# EXPLORATION VERSUS EXPLOITATION TRADEOFF IN THE PARTIAL FEEDBACK PARADIGM IN DECISION MAKING FROM EXPERIENCE.

*An analysis on the underlying reasons and behavior in a tradeoff where subjects know nothing.*

When facing two options with unknown characteristics, we have to explore to learn the properties. This gives two options: settle down on one choice, or explore both to find the true probabilities. This is known as the exploration-exploitation trade-off. In this research it is found that small-sample problems result in the distortion of probabilities, and settling down behavior is affected by experienced variance and the cost of learning. The hot stove effect and the gamblers fallacy are present. No conclusions on choice behavior in different type of problems is found, leaving room for the analysis of risk preferences in the trade-off.

Keywords: Exploration-exploitation tradeoff, Partial, Feedback paradigm, Decision from experience, Decision making theory, Behavioral economics

Henri van Dijk

Student nummer 376643

Supervisor I. Aydogan

Master Behavioral Economics

Erasmus University Rotterdam

August 2014

*Special thanks go out to:*

I. Erev, E. Ert, and A.E. Roth, for performing the experiments and the availability of the data from this.

I. Aydogan, for the support, critical remarks, feedback, and suggestions on my work given.

Lisette Westerhof, for her support on every level and believing in me.

# Table of contents

## Introduction

There is currently an academic debate about the differences between two major types of decision making: decision from experience and decision from description. This deals with the difference in how people make their decision, either based on available descriptive information or based on their own personal experiences. These two can sometimes lead to the same decision. Imagine the difference between a soccer analyst and a soccer supporter. The analyst can argue that team X will win because of its objective statistics. The supporter on the other hand will argue that team X will win based on his own personal observations about a favorite team. Both can thus reach the same decision, but in a different way. There are also many cases where these two types of decision making result in two diverging choices. A famous example within this field is the doctor-patient conflict. Here, doctors have to make decisions on a daily bases, hence relying on their experience. Patients however rely on the descriptive information given and known to them. This can result in conflict: patients might argue that they want to try a cure, as they believe it might be effective in 60% of the cases, whereas a doctor might not want to try the cure, as he might have experienced a 0% success rate. (Hertwig, Barron, & Weber, 2004)

Erev and Barron (2003) researched the differences in behavior between these two types of decision making and describe the two different experimental designs to measure decision making. For the descriptive decision making, subjects receive the complete description about two prospects they are facing. In the experience decision making, subjects face two options, and have to discover the underlying characteristics based on the feedback of their choices. Subjects thus do not know the possible outcome beforehand in this, which is the main distinction between these two. In experiments by Erev and Barron (2005), it is found that both decision from experience and decision from description might not result in maximization behavior of the expected payoffs. The reasoning behind this is that rational subjects should be able to follow maximization behavior as the experienced probabilities should be equal to the true probabilities. This implies that there is a linear probability function. However, decision making from description is found to be in line with prospect theory, and therefore results in overweighting of small probabilities. For decision making from experience the opposite holds: subjects underweight the small probabilities. This means that both types of decision making can move away from each other in the weighting of probabilities.

In 2005, Erev and Barron published another research on decision making. Here the differences between experience and description based decision making are investigated in more detail, and more proof on the difference in maximization is given. The main finding is that feedback does not bring decision from experience closer to the maximization of payoffs.

Over time, the differences between these two types of decision making have become known as the Description-Experience gap, which is described in detail by Hertwig and Erev (2009). This description-experience gap is derived from the differences in behavior based on how choices are defined; either based on the description or on the experienced. This difference in behavior is caused by overweighting and underweighting of probabilities, mainly because the perceived probability is different in the experienced based choice. Prospect theory found to be driving the description based choice, as the outcome of a previous choice is driving the beliefs for the next period, which means that the perceived probability differs from the true probability. Hertwig and Erev (2009) analyzed this gap for subjects facing risky choices. They find that differences between the two types of decision making exist because in practice descriptions are not always available. Decision from experience can be seen as the opposite of decision making from description because of the difference in weighting in small probabilities. Fujikawa (2009) explains this gap trough the hot stove effect, which occurs in both types of decision making. The hot stove effect states that once subjects experience a bad outcome in an option, they become averse to it. This is found to persist in decision from experience, especially in repeated games.

Hertwig and Erev (2009) analyzed the gap and differences based on three paradigms in experience based decision making; sampling, full-, and partial feedback. Sampling means that subjects determine how often they want to try the prospects before making a final choice. The full and partial feedback paradigms have a fixed amount of samples, where feedback is given either on both options (full feedback) or only on the chosen prospect (partial feedback). In all paradigms the gap between description and experience is found; subjects are risk loving over high probability gains and risk averse over small probabilities. This results in the underweighting of rare events.

Of these three paradigms, the partial feedback paradigm will be the paradigm of interest in this paper. The main difference with respect to the full feedback paradigm here lies in the costs of exploring the options. In the full feedback paradigm there are no costs as the payoff of both options is made known after a choice. In the partial feedback, choosing one option goes at the cost of getting to know the other option. This is important because if subjects do not know both options of the risky choice, they base their decisions on a small sample. This is caused by insufficient exploration, hence moving the decision making away from the true probabilities.

This paradigm is also relevant for real life decision making. For example, imagine an investor who can invest in either stocks or bonds. If he invests in bonds, he gets a certain return. If he invests in a stock, he can either gain or lose relative to the original stock price. This is relevant for the paradigm because you do not know the underlying probabilities or payoffs. The difference is that you should know whether you are investing in a bond or a stock.

The aim of this research is to give more insight in the understanding of how people reach their decisions, specifically in decision situations similar to the partial feedback paradigm. To give more insight on this the tradeoff between the risky and safe options will be investigated. This tradeoff also deals with settling down on one choice. This means that you either exploit the option to yield a favorable outcome, or switch around between the choices to explore more. To see this, imagine the choice between a risky and a safe option. The safe option yields an amount of money with certainty, whereas the risky option yields either a high or a low amount of money. In this risky option, one has the high probability and the other has the low probability. The safe option yields a medium amount of money, lying between the high and the low amount of money.

Figure 1 shows what the options of behavior are in this problem. Initially there is a distribution of probabilities, giving either the high or the low outcome a small probability. This results in a choice made by the subject. If the subject selects the risky option, one of the outcomes occurs. This can be either the low or the high probability outcome. Thus if the subject selects the risky option where the good outcome has a small probability, it is very likely that the good outcome will not occur. This means that the subject will create a dislike for the risky option, and will probably stick to the safe option. The alternative is that the subject does experience the rare outcome. If he also experienced the safe outcome, a tradeoff is possible. If the subjects wants to explore the risky option more, the medium outcome of the safe option (which is higher than the low outcome), is forgone. That is, the subject faces an exploration-exploitation tradeoff. This is the tradeoff of interest in this research.

Let us continue down the figure to explain the rest. If the subject explored enough and the rare good outcome occurs, the subject gains compared to the safe option. This results in another tradeoff for the subject in the next choice he/she has to make. Now if the subject selects the risky option he knows that he can either gain with a small probability (because the good outcome is rarely observed) or lose with a high probability (because the bad outcome has a high probability). In this case the subject can thus try to maximize his payoffs. A similar reasoning holds in the bottom part of the figure where the bad outcome is rare.

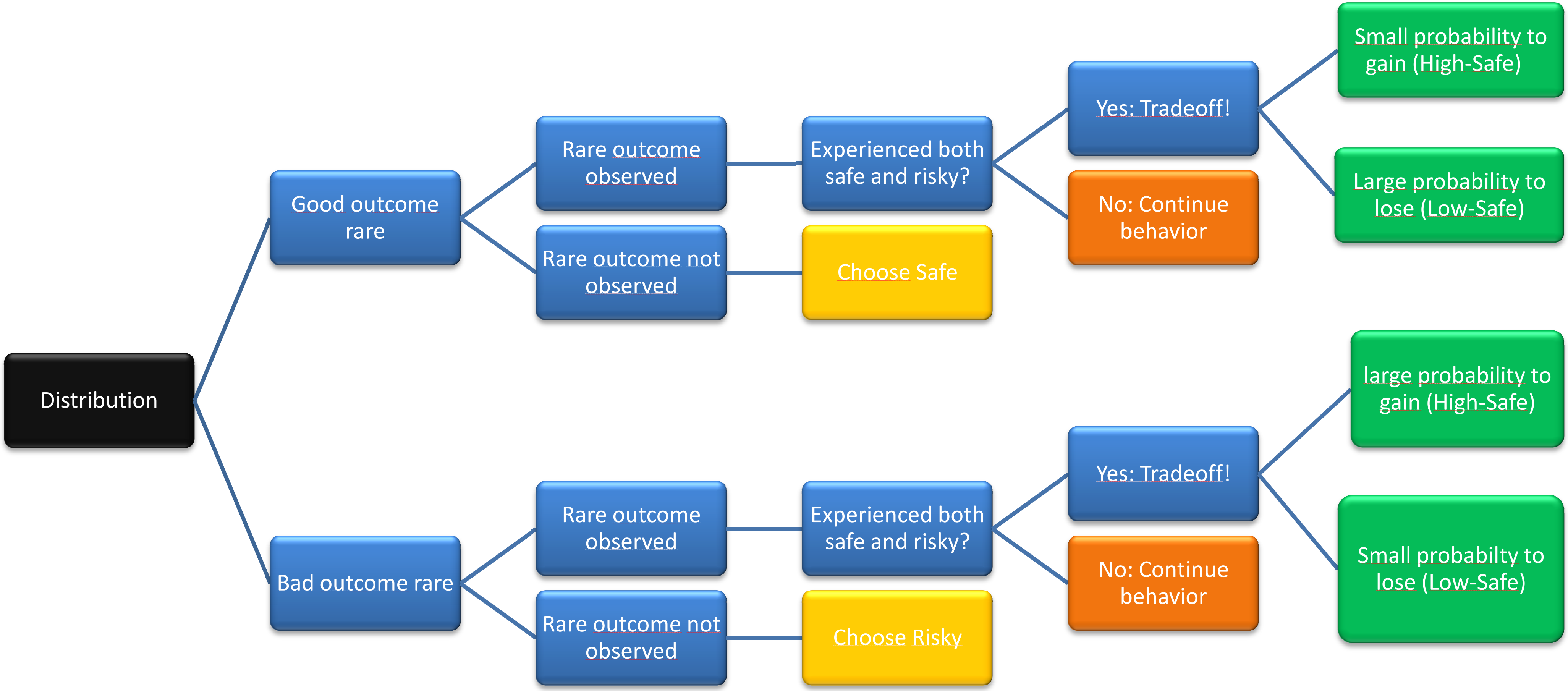



**Figure 1: Exploration - Exploitation tradeoff**

The previous example of stock and bond investing is very interesting in this tradeoff described above because of the so-called small-sample problems. If a stock went down in the previous two months, it might go up again in the next months. But if you switch to bonds after the initially two months, you did not experience the true probability of the increase/decrease of the stock. This means that the exploration-exploitation tradeoff could also be present in a tradeoff between stocks and bonds.

This research thus builds forth on the previous finding regarding the description-experience gap, but tries to differentiate itself by the explanation of behavior in the tradeoff described above and giving more insights about it. In other words, this research tries to explain the behavior in Figure 1. This means that the tradeoff will be investigated with respect to the choice made and the factors underlying this choice. Along with others, some of the investigated factors are the probabilities of outcomes, the amount of payoff, etc. Underlying behavioral biases will also be present. This will be tested trough a dataset available from Erev et al. (2010) which covers the partial feedback paradigm. It is found that the exploration-exploitation tradeoff is mainly affected by the small sample problem, the experienced variance, and the hot stove effect/gamblers fallacy. There is evidence of the cost of learning affecting the tradeoff, but mixed evidence on how the behavior differs, dependant on the type of problems faced.

In this research the following setup is used. First, theory and literature are discussed, from which hypotheses are derived. After this, the insights on the data are given, resulting in primary findings. These findings are then analyzed in more detail per hypothesis in the actual results analysis. These will be checked for robustness, leaving the final part dealing with the conclusions.

# Literature Review

In this section previous literature and theories of interest will be discussed. The first part will focus on decision from experience, the second part will elaborate on the partial feedback paradigm of decision from experience, and the last part will zoom in on the exploration-exploitation tradeoff, covering all aspects of the title. From these findings, conclusions can be drawn on how to analyze the tradeoff.

