0,0\}$ and $\{h_1, h_2\} = \{0,1\}$.

We need therefore to distinguish two cases. If $q \leq \frac{1}{4}$, the agent will always experiment and has a combined expected payoff from periods 2 and 3 equal to $\frac{1}{3}\left(1 + \frac{1}{2}\right) = \frac{1}{2}$.

If $\frac{1}{4} < q < \frac{1}{2}$, the agent will play risky in $N = 3$ if a success occurs in $N = 2$ and switches to safe project in $n = 3$ if a failure occurs in $N = 2$. In this case the combined expected payoff from periods 2 and 3 is equal to $\frac{1}{3}\left(1+\frac{1}{2}\right)+\frac{2}{3}q$. Henceforth the agent will find optimal to choose to experiment in period 2 if the expected combined payoff $\frac{1}{2}+\frac{2}{3}q$ is higher than the payoff from playing twice on the safe arm, equal to $2q$. The critical value of q below which it is optimal to experiment in period 2 following a failure in period 1 will be therefore:

$$q \leq \frac{3}{8}$$

Suppose $\{h_1\}=\{1\}$. The optimal experimentation strategy is characterized in terms of a q threshold in an analogous way. We should not distinguish whether $q \leq \frac{1}{4}$ or $\frac{1}{4}<q<\frac{1}{2}$. The expected payoff combined from period 2 and 3 is equal to $\frac{1}{3}\frac{1}{2}+\frac{2}{3}\left(1+\frac{3}{4}\right)=\frac{4}{3}$. The agent will find optimal to experiment in period 2 after a success in period 1 if the expected combined payoff is higher than the payoff from playing twice along the safe arm, equal to $2q$. This implies that the critical value of q below which it is optimal to experiment in period 2 following a success in period 1 becomes:

$$q \leq \frac{2}{3}$$

Summarizing, the agent's optimal strategy in period 2 will be to experiment if:

$$q \leq \frac{3}{8} \text{ if } \{h_1\}=\{0\}$$

$$q \leq \frac{2}{3} \text{ if } \{h_1\}=\{1\}$$

Proceeding by backward induction, we need to characterize the player's optimal experimentation strategy in period 1 ($N=1$). Given the prior ϕ, the expected payoff will be $\frac{1}{2}\left(1+\frac{4}{3}\right)+\frac{1}{2}\left[\frac{1}{2}+\frac{2}{3}q\right]=$. In the first period, the agent will choose optimally to experiment if the payoff from optimal experimentation is higher than the payoff from playing three times on the safe arms, equal to $3q$, i.e. if:

$$q \leq \frac{17}{32}$$

3.1.2 General method

In the first stage of the game ($N=1$), with three periods of experimentation ($n=3$), the history is an empty set $\{h_1\}=\{\varnothing\}$. The individual will choose the strategy as a pair of $\{k_1,k_2\}$ in order to maximize the dynamic allocation index:

$$Q(3,\phi,h=\{\varnothing\})=\max_{\{k_1,k_2\}}\left\{\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}\right\}$$

The individual will decide to experiment if and only if:

$$q\leq Q(3,\phi,h=\{\varnothing\})$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=0\}$ is calculated as follows. Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be $\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}}=\frac{1}{3}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]\right\}$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=1\}$ is calculated as follows. Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be $\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)q\right]$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}}=\frac{1}{2+\psi_{2,\{0\}}+\phi\left(1-\psi_{2,\{0\}}\right)}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\right\}$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=2\}$ is calculated as follows. Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be $\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)q\right]$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}} = \frac{1}{2+\psi_{2,\{0\}}+\phi\left(1-\psi_{2,\{0\}}\right)}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+\left(1-\phi\right)\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\right\}$$

The dynamic allocation index associated with strategy $\{k_1=1;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be $\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+\left(1-\phi\right)2q$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}} = \frac{1}{1+2\phi}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\right\}$$

The dynamic allocation index associated with strategy $\{k_1=1;k_2=1\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be $\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+\left(1-\phi\right)2q$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}} = \frac{1}{1+2\phi}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\right\}$$

The dynamic allocation index associated with strategy $\{k_1=2;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be $\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)q\right]+\left(1-\phi\right)2q$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}} = \frac{1}{1+\phi\left(1+\psi_{2,\{1\}}\right)}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)\right]\right\}$$

In case the prior $\phi=\frac{1}{2}$, $Q\left(3,\phi,h=\{\varnothing\}\right)=\frac{17}{32}$.

In the second stage of the game ($N=2$), with two periods of experimentation ($n=2$), the agent will keep playing on the risky arm following a history $\{h_1\}=\{1\}$. In case of a history $\{h_1\}=\{0\}$, the individual will calculate the dynamic allocation index $Q(2,\phi,h_1=0)$, given a failure in $N=1$. The optimal strategy is defined in terms of k_1, that can take only two possible values, either 0 or 1.

The dynamic allocation index under the strategy $k_1=0$ is calculated as follows. The combined expected payoff in the two remaining periods of experimentation is equal to $\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}$, since he will play risky in both periods. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $2q$.

The dynamic allocation index under the strategy $k_1=1$ is calculated as follows. The combined expected payoff in the remaining periods ($N=2$ and $N=3$) is equal to $\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)q$, since he will play safe in case of a failure in period $N=2$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $2q$.

Hence, the agent will experiment on the risky arm at $N=2$, if :

$$q \le Q(n=2,\phi,h_1=0)$$

where the dynamic allocation index takes the following form:

$$Q(2,\phi,h_1=0)=\max_{k_1=\{0,1\}}\left\{\tilde{q}\Big|_{k_1=0};\tilde{q}\Big|_{k_1=1}\right\}$$

where:

$$\tilde{q}\Big|_{k_1=0}=\frac{1}{2}\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]$$

$$\tilde{q}\Big|_{k_1=1}=\frac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}\left(1+\psi_{3,\{0,1\}}\right)$$

In case the prior $\phi=\frac{1}{2}$, $Q\left(2,\phi=\frac{1}{2},h_1=0\right)=\max\left\{\frac{1}{3};\frac{3}{8}\right\}=\frac{3}{8}$ associated with strategy $k_1=1$.

Hence the agent will choose to experiment if $q\le\frac{3}{8}$, which coincides with the optimal experimentation strategy calculated by backward induction.

In the third stage of the game ($N=3$), with one period of experimentation ($n=1$), the agent will keep playing on the risky arm following a history $\{h_1,h_2\}=\{1,1\}$ or $\{h_1,h_2\}=\{0,1\}$. In case of a history $\{h_1,h_2\}=\{0,0\}$, the individual will calculate the updated dynamic allocation index to attach to the risky arm and compare q with the expected payoff along the risky arm, given that he experienced a failure in the first two runs of trial. The decision problem simplifies to a one-stage decision problem. Hence, the agent will play risky if:

$$q \le Q(n=1,\phi,h_1=h_2=0)$$

where the dynamic allocation index simplifies to:

$$Q(n=1,\phi,h_1=h_2=0)=\psi_{3,\{0,0\}}$$

and

$$\psi_{3,\{0,0\}}=\frac{\mu_{\{0,0\}}+\rho_{\{0,0\}}}{1+\rho_{\{0,0\}}} \text{ if } \{h_1,h_2\}=\{0,0\}$$

The experimentation strategy coincides with the one calculated by backward induction. In case the prior $\phi=\frac{1}{2}$, $Q\left(n=1,\phi=\frac{1}{2},h_1=h_2=0\right)=\frac{1}{4}$.

3.1.3 Introducing risk aversion

In the first stage of the game, the history is the empty space, $\{h_1\}=\{\varnothing\}$. The individual will choose the strategy as a pair of $\{k_1,k_2\}$ in order to maximize the index attached to the risky arm (representing the expected payoff from experimenting). We denote with Q^{Ra} the dynamic allocation index of a risk averse agent. Analytically, the experimenter will choose:

$$Q^{Ra}(3,\phi,h=\{\varnothing\})=\max_{\{k_1,k_2\}}\left\{\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}\right\}$$

The individual will decide to experiment if and only if:

$$q \le Q^{Ra}(3,\phi,h=\{\varnothing\})$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]\right\}(1-r).$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}}=\frac{1}{3}(1-r)\left\{\begin{array}{l}\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+\\+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]\end{array}\right\}$$

The dynamic allocation index associated with strategy $\{k_1 = 0; k_2 = 1\}$ is calculated as follows. Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be $\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right](1-r)+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)(1-r)+\left(1-\psi_{2,\{0\}}\right)q\right]$ This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}} = \frac{1-r}{2+\psi_{2,\{0\}}+\phi\left(1-\psi_{2,\{0\}}\right)}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\right\}$$

The dynamic allocation index associated with strategy $\{k_1 = 0; k_2 = 2\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right](1-r)+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)(1-r)+\left(1-\psi_{2,\{0\}}\right)q\right].$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}} = \frac{1-r}{2+\psi_{2,\{0\}}+\phi\left(1-\psi_{2,\{0\}}\right)}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\right\}$$

The dynamic allocation index associated with strategy $\{k_1 = 1; k_2 = 0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be $\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right](1-r)+(1-\phi)2q$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}} = \frac{1}{1+2\phi}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\right\}$$

The dynamic allocation index associated with strategy $\{k_1 = 1; k_2 = 1\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be $\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right](1-r)+(1-\phi)2q$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}} = \frac{1-r}{1+2\phi}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\right\}$$

The dynamic allocation index associated with strategy $\{k_1 = 2; k_2 = 0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be $\phi\left\{\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)\right](1-r)+\left(1-\psi_{2,\{1\}}\right)q\right\}+(1-\phi)2q$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}} = \frac{1-r}{1+\phi\left(1+\psi_{2,\{1\}}\right)}\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)\right]\right\}$$

In case the prior $\phi = \frac{1}{2}$, $Q^{Ra}\left(3,\phi,h=\{\varnothing\}\right)=\frac{17}{32}(1-r)$.

In the second stage of the game, following a history $\{h_1\}=\{1\}$, the agent will keep playing on the risky arm. Following a history $\{h_1\}=\{0\}$, the player will experiment on the risky arm at $n=2$ if:

$$q \le Q^{Ra}\left(2,\phi,h_1=0\right)$$

Given a time horizon of the decisional problem reduced to $n=2$, the optimal strategy can be defined in terms of k_1, that can either take two possible values, 0 or 1.

The dynamic allocation index under the strategy $k_1=0$ is calculated as follows. The combined expected payoff in the two remaining periods of experimentation is equal to $\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right](1-r)$, since he will play risky in both periods. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $2q$.

The dynamic allocation index under strategy $k_1=1$ is calculate as follows. The combined expected payoff in the two remaining periods ($N=2$ and $N=3$) is equal to $\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)(1-r)+\left(1-\psi_{2,\{0\}}\right)q$, since he will play safe in case of a failure in period $n=2$.

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $2q$. Hence, the agent will experiment on the risky arm at $n=2$, if :

$$q \le Q^{Ra}\left(n=2,\phi,h_1=0\right)$$

where the dynamic allocation index takes the following form:

$$Q^{Ra}(2,\phi,h_1=0)=\max_{k_1=\{0,1\}}\left\{\tilde{q}\Big|_{k_1=0};\tilde{q}\Big|_{k_1=1}\right\}$$

where:

$$\tilde{q}\Big|_{k_1=0}=\frac{1}{2}\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right](1-r)$$

$$\tilde{q}\Big|_{k_1=1}=\frac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}\left(1+\psi_{3,\{0,1\}}\right)(1-r)$$

Assuming a prior $\phi=\frac{1}{2}$, $Q^{Ra}\left(2,\phi,h_1=\{0\}\right)=\frac{3}{8}(1-r)$.

In the last stage of the game ($N=3$), the agent will keep playing on the risky arm following histories $\{h_1,h_2\}=\{1,1\}$ or $\{h_1,h_2\}=\{0,1\}$.

Following a history $\{h_1,h_2\}=\{0,0\}$, the agent's optimal strategy in period 3 will be to experiment if:

$$q\le Q^{Ra}\left(1,\phi,h_1=h_2=0\right)$$

where $Q^{Ra}\left(1,h_1=h_2=0\right)=\psi_{3,\{0,0\}}(1-r)$.

Assuming a prior $\phi=\frac{1}{2}$, $Q^{Ra}\left(1,\phi=\frac{1}{2},h_1=h_2=0\right)=\frac{1}{4}(1-r)$

3.2.1 Bandit Problem and Loss Aversion: the "No Updating" case

In the first stage of the game, the individual will choose the strategy as a pair of $\{k_1,k_2\}$ in order to maximize the index attached to the risky arm (representing the expected payoff from experimenting). We denote with Q^{La} the dynamic allocation index of a loss averse agent. Analytically, the experimenter will choose:

$$Q^{La}\left(3,\phi,h=\{\varnothing\}\right)=\max_{\{k_1,k_2\}}\left\{\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}\right\}$$

The individual will decide to experiment if and only if:

$$q\le Q^{La}\left(3,\phi,h=\{\varnothing\}\right)$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]\right\}(2-\phi)-$$

$$-\left\{\phi\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right]+(1-\phi)\begin{bmatrix}2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)+\\+\left(1-\psi_{2,\{0\}}\right)\left(1-\psi_{3,\{0,1\}}\right)\end{bmatrix}\right\}(\lambda\phi)\qquad T$$

his payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\left.\tilde{q}\right|_{\substack{k_1=0\\k_2=0}} =$$

$$\frac{\left\{\begin{matrix}\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+\\+\left(1-\phi\right)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]\end{matrix}\right\}\left(2-\phi\right)-\left\{\begin{matrix}\phi\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right]+\\+\left(1-\phi\right)\begin{bmatrix}2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)+\\\left(1-\psi_{2,\{0\}}\right)\left(1-\psi_{3,\{0,1\}}\right)\end{bmatrix}\end{matrix}\right\}\left(\lambda\phi\right)}{3}$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=1\}$ is calculated as follows. Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\left(2-\phi\right)+\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right]\left(-\lambda\phi\right)\right]+$$
$$+\left(1-\phi\right)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\left(2-\phi\right)-\left[2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\right]\lambda\phi+\left(1-\psi_{2,\{0\}}\right)q\right]$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\left.\tilde{q}\right|_{\substack{k_1=0\\k_2=1}}=\frac{\begin{matrix}\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\left(2-\phi\right)+\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right]\left(-\lambda\phi\right)\right]+\\+\left(1-\phi\right)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\left(2-\phi\right)-\left[2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\right]\lambda\phi\right]\end{matrix}}{2+\phi+\left(1-\phi\right)\psi_{2,\{0\}}}$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=2\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\left(2-\phi\right)+\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right]\left(-\lambda\phi\right)\right]+$$
$$+\left(1-\phi\right)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\left(2-\phi\right)-\left[2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\right]\lambda\phi+\left(1-\psi_{2,\{0\}}\right)q\right]$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}} = \frac{\begin{array}{c}\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right](2-\phi)+\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right](-\lambda\phi)\right]+\\+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)(2-\phi)-\left[2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\right]\lambda\phi\right]\end{array}}{2+\phi+(1-\phi)\psi_{2,\{0\}}}$$

The dynamic allocation index associated with strategy $\{k_1=1;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right](2-\phi)+\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right](-\lambda\phi)\right]$$
$$+(1-\phi)(-\lambda\phi+2q)$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}} = \frac{\phi\left\{\begin{array}{l}\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right](2-\phi)+\\\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right](-\lambda\phi)\end{array}\right\}-(1-\phi)\lambda\phi}{1+2\phi}$$

The dynamic allocation index associated with strategy $\{k_1=1;k_2=1\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right](2-\phi)+\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right](-\lambda\phi)\right]$$
$$+(1-\phi)(-\lambda\phi+2q)$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}} = \frac{\phi\left\{\begin{array}{l}\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right](2-\phi)+\\\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right](-\lambda\phi)\end{array}\right\}-(1-\phi)\lambda\phi}{1+2\phi}$$

The dynamic allocation index associated with strategy $\{k_1=2;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)\right](2-\phi)+\left(1-\psi_{2,\{1\}}\right)[-\lambda\phi+q]+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)(-\lambda\phi)\right]+(1-\phi)(-\lambda\phi+2q).$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}} = \frac{\phi\left\{\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)\right](2-\phi)-\left(1-\psi_{3,\{1,1\}}\right)\lambda\phi\right\}-(1-\phi)\lambda\phi}{1+\phi\left(1+\psi_{2,\{1\}}\right)}$$

In case the prior $\phi=\frac{1}{2}$, $Q^{La}\left(3,\phi,h=\{\varnothing\}\right)=\max\left\{\frac{39}{44}-\frac{15}{88}\lambda;\frac{7}{8}-\frac{5}{24}\lambda;\frac{51}{64}-\frac{15}{64}\lambda;\frac{1}{18}-\frac{17}{72}\lambda\right\}$.

In the second stage of the game ($N=2$), the agent will keep playing on the risky arm following a history $\{h_1\}=\{1\}$. In case of a history $\{h_1\}=\{0\}$, the player will experiment on the risky arm at $n=2$, if:

$$q \le Q^{La}\left(n=2,\phi,h_1=0\right)$$

The dynamic allocation index under the strategy $k_1=0$ is calculated as follows. The combined expected payoff in the two remaining periods of experimentation is equal to
$\psi_{2,\{0\}}\left[\left(1+\psi_{3,\{0,1\}}\right)\left(2-\psi_{2,\{0\}}\right)-\left(1-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)\right]+\left(1-\psi_{2,\{0\}}\right)\left[-\left(2-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)+\psi_{3,\{0,1\}}\left(2-\psi_{2,\{0\}}\right)\right]$
since he will play risky in both periods. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $2q$.

The dynamic allocation index under the strategy $k_1=1$ is calculated as follows. The combined expected payoff in the remaining periods ($N=2$ and $N=3$) is equal to
$\psi_{2,\{0\}}\left[\left(1+\psi_{3,\{0,1\}}\right)\left(2-\psi_{2,\{0\}}\right)-\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{2,\{0\}}\right]+\left(1-\psi_{2,\{0\}}\right)\left(-\lambda\psi_{2,\{0\}}+q\right)$, since he will play safe in case of a failure in period $n=2$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $2q$.

Hence the agent will not experiment at $n=2$, if:

$$q \le Q^{La}\left(n=2,\phi,h_1=0\right)$$

where the dynamic allocation index takes the following form:

$$Q^{La}(2,\phi,h_1=0)=\max_{k_1=\{0,1\}}\left\{\tilde{q}\Big|_{k_1=0};\tilde{q}\Big|_{k_1=1}\right\}$$

and

$$\tilde{q}^{La}\Big|_{k_1=0}=\frac{1}{2}\begin{bmatrix}\psi_{2,\{0\}}\left[\left(1+\psi_{3,\{0,1\}}\right)\left(2-\psi_{2,\{0\}}\right)-\left(1-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)\right]+\\+\left(1-\psi_{2,\{0\}}\right)\left[-\left(2-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)+\psi_{3,\{0,1\}}\left(2-\psi_{2,\{0\}}\right)\right]\end{bmatrix}$$

$$\tilde{q}^{La}\Big|_{k_1=1} = \frac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}\left[\left(2-\psi_{2,\{0\}}\right)\left(1+\psi_{3,\{0,1\}}\right)-\lambda\left(1-\psi_{2,\{0\}}\psi_{3,\{0,1\}}\right)\right]$$

Assuming a prior $\phi=\frac{1}{2}$, $Q^{La}\left(2,\phi=\frac{1}{2},h_1=0\right)=\max\left\{\frac{25}{36}-\frac{7}{36}\lambda;\frac{5}{8}-\frac{5}{24}\lambda\right\}$

In the last stage of the game, following history $\{h_1,h_2\}=\{1,1\}$ or $\{h_1,h_2\}=\{0,1\}$, the agent will keep playing on the risky arm. In case of a history $\{h_1,h_2\}=\{0,0\}$, the agent will play risky if:

$$q \le Q^{La}(n=1,\phi,h_1=h_2=0)$$

where the dynamic allocation index simplifies to:

$$Q^{La}(n=1,\phi,h_1=h_2=0)=\psi_{3,\{0,0\}}\left(2-\psi_{3,\{0,0\}}\right)-\left(1-\psi_{3,\{0,0\}}\right)\left(\lambda\psi_{3,\{0,0\}}\right)$$

In case the prior $\phi=\frac{1}{2}$, $Q^{La}(n=1,\phi,h_1=h_2=0)=\frac{1}{16}(7-3\lambda)$, where $\bar{x}=\psi_{3,\{0,0\}}=\frac{1}{4}$.

3.2.1.1 A clarification: the case of loss-neutrality

We use backward induction. Consider stage *N=3* (with only one period of experimentation ahead). After a history of two failures: $h=\{0;0\}$, the "updated" probability of a success=$\frac{1}{4}$ and the expected payoff is equal to $\bar{x}_{\{0,0\}}=\frac{1}{4}$

The utility function of a loss neutral agent is equal to

$$U(x)=\begin{cases} 2-\bar{x}_{\{0,0\}}=\frac{7}{4} & \textit{if} \quad \textit{success} \\ -\bar{x}_{\{0,0\}}=-\frac{1}{4} & \textit{if} \quad \textit{fail} \end{cases}$$

Along the risky arm, the expected payoff is:

$$Q(3,h=\{0,0\})=\frac{1}{4}\frac{7}{4}+\frac{3}{4}\left(-\frac{1}{4}\right)=\frac{1}{4}$$

Hence, the agent experiments if:

$$q\le\frac{1}{4}$$

Consider stage *N=2* (with two periods of experimentation ahead). After a history of one failures: $h=\{0\}$, the "updated" probability of a success is equal to $\frac{1}{3}$ and the expected payoff is equal to $\bar{x}_{\{0\}}=\frac{1}{3}$.

The utility function of a loss neutral agent is equal to

$$U(x)=\begin{cases} 2-\bar{x}_{\{0\}}=\frac{5}{3} & if \quad success \\ -\bar{x}_{\{0\}}=-\frac{1}{3} & if \quad fail \end{cases}$$

We should distinguish two cases along the risky arm:

$q \le \frac{1}{4}$: the agent chooses to experiment in both periods

$\frac{1}{4} < q \le \frac{1}{2}$: the agent chooses to experiment in period N=2 and switches to the safe arm in period N=3.

If $q \le \frac{1}{4}$, the expected payoff of a loss-neutral agent becomes equal to:

$$\frac{1}{3}\frac{5}{3}+\frac{2}{3}\left(-\frac{1}{3}\right)+\frac{1}{3}\left[\frac{1}{2}\frac{5}{3}+\frac{1}{2}\left(-\frac{1}{3}\right)\right]+\frac{2}{3}\left[\frac{1}{4}\frac{5}{3}+\frac{3}{4}\left(-\frac{1}{3}\right)\right]=\frac{2}{3}$$

In alternative, the agent can play twice on safe arm, with a payoff equal to 2q.

Hence, the cutoff value of the agent becomes at *N=2*:

$$q \le \frac{1}{3}$$

Hence, the loss-neutral agent experiments if $q \le \frac{1}{3}$

If $\frac{1}{4} < q \le \frac{1}{2}$, the expected payoff of a loss-neutral agent is:

$$\frac{1}{3}\frac{5}{3}+\frac{2}{3}\left(-\frac{1}{3}\right)+\frac{1}{3}\left[\frac{1}{2}\frac{5}{3}+\frac{1}{2}\left(-\frac{1}{3}\right)\right]+\frac{2}{3}q=\frac{5}{9}+\frac{2}{3}q$$

In alternative, the agent can play twice on safe arm, with a payoff equal to 2q. Hence:

$$q \le \frac{5}{12}$$

The optimal cutoff value of a a loss-neutral agent takes the following form:

$$Q(2,h=\{0\})=\max\left\{\frac{1}{3};\frac{5}{12}\right\}=\frac{5}{12}$$

Consider stage *N=1* (with three periods of experimentation ahead). After a null history (no prior experimentation decisions): $h=\{\varnothing\}$. The agent's prior probability of a success along the risky arm is equal to $\frac{1}{2}$ and the expected payoff of the risky arm is equal to $\bar{x}_{\{\varnothing\}}=\frac{1}{2}$.

The utility function of the loss-neutral agent becomes equal to:

$$U(x)=\begin{cases} 2-\bar{x}_{\{\varnothing\}}=\dfrac{3}{2} & if \quad success \\ -\bar{x}_{\{\varnothing\}}=-\dfrac{1}{2} & if \quad fail \end{cases}$$

The expected payoff is

$$\frac{1}{2}\left[\frac{3}{2}-\frac{1}{2}\right]+\frac{1}{2}\left[\frac{2}{3}\frac{3}{2}+\frac{1}{3}\left(-\frac{1}{2}\right)\right]+\frac{1}{2}\frac{2}{3}\left[\frac{3}{4}\frac{3}{2}+\frac{1}{4}\left(-\frac{1}{2}\right)\right]+\frac{1}{2}\frac{1}{3}\left[\frac{1}{2}\frac{3}{2}+\frac{1}{2}\left(-\frac{1}{2}\right)\right]+$$
$$+\frac{1}{2}\left[\frac{1}{3}\frac{3}{2}+\frac{2}{3}\left(-\frac{1}{2}\right)\right]+\frac{1}{2}\frac{1}{3}\left[\frac{1}{2}\frac{3}{2}+\frac{1}{2}\left(-\frac{1}{2}\right)\right]+\frac{1}{2}\frac{2}{3}q=$$
$$=\frac{3}{2}+\frac{1}{3}q$$

In alternative, the agent can play twice on safe arm, with a payoff equal to 3q.

Hence, the optimal cutoff of the loss-neutral agent is:

$$q\leq\frac{9}{16}$$

Hence:

$$Q(3,h=\{\varnothing\})=\frac{9}{16}$$

3.2.2 Loss aversion and updating the reference point

In the first stage of the game, the individual will choose the strategy as a pair of $\{k_1,k_2\}$ in order to maximize the index attached to the risky arm (representing the expected payoff from experimenting). We denote with Q^{La} the dynamic allocation index of a loss averse agent. Analytically, the experimenter will choose:

$$Q^{La,Updating}(3,\phi,h=\{\varnothing\})=\max_{\{k_1,k_2\}}\left\{\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}\right\}$$

The individual will decide to experiment if and only if:

$$q\leq Q^{La,Updating}(3,\phi,h=\{\varnothing\})$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\left\{\phi\begin{bmatrix}2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)\\+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\left(2-\psi_{3,\{1,0\}}\right)\end{bmatrix}+(1-\phi)\begin{bmatrix}\psi_{2,\{0\}}\left(2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\right)\\+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\left(2-\psi_{3,\{0,0\}}\right)\end{bmatrix}\right\}-$$
$$-\lambda\left\{\phi\begin{bmatrix}\left(1-\psi_{2,\{1\}}\right)\left(\psi_{2,\{1\}}+\left(1-\psi_{3,\{1,0\}}\right)\psi_{3,\{1,0\}}\right)\\+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\psi_{3,\{1,1\}}\end{bmatrix}+(1-\phi)\begin{bmatrix}\phi+\left(1-\psi_{2,\{0\}}\right)\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\psi_{3,\{0,1\}}+\\+\left(1-\psi_{2,\{0\}}\right)\left(1-\psi_{3,\{0,0\}}\right)\psi_{3,\{0,0\}}\end{bmatrix}\right\}$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}}=$$
$$\frac{\left\{\phi\begin{bmatrix}2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)\\+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\left(2-\psi_{3,\{1,0\}}\right)\end{bmatrix}+(1-\phi)\begin{bmatrix}\psi_{2,\{0\}}\left(2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\right)\\+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\left(2-\psi_{3,\{0,0\}}\right)\end{bmatrix}\right\}-\lambda\left\{\phi\begin{bmatrix}\left(1-\psi_{2,\{1\}}\right)\left(\psi_{2,\{1\}}+\left(1-\psi_{3,\{1,0\}}\right)\psi_{3,\{1,0\}}\right)\\+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\psi_{3,\{1,1\}}\end{bmatrix}+(1-\phi)\begin{bmatrix}\phi+\left(1-\psi_{2,\{0\}}\right)\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\psi_{3,\{0,1\}}+\\+\left(1-\psi_{2,\{0\}}\right)\left(1-\psi_{3,\{0,0\}}\right)\psi_{3,\{0,0\}}\end{bmatrix}\right\}}{3}$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=1\}$ is calculated as follows. Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\begin{bmatrix}2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)\\+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\end{bmatrix}-\lambda\begin{bmatrix}\left(1-\psi_{2,\{1\}}\right)\left(\psi_{2,\{1\}}+\left(1-\psi_{3,\{1,0\}}\right)\psi_{3,\{1,0\}}\right)+\\+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\psi_{3,\{1,1\}}\end{bmatrix}\right]+$$
$$+(1-\phi)\left[\psi_{2,\{0\}}\left(2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\right)-\lambda\left[\phi+\left(1-\psi_{2,\{0\}}\right)\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\psi_{3,\{0,1\}}\right]+\left(1-\psi_{2,\{0\}}\right)q\right]$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}}=\frac{\phi\left[\begin{bmatrix}2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)\\+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\end{bmatrix}-\lambda\begin{bmatrix}\left(1-\psi_{2,\{1\}}\right)\left(\psi_{2,\{1\}}+\left(1-\psi_{3,\{1,0\}}\right)\psi_{3,\{1,0\}}\right)+\\+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\psi_{3,\{1,1\}}\end{bmatrix}\right]+\\+(1-\phi)\left[\psi_{2,\{0\}}\left(2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\right)-\lambda\left[\phi+\left(1-\psi_{2,\{0\}}\right)\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\psi_{3,\{0,1\}}\right]\right]}{2+\phi+(1-\phi)\psi_{2,\{0\}}}$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=2\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\left[\begin{array}{l}2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)\\+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\end{array}\right]-\lambda\left[\begin{array}{l}\left(1-\psi_{2,\{1\}}\right)\left(\psi_{2,\{1\}}+\left(1-\psi_{3,\{1,0\}}\right)\psi_{3,\{1,0\}}\right)+\\+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\psi_{3,\{1,1\}}\end{array}\right]\right]+$$
$$+(1-\phi)\left[\psi_{2,\{0\}}\left(2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\right)-\lambda\left[\phi+\left(1-\psi_{2,\{0\}}\right)\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\psi_{3,\{0,1\}}\right]+\left(1-\psi_{2,\{0\}}\right)q\right]$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}}=\frac{\phi\left[\left[\begin{array}{l}2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)\\+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\end{array}\right]-\lambda\left[\begin{array}{l}\left(1-\psi_{2,\{1\}}\right)\left(\psi_{2,\{1\}}+\left(1-\psi_{3,\{1,0\}}\right)\psi_{3,\{1,0\}}\right)+\\+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\psi_{3,\{1,1\}}\end{array}\right]\right]+ \\ +(1-\phi)\left[\psi_{2,\{0\}}\left(2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\right)-\lambda\left[\phi+\left(1-\psi_{2,\{0\}}\right)\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\psi_{3,\{0,1\}}\right]\right]}{2+\phi+(1-\phi)\psi_{2,\{0\}}}$$

The dynamic allocation index associated with strategy $\{k_1=1;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\begin{array}{l}\left[2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\left(2-\psi_{3,\{1,0\}}\right)\right]+\\\left[-\left(1-\psi_{2,\{1\}}\right)\left[\lambda\psi_{2,\{1\}}+\left(1-\psi_{3,\{1,0\}}\right)\lambda\psi_{3,\{1,0\}}\right]-\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\lambda\psi_{3,\{1,1\}}\right]\end{array}\right]+(1-\phi)(-\lambda\phi+2q)$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}}=\frac{\phi\left[\begin{array}{l}\left[2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\left(2-\psi_{3,\{1,0\}}\right)\right]+\\\left[-\left(1-\psi_{2,\{1\}}\right)\left[\lambda\psi_{2,\{1\}}+\left(1-\psi_{3,\{1,0\}}\right)\lambda\psi_{3,\{1,0\}}\right]-\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\lambda\psi_{3,\{1,1\}}\right]\end{array}\right]-(1-\phi)\lambda\phi}{1+2\phi}$$

The dynamic allocation index associated with strategy $\{k_1=1;k_2=1\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\begin{array}{l}\left[2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\left(2-\psi_{3,\{1,0\}}\right)\right]+\\\left[-\left(1-\psi_{2,\{1\}}\right)\left[\lambda\psi_{2,\{1\}}+\left(1-\psi_{3,\{1,0\}}\right)\lambda\psi_{3,\{1,0\}}\right]-\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\lambda\psi_{3,\{1,1\}}\right]\end{array}\right]+(1-\phi)(-\lambda\phi+2q)$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}}=\frac{\phi\begin{bmatrix}\left[2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\left(2-\psi_{3,\{1,0\}}\right)\right]+\\ \left[-\left(1-\psi_{2,\{1\}}\right)\left[\lambda\psi_{2,\{1\}}+\left(1-\psi_{3,\{1,0\}}\right)\lambda\psi_{3,\{1,0\}}\right]-\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\lambda\psi_{3,\{1,1\}}\right]\end{bmatrix}-\left(1-\phi\right)\lambda\phi}{1+2\phi}$$

The dynamic allocation index associated with strategy $\{k_1=2;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\left[2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)\right]+\left(1-\psi_{2,\{1\}}\right)\left[-\lambda\psi_{2,\{1\}}+q\right]-\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\lambda\psi_{3,\{1,1\}}\right]+$$
$$+\left(1-\phi\right)\left(-\lambda\phi+2q\right)$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}=\frac{\phi\begin{bmatrix}\left[2-\phi+\psi_{2,\{1\}}\left(2-\psi_{2,\{1\}}+\psi_{3,\{1,1\}}\left(2-\psi_{3,\{1,1\}}\right)\right)\right]\\ -\left(1-\psi_{2,\{1\}}\right)\lambda\psi_{2,\{1\}}-\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\lambda\psi_{3,\{1,1\}}\end{bmatrix}-\left(1-\phi\right)\lambda\phi}{1+\phi\left(1+\psi_{2,\{1\}}\right)}$$

Assuming a prior $\phi=\dfrac{1}{2}$:

$$Q^{La,Updating}\left(3,\phi,h=\{\varnothing\}\right)=\max\left\{\frac{293}{408}-\frac{89}{408}\lambda;\frac{586}{816}-\frac{112}{816}\lambda;\frac{235}{288}-\frac{67}{288}\lambda;\frac{217}{264}-\frac{61}{264}\lambda\right\}$$

In the second stage of the game ($N=2$), the agent will keep playing on the risky arm following a history $\{h_1\}=\{1\}$. In case of a history $\{h_1\}=\{0\}$, the player will experiment on the risky arm at $n=2$, if:

$$q\le Q^{La,Updating}\left(n=2,\phi,h_1=0\right)$$

The dynamic allocation index under the strategy $k_1=0$ is calculated as follows. The combined expected payoff in the two remaining periods of experimentation is equal to $\psi_{2,\{0\}}\begin{bmatrix}2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\\-\left(1-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{3,\{0,1\}}\right)\end{bmatrix}+\left(1-\psi_{2,\{0\}}\right)\begin{bmatrix}-\lambda\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)\\-\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{3,\{0,1\}}\end{bmatrix}$ since he will play risky in both periods. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $2q$.

The dynamic allocation index under the strategy $k_1 = 1$ is calculated as follows. The combined expected payoff in the remaining periods ($N = 2$ and $N = 3$) is equal to $\psi_{2,\{0\}}\begin{bmatrix} 2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)- \\ -\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{3,\{0,1\}} \end{bmatrix}+\left(1-\psi_{2,\{0\}}\right)\left(-\lambda\psi_{2,\{0\}}+q\right)$, since he will play safe in case of a failure in period $n = 2$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $2q$.

Hence the agent will not experiment at $n = 2$, if:

$$q \le Q^{La,Updating}\left(n = 2, \phi, h_1 = 0\right)$$

where the dynamic allocation index takes the following form:

$$Q^{La,Updating}(2,\phi,h_1 = 0) = \max_{k_1=\{0,1\}}\left\{\tilde{q}\Big|_{k_1=0};\tilde{q}\Big|_{k_1=1}\right\}$$

and

$$\tilde{q}^{La}\Big|_{k_1=0} = \frac{\psi_{2,\{0\}}\begin{bmatrix} 2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right) \\ -\left(1-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{3,\{0,1\}}\right) \end{bmatrix}+\left(1-\psi_{2,\{0\}}\right)\begin{bmatrix} -\lambda\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right) \\ -\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{3,\{0,1\}} \end{bmatrix}}{2}$$

$$\tilde{q}\Big|_{k_1=1} = \frac{\psi_{2,\{0\}}\begin{bmatrix} 2-\psi_{2,\{0\}}+\psi_{3,\{0,1\}}\left(2-\psi_{3,\{0,1\}}\right)- \\ -\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{3,\{0,1\}} \end{bmatrix}-\left(1-\psi_{2,\{0\}}\right)\lambda\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}$$

Assuming a prior $\phi = \frac{1}{2}$, $Q^{La,Updating}\left(2,\phi,h_1 = \{0\}\right) = \max\left\{\frac{47}{72}-\frac{17}{72}\lambda;\frac{2}{3}-\frac{11}{48}\lambda\right\}$

In the last stage of the game, following history $\{h_1,h_2\} = \{1,1\}$ or $\{h_1,h_2\} = \{0,1\}$, the agent will keep playing on the risky arm. In case of a history $\{h_1,h_2\} = \{0,0\}$, the agent will play risky if:

$$q \le Q^{La}\left(n = 1,\phi,h_1 = h_2 = 0\right)$$

where the dynamic allocation index simplifies to:

$$Q^{La}\left(n = 1,\phi,h_1 = h_2 = 0\right) = \psi_{3,\{0,0\}}\left(2-\psi_{3,\{0,0\}}\right)-\left(1-\psi_{3,\{0,0\}}\right)\left(\lambda\psi_{3,\{0,0\}}\right)$$

Assuming a prior $\phi = \frac{1}{2}$, $Q^{La}\left(n = 1,\phi,h_1 = h_2 = 0\right) = \frac{1}{16}(7-3\lambda)$, where $\bar{x} = \psi_{3,\{0,0\}} = \frac{1}{4}$.