## Decision from Experience

In the real world the descriptions of prospects, such as possible outcomes and probabilities, are not always available. Therefore people need to rely on their own experience and perceptions. As mentioned in the introduction, Erev and Barron (2003) investigated this type of decision making and labeled them as decision making from experience. To test this experience-based decision making in the lab, subjects are typically faced with two options, represented by two unlabelled buttons on a computer screen. Subjects can discover the outcome distribution of the options by making repeated trials from each of them. This simply means that they explore the options given in the problem by drawing samples from them or making repeated choice. This is known as the so-called clicking paradigm (Hertwig and Erev, 2009). This exploration aspect of the decision from experience differentiates it from decision from description case.

Within this clicking-paradigm, there are three major experimental paradigms given in decision from experience. The first one is the sampling paradigm. Here, subjects can sample as many times as they want, and have control over when they want to make the definitive choice.  It is important to note that sampling does not affect the final payoff (Hertwig, Barron, & Weber, 2004). Next to this the full paradigm feedback exists. Here, a subject makes a fixed amount of trials in a problem involving repeated choice between two options, and after each trial both the achieved and the forgone (the option that was not selected) payoffs are shown to the subject. The main difference with the sampling paradigm is that one of the trials is randomly selected to be made true; therefore all choices might affect the payoff. (Yechiam & Busemeyer, 2006) There is no exploration-exploitation tradeoff involved here because subjects do not actually have to explore the options in their choice. This is caused by the fact that the forgone payment from the unselected option is also given in the feedback.

The third and final paradigm is known as the partial feedback paradigm. This is similar to the full feedback paradigm, but differs in the feedback given. In the partial feedback paradigm only the achieved payoff of the selected option is made known in the partial feedback paradigm. (Barron & Erev, 2003) Therefore the partial feedback paradigm introduces the exploration-exploitation tradeoff, as subjects have to explore the options by themselves in this paradigm. The exploring of the options usually goes along with a cost in choosing, having to give up one option at the expense of the other.

Hertwig and Erev (2009) did a comprehensive review of the literature discussing the reasons for the description-experience gap. They analyzed the differences in behavior between the two types of decision making based on the three paradigms described above, and point out that experience-based decision making does not result in the (expected) maximization of payoffs. The deviation from rational choice seems to be in the opposite direction compared to the description-based decision making. This is caused by the fact that in all of the three experimental paradigms, subjects tend to underweight small probabilities in experienced based choice, opposite to the overweighting found in description based decision making. The behavior in description based decision making are found to be contradicting with the main findings from prospect theory. (Kahneman & Tversky, 1979) The causes for the description-experience gap are described below.

The first one is the *small-sample problems*. This states that subjects base their choice on a small sample because they do not explore the options enough, which can possibly explain the underweighting of small probabilities. Another cause is the *recency* effects, which state that subjects put the most weight in their decision making process on the most recent observations. This implies that subject unevenly weigh the received feedback. Next to this are *estimation errors*, which state that subjects might systematically underestimate the probability of rare outcomes. Next to this *contingent sampling* exists, which states that subjects base their decisions on similar experiences they might have had. This is especially relevant in the feedback paradigms. The final cause deals with the *information format and the cognitive algorithms*. This states that subjects create their own algorithm based on a series of experiences, which mainly explains the difference in the gap between sequential and single events. (Hertwig & Erev, 2009)[1]

[1] Relevant theory related to the exploration-exploitation gap are discussed later on in more detail.

Several researchers found aspects which could explain the description-experience gap. Hertwig, Barron, and Weber (2004) found that underweighting is mainly a result of three factors:  the event did not occur before, the event did not occur recently, or the event occurred less than expected. This results in recency and small sample problems. In addition it has also been found that choices involving rare events do not explain all fundamental differences between description and experience, which might be attributable to estimation errors, or the information format/cognitive algorithms.

An interesting recent finding is that the description-experience gap is not only present over small probability events, but also for medium probability events (Ludvig & Spetch, 2011). This implies that utility might be achieved different in experience-based decision making relative to the description-based decision making, as the risk preferences of the decision makers do not hold. Specific findings and theories on the partial feedback paradigm are discussed below in the next subsection.

## Partial Feedback Paradigm

Barron and Erev (2003) were the first to use the partial feedback paradigm in their research and experiments. In their research they analyzed small-feedback based decision making. This means that subjects were facing a series of choices, dealing with small probabilities. The small probabilities are used to reduce the importance of choice. From these series of choices made by the subjects, one choice is randomly selected to be paid out, giving equal weight to all the choices made. Feedback on the achieved payoff from the option selected in the previous trial is given to the decision maker before each choice. This is expected to bring subjects closer to the payoff maximizations, as they should learn the distributions trough the feedback. In these experiments subjects were asked to choose between two unlabelled buttons, repeated as much as 200 to 400 times per problem. The buttons represented different payoff distributions, which were unknown to the subject beforehand. The most common underlying distribution has one “safe” option which always yields the same, and a “risky” option which yields either a payoff higher or lower than the safe option. The payoff is determined by an underlying probability.

Barron and Erev (2003) used five different experiments in their research on the partial feedback paradigm, where each experiment was designed to test for one specific effect, either: loss aversion, the reversed certainty/common ratio effect, reversed payoff domain effects, underweighting of small probabilities, and description based choice. Over these five experiments it was found that feedback in decision making does not bring the subject closer to payoff maximization behavior. Interestingly enough, feedback actually drove the subject away from maximization of payoffs. This causes underweighting of small probabilities as well as a reversed common-ratio pattern/reversed payoff domain effects.

The model Barron and Erev (2003) used to find these effects tries to estimate the “adjusted” value of a gamble, based on the subjective weighting and valuation. In other words, the value of the next gamble is determined by the weight put on the subjective value from the obtained payoff in the previous trial (the feedback), and the weight put on the subjective expected value from the previous trial. Barron and Erev (2003) argued that these effects are a result of two psychological biases: loss aversion and recency effects. Loss aversion is found in this estimation by multiplying the subjective value of losses with the measure for loss aversion from Kahneman and Tversky (1979), the lambda. For the recency effects Barron and Erev (2003) assumed that the weight on feedback is determined differently for exploration trials. Subjects put more weight on feedback in exploration trials compared to other trials.  With this differentiation in weights it was possible to differentiate between exploration and exploitation, hence finding the tradeoff of interest in this paper.

More proof for this finding is given by Erev and Barron (2005) when they focused their research on the effect of feedback. They attribute the initial failure of the maximization of payoff to three factors: the variance in the payoffs, the underweighting of rare events, and loss aversion. It is concluded that the three factors hold when feedback is introduced in decision making from experience. From this research Erev and Barron (2005) concluded that behavior in experience-based decision making seems to be close to probability matching. Probability matching states that subjects use the experienced probability in the previous trials as the expected probability for the next trial (Shanks, Tunney, & McCarthy, 2002). Next to the probability matching, Erev and Barron (2005) found that if a subject receives feedback on the earnings in the previous sample, these earnings do not seem to affect his decision making. This leads Erev and Barron (2005) to conclude that a reinforcement learning model should be the best predictor for behavior within decision making from experience. The reasoning behind this is that subjects in repeated decision making start with a learning period, followed by the settling down on one choice in which they seem to neglect feedback. The reasons for this is that subjects have already made their final choice based on the learning period, and are not affected by shocks any more. This model is known as the reinforcement learning amongst cognitive strategies model (RELACS), which assumes that learning occurs during the cognitive strategies played by subjects, and subjects play one of the three possible cognitive strategies described below. The first strategy implies that subjects decide based on the highest recent outcome. The second one implies that subjects want to minimize losses, and therefore update beliefs in every trial. The final strategy is found to be used the most, and is known as the diminishing random choice/slow best reply. This strategy states that subjects choose random in the first trials, after which they slowly learn which strategy is the best to maximize their payoff. This strategy is the closest to reinforcement learning and Erev and Barron (2005) conclude this to explain their findings.

Other relevant effects in the partial feedback paradigm will be discussed below in detail, to show their importance for this research.

In general it is found that there is a high sensitivity to feedback in repeated games and problems as found by multiple researchers. (Erev & Roth, 1998) (Friedman, 1998) (Camerer & Ho, 1999) However, as shown by Barron and Erev (2005) this does not seem to seem to explain the differences between decision from experience and decision from description.  Jessup et al. (2008) argue that feedback can make experience-based decision making more pure as feedback can explain the underweighting of small probabilities. This was concluded from an experiment over 29 students in 120 trials per problem, facing a risky and a safe option over gains. The students were randomly assigned to problems either with or without feedback. The problems included both high and low probability problems.  It was concluded that feedback made the subjects more likely to pick the safe option in low probability gains, compared to subjects without feedback. The reverse holds for high probabilities. Subjects who received feedback were more likely to pick the risky option in the high probability problem compared to subjects without feedback. This results in the conclusion that feedback affects decision making. (Jessup, Bishara, & Busemeyer, 2008)

An interesting finding here is based on research by Shafran (2011). In his experiments, subjects were facing a loss with a probability, where the probability could be reduced by giving up a sum of money from the final payoff. This was repeated over 100 trials with and without feedback. It was found that subjects tend to switch their preferences when they made a loss in a previous choice or trial. Shafran (2011) also found proof of exploration; switching is more likely to occur when the probability of losing was high (larger than 90%). Next to this it is concluded that switching is less likely to occur in rare events (small probabilities).  Shafran concluded that there is an updating of beliefs in decision making. To which degree this updating exist is hard to estimate. Regarding the updating it is found that subjects tend to attach the most weight to the more recent outcomes. This is proof of the recency effect being present when feedback is included. (Shafran, 2011) Other research states that giving feedback after every trial or round makes a subject less likely to picking the safe option, as subjects want to correct against the previous loss. This finding implies that feedback is dependent on the updating frequency. (Yechiam & Busemeyer, 2006)

As mentioned in the research by Shafran (2011), the recency effect seems to be of importance, as subjects quickly switch after a negative experience and update their beliefs based on the most recent outcomes. The recency effect states that more weight is put on recent outcomes. The underlying implication here is that the subjects makes their decision on a smaller sample than they have access to (Hertwig, Barron, & Weber, 2004).

Opposite and related to this are primacy and order effects. Primacy effects state that the first trials of a problem have the most impact on a subjects' decision making, the first experiences are the most important. Next to this, the order effects state that the order of the problems might have an effect. The order effect refers to subjects facing multiple repeated-choice problems over time. In this, subjects might take more risk in the first problem compared to the final problem (Hogarth & Einhorn, 1992).

The final and main bias of interest in the partial-feedback paradigm deals with small sample problems. This is discussed in the exploration-exploitation tradeoff section of this chapter, as this is mainly present here.

## Background for the exploration/exploitation tradeoff

In this section the underlying theories and behavioral biases for the exploration and exploitation tradeoff will be discussed, as well as relevant findings specific for this tradeoff.

### Small sample problems

As mentioned before, one problem in the tradeoff is caused by the small sample selection problem. The underlying idea is that subjects might at first explore the options, after which they choose one path on which they continue. The small sample problem state that subjects do not explore these options enough to see all the possible outcomes occur, or do not experience the true probabilities, mainly resulting in underweighting of the rare outcomes. This results in the tradeoff between exploration and exploitation, as described before. Erev and Barron (2005) analyzed the tradeoff by looking at 40 experiments consisting of at least 200 trials. In all the experiments and trials, subjects were asked to pick between a risky and a safe option, yielding an immediate payoff. It was found that rare outcomes were underweighted in most of the experiments. This is caused by subjects who tend to rely on the safe option. This also results in undervaluation of the risky outcome because it is not selected enough. Too little risky outcomes are selected to give the subject an idea of the true probability. The sample in the risky prospect over which subjects make their decisions becomes relatively small, causing the risky prospect to be underweighted even more (Erev & Barron, 2005). The above is therefore known as the small-sample selection problem.