3.3 *Risk Aversion and Loss Aversion*

In the first stage of the game, the individual will choose the strategy as a pair of $\{k_1, k_2\}$ in order to maximize the index attached to the risky arm (representing the expected payoff from experimenting). We denote with $Q^{Ra,La}$ the dynamic allocation index of a risk and loss averse player. Analytically, the experimenter will choose:

$$Q^{Ra,La}\left(3,\phi,h=\{\varnothing\}\right)=\max_{\{k_1,k_2\}}\left\{\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}},\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}},\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}\right\}$$

The individual will decide to experiment if and only if:

$$q\le Q^{Ra,La}\left(3,\phi,h=\{\varnothing\}\right)$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\left\{\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]\right\}(2-r-\phi)-$$

$$-\left\{\phi\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right]+(1-\phi)\begin{bmatrix}2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)+\\+\left(1-\psi_{2,\{0\}}\right)\left(1-\psi_{3,\{0,1\}}\right)\end{bmatrix}\right\}(\lambda\phi)$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=0}}=$$

$$\frac{\left\{\begin{matrix}\phi\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]+\\+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)+\left(1-\psi_{2,\{0\}}\right)\psi_{3,\{0,0\}}\right]\end{matrix}\right\}(2-r-\phi)-\left\{\begin{matrix}\phi\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right]\\+(1-\phi)\begin{bmatrix}2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)+\\\left(1-\psi_{2,\{0\}}\right)\left(1-\psi_{3,\{0,1\}}\right)\end{bmatrix}\end{matrix}\right\}(\lambda\phi)}{3}$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=1\}$ is calculated as follows. Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right](2-\phi)+\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right](-\lambda\phi)\right]+$$

$$+(1-\phi)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)(2-\phi)-\left[2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\right]\lambda\phi+\left(1-\psi_{2,\{0\}}\right)q\right]$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=1}}=\frac{\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\left(2-r-\phi\right)+\begin{bmatrix}\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\\+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\end{bmatrix}\left(-\lambda\phi\right)\right]+ \atop +\left(1-\phi\right)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\left(2-r-\phi\right)-\left[2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\right]\lambda\phi\right]}{2+\phi+\left(1-\phi\right)\psi_{2,\{0\}}}$$

The dynamic allocation index associated with strategy $\{k_1=0;k_2=2\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\left(2-r-\phi\right)+\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right]\left(-\lambda\phi\right)\right]+$$
$$+\left(1-\phi\right)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\left(2-r-\phi\right)-\left[2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\right]\lambda\phi+\left(1-\psi_{2,\{0\}}\right)q\right]$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=0\\k_2=2}}=\frac{\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\left(2-r-\phi\right)+\begin{bmatrix}\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\\+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\end{bmatrix}\left(-\lambda\phi\right)\right]+ \atop +\left(1-\phi\right)\left[\psi_{2,\{0\}}\left(1+\psi_{3,\{0,1\}}\right)\left(2-r-\phi\right)-\left[2-\psi_{2,\{0\}}+\psi_{2,\{0\}}\left(1-\psi_{3,\{0,1\}}\right)\right]\lambda\phi\right]}{2+\phi+\left(1-\phi\right)\psi_{2,\{0\}}}$$

The dynamic allocation index associated with strategy $\{k_1=1;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\left[\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\left(2-r-\phi\right)+\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right]\left(-\lambda\phi\right)\right]$$
$$+\left(1-\phi\right)\left(-\lambda\phi+2q\right)$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=1\\k_2=0}}=\frac{\phi\begin{Bmatrix}\left[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\right]\left(2-r-\phi\right)+\\\left[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\right]\left(-\lambda\phi\right)\end{Bmatrix}-\left(1-\phi\right)\lambda\phi}{1+2\phi}$$

The dynamic allocation index associated with strategy $\{k_1=1;k_2=1\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\Big[\Big[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\Big](2-r-\phi)+\Big[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\Big](-\lambda\phi)\Big]$$
$$+(1-\phi)(-\lambda\phi+2q)$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=1\\k_2=1}}=\frac{\phi\left\{\begin{aligned}&\Big[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)+\left(1-\psi_{2,\{1\}}\right)\psi_{3,\{1,0\}}\Big](2-r-\phi)+\\&\Big[\left(1-\psi_{2,\{1\}}\right)\left(2-\psi_{3,\{1,0\}}\right)+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)\Big](-\lambda\phi)\end{aligned}\right\}-(1-\phi)\lambda\phi}{1+2\phi}$$

The dynamic allocation index associated with strategy $\{k_1=2;k_2=0\}$ is calculated as follows.

Given the prior ϕ, the player's expected payoff over the time horizon of three periods will be

$$\phi\Big[\Big[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)\Big](2-r-\phi)+\left(1-\psi_{2,\{1\}}\right)[-\lambda\phi+q]+\psi_{2,\{1\}}\left(1-\psi_{3,\{1,1\}}\right)(-\lambda\phi)\Big]+(1-\phi)(-\lambda\phi+2q)$$

This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $3q$.

Hence:

$$\tilde{q}\Big|_{\substack{k_1=2\\k_2=0}}=\frac{\phi\left\{\Big[1+\psi_{2,\{1\}}\left(1+\psi_{3,\{1,1\}}\right)\Big](2-r-\phi)-\left(1-\psi_{3,\{1,1\}}\right)\lambda\phi\right\}-(1-\phi)\lambda\phi}{1+\phi\left(1+\psi_{2,\{1\}}\right)}$$

In case the prior $\phi=\frac{1}{2}$,

$$Q^{La}\left(3,\phi,h=\{\varnothing\}\right)=\max\left\{\frac{39}{44}-\frac{13}{22}r-\frac{15}{88}\lambda;\frac{7}{8}-\frac{7}{12}r-\frac{5}{16}\lambda;\frac{51}{64}-\frac{51}{96}r-\frac{15}{64}\lambda;\frac{1}{18}-\frac{1}{9}r-\frac{17}{72}\lambda\right\}$$

associated with strategy $\{k_1=1,k_2=0\}$.

In the second stage of the game ($N=2$), the agent will keep playing on the risky arm following a history $\{h_1\}=\{1\}$. In case of a history $\{h_1\}=\{0\}$, the player will experiment on the risky arm at $n=2$, if:

$$q\leq Q^{Ra,La}\left(n=2,\phi,h_1=0\right)$$

The dynamic allocation index under the strategy $k_1=0$ is calculated as follows. The combined expected payoff in the two remaining periods of experimentation is equal to

$\psi_{2,\{0\}}\begin{bmatrix}\left(1+\psi_{3,\{0,1\}}\right)\left(2-r-\psi_{2,\{0\}}\right)- \\ -\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{2,\{0\}}\end{bmatrix}+\left(1-\psi_{2,\{0\}}\right)\begin{bmatrix}-\left(2-\psi_{3,\{0,1\}}\right)\lambda\psi_{2,\{0\}}+ \\ +\psi_{3,\{0,1\}}\left(2-r-\psi_{2,\{0\}}\right)\end{bmatrix}$ since he will play risky in both periods. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $2q$.

The dynamic allocation index under the strategy $k_1=1$ is calculated as follows. The combined expected payoff in the remaining periods ($N=2$ and $N=3$) is equal to $\psi_{2,\{0\}}\left[\left(1+\psi_{3,\{0,1\}}\right)\left(2-r-\psi_{2,\{0\}}\right)-\left(1-\psi_{3,\{0,1\}}\right)\lambda\psi_{2,\{0\}}\right]+\left(1-\psi_{2,\{0\}}\right)\left(-\lambda\psi_{2,\{0\}}+q\right)$, since he will play safe in case of a failure in period $n=2$. This payoff should be compared with the alternative payoff, which is to play safe in both periods, which yields $2q$.

Hence the agent will not experiment at $n=2$, if:

$$q\leq Q^{Ra,La}\left(n=2,\phi,h_1=0\right)$$

where the dynamic allocation index takes the following form:

$$Q^{Ra,La}(2,\phi,h_1=0)=\max_{k_1=\{0,1\}}\left\{\tilde{q}\Big|_{k_1=0};\tilde{q}\Big|_{k_1=1}\right\}$$

and

$$\tilde{q}\Big|_{k_1=0}=\frac{1}{2}\begin{bmatrix}\psi_{2,\{0\}}\left[\left(1+\psi_{3,\{0,1\}}\right)\left(2-r-\psi_{2,\{0\}}\right)-\left(1-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)\right]+ \\ +\left(1-\psi_{2,\{0\}}\right)\left[-\left(2-\psi_{3,\{0,1\}}\right)\left(\lambda\psi_{2,\{0\}}\right)+\psi_{3,\{0,1\}}\left(2-r-\psi_{2,\{0\}}\right)\right]\end{bmatrix}$$

$$\tilde{q}\Big|_{k_1=1}=\frac{\psi_{2,\{0\}}}{1+\psi_{2,\{0\}}}\left[\left(2-r-\psi_{2,\{0\}}\right)\left(1+\psi_{3,\{0,1\}}\right)-\lambda\left(1-\psi_{2,\{0\}}\psi_{3,\{0,1\}}\right)\right]$$

Assuming a prior $\phi=\frac{1}{2}$, $Q^{Ra,La}\left(2,\phi=\frac{1}{2},h_1=0\right)=\max\left\{\frac{25}{36}-\frac{5}{12}r-\frac{7}{36}\lambda;\frac{5}{8}-\frac{3}{8}r-\frac{5}{24}\lambda\right\}$

In the last stage of the game, following history $\{h_1,h_2\}=\{1,1\}$ or $\{h_1,h_2\}=\{0,1\}$, the agent will keep playing on the risky arm. In case of a history $\{h_1,h_2\}=\{0,0\}$, the agent will play risky if:

$$q\leq Q^{Ra,La}\left(n=1,\phi,h_1=h_2=0\right)$$

where the dynamic allocation index simplifies to:

$$Q^{Ra,La}\left(n=1,\phi,h_1=h_2=0\right)=\psi_{3,\{0,0\}}\left(2-r-\psi_{3,\{0,0\}}\right)-\left(1-\psi_{3,\{0,0\}}\right)\left(\lambda\psi_{3,\{0,0\}}\right)\triangleright$$

In case the prior $\phi=\frac{1}{2}$, $Q^{Ra,La}\left(n=1,\phi,h_1=h_2=0\right)=\frac{1}{16}\left(7-r-3\lambda\right)$, where $\bar{x}=\psi_{3,\{0,0\}}=\frac{1}{4}$.

Chapter 3: Experimentation and Loss Aversion – An Experiment

1. Introduction

Chapter 2 investigated the optimal experimentation choice of a risk-averse and a loss-averse player in the setting of a one-armed bandit problem. In Chapter 2 we showed that in the presence of both risk aversion and loss aversion new qualitative features of the prediction emerge. Chapter 3 addresses the issue of optimal experimentation in a traditional one-armed bandit problem adopting an experimental perspective. We investigate in which direction the presence of risk aversion and loss aversion affects the experimentation choices in experimental settings.

In total 254 subjects – first year undergraduate students in Microeconomics - took part in the experiment. The experimental study involved for each subject the elicitation of individual preferences (loss aversion and risk aversion) and the attitude to participate to experiments. In the first part of the study we measure loss aversion at the individual level in a within-subject design. We adopt an individual measure of loss aversion in a risky choice task, adapting a lottery choice similar to Fehr and Goette (2002). We complement this measure with a risk aversion one, in order to disentangle possible interlinked effects on the experimentation choices of the subjects.

In the second part we investigate the experimentation decision of each subject in a one armed-bandit setting. The experimental design is based on the three-stage game example illustrated in Section 3 of Chapter 2 – where the optimal experimentation strategy is fully characterized for a risk-neutral, loss-neutral player and for one with positive levels of risk aversion and loss aversion. The experiment is designed to replicate a traditional one-armed bandit problem, where the player faces the decision between an unknown arm and a safe one in each of the three rounds of the game. The player should decide in each round whether to experiment on the risky arm or to opt out of the game, winning a sure prize on the safe arm. The experiment involves an optimal stopping problem, i.e. each subject in the experiment has to decide how many times to experiment with the unknown arm before switching to the safe one, so that once he has switched the agent will not return to the unknown arm.

The main experimental finding is that there is evidence of both over-experimentation and under-experimentation, and this is true for subjects of all level of risk aversion and loss aversion. At some stages in the process subjects of all types display over-experimentation, while at other stages there is a systematic bias towards under-experimentation. The experimental findings provide support to the non-equivalence of risk aversion and loss aversion presented in Chapter 2.

Chapter 3 is organized as follows. Section 2 discussed the literature. Section 3 focuses on the design of the experiment, outlines the procedure for the elicitation of individual loss aversion and risk aversion and explains the set-up of the three-stage one-armed bandit problem. Section 4 discusses the main results from the experiment, concerning the distribution of preferences in the population of subjects and the experimental choices of the subjects as a function of loss aversion and risk aversion. Section 5 concludes.

2. Related literature

The present work relates to two different strands of literature: experimental literature on bandit problems and the role of loss aversion in decision-making.

Regarding the first strand, a number of experiments have been conducted on the experimentation strategies of agents. The main evidence collected from field and laboratory studies on search problems show that agents may not collect enough information during the experimentation, thereby stopping too soon and not searching enough. Banks, Olson and Porter (1997) run a laboratory experiment to determine whether people act optimally. They design an experiment with two separate treatments: in the first one they study the myopic behaviour of agents, who chooses the arm with the highest payoff. In the second treatment it is sometimes optimal to choose the arm with the lowest expected payoff. The main conclusion drawn from the experiment is that people do not engage in enough experimentation, since they observe a higher rate of adoption of myopic behaviour in the myopic behaviour treatment. Cox and Oxaca (1989, 1990, 1992, 1996) focused on an application of optimal search in the labour market. Subjects were told that they had a 20-period horizon to search for a job. In each period they receive a wage offer, drawn from a known uniform distribution on the interval of 1 to 10. If a subject accepts the offer and stops searching, he would be paid the agreed wage for all the remaining periods of the experiment. Cox and Oxaca show that people tend not to search enough and may accept a wage lower than the optimal reservation value, arguing that risk aversion may be the driving factor for a suboptimal search. The results are robust to alternative specifications of the experiment, such as the length of the search period, the variance of the wage offer distribution and the ability to recall past offers to reduce the risk of continued search. Schotter and Braunstein (1981) confirm these results, showing that even if subjects state reservation values close to the optimal ones, they search for a lower number of periods than would be optimal. Anderson (2001) investigates the role of behavioural biases in experimentation problems, focusing on the role of hyperbolic discounting and risk aversion and showing that these behavioural biases may lead to suboptimal experimentation and welfare loss.[42]

The second strand of literature relevant to the present work focuses on the role of loss aversion – the psychological propensity that makes losses loom larger than equal-sized gains relative to a reference point – in experimental setting. In two seminal papers Tversky and Kahneman (1979, 1991) argued that loss aversion can occur both in riskless and in risky choices. An example of loss aversion in a riskless choice is the well known 'endowment effect' – the observation that experimental subjects who are randomly assigned with a commodity ask for a selling price substantially exceeding the buying price of subjects who have only the possibility of buying the same good (among others, see Kahneman, Knetsch and Thaler, 1990). An example of loss aversion in risky choices is the observation that people reject small-scale gambles with a positive expected value, which involve losses (Rabin, 2000; Fehr and Goette, 2007; Tom, Fox, Trepel and Poldrack, 2007).[43]

Loss aversion has been invoked to explain a number of naturally occurring phenomena that cannot be explained under the assumption of reference-point independence. Among others, application of loss aversion has been successful in explaining behaviour in financial markets (Benartzi and Thaler, 1995; Gneezy and Potters, 1997; Odean, 1998; Haigh and List, 2005); sellers' behaviour in housing markets (Genesove and Mayer, 2001); consumption behaviour (Bowman, Minehart and Rabin, 1999; Chua and Camerer, 2004); labour supply (Camerer, Babcock, Loewenstein and Thaler, 1997; Goette, Huffman and Fehr, 2004; Fehr and Goette, 2007); trade policy (Tovar, 2006); and the importance of defaults and the status-quo bias in decision making (Samuelson and Zeckhauser, 1988; Johnson and Goldstein, 2003).

---

[42] A different explanation for inefficient search is constituted by unobservable search cost. See among others Pratt, Wise and Zeckerhaus (1979), in which they showed that the observed variance of retail prices (for 39 goods randomly selected on the yellow pages in Boston) could not be sustained by optimal search and concluded that consumers do not search optimally.

[43] Camerer (2004) provides an overview of the field evidence, and Starmer (2000) a survey of theoretical explanations. See Sugden (2003), Schmidt, Starmer and Sugden (2005), and Köszegi and Rabin (2006) for recent theoretical frameworks of reference-dependent preferences.

3.1 Individual preferences: loss aversion and risk aversion

We aim to measure loss aversion in risky choices. We adapt a simple lottery choice task from Fehr and Goette (2002).[45] In this choice task individuals are asked to choose for each of six lotteries whether they want to accept (in other words, play it) or reject it (and receive nothing). In each lottery the winning price is fixed at 6 and only the losing price is varied (between 2 and 7). At the end of the experiment we randomly selected one lottery for pay (Cubitt, Starmer and Sugden, 1998). Table I reproduces the decision sheet of the lottery choice task as presented to subjects.

Lottery	Accept	Reject
#1. If the coin turns up heads, then you lose €2; if the coin turns up tails, you win €6		
#2. If the coin turns up heads, then you lose €3; if the coin turns up tails, you win €6		
#3. If the coin turns up heads, then you lose €4; if the coin turns up tails, you win €6		
#4. If the coin turns up heads, then you lose €5; if the coin turns up tails, you win €6		
#5. If the coin turns up heads, then you lose €6; if the coin turns up tails, you win €6		
#6. If the coin turns up heads, then you lose €7; if the coin turns up tails, you win €6		

Table I

The lottery choice task: Loss Aversion

The lottery choice task presented here measures loss aversion rather than risk aversion, following Rabin (2000), Rabin and Thaler (2001), Wakker (2005), Köbberling and Wakker (2005) and Fehr and Goette (2007). Rabin (2000) provides evidence that risk aversion cannot plausibly explain choice behaviour in small-stake risky lotteries. We observe the rejection of small-stake risky lotteries that offer positive expected value and risk aversion (modeled as a concave utility of wealth function). In such small-stakes lotteries risk aversion cannot be invoked to explain the

[45] Gachter, Johnson and Hermann (2007) provide additional support for the idea that loss aversion can be measured by the simple lottery choice task from Fehr and Goette. They run two separate sets of experiments, a risky small stake lottery and a riskless lottery for the elicitation of the willingness-to-pay and willingness-to-accept. They provide evidence for a positive strong correlation between the measure of loss aversion in the risky lottery and the measure of loss aversion implied in a riskless choice.

3. Experimental Design

Subjects in the study were 254 undergraduate students enrolled in an Introductory Course of Microeconomics at the University of Modena and Reggio Emilia in Italy. All participants were Italian speaking.

The students answered the questions during a lab. Each respondent was asked to always answer alone, undisturbed by other students. The interviews, including the experiments, lasted about one hour. While familiar with the experimental protocol, all students were naïve about the experimental hypotheses. All the students were enrolled in a first-year undergraduate program with a major in economics and received no training on either decision making under uncertainty or behavioural economics during the microeconomics lectures[44].

We conducted an experimental study with the aim to insulate the role of loss aversion in the experimental decisions. The experimental study involved the elicitation of individual preferences (loss aversion and risk aversion) and the experimental choices of the subjects. The experiment set-up is divided into two distinct parts. In the first one we measure loss aversion at the *individual* level in a within-subject design. We complement our individual measure of loss aversion in risky choice with one from a standard measure of risk aversion, to analyze how these measures are correlated within subjects. In the second part of the study all subjects who participated in this within-subject design study also took part in a three-stage experimentation game replicating the setting of a one-armed bandit problem. The goal of the second part of the study is to determine the experimentation strategy of each individual as a function of the loss aversion attitude isolated in the first part of the experiment, controlling for risk aversion. We now describe the experimental design in detail.

[44] Subjects in the experiment did not receive any previous training in Microeconomics on the topic of decision-making under uncertainty and behavioural economics. The experiment took place in the winter semester of their first year of an undergraduate program in Economics. The Introductory course of Microeconomics constituted the only economic background of the subjects at the time of the experiment.

3.1 Individual preferences: loss aversion and risk aversion

We aim to measure loss aversion in risky choices. We adapt a simple lottery choice task from Fehr and Goette (2002).[45] In this choice task individuals are asked to choose for each of six lotteries whether they want to accept (in other words, play it) or reject it (and receive nothing). In each lottery the winning price is fixed at 6 and only the losing price is varied (between 2 and 7). At the end of the experiment we randomly selected one lottery for pay (Cubitt, Starmer and Sugden, 1998). Table I reproduces the decision sheet of the lottery choice task as presented to subjects.

Lottery	Accept	Reject
#1. If the coin turns up heads, then you lose €2; if the coin turns up tails, you win €6		
#2. If the coin turns up heads, then you lose €3; if the coin turns up tails, you win €6		
#3. If the coin turns up heads, then you lose €4; if the coin turns up tails, you win €6		
#4. If the coin turns up heads, then you lose €5; if the coin turns up tails, you win €6		
#5. If the coin turns up heads, then you lose €6; if the coin turns up tails, you win €6		
#6. If the coin turns up heads, then you lose €7; if the coin turns up tails, you win €6		

Table I

The lottery choice task: Loss Aversion

The lottery choice task presented here measures loss aversion rather than risk aversion, following Rabin (2000), Rabin and Thaler (2001), Wakker (2005), Köbberling and Wakker (2005) and Fehr and Goette (2007). Rabin (2000) provides evidence that risk aversion cannot plausibly explain choice behaviour in small-stake risky lotteries. We observe the rejection of small-stake risky lotteries that offer positive expected value and risk aversion (modeled as a concave utility of wealth function). In such small-stakes lotteries risk aversion cannot be invoked to explain the

[45] Gachter, Johnson and Hermann (2007) provide additional support for the idea that loss aversion can be measured by the simple lottery choice task from Fehr and Goette. They run two separate sets of experiments, a risky small stake lottery and a riskless lottery for the elicitation of the willingness-to-pay and willingness-to-accept. They provide evidence for a positive strong correlation between the measure of loss aversion in the risky lottery and the measure of loss aversion implied in a riskless choice.

pattern of rejection, since it would imply absurdly high degrees of risk aversion in high-stake gambles, as argued by Rabin (2000). He additionally sustains that under an expected utility framework people in such small-stake risky lotteries should be risk neutral. Hence, on the basis of this argument, in the risky lottery choice task in Table 1, subjects should therefore accept lotteries #1 to #5, which all present a non-negative expected value. The observation of rejections of low-stake gambles with a positive expected value might imply the presence of loss aversion rather than risk aversion.

We can determine loss aversion in the risky choice task by applying cumulative prospect theory (Tversky and Kahneman, 1992). A decision maker will be indifferent between accepting and rejecting the lottery if:

$$p_G(0.5)v(G) = p_L(0.5)\lambda v(L)$$

where G indicates the gain and L the loss in the lottery; $v(x)$ is the utility of the lottery outcome $x \in \{G, L\}$; λ denotes the loss aversion coefficient in the risky choice task; $p_G(0.5)$and $p_L(0.5)$denote the probability weights respectively in the case of positive and negative outcomes associated with a lottery that offers an equal chance of gaining G or losing L. According to the probability weighting function proposed by Prelec (1998), we can assume that $p_G(0.5) = p_L(0.5)$. This simplifies the estimation of the parameter of loss aversion, which reduces to the ratio of the utility associated with the gain relative to the disutility associated with a loss. In analytical terms, $\lambda = \frac{v(G)}{v(L)}$ defines the implied loss aversion of the individual displayed in the lottery choice task. We additionally assume the linearity of the preference over the space of gains and losses, according to which $v(x) = \begin{cases} x & if \quad x \geq 0 \\ -\lambda x\, if & x < 0 \end{cases}$.[46] This assumption provides very simple measure of loss aversion: $\lambda = \frac{G}{L}$.[47]

We measure risk aversion in a risky choice lottery. In this choice task individuals decide for each of seven lotteries whether they want to accept (that is, play it) or reject it. In the alternative, they can accept a sure amount of money of euro 1. In each lottery, the expected value is fixed at euro 5, while the winning prize and the probability of winning the prize change in each lottery with an associated different variance. We elicit the risk aversion coefficient implied in the decision to accept the risky lottery as the difference between the expected value of the lottery and the sure cost

[46] The piecewise linearity of the prospect theory utility function is seen frequently in many economic applications, especially in the case of small stake lotteries. See among other Bernartzi and Thaler (1995).
[47] This formulation is entirely consistent with the analytical model discussed in Chapter 2.

of the lottery[48]. The lotteries are ordered such that the acceptance of the lottery denotes a decreasing level of the parameter of risk aversion in Table II.

Sure amount	Lottery
euro 1.00 for sure	#1. 80% chance of euro 6.25 and 20% chance of euro 0
euro 1.00 for sure	#2. 50% chance of euro 10 and 50% chance of euro 0
euro 1.00 for sure	#3. 40% chance of $12.50 and 60% chance of euro 0
euro 1.00 for sure	#4. 20% chance of euro 25 and 80% chance of euro 0
euro 1.00 for sure	#5. 10% chance of euro 50 and 90% chance of euro 0
euro 1.00 for sure	#6. 5% chance of euro 100 and 95% chance of euro 0
euro 1.00 for sure	#7. 1% chance of euro 500 and 99% chance of euro 0

Table II

The lottery choice task: Risk Aversion

3.2 One-armed bandit with three stages of experimentation

In the second part of the experiment we investigate the experimental decisions of the subjects in a bandit setting. The experiment was designed to simulate the experimentation decision over time in the setting of a bandit problem with a finite horizon time. The experimental design replicates exactly the three-stage game illustrated in Section 3 of Chapter 2.

Subjects were asked to play on a slot machine. They were told that the slot machine pays out 10 euro with some probability. They were told they did not know the probability for sure, but it was reasonable to think that the slot machine gave a 50-50 chance to win. Subjects were told that they could play on the slot machine, but that they should pay a ticket at a cost of 3.5 euro to play on the slot machine. Subjects responded yes or no when asked whether they would play on the slot

[48] The parameter of risk aversion is estimated according to the formula $E(x)-\frac{1}{2}rVar(x)$, where $E(x)-\frac{1}{2}rVar(x)=c$ denotes the expected utility of the lottery, $Var(x)$ the variance of the lottery, c the utility of certainty equivalent and r is the measure of risk aversion.

pattern of rejection, since it would imply absurdly high degrees of risk aversion in high-stake gambles, as argued by Rabin (2000). He additionally sustains that under an expected utility framework people in such small-stake risky lotteries should be risk neutral. Hence, on the basis of this argument, in the risky lottery choice task in Table 1, subjects should therefore accept lotteries #1 to #5, which all present a non-negative expected value. The observation of rejections of low-stake gambles with a positive expected value might imply the presence of loss aversion rather than risk aversion.

We can determine loss aversion in the risky choice task by applying cumulative prospect theory (Tversky and Kahneman, 1992). A decision maker will be indifferent between accepting and rejecting the lottery if:

$$p_G(0.5)v(G) = p_L(0.5)\lambda v(L)$$

where G indicates the gain and L the loss in the lottery; $v(x)$ is the utility of the lottery outcome $x \in \{G, L\}$; λ denotes the loss aversion coefficient in the risky choice task; $p_G(0.5)$and $p_L(0.5)$denote the probability weights respectively in the case of positive and negative outcomes associated with a lottery that offers an equal chance of gaining G or losing L. According to the probability weighting function proposed by Prelec (1998), we can assume that $p_G(0.5) = p_L(0.5)$. This simplifies the estimation of the parameter of loss aversion, which reduces to the ratio of the utility associated with the gain relative to the disutility associated with a loss. In analytical terms, $\lambda = \frac{v(G)}{v(L)}$ defines the implied loss aversion of the individual displayed in the lottery choice task. We additionally assume the linearity of the preference over the space of gains and losses, according to which $v(x) = \begin{cases} x & if \quad x \geq 0 \\ -\lambda x \, if & x < 0 \end{cases}$.[46] This assumption provides very simple measure of loss aversion: $\lambda = \frac{G}{L}$.[47]

We measure risk aversion in a risky choice lottery. In this choice task individuals decide for each of seven lotteries whether they want to accept (that is, play it) or reject it. In the alternative, they can accept a sure amount of money of euro 1. In each lottery, the expected value is fixed at euro 5, while the winning prize and the probability of winning the prize change in each lottery with an associated different variance. We elicit the risk aversion coefficient implied in the decision to accept the risky lottery as the difference between the expected value of the lottery and the sure cost

[46] The piecewise linearity of the prospect theory utility function is seen frequently in many economic applications, especially in the case of small stake lotteries. See among other Bernartzi and Thaler (1995).
[47] This formulation is entirely consistent with the analytical model discussed in Chapter 2.

of the lottery[48]. The lotteries are ordered such that the acceptance of the lottery denotes a decreasing level of the parameter of risk aversion in Table II.

Sure amount	Lottery
euro 1.00 for sure	#1. 80% chance of euro 6.25 and 20% chance of euro 0
euro 1.00 for sure	#2. 50% chance of euro 10 and 50% chance of euro 0
euro 1.00 for sure	#3. 40% chance of $12.50 and 60% chance of euro 0
euro 1.00 for sure	#4. 20% chance of euro 25 and 80% chance of euro 0
euro 1.00 for sure	#5. 10% chance of euro 50 and 90% chance of euro 0
euro 1.00 for sure	#6. 5% chance of euro 100 and 95% chance of euro 0
euro 1.00 for sure	#7. 1% chance of euro 500 and 99% chance of euro 0

Table II

The lottery choice task: Risk Aversion

3.2 One-armed bandit with three stages of experimentation

In the second part of the experiment we investigate the experimental decisions of the subjects in a bandit setting. The experiment was designed to simulate the experimentation decision over time in the setting of a bandit problem with a finite horizon time. The experimental design replicates exactly the three-stage game illustrated in Section 3 of Chapter 2.

Subjects were asked to play on a slot machine. They were told that the slot machine pays out 10 euro with some probability. They were told they did not know the probability for sure, but it was reasonable to think that the slot machine gave a 50-50 chance to win. Subjects were told that they could play on the slot machine, but that they should pay a ticket at a cost of 3.5 euro to play on the slot machine. Subjects responded yes or no when asked whether they would play on the slot

[48] The parameter of risk aversion is estimated according to the formula $E(x)-\frac{1}{2}rVar(x)$, where $E(x)-\frac{1}{2}rVar(x)=c$ denotes the expected utility of the lottery, $Var(x)$ the variance of the lottery, c the utility of certainty equivalent and r is the measure of risk aversion.

machine. Additionally, subjects were asked to state the maximum ticket price at which they would agree to play on the slot machine. Responses were given as an integer from 2 up to 5.5 euro (with increments of 50 cents).

Each trial in the experiment consisted of deciding whether to pay the ticket to play on the slot machine or not. The cost of the ticket plays the role of a sure prize awarded on the safe arm in the one-armed bandit problem illustrated in Chapter 2. By deciding not to play along the slot machine, subjects declare to have a reservation value at least equal to the cost of the ticket. If subjects agree to play at the stated price of 3.5 euro, the reservation value of the player can be elicited in each period of the game through the additional enquiry as to the maximum ticket price at which subjects are willing to participate in the experimentation problem.

Subjects were told that they could play at most three times on the slot machine. Each time they played, they were told whether they had a failure or a success on the slot machine in the previous stage of the game. The experiment is designed to replicate some (not all) of the possible patterns associated with a three-stage game illustrated in Chapter 2. After the first stage of experimentation, if players decided to play on the slot machine in stage 1, subjects were told that a) a failure was scored in the first stage b) a history of one success in first stage and a failure in the second stage was recorded c) the (reverse) history of a failure in the first stage and a success in the second stage. The design of the game allows us to study the experimentation pattern for each individual in four possible scenarios: the first stage (when the history is the empty set) and the three scenarios a), b), and c) described above.

4. Results

We organize the presentation of the experimental results as follows. In the first section we analyze the distribution of preferences in the population of subjects who have taken part to the experiment. In the second section we present the results on the experimental choices of the subjects as a function of the loss aversion and risk aversion parameters.

4.1 Loss Aversion and Risk Aversion

We will first analyze the distribution of loss aversion elicited from the within- subject design in the risky choice task illustrated in Table I. The second step will be to estimate the individual-level of risk aversion from the within-subject design illustrated in Table II and its distribution in the machine. Additionally, subjects were asked to state the maximum ticket price at which they would agree to play on the slot machine. Responses were given as an integer from 2 up to 5.5 euro (with increments of 50 cents).

Each trial in the experiment consisted of deciding whether to pay the ticket to play on the slot machine or not. The cost of the ticket plays the role of a sure prize awarded on the safe arm in the one-armed bandit problem illustrated in Chapter 2. By deciding not to play along the slot machine, subjects declare to have a reservation value at least equal to the cost of the ticket. If subjects agree to play at the stated price of 3.5 euro, the reservation value of the player can be elicited in each period of the game through the additional enquiry as to the maximum ticket price at which subjects are willing to participate in the experimentation problem.

Subjects were told that they could play at most three times on the slot machine. Each time they played, they were told whether they had a failure or a success on the slot machine in the previous stage of the game. The experiment is designed to replicate some (not all) of the possible patterns associated with a three-stage game illustrated in Chapter 2. After the first stage of experimentation, if players decided to play on the slot machine in stage 1, subjects were told that a) a failure was scored in the first stage b) a history of one success in first stage and a failure in the second stage was recorded c) the (reverse) history of a failure in the first stage and a success in the second stage. The design of the game allows us to study the experimentation pattern for each individual in four possible scenarios: the first stage (when the history is the empty set) and the three scenarios a), b), and c) described above.

4. Results

We organize the presentation of the experimental results as follows. In the first section we analyze the distribution of preferences in the population of subjects who have taken part to the experiment. In the second section we present the results on the experimental choices of the subjects as a function of the loss aversion and risk aversion parameters.

4.1 Loss Aversion and Risk Aversion

We will first analyze the distribution of loss aversion elicited from the within- subject design in the risky choice task illustrated in Table I. The second step will be to estimate the individual-level of risk aversion from the within-subject design illustrated in Table II and its distribution in the

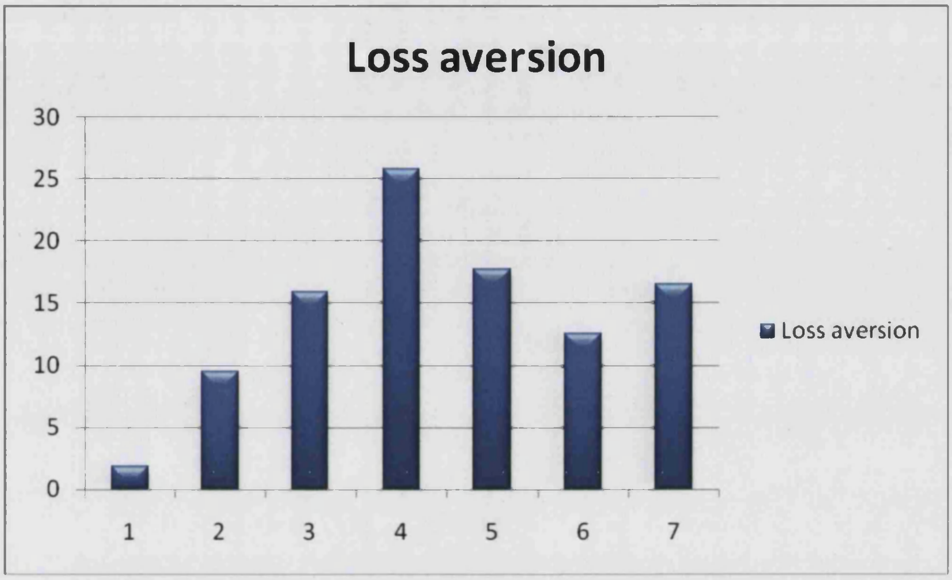


Figure I

Distribution of subjects on the basis of

loss aversion parameter λ in the lottery choice task

(ordered from HIGH to LOW level of loss aversion according to Table III)

According to Table III, 12.58 percent behaved like risk-neutral decision makers because they accepted all lotteries with a non-negative expected value and only rejected lottery #6, which has a negative expected value. Hence, their implied $\lambda = 1$. Slightly more than sixteen percent of our respondents also accepted lottery #6, which has a negative expected value, i.e., their $\lambda < 0.87$. Most participants rejected gambles with a positive expected value. Specifically, 70.86 percent of our respondents rejected lottery #5 or some lottery before #5(#1 to #4). A few respondents (1.84 percent) rejected all six lotteries; for these people $\lambda > 3$. The median respondent's cutoff lottery was #4: he or she accepted lotteries #1 to #4 and rejected lotteries #5 and #6, which implies $\lambda = 1.2$.

The second step is to look at the distribution of risk aversion in the population of subjects. In an analogous way to Table III – in Table IV we indicate the acceptance rates of subjects, which are ranked from the more risk averse subjects (associated with group 1 in Table IV– those subjects who reject all lotteries) to groups of subjects with a decreasing level of risk aversion.

population. Finally the joint distribution of loss aversion and risk aversion in the risky choice task is examined.

The first step is to look at loss aversion in risky choices. Table III records the results[49].

Acceptance behaviour (lottery choice)	Percentual acceptance rate	Implied acceptable loss (in euro)	Implied λ
1) Reject all lotteries	1.84	<2	>3
2) Accept lottery #1, reject lotteries #2 to #6	9.51	2	3
3) Accept lotteries #1 and #2, reject lotteries #3 to #6	15.95	3	2
4) Accept lotteries #1 to #3, reject lotteries #4 to #6	25.77	4	1.5
5) Accept lotteries #1 to #4, reject lotteries #5 to #6	17.79	5	1.2
6) Accept lotteries #1 to #5, reject lotteries #6	12.58	6	1
7) Accept all lotteries	16.56	≥7	≤0.87

Table III

Acceptance rates of the alternative lotteries in the lottery choice task and implied λ

Figure I illustrates the distribution of loss aversion in the population of subjects taking part to the experiment – ranked from the more loss averse (associated with group 1 in Table III, i.e. the subjects rejecting all lotteries in Table I, with a loss aversion parameter $\lambda > 3$) to the those subjects with a decreasing level of loss aversion (from group 2, with a loss aversion parameter $\lambda = 3$ up to group 7, with $\lambda < 1$).

[49] Booij and Van de Kuilen (2006) propose more recent estimates, using data from a large representative pool of subjects, suggesting a value of λ riskless = 2.47.

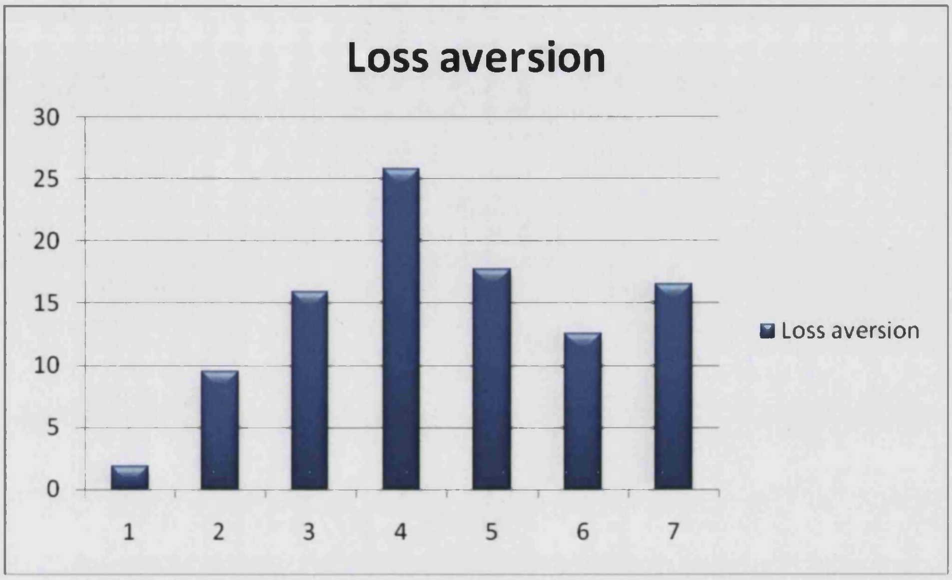


Figure I

Distribution of subjects on the basis of

loss aversion parameter λ in the lottery choice task

(ordered from HIGH to LOW level of loss aversion according to Table III)

According to Table III, 12.58 percent behaved like risk-neutral decision makers because they accepted all lotteries with a non-negative expected value and only rejected lottery #6, which has a negative expected value. Hence, their implied $\lambda = 1$. Slightly more than sixteen percent of our respondents also accepted lottery #6, which has a negative expected value, i.e., their $\lambda < 0.87$. Most participants rejected gambles with a positive expected value. Specifically, 70.86 percent of our respondents rejected lottery #5 or some lottery before #5(#1 to #4). A few respondents (1.84 percent) rejected all six lotteries; for these people $\lambda > 3$. The median respondent's cutoff lottery was #4: he or she accepted lotteries #1 to #4 and rejected lotteries #5 and #6, which implies $\lambda = 1.2$.

The second step is to look at the distribution of risk aversion in the population of subjects. In an analogous way to Table III – in Table IV we indicate the acceptance rates of subjects, which are ranked from the more risk averse subjects (associated with group 1 in Table IV– those subjects who reject all lotteries) to groups of subjects with a decreasing level of risk aversion.

Acceptance behaviour (lottery choice)	Percentual acceptance rate	Risk aversion parameter
1) Reject all lotteries	2.15	6.4
3) Accept lotteries #1, reject lotteries #2 to #7	8.45	0.32
3) Accept lotteries #1 to #2, reject lotteries #3 to #7	18.43	0.213
4) Accept lotteries #1 to #3, reject lotteries #4 to #7	22.76	0.08
5) Accept lotteries #1 to #4, reject lotteries #5 to #7	22.01	0.0355
6) Accept lotteries #1 to #5, reject lotteries #6 to #7	14.16	0.0168
7) Accept all lotteries #1 to #6, reject lotteries #7	12.02	0.0032

Table IV
Population distribution on the basis of risk aversion measure

According to Table IV, 12.02 percent behaved like risk-neutral decision makers because they accepted all lotteries with an increasing level of risk. A few respondents (2.15 percent) rejected all six lotteries; for these people the degree of risk aversion is the highest. The median respondent's cutoff lottery was #4: he or she accepted lotteries #1 to #3 and rejected lotteries #4 to #7. The population of subjects is distributed across all seven classes as illustrated in Figure II.

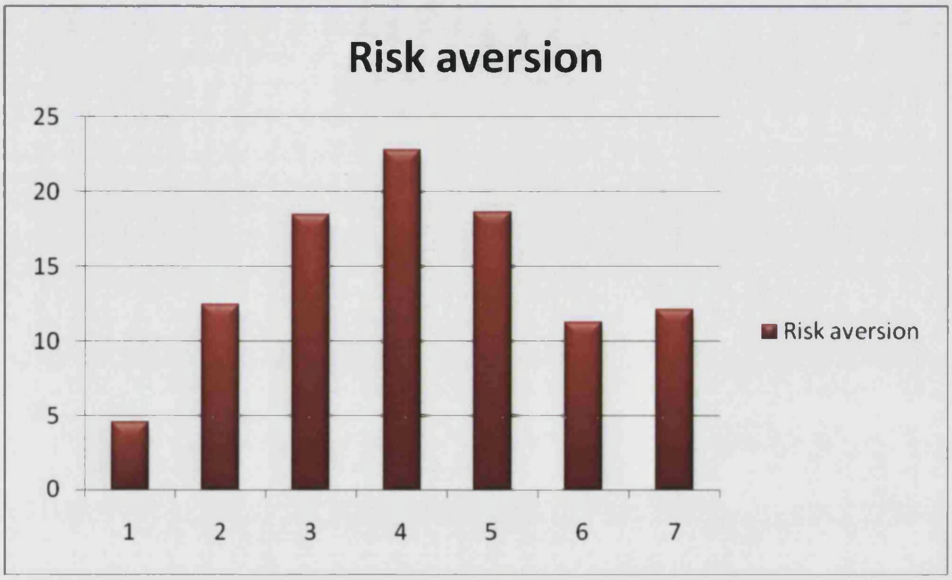


Figure II
Population distribution on the basis of risk aversion measure RA
(ordered from HIGH to LOW risk aversion)

Finally, we study the joint distribution of loss aversion and risk aversion in the population of subjects who participated to the experiment – in order to investigate how the dimensions of loss aversion and risk aversion are interlinked within subjects[50]. For the construction of the joint distribution, we aggregate the seven classes of loss aversion and risk aversion specified in Table III and IV- thereby reducing them from seven to four – labelled respectively neutral, low, medium and high levels of respectively loss aversion (in columns) and risk aversion (in rows).

Table Va) and Vb) illustrate the distribution of risk aversion and loss aversion across subjects respectively in terms of the number of subjects and the percentage of subjects interviewed in the experiment. We decided to eliminate from the analysis all those subjects who responded in an inconsistent manner to questions in Section 1 of the experiment, showing deviations from rational choices. For example, we do not include in the analysis those subjects willing to reject lottery #1 in Table I (i.e. the one imposing a loss of euro 2 in case the coin turns heads and a gain of euro 6 in case of tails), but willing to accept lottery #6 (i.e. the one imposing a loss of euro 7 in case the coin turns heads and a gain of euro 6 in case of tails). By excluding those subjects displaying inconsistent choices, the sample reduces to 203 subjects.

[50] We do not have any hypothesis regarding how loss aversion and risk aversion are jointly distributed across population.

Acceptance behaviour (lottery choice)	Percentual acceptance rate	Risk aversion parameter
1) Reject all lotteries	2.15	6.4
3) Accept lotteries #1, reject lotteries #2 to #7	8.45	0.32
3) Accept lotteries #1 to #2, reject lotteries #3 to #7	18.43	0.213
4) Accept lotteries #1 to #3, reject lotteries #4 to #7	22.76	0.08
5) Accept lotteries #1 to #4, reject lotteries #5 to #7	22.01	0.0355
6) Accept lotteries #1 to #5, reject lotteries #6 to #7	14.16	0.0168
7) Accept all lotteries #1 to #6, reject lotteries #7	12.02	0.0032

Table IV

Population distribution on the basis of risk aversion measure

According to Table IV, 12.02 percent behaved like risk-neutral decision makers because they accepted all lotteries with an increasing level of risk. A few respondents (2.15 percent) rejected all six lotteries; for these people the degree of risk aversion is the highest. The median respondent's cutoff lottery was #4: he or she accepted lotteries #1 to #3 and rejected lotteries #4 to #7. The population of subjects is distributed across all seven classes as illustrated in Figure II.

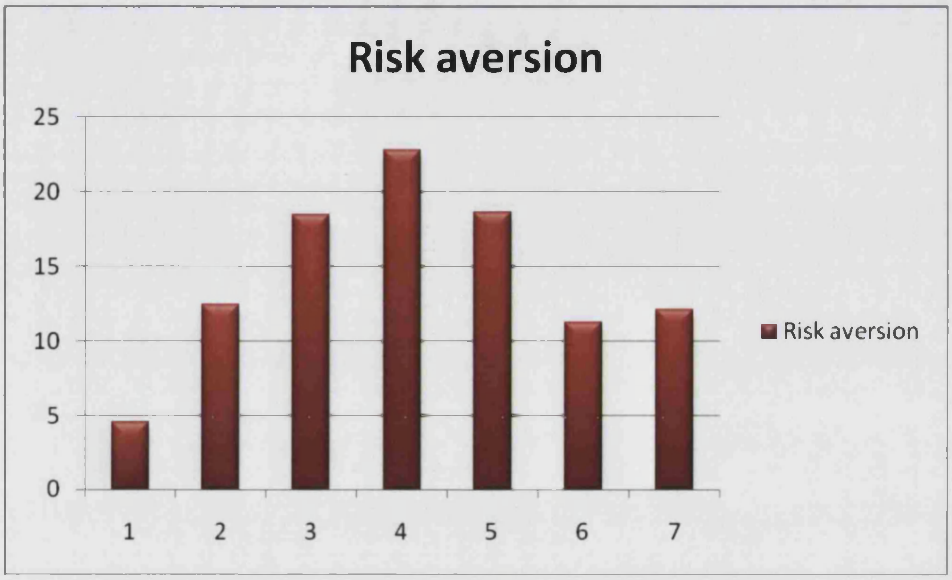


Figure II

Population distribution on the basis of risk aversion measure RA

(ordered from HIGH to LOW risk aversion)

Finally, we study the joint distribution of loss aversion and risk aversion in the population of subjects who participated to the experiment – in order to investigate how the dimensions of loss aversion and risk aversion are interlinked within subjects[50]. For the construction of the joint distribution, we aggregate the seven classes of loss aversion and risk aversion specified in Table III and IV- thereby reducing them from seven to four – labelled respectively neutral, low, medium and high levels of respectively loss aversion (in columns) and risk aversion (in rows).

Table Va) and Vb) illustrate the distribution of risk aversion and loss aversion across subjects respectively in terms of the number of subjects and the percentage of subjects interviewed in the experiment. We decided to eliminate from the analysis all those subjects who responded in an inconsistent manner to questions in Section 1 of the experiment, showing deviations from rational choices. For example, we do not include in the analysis those subjects willing to reject lottery #1 in Table I (i.e. the one imposing a loss of euro 2 in case the coin turns heads and a gain of euro 6 in case of tails), but willing to accept lottery #6 (i.e. the one imposing a loss of euro 7 in case the coin turns heads and a gain of euro 6 in case of tails). By excluding those subjects displaying inconsistent choices, the sample reduces to 203 subjects.

[50] We do not have any hypothesis regarding how loss aversion and risk aversion are jointly distributed across population.

Risk Aversion/Loss Aversion	Neutral	Low	Medium	High	Total
Neutral	15	4	13	4	36
Low	11	21	21	3	56
Medium	4	13	48	10	75
High	2	5	24	5	36
Total	32	43	106	22	203

Table Va – *Number of subjects according to the degree of loss aversion and risk aversion*

Risk Aversion/Loss Aversion	Neutral	Low	Medium	High	
Neutral	7.38	1.97	6.40	1.97	17.73
Low	5.41	10.34	10.34	1.47	27.58
Medium	1.97	6.40	23.64	4.93	36.96
High	0.98	2.46	11.82	2.46	17.73
Total	15.78	21.18	52.21	10.83	100

Table Vb – *The percentage of subjects according to the degree of loss aversion and risk aversion*

4.2 Experimentation strategies

The design of the experiment follows the set-up of the analytical example discussed in Chapter 2. We investigate the results on the experimental choices of the subjects as a function of the parameters of loss aversion and risk aversion. We follow the same structure of the analysis discussed in Chapter 2 and we test the Proposition 1a,1b,2 and 3 illustrated in Chapter 2 on the basis of the experimental data.

Note that the design of the experiment involves only subjects reporting what they would do in response to a series of histories of outcomes.

The essence of a bandit problem is that when a subject chooses not to experiment, no outcome is generated. This raises an issue regarding subjects who report that they will not experiment in stage 1, but who in response to questions about experimentation in later stages say they will experiment following certain histories. To mimic what happens in a physical bandit environment, where such subjects would not observe any such history, we identify these subjects for the subjects' pool, so that the results can be reported both for the full pool and for the "net" pool that excludes this group.

Risk Aversion/Loss Aversion	Neutral	Low	Medium	High	Total
Neutral	15	4	13	4	36
Low	11	21	21	3	56
Medium	4	13	48	10	75
High	2	5	24	5	36
Total	32	43	106	22	203

Table Va – *Number of subjects according to the degree of loss aversion and risk aversion*

Risk Aversion/Loss Aversion	Neutral	Low	Medium	High	
Neutral	7.38	1.97	6.40	1.97	17.73
Low	5.41	10.34	10.34	1.47	27.58
Medium	1.97	6.40	23.64	4.93	36.96
High	0.98	2.46	11.82	2.46	17.73
Total	15.78	21.18	52.21	10.83	100

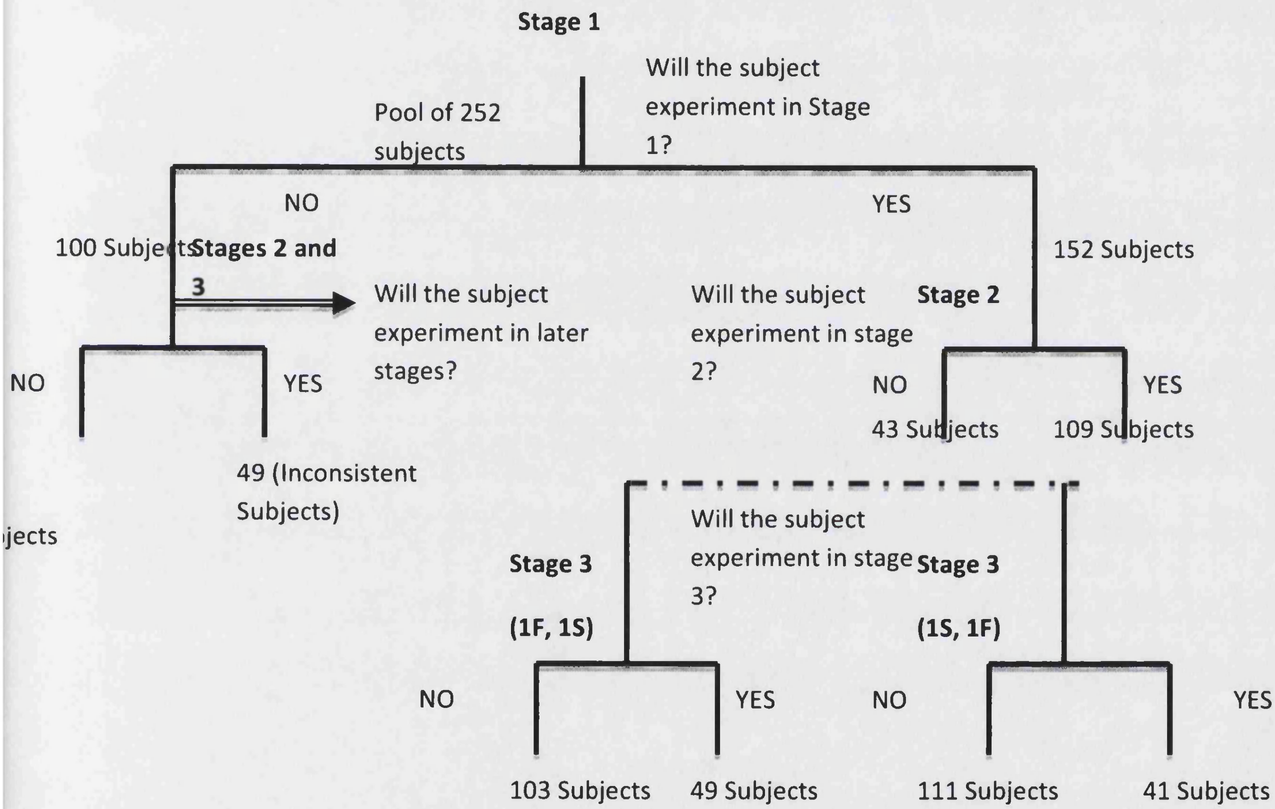
Table Vb – *The percentage of subjects according to the degree of loss aversion and risk aversion*

4.2 Experimentation strategies

The design of the experiment follows the set-up of the analytical example discussed in Chapter 2. We investigate the results on the experimental choices of the subjects as a function of the parameters of loss aversion and risk aversion. We follow the same structure of the analysis discussed in Chapter 2 and we test the Proposition 1a,1b,2 and 3 illustrated in Chapter 2 on the basis of the experimental data.

Note that the design of the experiment involves only subjects reporting what they would do in response to a series of histories of outcomes.

The essence of a bandit problem is that when a subject chooses not to experiment, no outcome is generated. This raises an issue regarding subjects who report that they will not experiment in stage 1, but who in response to questions about experimentation in later stages say they will experiment following certain histories. To mimic what happens in a physical bandit environment, where such subjects would not observe any such history, we identify these subjects for the subjects' pool, so that the results can be reported both for the full pool and for the "net" pool that excludes this group.



Stage 1
Will the subject experiment in Stage 1?
Pool of 252 subjects
NO
YES
100 Subjects
Stages 2 and 3
152 Subjects
Will the subject experiment in later stages?
Will the subject experiment in stage 2?
Stage 2
NO
YES
NO
YES
43 Subjects
109 Subjects
49 (Inconsistent Subjects)
jects
Will the subject experiment in stage 3?
Stage 3
(1F, 1S)
Stage 3
(1S, 1F)
NO
YES
NO
YES
103 Subjects
49 Subjects
111 Subjects
41 Subjects

The idea behind the diagram is to mimic the physical bandit setting, where no information regarding paths who have not been effectively experimented is available to subjects, we ignore what they say they would do under subsequent histories.
The main source of inconsistency in our data set was: "a subject said no to experimentation in period 1 (they were 49), but then he experimented in late stages. This is inconsistent with histories $\{h_1 = 0\}$, $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$ where the posterior is less or equal to the prior at stage 1.

4.2.1 Participation rate in the game

Table VI and VII show the decision of participation (measured as the frequency of participation at the cost of 3.5 euro) of the population according to the distribution of loss aversion. Specifically, the following tables indicate the total individuals who choose to experiment in each stage of the game and the associated participation rate, classified according to:
a) Risk aversion (Table VI a,b)
b) Loss aversion (Table VII a,b)
The statistics presented in the following table contains all the individuals who do not exhibit inconsistent behaviour. Results exhibit the presence of monotonicity of the participation rate as a function of risk aversion and loss aversion. Depending on the history of the game, participation rates tend to decrease as a consequence of failures in the previous round of experimentation.

The idea behind the diagram is to mimic the physical bandit setting, where no information regarding paths who have not been effectively experimented is available to subjects, we ignore what they say they would do under subsequent histories.

The main source of inconsistency in our data set was: "a subject said no to experimentation in period 1 (they were 49), but then he experimented in late stages. This is inconsistent with histories $\{h_1 = 0\}$, $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$ where the posterior is less or equal to the prior at stage 1.

4.2.1 Participation rate in the game

Table VI and VII show the decision of participation (measured as the frequency of participation at the cost of 3.5 euro) of the population according to the distribution of loss aversion. Specifically, the following tables indicate the total individuals who choose to experiment in each stage of the game and the associated participation rate, classified according to:

a) Risk aversion (Table VI a,b)

b) Loss aversion (Table VII a,b)

The statistics presented in the following table contains all the individuals who do not exhibit inconsistent behaviour. Results exhibit the presence of monotonicity of the participation rate as a function of risk aversion and loss aversion. Depending on the history of the game, participation rates tend to decrease as a consequence of failures in the previous round of experimentation.

Risk aversion	Stage 1	Stage 2	Stage 3 (1F, 1S)	Stage 3 (1S, 1F)
6.4	5	2	1	1
0.32	9	5	3	2
0.213	15	10	4	4
0.08	27	18	8	6
0.0355	29	24	10	9
0.0168	33	25	11	9
0.0032	34	25	12	10
Totals	152	109	49	41

Table VIa - Risk aversion: Total number of individuals

Risk aversion	Stage 1	Stage 2	Stage 3 (1F, 1S)	Stage 3 (1S, 1F)
6.4	29%	14%	50%	50%
0.32	47%	55%	60%	40%
0.213	71%	66%	40%	40%
0.08	76%	67%	44%	33%
0.0355	77%	82%	41%	37.5%
0.0168	89%	75%	44%	36%
0.0032	94%	73%	48%	40%

Table VIb - Risk aversion: Participation Rate

Loss aversion	Stage 1	Stage 2	Stage 3 (1F, 1S)	Stage 3 (1S, 1F)
>3	5	3	2	2
3	12	6	3	3
2	21	13	5	4
1.5	33	23	10	8
1.2	39	28	14	11
1	30	23	8	7
0.87	12	13	7	6
Totals	152	109	49	41

Table VIIa: Loss aversion: Total number of individuals

Loss aversion	Stage 1	Stage 2	Stage 3 (1F, 1S)	Stage 3 (1S, 1F)
>3	35%	60%	66%	66%
3	58%	50%	50%	50%
2	66%	61%	38%	38%
1.5	82%	69%	43%	34%
1.2	81%	71%	50%	39%
1	90%	80%	53%	39%
0.87	85%	100%	58%	50%

Table VIIb: Loss aversion: Participation Rate

We proceed by disaggregating the data on the basis of the risk aversion and loss aversion parameter and we examine the participation rate according to both coefficients.

Risk aversion/ Loss aversion	0.87	1	1.2	1.5	2	3	>3	Totals
6.4	1,1	0,1	1,3	1,4	1,2	1,1	0,5	5,17
0.32	0,1	2,2	2,4	3,4	0,7	0,1	1,1	9,19
0.213	1,1	6,7	4,6	3,4	1,1	0,1	0,1	15,21
0.08	1,2	5,5	4,5	5,5	3,8	7,7	2,3	27,35
0.0355	0,1	7,8	8,8	5,5	7,8	1,5	1,3	29,38
0.0168	3,3	7,7	9,10	7,8	6,7	1,1	0,1	33,37
0.032	5,5	3,3	11,12	9,10	3,3	2,2	1,1	34,36
Totals	12,14	30,33	39,48	33,40	21,36	12,18	5,14	152,203

Table VIIIa First-stage of experimentation

Risk aversion/ Loss aversion	0.87	1	1.2	1.5	2	3	>3	Totals
6.4	1,1	0,0	1,1	0,1	0,1	0,1	0,0	2,5
0.32	0,0	2,2	2,2	1,3	0,0	0,0	0,1	5,9
0.213	1,1	3,6	2,4	3,3	1,1	0,0	0,0	10,15
0.08	1,1	3,5	2,4	5,5	2,3	3,7	1,2	18,27
0.0355	0,0	7,7	7,8	3,5	5,7	1,1	1,1	24,29
0.0168	3,3	6,7	8,9	4,7	3,6	1,1	0,0	25,33
0.032	5,5	3,3	7,11	6,9	2,3	1,2	1,1	25,34
Totals	12,12	24,30	28,39	23,33	13,21	6,12	3,5	109,152

Table VIIIb Second-stage of experimentation

Risk aversion/ Loss aversion	0.87	1	1.2	1.5	2	3	>3	Totals
6.4	1,1	0,0	0,1	0,0	0,0	0,0	0,0	1,2
0.32	0,0	2,2	1,2	0,1	0,0	0,0	0,0	3,5
0.213	1,1	1,3	1,2	1,3	0,1	0,0	0,0	4,10
0.08	1,1	1,3	1,2	2,5	1,2	1,3	1,1	8,18
0.0355	0,0	1,7	3,7	2,3	2,5	1,1	1,1	10,24
0.0168	1,3	1,7	3,8	3,4	2,3	1,1	0,0	11,25
0.032	3,5	2,3	5,7	2,6	0,2	0,1	0,1	12,25
Totals	7,12	8,24	14,28	10,23	5,13	3,6	2,3	49,109

Table VIIIc: Third-stage of experimentation (1 Failure, 1 Success)

Figure III exhibits the frequency of participation to the experiment in Stage 1 as a function of risk aversion. This figure exhibits a monotonicity that matches exactly the analysis done in the three-stage game illustrated in Chapter 2. The same monotonic shape of the participation rate is confirmed in Stage 2 of the game. Depending on the nature of the game, participation rates tend to decrease in Stage 2 as a consequence of failures in the previous round of experimentation.

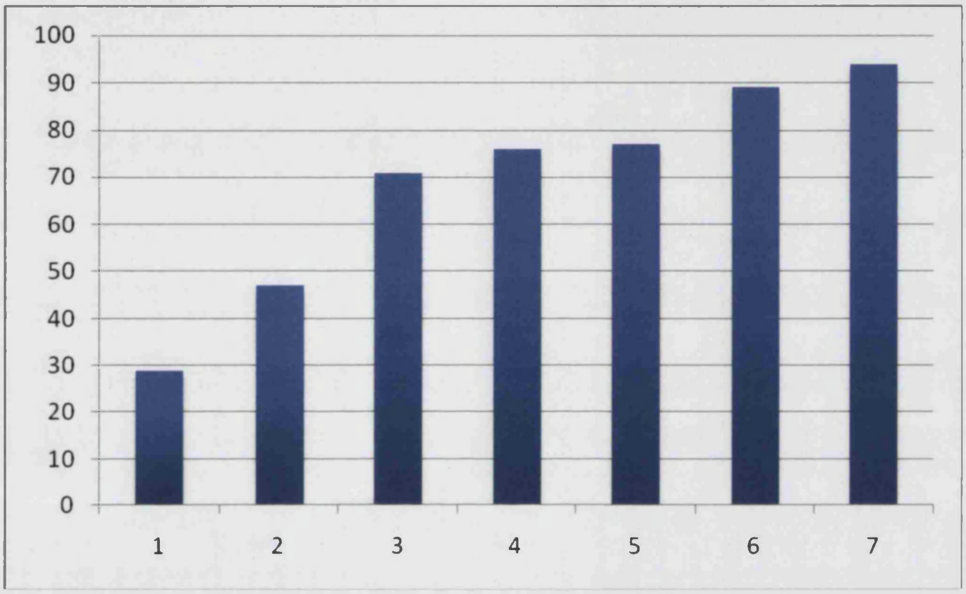


Figure III - *Participation rate and risk aversion*
(Ordered from HIGH to LOW risk aversion)

Test 1b: Is the level of experimentation a monotonic decreasing function in the parameter of loss aversion r?

Figure IV exhibit the frequency of participation to the experiment in Stage 1 as a function of loss aversion. The experimental decisions support the monotonicity prediction of Proposition 1b. The result is confirmed in Stage 2 of the game.

Risk aversion/ Loss aversion	0.87	1	1.2	1.5	2	3	>3	Totals
6.4	0,1	0,0	1,1	0,0	0,0	0,0	0,0	1,2
0.32	0,0	1,2	1,2	0,1	0,0	0,0	0,0	2,5
0.213	0,1	2,3	1,2	1,3	0,1	0,0	0,0	4,10
0.08	1,1	0,3	2,2	2,5	0,2	0,3	1,1	6,18
0.0355	0,0	0,7	4,7	2,3	1,5	1,1	1,1	9,24
0.0168	1,3	1,7	0,8	4,4	2,3	1,1	0,0	9,25
0.032	3,5	3,3	2,7	0,6	1,2	1,1	0,1	10,25
Totals	6,12	7,24	14,28	8,23	4,13	3,6	2,3	41,109

Table VIIId: Third-stage of experimentation (1 Success, 1 Failure)

4.2.2 Testing the theory on experimental data

In this section we test the theoretical predictions derived in Chapter 2 on the experimental data. Specifically, we test Propositions 1a, 1b, 2 and 3 of Chapter 2 and we discuss the experimental results following closely the structure of the theoretical analysis developed in Chapter 2.

First, we use experimental data to test whether the experimentation decisions of the subjects are a function of the risk aversion according to Propositions 1a, and as a function of the loss aversion parameter according to Proposition 1b.

Test 1a: Is the level of experimentation a monotonic decreasing function in the parameter of risk aversion *r*?

Figure III exhibits the frequency of participation to the experiment in Stage 1 as a function of risk aversion. This figure exhibits a monotonicity that matches exactly the analysis done in the three-stage game illustrated in Chapter 2. The same monotonic shape of the participation rate is confirmed in Stage 2 of the game. Depending on the nature of the game, participation rates tend to decrease in Stage 2 as a consequence of failures in the previous round of experimentation.

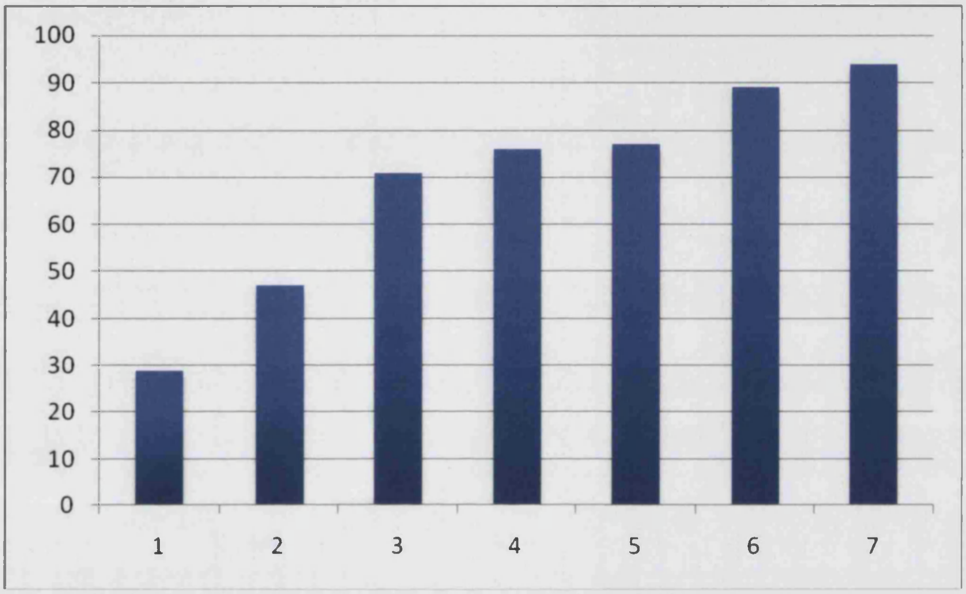


Figure III - *Participation rate and risk aversion*

(Ordered from HIGH to LOW risk aversion)

Test 1b: Is the level of experimentation a monotonic decreasing function in the parameter of loss aversion r?

Figure IV exhibit the frequency of participation to the experiment in Stage 1 as a function of loss aversion. The experimental decisions support the monotonicity prediction of Proposition 1b. The result is confirmed in Stage 2 of the game.

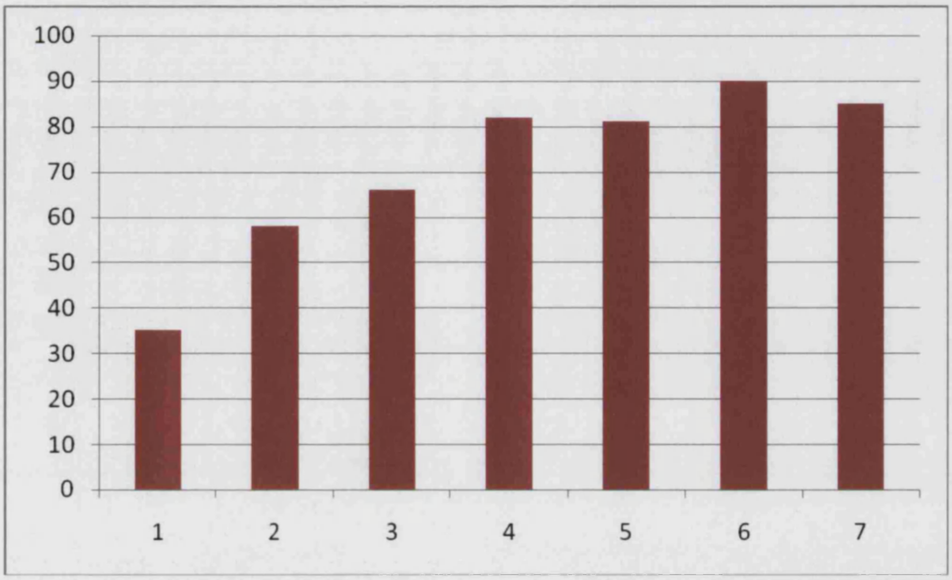


Figure IV - *Participation rate and loss aversion*
(Ordered from HIGH to LOW loss aversion)

We study the experimental decisions in more detail by estimating the optimal amount of experimentation. Subjects in the sample may decide to experiment more or less with respect to the optimal experimentation choice predicted by the theoretical analysis in Chapter 2. We calculate separately the range of values for each parameter (of either risk aversion or loss aversion) that supports experimentation as the optimal choice, when the ticket price to play on the slot machine is set equal to 3.5 euros[51]. We compare the experimental decision undertaken by subjects for each parameter value and the optimal experimentation choice predicted according to the analysis developed in Chapter 2 for the three-stage example. We examine separately the role of risk aversion and loss aversion in the prediction of experimental decisions.

Figures Va and Vb depict the actual participation rate of subjects for each value of *r*, the risk aversion parameter. Figure Va and Vb plot the experimentation choices respectively in the first stage of the game and in the second stage after one failure is realized in Stage one. The dotted line indicates the maximum level of risk aversion *r* that supports experimentation as the optimal choice according to the theoretical analysis developed in Chapter 2. Risk averse subjects located to the left of the dotted line in Figure V and who choose to experiment behave optimally according to theory, while those located to the right of the dotted line in Figure III choosing to experiment do not adopt an optimizing behaviour.

[51] The threshold value for each coefficient respectively of risk aversion and loss aversion is calculated by imposing that the subject is indifferent between the decision to experiment on the slot machine and the decision to spare the ticket cost, equal to 3.5 euro and play on the safe arm. We calculate the threshold value on the basis of the Gittins index calculated in Chapter 2 for the three-stage example in two alternative cases: when the subject is risk averse, but not loss averse and when the subject is loss averse, but not loss averse.

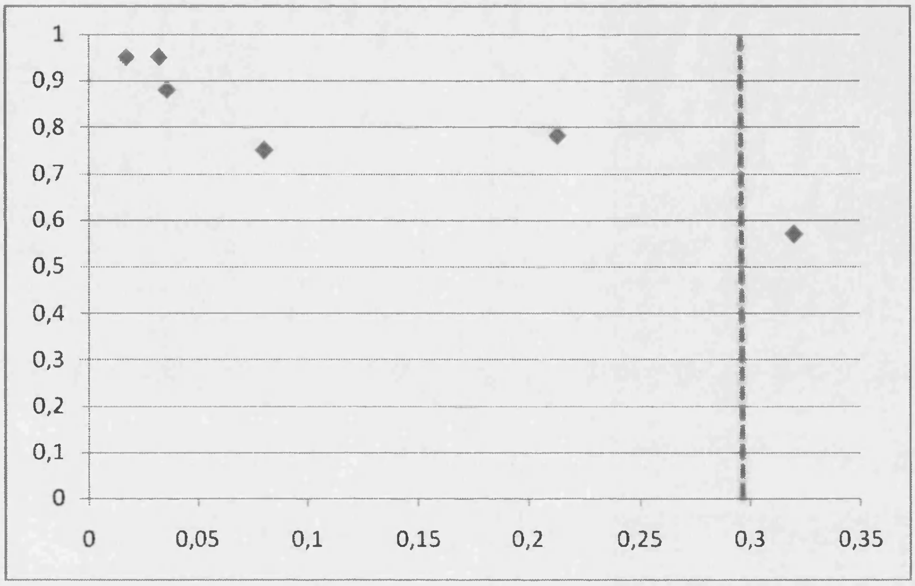


Figure Va
Experimentation at Stage 1
Participation rate on the basis of risk aversion measure RA

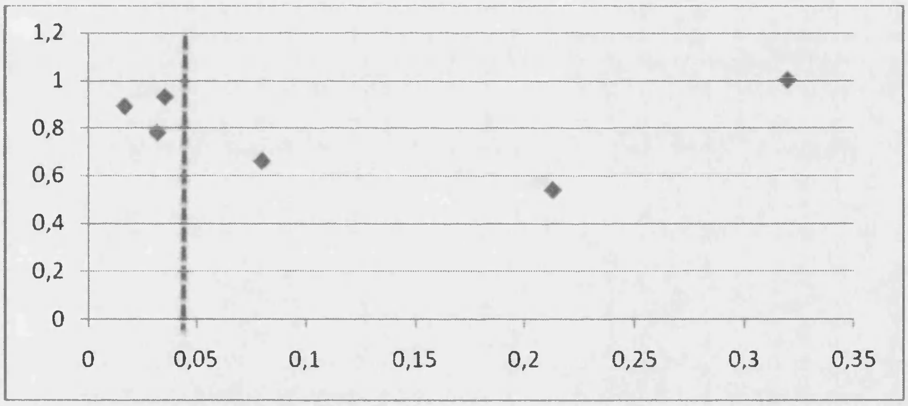


Figure Vb
Experimentation at Stage 2 after history $\{h_1 = 0\}$
Participation rate on the basis of risk aversion measure RA

Figures VIa and VIb illustrate the actual participation rate of subjects for each value of λ, the loss aversion parameter. Figure VIa and VIb plot the experimentation choices respectively in the first stage of the game and in the second stage after one failure is realized in Stage one. Similarly to the previous analysis, the dotted line indicates the maximum level of loss aversion λ that supports experimentation as the optimal choice according to the theoretical analysis developed in Chapter 2.