## Factors affecting the small sample problem

Several factors affect the small sample selection described above. Hertwig and Pleskac (2010) suggest that subjects make their decisions on the experienced difference between risky and safe options. The smaller the sample size of a subject, the larger the difference between descriptions and the experience might be. This is caused by the distorted representation of the true probabilities in these small samples.[2] Furthermore, gathering experience is costly because the choice a subject made might affect the final payment. Another factor of importance in this small sample problem is the recency effect, the idea that subjects put more weight on recent observations. This is of importance because subjects combine their current findings with previous findings for the future choices. This recency effect might have a stronger effect when samples are small because of the relative weight. There is also an advantage of small samples, as it might help to simplify a choice. (Hogarth & Einhorn, 1992)

Another subject of interest deals with learning within these trials and the associated costs. Theoretical research by Aghion, Bolton, Harris, and Jullien (1991) analyzed learning models and tries to find the optimal path in this. The conclusions were that the local properties of the choice are the main determinant of whether the subject would reach the true maximum achievable payoff. Furthermore there seems to be a negative relation between the costs of learning and the distance to the maximum payoff. If learning is cheaper, it is easier to reach the maximum payoff. Empirical research by Gureckis and Love (2009) tested learning in decision making trough a “farming on mars” experiment. Here, subjects have to select one out of two robots to do the mining on mars, trying to maximize oxygen, where more oxygen would result in a higher payoff for the subject. This is repeated in five sets of 100 trials.  Feedback is given on the amount of oxygen farmed. The underlying difference between the robots is that one robot is more suitable for the short run, which has a higher current yield but decreases future yields. The opposite holds for the other robot, being better suited for the long run. This was not told to the subjects. To maximize the payoff of the subjects, subjects would thus have to show learning behavior. It is found that the best option was selected more often when there was little noise. If there is a lot of noise, the worst option is selected more often. This let the researchers to believe that there is no learning when there is a lot of noise in the experiment. However, a small amount of noise makes the subject more eager to learn in decision making. (Gureckis & Love, 2009)

[2] Imagine a risky prospect with (0.2; 10; 0.8;0) and safe prospect with (1;2). If the risky prospect is chosen 5 times, and yields 0 in all 5 choices, the experienced mean will be smaller than the descriptive mean over 5 choices. (0 vs. 10). The safe option will have always yielded 10 after 5 choices. As can be seen, the difference between the means is very large. (0 against 10 for experience; 10 against 10 for descriptive). Increasing your sample size should make the experienced mean closer to the descriptive mean.

Related to the tradeoff, this means that the exploration – exploitation tradeoff is affected by the noise and learning mechanisms can be used to reduce the over and under weighting of choices. In the tradeoff of interest there is some noise: subjects do not know the characteristics of the problem beforehand.

Related to learning is the switching behavior and patterns of subjects, which has been researched by Hills and Hertwig (2010). Here they pooled the data of four experiments and found that most of the switching occurred in the first 20% of the trials. Switching also slightly increases halfway during the problems. Frequent switchers pick more "winning" options, but are also more likely to underweight the rare probability. Against this, subjects who switch less are more likely to win in the long run. Therefore the authors conclude that searching behavior seems to determine the choices made. Next to this, Lejarrage, Hertwig, and Gonzalez (2012) also researched pooled data of experiments and found that subjects search longer when faced with a loss relative to searching shorter when they face a gain. Important to note is that they found that the variance affects the amount of searching here; more variance should result in more searching. (Lejarraga, Hertwig, & Gonzalez, 2012)[3] Since the variance is dependent on the characteristics of the problems faced, we know that there should be problem-specific effects in the tradeoff.

One important aspect in the tradeoff is the learning effect. There are two interesting aspects related to this. The first one is the gamblers fallacy/hot hand effect. This is tested by Ayton and Fischer (2004). In an experiment they divided subjects between gamblers and forecasters. Both were asked to predict the correct color (red or blue). Gamblers are asked to gamble knowing that the outcome was generated based on a computerized roulette outcome. Forecasters were asked to forecast the outcome of this computerized algorithm. Rewards were given for correct predictions in both cases. It was found that there are negative recency effects present in the outcomes. This is known as the gamblers fallacy; after a series of winnings, the subject was less likely to predict correct. The second finding was that there are positive recency effects present in the expectations. This is known as the hot hand effect; subjects were more confident after a win. Between the two types, it is found that gamblers achieve a longer streak of correct *and* incorrect predictions, compared to the forecasters. From this experiment it is concluded that subjects might anchor their choices based on a series of good or bad outcomes. Winning in one option over time makes you more likely to select it again, vise versa for losses. (Ayton & Fischer, 2004)

[3] Note: these papers tested this in the sampling paradigm

The other effect is the “hot stove” effect. This means that when the subject experienced a bad outcome from a certain type of choice, it becomes averse to that option and might not select it again. Fukijawa (2009) investigated this trough an experiment. Subjects were invited to choose between a risky and a safe option in two sequential problems. The first problem had a high probability with a low payoff, and the second problem low probabilities with high payoff. This means that the bad outcome is very rare in the first problem, and less rare in the second problem. In the results it was found that there is less maximization in the second problem. From this, Fujikawa (2009) concluded that the hot stove effect is present in repeated games, as more safe options are selected in the first problem after the bad rare outcome occurred. The small sample problem is also found by Fujikawa (2009), as rare probabilities were underweighted. This was concluded based on the average payoffs. The problems were tested over 400 trials. The hot stove effect is concluded to be important when there is underweighting of probabilities and the payoff varies over the choices made. (Fujikawa, 2009).

A large part of the literature focuses on the “armed-bandit” problems in estimating the exploration-exploitation tradeoff. In this, a gambler is facing a row of slot machines (the one-armed bandit), and has to decide the order, length, and the amount of time he wants to spend on these machines. Each slot machine has its own different distribution and payoffs (Berry & Fristedt, 1985) To maximize the payoff in these problems, the gambler should make use of the Gittins index. This index is a measure for the payoff which takes the probability that the payoff might be cancelled in the future into account, relatively to the present state of the game. (Gittins & Jones, 1979) These armed-bandit problems are a good representation of the exploration exploitation tradeoff and seem very suitable to estimate this tradeoff, as the gamblers can explore multiple machines, or settle down on one.

Empirical research shows that that variance is a good measure to counter this tradeoff in the one armed bandit problems; higher variance of the payoffs (the arms of the bandits) makes a subject more likely to stay with one choice, risky or safe. This means that variance can help to reduce the regret of subjects picking one choice. The settling down to one choice is determined based on the learning. Subjects settle down after they have reduced the noise in the payoffs enough. Furthermore maximizing the payoff will also reduce the regret. This was tested trough algorithms trying to estimate computer simulations of these problems. (Audibert, Munos, & Szepesvari, 2009)

Other research focused on what would happen when the set of arms in bandit problems could vary, when more options are possible per machine. The amount of arms does not seem to affect regret. It is also found that there might be large differences in the exploration-exploitation tradeoff, dependant on the phase where regret occurs. (Bubeck, Munos, & Stoltz, 2011) There are not that many experiments which estimate these armed bandit problems.

One final interesting case of application of this tradeoff is in organizational learning. March (1991) applied this by looked at the tradeoff between new possibilities (explore) and old certainties (exploit). Both learning within a company and competitive learning are assumed. It is stated that exploring can be useful in the short run for a company, but will have negative effects in the long run. This is because the effectiveness of organization learning is marginally decreasing. The preference between exploration and exploitation also depends on what is needed in the organization; exploration is more flexible, whereas exploitation is better suited for performing at the top. (March, 1991)

The findings reported above are used as a background, to form the hypotheses to be used. These are described in the next section.

## Hypotheses

Based upon the previous research and literature, hypotheses for this research are derived. These will be depicted below. The first and main hypothesis regarding choice behavior is that the description-experience gap in the partial feedback paradigm is caused by the exploration-exploitation tradeoff. This translates itself to the following hypothesis:

*Hypothesis 1: The lack of exploration is the main reason of the description-experience gap in the feedback paradigm.*

We already know that this should be true, as it is in line with the findings from the competition paper from Erev et al (2010). In this competition they conclude that the best prediction of experience based decision making are explained by the small-sample problems. This means that exploration and exploitation should be causing the description-experience gap. Since the data used for this analysis is based on the experiments from this competition paper, similar conclusions should be found.

Next to this hypothesis, the underlying reasons for exploration and exploitation should be investigated. To see this, we will look for reasons which cause subjects to either explore more or less. From the findings from Lejarraga, Hertwig, and Gonzalez (2012), it is known that subjects explore more when the variance in a problem and the samples seems to be high. This translates itself to the following hypothesis to test related to this:

*Hypothesis 2: Subjects explore more when they experience more variance*.

To test this hypothesis it is important to use the experienced variance. This means that it has to be based on the previous samples within a problem from a subject, and needs to be compared with the pattern of selection from the subjects. If the selection of the risky choice is significantly higher or lower related to the experienced variance, we can conclude that hypothesis two is true.

Opposite to the factors causing more exploration, there are factors which cause subjects to explore less. The main factor reducing exploration seems to be the cost of learning. When subjects face high costs of learning ( the gap between the risky and the safe option is found to be large), the subjects will explore less. These costs of learning are found to be of importance in several of the discussed theory and literature. (Lejarraga, Hertwig, and Gonzalez) (Aghion, Bolton, Harris, & Jullien, 1991)(Hogarth & Einhorn, 1992) This translates to the following hypothesis to be analyzed:

*Hypothesis 3: Subjects explore less when they experience higher costs of learning*

This means that a negative relation between exploration and the cost of learning is expected. In testing this, the difference between the risky and the safe payoffs must be analyzed. It is also important to correct for the experienced variance, as this might introduce noise into testing this hypothesis.

Next to the underlying behavioral reasons for the exploration-exploitation gap, it is found that several behavioral biases seem to have an effect. This is affecting the settling down behavior, as seen by the hot stove effect and the gamblers fallacy, which is found in the findings from Ayton and Fischer (2004) and Fujikawa (2009). The hot stove will drive subjects away from the risky choice after experiencing the bad outcome once, whereas the gamblers fallacy will draw subjects towards the risky choice after a winning. This translates itself to the following hypothesis to be tested:

*Hypothesis 4: The hot stove effect and the gamblers fallacy affect the exploration behavior, as reflected in the settling down behavior of the subjects.*

It important to take the timing of the rare outcome into account, as settling down is more likely in the beginning of a problem compared to the final part because more learning has been possible. These biases are present if the rare outcome of the risky choice occurred (the option in the risky choice with the lowest probability) and the subjects who experienced this are less likely to select the risky choice compared to subjects who did not experience the rare outcome.

It is also possible that external factors might affect the exploration-exploitation tradeoff. It is hypothesized that the domain of payoffs is found to be of importance. As learned from Lejarraga, Hertwig, and Gonzalez (2012), problem specific difference can arise because subjects search longer when facing losses compared to facing gains, hence they settle down later. This results in the following hypothesis to be analyzed:

*Hypothesis 5: the choice behavior of subjects differs when facing losses compared to gains.*

In testing for the differences in searching behavior here, it is known that there are three types of problems. Mixed (gains and losses in the risky and safe option), losses (risky and safe option always yields a loss), and gains (risky and safe option always yield a gain). The probabilities also need to be taken into account here, as the weighting might differ between losses and gains.

In the data description and the data analysis these hypotheses will be answered. To be able to do this it is firstly needed to describe the experiment and the data, which is done in the next chapter.

## The data

To analyze the research question, an experiment testing the partial-feedback paradigm is needed. Erev, Ert, and Roth (2010) did a series of experiments on multiple paradigms, including the partial feedback paradigm. Their goal was to create a competition for researchers to come up with the best model to predict behavior in these paradigms. This data is publicly accessible for free on their website.[4] Two datasets are available on the site, estimation and a competition dataset. In each of these sets, 160 students were invited for this, which were randomly allocated to one of the paradigms tested. Our initial analysis is based on the estimation set. To test the robustness of the findings, the analysis is checked in the competition dataset to see whether the findings are consistent. Below the partial-feedback experiment is described in detail.

### Experimental design

The partial feedback experiment was set up as following: 100 students participated and were paid a show-up fee of 40 Shekels (€8.35), which increased or decreased based on the results from their choices. These students are split in five sessions with randomly assigned, independent problems per session. These problems are equally divided in the different payoff domains. The subjects played twelve problems with 100 trials per problem, making 1200 observations per subject in total. Per problem, one option was randomly selected to be played and paid. In the trials subjects were asked to choose between two unlabelled buttons, where one is the safe option with a certain payoff, yielding a medium amount of money. The other is the risky option, with either a low payoff or a high payoff with a high or low probability. The distribution of these factors differs per problem. After the subject made their choice in a trial, the payoff of that trial is made known to the subject, and the next trial starts. Important to note is that subjects do not know anything except the payoffs they get as feedback, as the buttons are unlabeled.

Combining the five sessions in this paradigm results in a dataset of 120000 observations. (1200 per subject, 20 subjects per session, five sessions) The output of this experiment consists of the subject's *ID*, which is a number from 1-100, corresponding to the subjects. The variable *problem* is the number of the problem on a scale from 1-60. The variable *trial* corresponds to the number of the trial in the problem. These range from 1-100. The variable *order* measures the order in which problems are shown to the subjects. This is on a range from 1-12. Next is the variable *p(high).* This is the probability of getting the high payoff in the risky option, ranging from 1-99%. Furthermore the variables *high, medium,* and *low* measure the payoff for the corresponding options. Note that *medium* measures the payoff for the safe option. These payoffs range from -29.2 until 26.5 shekel.

[4] http://tx.technion.ac.il/~erev/Comp/Comp.html

The most important variable is the variable *choice*, which measures the choice of a subject in a trial. This is either 0 (the subject picked the safe option) or 1 (the subject picked the risky option). Finally, the variable *payoff* measures the realized payoff for that specific trial. This covers the same range as the variables *high medium* and *low.* Since the payoff in the previous period might impact the choice, the lagged payoff is generated, labeled *l1payoff.* There is no lag included for the first trial of the problems, as this would capture the choice in the 100$^{th}$ trial of the previous problem.

Next to this more variables are added. The first one is *session*, which indicates what session the subject participated in. This is on a scale from 1-5. The second is the *p(low)*. This is the probability for the low outcome to occur and is calculated by taking 1-*p(high).* This is thus also on a scale from 1-99%. Next to this, the *expected payoff* is generated. This is the expected payoff of the problem the subject is facing, and calculated with the formula: *expected payoff = p(low) * low + p(high)*high*

For the tradeoff between exploration and exploitation, it is important to know the tradeoff from the exploration and exploitation. The first thing needed here is which payoff is a rare payoff, low or high. Two dummy variables are generated, labeled *low_rare* and *high_rare*. If these variables have a value of 1, the name holds. This is the case when the probability of low (high) is smaller than the probability of high(low). The next variable needed are the gains and losses from the exploration vs. exploitation tradeoff scheme. These variables are generated by taking the difference between the high or the low outcome and the safe outcome, and are labeled respectively *gain* and *lose.* Furthermore, *trial* is split in groups of ten and twenty trials. These are labeled *trialgroup10* and *trialgroup20*, and are categorical on a scale from 1-10 and 1-5. The payoff domains of the problems are captured by three dummy variables: *gainproblem, lossproblem, mixedproblem*, which are 1 if the problem deals with the corresponding domain of the payoffs.

From this, the final data transformation can be made. Variables are generated which define whether the rare outcome is observed in two steps. First, the achieved payoff is compared with the high outcome and given a value of 1 when the high outcome was rare, using the variable *high_rare.* This is repeated for the low outcome, and thus given a value of 1 when the low outcome was rare, and the payoff matches the low outcome. This results in two new variables: *rare_low_outcome_observed* and *rare_high_outcome_observed,* which are 1 if the name holds: experienced outcome = low & rare (or high & rare) and 0 if otherwise. These two are also combined in the variable *rare_outcome_observed*. For all three groups, the lagged variables are created as well, labeled *l.rareoutcome, l.highrareoutcome,* and *l.lowrareoutcome.* This is repeated for a total of twenty lags. Again, the first trials are not included to avoid the use of the 100$^{th}$ trial of the previous problem. With all these variables, the data can be described in the next subsection.

## Data description

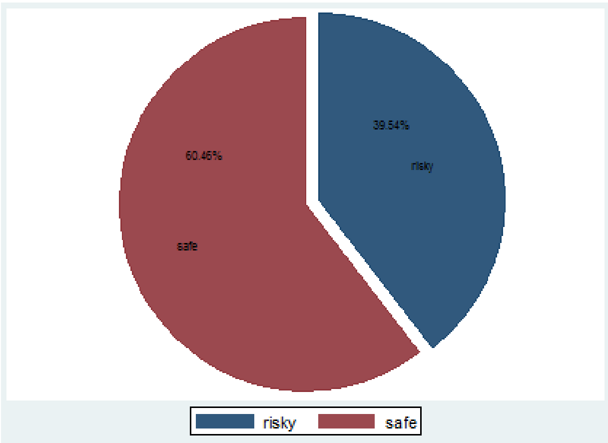
This subsection will describe the main data from the experiment. The specific descriptions of the data related to the main hypotheses are described in the specific chapters for the analysis of these hypotheses. All of the 60 used problems in the experiments and their payoffs and probabilities are shown in Table 37 in the appendix. From this it is known that the problems are equally divided between the payments. 20 problems are on gains, 20 on losses, and 20 on mixed payoffs. The expected payoffs from the risky choice per problem is also shown here. As can be seen, the risky choices are close but not equal to the certain outcome.

Summary statistics are given in Table 1 for all the relevant variables, both from the experiment and the created variable. From this table it can be seen that in about 60% of the cases, the subjects picked the safe option, opposite to the risky case in about 40% of all cases. In about 56% of all the faced problems, the low outcome was the rare outcome, opposite to 42% where the high outcome is rare. Of *all* the 120000 made choices (risky *and* safe), the rare outcome occurred in only 4% of this. To see whether this means whether there is too little exploration requires testing, which will be done in the specific subsections and chapters for the four main hypotheses.

| Variable | Observations | Mean | Standard Error | Min | Max |
|---|---|---|---|---|---|
| *choice* | 120000 | 0.3954 | 0.4889 | 0 | 1 |
| *risky* | 120000 | 0.3954 | 0.4889 | 0 | 1 |
| *safe* | 120000 | 0.6046 | 0.4889 | 0 | 1 |
| *high (risky)* | 120000 | 5.2317 | 9.3011 | -10 | 26.5 |
| *low (risky)* | 120000 | -5.0767 | 9.4692 | -29.2 | 9.7 |
| *medium (safe)* | 120000 | 0.2933 | 10.7588 | -25.6 | 25.2 |
| *rare outcome occurred* | 120000 | 0.0429 | 0.2027 | 0 | 1 |
| *rare high outcome occurred* | 120000 | 0.0139 | 0.1169 | 0 | 1 |
| *rare low outcome occurred* | 120000 | 0.0291 | 0.1680 | 0 | 1 |
| *phigh* | 120000 | 0.5522 | 0.3953 | 0.01 | 0.99 |
| *plow* | 120000 | 0.4478 | 0.3953 | 0.01 | 0.99 |
| *achieved payoff* | 120000 | 0.3775 | 10.8851 | -29.2 | 26.5 |
| *expected payoff* | 120000 | 0.2902 | 10.6472 | -25.36 | 25.408 |
| *low_rare* | 120000 | 0.5667 | 0.4955 | 0 | 1 |
| *high_rare* | 120000 | 0.4167 | 0.4930 | 0 | 1 |
| *gain* | 120000 | 4.9383 | 5.3879 | 0 | 16.7 |
| *lose* | 120000 | 5.3700 | 4.5859 | -0.4 | 16.9 |

**Table 1: summary statistics**

Now let us look at the distribution between the safe and risky choices in more detail. The pie diagram in Figure 2 shows the division between the safe and risky choices over all the trials and problems. This shows the same distribution as in the summary statistics, 40% vs. 60%. Figure 3 includes the outcomes of the risky choice. Here it can be seen that in 13% of all the choices the subjects achieved a loss relative to the safe option. In about 27% of all the choices, the subjects achieved a gain relative to the safe option. When we compare this to the 40% of the risky choice it is found that when the subjects choose the risky option they lose in 32.5% and win in 67.5% of the risky choices.[5] Note that the percentages between Figure 2 and Figure 3 are not exactly the same. This is caused by problem 43 were the safe option is dominated by the risky option; the risky choice always yields more than safe option, as can be seen in Table 37.



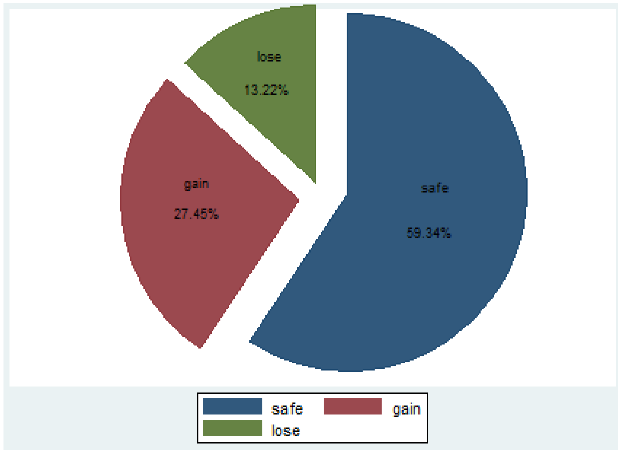




Figure 2: Risky vs. Safe choice.

Figure 3: Gain or Lose vs. Safe

## Preliminary analysis

Next to the five main hypotheses, other effects might be present in the data. There might be differences based on the session in which the subjects participated. Subjects might also change their behavior based on the order of the problems they are facing, or the characteristics of the payoffs. This is analyzed below.

[5] Total percentage risky choice = 13.22+ 27.45 = 40.67. For loss: 13.22/40.67 * 100= 32.5%. For gain: 27.45/40.67 * 100 = 67.5%

## Testing on the sessions

The first analysis will check whether there are differences in the choice behavior of subjects over the five different sessions used in the analysis. To see whether there are differences present the fitted curve for the average risky choice over the trials per session are shown in Figure 4. From this it hints that there are some differences in the risky choice amongst the session. However, this might also be due to the differences in the problem selection. To see whether this is the case, the average choices over time for all the problems are graphed in the appendix in Figure 23 and Figure 24. From this it seems that the volatility of choice seems quite different per problem, which is found to be of importance according to Lejarraga, Hertwig, & Gonzalez (2012). To be able to conclude whether there are significant differences between the sessions, a statistical test is needed.

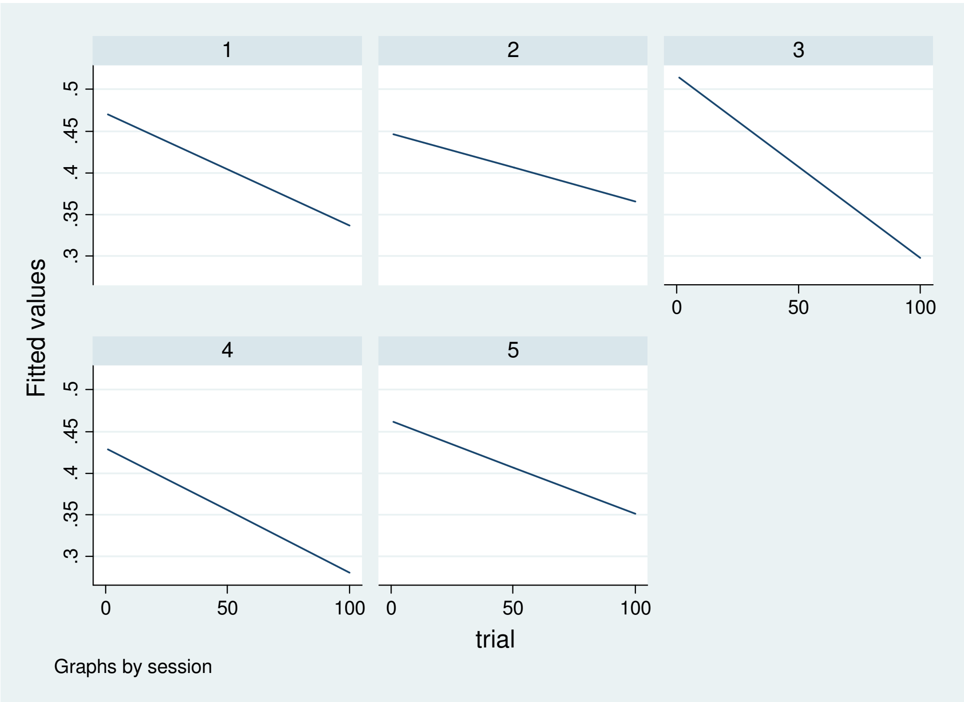



Figure 4: Fitted average choice over trial over sessions

To test for the difference in choice between the five sessions, a simple pair wise mean comparison will be used. This compares the means in choice for all the possible combinations between the five sessions using the variable *session*. The results for this are shown in Table 2. The tested variable is the variable *choice*. From the table below it can be concluded that there are very small differences in the mean choice between the sessions, which are not always significant. There are only significant differences between sessions one and four, two and four, three and four, and four and five, all significant at a 1% significance level. From this it can be concluded that the proportion of risky choice in session four is significantly lower than all other sessions, at about 0.05 (5%). It is known that the sessions differ based upon the problems faced. In Table 39 in the appendix, the used problems per session are shown. If we combine this with Table 37, we can see that there are some differences in the payoff of the problems. Because the difference is very small and only for one session, no differentiation for the sessions will be made in the rest of the analysis.

The analysis of the differences in session thus tell us that the choice behavior between the sessions seems equal. Subjects in session four are slightly less likely to pick the risky choice compared to the other sessions. This might be attributable to the problems faced. The effect of the type of problems will be tested later on in hypothesis 5.

| Session | Difference | Standard error |
|---|---|---|
| 1 to 2 | 0.0025 | 0.0045 |
| 1 to 3 | 0.0030 | 0.0045 |
| 1 to 4 | -0.0489*** | 0.0045 |
| 1 to 5 | 0.0028 | 0.0045 |
| 2 to 3 | 0.0005 | 0.0045 |
| 2 to 4 | -0.0514*** | 0.0045 |
| 2 to 5 | 0.0003 | 0.0045 |
| 3 to 4 | -0.0519*** | 0.0045 |
| 3 to 5 | -0.0003 | 0.0045 |
| 4 to 5 | 0.05167*** | 0.0045 |

**Table 2: Difference between sessions**[6]

[6] *** = significant at 1%, ** = significant at 5%, * = significant at 10%. This distinction is used in all the other tables and regressions.