Loss averse subjects located to the left of the dotted line in Figure VI and who choose to experiment behave optimally according to theory, while those located to the right of the dotted line in Figure III choosing to experiment do not adopt an optimizing behaviour.

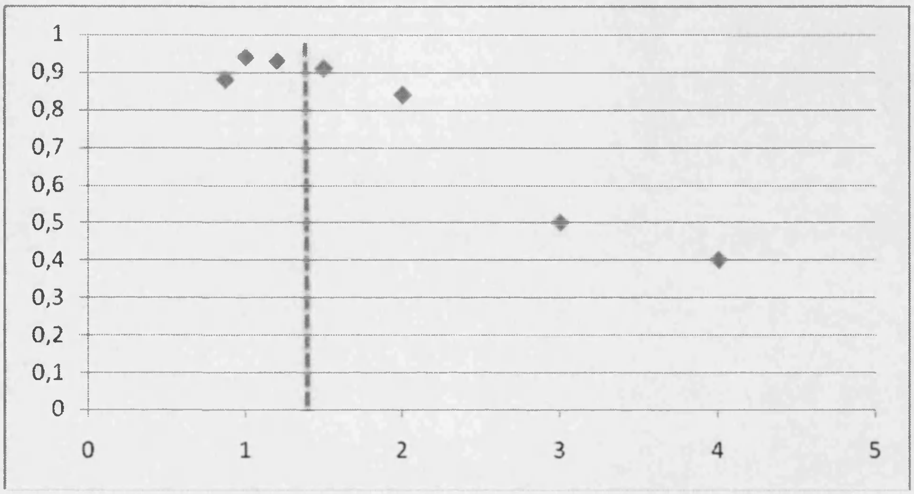


Figure VIa
Experimentation at Stage 1
Participation rate on the basis of risk aversion measure LA

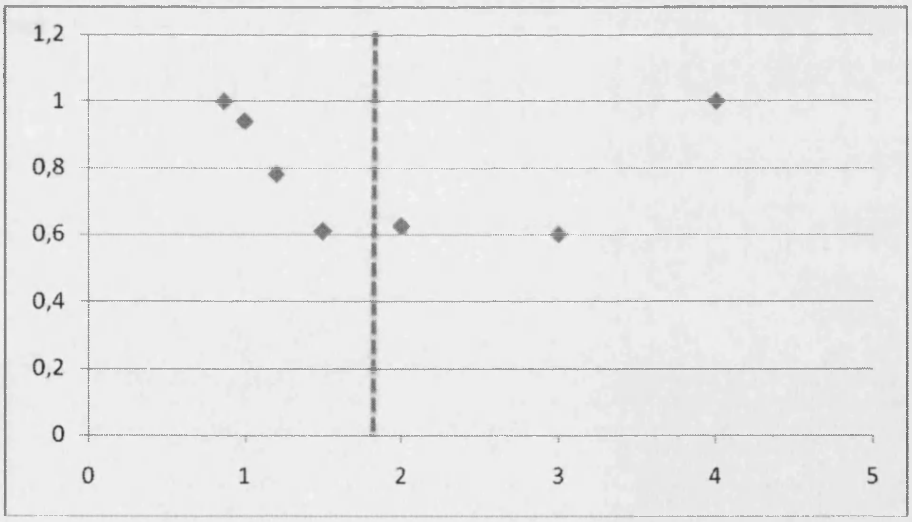


Figure VIb
Experimentation at Stage 2 after history $\{h_1 = 0\}$
Participation rate on the basis of risk aversion measure LA

The bias towards excess experimentation given the degree of loss aversion and risk aversion occurs:

- in stage 1 of experiment
- in stage 2 of the experiment, where agents continue experimenting

From the analysis of Figures Va and VIa it emerges that the curve of subjects' participation to the experiment in stage 1 has a common shape that tends downward as the degree of risk aversion (Figure Va) increases or as the degree of loss aversion (Figure VIa) increases.

It is important to note that Figures Vb and VIb show the presence of excess experimentation (above optimal experimentation) after a history $h_1 = \{0\}$. Hence, Figures VIb and VIb show the persistence of the bias to experiment at stage 2 after a history, showing those who experiment in stage 2 and should have not experimented even in stage 1.

There is a bias towards excess experimentation even for those subjects who are intensely risk averse (high level of parameter *r*) or intensely loss averse (high level of parameter λ). Those subjects characterized by either high levels of risk aversion or loss aversion are less likely to experiment in stage 1, but conditional on the experimentation decision in stage 1, subjects are likely to continue at stage 2 after $h_1 = \{0\}$. In other words, this means that the bias to experiment comes on to stage 2 after $h_1 = \{0\}$, i.e. subjects are likely to continue with experimentation conditional on the suboptimal decision of experimenting in stage 1.

We proceed by examining the experimental data regarding the subjects' experimentation choices in Stage 3 of the game. According to the theoretical predictions discussed in Chapter 2, if a subjectdecides to experiment in stage 1, then he should find it optimal to experiment in Stage 3, following a history $\{h_1 = 0, h_2 = 1\}$ or $h = \{1, 0\}$, independently of the degree of risk aversion, *r* and the degree of loss aversion, λ. This is because the posterior associated with the two histories is the same - so symmetry in experimentation strategies.

Observation 1: A history of $\{h_1 = 0, h_2 = 1\}$ must lead to experimentation in Stage 3. But a history of $\{h_1 = 1, h_2 = 0\}$ gives same posterior probability of a history $\{h_1 = 0, h_2 = 1\}$. Therefore the agent must experiment after a history of $\{h_1 = 1, h_2 = 0\}$.

Observation 2: The 109 subjects out of the 203 (constituting the initial pool of subjects we start with) are coming through stage 3.[52]

[52] Experimental data can be summarized as follows. 252 subjects take part to the experiment, among which 203 subjects behave coherently with rationality and 49 adopt an inconsistent behaviour. 152 subjects (out of 203) choose to experiment and 51 to stop at stage 1. In stage 2 following history $\{h_1 = 0\}$, 109 subjects (out of 152 who experimented in stage 1) choose to experiment and 43 to stop. In stage 3 following history $\{h_1 = 1, h_2 = 0\}$, 49 subjects (out of 109

From the analysis of Figures Va and VIa it emerges that the curve of subjects' participation to the experiment in stage 1 has a common shape that tends downward as the degree of risk aversion (Figure Va) increases or as the degree of loss aversion (Figure VIa) increases.

It is important to note that Figures Vb and VIb show the presence of excess experimentation (above optimal experimentation) after a history $h_1 = \{0\}$. Hence, Figures VIb and VIb show the persistence of the bias to experiment at stage 2 after a history, showing those who experiment in stage 2 and should have not experimented even in stage 1.

There is a bias towards excess experimentation even for those subjects who are intensely risk averse (high level of parameter *r*) or intensely loss averse (high level of parameter λ). Those subjects characterized by either high levels of risk aversion or loss aversion are less likely to experiment in stage 1, but conditional on the experimentation decision in stage 1, subjects are likely to continue at stage 2 after $h_1 = \{0\}$. In other words, this means that the bias to experiment comes on to stage 2 after $h_1 = \{0\}$, i.e. subjects are likely to continue with experimentation conditional on the suboptimal decision of experimenting in stage 1.

We proceed by examining the experimental data regarding the subjects' experimentation choices in Stage 3 of the game. According to the theoretical predictions discussed in Chapter 2, if a subjectdecides to experiment in stage 1, then he should find it optimal to experiment in Stage 3, following a history $\{h_1 = 0, h_2 = 1\}$ or $h = \{1, 0\}$, independently of the degree of risk aversion, *r* and the degree of loss aversion, λ. This is because the posterior associated with the two histories is the same - so symmetry in experimentation strategies.

Observation 1: A history of $\{h_1 = 0, h_2 = 1\}$ must lead to experimentation in Stage 3. But a history of $\{h_1 = 1, h_2 = 0\}$ gives same posterior probability of a history $\{h_1 = 0, h_2 = 1\}$. Therefore the agent must experiment after a history of $\{h_1 = 1, h_2 = 0\}$.

Observation 2: The 109 subjects out of the 203 (constituting the initial pool of subjects we start with) are coming through stage 3.[52]

[52] Experimental data can be summarized as follows. 252 subjects take part to the experiment, among which 203 subjects behave coherently with rationality and 49 adopt an inconsistent behaviour. 152 subjects (out of 203) choose to experiment and 51 to stop at stage 1. In stage 2 following history $\{h_1 = 0\}$, 109 subjects (out of 152 who experimented in stage 1) choose to experiment and 43 to stop. In stage 3 following history $\{h_1 = 1, h_2 = 0\}$, 49 subjects (out of 109

A (0,1) or (1,0) history is interesting theoretically since the posterior probability of a win along the risky arm in Stage 3 corresponds to the prior probability of 1/2 in Stage 1. This implies that the gamble at Stage 3 should be equivalent to the Stage 1 gamble. We investigate this statement by breaking the questions into two tests, namely Test 2 and Test 3.

Test 2: Test the equivalence between experimentation following a history (0,1) and a history (1,0) in stage 3 of the game.

Are subjects more likely to experiment after $\{h_1 = 0, h_2 = 1\}$ than after $\{h_1 = 1, h_2 = 0\}$?It could be that subjects choose to experiment more following a success than after a failure. In the presence of a statistical significant difference of experimentation choices after a history $\{h_1 = 0, h_2 = 1\}$ and a history $\{h_1 = 1, h_2 = 0\}$, we may have a "recency" bias, i.e. the tendency of subjects to recall the more recent piece of information and adapt the experimentation accordingly. In the presence of a "recency" bias, after histories leading to the same posterior subjects will tend to experiment more if a success is recorded in the last period and less if a failure occurred in the last period.
Table IX illustrates the differences between experimental decision after (0,1) and (1,0).

Risk aversion/ Loss aversion	0.87	1	1.2	1.5	2	3	4
6.4	1	0	-1	0	0	0	0
0.32	0	1	0	0	0	0	0
0.213	1	-1	0	0	0	0	0
0.08	0	1	-1	0	1	1	0
0.0355	0	1	-1	0	0	0	0
0.0168	0	0	3	-1	0	0	0
0.032	0	-1	3	2	-1	-1	0

Table IX:*A comparison between experimental decisions following history (0,1) and (1,0)*

who experimented in stage 1) choose to experiment and 50 to stop. In stage 3 following history $\{h_1 = 1, h_2 = 0\}$, 41 subjects (out of 109 who experimented in stage 1) choose to experiment and 58 to stop.

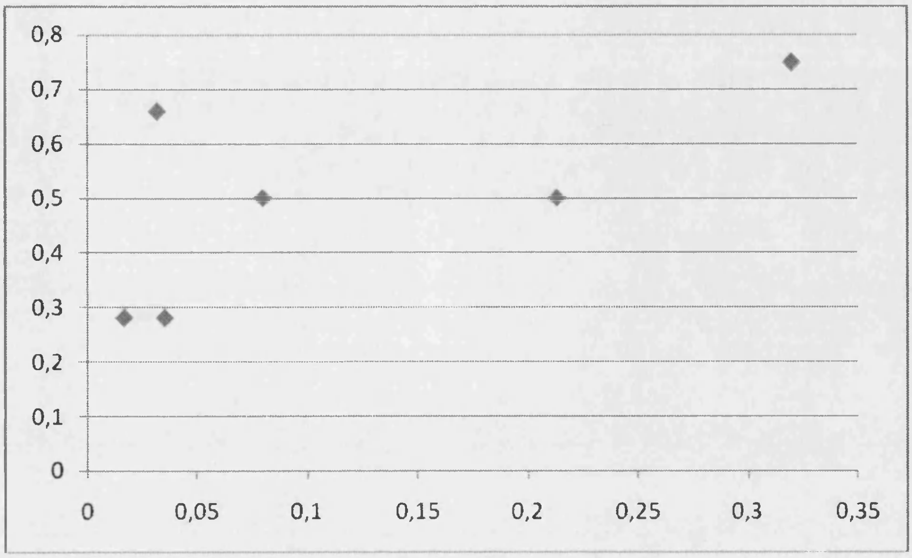


Figure Va

Experimentation at Stage 3 after history $\{h_1 = 0, h_2 = 1\}$

Participation rate on the basis of risk aversion measure RA

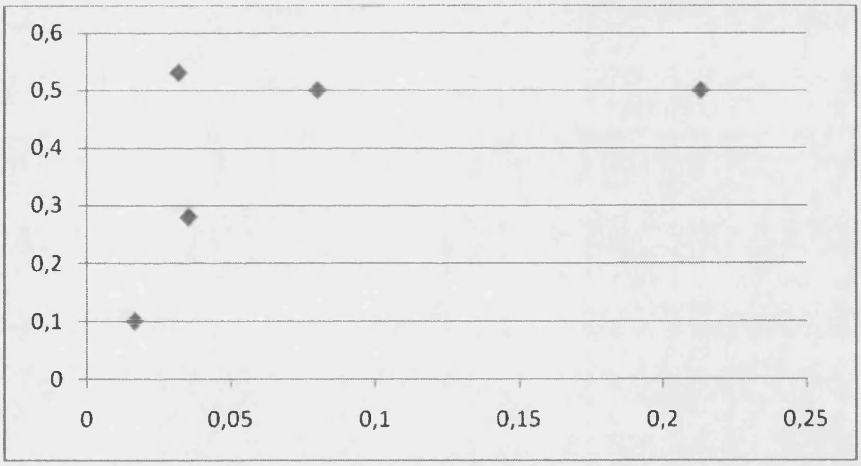


Figure Vb

Experimentation at Stage 3 after history $\{h_1 = 1; h_2 = 0\}$

Participation rate on the basis of risk aversion measure RA

From inspection of Table IX, the sum of absolute differences of experimental decisions is 7 out of 109 subjects. Additionally, we have 15 positive differences and 8 negative differences.

The equivalence between experimentation choices following history $\{h_1 = 0, h_2 = 1\}$ and history $\{h_1 = 1, h_2 = 0\}$ is tested under the null hypothesis that the probability of positive errors should be equal to the probability of negative errors. Note that this is indeed the case since $\frac{15-8}{\sqrt{109}}$ is less than 1. We can thereby conclude that there is no statistical significant difference of experimentation strategy continuing history $\{h_1 = 0, h_2 = 1\}$ or history $\{h_1 = 1, h_2 = 0\}$.

Test 3: Test the presence of over or under experimentation, i.e. the equivalence of $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$ with history $\{h = \varnothing\}$.

We compare the experimental decisions in stage 3 following histories $\{h_1 = 0, h_2 = 1\}$ and $\{h_1 = 1, h_2 = 0\}$ with history $\{h = \varnothing\}$. We have shown the equivalence of experimentation strategies in Stage 3 following two alternative histories. However, subjects in stage 3 of the game may decide to experiment more or less than in stage 1, despite the same probability of a win - which should lead to the same experimentation choice predicted by the theoretical analysis in Chapter 2. This implies that subjects that decide to experiment in Stage 2 after one failure should decide to experiment at Stage 3.

Figures VIIa and VIIb depict the actual participation rate of subjects for each value of r, the risk aversion parameter. Figure Va and Vb plot the experimentation choices in stage 3 of the game respectively following history $\{h_1 = 0, h_2 = 1\}$ and history $\{h_1 = 1, h_2 = 0\}$. Figures VIa and VIb illustrate the actual participation rate of subjects for each value of λ, the loss aversion parameter. Figure VIa and VIb plot the experimentation choices in stage 3 of the game respectively following history $\{h_1 = 0, h_2 = 1\}$ and history $\{h_1 = 1, h_2 = 0\}$.

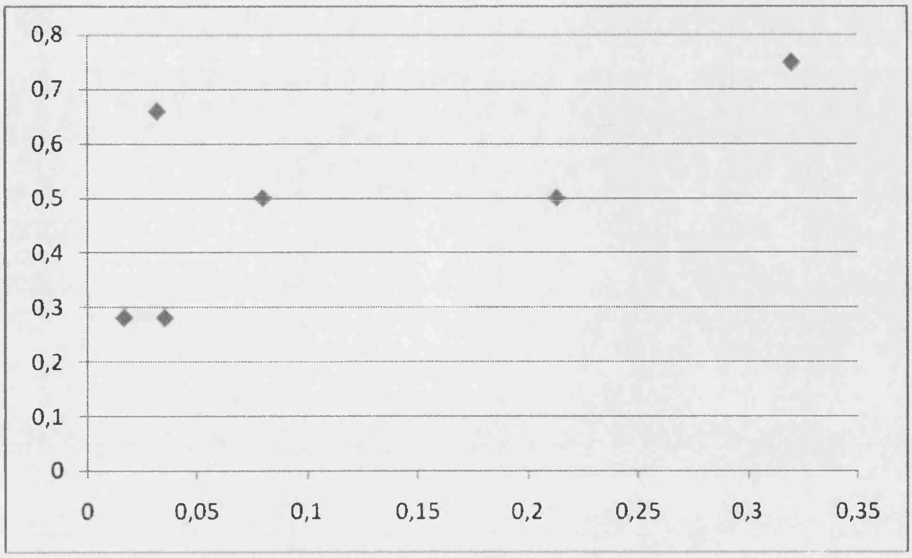


Figure Va

Experimentation at Stage 3 after history $\{h_1 = 0, h_2 = 1\}$

Participation rate on the basis of risk aversion measure RA

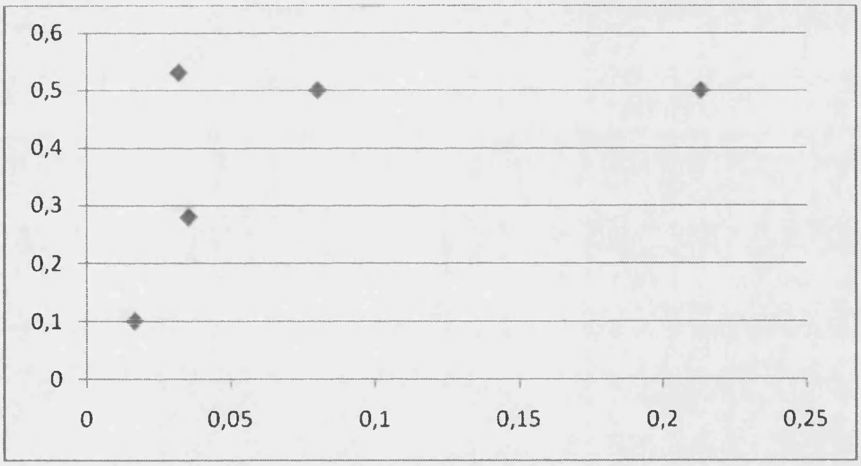


Figure Vb

Experimentation at Stage 3 after history $\{h_1 = 1; h_2 = 0\}$

Participation rate on the basis of risk aversion measure RA

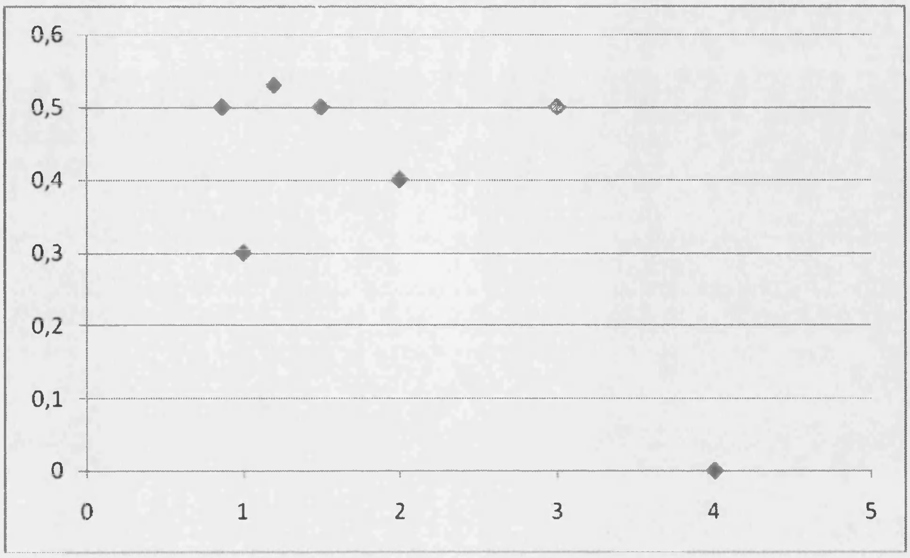


Figure VIa

Experimentation at Stage 3 after history $\{h_1 = 0, h_2 = 1\}$

Participation rate on the basis of risk aversion measure LA

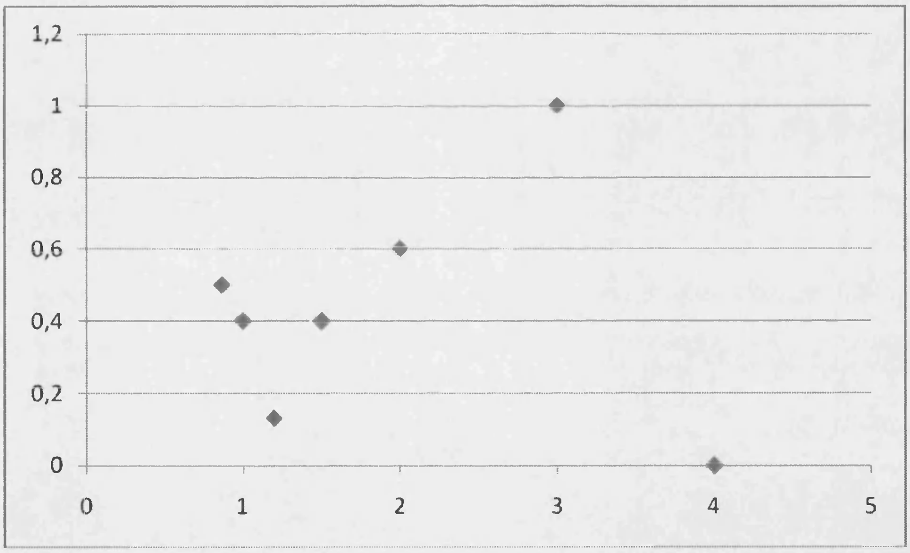


Figure VIb

Experimentation at Stage 3 after history $\{h_1 = 1, h_2 = 0\}$

Participation rate on the basis of risk aversion measure RA

We compare the experimental decisions undertaken by subjects in Stage 3 (Figure V and VI) and in Stage 1 (Figure III) for each parameter value. We examine separately the role of risk aversion and loss aversion in the prediction of experimental decisions.

Our results show the presence of under-experimentation relative to history $\{h_1 = 0\}$, i.e. amount of those who chose to experiment in stage 2 after $\{h_1 = 0\}$, but refuse to experiment in period 3, after a "better" history $\{h_1 = 0; h_2 = 1\}$ or $\{h_1 = 1, h_2 = 0\}$.

The question we investigate is whether we should have a fall in experimentation after history $\{h_1 = 0; h_2 = 1\}$ or $\{h_1 = 1, h_2 = 0\}$, independently of the ranges of r and λ. Experimental data shows that independently of the combination of loss aversion and risk aversion, only half of subjects who experimented in stage 2 choose to continue in stage 3. Hence, in stage 3 we observe a sharp decrease in experimentation when it is optimal to experiment. This contradicts with the theoretical prediction, according to which given the combination of risk aversion and loss aversion, the optimal strategy predicts thata subject finds it optimal to experiment in stage 3 if the subject has experimented in stage 1 and continued to stage 2. Experimental data shows the presence of a bias to stop in stage 3.

Additionally, we might expect that failures to continue experimentation in stage 3 following history (1,0) or (0,1) are more likely in the case of subjects with high levels of risk aversion and/or loss aversion. On the contrary, experimental data show that those subjects who have low levels of risk aversion and loss aversion are the most likely to discontinue experimenting at stage 3 following history (1,0) or (0,1). Figure V and VI show this upward trend. We can check the presence of this bias to stop by splitting into halves the subjects pool (low r, low λ versus high r, high λ) and check the continuation probability for both subgroups.

4.3 The role of loss aversion

We aim to test whether a model that makes use of both risk aversion and loss aversion may explain the experimental data better than a model that makes use only of the information regarding the individual level of risk aversion.

The theoretical model developed in Chapter 2 allows us to explicitly characterize the optimal procedure of experimentation of each subject for any degree of risk aversion and any degree of loss aversion and make a theoretical prediction regarding the individual decision to continue experimentation for any ticket price. We compute the dynamic allocation index in each stage of the game following the analytical procedure developed in Section 3 of Chapter 2. The dynamic allocation index for each subject in the experiment is computed as a function of the loss aversion parameter λ and the risk aversion parameter r[53]. The dynamic allocation index acts as a reservation price: if the ticket price of 3.5 euro falls below the "theoretical" dynamic allocation index, subjects will find it optimal to experiment (we label this outcome "continue"). If the ticket price of 3.5 euro is above the "theoretical" dynamic allocation index, subjects will find it optimal to stop experimenting (we label this outcome "stop"). This is what we call the "predicted" outcome of the model. For each subject we can compare the predicted outcome and the "real" outcome, as a result of the experimental choice of subjects.

We compare two alternative models: a) the model that uses only risk aversion and b) the model that uses both the risk aversion and loss aversion. Tables X, XI, XII and XIII compare the predicted outcomes (respectively of option a) and b)) and the experimental outcomes in the four histories of the game. Results are classified in a 2×2 matrix as follows. Both the "predicted" and the "experimental" outcome can either be "continue" (if the subject finds optimal to experiment and play on the risky arm at the ticket price of 3.5 euro, given the prior history of the game) or "stop" in the opposite case. If the predicted outcome coincides with the experimental one, the theoretical model makes the correct prediction. On the contrary, the model produces a wrong prediction if the classification produces a mismatch ending in the matrix cells (stop, continue) and (continue, stop).

We perform this analysis in each stage of the game, conditional on the history of play. Note that subjects who choose to opt out in, say, Stage 2 of the game will not be counted in Stage 3, since the one-armed bandit problem is an optimal stopping game.

We perform a Fisher exact test to evaluate the predictive ability of the two alternative models, one that uses only risk aversion (labelled RA in the matrix) and the one that uses both risk aversion and loss aversion (labelled RA-LA in the matrix). The Fisher exact test is based on the hyper-geometric distribution that any of the two contingency tables should follow, under the null hypothesis of perfect independence among the two variables (predicted and experimental outcome). If the p-value of the test is large, then there is more evidence in favour of independence. We report the test statistics at the bottom end of each table for each of the two models.

[53] In order to compute the dynamic allocation index at the individual level in each stage of the game, we need first to reconcile the risk aversion measure computed in Table III with the functional form of the utility function assumed in Section 3 of Chapter 2. We interpret the risk aversion estimate in table III as the absolute measure of risk aversion and we compute the parameter r.

aversion parameter λ and the risk aversion parameter r[53]. The dynamic allocation index acts as a reservation price: if the ticket price of 3.5 euro falls below the "theoretical" dynamic allocation index, subjects will find it optimal to experiment (we label this outcome "continue"). If the ticket price of 3.5 euro is above the "theoretical" dynamic allocation index, subjects will find it optimal to stop experimenting (we label this outcome "stop"). This is what we call the "predicted" outcome of the model. For each subject we can compare the predicted outcome and the "real" outcome, as a result of the experimental choice of subjects.

We compare two alternative models: a) the model that uses only risk aversion and b) the model that uses both the risk aversion and loss aversion. Tables X, XI, XII and XIII compare the predicted outcomes (respectively of option a) and b)) and the experimental outcomes in the four histories of the game. Results are classified in a 2×2 matrix as follows. Both the "predicted" and the "experimental" outcome can either be "continue" (if the subject finds optimal to experiment and play on the risky arm at the ticket price of 3.5 euro, given the prior history of the game) or "stop" in the opposite case. If the predicted outcome coincides with the experimental one, the theoretical model makes the correct prediction. On the contrary, the model produces a wrong prediction if the classification produces a mismatch ending in the matrix cells (stop, continue) and (continue, stop).

We perform this analysis in each stage of the game, conditional on the history of play. Note that subjects who choose to opt out in, say, Stage 2 of the game will not be counted in Stage 3, since the one-armed bandit problem is an optimal stopping game.

We perform a Fisher exact test to evaluate the predictive ability of the two alternative models, one that uses only risk aversion (labelled RA in the matrix) and the one that uses both risk aversion and loss aversion (labelled RA-LA in the matrix). The Fisher exact test is based on the hyper-geometric distribution that any of the two contingency tables should follow, under the null hypothesis of perfect independence among the two variables (predicted and experimental outcome). If the p-value of the test is large, then there is more evidence in favour of independence. We report the test statistics at the bottom end of each table for each of the two models.

[53] In order to compute the dynamic allocation index at the individual level in each stage of the game, we need first to reconcile the risk aversion measure computed in Table III with the functional form of the utility function assumed in Section 3 of Chapter 2. We interpret the risk aversion estimate in table III as the absolute measure of risk aversion and we compute the parameter r.

RA			RA-LA		
	Experimental outcome			Experimental outcome	
Predicted Outcome	Continue	Stop	Predicted Outcome	Continue	Stop
Continue	122	21	Continue	128	16
Stop	26	34	Stop	21	38
p-value = 3.128e-09 95% confidence interval: 3.607375 16.051367			p-value= 5.171e-14 95 % confidence interval: 6.473132 32.716647		

Table X: Stage 1, $\{h=\varnothing\}$

RA			RA-LA		
	Experimental outcome			Experimental outcome	
Predicted Outcome	Continue	Stop	Predicted Outcome	Continue	Stop
Continue	93	15	Continue	98	20
Stop	16	28	Stop	14	30
p-value = 2.658e-09 95% confidence interval: 4.429928 26.813094			p-value = 1.26e-09 95 % confidence interval: 4.420605 25.235523		

Table XI: Stage 2,h_1=0

The Fisher exact test shows that both the RA and LA-RA model have predictive power on the experimental choices of subjects. Additionally, the Fisher test indicates that the predictive ability of the model making use of both risk aversion and loss aversion is higher than the one making use only of the "risk aversion" model. Consider Table IV: both models produce fairly small p-values, showing that both RA and RA-LA models have the ability to predict the outcome. However, the RA-LA model gives a p-value of 6 orders of magnitude smaller than the RA model. Furthermore, the confidence interval for the true odds ratio for the RA model (3.91858 64.03977) does not contain the estimated odds ratio for the RA-LA method (162.4929). Conversely, the confidence interval built for the RA-model does not contain the estimated odds ratio for the RA model. The same analysis applies for all histories of the game. The same conclusions can be reached by comparing the odds ratio associated with the RA and RA-LA models.

We have focused our analysis on the comparison of the predictive power of the RA-LA model and the RA model. This appears to be the most sensible comparison because the presence of risk aversion can lead players to stop experimenting before risk neutral agents and we aimed to insulate the additional explicative power of loss aversion in bandit settings. Nonetheless, we have further compared the 'LA-RA' model with the model that did not include risk aversion and loss aversion. Not surprisingly, we can conclude that the LA-RA model performs better and has a better predictive power than the standard bandit model with risk-neutral and loss-neutral agents.

4.4 The main differences between theoretical and empirical analysis

The analysis of the empirical results shows some differences with the theoretical predictions in Chapter 2. The main differences between the prediction and the experimental results lies in the bias toward excess experimentation in stages 1 and 2 of the game and towards less experimentation in stage 3 of the game.

We consider two alternative "behavioural" explanations, which we argue cannot rationalize these outcomes.

Candidate Interpretation 1: A failure induces subjects to excessively reduce the subjective probability of a success in the experimentation process.

Following this line of explanation we can rationalize the bias towards under-experimentation in stage 2.Note, however, that there should then be a sharp fall in expectation of p in stage 2 following a failure in stage 1, leading to under-experimentation. This is not the case in our experimental results. Hence, the proposed "behavioural" explanation can rationalize the tendency to under-experimentation in stage 3, but it does not appear to be consistent with the experimental data in stage 2, which shows excess experimentation.

RA			RA-LA		
	Experimental outcome			Experimental outcome	
Predicted Outcome	Continue	Stop	Predicted Outcome	Continue	Stop
Continue	41	8	Continue	45	4
Stop	6	14	Stop	4	16
p-value = 3.208e-05 95 % confidence interval: 3.058603 48.914165			p-value = 1.182e-12 95 % confidence interval: 19.58121 7465.21302		

Table XII: Stage 3, h_1=0, h_2=1

RA			RA-LA		
	Experimental outcome			Experimental outcome	
Predicted Outcome	Continue	Stop	Predicted Outcome	Continue	Stop
Continue	36	9	Continue	41	5
Stop	5	19	Stop	1	23
p-value = 2.453e-06 95% confidence interval: 3.91858 64.03977			p-value = 1.182e-12 95 % confidence interval: 19.58121 7465.21302		

Table XIII: Stage 3, h_1=1, h_2=0

The Fisher exact test shows that both the RA and LA-RA model have predictive power on the experimental choices of subjects. Additionally, the Fisher test indicates that the predictive ability of the model making use of both risk aversion and loss aversion is higher than the one making use only of the "risk aversion" model. Consider Table IV: both models produce fairly small p-values, showing that both RA and RA-LA models have the ability to predict the outcome. However, the RA-LA model gives a p-value of 6 orders of magnitude smaller than the RA model. Furthermore, the confidence interval for the true odds ratio for the RA model (3.91858 64.03977) does not contain the estimated odds ratio for the RA-LA method (162.4929). Conversely, the confidence interval built for the RA-model does not contain the estimated odds ratio for the RA model. The same analysis applies for all histories of the game. The same conclusions can be reached by comparing the odds ratio associated with the RA and RA-LA models.

We have focused our analysis on the comparison of the predictive power of the RA-LA model and the RA model. This appears to be the most sensible comparison because the presence of risk aversion can lead players to stop experimenting before risk neutral agents and we aimed to insulate the additional explicative power of loss aversion in bandit settings. Nonetheless, we have further compared the 'LA-RA' model with the model that did not include risk aversion and loss aversion. Not surprisingly, we can conclude that the LA-RA model performs better and has a better predictive power than the standard bandit model with risk-neutral and loss-neutral agents.

4.4 The main differences between theoretical and empirical analysis

The analysis of the empirical results shows some differences with the theoretical predictions in Chapter 2. The main differences between the prediction and the experimental results lies in the bias toward excess experimentation in stages 1 and 2 of the game and towards less experimentation in stage 3 of the game.

We consider two alternative "behavioural" explanations, which we argue cannot rationalize these outcomes.

Candidate Interpretation 1: A failure induces subjects to excessively reduce the subjective probability of a success in the experimentation process.

Following this line of explanation we can rationalize the bias towards under-experimentation in stage 2.Note, however, that there should then be a sharp fall in expectation of p in stage 2 following a failure in stage 1, leading to under-experimentation. This is not the case in our experimental results. Hence, the proposed "behavioural" explanation can rationalize the tendency to under-experimentation in stage 3, but it does not appear to be consistent with the experimental data in stage 2, which shows excess experimentation.

Candidate Interpretation 2: Subjects have an optimistic prior p>1/2 to begin with. The presence of an optimistic probabilistic bias can rationalize the bias towards excess experimentation in stages 1 and 2 of the game. However, the presence of the optimistic bias is inconsistent with the bias towards under-experimentation in stage 3.

5. Conclusions

In Chapter 3 we address the issue of optimal experimentation in a traditional one-armed bandit problem adopting an experimental perspective. The main aim of the analysis is to investigate in which direction the presence of loss aversion affects the experimentation choices in experimental settings.

In total 254 subjects – first year undergraduate students in Microeconomics - took part in the experiment. The experimental study involved for each subject involved the elicitation of individual preferences (measuring loss aversion and risk aversion in a within-subject design) and the agreement to participate in experimental decisions. The set-up of the experiment is designed as a one armed-bandit problem with a three-period horizon.

The main experimental finding is that there is evidence of both over-experimentation and under-experimentation, and this is true for subjects of all level of risk aversion and loss aversion. At some stages in the process subjects of all types display over-experimentation, while at other stages there is a systematic bias towards under-experimentation.

Candidate Interpretation 2: Subjects have an optimistic prior p>1/2 to begin with. The presence of an optimistic probabilistic bias can rationalize the bias towards excess experimentation in stages 1 and 2 of the game. However, the presence of the optimistic bias is inconsistent with the bias towards under-experimentation in stage 3.

5. Conclusions

In Chapter 3 we address the issue of optimal experimentation in a traditional one-armed bandit problem adopting an experimental perspective. The main aim of the analysis is to investigate in which direction the presence of loss aversion affects the experimentation choices in experimental settings.

In total 254 subjects – first year undergraduate students in Microeconomics - took part in the experiment. The experimental study involved for each subject involved the elicitation of individual preferences (measuring loss aversion and risk aversion in a within-subject design) and the agreement to participate in experimental decisions. The set-up of the experiment is designed as a one armed-bandit problem with a three-period horizon.

The main experimental finding is that there is evidence of both over-experimentation and under-experimentation, and this is true for subjects of all level of risk aversion and loss aversion. At some stages in the process subjects of all types display over-experimentation, while at other stages there is a systematic bias towards under-experimentation.

Appendix 1

INSTRUCTIONS OF THE EXPERIMENT

The questionnaire is part of a scientific research project on how people make economic choices.

In the following questions there are no "right" or "wrong" answers. Your response should only reflect your own preferences. As the other parts of the questionnaire this following question is part of a scientific research project on how people make economic choices.

Part I

1. Today you can participate to each of the following lottery in turn that offers you a win a certain amount of money (in euro) with a given probability and 0 euro otherwise. Alternatively you can decide not to participate to the lottery and receive a sure payment of 1 euro.