### Testing on the order

One might argue that the behavior of the subjects differs over time. As learned in the theory, subjects might take more risk in the first problem, whereas in the tenth problem they might want to play it safe. To see whether this is the case, the proportion of risky choice for the 100 trials are plotted separately for all the twelve problems. This means that the first graph deals with the mean risky choice for the first problem everybody faced, and the twelfth graph deals with the mean risky choice for the last problem everybody faced. This outcome is shown in Figure 5. From the fitted curve in these graphs it can be seen that there is a decrease in choice over time *within* the problems, but there seem to be no clear change or differences visible over time *between* the problems. This is another interesting finding of this research, as it is expected that the variance should differ between the beginning and the end of the experiment. However, this is also in line with theory predicting that there should be no differences in behavior amongst problems, only within problems. A statistical test needs to be used to give definitive proof on this.

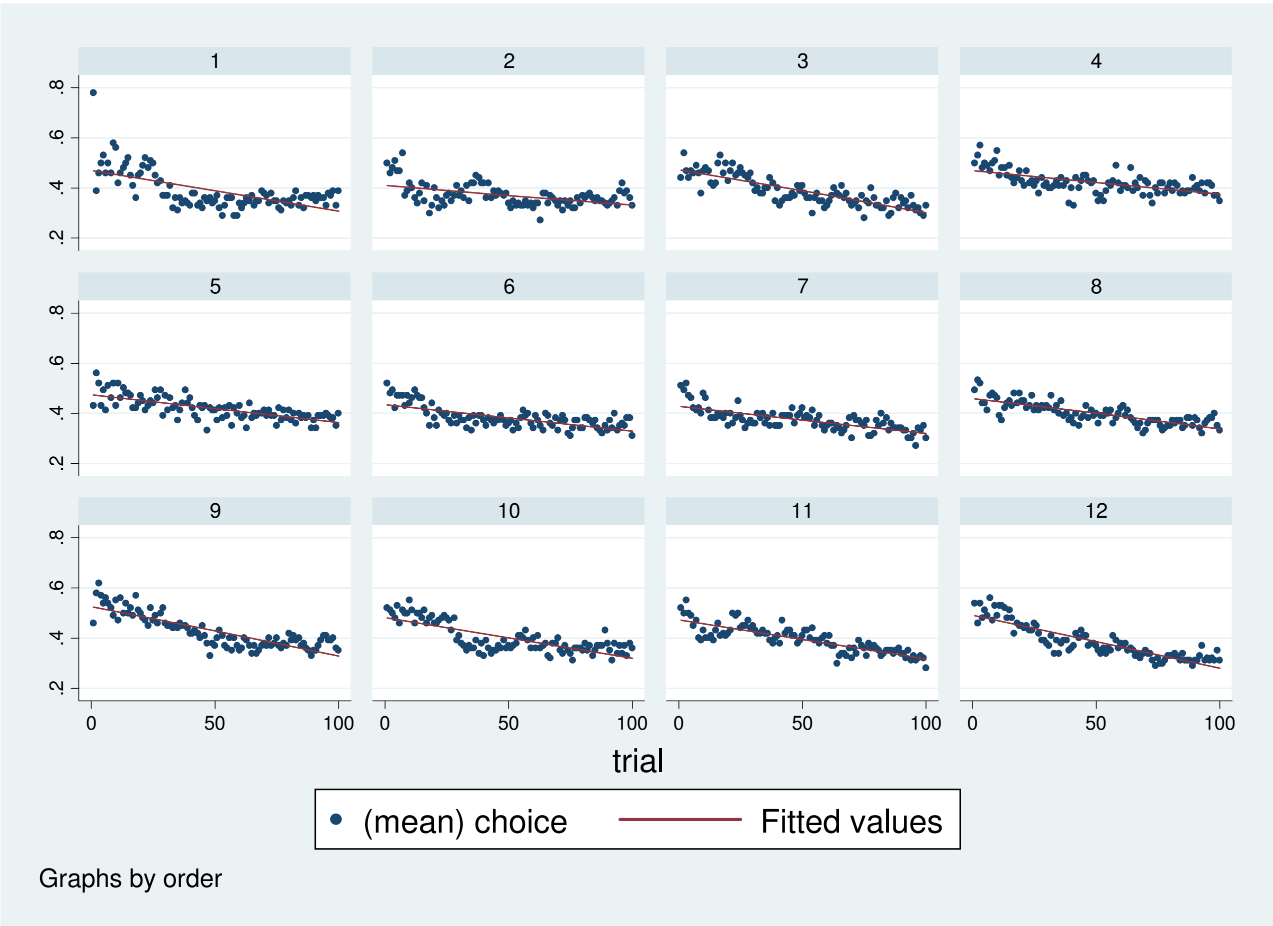



Figure 5: average choice over trials over order

To see whether there are also differences over time in the order of the problem, the mean choice over the order of the problems will be tested. The variable *order* will be used here as this measures the order of the problems the subjects faced. If the order equals five, it deals with the fifth problem out of the twelve a subject is facing.  Again, a pair wise mean comparison is done over the variable *choice*. The results for this are shown in Table 3. Here the difference in the proportion of risky choice between the order of the problems is given. As can be seen, there are some differences in the average mean. After the first problem faced by the subjects, the mean average choice is significantly decreasing, at a 5% significance level. But after the second problem, the average mean choice is significantly increasing, at a 5% significance level. The main conclusion from this table, is that there is no trend visible, the mean is both decreasing and increasing over time with the order of the problems.

Combining the findings from the graphs and the statistical tests, it is concluded that there are little to no differences over the order of the problems faced in the choice behavior of subjects. The difference which are found are very small, and never larger than 0.05 (5%). Therefore it is concluded that the order does not seem to matter.

| Order | Difference | Standard error |
|---|---|---|
| 1 to 2 | -0.0162** | 0.0069 |
| 2 to 3 | 0.0158** | 0.0069 |
| 3 to 4 | 0.0342*** | 0.0069 |
| 4 to 5 | -0.0044 | 0.0069 |
| 5 to 6 | -0.0359*** | 0.0069 |
| 6 to 7 | -0.0085 | 0.0069 |
| 7 to 8 | 0.0252*** | 0.0069 |
| 8 to 9 | 0.0294*** | 0.0069 |
| 9 to 10 | -0.0267*** | 0.0069 |
| 10 to 11 | -0.0053 | 0.0069 |
| 11 to 12 | -0.0101 | 0.0069 |

**Table 3: Difference over time (order)**

## Testing the characteristics of the problems

To see how a variable affects the risky choice behavior, a logit model can be estimated over the variables of interest, as a logit model predicts the change in a binary variable based on the change in the explanatory variables. The logit model is preferred over the probit model in this analysis because of its advantages in implementation. Furthermore the logistic distribution is preferred over the standardized normal distribution of the probit model.

To test change in the risky choice behavior, a panel logit model will thus be used. In the panel, the time is determined by the combination of trials and order[7], and the ID's of the subjects are the identifiers. The results for the simple logit model are shown in Table 3. Here the choice is explained by the high, medium, and low payoffs, as well as the probability of the high outcome: the characteristics of the problem. Note that the relation of the probability of the low outcome is the reverse of the relation of the probability of the high outcome, since p(low) = 1-p(high). From Table 3 it can be concluded that there is a positive significant relation between the p(high), high, low and the dependant variable choice. Since we are dealing with a logit model, this positive relation means that an increase in these variable makes the subject more likely to select the risky outcome (risky = 1).This thus means that an increase in both the high and the low option of the risky choice make the subject more likely to select the risky choice. If the high outcome is more likely (phigh increases), the subject is also more likely to select the risky choice. For the safe(medium) outcome, it can be seen that there is a negative relation, but not significant. To shed more light on this model, control variables need to be added.

| Dependant Variable Choice | Coefficient | Standard Deviation |
|---|---|---|
| phigh | 3.046*** | 0.371 |
| high | 0.061*** | 0.022 |
| low | 0.058** | 0.023 |
| medium | -0.104*** | 0.033 |
| _cons | -2.395*** | 0.261 |

**Table 4: Logit model of the options (restricted)**

In Table 5, the model from Table 4 is expanded with the achieved payoff in the previous trial, the order of the problems, and the number of the trial. This thus shows how the feedback and the timing of the choice affect the behavior of the subject in making a choice. From this table it can be concluded that for the p(high), high, and low, the same relations hold as before. Furthermore it can be seen that the relation for the safe outcome is also the same, but now it is significant. This means that if the payoff of the safe option increases, the subject is less likely to select the risky choice, hence more likely to pick the safe option. For the lagged payoff, it can be seen that this shows a positive significant relation. This implies that feedback will affect the choice in the next outcome.  To see the actual direction of feedback, it must be linked with whether the choice made in the previous trial was a risky or a safe option, and also whether it was a rare or non-rare option. This will be done in the chapters dealing with the specific hypotheses.

---

[7] E.g. if time = 565: order = 6 and trial = 65: 6th problem, 65th trial. Similar, 041: problem = 1, trial = 41.

For the trial it can be seen that there is a negative significant relation. This implies that over time (trials) the subjects are more likely to pick the safe option, and on average there is a settling down behavior on the safe choice. Finally, the order is positive but not significant. This is in line with the analysis on differences in choice behavior dependant on the order of the problems above.

From these models it can thus be concluded that the characteristics of the problem a subject is facing all affect the choice behavior of the subject. An increase in the likelihood of the high option, or in the payoffs of the high and the low option make a subject more likely to select the risky choice, whereas an increase in the payoff of the safe option makes the subject more likely to select the safe option.

| Dependant Variable Choice | Coefficient | Standard Deviation |
|---|---|---|
| phigh | 3.163*** | 0.400 |
| high | 0.058** | 0.024 |
| low | 0.059** | 0.025 |
| medium | -0.162*** | 0.035 |
| trial | -0.011*** | 0.000 |
| order | 0.005 | 0.019 |
| lagged payoff | 0.062*** | 0.004 |
| _cons | -1.959*** | 0.309 |

**Table 5: Logit model of the options (unrestricted)**

## Preliminary findings

The data section shows already some interesting findings regarding the tradeoff and the hypotheses. 40 percent of the subjects pick the risky choice, whereas 60 percent chooses the safe option. Furthermore it is known that the rare outcome only occurred in four percent of all cases, which seems quite low. The primary analysis showed us that the choice can be explained by the payoff and probabilities of the options, as well by the received feedback and the timing of the choice. It is found that there are no significant differences between the sessions, nor amongst the order of the twelve problems a subject is facing.

What these findings mean and how they relate to the five hypotheses can be seen in the next chapters. Each chapter will give findings and analyze these related to the five hypotheses of interest. The combination of these findings will result in conclusions and lessons regarding to the understanding of the tradeoff.

## Analysis of hypothesis 1 – lack of exploration?

The first hypotheses of interest states that the exploration-exploitation tradeoff is caused by the lack of exploration in the subjects: "*The lack of exploration is the main reason of the description-experience gap in the feedback paradigm.*" In this chapter the hypothesis will be tested and investigated in detail. The first step is to see what patterns are present in this exploration by looking at the data. The second step will be to analyze the data trough statistical tests and analyses.

### Data related to hypothesis 1

From the data description we know that the subjects picked the risky choice in 40% of the cases. In all the cases, the rare outcome occurred in 4% of the cases. To see whether subjects have explored enough, it is important to know more about the occurrence of the rare outcome. In Figure 6 the rare outcome is split out into more detail. This figure covers only the trials where the subjects chose the risky outcome. From this it can be seen that in about 10% of the risky choices the rare outcome occurred. This occurred more often for the low outcome than the high outcome (7.3% vs. 3.5%). This is already the first interesting finding of this chapter: the rare outcome is not observed in 90% of all risky choices. This implies that underweighting of rare outcomes is present in the risky choice of the subjects of interest. As depicted in the theory, this underweighting might result in small sample problems if subjects do not explore enough.

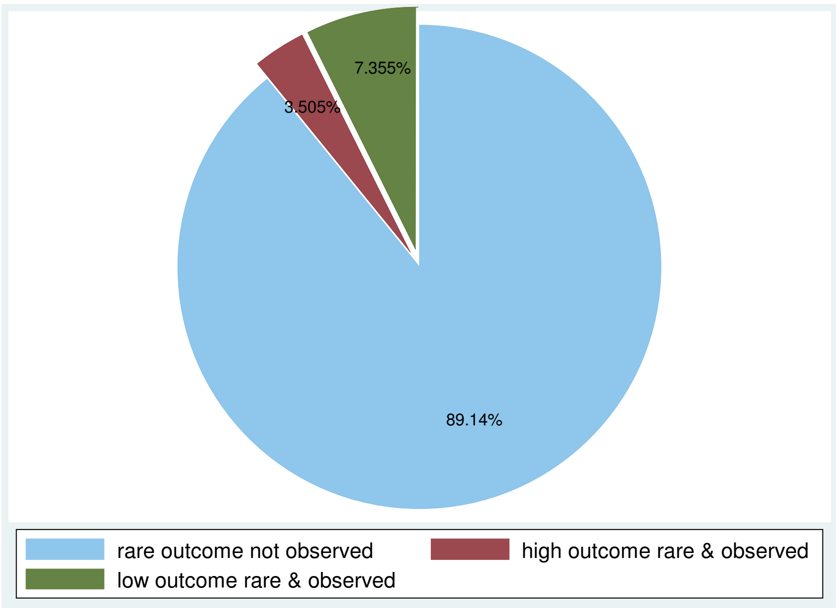



**Figure 6: Rare outcome observed/not observed**

To see how the risky choice evolves over time, the average choice over the trials are plotted over time. With this it can be seen whether there are indications of settling down and whether the subjects explore enough. The first step will be to look at the whole dataset. This is shown in Figure 7, where the average risky choice is set out against the trials, as well as the fitted curve. It seems that there is a negative relation over time, the proportion of risky choice seems to be decreasing with the trials. After about 50 trials a large part of the subjects seem to be settling down on one choice, with a slight bump, as the average choice seems to become stable. This indicates that there is some sort of settling down. This is the second interesting finding of this research: subjects seem to settle down after a certain period on the safe option. This implies that subjects seem to have learned enough after a certain amount of trials to make one choice with which they want to continue.

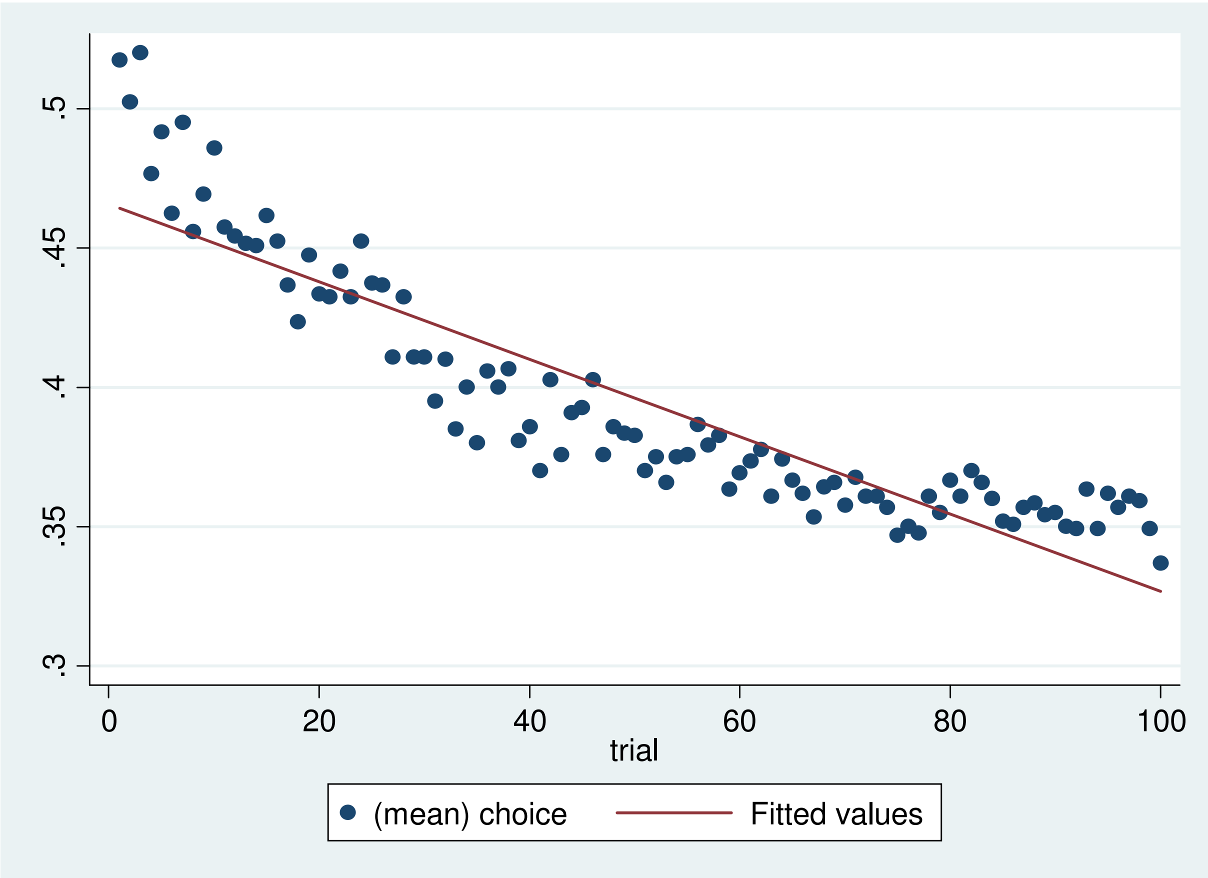



Figure 7: average choice over trials

To zoom in on these differences in more detail and to relate them to the rare outcomes, Figure 7 will be split out for subjects who did or did not experienced the rare outcomes. This will show whether the settling down  and exploration differs regarding the occurrence of the rare outcomes. Figure 8 shows the mean choice for subjects who experienced a rare outcome in any of their trials, Figure 9 for subjects who never experienced a rare outcome. As can be seen from the figures, the fitted curve seems much steeper for subjects who experienced a rare outcome compared to subjects who never experience a rare outcome. Furthermore, the proportion of risky choice for subjects who never experience the rare outcome seems to be much lower. However, subjects who never experience a rare outcome seem to stop exploring after about ten to twenty trials. This implies that subjects who never experience the rare outcome are more risk averse. These findings hint that the settling down is not completely determined by the occurrence of a rare outcome, but that the length of exploration and settling down is affected by the occurrence of the rare outcomes.

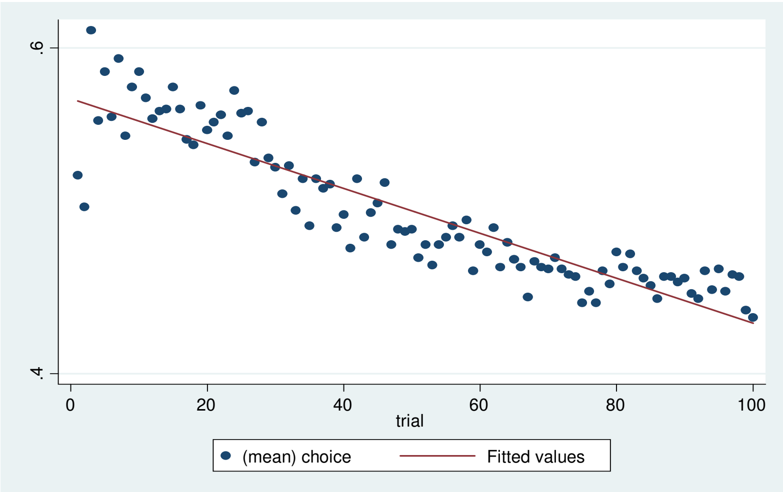



**Figure 8: Mean choice for subjects who experienced any rare outcome.**

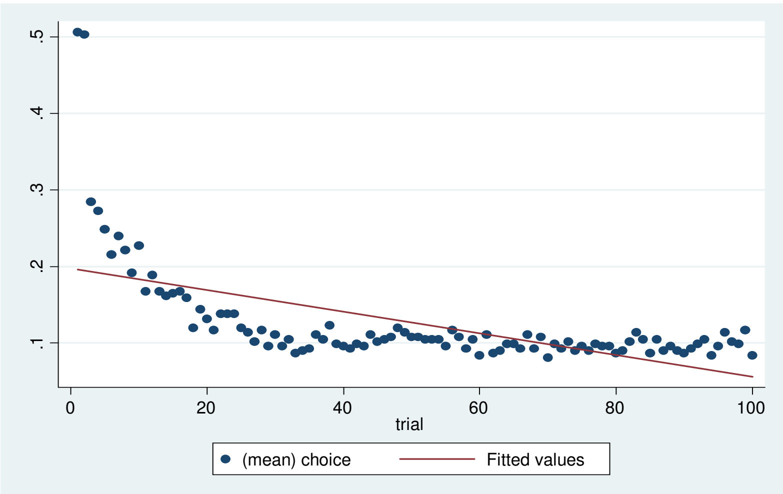



**Figure 9: Mean choice for subjects who did not experienced any rare outcome**

To see how the occurrence of the rare outcome affects this, Figure 9 will be split out separately for subjects who did not experience a good rare or a bad rare outcome. This will help us see whether the settling down is affected if the subject never experienced a positive or negative rare outcome. This is important because subjects who only experience one option of the risky choice did not explore enough. The choice behavior over time for this is shown in Figure 10 and Figure 11. For subjects who never experience the good rare outcome, a clear constantly decreasing settling down pattern over time is visible. Looking at the subjects who never experienced the bad rare outcome, it can be seen that there is only seems to be an immediately settling down in the first 10 trials, after which the proportion of risky choice remains low. The findings from these graphs is interesting because it implies that subjects who do not experience the good rare outcome keep exploring longer and have a higher proportion of risky choice compared to subjects who did not experience the bad rare outcome.

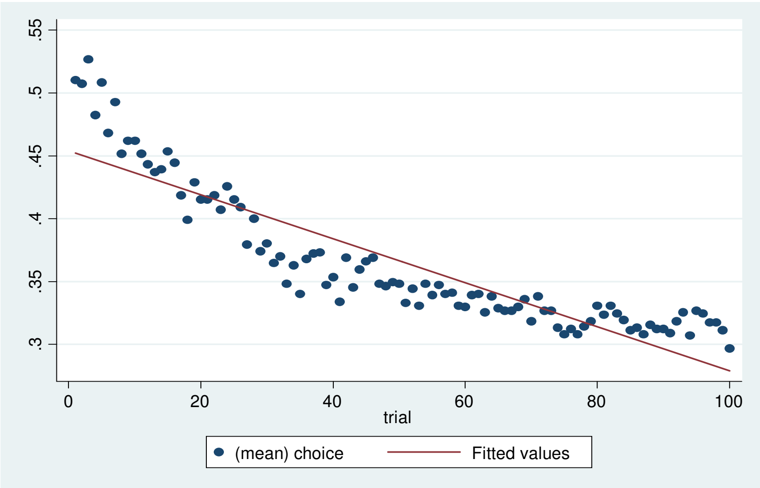



Figure 10: Mean choice for subjects who never experienced a good rare outcome

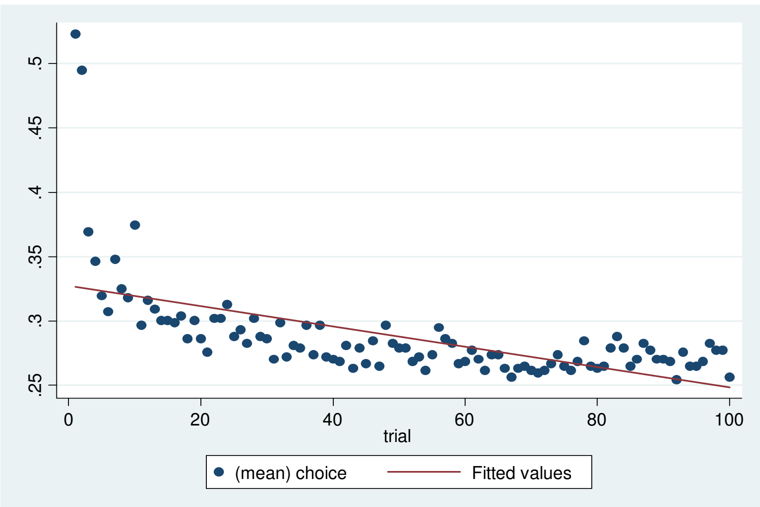



Figure 11: Mean choice for subjects who never experienced a bad rare outcome

## Testing hypotheses 1

The second step is to test the findings from above. The first factor to analyze will be to look at the difference over time in the trials a subject is facing. Recalling Figure 7, the proportion of risky choices seemed to be higher in the beginning of the trials. To see whether this is a significant difference, the trials have been split in groups of both 10 and 20 trials, to see whether the difference in the proportion of risky choices in the variable *choice* is significant. This will make use of the variable *trials*, and is done trough a pair wise mean comparison. The results for this are shown in Table 6. The top part looks at the difference per ten trials. Here, 10-20 compares the difference in the proportion of risky choice between the trials one and ten with the trials eleven till twenty. From this table, it can be seen that the proportion of risky choice is decreasing for the first 50 trials. This means that more subjects are settling down on the safe option over time. This is only significant for the difference between the 30 to 40 trials. The final 50 trials have mixed results, with both small increases and decreases in the proportion of risky choice between the sets of trials. The only significant case here is the 60-70 trials, which shows a significant decrease of the average mean choice.