Please circle your choice in each line depending on whether you choose to play the lottery or to collect the sure amount.

Sure amount	Lottery
euro 1.00 for sure	#1. 80% chance of euro 6.25 and 20% chance of euro 0
euro 1.00 for sure	#2. 50% chance of euro 10 and 50% chance of euro 0
euro 1.00 for sure	#3. 40% chance of euro 12.50 and 60% chance of euro 0
euro 1.00 for sure	#4. 20% chance of euro 25 and 80% chance of euro 0
euro 1.00 for sure	#5. 10% chance of euro 50 and 90% chance of euro 0
euro 1.00 for sure	#6. 5% chance of euro 100 and 95% chance of euro 0
euro 1.00 for sure	#7. 1% chance of euro 500 and 99% chance of euro 0

2. Will you take part to the following lottery?

   Please check a box for each line indicating whether you accept to participate in this gamble.

Lottery	Accept	Reject
1. If the coin turns up heads, you lose euro 2; if the coin turns up tails, you win euro 6		
2. If the coin turns up heads, you lose euro 3; if the coin turns up tails, you win euro 6		
3. If the coin turns up heads, you lose euro 4; if the coin turns up tails, you win euro 6		
4. If the coin turns up heads, you lose euro 5; if the coin turns up tails, you win euro 6		
5. If the coin turns up heads, you lose euro 6; if the coin turns up tails, you win euro 6		
6. If the coin turns up heads, you lose euro 7; if the coin turns up tails, you win euro 6		

Part II

You are in front of a slot machine. The slot machine pays out 10 euro with some probability. We don't know what the probability is, but we think it is ½.

You are allowed to have up to 3 tries on the slot machine, but you must pay for a ticket each time you do it. Tickets cost 3,5 euro each.

You can, of course, stop playing after one go or after two goes, if you wish.

1. Will you buy a ticket for 3,5 euro?

YES	NO

2. What is the highest price you would pay for a ticket?

   Note: the highest price means that if the ticket costs less or equal than this value, you are going to buy the ticket and play on the slot machine. If the ticket costs more than the highest value you wish to pay, you are not going to buy the ticket and you will not play

   Please circle your choice.

Ticket price (euro)
2
2,5
3
3,5
4
4,5
5
5,5

3. Having paid that highest price (selected in Question 2), you play on the slot machine. Suppose you played and at first trial you failed. Are you going to play on the slot machine a second time?

YES	NO

4. You played once along the slot machine. In the first trial you FAIL. At which price are you going to buy the second ticket to play with the slot machine?
   The following table contains a list of possible ticket prices. Please indicate for each ticket price whether you are going to buy the ticket and play on the slot machine a second time at the price stated.

   On the table please circle your choice in each line:

   YES means you buy the ticket at the stated price and play on the slot machine

   NO means you do not buy the ticket.

Ticket price (euro)		
2	YES	NO
2,5	YES	NO
3	YES	NO
3,5	YES	NO
4	YES	NO
4,5	YES	NO
5	YES	NO
5,5	YES	NO

5. You played twice along the slot machine. In the first trial you WIN and in the second trial you FAIL. Are you going to buy the third ticket to play with the slot machine?
   The following table contains a list of possible ticket prices. Please indicate for each ticket price whether you are going to buy the ticket and play on the slot machine a third time at the price stated.

   On the table please circle your choice in each line:

   YES means you buy the ticket at the stated price and play on the slot machine

   NO means you do not buy the ticket.

Ticket price (euro)		
2	YES	NO
2,5	YES	NO
3	YES	NO
3,5	YES	NO
4	YES	NO
4,5	YES	NO
5	YES	NO
5,5	YES	NO

6. You played twice along the slot machine. In the first trial you FAIL and in the second trial you WIN. Are you going to buy the third ticket to play with the slot machine?
   The following table contains a list of possible ticket prices. Please indicate for each ticket price whether you are going to buy the ticket and play on the slot machine a third time at the price stated.

   On the table please circle your choice in each line:

   YES means you buy the ticket at the stated price and play on the slot machine

   NO means you do not buy the ticket.

Ticket price (euro)		
2	YES	NO
2,5	YES	NO
3	YES	NO
3,5	YES	NO
4	YES	NO
4,5	YES	NO
5	YES	NO
5,5	YES	NO

7. You played twice along the slot machine. In the first trial you FAIL and in the second trial you FAIL. Are you going to buy the third ticket to play with the slot machine?
   The following table contains a list of possible ticket prices. Please indicate for each ticket price whether you are going to buy the ticket and play on the slot machine a third time at the price stated.

   On the table please circle your choice in each line:

   YES means you buy the ticket at the stated price and play on the slot machine

   NO means you do not buy the ticket.

Ticket price (euro)		
2	YES	NO
2,5	YES	NO
3	YES	NO
3,5	YES	NO
4	YES	NO
4,5	YES	NO
5	YES	NO
5,5	YES	NO

Appendix 2
The maximum ticket price

Tables XIV and XV show the average willingness to pay to play on the slot machine according to the distribution of loss aversion and risk aversion in the population of subjects. By inspecting Table XIV, subjects who are moderately loss averse (with a parameter λ equal to 1.5 and 1.2) are willing to pay more. This result does not hold when we study the distribution of maximum willingness to pay for ticket price as a function of risk aversion: the maximum willingness to pay is decreasing in the risk aversion parameter[54]. Table XVI shows the average maximum willingness to pay of subjects as a function of the two parameters of loss aversion and risk aversion. We partition subjects into three classes (low, medium and high) for risk aversion and loss aversion. Results show the tendency for a moderate loss averse subject to experiment more for any degree of risk aversion[55].

	Ticket price			
Loss Aversion parameter λ	Stage 1	Stage 2 (after 1 failure)	Stage 3 (after 1 failure and 1 success)	Stage 3 (after 1 success and 1 failure)
> 3	3.45	2.6	3.2	3.1
3	3.5	2.8	3.4	3.4
2	3.8	3.1	3.7	3.9
1.5	4.6	4.1	4.4	4.5
1.2	4.9	4.3	4.8	4.95
1	4.3	3.7	4.2	4.6

Table XIV - *Distribution of average willingness to pay on the basis of loss aversion λ*

[54] Anderson (2001) examined the issue of risk aversion in a bandit setting taking a rather different perspective. He focused the experimental design to study the suboptimal level of experimentation in a bandit setting as a function of the degree of risk aversion and to quantify the associated welfare loss. Here we do not investigate this issue e we do not claim subjects are undertaking optimal level of experimentation.

[55] We grouped for simplicity the seven classes of risk aversion and loss aversion identified in Section and we reduce them to "low", "medium" and "high" levels of r and λ. Partitioning subjects in three classes helps us to deliver results in a more intuitive way. Additionally, each class has a significant density, avoiding excessive dispersion of subjects across classes.

	Participation rate (ticket at 3.5)			
Risk Aversion parameter r	Stage 1	Stage 2 (after 1 failure)	Stage 3 (after 1 failure and 1 success)	Stage 3 (after 1 success and 1 failure)
5	2.7	2.3	3.25	3.4
4	3.3	2.8	3.5	3.5
3	3.6	3	3.6	3.7
2	3.8	3.3	3.8	4
1	4	3.5	4	4.1
0	4.2	3.7	4.3	4.4

Table XV - *Distribution of average willingness to pay on the basis of risk aversion*

		Loss aversion		
		Low	Medium	High
Risk aversion	High	2.96	3.75	2.7
	Medium	3.8	4.15	3.2
	Low	4.2	4.65	3.62

Table XVI - *Distribution of average willingness to pay on the basis of loss aversion (row) and risk aversion (column)*

Chapter 4: Overconfidence in Tort Law: General Principles

"Human beings tend to be optimistic.
By itself this seems to be good news;
but it can lead them to make big mistakes."

Cass Sunstein[1]

1. Introduction

Behavioural biases – such as loss aversion, endowment effect and others – have important effects on the optimality of economic choices. In the previous chapters we have investigated the role of loss aversion in experimentation settings and we have characterized the optimal experimentation choice of a loss averse experimenter.

In Chapter 4[2] we consider an application of behavioural biases to the field of tort law, adopting a law and economics approach. The economic analysis of tort law characterizes the care incentives of a potential tortfeasor, defined as an economic agent who undertakes a risky activity that may cause an injury to another agent, called victim, and chooses which level of care to undertake in order to reduce the likelihood of an injury. In many situations, not only potential injurers but also potential victims may undertake care in order to make the injury less likely. Liability regimes provide different care incentives to potential injurers and victims, since they shield parties from liability in different ways. The economic analysis of tort law studies how the liability regimes affect the optimal care incentives of tortfeasors (in a unilateral care model) and victims (in the bilateral one).

Loss aversion plays a role in the field of tort law. Loss aversion means that a negative deviation to a certain reference point, a loss, counts more than a positive deviation (gain) even though in absolute terms they are the same. An implication is that people treat opportunity costs differently than "out-of-pocket" costs. Foregone gains are less painful than perceived losses. This

[1] Sunstein (2000, p. 4)

[2] This chapter is based on joint work with Francesco Parisi.

perception is strongly manifested in people's judgments about fair behaviour. Along this line, Sunstein (2000) points out that "loss aversion raises questions about the goals of the tort system. Should damages measure the amount that would restore an injured party to the status quo ex ante, or should they reflect the amount an injured party would demand to be subject to the injury before an act? Juries appear to believe that the amount that would be demanded pre-injury is far greater than the amount that would restore the status quo ex ante. The legal system appears generally to see compensation question as the latter one, though it does not seem to have made this choice in any systematic ways." Cohen and Knetsch (1990) showed that this principle, embodied in the old expression that "possession is nine tenths of the law," is reflected in many judicial opinions. For example, in tort law judges make the distinction between "loss by way of expenditure and failure to make gain." In one case, several bales fell from the defendant's truck and hit a utility pole, cutting off power to the plaintiff's plant. The plaintiff was able to recover wages paid to employees which were considered "positive outlays" but could not recover lost profits which were merely "negative losses consisting of a mere deprivation of an opportunity to earn an income" (p. 18). A similar distinction is made in contract law. A party that breaches a contract is more likely to be held to the original terms if the action is taken to make an unforeseen gain than if it is taken to avoid a loss.

The definition of loss matters when potential tortfeasors display loss aversion. Bigus (2005) shows that the care incentives of a potential injurer remain optimal if only damage payments are considered as a loss, thereby obtaining similar results to Expected Utility Theory. Additional efficient care incentives are preserved even if both liability payments and costs of care are considered as a loss. This occurs even under the rule of strict liability, while it is not true in the case of risk aversion in an Expected Utility Theory framework. The presence of loss aversion does not undermine care incentives in neither liability regimes, while on the contrary it enhances them above the social optimum. Hence loss aversion is not a primary source of inefficiency in a tort problem.

We turn our attention to other behavioural biases that may impact on optimal choices in the field of tort law. Psychological research shows that there is systematic overconfidence in risk judgments. Data shows that overconfidence bias is pervasive in judgments that individuals make regularly in everyday life, rather than in activities that are seldom carried out (Jolls and Sunstein, 2006). Overconfidence creates a distinctive problem for legal policymakers: even factually informed people tend to think that risks are less risky to materialize for themselves than for others (Sunstein, 2000).[3] Overconfident people therefore inadequately react to legal threats and incentives such as liability rules in a number of areas of law, showing some surprising implications.

[3] Sunstein (2000) considers overconfidence and the role of law in debiasing this judgment error, considering the special challenges posed by the fact that the vast majority of people believe that they are less likely than other people to be

Behavioural law and economics scholars have addressed the issue of overconfidence, considering the possible role of law in restraining or correcting this judgment error.[4] In Chapter 4 we consider the role of tort rules in debiasing overconfidence, unveiling another surprising and counterintuitive implication. The most effective way to correct overconfidence in tort law could well be to forgive it, rather than to penalize it through liability. Chapter 4 is structured as follows. In Section 2, we provide a brief introduction to the issue if overconfidence with reference to the psychological and behavioural literature and review the previous work in law and economics considering the effect of judgment errors under different liability regimes. In Section 3, we consider alternative legal strategies to correct overconfidence problems in tort law. We illustrate the possible use of threat strategies (threatening liability when overconfidence leads to an accident) and forgiveness strategies (forego liability when the accident is solely caused by a biased perception of risk) as alternative ways in which law could be used to debias overconfidence. In Section 4, we build a simple model of bilateral care to describe the effects of a bias in the perception of the probability of harm and consider the effect of overconfidence on parties' incentives. In Section 5, we compare the alternative threat and forgiveness strategies in reducing the cost of accidents due to overconfidence, allowing for the possibility that government investment in information fails to guarantee full debiasing of agents. Under each liability rule we characterize the care and activity levels chosen by injurer and victim in the presence of overconfidence and we rank each combination of rule/strategy according to the efficiency level that it will induce. The model highlights the role of tort law and the optimal design of liability rules for correcting overconfidence biases. We compare the effectiveness of alternative legal strategies under alternative liability rules, unveiling an interesting paradox: legal forgiveness of overconfidence may be a valuable second-best solution, when debiasing through information and threat strategies proves ineffective. The most effective way to correct overconfidence in tort law may well be to forgive it, rather than to penalize it through liability. Products regulations that impose liability on producers for not increasing the safety of their products in anticipation of consumer's overconfidence can be viewed as a legal strategy germane to legal forgiveness of overconfidence (consumer are effectively shielded against

subject to automobile accidents, infection from AIDS, heart attacks, asthma, and many other health risks, even though they do not lack statistical information about these risks in general.

[4] A growing body of law and economics literature focuses on the departures of human behavior from full rationality and attempt to explain the positive and normative implications of bounded rationality in the formulation of legal policy. Sunstein (1997) and Jolls, Sunstein, and Thaler (1998) points out the need of a more accurate understanding of behavior and individual choice in legal context, in order to take into account the shortcomings in human behavior when structuring the law. Jolls and Sunstein (2006) discuss the idea of "debiasing through law", instead of "debiasing law", i.e. to insulate legal outcomes from the effects of boundedly rational behavior. Debiasing through law is instead aimed at developing legal strategies attempting to reduce or eliminate boundedly rational behavior. Jolls and Sunstein provide a general description of debiasing through law with application to many areas, as consumer safety law, corporate law and property law.

the consequences of their overconfident errors). Legal forgiveness of overconfidence has not yet been considered, however, as a general rule in tort law, something that this chapter attempts to do. Section 6 concludes considering the optimal scope of forgiveness under different tort scenarios.

2. Overconfidence and Unrealistic Optimism

Overconfidence is known to psychologists as one of the most widespread psychologically-generated biases in human judgment. Overconfident individuals overestimate their own ability underestimating the risk that they face. Overconfidence is one of the two biases that result from what psychologists know as optimism bias.[5] Optimism bias affects people's subjective estimates of the likelihood of future events, and causes them to overestimate the likelihood of positive or desirable events and to underestimate the likelihood of negative or undesirable events (Colman, 2001). Optimism bias was reported by psychologists starting from the early twentieth-century (Lund, 1925; and Cantril 1938) and was rigorously studied and documented by the US psychologist Neil David Weinstein, who gave it the name of "unrealistic optimism" in a 1980 article published in the *Journal of Personality and Social Psychology*.[6]

Several subsequent psychological studies confirm the stylized fact that people exhibit an unrealistic optimism bias (see the survey of the literature in Wenglert and Rosén, 2000).[7] An interesting example of overconfidence for the purpose of this chapter is the finding by Svenson (1981) where most survey respondents rated themselves as better and more competent drivers than average. Overconfidence appears to be robust with respect to a variety of accident risks (see, among others, Sunstein, 1997 and Jolls, 1998).[8] Studies on traffic accidents show that people's assessment of accident risks faced by others is fairly accurate (Lichtenstein et al. 1978) and at the same time people tend to be unrealistically optimistic about themselves, underestimating the likelihood that

[5] There are two forms of judgment bias that follow from unrealistic optimism: overconfidence bias which implies an overestimation of one's own ability, and self-serving bias which is a tendency to evaluate evidence or make judgments in a way that benefits oneself (Muren, 2004).

[6] Weinstein (1980) asked students to estimate the likelihoods of various events happening to them and showed that they rated their chances of experiencing positive (negative) events significantly above (below) the average for their peers.

[7] Economic experiments confirm the existence of optimism and related overconfidence bias. Forsythe, Rietz and Ross (1999) find experimental evidence of the human tendency to overestimate the probability of desirable events. Babcock, Loewenstein, Issacharoff and Camerer (1995) find evidence of a self-serving bias in an experiment where subjects were given roles as plaintiffs and defendants in a legal dispute over a claim for damages. Kaplan and Ruffle (2001) argue that strategic behavior may affect the measure of these biases.

[8] Empirical evidence suggests that optimism bias affects not only the perception of risky events, but it affects other choices, such as litigation choices. It has been shown that lawyers and litigants are systematically optimistic with respect to the outcome at trial. Bar-Gill (2006) studies the persistence of optimism bias in litigation using a setting of evolutionary game theory. The adaptive force of optimism is linked to its instrumental value in the pre-trial bargaining: optimistic lawyers can credibly threaten to resort to costly litigation, and are therefore more successful in extracting bargaining surplus from settlements.

they will cause an accident (see Svenson 1981; Svenson, Frischhoff, and MacGregor 1985; Finn and Bragg 1986; Matthews and Moran 1986; DeJoy 1989; McKenna, Stanier and Lewis 1991; Guppy 1992; Jolls, 1998).[9]

The traditional law and economics view on issues of overconfidence is that the problem is caused by imperfect information and can be appropriately corrected through the provision of additional information (see, e.g., Stiglitz, 1986 on consumer optimism). However, as the extensive evidence suggests, overconfidence leads many individuals to underestimate their personal risks even if they receive accurate information about average risks. Evidence indeed suggests that debiasing strategies through risk education and information are only partially effective. As pointed out by Jolls and Sunstein (2006), educated people could hold accurate statistical information about the probability of contracting cancer, and yet believe that they are less likely than most people to contract it. As Viscusi (2002) puts it, people may accurately learn the aggregate statistical risk, but their information only has limited effect on the optimistic "above average" illusion that leads them to underestimate the actual risk that they face.

Legal scholars have introduced the concept of debiasing through law, both with reference to procedural rules governing adjudication (Babcock, Loewenstein, Issacharoff and Camerer, 1995; and Babcock, Loewenstein and Issacharoff, 1997) and substantive law (Jolls and Sunstein, 2006). The idea of debiasing through law, as most notably presented by Jolls and Sunstein (2006), is that the design of substantive law should consider the need to correct the systematic judgment errors of individuals.[10] A number of papers studies tort law models in non-expected utility framework. Surprisingly, even though there is a vast literature on the economics of tort law, new models of decision-making have hardly been incorporated yet (Shavell, 1987 and 2004). Specifically, Eide (2007) analyzes the basic tort law model under rank-dependent expected utility theory developed by Quiggin (1982 and 1993). Bigus (2006) analyzes the basic tort law model using Kahneman and Tversky's (1979) prospect theory. Teitelbaum (2007) develops a tort model based on Choquet's (1954) expected utility. In all these models inefficient levels of cares may be undertaken depending on the slope of the probability function under both strict liability and negligence. Teitelbaum finds that the injurer's level of care decreases (increases) with ambiguity if he is optimistic (pessimistic)

[9] Hanson and Kysar (1999a and 1999b) report evidence of the resilience of consumer optimism, also in the face of risk disclosure and product warnings. Jolls (1998) further reports that people underestimate their own likelihood of being victims of natural disasters, as evidenced by their frequent failure to buy insurance for floods and earthquakes. Armour and Taylor (2002) discuss the case of professional financial experts, who consistently overestimate expected earnings, and business school students who overestimate starting salaries and probabilities of employment at graduation. According to Jolls and Sunstein (2006), this evidence suggests that people sometimes exhibit optimism bias not only with respect to relative risk, but also in the estimation of actual probabilities in their category of profession.

[10] An important concern in the legal literature is to avoid the adoption of legal solutions that restrict choices or impose significant costs on unbiased individuals in order to improve the outcome for those who exhibit cognitive biases (Camerer, Issacharoff, Loewenstein, O'Donoghue and Rabin 2003; Mitchell 2002).

and decreases (increases) with his degree of optimism (pessimism). The results suggest that negligence is more robust to ambiguity and, therefore, may be superior to strict liability in unilateral accident cases.[11] A relevant study of overconfidence errors in the field of tort law comes from Posner (2003), who develops a positive economic model to study the effect of optimism for rare events (i.e., low probability accidents) in unilateral care cases. In Posner's framework, agents know the probability of an accident when it is above some threshold but set accident probabilities to zero if it is below the threshold. Posner shows that under both strict liability and negligence, agents take inefficient level of care (either too much or too little) for sufficiently high levels of optimism and take optimal care for sufficiently low levels of optimism. Posner additionally shows that the difference on the effect on activity levels between strict liability and negligence rule tends to disappear in case of optimism, due to the fact that the optimistic agent treats the rare events as a zero-probability event.

Bigus (2006) analyzes unilateral tort liability assuming that the tortfeasor evaluates probabilities according to Prospect Theory (Kahneman and Tversky, 1979), in the sense that the tortfeasor underestimates very high probabilities and overestimates very low ones. When the liability rule imposes a precise standard of due care, the tortfeasor faces lower incentives to take care than in the EU-model, since the probability weighting decreases the *marginal* benefits of taking care. This holds for strict liability and for a negligence rule with a precise standard of due care.

Often, the standard of due care is not precisely defined. With a vague standard, probability weighting mitigates the problem of overdeterrence that exists with the EU-model, but there might be an opposite second effect which induces the tortfeasor to choose a care level where there is no liability *for sure* but which is inefficiently high. The overall outcome is not clear.

Empirical evidence suggests that individuals do not necessarily follow this concept of rationality. Still, one might argue that institutional actors may provide mechanisms to induce their agents to behave rationally and to learn over time. Yet, the evidence for this claim is mixed at best and shows that even institutional actors might be subject to bounded rationality (see e.g., Odean, 1998). Rabin (1998) points out that learning effects in reality are considerably more limited than economists usually assume.

[11] From a different non-behavioral perspective other law and economics scholars have dealt with the issue of court errors and the effect of misperception and errors on expected liability and due care. Among others, Dari-Mattiacci (2005) shows that errors may distort care incentives in different ways, depending on whether they occur under rules (when a regulator defines due care and courts set the amount of damages) or standards (when courts set both damages and due care on a case-by-case basis). The two regimes differ in their capacity to insulate the effects of errors. Under rules, errors in determining damages and in setting due care occur independently of each other, while, under standards, an error in damages may trigger a corresponding error in due care.

decreasing returns. Hence we assume that the tortfeasor's precautions decrease the probability of an environmental damage, $p_x < 0$, at a decreasing rate, $p_{xx} > 0$, for all values of w. Similarly, the victim can decrease the likelihood of harm through precaution, $p_y < 0$, at a decreasing rate, $p_{yy} > 0$, for all values of z. Hence, increasing care is always costly to the parties and, in the relevant range, increasing activity levels is beneficial. When an accident occurs, it creates an (exogenous) loss denoted by L,[13] where $L > 0$. Let $p(x,y)L$ be the expected harm per unit of activity. Total expected harm is assumed to be $wzp(x,y)L$.

The lawmaker cannot intervene directly by taking precautions on its own in order to reduce the accident risk. However it can affect the tortfeasor's and victim's level of precautions and activity by recurring to alternative strategies. We shall examine these alternative strategies in Section 3.

3.1 Overconfidence: An Analytical Outlook

Overconfidence is a distortion of perception, characterized by an overestimation of one's own ability and an underestimation of risk. In a typical tort law problem, overconfidence can be formally described by reference to the individual's distorted perception of a probability of an accident. Let us indicate with $p(x,y)$ the unbiased probability function of an accident. Overconfidence induces an optimistic bias in the perception of the accident probability. An overconfident individual underestimates risk and believes that he faces a lower probability $\underline{p}(x,y)$ of being involved in an accident than the unbiased probability:

$$\underline{p}(x,y) < p(x,y)$$

For any level of care of tortfeasor and victim, an overconfident agent considers less likely for an accident to occur compared to an unbiased agent.[14] Additionally, the biased probability function obeys the following relationship:

$$\left|\underline{p_x}(x,y)\right| < |p_x(x,y)|$$

That is, given the lower probability estimate of an overconfident agent, care investments are perceived to be less valuable, yielding a lower absolute reduction in the probability of an accident

[13] As it is standard in the literature, we are assuming that the injurer can only affect the probability of the harm, but not its magnitude. This assumption simplifies the model without loss of generality. As shown by Dari Mattiacci and De Geest (2005), the impact of insolvency on the incentives to take precaution is not qualitatively changed when the magnitude of the harm is endogenous.

[14] Qualitatively similar results would be reached by assuming that overconfidence affected the perception of the gravity of the loss, rather than its probability, with endogenous harm.

In the following, we build on this literature, developing a model to highlight the role of overconfidence bias in bilateral accident cases. We use these results to formulate some normative corollaries for the optimal design of liability rules. We consider the alternative use of "threat strategies" (imposing liability when accidents are the result of a biased perception of risk) and "forgiveness strategies" (foregoing liability when accidents are caused by such perception bias) in order to reduce the cost of accidents due to overconfidence. These two strategies are considered in combination with possible debiasing through regulation, where the government attempts to correct the parties' perceptions of risk through education, disclosure and warning about statistical risk and severity of harm. Our results suggest that legal forgiveness of overconfidence may be a valuable second-best solution, when debiasing through information and threat strategies proves ineffective.

3. A Bilateral Accident Problem with Overconfidence

In this section, we develop a simple model of bilateral care as in Shavell (1987) to describe the effects of a bias in the perception of the probability of harm, and to study the effectiveness of alternative legal strategies under alternative liability rules.[12] The tortfeasor carries out an activity, with a value equal to $V_T(w)$, where w denotes the injurer's activity level. We assume that the value of the activity increases with the activity level in the relevant range, $V_w > 0$, at a decreasing rate, $V_{ww} < 0$. Likewise, the victim carries out an activity, with a value equal to $V_V(z)$, where z denotes the victim's activity level. Assume that $V_V(z)$ has similar properties to the injurer's benefit function: the value of the victim's activity increases with the activity level in the relevant range, $V_z > 0$, at a decreasing rate, $V_{zz} < 0$.

The activity of the tortfeasor may cause harm. The tortfeasor can invest in precautions to reduce the probability of such harm. Denote with x the tortfeasor's level of precaution per unit of activity w, where $x \in [0, \infty)$. Likewise, the victim bears the harm but is able to reduce the probability of its occurrence with her own precautions. Denote with y the victim's level of precaution per unit of activity z, where $y \in [0, \infty)$. With a level of tortfeasor's precaution x and a level of victim's precaution y, damage occurs with probability $p(x, y)$, where $p(x, y) \in (0,1)$. We assume bilateral precautions, such that the probability of the environmental damage is effectively controlled by the tortfeasor's and victim's level of precautions, and alternative care where the parties' precautions are substitutes, $p_{xy} < 0$. Increasing care is costly to the injurer and leads to

[12] The model is based on Shavell (1987), with one main difference: the expected loss is written as the product of a given loss L and the probability of an accident, as in Brown (1973).

decreasing returns. Hence we assume that the tortfeasor's precautions decrease the probability of an environmental damage, $p_x < 0$, at a decreasing rate, $p_{xx} > 0$, for all values of w. Similarly, the victim can decrease the likelihood of harm through precaution, $p_y < 0$, at a decreasing rate, $p_{yy} > 0$, for all values of z. Hence, increasing care is always costly to the parties and, in the relevant range, increasing activity levels is beneficial. When an accident occurs, it creates an (exogenous) loss denoted by L,[13] where $L > 0$. Let $p(x,y)L$ be the expected harm per unit of activity. Total expected harm is assumed to be $wzp(x,y)L$.

The lawmaker cannot intervene directly by taking precautions on its own in order to reduce the accident risk. However it can affect the tortfeasor's and victim's level of precautions and activity by recurring to alternative strategies. We shall examine these alternative strategies in Section 3.

3.1 Overconfidence: An Analytical Outlook

Overconfidence is a distortion of perception, characterized by an overestimation of one's own ability and an underestimation of risk. In a typical tort law problem, overconfidence can be formally described by reference to the individual's distorted perception of a probability of an accident. Let us indicate with $p(x,y)$ the unbiased probability function of an accident. Overconfidence induces an optimistic bias in the perception of the accident probability. An overconfident individual underestimates risk and believes that he faces a lower probability $\underline{p}(x,y)$ of being involved in an accident than the unbiased probability:

$$\underline{p}(x,y) < p(x,y)$$

For any level of care of tortfeasor and victim, an overconfident agent considers less likely for an accident to occur compared to an unbiased agent.[14] Additionally, the biased probability function obeys the following relationship:

$$\left|\underline{p_x}(x,y)\right| < |p_x(x,y)|$$

That is, given the lower probability estimate of an overconfident agent, care investments are perceived to be less valuable, yielding a lower absolute reduction in the probability of an accident

[13] As it is standard in the literature, we are assuming that the injurer can only affect the probability of the harm, but not its magnitude. This assumption simplifies the model without loss of generality. As shown by Dari Mattiacci and De Geest (2005), the impact of insolvency on the incentives to take precaution is not qualitatively changed when the magnitude of the harm is endogenous.

[14] Qualitatively similar results would be reached by assuming that overconfidence affected the perception of the gravity of the loss, rather than its probability, with endogenous harm.

loss.[15] This means that from the point of view of a biased agent, the marginal benefit that is perceived to be gained from an additional unit of care falls below that of an unbiased agent. Additionally, in our model we incorporate the stylized fact shown in the evidence discussed in the Section 1, according to which a representative agent is aware of the overconfidence bias of the population, but not fully aware of his own overconfidence bias.[16]

3.2 Information and Other Debiasing Techniques

To the extent that it results from imperfect information, overconfidence can be corrected by providing individuals with additional information about the specific risk they face (Stiglitz, 1986). The government may invest in policies and regulations aimed at correcting overconfidence, through education about risk, disclosure of information and other debiasing techniques.[17] Most systems undertake *debiasing strategies*, according to which the policymaker provides education about risk and statistical information about riskiness of activities, likelihood of cognitive lapses (e.g., falling asleep while driving, etc.) in order to eliminate or reduce the overconfidence bias. However, as the extensive evidence suggests, debiasing strategies through risk education and information are costly and only partially effective: knowledge of the statistical incidence of overconfidence only minimally affects the awareness of one's own overconfidence bias (Viscusi, 2002; Jolls and Sunstein, 2006). Hence governmental investments in information give only small returns in accident reduction.

In the following, we allow for the possibility that, when provided with statistical information and education about risks, parties correct their estimation of the likelihood of the accident only partially. Let's denote as α, the level of effectiveness of debiasing through information. After receiving information, the probability estimation of the overconfident individual is partially

[15] The opposite case of pessimism bias would be characterized by a higher estimation of the likelihood of a negative event. The case of pessimism bias is not treated in the chapter inasmuch as it does not represent a frequently observed bias in standard situations. As it will be briefly discussed in the conclusions, the application of a framework of legal forgiveness of underconfidence would not be readily applicable and would not correct the inefficiency induced by the excessively high levels of care exerted by pessimistic agents (unless we could conceive of a liability rule penalizing agents that undertake excessively high levels of care).

[16] Note that here that agents have a correct perception of the average bias in the population. Henceforth the distortion in beliefs has to be interpreted as the one exhibited on average in the population.

[17] For example, social scientists have extensively examined the prospects for debiasing of individuals (Fischhoff, 1982; Sanna, Schwarz and Stocker, 2002; and Weinstein and Klein, 2002). Among various approaches, debiasing with the instrumental use of other cognitive biases has been considered in the literature. Sherman, Cialdini, Schwartzman and Reynolds (2002) note that making an occurrence "available" increases individuals' estimates of the likelihood of the occurrence, hence the availability bias could be used to correct for the optimism bias in some instances. Likewise, framing the information to emphasize the gravity of prospective losses may counteract excessive optimism. An opposite pessimism bias is instead exhibited with respect to the risk of accidents that are either salient, catastrophic or technological in nature (see Sunstein 1997; Jolls 1998; Jolls, Sunstein, and Thaler 1998; see also Slovic, Frischhoff, and Lichtenstein 1982; Viscusi and Magat 1987; Viscusi 1992).

debiased, becoming a weighted average of the prior biased estimate, $\underline{p}$, and the unbiased probability, p:

$$p^I = \alpha p + (1 - \alpha)\underline{p}$$

When debiasing through information is perfectly effective, $\alpha = 1$, the agent makes his optimal choices of care and activity based on an unbiased estimate of probability, p. If the government makes no investment in information or debiasing strategies are totally ineffective, $\alpha = 0$, the agent chooses care and activity levels based on the biased estimate of probability, $\underline{p}$. When $0 < \alpha < 1$, debiasing through information is only partial and the agent makes his optimal choices of care and activity based on a partially biased estimate of the risk, p^I. Higher values of α denote higher levels of debiasing, with an upper threshold of perfectly effective debiasing, where the subjective estimate of the risk corresponds to the actual probability, $p^I = p$.

4. Two Legal Strategies to Correct Overconfidence: Threat and Forgiveness

As discussed in Section 1 above, according to empirical and experimental evidence, the knowledge about statistical incidence of overconfidence only has a limited effect on the belief about one's own overconfidence bias. In this section we illustrate the possible use of threat and forgiveness strategies, allowing for the possibility that debiasing through information may be partially or entirely ineffective. We consider the effect of alternative liability rules under both threat and forgiveness regimes in correcting the inefficiency in care and activity levels caused by the residual overconfidence bias. Under each liability rule we characterize the care and activity levels chosen by injurer and victim and we rank each combination of rule/strategy according to the efficiency level that it will induce. We assume that parties (tortfeasor and victim) are risk neutral, rational and utility maximizing.

We consider two alternative strategies to address overconfidence in tort law: (a) threat strategies, and (b) forgiveness strategies. Most legal systems utilize debiasing strategies in combination with a *threat strategy*, according to which a prospective tortfeasor faces a threat of liability if an accident occurs due to an optimism bias when statistical facts and riskiness of the activity were known (or should have been known) by the tortfeasor. Threat strategies construe overconfidence errors as negligence: if the tortfeasor falls short of adopting the level of care of a fully rational/unbiased agent, liability is imposed.

debiased, becoming a weighted average of the prior biased estimate, $\underline{p}$, and the unbiased probability, p:

$$p^I = \alpha p + (1 - \alpha)\underline{p}$$

When debiasing through information is perfectly effective, $\alpha = 1$, the agent makes his optimal choices of care and activity based on an unbiased estimate of probability, p. If the government makes no investment in information or debiasing strategies are totally ineffective, $\alpha = 0$, the agent chooses care and activity levels based on the biased estimate of probability, $\underline{p}$. When $0 < \alpha < 1$, debiasing through information is only partial and the agent makes his optimal choices of care and activity based on a partially biased estimate of the risk, p^I. Higher values of α denote higher levels of debiasing, with an upper threshold of perfectly effective debiasing, where the subjective estimate of the risk corresponds to the actual probability, $p^I = p$.

4. Two Legal Strategies to Correct Overconfidence: Threat and Forgiveness

As discussed in Section 1 above, according to empirical and experimental evidence, the knowledge about statistical incidence of overconfidence only has a limited effect on the belief about one's own overconfidence bias. In this section we illustrate the possible use of threat and forgiveness strategies, allowing for the possibility that debiasing through information may be partially or entirely ineffective. We consider the effect of alternative liability rules under both threat and forgiveness regimes in correcting the inefficiency in care and activity levels caused by the residual overconfidence bias. Under each liability rule we characterize the care and activity levels chosen by injurer and victim and we rank each combination of rule/strategy according to the efficiency level that it will induce. We assume that parties (tortfeasor and victim) are risk neutral, rational and utility maximizing.

We consider two alternative strategies to address overconfidence in tort law: (a) threat strategies, and (b) forgiveness strategies. Most legal systems utilize debiasing strategies in combination with a *threat strategy*, according to which a prospective tortfeasor faces a threat of liability if an accident occurs due to an optimism bias when statistical facts and riskiness of the activity were known (or should have been known) by the tortfeasor. Threat strategies construe overconfidence errors as negligence: if the tortfeasor falls short of adopting the level of care of a fully rational/unbiased agent, liability is imposed.

In the following we consider the possible combined use of tort law remedies and *forgiveness strategies* in correcting judgment error. Forgiveness strategies do not construe overconfidence as per se negligence. Overconfidence is taken into consideration as an objective and resilient phenomenon in human nature. With legal forgiveness of overconfidence, a prospective tortfeasor does not face liability if he adopts reasonable care, in response to his assessment of the probability of an accident, even if the subjective assessment of the risk was distorted by some natural degree of overconfidence. It is important to note that in order to avoid opportunistic reliance on legal forgiveness, forgiveness strategies should only adjust the legal expectation of due care in light of the average level of overconfidence exhibited by the average person under similar circumstances. Similar to the evaluation of objective fault under negligence, courts should not engage in any subjective inquiry into the parties' state of mind at the time of the accident, but only account for the average overconfidence bias in the population to avoid a moral hazard problem or a strategic exacerbation of one's own overconfidence.[18] Under forgiveness strategies, liability arises only when the tortfeasor's care falls short of the level of care that would have been reasonable, given an assessment of the likelihood of an accident carried out by a person with an average level of overconfidence. Forgiveness strategies, albeit counterintuitive, already finds some incarnations in existing law and are aimed at shifting the burden of correcting for the overconfidence bias on those who are in the best position to do so. For example, products liability regulation shifts the burden of correcting consumers' overconfidence bias on producers. In the face of resilient overconfidence biases on the part of consumers, heightened standards of products liability have been chosen, and liability is imposed on producers who do not adjust the safety of their products in anticipation of consumer's overconfidence errors.[19]

In the following, we compare the effect of threat and forgiveness strategies when used in combination with debiasing strategies, showing the comparative advantage of each strategy in correcting the distortions caused by the overconfidence bias under different liability regimes.

5. Results under Alternative Liability Regimes

In order to establish a benchmark of comparison, we derive the equilibrium configuration for a social optimum and we state the results of the model in absence of overconfidence, which

[18] Subjective test of overconfidence, if at all desirable should be limited to the proof of absence of bias, to avoid that unbiased parties could opportunistically rely on forgiveness and undertake a lower level of due care.

[19] A vast number of federal and state laws that regulate the safety of consumer products are driven by desire to protect consumers who do not adequately appreciate the risk of products. For further discussion, see Jolls and Sunstein (2006).

5.2 Results without Overconfidence

As stated in Section 2.2, in the absence of overconfidence (i.e., lack of individual bias or as a result of perfect debiasing through information, $\alpha = 1$), the model reduces to the standard bilateral accident model, as in Shavell (1987). The results of the model without overconfidence coincide with those of the standard bilateral accident model and are summarized in Proposition 1. This provides a benchmark to analyze the role of overconfidence on the privately optimal choice of care and activity.[20]

Proposition 1. *(i) Care Level. Under all liability rules, except for the case of strict liability, both parties exercise the socially optimal level of care. Under a strict liability only the prospective tortfeasor exercises care; (ii) Activity Level. Under all negligence-based rules, the victim undertakes the socially optimal activity level, and the tortfeasor undertakes an inefficiently high activity level. Under all strict-liability-based rules, the tortfeasor undertakes the socially optimal activity level, and the victim undertakes an inefficiently high activity level.*

Except for the case of strict liability, the level of care exerted by both parties is socially efficient at equilibrium under any liability rule. The level of activity undertaken instead is socially efficient only for the residual bearer of liability and is inefficiently high for the other party. Henceforth, in negligence-based regimes, the victim bears the residual liability and undertakes a socially optimal activity level, and in strict-liability-based regimes, the tortfeasor is instead the residual bearer who will have incentive to mitigate its activity level.

5.3 Results with Overconfidence

We characterize the optimal behaviour of the tortfeasor and the victim under overconfidence. When information strategies are perfectly effective, debiasing of overconfidence takes place and liability rules will generate the standard results, as in Section 4.2. In the following, we assume that debiasing through information can only be partially effective ($0 \leq \alpha < 1$). We analyze the role of the alternative threat and forgiveness strategies defined in Section 3 on privately optimal care and activity levels under alternative liability regimes. Specifically, we will consider the comparative advantage of threat and forgiveness strategies in inducing efficient incentives under (a) simple

[20] The model without overconfidence coincides with the standard model of bilateral accident: proofs are therefore omitted for brevity. See Shavell (1987) for details of the analytical proof.

correspond to the standard model of bilateral accident. We characterize the results of the model in the presence of overconfidence under different liability regimes.

5.1 Benchmark Case: Socially Optimal Level of Care and Activity

According to Shavell (1987), the social goal is to maximize total welfare, defined as the sum of the value of the activities for tortfeasor and victim at the net of the expected accident costs and precaution costs for both parties:

$$S = V_T(w) + V_V(z) - wzp(x, y)L - wx - zy$$

The socially optimal values of care x and y are implicitly defined by the first order conditions

$$-zp_x(x^*, y)L = 1 \qquad (1)$$

$$-wp_y(x, y^*)L = 1 \qquad (2)$$

where the left-hand-side of (1) and (2) represent the marginal social benefit of care (in terms of reduced probability of an accident loss) and where the right-hand-side represents the marginal social cost of care, respectively for the tortfeasor and for the victim. At the social optimum equations (1) and (2) require that the marginal reduction in the expected accident loss equals the marginal cost of care.

The socially optimal levels of activity w and z satisfy the following first order conditions:

$$V_w = zp(x, y)L \qquad (3)$$

$$V_z = wp(x, y)L \qquad (4)$$

Equations (3) and (4) require that the marginal benefit from an increase in activity level equals the marginal cost of the activity, given by an increase in expected accident loss.

In the following, we shall denote x^{**} and w^{**} as the socially optimal levels of precaution and activity for the tortfeasor, and y^{**} and z^{**} as the socially optimal levels of precaution and activity for the victim.

5.2 Results without Overconfidence

As stated in Section 2.2, in the absence of overconfidence (i.e., lack of individual bias or as a result of perfect debiasing through information, $\alpha = 1$), the model reduces to the standard bilateral accident model, as in Shavell (1987). The results of the model without overconfidence coincide with those of the standard bilateral accident model and are summarized in Proposition 1. This provides a benchmark to analyze the role of overconfidence on the privately optimal choice of care and activity.[20]

Proposition 1. *(i) Care Level. Under all liability rules, except for the case of strict liability, both parties exercise the socially optimal level of care. Under a strict liability only the prospective tortfeasor exercises care; (ii) Activity Level. Under all negligence-based rules, the victim undertakes the socially optimal activity level, and the tortfeasor undertakes an inefficiently high activity level. Under all strict-liability-based rules, the tortfeasor undertakes the socially optimal activity level, and the victim undertakes an inefficiently high activity level.*

Except for the case of strict liability, the level of care exerted by both parties is socially efficient at equilibrium under any liability rule. The level of activity undertaken instead is socially efficient only for the residual bearer of liability and is inefficiently high for the other party. Henceforth, in negligence-based regimes, the victim bears the residual liability and undertakes a socially optimal activity level, and in strict-liability-based regimes, the tortfeasor is instead the residual bearer who will have incentive to mitigate its activity level.

5.3 Results with Overconfidence

We characterize the optimal behaviour of the tortfeasor and the victim under overconfidence. When information strategies are perfectly effective, debiasing of overconfidence takes place and liability rules will generate the standard results, as in Section 4.2. In the following, we assume that debiasing through information can only be partially effective ($0 \leq \alpha < 1$). We analyze the role of the alternative threat and forgiveness strategies defined in Section 3 on privately optimal care and activity levels under alternative liability regimes. Specifically, we will consider the comparative advantage of threat and forgiveness strategies in inducing efficient incentives under (a) simple

[20] The model without overconfidence coincides with the standard model of bilateral accident: proofs are therefore omitted for brevity. See Shavell (1987) for details of the analytical proof.

negligence; (b) negligence with contributory negligence; (c) negligence with comparative negligence; (d) strict liability; (e) strict liability with contributory negligence; and (f) strict liability with comparative negligence. As it will be shown in Table 1 at the end of this section, forgiveness and threat strategies correct the inefficiency induced by overconfidence in quite different ways.

Analytically, the crucial distinction between threat and forgiveness strategies lies in the different choice of the level of due care used to establish the parties' negligence.

Under threat strategy, if a party that fails to take optimal care is held negligent (or contributorily negligent) even if his failure to exercise due care is due to a natural level of overconfidence under the circumstances. We denote the level of due care for the injurer and for the victim under threat strategies respectively as $\overline{x}^T$ and $\overline{y}^T$. Under a threat strategy, the court sets a level of due care equal to the one satisfying the FOC of the social problem in absence of an overconfidence bias, according to equations (1) and (2). Hence, the level of due care will be fixed at the socially optimal level of care, x^{**} and y^{**}.

Under a forgiveness strategy, the court lowers the standard of due care to account for the margin of overconfidence error that an average individual would have under the circumstances. The lower levels of due care for tortfeasor and victim under a forgiveness strategy are denoted $\overline{x}^F$ and $\overline{y}^F$, respectively. Under a forgiveness strategy, the court sets a level of due care according to the FOC of the social problem incorporating the presence of the overconfidence bias and the consequent effect on the optimal choice of the agent. Conditions (1) and (2) hence become:

$$-zp_x^I(x^*, y)L = 1 \qquad (5)$$

$$-wp_y^I(x, y^*)L = 1 \qquad (6)$$

By incorporating the level of overconfidence bias that affect the estimation of risk of the average person, the court holds both parties to a lower level of due care, $\overline{x}^F < \overline{x}^T$ and $\overline{y}^F < \overline{y}^T$.

Simple Negligence. Under a simple negligence rule the tortfeasor is held liable if he exerts a level of care that is lower than the level of due care. If the tortfeasor undertakes a level of care equal or higher than the legal standard of due care, he avoids liability and the accident loss falls on his victim. Analytically, the tortfeasor is held liable only if the chosen level of care is lower than $\overline{x}$. Accordingly, the injurer's problem is:

$$max_{(w,x)}F^T = \begin{cases} V_T(w) - wzp(x,y)L - x \text{ } if \text{ } x < \overline{x} \\ V_T(w) - x \text{ } if \text{ } x \geq \overline{x} \end{cases}$$

and the victim's problem is:

$$max_{(z,y)}F^V = \begin{cases} V_V(z) - y \text{ } if \text{ } x < \overline{x} \\ V_V(z) - wzp(x,y)L - y \text{ } if \text{ } x \geq \overline{x} \end{cases}$$

negligence; (b) negligence with contributory negligence; (c) negligence with comparative negligence; (d) strict liability; (e) strict liability with contributory negligence; and (f) strict liability with comparative negligence. As it will be shown in Table 1 at the end of this section, forgiveness and threat strategies correct the inefficiency induced by overconfidence in quite different ways.

Analytically, the crucial distinction between threat and forgiveness strategies lies in the different choice of the level of due care used to establish the parties' negligence.

Under threat strategy, if a party that fails to take optimal care is held negligent (or contributorily negligent) even if his failure to exercise due care is due to a natural level of overconfidence under the circumstances. We denote the level of due care for the injurer and for the victim under threat strategies respectively as $\overline{x}^T$ and $\overline{y}^T$. Under a threat strategy, the court sets a level of due care equal to the one satisfying the FOC of the social problem in absence of an overconfidence bias, according to equations (1) and (2). Hence, the level of due care will be fixed at the socially optimal level of care, x^{**} and y^{**}.

Under a forgiveness strategy, the court lowers the standard of due care to account for the margin of overconfidence error that an average individual would have under the circumstances. The lower levels of due care for tortfeasor and victim under a forgiveness strategy are denoted $\overline{x}^F$ and $\overline{y}^F$, respectively. Under a forgiveness strategy, the court sets a level of due care according to the FOC of the social problem incorporating the presence of the overconfidence bias and the consequent effect on the optimal choice of the agent. Conditions (1) and (2) hence become:

$$-zp_x^I(x^*,y)L = 1 \qquad (5)$$

$$-wp_y^I(x,y^*)L = 1 \qquad (6)$$

By incorporating the level of overconfidence bias that affect the estimation of risk of the average person, the court holds both parties to a lower level of due care, $\overline{x}^F < \overline{x}^T$ and $\overline{y}^F < \overline{y}^T$.

Simple Negligence. Under a simple negligence rule the tortfeasor is held liable if he exerts a level of care that is lower than the level of due care. If the tortfeasor undertakes a level of care equal or higher than the legal standard of due care, he avoids liability and the accident loss falls on his victim. Analytically, the tortfeasor is held liable only if the chosen level of care is lower than $\overline{x}$. Accordingly, the injurer's problem is:

$$max_{(w,x)}F^T = \begin{cases} V_T(w) - wzp(x,y)L - x \; if \; x < \overline{x} \\ V_T(w) - x \; if \; x \geq \overline{x} \end{cases}$$

and the victim's problem is:

$$max_{(z,y)}F^V = \begin{cases} V_V(z) - y \; if \; x < \overline{x} \\ V_V(z) - wzp(x,y)L - y \; if \; x \geq \overline{x} \end{cases}$$

We denote with x^* the privately optimal care of the tortfeasor in the absence of overconfidence and $\underline{x}^*$ the privately optimal level in the presence of overconfidence. We indicate with the superscript T the equilibrium outcome associated to threat strategy and with F the one associated with forgiveness strategy.[21]

Proposition 2. *Under a negligence rule, overconfidence leads injurers to undertake a suboptimal level of care and an excessive activity level. The injurer's care decreases in the degree of overconfidence. Threat and forgiveness strategies have quite different effects on the tortfeasor's optimal choices. Under a threat strategy, the victim exercises no care and undertakes an inefficiently high level of activity. Under a forgiveness strategy, the victim exercises a higher level of care and a lower level of activity than socially optimal, in order to compensate the inefficient choices of the injurer.*

Proof: See Appendix

The intuition behind Proposition 2 is straightforward. Under a rule of simple negligence, the tortfeasor wishes to take a level of care equal to the standard of due care to avoid liability. However, in presence of overconfidence, the tortfeasor undertakes a privately optimal level of care, which is lower than the socially optimal level, i.e.:

$$\underline{x}^{i*} < x^{**}, \quad i = T, F$$

Since the injurer is not aware of his overconfidence bias, he will not correct for his bias when choosing his care level. Irrespective of the strategy adopted by the courts in setting liability, the injurer believes to have taken precautions that satisfy the requirement of due care, given his perception of risk, and thus he discounts his liability in case an accident occurs. Henceforth, raising the level of precaution above the perceived level of due care only imposes a cost and does not generate any perceived benefit in terms of a decrease in expected liability. As in the standard case, a negligence rule will also lead the tortfeasor to undertake a level of activity exceeding the socially optimal, i.e. $\underline{w}^* > w^{**}$. Note that overconfidence introduces inefficiency in the tortfeasor's choice of care, but it is not responsible for the inefficient activity level, which is observed under a negligence rule also in the absence of overconfidence.

The presence of overconfidence affects the victim's privately optimal choices drastically. It turns out that a threat strategy induces a higher inefficiency than a forgiveness strategy in the privately optimal care level chosen by the victim. The reason behind the result is rather intuitive. If

---

[21] Similar notations are used for the other variables.

the court follows a threat strategy when setting the required level of due care, the victim anticipates that the tortfeasor will be held liable by the court. The victim will therefore be repaid in case an accident occurs and finds optimal to exert no care since he will be compensated in case of a loss, i.e. $\underline{y}^{T*} = 0$. In an analogous way, the level of activity chosen by the victim will be the highest possible and higher than the socially optimal one. In terms of notation, we label with superscript max the level of activity chosen in case of zero care, i.e. $\underline{z}^{T*} = z^{max}$. As it will be seen, we will observe a mirror-image equilibrium in presence of overconfidence and threat strategies under a strict liability rule.

Forgiveness strategies work in the opposite direction: the victim is aware that the tortfeasor will not be held liable in equilibrium and realizes that he needs to compensate the lower care level chosen by the tortfeasor due to overconfidence. Hence, the victim will find optimal to choose a higher level of care, $\underline{y}^{T*} > y^{**}$ and a lower activity level, $\underline{z}^{T*} < z^{**}$.

Threat strategies, aimed as they are at "punishing overconfidence" through liability not only may fail to debias overconfidence, but may exacerbate inefficiency – inducing an opportunistic behaviour. Parties may opportunistically rely on the overconfidence error and liability of the other party and engage in moral hazard. As it will be discussed, this result is stable across different liability rules. On the contrary, under a simple negligence regime, forgiveness strategies help restore efficiency for the victim. Forgiveness strategies fail to correct the tortfeasor's bias and resulting inefficiency in care and activity level, but avoid the victim's moral hazard, actually incentivizing the victim to compensate the tortfeasor's (forgiven) error with their own precautions.

Negligence with a Defense of Contributory Negligence. Under the negligence rule with the defense of contributory negligence, a injurer is held liable for accident losses caused by his activity only if two conditions are met simultaneously: he is negligent and exercises a level of care less than due care $\bar{x}$ and the victim is diligent, exerting a care level higher than due care $\bar{y}$. Accordingly, the injurer's problem becomes:

$$max_{(w,x)}F^T = \begin{cases} V_T(w) - wzp(x,y)L - x \; if \; x < \bar{x} and \; y \geq \bar{y} \\ V_T(w) - x \;\; if \; x < \bar{x} \; and \; y < \bar{y} \; or \; if \; x \geq \bar{x} \end{cases}$$

and the one of the victim is:

$$max_{(z,y)}F^V = \begin{cases} V_V(z) - y \; if \; x < \bar{x} \; and \; y \geq \bar{y} \\ V_V(z) - wzp(x,y)L - y \quad if \; x < \bar{x} \; and \; y < \bar{y} \; or \; if \; x \geq \bar{x} \end{cases}$$

Proposition 3. *Under a negligence rule with a defense of contributory negligence, the victim will undertake a suboptimal level of care and an excessive activity level. The injurer's care decreases in* *the degree of overconfidence. Threat and forgiveness strategies have quite different effects on the injurer's optimal choices. Under a threat strategy, the injurer exercises no care and undertakes an inefficiently high level of activity. Under a forgiveness strategy, the injurer exercises a higher level of care and a lower level of activity than socially optimal, in order to compensate the inefficient choices of the victim.*

Proof: See Appendix

The equilibria generated by threat and forgiveness strategies in a regime of negligence with a defense of contributory negligence mirror those generated by a rule of simple negligence. Under both strategies, the effects of contributory negligence on the victim's (injurer's) incentives are symmetric to the incentives faced by the injurer (victim) in a regime of simple negligence.

The logic behind this symmetry is given by the fact that under contributory negligence the victim receives compensation only if he exercises due care. Therefore he will find it optimal to choose the level of care that satisfies the social problem. However, overconfidence affects his choice, yielding a lower level of care:

$$\underline{y}^{i*} < y^{**}, i = T, F$$

The symmetries between these two liability rules are two-fold.

First, similarly to the injurer's behaviour under simple negligence, the victim overconfidently believes to have undertaken due care, hence avoiding liability. As a result, the victim carries out his activity beyond the socially optimal level, as generally observed under negligence. This equilibrium outcome is insensitive to threat or forgiveness strategies.

Second, similarly to the victim's behaviour under simple negligence, the injurer realizes that if a threat strategy is utilized, the overconfident victim will be barred from compensation and will bear the full accident loss. Hence he will opportunistically rely on the avoidance of his liability, adopting care $\underline{x}^{T*} = 0$. As a result, the injurer chooses in equilibrium an activity level higher than socially optimal.

Forgiveness strategies create incentives for tortfeasors to compensate the (forgiven) errors of their victims. Since under a forgiveness strategy the victim will not be held liable if he adopts the due care level (calculated accounting for the natural level of overconfidence possessed by the average person under similar circumstances), the injurer will face restored incentives to adopt optimal care. Additionally, the tortfeasor will adjust his care in response to the victim's overconfidence errors, compensating the victim's lack of care with his own. The tortfeasor's level of care will be higher than the second best efficient equilibrium of the victim in absence of the bias,

$\underline{x}^{F*} > x^{**}$. Consequently, he will reduce activity level below the socially optimal level in the absence of overconfidence, $\underline{w}^{F*} < w^{**}$.

Negligence Rule with a Defense of Comparative Negligence. Under a negligence rule with a defense of comparative negligence, the victim receives a share α of compensation if the tortfeasor is negligent and the victim is diligent, according to the standard of due care defined above. Accordingly, the injurer's problem is:

$$max_{(w,x)}F^T = \begin{cases} V_T(w) - wzp(x,y)L - x \; if \; x < \bar{x} \; and \; y \geq \bar{y} \\ V_T(w) - \alpha wzp(x,y)L - x \;\; if \; x < \bar{x} \; and \; y < \bar{y} \\ V_T(w) - x \;\; if \; x \geq \bar{x} \end{cases}$$

and the victim's one is:

$$max_{(z,y)}F^V = \begin{cases} V_V(z) - y \; if \; x < \bar{x} \; and \; y \geq \bar{y} \\ V_V(z) - (1-\alpha)wzp(x,y)L - y \;\; if \; x < \bar{x} \; and \; y < \bar{y} \\ V_V(z) - wzp(x,y)L - y \; if \; x \geq \bar{x} \end{cases}$$

Proposition 4. *Under a rule of comparative negligence, both the injurer and the victim exert positive level of care. The level of care decreases in the degree of overconfidence. Under a threat strategy, both the victim and the injurer undertake an inefficiently high level of activity. Under forgiveness strategy, the victim undertakes a lower level of activity to compensate for injurer's overconfidence.*

Proof: See Appendix.

The result of Proposition 4 is rather intuitive. Both the victim and the tortfeasor have incentives to exert due care to avoid liability. Overconfidence bias affects both parties, thereby inducing a lower level of care for both in case they act according to socially optimal incentives, i.e. $\underline{x}^* < x^{**}$ and $\underline{y}^* < y^{**}$. Due to the defense of comparative negligence, both the tortfeasor and the injurer know a positive level of care is required to be shielded from liability. Hence, they exert a positive, even if inefficiently low level of care due to overconfidence. Ex-post, in case an accident occurs, both the injurer and the victim will be found liable and will share the cost of accident. Since residual liability is on the victim, the injurer chooses an inefficiently high level of activity.

By a line of reasoning analogous to the one established in case of a rule of simple negligence, threat and forgiveness strategies affect only the equilibrium of the victim. Under a threat strategy the victim anticipates the mistake of the tortfeasor and that he will receive compensation in case of any accident. Thereby the victim chooses to perform a higher level of activity. Forgiveness strategies, instead, impose to the victim to restore a choice of care as a second best efficient level in response of the tortfeasor's choice. Due to the inefficiently low care level, the victim ends up compensating the mistake of the other party, thereby increasing his care level above the second best efficient level in absence of the bias, i.e. $\underline{y}^{F*} > y^{**}$.

Strict Liability. Under strict liability, the tortfeasor must pay for all accident losses caused by his activity, independently of the level of care exercised. The objective function of the tortfeasor is:

$$max_{(w,x)}F_T = V_T(w) - wzp^I(x,y)L - wx$$

The victim bears no residual liability and his objective function is:

$$max_{(z,y)}F_V = V_V(z) - zy$$

Proposition 5. *Under strict liability, the equilibrium of the victim is unchanged: threat and forgiveness strategies do not correct the inefficiency due to overconfidence. The level of care exercised by the injurer is less than the socially optimal level of care and the activity level is above the socially optimal one chosen in absence of overconfidence. In addition, the injurer's care decreases in the size of overconfidence. The equilibrium outcome of the victim is unaffected by the presence of overconfidence.*

Proof: See Appendix

In presence of overconfidence, the equilibrium of the victim does not change: as in standard analysis, the victim will exert no care since he will be compensated in any case of a loss. The same is true for the level of activity chosen by the victim, that will be higher than socially optimal, and chosen according to first order condition: $V_z = 0$, i.e. $z^* = \underline{z}^{i*} > z^{**}$, where $i = T, F$. Threat and forgiveness strategies do not make any difference in the optimal choice of the victim.

Under the strict liability rule, the tortfeasor has an objective function identical to social optimum. However, the presence of overconfidence induces an inefficiency in the choice of the level of care: since the injurer underestimates the probability of an accident to occur he will find optimal to exert a lower level of care at equilibrium. Additionally this induces a second inefficiency: given a lower care level the injurer finds optimal to increase activity above the social optimum.

Threat and forgiveness strategies lead to the same equilibrium choices of both the tortfeasor and the victim. The reason is rather intuitive: the strategy of forgiving the fact that the tortfeasor exercises a level of care lower than the socially optimal plays no role under a strict liability rule since the tortfeasor will be held liable irrespective of the care level chosen by the tortfeasor.

Symmetrically, the activity level chosen by the tortfeasor is second best efficient given the choice of the victim and taking into consideration the lower care, $\underline{w}^{F*} < w^{**}$.

Strict Liability with a Defense of Comparative Negligence. Under the defense of comparative negligence, the victim receives a share α of the compensation if the tortfeasor is negligent and the victim is diligent, i.e. $x < \overline{x}$ and $y \geq \overline{y}$. Accordingly, the injurer's problem is:

$$max_{(w,x)}F^T = \begin{cases} V_T(w) - \alpha wzp(x,y)L - x \quad if\ x < \overline{x}\ and\ \ y < \overline{y} \\ V_T(w) - wzp(x,y)L - x \quad if\ \ y \geq \overline{y} \\ V_T(w) - x\ if\ x \geq \overline{x}\ and\ \ y < \overline{y} \end{cases}$$

and the victim's one is:

$$max_{(z,y)}F^V = \begin{cases} V_V(z) - (1-\alpha)wzp(x,y)L - y\ \ if\ x < \overline{x}\ and\ y < \overline{y} \\ V_V(z) - y \quad if\ \ y \geq \overline{y} \\ V_V(z) - wzp(x,y)L - y\ if\ x \geq \overline{x}\ and\ \ y < \overline{y} \end{cases}$$

Proposition 7. *In a regime of strict liability with a defense of comparative negligence, the care equilibrium outcome coincides with that observed under a rule of negligence with a defense of comparative negligence. Under a threat strategy, both the victim and the injurer undertake an inefficiently high level of activity. Under forgiveness strategy, the injurer undertakes a lower level of activity to compensate for injurer's overconfidence.*

Proof: See Appendix.

The intuition behind Proposition 7 is best explained by analogy with the case of negligence with the defense of comparative negligence. The presence of the defense of contributory negligence pushes both parties to wish to exercise due care to avoid liability – as under the rule of negligence with a defense of contributory negligence. Note, however, that activity incentives work in the opposite direction with respect to the rule of negligence with a defense of comparative negligence and the injurer will be the one exerting optimal activity under forgiveness strategy.

5.4 Further Discussion of the Results

The presence of overconfidence affects equilibrium outcomes in different ways in different liability regimes. In particular, the presence of overconfidence bias breaks down the equivalence among the defense of contributory and comparative negligence augmenting either simple negligence or strict liability rule. This is proved to be true despite of the due care level chosen by

Strict Liability with a Defense of Contributory Negligence. Under strict liability with the defense of contributory negligence, compensation is paid to the victim in all cases, except when the victim himself acted negligently, i.e. $y < \overline{y}$. Accordingly, the injurer's problem is:

$$max_{(w,x)}F^T = \begin{cases} V_T(w) - wzp(x,y)L - x \; if \; y \geq \overline{y}. \\ V_T(w) - x \; if \;\; y < \overline{y} \end{cases}$$

and the victim's one is:

$$max_{(z,y)}F^V = \begin{cases} V_V(z) - y \; if \;\; y \geq \overline{y} \\ V_V(z) - wzp(x,y)L - y \quad if \;\; y < \overline{y} \end{cases}$$

Proposition 6. *In a regime of strict liability with a defense of contributory negligence, the equilibrium coincides with that observed under a rule of negligence with a defense of contributory negligence. Proposition 3 holds.*

Proof: See Appendix.

The intuition behind Proposition 6 is straightforward: in the presence of the defense of contributory negligence, the victim will wish to exercise due care to avoid liability – as under the rule of negligence with a defense of contributory negligence. This is true since the victim becomes the one who bears the burden of the proof regarding his care level. Hence the same reasoning explained for the case of negligence with the defense of contributory negligence applies and lead to the same equilibrium under threat and forgiveness strategy.

The equilibrium of the victim does not change irrespective of the threat or forgiveness strategy. Due to the overconfidence bias, however, the victim fails to choose the socially efficient level of care in absence of overconfidence and will exert a lower level of care, $\underline{y}^{i*} < y^{**}, i = T, F$. However, the victim does not realize her own mistake and believes he will receive compensation in case of an accident. Thereby, the victim chooses a higher level of activity than the social optimum , $\underline{z}^{i*} > z^{**}, i = T, F$.

Again, the equilibrium changes drastically for the tortfeasor: under threat strategy, the tortfeasor realizes victim's mistake due to bias and set a level of care equal to zero, i.e. $z^{*} = 0$. Forgiveness strategies restore the incentive of the injurer, implying for him that he will be held liable if exerts no care and corrects the choice of care to be second best efficient in response of the victim's choice. This however produces an overcompensation of the lower level of care of the victim, thereby increasing his care level above the second best efficient level in absence of the bias, i.e. $\underline{x}^{F*} > x^{**}$.

Symmetrically, the activity level chosen by the tortfeasor is second best efficient given the choice of the victim and taking into consideration the lower care, $\underline{w}^{F*} < w^{**}$.

Strict Liability with a Defense of Comparative Negligence. Under the defense of comparative negligence, the victim receives a share α of the compensation if the tortfeasor is negligent and the victim is diligent, i.e. $x < \overline{x}$ and $y \geq \overline{y}$. Accordingly, the injurer's problem is:

$$max_{(w,x)}F^T = \begin{cases} V_T(w) - \alpha wzp(x,y)L - x \quad if\ x < \overline{x}\ and\ \ y < \overline{y} \\ V_T(w) - wzp(x,y)L - x \quad if\ \ y \geq \overline{y} \\ V_T(w) - x\ if\ x \geq \overline{x}\ and\ \ y < \overline{y} \end{cases}$$

and the victim's one is:

$$max_{(z,y)}F^V = \begin{cases} V_V(z) - (1-\alpha)wzp(x,y)L - y\ \ if\ x < \overline{x}\ and\ y < \overline{y} \\ V_V(z) - y \quad if\ \ y \geq \overline{y} \\ V_V(z) - wzp(x,y)L - y\ if\ x \geq \overline{x}\ and\ \ y < \overline{y} \end{cases}$$

Proposition 7. *In a regime of strict liability with a defense of comparative negligence, the care equilibrium outcome coincides with that observed under a rule of negligence with a defense of comparative negligence. Under a threat strategy, both the victim and the injurer undertake an inefficiently high level of activity. Under forgiveness strategy, the injurer undertakes a lower level of activity to compensate for injurer's overconfidence.*

Proof: See Appendix.

The intuition behind Proposition 7 is best explained by analogy with the case of negligence with the defense of comparative negligence. The presence of the defense of contributory negligence pushes both parties to wish to exercise due care to avoid liability – as under the rule of negligence with a defense of contributory negligence. Note, however, that activity incentives work in the opposite direction with respect to the rule of negligence with a defense of comparative negligence and the injurer will be the one exerting optimal activity under forgiveness strategy.

5.4 Further Discussion of the Results

The presence of overconfidence affects equilibrium outcomes in different ways in different liability regimes. In particular, the presence of overconfidence bias breaks down the equivalence among the defense of contributory and comparative negligence augmenting either simple negligence or strict liability rule. This is proved to be true despite of the due care level chosen by courts, either in the case of threat or forgiveness strategy. We establish the equivalence between negligence with a defense of comparative negligence and strict liability with a defense of comparative negligence. In a parallel way, the defense of contributory negligence under negligence or strict liability leads to the same equilibrium outcomes.

	CARE LEVELS			
	Injurer (X)		Victim (Y)	
	T	F	T	F
Simple Negligence	$\underline{x}^{T*} < x^{**}$	$\underline{x}^{F*} < x^{**}$	$\underline{y}^{T*} = 0$	$\underline{y}^{T*} < \underline{y}^{F*} \approx y^{**}$
Negligence + Contributory. Negligence	$\underline{x}^{T*} = 0$	$\underline{x}^{T*} < \underline{x}^{F*} \approx x^{**}$	$\underline{y}^{T*} < y^{**}$	$\underline{y}^{F*} < y^{**}$
Negligence + Comparative Negligence	$\underline{x}^{T*} < x^{**}$	$\underline{x}^{F*} < x^{**}$	$\underline{y}^{T*} < y^{**}$	$\underline{y}^{T*} < \underline{y}^{F*} \approx y^{**}$
Strict Liability	$\underline{x}^{T*} < x^{**}$	$\underline{x}^{F*} = \underline{x}^{T*} < x^{**}$	**0**	**0**
Strict Liability + Contributory Negligence	$\underline{x}^{T*} = 0$	$\underline{x}^{T*} < \underline{x}^{F*} \approx x^{**}$	$\underline{y}^{T*} < y^{**}$	$\underline{y}^{F*} < y^{**}$
Strict Liability + Comparative Negligence	$\underline{x}^{T*} < x^{**}$	$\underline{x}^{F*} < x^{**}$	$\underline{y}^{T*} < y^{**}$	$\underline{y}^{T*} < \underline{y}^{F*} \approx y^{**}$

Table 1a: *Threat vs. Forgiveness Strategies: Care Level*

	ACTIVITY LEVELS			
	Injurer (X)		**Victim (Y)**	
	T	F	T	F
Simple Negligence	$\underline{w}^{T*} > w^{**}$	$\underline{w}^{F*} > w^{**}$	$\underline{z}^{*} = z^{max}$	$\underline{z}^{T*} > \underline{z}^{F*} \approx z^{**}$
Negligence + Contributory. Negligence	$\underline{w}^{*} = w^{max}$	$\underline{w}^{T*} > w^{F*} \approx w^{**}$	$\underline{z}^{T*} > z^{**}$	$\underline{z}^{F*} > z^{**}$
Negligence + Comparative Negligence	$\underline{w}^{T*} > w^{**}$	$\underline{w}^{F*} > w^{**}$	$\underline{z}^{*} = z^{max}$	$\underline{z}^{T*} > \underline{z}^{F*} \approx z^{**}$
Strict Liability	$\underline{w}^{T*} > w^{**}$	$\underline{w}^{F*} = \underline{w}^{T*} < w^{**}$	$\underline{z}^{*} = z^{max}$	$\underline{z}^{*} = z^{max}$
Strict Liability + Contributory Negligence	$\underline{w}^{*} = w^{max}$	$\underline{w}^{T*} > w^{F*} \approx w^{**}$	$\underline{z}^{T*} > z^{**}$	$\underline{z}^{F*} > z^{**}$
Strict Liability + Comparative Negligence	$\underline{w}^{T*} > w^{**}$	$\underline{w}^{F*} > w^{**}$	$\underline{z}^{*} = z^{max}$	$\underline{z}^{T*} > \underline{z}^{F*} \approx z^{**}$

Table 1b: *Threat vs. Forgiveness Strategies: Activity Level*

5.5 Summing Up

Our results suggest that legal forgiveness of overconfidence may be a valuable second-best solution, when debiasing through tort law cannot be effectively achieved. We sum up the main results on the comparison between threat strategies versus forgiveness strategies in terms of information incentives, deterrence and moral hazard, and activity levels.

5.5.1 Care incentives

The legal system's treatment of overconfidence errors has an impact on the parties' incentives to invest in precautions. Imposing liability for overconfidence errors creates incentives on the parties to correct their own overconfidence bias though additional precautions, but may undermine the incentives to invest in precaution to compensate for the other party's overconfidence. In the following, we identify the main features of these care incentives.

courts, either in the case of threat or forgiveness strategy. We establish the equivalence between negligence with a defense of comparative negligence and strict liability with a defense of comparative negligence. In a parallel way, the defense of contributory negligence under negligence or strict liability leads to the same equilibrium outcomes.

	CARE LEVELS			
	Injurer (X)		Victim (Y)	
	T	F	T	F
Simple Negligence	$\underline{x}^{T*} < x^{**}$	$\underline{x}^{F*} < x^{**}$	$\underline{y}^{T*} = 0$	$\underline{y}^{T*} < \underline{y}^{F*} \approx y^{**}$
Negligence + Contributory. Negligence	$\underline{x}^{T*} = 0$	$\underline{x}^{T*} < \underline{x}^{F*} \approx x^{**}$	$\underline{y}^{T*} < y^{**}$	$\underline{y}^{F*} < y^{**}$
Negligence + Comparative Negligence	$\underline{x}^{T*} < x^{**}$	$\underline{x}^{F*} < x^{**}$	$\underline{y}^{T*} < y^{**}$	$\underline{y}^{T*} < \underline{y}^{F*} \approx y^{**}$
Strict Liability	$\underline{x}^{T*} < x^{**}$	$\underline{x}^{F*} = \underline{x}^{T*} < x^{**}$	**0**	**0**
Strict Liability + Contributory Negligence	$\underline{x}^{T*} = 0$	$\underline{x}^{T*} < \underline{x}^{F*} \approx x^{**}$	$\underline{y}^{T*} < y^{**}$	$\underline{y}^{F*} < y^{**}$
Strict Liability + Comparative Negligence	$\underline{x}^{T*} < x^{**}$	$\underline{x}^{F*} < x^{**}$	$\underline{y}^{T*} < y^{**}$	$\underline{y}^{T*} < \underline{y}^{F*} \approx y^{**}$

Table 1a: *Threat vs. Forgiveness Strategies: Care Level*

	ACTIVITY LEVELS			
	Injurer (X)		**Victim (Y)**	
	T	F	T	F
Simple Negligence	$\underline{w}^{T*} > w^{**}$	$\underline{w}^{F*} > w^{**}$	$\underline{z}^{*} = z^{max}$	$\underline{z}^{T*} > \underline{z}^{F*} \approx z^{**}$
Negligence + Contributory. Negligence	$\underline{w}^{*} = w^{max}$	$\underline{w}^{T*} > w^{F*} \approx w^{**}$	$\underline{z}^{T*} > z^{**}$	$\underline{z}^{F*} > z^{**}$
Negligence + Comparative Negligence	$\underline{w}^{T*} > w^{**}$	$\underline{w}^{F*} > w^{**}$	$\underline{z}^{*} = z^{max}$	$\underline{z}^{T*} > \underline{z}^{F*} \approx z^{**}$
Strict Liability	$\underline{w}^{T*} > w^{**}$	$\underline{w}^{F*} = \underline{w}^{T*} < w^{**}$	$\underline{z}^{*} = z^{max}$	$\underline{z}^{*} = z^{max}$
Strict Liability + Contributory Negligence	$\underline{w}^{*} = w^{max}$	$\underline{w}^{T*} > w^{F*} \approx w^{**}$	$\underline{z}^{T*} > z^{**}$	$\underline{z}^{F*} > z^{**}$
Strict Liability + Comparative Negligence	$\underline{w}^{T*} > w^{**}$	$\underline{w}^{F*} > w^{**}$	$\underline{z}^{*} = z^{max}$	$\underline{z}^{T*} > \underline{z}^{F*} \approx z^{**}$

Table 1b: *Threat vs. Forgiveness Strategies: Activity Level*

5.5 Summing Up

Our results suggest that legal forgiveness of overconfidence may be a valuable second-best solution, when debiasing through tort law cannot be effectively achieved. We sum up the main results on the comparison between threat strategies versus forgiveness strategies in terms of information incentives, deterrence and moral hazard, and activity levels.

5.5.1 Care incentives

The legal system's treatment of overconfidence errors has an impact on the parties' incentives to invest in precautions. Imposing liability for overconfidence errors creates incentives on the parties to correct their own overconfidence bias though additional precautions, but may undermine the incentives to invest in precaution to compensate for the other party's overconfidence. In the following, we identify the main features of these care incentives.