The bottom part of the table looks at the difference per twenty trials. Here, 20-40 compares the differences in the proportion of choice between the trials one till twenty with the mean of the trials 21 till 40. From this it can be seen that there is a decrease of choice over all the trials. This implies that more and more subjects are switching away from the risky option towards the safe option over time. This is significant in all cases except for the last 20 trials. It also means that most exploration seems to occur in the first part of the trials. Since the division in groups of 10 gives more detail, this is preferred in the rest of the analysis.

| Group size | Trials | Difference | Standard Error |
|---|---|---|---|
| 10 trials | | | |
| | 10 to 20 | -0.0193* | 0.0108 |
| | 20 to 30 | -0.0133 | 0.0128 |
| | 30 to 40 | -0.0756*** | 0.0107 |
| | 40 to 50 | -0.0035 | 0.0094 |
| | 50 to 60 | 0.0014 | 0.0085 |
| | 60 to 70 | -0.0145* | 0.0078 |
| | 70 to 80 | -0.0099 | 0.0072 |
| | 80 to 90 | 0.0036 | 0.0068 |
| | 90 to 100 | -0.0028 | 0.0064 |
| 20 trials | | | |
| | 20 to 40 | -0.0564*** | 0.0059 |
| | 40 to 60 | -0.0272*** | 0.0064 |
| | 60 to 80 | -0.0260*** | 0.0054 |
| | 80 to 100 | -0.0018 | 0.0047 |

**Table 6: Difference over time (trial)**

The finding that there are large differences over time in the trials confirms that there seems to be an exploration exploitation tradeoff. Subjects seem to be settling down over time. To analyze these differences over time in more detail, the analysis will be split out for subjects who experienced a rare outcome in any of the trials, and subjects who did not. The results for this mean difference testing are shown in Table 7. From this it can be seen that the proportion of risky choice is always higher for subjects who experienced a rare outcome in any of the trials. This implies that subjects who experienced a rare outcome, pick the risky choice more often and thus explore more. This is significant in all the cases. The difference in the proportion of risky choice between the subjects who did and did not experience the rare outcome is the largest between the trials 20 and 40, and the smallest in the first ten trials. After 40 trials, the difference seems to remain constant.

Relating this to the hypothesis of exploration it is thus found that the exploration phase seems to differ dependant on whether a subject experiences a rare outcome. Subjects who experience a rare outcome settle down earlier. This means that they have access to a smaller sample compared to subjects who did not experience a rare outcome.

| Trials | Over | Difference | Standard Deviation |
|---|---|---|---|
| 1-10 | rare vs. no rare | 0.272*** | 0.009 |
| 11-20 | rare vs. no rare | 0.434*** | 0.024 |
| 21-30 | rare vs. no rare | 0.450*** | 0.018 |
| 31-40 | rare vs. no rare | 0.359*** | 0.015 |
| 41-50 | rare vs. no rare | 0.361*** | 0.014 |
| 51-60 | rare vs. no rare | 0.373*** | 0.012 |
| 61-70 | rare vs. no rare | 0.396*** | 0.011 |
| 71-80 | rare vs. no rare | 0.385*** | 0.010 |
| 81-90 | rare vs. no rare | 0.379*** | 0.010 |
| 91-100 | rare vs. no rare | 0.353*** | 0.009 |

**Table 7: Difference over rare outcomes**

To investigate this finding further, let us see how the choice behavior between subjects who did and did not experience the rare outcome develops over time. This means that the difference between trials is tested for subjects who did and did not experience a rare outcome. The results for this are shown in Table 8. Looking at the proportions of risky choice over time, it can be seen that there are little differences over time when the choice is split out between subjects who did and did not experience the rare outcome. The largest differences exist in the first 40 trials, specifically in the difference between the trials 1-10 and 11-20. The proportion of risky choice is decreasing here in both cases, but is stronger and significant for the subjects who did not experience the rare outcome. This means that subjects who never experience a rare outcome only explore in the first ten trials, because they believe that there are only two options: the safe and the likely option of the risky choice. Because the difference in proportion of risky choice is relatively stable after these trials, we know that they settle down on one of these choices.

This is not a surprising finding, because there are two effects at work here, both the good and bad rare outcome for the risky choice are included here. The results will need to be repeated for subjects who never experience a good rare outcome and never experience a bad rare outcome. This should also help to clarify the findings from Figure 10 and Figure 11.

| Trials | Rare outcome Observed | | Rare outcome not Observed | |
|---|---|---|---|---|
| | *Difference* | *Standard Deviation* | *Difference* | *Standard Deviation* |
| 11-20 to 1-10 | -0.016 | 0.011 | -0.177*** | 0.023 |
| 21-30 to 11-20 | 0.011 | 0.013 | -0.005 | 0.027 |
| 31-40 to 21-30 | -0.092*** | 0.012 | -0.001 | 0.021 |
| 41-50 to 31-40 | -0.001 | 0.010 | -0.002 | 0.018 |
| 51-60 to 41-50 | 0.006 | 0.009 | -0.006 | 0.016 |
| 61-70 to 51-60 | 0.003 | 0.009 | -0.020 | 0.014 |
| 71-80 to 61-70 | -0.007 | 0.008 | 0.004 | 0.013 |
| 81-90 to 71-80 | -0.002 | 0.008 | 0.004 | 0.012 |
| 91-100 to 81-90 | -0.015** | 0.007 | 0.012 | 0.011 |

**Table 8: differences over time rare observed and not observed**

In Table 9 the results for subjects who never experienced a rare outcome are split out separately for subjects who never experienced the good rare outcome or the bad rare outcome. Interesting to see here is that the significant decrease in the proportion of risky choice in the first ten trials is only present for subjects who do not observe the good rare outcome. This is in line with the expectations, because if you did experience the good rare outcome, you might want to select the rare outcome again. Furthermore it can be seen that the largest settling down occurs in the first 30 trials. There is a significant decrease over time between the trials 21-30 and 31-40 for both types of subjects here. This was not significant in Table 8, implying that subject who experience at least one type of rare outcome settle down more easily. Furthermore it can be concluded that the settling down on the safe choice persists the longest for subjects who never experience the good rare outcome, which is also in line with expectations. The decrease in the proportion of risky choice for subjects who never experience the bad rare outcome is not significant in most of the cases. The main finding here is thus that only subjects who do not experience the good rare outcome immediately settle down, subjects who do not experience the bad rare outcome settle down later on.

| Trials | Good Rare never Observed | | Bad Rare never Observed | |
|---|---|---|---|---|
| | *Difference* | *Standard Deviation* | *Difference* | *Standard Deviation* |
| 11-20 to 1-10 | -0.043*** | 0.012 | -0.012 | 0.015 |
| 21-30 to 11-20 | -0.020 | 0.014 | -0.030* | 0.018 |
| 31-40 to 21-30 | -0.076*** | 0.012 | -0.043*** | 0.015 |
| 41-50 to 31-40 | 0.001 | 0.010 | -0.017 | 0.013 |
| 51-60 to 41-50 | -0.005 | 0.009 | 0.005 | 0.012 |
| 61-70 to 51-60 | -0.013 | 0.008 | -0.033*** | 0.011 |
| 71-80 to 61-70 | -0.009 | 0.008 | -0.007 | 0.010 |
| 81-90 to 71-80 | -0.003 | 0.007 | 0.021** | 0.009 |
| 91-100 to 81-90 | -0.003 | 0.007 | 0.011 | 0.009 |

**Table 9: difference over time good rare or bad rare not observed**

## Conclusions on hypothesis 1

With respect to hypothesis 1 it can be stated that the differences in choice behavior can be explained by the settling down behavior of the subjects. This because the proportion of risky choice is decreasing over time, which is stronger for subjects who experienced the rare outcome opposite to subjects who did not experience the rare outcome. Specifically, subjects who do not experience the rare outcome in the first ten trials immediately become averse to the risky choice. When this was investigated in more detail, it was found that this holds for subjects who do not experience the good rare outcome in the first tent trials. Subjects who do experience the rare outcome seem to explore more and settle down later on. This means that hypothesis 1 is found to be true because subjects move away from the risky choice over time, and settle down. This causes them to reduce their sample size, hence giving room for the gap. The rare outcome also explains these differences.

## Analysis of hypothesis 2 - variance

Now that we know that the small sample problems, caused by the lack of exploration is causing the description-experience gap, we can zoom in on the other factors. The second hypothesis of interest deals with a factor which should cause an increase in the exploration phase, the variance. This was captured with the following hypothesis: "*Subjects explore more when they experience more variance*." This chapter will thus test if and how the variance in the payoffs from the previous samples affects the exploration and risky choice.

### Data related to hypothesis 2

The experienced variance is based on the outcome of the previous variables, and how this differs from the expectations. Specifically, this means that we are interested in what happens to the exploration when the subject experienced a rare outcome in the previous trial(s). In the data we saw this already in the analysis of hypothesis 1, specifically in Figure 8 and Figure 9, where we saw that the exploration was less for subjects who experienced a rare outcome. In these graphs it is not clear when the rare outcome occurred and how many trials are taken into account. Therefore a statistical analysis on the lagged outcomes are required.

### Testing hypothesis 2

To test the effect of the variance, a logit model will be estimated to analyze the relation of the lagged previous outcome on the risky choice. When this is combined with a variable for whether a rare outcome occurred in the previous period, we can conclude what the effect of variance on the proportion of risky choice is. This means that the variables for the rare outcomes for both the high and the low outcome are regressed on the choice, in a panel logit regression. The first step will be to estimate a restricted model. This is the lagged dummy variables on the variable choice, which has a value of 1 if the outcome is the highest possible payoff and has the smallest probability. The results for this are shown in Table 11. From this it can be concluded that there is a significant and positive relation for both cases. If the rare outcome occurred in the previous trial, the subjects are more likely to select the risky prospect in the next trial. In other words, if the variance increases the exploration seems to continue, as the subject is likely to select the risky choice again. It was expected that when the rare outcome with the low payoff occurred, the subjects would become averse to the risky prospect. However, this is not found here. Opposite to this it does seem to confirm that variance increases the exploration. Two effects might work against each other here. What can be seen is that the coefficient for the high outcome is much higher than the low outcome, implying that the positive or negative variance might have an effect. Nothing can be concluded about that in this model. It implies that subjects keep exploring when there is variation and ignore the associated costs.

| Dependant Variable Choice | Coefficient | Standard Deviation |
|---|---|---|
| 1 lag (high) | 1.212*** | 0.077 |
| 1 lag (low) | 0.095** | 0.047 |
| _cons | -0.779*** | 0.069 |

**Table 10: Logit model of previous outcome (restricted)**

To be able to give definitive answers about hypothesis two, the model needs to be expanded with control factors to correct for the size and likelihood of the previous option. This includes the payoffs of the options and the probabilities, as well as the variables measuring the trial and the order of the problem. The unrestricted model from this is shown in Table 10. As can be seen from this table the control variables *phigh, high, low, medium,* and *trial* have the same relations as found in the data description. These control variables are all significant (at a 5%s significance level). The order of the problem does not have a significant effect on the risky choice. The main conclusion from this table is that when the rare outcome occurred in the previous choice, the subject is more likely to select the risky choice again in the next trial. This seems to be strong and significant for when the high rare outcome occurred in the previous sample. The lagged rare bad outcome is positive but does not seem to have a significant effect on the proportion of risky choice. This gives rise to one interesting finding however: exploration seems to be increasing based on the experienced variability even if it is results in costs. The analysis of hypothesis 3 will analysis this in more detail. The findings of this table do imply that there is information search present in the subjects' behavior. To reduce the noise in the model more lags will be added, as subjects might also take the rare outcome into account if it happened more trials ago. This will help to conclude how much of the sample size is used, dependant on the occurrence of rare outcomes.

| Dependant Variable Choice | Coefficient | Standard Deviation |
|---|---|---|
| 1 lag (high) | 1.202*** | 0.077 |
| 1 lag (low) | 0.053 | 0.047 |
| phigh | 3.215** | 0.398 |
| high | 0.059** | 0.024 |
| low | 0.060** | 0.025 |
| medium | -0.103*** | 0.035 |
| trial | -0.011*** | 0.000 |
| order | 0.005 | 0.019 |
| _cons | -2.005*** | 0.308 |

**Table 11: Logit model of previous outcome (unrestricted)**

In steps of five, lags will be added to see how far back into time the experienced variance (measured trough the occurrence of rare outcomes) has an effect on the exploration (measured trough the variable for the risky choice). The results for this analysis are shown in Table 12. The left column shows the model with five lags for both rare high and rare low outcomes.[8] Looking at the results, we can see that the lagged samples seem to have a positive effect for all the five samples. If the rare outcome occurred, regardless whether it was the high or the low outcome, it will increase the proportion of risky choice. Only the first lag for the low outcome is not significant, whereas all the other lags are. Looking at the sizes of the lags, it can be seen that the lags seem to have less effect over time as the coefficients are decreasing when the lags are increasing. Furthermore, the good rare outcome has a larger effect than the bad rare outcome, in line with previous findings.