Proposition 1A: *With Forgiveness Strategies [Threat Strategies], parties will not [will] be deterred by their own liability, and will not [will] adjust behaviour accordingly.*

Proposition 1B: *With Forgiveness Strategies [Threat Strategies], parties will not [will] rely on the other party's liability, and will not [will] engage in moral hazard.*

Corollary 1: *With Forgiveness Strategies [Threat Strategies], parties will [will not] be induced to compensate the other party's overconfidence with their own care.*

As it has been shown, in the limiting case where tortfeasor and victim are similarly affected by an overconfidence bias (symmetric overconfidence), legal forgiveness of overconfidence will induce both parties to undertake precaution levels that approach the socially optimal level. Undershooting and overshooting are possible when asymmetries are introduced between the levels of overconfidence of tortfeasors and victims.

The legal system's treatment of overconfidence errors also has an impact on the parties' incentives to acquire and to take into account information that may reduce or correct their bias. Liability regimes create information incentives on the parties' own overconfidence bias, but may create reverse effects on the other party's incentives. For example, forgiveness strategies do not induce the parties to take into consideration the liability impact of their own overconfidence errors, but induce the parties to take into consideration the other party's overconfidence errors. Conversely, threat strategies induce the parties to take into consideration the impact of their own overconfidence, but do not induce the parties to take into consideration the other party's overconfidence errors.

When parties can only be partially debiased through information ($\alpha > 0$) the choice between forgiveness and liability for overconfidence depends on two countervailing effects. The first effect is given by the threat of liability for accidents caused by overconfidence errors. This threat of liability will only be partially internalized by the parties, inasmuch as the parties can only be made partially aware of their bias through information and education. Internalization will increase with α, the degree of effectiveness of debiasing strategies. The second effect is given by the residual liability effect. If overconfident actors are held liable for their overconfidence errors, the other party's residual liability for overconfidence is eliminated. Prospective victims of overconfidence errors (and injurers of overconfident victims in contributory negligence regimes) may face a moral hazard problem, failing to adjust their precautions to compensate for the other party's overconfidence errors. The incentive to compensate for the other party's overconfidence will be undermined by the rule of liability for overconfidence.

Proposition 2A: *When debiasing can [cannot] be achieved through debiasing strategies, imposing liability for overconfidence errors would [would not] be desirable.*

Proposition 2B: *Under a strict liability regime, threat and forgiveness strategies lead to identical result, regardless of the effectiveness of debiasing strategies, α.*

When debiasing cannot be achieved through debiasing strategies, imposing liability for overconfidence errors is: (a) ineffective for the correction of the parties' own bias, and (b) would undermine the parties' incentives to compensate for the overconfidence of the other party. Hence, with ineffective debiasing strategies forgiveness strategies may be preferable in order to create optimal care incentives.

5.5.2 Overconfidence and Activity Levels

The alternative threat and forgiveness strategies shift the expected loss due to overconfidence errors from one party to the other, affecting the parties' cost of the activity and activity level incentives.

Proposition 3: *Legal forgiveness corrects the inefficient level of activity chosen by the "residually liable agent", restoring the second best efficient activity level.*

Corollary 2: *In cases of unilateral overconfidence, activity level is not corrected when the overconfident party is held residually liable.*

Proposition 1A: *With Forgiveness Strategies [Threat Strategies], parties will not [will] be deterred by their own liability, and will not [will] adjust behaviour accordingly.*

Proposition 1B: *With Forgiveness Strategies [Threat Strategies], parties will not [will] rely on the other party's liability, and will not [will] engage in moral hazard.*

Corollary 1: *With Forgiveness Strategies [Threat Strategies], parties will [will not] be induced to compensate the other party's overconfidence with their own care.*

As it has been shown, in the limiting case where tortfeasor and victim are similarly affected by an overconfidence bias (symmetric overconfidence), legal forgiveness of overconfidence will induce both parties to undertake precaution levels that approach the socially optimal level. Undershooting and overshooting are possible when asymmetries are introduced between the levels of overconfidence of tortfeasors and victims.

The legal system's treatment of overconfidence errors also has an impact on the parties' incentives to acquire and to take into account information that may reduce or correct their bias. Liability regimes create information incentives on the parties' own overconfidence bias, but may create reverse effects on the other party's incentives. For example, forgiveness strategies do not induce the parties to take into consideration the liability impact of their own overconfidence errors, but induce the parties to take into consideration the other party's overconfidence errors. Conversely, threat strategies induce the parties to take into consideration the impact of their own overconfidence, but do not induce the parties to take into consideration the other party's overconfidence errors.

When parties can only be partially debiased through information ($\alpha > 0$) the choice between forgiveness and liability for overconfidence depends on two countervailing effects. The first effect is given by the threat of liability for accidents caused by overconfidence errors. This threat of liability will only be partially internalized by the parties, inasmuch as the parties can only be made partially aware of their bias through information and education. Internalization will increase with α, the degree of effectiveness of debiasing strategies. The second effect is given by the residual liability effect. If overconfident actors are held liable for their overconfidence errors, the other party's residual liability for overconfidence is eliminated. Prospective victims of overconfidence errors (and injurers of overconfident victims in contributory negligence regimes) may face a moral hazard problem, failing to adjust their precautions to compensate for the other party's overconfidence errors. The incentive to compensate for the other party's overconfidence will be undermined by the rule of liability for overconfidence.

Proposition 2A: *When debiasing can [cannot] be achieved through debiasing strategies, imposing liability for overconfidence errors would [would not] be desirable.*

Proposition 2B: *Under a strict liability regime, threat and forgiveness strategies lead to identical result, regardless of the effectiveness of debiasing strategies, α.*

When debiasing cannot be achieved through debiasing strategies, imposing liability for overconfidence errors is: (a) ineffective for the correction of the parties' own bias, and (b) would undermine the parties' incentives to compensate for the overconfidence of the other party. Hence, with ineffective debiasing strategies forgiveness strategies may be preferable in order to create optimal care incentives.

5.5.2 Overconfidence and Activity Levels

The alternative threat and forgiveness strategies shift the expected loss due to overconfidence errors from one party to the other, affecting the parties' cost of the activity and activity level incentives.

Proposition 3: *Legal forgiveness corrects the inefficient level of activity chosen by the "residually liable agent", restoring the second best efficient activity level.*

Corollary 2: *In cases of unilateral overconfidence, activity level is not corrected when the overconfident party is held residually liable.*

6. Conclusions

Biased perceptions of reality can generally be (at least partially) corrected by providing a better knowledge of the statistical probability of harm. Overconfidence poses a special threat to tort rules: according to empirical and experimental evidence, the knowledge about statistical incidence of overconfidence only has a limited effect on the belief about one's own overconfidence bias. Hence, even if parties are provided accurate information about statistical facts, overconfidence may lead to the assumption of excessive risks, undermining the deterrent effect of liability rules. In this chapter, we have considered the role of tort rules in debiasing overconfidence, showing that in many situations the most effective way to correct overconfidence in tort law may be to forgive it, rather than to penalize it through liability.

What if the courts adopt liability rules indicating clear actions needed to be taken in order to avoid liability? This could be a solution (except for the presence of extreme overconfidence phenomena) – but it is very difficult to implement. What is the meaning for example on product liability? Could the list be complete?

Future extensions with two or more types of agents should be developed to consider the effect of forgiveness and threat strategies when both biased and unbiased agents are present. In the presence of unbiased agents in the population, forgiveness strategies will create an opportunity for unbiased individuals to rely strategically on the lower standard of care set on the basis of the average bias. The respective advantages of threat and forgiveness strategies will therefore have to be evaluated in light of the undesirable effects of the rule on unbiased agents.

Future extensions should also consider cases where information strategies have asymmetric effectiveness on the parties. These extensions should consider the possibility of a mixed application of forgiveness and threat strategies. For example, in the field of products liability it may be much easier to correct for the producer's overconfidence through information or ex ante safety regulation, than to debias the consumers' overconfidence through information or warnings. A hybrid rule that threatens liability for the producers' overconfidence errors, yet forgiving the consumers' natural level of overconfidence may be desirable. Such mixed application of threat and forgiveness strategies already finds some incarnations in products liability (e.g., by imposing liability on producers that do not increase the safety of their products in anticipation of consumer's overconfidence), though it has not been considered as a possible approach in general tort law.

6. Conclusions

Biased perceptions of reality can generally be (at least partially) corrected by providing a better knowledge of the statistical probability of harm. Overconfidence poses a special threat to tort rules: according to empirical and experimental evidence, the knowledge about statistical incidence of overconfidence only has a limited effect on the belief about one's own overconfidence bias. Hence, even if parties are provided accurate information about statistical facts, overconfidence may lead to the assumption of excessive risks, undermining the deterrent effect of liability rules. In this chapter, we have considered the role of tort rules in debiasing overconfidence, showing that in many situations the most effective way to correct overconfidence in tort law may be to forgive it, rather than to penalize it through liability.

What if the courts adopt liability rules indicating clear actions needed to be taken in order to avoid liability? This could be a solution (except for the presence of extreme overconfidence phenomena) – but it is very difficult to implement. What is the meaning for example on product liability? Could the list be complete?

Future extensions with two or more types of agents should be developed to consider the effect of forgiveness and threat strategies when both biased and unbiased agents are present. In the presence of unbiased agents in the population, forgiveness strategies will create an opportunity for unbiased individuals to rely strategically on the lower standard of care set on the basis of the average bias. The respective advantages of threat and forgiveness strategies will therefore have to be evaluated in light of the undesirable effects of the rule on unbiased agents.

Future extensions should also consider cases where information strategies have asymmetric effectiveness on the parties. These extensions should consider the possibility of a mixed application of forgiveness and threat strategies. For example, in the field of products liability it may be much easier to correct for the producer's overconfidence through information or ex ante safety regulation, than to debias the consumers' overconfidence through information or warnings. A hybrid rule that threatens liability for the producers' overconfidence errors, yet forgiving the consumers' natural level of overconfidence may be desirable. Such mixed application of threat and forgiveness strategies already finds some incarnations in products liability (e.g., by imposing liability on producers that do not increase the safety of their products in anticipation of consumer's overconfidence), though it has not been considered as a possible approach in general tort law.

Appendix

Proof of Proposition 2

Under the rule of simple negligence the tortfeasor will wish to take due care to avoid liability, i.e. the care level is chosen according to the following first order conditions:

$$-zp_x^I(\underline{x}^*, y)L = 1 \qquad (3)$$

However, in presence of overconfidence, equation (3) yields a lower level of care with respect to socially optimal one, as established before, i.e.: $\underline{x}^* < x^{**}$

The level of activity chosen by the tortfeasor will exceed the social optimum, and will be chosen according to first order condition: $V_w = 0$, i.e. $\underline{w}^* > w^{**}$.

The victim realizes that, applying a threat strategy, a court will hold the tortfeasor liable – something that the tortfeasor fails to realize. The victim will therefore exert no care since he will be compensated in case of a loss: $y^* = \underline{y}^* = 0$. In an analogous way, the level of activity chosen by the victim will be higher than socially optimal, and chosen according to first order condition: $V_z = 0$, i.e. $z^* = \underline{z}^* > z^{**}$.

Under forgiveness strategies, the tortfeasor will not be held liable if he exercises a level of care according to (3), i.e. the social best response to the victim given the biased perception of the likelihood of the accident. The choice of care and activity of the tortfeasor will not change under forgiveness, while the equilibrium for the victim restores to the simple negligence case, i.e. the victim will choose care and activity level according to the following first order conditions:

$$-zp_y^I\left(x, \underline{y}^*\right)L = 1$$

$$V_z = wp^I\left(x, \underline{y}^*\right)L$$

Due to the presence of overconfidence, the care and activity level chosen by the victim will be respectively higher and lower with respect to the social efficient equilibrium of the victim, i.e. $\underline{y}^* > y^{**}$ and $\underline{z}^* < z^{**}$.

Proof of Proposition 3

Under the rule of negligence with the defense of contributory negligence, the victim receives compensation only if he exercises due care to avoid liability, i.e. the care level is chosen according to the following first order conditions:

$$-wp_y^I\left(x, \underline{y}^*\right)L = 1 \qquad (4)$$

However, in presence of overconfidence, equation (4) yields a lower level of care with respect to socially optimal one, as established before, i.e.:

$$\underline{y}^* < y^{**}$$

At the same time, the level of activity chosen by the victim will exceed the social optimum, and will be chosen according to first order condition: $V_z = 0$, i.e. $\underline{z}^* > z^{**}$.

The tortfeasor – knowing he will not pay compensation to the victim - will exert no care: $\underline{x}^* = 0$. In an analogous way, the level of activity chosen by the tortfeasor will be higher than socially optimal, and chosen according to first order condition: $V_w = 0$, i.e. $\underline{w}^* > w^{**}$.

Under forgiveness strategies, the victim will not be held liable if he exercises a level of care according to (4), i.e. the second best response to the tortfeasor under the biased perception of the likelihood of the accident. The choice of care and activity of the victim will not change under forgiveness, while the equilibrium for the tortfeasor restores to the one under negligence rule with the defense of contributory negligence, i.e. the tortfeasor will choose care and activity level according to the following first order conditions:

$$-zp_x^I(\underline{x}^*, y)L = 1$$

$$V_w = zp^I(\underline{x}^*, y)L$$

Due to the presence of overconfidence, the care and activity level chosen by the tortfeasor will be respectively higher and lower with respect to the second best efficient equilibrium of the victim in absence of the bias, i.e. $\underline{x}^* > x^*$ and $\underline{w}^* < w^*$

Proof of Proposition 4

In case of threat strategies, both the victim and the tortfeasor will wish to exercise due care to avoid liability, i.e. the care level is chosen according respectively to first order condition (3) and (4), yielding a lower level of care than social optimum for both of them, i.e. $\underline{x}^* < x^{**}$ and $\underline{y}^* < \overline{y}$. Since residual liability is on the victim and the tortfeasor does not realize his bias, the level of activity chosen by the tortfeasor will exceed the social optimum, and will be chosen according to first order condition: $V_w = 0$, i.e. $\underline{w}^* > w^{**}$. Symmetrically, the victim anticipates the mistake of the tortfeasor and under the anticipation he will be compensated the victim chooses a higher level of activity according to first order condition: $V_z = 0$, i.e. $\underline{z}^* > z^{**}$.

Under bias, forgiveness does not change the optimal behaviour of tortfeasor (since he will not be held liable under forgiveness strategies), i.e. $\underline{x}^* < x^{**}$ and $\underline{w}^* > w^{**}$, while the victim corrects the choice of the activity level, as a second best efficient level in response of the tortfeasor's choice and overcompensate the lower level of care of the tortfeasor by increasing his care level above the second best efficient level in absence of the bias, i.e. $\underline{y}^* < y^*$.

Proof of Proposition 5

In presence of overconfidence, the tortfeasor will choose care and activity level according to the following first order conditions:

$$-zp_x^I(\underline{x}^*, 0)L = 1$$

$$V_w = zp^I(\underline{x}^*, 0)L$$

The tortfeasor will exert a lower care: $\underline{x}^* < x^*$. Symmetrically, the level of activity chosen by the tortfeasor is higher than socially optimal, i.e. $\underline{w}^* > w^*$.

$$\underline{x}^T = \underline{x}^F < x^{**}$$

$$\underline{w}^T = \underline{w}^F > w^{**}$$

For the victim

$$\underline{y}^T = \underline{y}^F = y^* = 0 < y^{**}$$

$$\underline{z}^T = \underline{z}^F = z^* > z^{**}$$

Proof of Proposition 6

In case of threat strategies, the victim will wish to exercise due care to avoid liability, i.e. the care level is chosen according respectively to first order condition (4). Due to the overconfidence bias, however, the victim chooses a level of care lower than social optimum, i.e. $\underline{y}^* < y^{**}$. Because the victim does not realize her own mistake and feeling he will be compensated, the victim chooses a higher level of activity according to first order condition: $V_z = 0$, i.e. $\underline{z}^* > z^{**}$. The tortfeasor realizes victim's mistake due to bias and set a level of care equal to zero, i.e. $z^* = 0$ and chooses a level of activity exceeding the social optimum, according to first order condition: $V_w = 0$, i.e. $\underline{w}^* > w^{**}$.

Under bias, forgiveness does not change the optimal behaviour of victim (since he will not be held liable under forgiveness strategies), i.e. $\underline{y}^* < y^{**}$ and $\underline{z}^* > z^{**}$. This implies that the tortfeasor realizes fully he will be held liable and corrects the choice of care to be second best efficient in response of the victim's choice, inducing an overcompensation of the lower level of care of the victim, thereby increasing his care level above the second best efficient level in absence of the bias, i.e. $\underline{x}^* > x^*$. Symmetrically, the activity level chosen by the tortfeasor is second best efficient given the choice of the victim and taking into consideration the lower care, $\underline{w}^* < w^*$.

Proof of Proposition 7

In case of threat strategies, both the victim and the tortfeasor will wish to exercise due care to avoid liability, i.e. the care level is chosen according respectively to first order conditions (3) and (4). The equilibrium is analogous to the rule of negligence with the defense of comparative negligence, but here the residual liability is on the tortfeasor. Due to the overconfidence bias, however, both the victim and the tortfeasor choose a level of care lower than social optimum, i.e. $\underline{y}^* < y^{**}$ and $\underline{x}^* < x^{**}$. Because the victim does not realize her own mistake and feeling he will be compensated, the victim chooses a higher level of activity according to first order condition: $V_z = 0$, i.e. $\underline{z}^* > z^{**}$. The tortfeasor chooses a level of activity exceeding the social optimum, according to first order condition: $V_w = 0$, i.e. $\underline{w}^* > w^{**}$.

Under bias, forgiveness does not change the optimal behaviour of victim (since he will not be held liable under forgiveness strategies), i.e. $\underline{y}^* < y^{**}$ and $\underline{z}^* > z^{**}$. This implies that the tortfeasor realizes fully he will be held liable and corrects the choice of care to be second best efficient in response of the victim's choice, inducing an overcompensation of the lower level of care of the victim, thereby increasing his care level above the second best efficient level in absence of the bias, i.e. $\underline{x}^* > x^*$. Symmetrically, the activity level chosen by the tortfeasor is second best efficient given the choice of the victim and taking into consideration the lower care, $\underline{w}^* < w^*$.

Chapter 5: Tort Law and Overconfidence Applications in Products Liability

"Some degree of negligence by users is always foreseeable ... a producer has a duty to safeguard all users, including 'klutzes and idiots'."

Klopp v. Wackenhut, 113 N.M. 153

1. Introduction

Although the terminology of behavioural economics has not yet permeated case law, the problems identified by behavioural psychologists and addressed by behavioural economists are well known by courts and in many ways influence contemporary judicial thinking. In this chapter, I survey cases from US in which the courts address situations where one or more parties were involved in a tort accident due to their optimism bias. The cases discussed in this chapter, although not directly utilizing the terminology of behavioural economists, clearly address situations where an individual's unrealistic optimism or excessive confidence in his skills was at the origin of a tort accident. Examples range from situations where a warning was not heeded because the person involved thought himself to be very experienced, or just optimistically dismissed the thought that they could be injured by whatever the risk was. These situations seem to come up most frequently in cases where comparative negligence and contributory negligence are involved. Some of these cases have also attracted the attention of legal academics and been cited in law review articles considering the legal remedies for optimism bias.

In Section 2 below, I shall discuss cases where construed errors due to overconfidence and unrealistic optimism as negligence, imposing liability for such biases. These are cases, where the court chooses "treat strategies" to induce the parties to correct for such biases. In Section 3, I discuss the gradual trend of case law towards "forgiveness strategies". In Section 4, I discuss the cases where forgiveness strategies are fully implemented and extended to areas outside of products liability, including premises liability, wrongful death, and negligence. In Section 5, I will also look at the different treatment of overconfidence in other areas of tort law involving unilateral care, such as medical malpractice law.

2. "Threat Strategies" for Consumer's Overconfidence in Products Liability Law

In the case *Skyhook Corporation v. John Jasper*, 90 N.M 143, the rule of strict liability was not applied in this case, because the court found that the product wasn't unreasonably dangerous when used according to the instructions and taking into account the warnings. The warnings were probably not read because the users were experienced crane operators who assumed they could use the crane correctly. However, the court does not discuss this directly. It did not discuss the possibility that nobody read the signs due to optimism bias but we can infer this because of the comments made about the decedents being experienced crane operators who ignored the signs.

This was an action for wrongful death, and the trial court found for the decedent while the appeals court overturned this and found for Skyhook.[22] The decedent worked as a sign installer.[23] He was using a crane, manufactured by Skyhook, to lift the pole for a sign into a hole.[24] There was a clear warning on the crane which said: "All equipment shall be so positioned, equipped or

[22] Skyhook v. Jasper at 144.
[23] Id.
[24] Id.

protected so no part shall be capable of coming within ten feet of high voltage lines."[25] There was evidence that the decedent knew of the sign and that he knew of the presence of high voltage lines.[26]There were also other means, such as use of a non-conductive guide rope, that were well known in the industry and which could have prevented the electrocution, which killed the decedent.[27] The decedent's estate sued, arguing that Skyhook could have installed safety equipment to prevent this accident.[28] Evidence showed that no other manufacturer in the industry installed this equipment as a matter of course, and that the decedent could have purchased and installed it himself post-market.[29]

The only issue the court considered was whether the crane was in a defective condition, making it unreasonably dangerous to use.[30] To do this, it decided whether failure to include an optional safety device could be considered a "defective condition."[31] The test used is whether, without the safety device, the product is unreasonably dangerous.[32] The court noted that almost everything is dangerous if used improperly and that "hindsight bias" must be avoided.[33] Since the crane had been used for five years with no problems, the court decided it was not unreasonably dangerous.[34] Also, there was a clear warning not to use the crane around high voltage lines, which caused this accident.[35]

In the case *Halvorson v. American Hoist and Derrick Company, et al.*, 307 Minn. 48; 240 N.W.2d 303, the court decides in a way that is very similar to the Skyhook holding, affirming that there is no negligence if the plaintiff is a victim of his own overconfidence and fails to read, or else to adjust behaviour, according to warnings that could have prevented the accident. Although the

---

[25] Id.
[26] Id. at 145.
[27] Id.
[28] Id.
[29] Id.
[30] Id. at 147.
[31] Id.
[32] Id.
[33] Id.
[34] Id.
[35] Id.

court doesn't use explicit terminology from behavioural economics, it seems clear from the facts of the case that the plaintiff was overly optimistic and failed to read the user's manual and to pay attention to the warnings, due to his overly confident attitude.

The specific facts of the case are as follows. On June 7, 1968, the plaintiff was employed on a highway construction crew, which was surfacing a rural highway turn lane near Duluth, Minnesota.[36] The surfacing was accomplished by pouring concrete from a portable mixer and levelling it with a machine called a screed, which was lifted in and out of the equipment by a crane manufactured by the defendant.[37] Halvorson was near the crane and noticed the screed was about to strike something, so he touched the screed and received an electric shock because the crane was contacting a power line.[38] He was severely injured but did not die, and he sued the defendant saying that there was insufficient safety equipment.[39] There was conflicting expert testimony and experimental evidence as to the effectiveness of these devices in preventing electrocution.[40] The plaintiff testified that he knew power lines were dangerous and that he should check for them, but he did not.[41] The instruction manual for the crane warned about power lines.[42]The jury found no strict liability but did find the defendant was 25% at fault.[43] Jury instructions stated that there should be strict liability if a condition was unreasonably dangerous, meaning that it is dangerous when used by an "ordinary user who uses it with knowledge common to the community as to the products' characteristics and common usage."[44] A product is not unreasonable dangerous if it is safe when used in accordance with all instructions and warnings.[45]The appeals court held that it was inconsistent for the jury to say that the product was not dangerous and defective in the absence of

[36] Halvorson at 304.
[37] Id.
[38] Id. at 305.
[39] Id.
[40] Id.
[41] Id.
[42] Id.
[43] Id.
[44] Id. at 306
[45] Id.

safety devices (as would be required to find strict liability) but at the same time to say that it was negligent to manufacture the crane without the devices.[46] The appeals court held that the plaintiff could not recover as a matter of law on either the theory of strict liability or of negligence, so no new trial was necessary.[47]

3. Towards a "Forgiveness Rule" in Products Liability Law

In the case *Arnold Holm, Appellant, v. Sponco Mfg., Inc.*, 324 N.W.2d 207, the court took a first step away from the traditional treat strategies, holding that a consumer should get the benefit of the doubt when there is a new technology, even though the consumer may have negligently failed to read the manual where warnings were given. This means that the manufacturer should take more care when introducing a new technology to ensure that there are fewer dangerous defects, or to warn as much as possible about any risks. This case mentions that people think they will be able to use a new technology safely, but again the words "optimism bias" are not used. This was not a new technology to the user, but the court also discusses the fact that this user had used this technology hundreds of times and was aware of the risks. Holm was severely injured on September 10, 1973, when he came in contact with a high voltage power line while operating an aerial ladder manufactured by the plaintiff.[48] Holm sued based on theories of negligence and product liability.[49] Holm admitted he knew of the danger and that the danger was obvious, and that there were warnings against the danger.[50] The trial court looked to Halvorson (above) to find for the defendant.[51] The aerial ladder had an automatic worker's platform at the top section of the ladder with a remote control device that permitted operation of the ladder from the platform.[52] Holm was

[46] Id. at 307
[47] Id. at 308
[48] Holm at 208
[49] Id.
[50] Id.
[51] Id.
[52] Id.

familiar with the ladder and had received training on its use.[53] He had used it thousands of times and claimed he could operate it without looking at the instructions.[54] Prior to the accident, Holm had had difficulty positioning the aerial ladder because it would drift after the power to it was turned off.[55] This condition was instrumental in causing the accident[56]. Holm filed an affidavit of an expert witness, testifying that the aerial ladder was defective and unreasonably dangerous because it lacked safety devices such as insulation, sensors, and other limiting or proximity warning devices which would have either warned appellant of the proximity of the electrical wires or would have prevented the electrical current from passing through appellant to ground.[57] The issue on appeal was whether the manufacturer of an aerial ladder in a defective condition unreasonably dangerous to the user is liable to the user if that defective condition is obvious.[58]The appeals court noted that Halvorson (discussed above) adopted the latent-patent danger rule that relieves a manufacturer from liability if the dangers of his product are obvious to the user.[59] Holm wanted Halvorson to be overruled, "contending (1) that it already has been implicitly overruled, (2) that the latent-patent rule is inconsistent with the current trend in products liability law, (3) that the policy considerations underpinning strict product liability are not served by adherence to Halvorson, and (4) that the latent-patent rule alters the basic allocation of liability under Minnesota's comparative fault statute.[60] Though the district court did use Halvorson, it wanted the appeals court to affirm the legitimacy or illegitimacy of Halvorson to provide more guidance.[61]

The case *Klopp v. Wackenhut*, 113 N.M. 153, changed the rule adopted in the previously discussed case of *Skyhook*. This case involves a combination of products liability and premises liability issues and sets an important precedent that applies to both areas of the law. The decision

---

[53] Id.
[54] Id.
[55] Id. at 209
[56] Id.
[57] Id.
[58] Id.
[59] Id. at 210.
[60] Id.
[61] Id.

stands for the proposition that some safeguards must always be taken, on behalf of prospective victims, no matter how obvious the risk may be for the population at large. This case is particularly important for our analysis, since it innovates in the law, adopting a forgiveness rule for the possible overconfidence errors of victims. As a result of this case, manufacturers and property owners need to take into account the optimistic and overconfident behaviour of users and visitors. Even if a hazard is obvious and can be easily recognized and avoided by the average person, the manufacturer and the owner of premises open to the public cannot assume all consumers and visitors to behave like an average person. Not every person can be average in cognitive skills and foresight. Hence, something that may seem obvious to most may not be obvious to a few. Producers and owners hence need to take precautions also for those few who will not take precautions due to erroneous beliefs driven by overconfidence or unrealistic optimism. As a result of this case, the law forgives those who become victims due to their own cognitive biases, and shifts the burden of correcting for the victims' shortcomings on producers and premises owners, who now face the burden of mitigate risk even when the hazards are obvious and avoidable for the most.

In *Klopp,* the plaintiff was suing for personal injuries sustained when tripping over a metal detector at an airport security station.[62] The issue is whether the obvious danger doctrine was abrogated by the comparative negligence of the plaintiff.[63] The plaintiff was proceeding through the Albuquerque International Airport to board an airplane when she passed through an airport security station.[64]While going through the metal detector, the plaintiff tripped over the protruding stanchion base of the metal detector.[65]She was preoccupied by the retrieval her things and was distracted at the time she tripped over the base.[66] The defendant argued that the base was open and obvious and there was no reason to believe it presented a danger.[67] The plaintiff contended that the "open and

[62] Klopp at 155.
[63] Id.
[64] Id.
[65] Id.
[66] Id.
[67] Id.

obvious" requirement was basically a contributory negligence rule that was not supposed to be appropriate in this jurisdiction.[68]

The plaintiff also contended that there was a question as to whether the base was an unreasonable risk of danger.[69]The court held that the defendant had the duty to safeguard all users, including "klutzes and idiots."[70]The court states that a risk cannot be made reasonable just by making it open and obvious, there also needs to be reasonable care to safeguard against the risk that may be undertaken by overconfident users.[71] Some degree of negligence by users is always foreseeable.[72]If the contributory negligence of a business visitor has been so extraordinary as to have been unforeseeable, then it would not have been a breach of duty for the occupier not to have taken precautions against an open and obvious risk.[73] If the contributory negligence was foreseeable, then the defendant would have had the duty to take ordinary care to protect even that person.[74]

Prior to Klopp's accident, not one person out of an estimated 540,000 people passing through this station had ever fallen, slipped, or tripped while passing through the metal detectors.[75] The court was not persuaded, however, that this meant the danger was unforeseeable by the defendant.[76] Also, the fact that she was distracted wasn't extraordinary enough to be considered unforeseeable by the defendant.[77]

The following trend in products liability law has been to find that the obviousness of the defect is only a factor to be considered as a mitigating defense in determining whether a defect is unreasonably dangerous and whether plaintiff used that degree of reasonable care required by the

---

68 Id. at 156.
69 Id.
70 Id. at 157
71 Id. THIS OVERRULES SKYHOOK which says if you warn enough, you don't need to safeguard against the risk
72 Id.
73 Id. at 158.
74 Id. at 159.
75 Id. at 160.
76 Id.
77 Id.

circumstances.[78] This court thinks that, given the pace of technology, people may not realize the scope of dangers as much as they could.[79] Manufacturers are in a superior position to recognize and cure defects, for improper conduct in the placement of finished products into the channels of commerce furthers the public interest.[80] The court held that "a manufacturer is obligated to exercise that degree of care in his plan or design so as to avoid any unreasonable risk of harm to anyone who is likely to be exposed to the danger when the product is used in the manner for which the product was intended, as well as an unintended yet reasonably foreseeable use."[81]

In 1970, Minnesota adopted a comparative negligence statute. Minn. Stat. § 604.01 (1970).[82] The statute also includes an assumption of the risk defense and covers strict liability.[83] As a result, contributory negligence (other than the failure to inspect a product or to guard against defects), misuse, and assumption of the risk were specifically recognized as valid defenses.[84] In Minnesota, contributory fault, including "unreasonable assumption of risk" and "unreasonable failure to avoid an injury," bars recovery only when it is greater than the fault of the party from whom recovery is being sought.[85] Previously there was a latent-patent rule that made obviousness of a danger a complete bar to recovery.[86] This result was found by this court to be contrary to public policy because it didn't apportion loss to blameworthy defendants.[87] The case is remanded for the trial court to determine if the defendant took reasonable care, even though the risk was obvious.[88]

4. Applications of a "Forgiveness Rule" in Recent Tort Cases

---

[78] Id. at 211.
[79] Id. at 212.
[80] Id.
[81] Id.
[82] Id. at 213.
[83] Id.
[84] Id.
[85] Id.
[86] Id.
[87] Id.
[88] Id.

The case of *Benavidez v. City of Gallup*, 131 N.M. 808, applies the forgiveness rule in a case of premises liability. The court states that, even if a risk is obvious, if the plaintiff had an optimism bias that would cause him to ignore the risk, and the property owner could have avoided the obvious risk by taking reasonable care, then the property owner will be liable for the injury. In *Benavidez,* the plaintiff sued the city of Gallup after tripping over a water meter and breaking her ankle.[89] The trial court ruled in favour of the city and plaintiff appealed.[90] The plaintiff requested a jury instruction that a property owner owes a visitor a duty of care whether or not a dangerous condition is obvious, if he could have discovered the danger with a reasonable inspection.[91] The court cited Klopp for the proposition that property owners cannot avoid liability for injuries for which the risk was obvious.[92] The duty to take reasonable care applies whether or not the hazard involved was obvious.[93]The court held that that jury instruction should have been made and that the omission of it did in fact harm the plaintiff.[94]

The case of *Stetz v. Skaggs Drug Centers*, 114 N.M. 465, extends the forgiveness rule and restates it in a language that signals the recognition of a broader principle of tort law, applicable to all negligence cases. In *Stetz,* the court stated that, even if the plaintiff has an optimism bias, and is therefore contributorily negligent, the defendant may still be liable. If the contributory negligence is foreseeable and not extraordinary, then the defendant should bear the burden of precaution of his overconfident victim, planning ahead to compensate for the untaken precautions of his victim. Failing to do so could constitute negligence on his part.

In *Stetz*, the plaintiff was injured when she fell on a sidewalk leading to the Skaggs store in the Fair Plaza Shopping Center while on her way into the store.[95] An area of the sidewalk was in

[89] Benavidez at 810.
[90] Id.
[91] Id.
[92] Id. at 814.
[93] Id. at 815.
[94] Id. at 815.
[95] Stetz at 466.

disrepair.[96] The plaintiff did not contest that this disrepair was obvious.[97] The trial court determined that the plaintiff was comparatively negligent in trying, but apparently failing, to walk around it.[98]However, the defendants also argue that because the disrepair was obvious, they had no duty to warn or to take measures to protect from the danger.[99] The court cited Klopp for the proposition that neither the open and obvious nature of a defect nor the injured party's own negligence constitutes an automatic bar to recovery.[100]The plaintiff's contributory negligence must be so extraordinary as to be unforeseeable in order for the defendant to escape liability for failure to protect against an open and obvious risk.[101] Therefore a court must determine whether contributory negligence of the plaintiff is so extraordinary as to be unforeseeable.[102] The trial court determined that the contributory negligence was foreseeable because the area had been in disrepair for a year and even the manager admitted the area was dangerous.[103]

In *Emery v. Federated Foods*, 262 Mont. 83, the court held that a manufacturer has to take into account optimism bias if the food that it manufactures has a special danger to it. The case follows the trend in case law by establishing that the manufacturer should take into account not only diligent consumers, but also the average "sloppy" consumer, using language that resembles that used by the other court with respect to the incident with an airport security detector. In *Emery,* the plaintiff lived with her two children.[104] On her way home from work, she bought a bag of marshmallows, and scanned the label before purchasing them.[105] She placed them out of reach of the children.[106]The next day a friend came over while the plaintiff was sleeping, and the children

[96] Id.
[97] Id.
[98] Id. at 467.
[99] Id.
[100] Id.
[101] Id.
[102] Id.
[103] Id.
[104] Emery at 86.
[105] Id.
[106] Id.

convinced her to give them marshmallows.[107] One of the children began choking and the friend was unable to dislodge the marshmallow.[108] They took the child to the hospital, but by the time they dislodged the marshmallow the child had suffered brain damage due to lack of oxygen.[109]The mother sued the manufacturer of the marshmallow, claiming that they were defective and dangerous, and that there was no warning about dangers of inhaling marshmallows.[110]The trial court found that the defendant was not under a duty to warn that infants and toddlers are in danger of choking when consuming large quantities of marshmallows at a time.[111]The reasoning was that warning is not required if the danger is generally known and recognized.[112]Medical experts testified that marshmallows are particularly dangerous because they swell when wet and can become lodged in the throat if eaten too quickly.[113] Children are particularly at danger because they are often unable to control themselves when eating candy.[114] The plaintiff claimed that she always read safety-warning labels, especially when they involved small children.[115] The plaintiff also claimed she had no idea marshmallows posed any particular danger of choking, and a doctor testified that this was a common misperception among parents.[116]

The case *Metzgar v. Playskool,* 30 F.3d 459 involved the overconfidence and unrealistic optimism of parents when assessing the abilities and maturity of their children. In *Metzgar,* the court applied the forgiveness rule also with respect to parents. The holding of this case should be interpreted restrictively inasmuch as it does not intend to forgive parents by shifting the residual loss on their children, but rather to forgive parents by shifting the liability on manufacturers who had an opportunity to anticipate and to correct for parents' biases. The bias

---

107 Id.
108 Id. at 87.
109 Id.
110 Id.
111 Id. at 90
112 Id.
113 Id. at 91.
114 Id.
115 Id. at 93.
116 Id.

considered in *Metzgar* is the one involving parents who systematically believe that their children are smarter than their age. The court holds that this perception bias by parents should be 'forgiven" and that manufacturers should account for this systematic bias by giving clearer warnings of dangers. In *Metzgar,* a toddler choked on a building block manufactured by the defendant.[117] The parents sued under negligence and strict product liability rules and the trial court found for the defendant.[118] The appeals court disagrees with the trial court's finding that the risk of choking was so small that a warning label was not necessary.[119] The appeals court also rejects the trial courts finding that the age guideline on the product packaging precludes the manufacturer's liability for safety when used by children who are within that age category.[120]Finally the appeals court disagrees with the trial court in finding that the danger of a small child choking on the block was obvious so as to negate any duty by Playskool to so warn.[121]One morning, the child's father placed the child in his crib, left for five minutes, and came back to find the child dead.[122] The block the child choked on was the smallest block in the kit, and there was no warning label about choking hazard.[123] The label said the blocks were for children older than 18 months, but the child who died was 15 months old.[124]The parents sued Playskool for negligent design and strict product liability.[125] The parents claim that the age warning wasn't sufficient to warn of choking hazard.[126] The trial court found that the danger of choking was foreseeable but so small that the risk was not unreasonable.[127]The trial court reasoned that no warning is necessary where a risk of danger is obvious, and the likelihood of a small child choking on a small block was obvious.[128]Playskool should have taken into account whether parents are adequately alerted to the danger by an age recommendation for use of a toy, since many parents

[117] Metzgar at 460.
[118] Id.
[119] Id.
[120] Id.
[121] Id.
[122] Id.
[123] Id. at 461.
[124] Id.
[125] Id.
[126] Id.
[127] Id.
[128] Id.

think their child is developmentally smarter than his age.[129]The hazard of choking on the blocks was not obvious enough that no warning should be required.[130]

In *Johnson v. Southern Minnesota Machinery Sales*, 442 N.W.2d 843, the Court held that a manufacturer has a duty to warn of dangers that an inexperienced person might not recognize but which an experienced person might. The inexperienced person exercised optimism bias because, on watching a brief demonstration, he thought he could do the task. But, as it turns out, he did not know of all the risks. The plaintiff alleged that a table saw was defective and unreasonably dangerous.[131] The plaintiff was asked to use some new equipment, including the table saw.[132] The saw had a standard blade guard and an optional blade guard, but the plaintiff's employer decided not to purchase the optional blade guard.[133] The blade guard was removed when the plaintiff used it.[134] The plaintiff received training right before use of the saw, but had little experience with table saws in general.[135] The plaintiff lost four fingers in the saw shortly thereafter.[136]Expert witnesses testified that, in their opinion, the Model 66 table saw was defective and unreasonably dangerous with the standard guard and that if the optional guard had been in place there would have been no injury.[137] The plaintiff was using the saw for freehand cutting that was warned against, yet foreseen, by the manufacturer.[138]The defendant wanted the court to find that the plaintiff was negligent in the use of the saw.[139] The appeals court found that the plaintiff was using the saw exactly as trained, but did not have the experience to recognize every possible risk.[140]

---

[129] Id. at 465.
[130] Id.
[131] Johnson at 845.
[132] Id.
[133] Id.
[134] Id.
[135] Id.
[136] Id.
[137] Id.
[138] Id. at 846.
[139] Id. at 847.
[140] Id.

In *Jonathan v. Kvaal*, 403 N.W.2d 256, the court addressed a situation of optimism bias of good swimmers and athletic people, who actually are more prone to diving injuries. Jonathan was one of three tenants in Kvaal's home at the time of the accident in August 1980.[141]There was a pool in the backyard, with depth markings, a minimum depth of four feet, and a maximum depth of seven feet.[142]The pool also contained a sign warning against jumping and diving.[143]Jonathan used the pool on a regular basis, at least 10 times prior to the accident. [144]Kvaal acknowledged that there was no nighttime lighting at the pool.[145]Jonathan acknowledged by deposition that he was aware of the warning sign and the depth markers.[146]The night of the accident, Jonathan had been drinking heavily.[147] Jonathan dove into the shallow end and sustained an injury that left him a quadriplegic.[148] Experts said that this sort of pool had risks unknown to the general public, but the trial court held that Jonathan was negligent in diving when he did and that he knew a lot about the construction of the pool.[149] It was well known in the industry that this sort of pool was especially prone to being involved in injuries such as Jonathan received.[150] There was also data showing that the more experienced and fit swimmers were more prone to these injuries because of their longer trajectory during dives.[151]The court followed *Holm* (above) in stating that a manufacturer is not relieved of liability under the latent-patent defect rule simply because dangers associated with the product are obvious to the user.[152]It also cited a New York case for the proposition that a "manufacturer is obligated to exercise that degree of care in his plan or design so as to avoid any unreasonable risk of harm to anyone who is likely to be exposed to the danger when the product is used in the manner for which the product was intended, as well as an unintended yet reasonable

141 Jonathan at 258.
142 Id.
143
144 Id.
145 Id.
146 Id.
147 Id.
148 Id.
149 Id. at 259.
150 Id.
151 Id. at 260.
152 Id. at 261.

foreseeable use."[153] The court found that Jonathan was not aware of the extent of the danger of this type of pool, even though he had helped install it.[154]

5. The Limits of Forgiveness: Medical Malpractice and Unilateral Care Cases

In addition to the cases discussed below, there are several cases in the areas of medical malpractice. The rulings in these cases are quite at odds with the "forgiveness" regime identified in the previous cases. The point frequently faced by courts is that doctors may be overconfident, leading the patient to sign consent forms that he might not have otherwise.[155]In these cases, the patient is led to underestimate or disregard a risk because of the doctor's overconfidence. Then, when the procedure does go wrong there is an issue of how much the doctor should have warned of every danger. These cases, although having some similarities with the cases examined in Section 3 and 4, are generally subjected to a different treatment by courts.

The application of a "threat strategy" in medical malpractice cases is explainable by the findings of our model of liability, inasmuch as medical malpractice generally represents a unilateral-care problem: patients cannot take precautions that can compensate for the untaken precautions of their doctors. For this reason, a forgiveness rule in medical malpractice cases would be undesirable, since it would undermine the incentives of doctors, without much an opportunity for patients to compensate their doctors' shortcomings with their own efforts. In the specific case of doctors' disclosure of risk to their patients, it is likewise true that doctors have an informational

[153] Id.
[154] Id. at 262.
[155] A law review article considers the problem of overconfidence of doctors when explaining the risks of alternative medical procedures to their patients: THE LAW AND POLITICS OF TORT REFORM: TORT REFORM SYMPOSIUM COMMENT: Not-So-Peaceful Coexistence: Inherent Tensions in Addressing Tort Reform, 4 Nev. L.J. 337

advantage over patients and it would not be wise to shift the burden of searching information from doctors to patients, by means of forgiveness of the doctor's overconfidence.

Conclusion

The aim of the thesis was to explore how certain behavioural biases affect decision-making. We focused on two contexts, the bandit problem scenario and the case of legal decision-making.

In regard to bandit problems, the focus of interest has been to examine the role of risk aversion and loss aversion, which are both excluded from the standard literature on bandit problems. In Chapter 2 we have examined the bandit problem under alternative models of behaviour, from the standard expected utility model with risk neutral agents to models with risk aversion, with loss aversion and with risk and loss aversion. We have departed from the classical setting of bandit problems by endowing the agent with a disappointment-elation utility function (Sugden and Loomes, 1986), assuming that the player aims at maximizing the expected utility of profits, instead of assuming the maximization of expected profits as an objective function. According to Sugden and Loomes (1986), the individual receives the utility derived directly from the actual consequence of an uncertain prospect, but he also feels some degree of disappointment and elation – measured as a deviation from the a priori expected payoff. If the actual consequence turns out to be worse than the expectation, he feels disappointment, while he experiences some degree of elation, if the actual consequence is better than the a priori expectation. We model the disappointment-elation utility function to capture the spirit of loss-aversion (Kahneman and Tversky, 1979): the disutility of a loss is greater than the elation associated with a same-sized gain.

Chapter 2 investigated the optimal experimentation strategy of a loss averse player and whether he will engage in more experimentation. This question is addressed in a traditional one-armed bandit problem, where the player's decision problem has the same characteristics of the standard one. The player faces the decision between an unknown arm and a safe one – in the TV show language the player should decide in each round whether to keep the case (risky arm) or accept the banker's offer (safe arm) and walk away with a sure prize. Analytically the player faces an optimal stopping problem, i.e. he has to decide how many times to experiment with the unknown arm before switching to the safe one, so that once switched the agent won't return to the unknown arm again. Such experimentation is costly: in the short run, the agent bears the loss in case of a negative result. This loss must be traded off against the potential informational gain associated with experimentation, in terms of a more correct estimate of the payoffs obtainable when undertaking the risky project. In the standard framework, the agent's optimal strategy is found by weighing these two factors.

Common intuition would lead us to think that a loss-averse player would experiment less and choose to opt out the game before a loss-neutral player. In Chapter 2 we show that this standard intuition holds true only for individuals who exhibit a high degree of loss aversion. On the contrary, players who are moderately loss-averse will choose to experiment more. At every stage of the game, the player trades off the immediate cost of experimentation – due to the possible disappointment experienced in the short run when the failure occurs – with the long-run benefits from experimentation, which are measured by the presence of the additional benefit of elation in case of a success. The understanding of the long-run benefits of experimentation is important at this point. There are three main benefits that the player can obtain from experimentation. First, experimentation provides information on the payouts of the risky arm. Second, experimentation may provide elation when the attempt proves successful. Third, experimentation may "spare" disappointment in the long run due to the choice of a more profitable arm. For high levels of loss-aversion, the immediate short run cost of disappointment weighs more and may induce the player to opt for the safe arm, avoiding any disappointment cost even in the short run. A moderately loss-averse player, however, may be willing to do more experimentation, because the additional gain represented by the elation component may outweigh the higher cost of disappointment due to the loss aversion in the short run.

In Chapter 3, an experimental study is conducted that involves a three-stage one-armed bandit problem. The main substantive finding in Chapter 3 is that there is a bias towards over-experimentation in stages 1 and 2, and a bias towards under-experimentation in stage 3, relative to the theoretical model, which is not explicable by reference to subjects' levels of risk or loss aversion. This remains a 'behavioural' puzzle. There is evidence of both over-experimentation and under-experimentation, and that this is true for subjects of all levels of risk aversion and loss aversion. At some stages in the process subjects of all types display over-experimentation, while at other stages there is a systematic bias towards under-experimentation.

In total 254 subjects – first year undergraduate students in Microeconomics - took part to the experiment. The experimental study involved for each subject the elicitation of individual preferences (measuring loss aversion and risk aversion in a within-subject design) and the attitude to participate to experimental decisions. The set-up of the experiment is designed as a one armed-bandit problem with a three-period horizon, where the player faces the decision between an unknown arm and a safe one and the player should decide in each round whether to keep experimenting on the risky arm or opting out of the game with a sure prize on the safe arm. In the experimental design we ask each subject whether they would accept to play on a slot machine at a pre-specified price and at which highest price they would accept to play. By doing so, we are able to characterize fully their experimentation decisions at each stage of the game for each participant at the experiment.

In regard to legal decision-making, we focused on the role of overconfidence in tort law. Psychological research shows that there is systematic overconfidence in risk judgments – which is one of the most widespread psychologically-generated biases in human judgment. Data shows that overconfidence bias is pervasive in judgments that individuals make regularly in everyday life, rather than in activities that are seldom carried out (Jolls and Sunstein, 2006). Overconfident individuals overestimate their own ability underestimating the risk that they face. Overconfidence is one of the two biases that result from what psychologists know as optimism bias. Optimism bias affects people's subjective estimates of the likelihood of future events, and causes them to overestimate the likelihood of positive or desirable events and to underestimate the likelihood of negative or undesirable events (Colman, 2001).

Overconfidence creates a distinctive problem for legal policymakers: even factually informed people tend to think that risks are less risky to materialize for themselves than for others (Sunstein, 2000). Overconfident people therefore inadequately react to legal threats and incentives such as liability rules in a number of areas of law, showing some surprising implications. The traditional law and economics view on issues of overconfidence is that the problem is caused by imperfect information and can be appropriately corrected through the provision of additional information (see, e.g., Stiglitz, 1986 on consumer optimism). However, as the extensive evidence suggests, overconfidence leads many individuals to underestimate their personal risks even if they receive accurate information about average risks. Evidence indeed suggests that debiasing strategies through risk education and information are only partially effective. Behavioural law and economics scholars have addressed the issue of overconfidence, considering the possible role of law in restraining or correcting this judgment error.

Chapter 4 considered the role of tort rules in debiasing overconfidence and illustrated the possible use of threat strategies (threatening liability when overconfidence leads to an accident) and forgiveness strategies (forego liability when the accident is solely caused by a biased perception of risk) as alternative ways in which law could be used to debias overconfidence. We compared the alternative threat and forgiveness strategies in reducing the cost of accidents due to overconfidence, allowing for the possibility that government investment in information fails to guarantee full debiasing of agents.

The model highlighted the role of tort law and the optimal design of liability rules for correcting overconfidence biases. The analysis suggested that the presence of this bias might affect outcomes in a manner that has practical consequence and that taken such potential bias into consideration when forming legislation may be desirable. Legal forgiveness of overconfidence may be a valuable second-best solution, when debiasing through tort law cannot be effectively achieved.

The most effective way to correct overconfidence in tort law may well be to forgive it, rather than to penalize it through liability. Products regulations that impose liability on producers for not increasing the safety of their products in anticipation of consumer's overconfidence can be viewed as a legal strategy analogous to legal forgiveness of overconfidence (consumer are effectively shielded against the consequences of their overconfident errors).

In Chapter 5 we looked at applications in product liability law to discuss the relevance of forgiveness and threat strategies in tort law. Although the terminology of behavioural economics has not yet permeated case law, the problems identified by behavioural psychologists and addressed by behavioural economists are well known by courts and in many ways influence contemporary judicial thinking. We found a number of cases from US in which the courts address situations where one or more parties were involved in a tort accident due to their optimism bias. Examples range from situations where a warning was not heeded because the person involved thought himself to be very experienced, or just optimistically dismissed the thought that they could be injured by whatever the risk was. These situations seem to come up most frequently in cases where comparative negligence and contributory negligence are involved. We discussed a gradual trend of case law towards "forgiveness strategies", that goes beyond the case of product liability law, to include premises liability, wrongful death, and negligence.

References

Abdellaoui, Mohammed, Han Bleichrodt, and Corina Paraschiv, "Loss Aversion under Prospect Theory: A Parameter-Free Measurement," *Management Science*, (in print)

Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David I. Laibson, 2007, "The Age of Reason: Financial Decisions Over the Lifecycle," *MIT Department of Economics Working Paper No. 07-11*

Agresti, A. (2002) Categorical data analysis, Second edition, New York: Wiley. Pages 91-101.

Ainslie G., 1992, Picoeconomics, the strategic interaction of successive motivational states within the person, Cambridge, Cambridge University Press

Anderson, Cristopher M. (2001), Behavioural Models of Strategies in Multi-Armed Bandit Problems, mimeo

Anscombe F.J., Aumann R., 1963, A Definition of Subjective Probability, *Annals of Mathematical Statistics*, 34, 199-205

Armour, David, and Shelley E. Taylor, 2002, "When Predictions Fail: The Dilemma of Unrealistic Optimism" in Heuristics and Biases: The Psychology of Intuitive Judgment, Thomas Gilovich, Dale Griffin, and Daniel Kahneman (eds.), New York, Cambridge University Press

Aumann, R. J. , 1990, "Nash Equilibria are Not Self-Enforcing", in Economic Decision Making: Games, Econometrics and Optimisation, ed. by J.J. Gabszewicz, J.F. Richard, and L. A. Wolsey, 201-206, Elsevier Science Publishers B.V.

Babcock, Linda, George Loewenstein, Samuel Issacharoff and Colin Camerer, 1995, "Biased Judgments of Fairness in Bargaining" 85 *American Economic Review* 1337-43.

Babcock, Linda, George Loewenstein and Samuel Issacharoff, 1997, "Creating Convergence: Debiasing Biased Litigants", 22 *Law and Social Inquiry* 913-925

Barberis, N. and M. Huang, 2001, "Mental accounting, loss aversion and individual stock returns", *Journal of Finance* 56:1247-1292

Barberis N., Huang, Santos, 2001, Prospect Theory and Asset Prices, *Quarterly Journal of Economics*, 116, 1-53

Barberis, N. and R. Thaler, 2003, Chapter 18 "A survey of behavioural finance", Handbook of the Economics of Finance, Elsevier Science

Bateman, Ian, Daniel Kahneman, Alistair Munro, Chris Starmer, and Robert Sugden, 2005, "Testing competing models of loss aversion: an adversarial collaboration," *Journal of Public Economics*, 89, 1561-1580

Bateman, Ian, Alistair Munro, Bruce Rhodes, Chris Starmer, and Robert Sugden, 1997, "A test of the theory of reference-dependent preferences," *Quarterly Journal of Economics*, 112, 479-505

Becker, Gordon M., Morris H. De Groot, and Jacob Marschak, 1964, "Measuring Utility by a Single-Response Sequential Method," *Behavioural Science*, 9, 226-232

Benartzi, Shlomo, and Richard H. Thaler, 1995, "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110, 73-92

Bernartzi S., Thaler R., Save More Tomorrow: Using Behavioural Economics in Increase Employee Savings, *Journal of Political Economy*

Berry, D and Fristedt B., 1985, "Bandit Problems: Sequential Allocation of Experiments" Chapman&Hall

Bigus, Jochen, 2006, "Tort Liability and Probability Weighting Function According to Prospect Theory", paper presented at the sixteenth annual meeting of the American Law and Economics Association, Berkeley, California

Bleichrodt, Han, and Jose Luis Pinto, 2000, "A parameter-free elicitation of the probability weighting function in medical decision analysis," *Management Science*, 46, 1485-1496

Booij, Adam S., and Gijs Van de Kuilen, 2006, "A Parameter-Free Analysis of the Utility of Money for the General Population under Prospect Theory," *mimeo, CREED, University of Amsterdam*

Bowman, D., D. Minehart, and Matthew. Rabin, 1999, "Loss aversion in a consumption-savings model," *Journal of Economic Behaviour & Organization*, 38, 155-178

Bradt, R. N., S. M. Johnson, and S. Karlin, 1956, "On Sequential Designs for Maximizing the Sum of n Observations", *Annals of Mathematical Statistics*, Volume 27, Issue 4 (Dec., 1956), 1060-1074

Brown, John P., 1973, "Toward an Economic Theory of Liability", 2 *Journal of Legal Studies*, 323–49

Camerer C., 1995, Individual Decision Making, in Kagel J., Roth A. (a cura di), Handbook of Experimental Economics, Princeton, Princeton, University Press, 587-703

Camerer C., 2003, Behavioural studies of strategic thinking in games, *Trends in Cognitive Sciences*, 7, 5, 225-231

Camerer C., 2003, Behavioural Game Theory, Princeton, Princeton University Press

Camerer, Colin F., "Prospect Theory in the Wild: Evidence from the Field," in Colin F. Camerer, George Loewenstein, and Matthew. Rabin, eds., *Advances in Behavioural Economics* (Princeton: Princeton University Press, 2004), 148-161

Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue and Matthew Rabin, 2003, "Regulation for Conservatives: Behavioural Economics and the Case for "Asymmetric Paternalism", 151 *University of Pennsylvania Law Review*, 1211-44

Camerer C., Thaler R.H., 2003, In Honor of Matthew Rabin: Winner of The John Bates Clark Medal, *Journal of Economic Perspectives*, 17, 3, 159-176

Camerer, Colin F., L. Babcock, G. Loewenstein, and Richard H. Thaler, 1997, "Labor supply of New York City cabdrivers: One day at a time," *Quarterly Journal of Economics*, 112, 407-441

Cantril, Hadley, 1938, "The Prediction of Social Events", 33 *Journal of Abnormal and Social Psychology*, 364-389

Carmon, Zvi, and Dan Ariely, "Focusing on the forgone: How value can appear so different to buyers and sellers," *Journal of Consumer Research*, 27 (2000), 360-370

Chen, M. Keith, Venkat Lakshminarayanan, and Laurie R. Santos, 2006, "How Basic Are Behavioural Biases? Evidence from Capuchin Monkey Trading Behaviour", *Journal of Political Economy*, 114, 517-537

Chua, Zhikang, and Colin F. Camerer, 2004, "Experiments on Intertemporal Consumption with Habit Formation and Social Learning," *mimeo, California Institute of Technology*

Coleman, Andrew W., 2001, "Unrealistic Optimism" A Dictionary of Psychology, Oxford University Press

Choquet, Gustave, 1954, "Theory of Capacities", 5 *Annales de l'institute Fourier*, 131–295.

Cubitt, Robin P., Chris Starmer, and Robert Sugden, 1998, "On the Validity of the Random Lottery Incentive System," *Experimental Economics*, 1, 115-131

Daniel, K., D. Hirshleifer and A. Subrahmanyam, 2001, "Overconfidence, Arbitrage and Equilibrium Asset Pricing", *Journal of Finance,* 56, 921-965

Dari Mattiacci, Giuseppe, 2005, "Errors and the Functioning of Tort Liability", 13 *Supreme Court Economic Review*, 165-187

Dari-Mattiacci, Giuseppe, and Gerrit De Geest, 2005, "Judgment Proofness Under Four Different Precaution Technologies," 161 *Journal of Institutional and Theoretical Economics* 38-56

De Groot, M. H, 1970, "Optimal Statistical Decisions" McGraw Hill

DeJoy, David M., (1989) "The Optimism Bias and Traffic Accident Risk Perception", 21 Accident Analysis and Prevention, 333–40.
Dickhaut, John Paolo Ghirardato, Kip Smithi, Jose' V. Pardo and Aldo Rustichini, 2002, "A Brain Imaging Study of Procedural Choice", mimeo

Dickhaut, John, Kevin McCabe, Jennifer C. Nagode, Aldo Rustichini, Kip Smithi, and Jose′ V. Pardo, 2004, "The impact of the certainty context on the process of choice. Forthcoming in the Proceedings of the National Academy of Sciences USA

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G.Wagner, 2005, "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey," *IZA Discussion Paper Series*

Edwards, W., 1968, "Conservatism in human information processing". In B. Kleinmutz (ed.), Formal Representation of Human Jugdment. New York: John Wiley and Sons

Eide, Erling, 2007, "Accident Liability with Rank Dependence", *Kritische Vierteljahresschrift fur Gesetzgebung und Rechtswissenschaft* 1/2:160–71

Eisenberger, Roselies, and Martin Weber, 1995, "Willingness-to-Pay and Willingness-to-Accept for Risky and Ambiguous Lotteries," *Journal of Risk and Uncertainty*, 10, 223-233

Fehr, Ernst, and Lorenz Goette, 2007, "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment," *American Economic Review*, 97 (2007), 298-317

Finn, Peter, and Barry W. E. Bragg, 1986, "Perception of the Risk of an Accident by Younger and Older Drivers", 18 *Accident Analysis and Prevention*, 289–98

Fischhoff, Baruch, 1982, "Debiasing" in Judgment under Uncertainty: Heuristics and Biases, Daniel Kahneman, Paul Slovic, and Amos Tversky (eds), New York, Cambridge University Press

Fischhoff, Baruch, 2002, "Heuristics and Biases in Application" in Heuristics and Biases: The Psychology of Intuitive Judgment, Thomas Gilovich, Dale Griffin, and Daniel Kahneman (eds.) New York, Cambridge University Press

Fisher, R. A., 1935, The logic of inductive inference, *Journal of the Royal Statistical Society Series A*, 98, 39-54.

Fisher, R. A., 1970, Statistical Methods for Research Workers, Oliver & Boyd

Forsythe, Robert, Thomas A. Rietz, Thomas W. Ross, 1999, "Wishes, Expectations and Actions: A Survey on Price Formation in Election Stock Markets" 39 *Journal of Economic Behaviour and Organization* 83-110

Genesove D., Mayer C., 2001, Loss Aversion and Seller Behaviour: Evidence from the Housing Market, *Quarterly Journal of Economics*, pp. 1233-1260

Gittins, J., 1979, "Bandit processes and dynamic allocation indices", *Journal of the Royal Statistical Society B*, Volume 41, 148-64

Gneezy, Uri, and Jan Potters, 1997, "An experiment on risk taking and evaluation periods", *Quarterly Journal of Economics*, 112, 631-645

Goette, Lorenz, David Huffman, and Ernst Fehr, 2004,"Loss Aversion and Labor Supply," *Journal of the European Economic Association*, 2, 216-228

Guppy, Andrew, 1992, "Subjective Probability of Accident and Apprehension in Relation to Self-Other Bias, Age, and Reported Behaviour", 25 *Accident Analysis and Prevention* 375–82

Jolls, Christine, 1998, "Behavioural Economic Analysis of Redistributive Legal Rules" 51 *Vanderbilt Law Review* 1653–77

Jolls, Christine, Cass R. Sunstein, and Richard Thaler, 1998, "A Behavioural Approach to Law and Economics" 50 *Stanford Law Review* 1471–1550

Jolls, Christine and Sunstein, Cass R., 2006, “Debiasing through Law”, 35 *Journal of Legal Studies* 199-241

Haigh, Michael S., and John A. List, 2005, "Do professional traders exhibit myopic loss aversion? An experimental analysis", *Journal of Finance*, 60, 523-534

Harbaugh, William, Kate Krause, and Lise Vesterlund, 2001, "Are adults better behaved than children? Age, experience, and the endowment effect," *Economics Letters*, 70, 175-181

Hardie, Bruce G. S., Eric J. Johnson, and Peter S. Fader, 1993, "Modeling Loss Aversion and Reference Dependence Effects on Brand Choice," *Marketing Science*, 12, 378-394

Harless, David W., 1989, "More Laboratory Evidence on the Disparity between Willingness to Pay and Compensation Demanded," *Journal of Economic Behaviour & Organization*, 11, 359-379

Hey, John D., Chapter 6, ”Experiments and the economics of individual decision making”

Horowitz, John K., and Kenneth E. McConnell, 2002, "A review of WTA/WTP studies", *Journal of Environmental Economics and Management*, 44, 426-447

Johnson, Eric J., Simon Gächter, and Andreas Herrmann, 2006, "Exploring the nature of loss aversion," *CeDEx Discussion Paper No. 2006-02*

Johnson, Eric J., and Daniel Goldstein, 2003, "Do Defaults Save Lives?", *Science*, 302, 1338-1339

Johnson, Eric J., Gerald Haeubl, and Anat Keinan, "Aspects of Endowment: A Query Theory of Value Construction," *Journal of Experimental Psychology - Learning Memory and Cognition*, (in press)

Hanson, Jon D. and Douglas A. Kysar, 1999a, “Taking Behavioralism Seriously: Some Evidence of Market Manipulation”, 112 *Harvard Law Review*, 1420-1572

Hanson, Jon D. and Douglas A. Kysar, 1999b, “Taking Behavioralism Seriously: The Problem of Market Manipulation”, 74 *New York University Law Review*, 630- 749

Kachelmeier, Steven J., and Mohamed Shehata, 1992, "Examining Risk Preferences under High Monetary Incentives - Experimental-Evidence from the Peoples-Republic-of-China," *American Economic Review*, 82, 1120-1141

Kahneman D., Knetsch J., Thaler R., 1986, Fairness and The Assumptions of Economics, *Journal of Business*, 285-300

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler, 1990, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 98, 1325-1348

Kahneman, D., Knetsch J., 1991, Anomalies. The Endowment Effect, Loss Aversion, and Status Quo Bias, *Journal of Economic Perspectives*, 5, 1,193-206

Kahneman D., Tversky A., 1973, On the Psychology of Prediction, *Psychological Review*, 80, 237-51

Kahneman, D. and A. Tversky, 1974, "Judgement under uncertainty: heuristics and biases", *Science* 185: 1124-1131

Kahneman D., Tversky A., 1979, A Prospect Theory: An Analysis of Decision Under Risk, *Econometrica*, 47, 2, pp. 263-291

Kahneman, D and A. Tversky, 1991, "Loss Aversion and Riskless Choice: A Reference Dependent Model", *Quarterly Journal of Economics*, Volume 104, 1039-61

Kahneman, D. and A. Tversky, 1992, "Advances in Prospect Theory: Cumulative Representation of Uncertainty", *Journal of Risk and Uncertainty*, Volume 5, 297-323

Kahneman, Daniel, Paul Slovic and Amos Tversky, 1982, Judgement under Uncertainty: Heuristics and Biases, edited by Daniel Kahnemann, Paul Slovic, and Amos Tversky. Cambridge: Cambridge University Press

Kahneman, D., Tversky A., Thaler R., 1990, Experimental Tests of the Endowment Effect and the Coase Theorem, *Journal of Political Economy*, 98, 1325-1348

Kahneman D.,Wakker P.P., Sarin R., 1997, Back to Bentham? Explorations of Experienced Utility, *Quarterly Journal of Economics*, 112, 375-405

Kaplan, Todd R. and Bradley J. Ruffle, 2004, "The Self-Serving Bias and Beliefs about Rationality" 42 *Economic Inquiry*, 237-46

Knetsch, Jack L., 1989, "The Endowment Effect and Evidence of Nonreversible Indifference Curves," *American Economic Review*, 79, 1277-1284

Köbberling, Veronika, and Peter P. Wakker, 2005, "An index of loss aversion," *Journal of Economic Theory*, 122, 119-131

Köszegi, Botond, and Matthew. Rabin, 2006, "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, CXXI, 1133-1165

Kovalchik, Stephanie, Colin F. Camerer, David M. Grether, Charles R. Plott, and John M. Allman, 2005, "Aging and decision making: a comparison between neurologically healthy elderly and young individuals," *Journal of Economic Behaviour & Organization*, 58, 79-94

Lichtenstein, Sarah, Paul Slovic, Baruch Frischhoff, Mark Layman, and Barbara Combs, 1978, "Judged Frequency of Lethal Events" 4 *Journal of Experimental Psychology: Human Learning and Memory* 551–78

List, John A., 2003, "Does market experience eliminate market anomalies?," *Quarterly Journal of Economics*, 118, 41-71.

Loewenstein, G, Adler D., 1995, A Bias in the Prediction of Tastes, *Economic Journal*, 105, 929-937

Loomes, G. and R. Sugden, 1982, "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty", *The Economic Journal*, Volume 92, 805-824

Loomes, G. and R. Sugden, 1986, "Disappointment and Dynamic Consistency in Choice under Uncertainty", *Review of Economic Studies*, Volume 53, 271-282

Loomes, G. and R. Sugden, 1987, "Some Implications of a more General Form of Regret Theory", *Journal of Economic Theory*, Volume 41, 270-287

Lund, Frederick H., 1925, "The Psychology of Belief", 20 *Journal of Abnormal and Social Psychology* 174-196

Matthews, Michael L., and Andrew R. Moran, 1986, "Age Differences in Male Drivers' Perception of Accident Risk: The Role of Perceived Driving Ability" 18 *Accident Analysis and Prevention* 299–313

McKenna, Frank P., Robert A. Stainer, and Clive Lewis, 1991, "Factors Underlying Illusory Self-Assessment of Driving Skill in Males and Females" 23 *Accident Analysis and Prevention* 45–52

Mehra, R. and E. Prescott, 1985, "The equity premium: a puzzle", *Journal of Monetary Economics* 15:145-161

Mitchell, Gregory, 2002, "Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioural Law and Economics' Equal Incompetence", 91 *Georgetown Law Journal*, 67-167

Mullainathan S., Thaler R., 2001, "Behavioural Economics", International Encyclopedia of Social Sciences, Pergamon Press, 1094-1100

Muren, Astri, 2004, "Unrealistic Optimism about Exogenous Events: An Experimental Test" *Research Papers in Economics* 2004:1, Stockholm University, Department of Economics

Novemsky, N., and D. Kahneman, 2005, "The boundaries of loss aversion," *Journal of Marketing Research*, 42, 119-128

Odean, Terrence, 1998. "Are Investors Reluctant to Realize Their Losses?", *Journal of Finance*, 53, 1775-1798

Parisi, Francesco and Smith, Vernon L. (eds.), 2005, The Law and Economics of Irrational Behaviour, Stanford University Press

Plott, Charles R., and Kathryn Zeiler, 2005, "The willingness to pay-willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations," *American Economic Review*, 95, 530-545

Posner, Eric A., 2003, "Probability Errors: Some Positive and Normative Implications for Tort and Contract Law", 11 *Supreme Court Economic Review*, 125–41

Prelec, D., 1998, "The probability weighting function," *Econometrica*, 66, 497-527

Quiggin, John, 1982, "A Theory of Anticipated Utility", 3 *Journal of Economic Behaviour and Organization,* 323–43

Quiggin, John, 1983, *Generalized Expected Utility Theory: The Rank Dependent Model.* Boston: Kluwer Academic Publishers

Rabin M., 1998, Psychology and Economics, *Journal of Economic Literature*, 36, 11-46

Rabin, M., 2000, Risk Aversion and Expected-Utility Theory: A Calibration Theorem, *Econometrica*, 68, 1281-1292

Rabin M., Thaler R., 2001, Anomalies: Risk Aversion, *Journal of Economic Perspectives*, 15, 1, 219-232

Roberts K. and M. Weitzman, 1981,"Funding criteria for research, development, and exploration projects", *Econometrica*, Volume 49, 1261-1288

Rothschild, M., 1974, "A Two-armed Bandit Theory of Market Pricing", Journal of Economic Theory, Volume 9, 185-202

Salthouse, Timothy A., 2004, "What and when of cognitive aging," *Current Directions in Psychological Science*, 13, 140-144

Sanna, Lawrence, Norbert Schwarz and Shavaun L. Stocker, 2002, "When Debiasing Backfires: Accessible Content and Accessibility Experiences in Debiasing Hindsight", 28, *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 497-502

Samuelson, William, and Richard Zeckhauser, 1998, "Status quo bias in decision making," *Journal of Risk and Uncertainty*, 1, 7-59

Sayman, Serdar, and Ayse Onculer, 2005, "Effects of study design characteristics on the WTAWTP disparity: A meta analytical framework," *Journal of Economic Psychology*, 26, 289-312

Savage, L.J., 1954, The Foundations of Statistics, New York, Wiley

Schmidt, Ulrich, Chris Starmer, and Robert Sugden, 2005, "Explaining preference reversal with third generation prospect theory," *CeDEx Discussion Paper 2005-19, University of Nottingham*

Shavell, Steven, 1987, *Economic Analysis of Accident Law*, Cambridge, Mass.: Harvard University Press

Sherman, Steven J., Robert B. Cialdini, Donna F. Schwartzman and Kim D. Reynolds, 2002, "Imagining Can Heighten or Lower the Perceived Likelihood of Contracting a Disease: The Mediating Effect of Ease of Imagery" in *Heuristics and Biases: The Psychology of Intuitive Judgment*, Thomas Gilovich, Dale Griffin, and Daniel Kahneman (eds.), New York: Cambridge University Press

Skinner B.F., 1953, Science and Human Behaviour, New York, MacMillan

Skinner B.F., 1979, The shaping of a behaviourist, New York, Knopf

Slovic, Paul, Baruch Frischhoff, and Sarah Lichtenstein, 1982, "Facts versus Fears: Understanding Perceived Risk" in *Judgement under Uncertainty: Heuristics and Biases*, edited by Daniel Kahnemann, Paul Slovic, and Amos Tversky. 462–92Cambridge: Cambridge University Press

Starmer, Chris, 2000, "Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk," *Journal of Economic Literature*, 38, 332-382

Stiglitz, Joseph E., 1986, *Economics of the Public Sector*, New York: W.W. Norton

Strotz, R., 1956, "Myopia and Inconsistency in Dynamic Utility Maximization", *Review of Economic Studies*, 23, 165-180

Sugden, R., 1991, "Rational Choice: A Survey of Contributions from Economics and Phylosophy", *Economic Journal*, 101, (497) 751-785

Sugden, Robert, 2003, "Reference-Dependent Subjective Expected Utility," *Journal of Economic Theory*, 111, 172-191

Sunstein, Cass R., 1997, "Behavioural Analysis of Law", 64 *University of Chicago Law Review,* 1175–95

Sunstein, Cass R.. 2000, *Behavioural Law and Economics*, Cambridge University Press

Svenson, Ola, 1981, "Are We All Less Risky And More Skillful Than Our Fellow Drivers?" 47 *Acta Psychologica* 143-148

Svenson, Ola, Baruch Frischhoff, and Donald MacGregor, 1985, "Perceived Driving Safety and Seatbelt Usage" 17 *Accident Analysis and Prevention,* 119–33

Teitelbaum , Joshua C., 2007, "A Unilateral Accident Model under Ambiguity", 36 *Journal of Legal Studies*, 431-477

Thaler, R. and E. Johnson, 1990, "Gambling with the house money and trying to break even: the effect of prior outcomes on risky choice" *Management Science* 36: 643-660

Tirole J., 2002, "Rational irrationality: Some economics of self-management", *European Economic Review*, 46, 633-655

Tom, Sabrina M., Craig R. Fox, Christopher Trepel, and Russell A. Poldrack, 2007, "The Neural Basis of Loss Aversion in Decision-Making Under Risk," *Science*, 315, 515-518

Tovar, Patricia, 2006, "The Effects of Loss Aversion on Trade Policy: Theory and Evidence", *mimeo, Brandeis University*

Tversky, Amos, and Daniel Kahneman, 1991, "Loss Aversion in Riskless Choice - a Reference-Dependent Model," *Quarterly Journal of Economics*, 106, 1039-1061

Tversky, Amos, and Daniel Kahneman, 1992, "Advances in Prospect-Theory – Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297-323

Viscusi, W. Kip, 1992, *Fatal Tradeoffs: Public and Private Responsibilities for Risk.* New York: Oxford University Press

Viscusi, W. Kip, 2002, *Smoke-Filled Rooms*, Chicago: University of Chicago Press

Viscusi, W. Kip, and Wesley A. Magat, 1987, *Learning about Risk: Consumer and Worker Responses to Hazard Warnings*, Cambridge, Mass.: Harvard University Press
Von Neumann, J. and Morgenstern, O., 1947, Theory of Games and Economic Behaviour (2nd edition), Princeton, Princeton University Press

Weitzman M., 1979, "Optimal search for the best alternative" *Econometrica*, Volume 47, 641-654

Wakker, Peter P., 2005, "Formalizing reference dependence and initial wealth in Rabin's calibration theorem," *Mimeo, University of Rotterdam*

Weber, Elke U., and Eric J. Johnson, 2006, "Constructing Preferences from Memories", in Sarah Lichtenstein, and Paul Slovic, eds., *The Construction of Preferences* (New York:Cambridge University Press, pp. 397-410

Weinstein, Neil D., 1980, "Unrealistic Optimism about Future Life Events" 39 *Journal of Personality and Social Psychology* 806-820.

Weinstein, Neil D. and William M. Klein, 2002, "Resistance of Personal Risk Perceptions to Debiasing Interventions", in *Heuristics and Biases: The Psychology of Intuitive Judgment*, Thomas Gilovich, Dale Griffin, and Daniel Kahneman (eds.), New York: Cambridge University Press

Wenglert, Leif and Anne-Sofie Rosén, 2000, "Measuring Optimism-Pessimism from Beliefs about Future Events" 28 *Personality and Individual Differences* 717-728