When a total of ten lags are included in the second column, it can be seen that the first five lags seem to remain significant. For the bad rare outcome it can be seen that the lags stop after about eight samples, the ninth and tenth lag are not significant. Interesting to note is that the first lag for bad rare outcomes is again not significant. This indicates that subjects do not alter their decision if the most recent outcome was negative. For the good rare outcome, all the ten lags are significant. Interesting to note here is that for the lags six till ten, the size of the effect is no longer negative, a high rare outcome six samples ago seems to have the same effect on the proportion of risky choice as a high rare outcome ten samples ago.

When a total of fifteen lags are included in the third column, it can be seen that for the lagged rare high outcome all the lags are significant. In the fourth column, five more lags are added, for a total of twenty lags. Again all the lagged rare high outcomes are significant at a 1% significance level, and have a positive relation. This indicates that subjects are more likely to select the risky choice if the rare high outcome occurred in any of the previous samples. A positive shock makes the subject more favorable to the risky option, he beliefs he can get the high amount again at any point of time. This result implies that the subject overvalue and therefore explores the risky option more, because they have experienced more positive variance in the risky choice. It can also be seen in the table that the size of the coefficients seems to be decreasing when the lags are increasing. This implies that the decision making is affected the most by a rare outcome in the most recent outcomes. From this it seems that there is some sort of recency effect present, but in a very weak form, as all the outcomes seem to have an effect.

[8] To give an example, Lag 4 (high) can be interpreted as the dummy variable for when the rare high outcome occurred four samples ago, which is positive and significant at a 1% significance level.

Looking at the lagged variables for where the bad outcome is rare, it can be seen that after eight lags the lags no longer have a significant relation in the models with fifteen and twenty lags. The first eight lags are significant at 1% significance level and have a positive relation. This means that if the bad rare outcome occurred in any of the most recent eight samples, the subjects will keep exploring because of the perceived variance. However if it occurred further back in time the subjects do not seem to take this into account when making their choice. This means that the experienced variance is only relevant for losses in the most recent outcomes. Again it can be seen that the first sample remains insignificant. Furthermore the coefficients for the lags seem to decrease over time, indicating that there are recency effects present for the subjects when facing losses. Only the last eight samples affect the choice, where the second sample has a stronger impact than the eight' sample. This is a stronger form of recency effects compared to the high rare outcomes, as the bad rare outcome only takes the last eight samples into account, opposite to all samples for the high rare outcome.

## Conclusions on hypothesis 2

The above analysis showed that the experience variance can be measured by investigating when the rare outcome occurred, and how long it affects the decision making process. Two different effects can be found here. For a positive shock it can be stated that subjects who experienced a positive shock in any of the previous samples always seem to become more favorable to the risky choice. This effect is decreasing over time, implying that there are recency effect present. Regarding the hypothesis of the experienced variance it is found that when the subjects experience more variance, measured by positive shocks, they are more likely to explore more. For a negative shock it seems that subjects neglect the occurrence of a rare outcome if it happened in the previous sample. Furthermore, subjects only take the last eight samples into account when making their decision with a decreasing effect over time, implying stronger recency effect compared to positive shocks. Intuitively this might make sense, because subjects might at first move away from an option after experience a bad outcome. But after a while, they might get curious what else could happen in that option, and they might select it again, explaining the increase in the proportion of risky choice.

For the experienced variance, this means that over bad rare outcomes, subjects use a smaller sample for their decision making. Subjects will explore longer if a rare negative shock occurred in the most recent outcomes. In general there thus seems to be supporting evidence for hypothesis 2, subjects seem to explore longer when they experience more variance. The most interesting finding is that they also explore more when they experience a bad shock, the rare low outcome, but the impact of a negative shock is shorter than a positive shock.

| Dependant variable: Choice | model with 5 lags | | model with 10 lags | |
|---|---|---|---|---|
| | Coefficient | Standard Deviation | Coefficient | Standard Deviation |
| trial | -0.011*** | 0.000 | -0.011*** | 0.000 |
| order | 0.005 | 0.018 | 0.004 | 0.018 |
| phigh | 3.156*** | 0.388 | 3.136*** | 0.382 |
| high | 0.055** | 0.023 | 0.052** | 0.023 |
| low | 0.056** | 0.024 | 0.054** | 0.024 |
| medium | 0.095*** | 0.007 | -0.088*** | 0.034 |
| | | | | |
| 1 lag (high) | 1.185*** | 0.078 | 1.185*** | 0.078 |
| 2 lag (high) | 0.889*** | 0.075 | 0.886*** | 0.076 |
| 3 lag (high) | 0.667*** | 0.073 | 0.651*** | 0.074 |
| 4 lag (high) | 0.850*** | 0.074 | 0.825*** | 0.075 |
| 5 lag (high) | 0.585*** | 0.072 | 0.524*** | 0.073 |
| 6 lag (high) | xx | xx | 0.634*** | 0.074 |
| 7 lag (high) | xx | xx | 0.485*** | 0.073 |
| 8 lag (high) | xx | xx | 0.720*** | 0.074 |
| 9 lag (high) | xx | xx | 0.490*** | 0.072 |
| 10 lag (high) | xx | xx | 0.617*** | 0.073 |
| 11 lag (high) | xx | xx | xx | xx |
| 12 lag (high) | xx | xx | xx | xx |
| 13 lag (high) | xx | xx | xx | xx |
| 14 lag (high) | xx | xx | xx | xx |
| 15 lag (high) | xx | xx | xx | xx |
| 16 lag (high) | xx | xx | xx | xx |
| 17 lag (high) | xx | xx | xx | xx |
| 18 lag (high) | xx | xx | xx | xx |
| 19 lag (high) | xx | xx | xx | xx |
| 20 lag (high) | xx | xx | xx | xx |
| | | | | |
| 1 lag (low) | 0.042 | 0.047 | 0.043 | 0.047 |
| 2 lag (low) | 0.432*** | 0.048 | 0.431*** | 0.048 |
| 3 lag (low) | 0.235*** | 0.047 | 0.238*** | 0.047 |
| 4 lag (low) | 0.241*** | 0.047 | 0.238*** | 0.047 |
| 5 lag (low) | 0.166*** | 0.047 | 0.165*** | 0.047 |
| 6 lag (low) | xx | xx | 0.171*** | 0.047 |
| 7 lag (low) | xx | xx | 0.159*** | 0.047 |
| 8 lag (low) | xx | xx | 0.118** | 0.047 |
| 9 lag (low) | xx | xx | 0.031 | 0.047 |
| 10 lag (low) | xx | xx | 0.073 | 0.047 |
| 11 lag (low) | xx | xx | xx | xx |
| 12 lag (low) | xx | xx | xx | xx |
| 13 lag (low) | xx | xx | xx | xx |
| 14 lag (low) | xx | xx | xx | xx |
| 15 lag (low) | xx | xx | xx | xx |
| 16 lag (low) | xx | xx | xx | xx |
| 17 lag (low) | xx | xx | xx | xx |
| 18 lag (low) | xx | xx | xx | xx |
| 19 lag (low) | xx | xx | xx | xx |
| 20 lag (low) | xx | xx | xx | xx |
| constant | -2.054*** | 0.301 | -2.099*** | 0.295 |

| | model with 15 lags | | | model with 20 lags | |
|---|---|---|---|---|---|
| Dependant variable: Choice | | | | | |
| | Coefficient | dard Deviation | | Coefficient | Standard Deviation |
| trial | -0.011*** | 0.000 | | -0.011*** | 0.000 |
| order | 0.004 | 0.018 | | 0.004 | 0.018 |
| phigh | 3.144*** | 0.379 | | 3.152*** | 0.377 |
| high | 0.050** | 0.022 | | 0.049** | 0.022 |
| low | 0.052** | 0.024 | | 0.051*** | 0.024 |
| medium | -0.084** | 0.033 | | -0.082** | 0.033 |
| | | | | | |
| 1 lag (high) | 1.192*** | 0.079 | | 1.193*** | 0.079 |
| 2 lag (high) | 0.877*** | 0.076 | | 0.870*** | 0.076 |
| 3 lag (high) | 0.629*** | 0.074 | | 0.621*** | 0.075 |
| 4 lag (high) | 0.800*** | 0.075 | | 0.791*** | 0.076 |
| 5 lag (high) | 0.499*** | 0.074 | | 0.493*** | 0.074 |
| 6 lag (high) | 0.607*** | 0.074 | | 0.593*** | 0.074 |
| 7 lag (high) | 0.457*** | 0.073 | | 0.439*** | 0.073 |
| 8 lag (high) | 0.691*** | 0.074 | | 0.666*** | 0.074 |
| 9 lag (high) | 0.444*** | 0.073 | | 0.418*** | 0.073 |
| 10 lag (high) | 0.559*** | 0.073 | | 0.532*** | 0.073 |
| 11 lag (high) | 0.274*** | 0.072 | | 0.249*** | 0.072 |
| 12 lag (high) | 0.412*** | 0.072 | | 0.385*** | 0.073 |
| 13 lag (high) | 0.330*** | 0.072 | | 0.295*** | 0.072 |
| 14 lag (high) | 0.460*** | 0.072 | | 0.417*** | 0.072 |
| 15 lag (high) | 0.428*** | 0.072 | | 0.374*** | 0.072 |
| 16 lag (high) | xx | xx | | 0.316*** | 0.072 |
| 17 lag (high) | xx | xx | | 0.213*** | 0.072 |
| 18 lag (high) | xx | xx | | 0.40*** | 0.073 |
| 19 lag (high) | xx | xx | | 0.289*** | 0.072 |
| 20 lag (high) | xx | xx | | 0.194*** | 0.072 |
| | | | | | |
| 1 lag (low) | 0.041 | 0.047 | | 0.039 | 0.047 |
| 2 lag (low) | 0.431*** | 0.048 | | 0.430*** | 0.048 |
| 3 lag (low) | 0.238*** | 0.047 | | 0.237*** | 0.047 |
| 4 lag (low) | 0.239*** | 0.047 | | 0.238*** | 0.047 |
| 5 lag (low) | 0.166*** | 0.047 | | 0.165*** | 0.047 |
| 6 lag (low) | 0.173*** | 0.047 | | 0.173*** | 0.047 |
| 7 lag (low) | 0.162*** | 0.047 | | 0.163*** | 0.047 |
| 8 lag (low) | 0.120*** | 0.047 | | 0.122*** | 0.047 |
| 9 lag (low) | 0.034 | 0.047 | | 0.036 | 0.047 |
| 10 lag (low) | 0.075 | 0.047 | | 0.077 | 0.047 |
| 11 lag (low) | 0.057 | 0.047 | | 0.060 | 0.047 |
| 12 lag (low) | 0.051 | 0.047 | | 0.053 | 0.047 |
| 13 lag (low) | 0.007 | 0.047 | | 0.011 | 0.047 |
| 14 lag (low) | -0.011 | 0.047 | | -0.008 | 0.047 |
| 15 lag (low) | -0.067 | 0.047 | | -0.062 | 0.047 |
| 16 lag (low) | xx | xx | | -0.011 | 0.047 |
| 17 lag (low) | xx | xx | | -0.020 | 0.047 |
| 18 lag (low) | xx | xx | | 0.013 | 0.047 |
| 19 lag (low) | xx | xx | | 0.001 | 0.047 |
| 20 lag (low) | xx | xx | | 0.034 | 0.047 |
| constant | -2.125*** | 0.293 | | -2.144*** | 0.292 |



**Table 12: logit model with multiple lagged rare outcomes**

## Analysis of hypothesis 3 – the cost of learning

Opposite to the expectations from hypothesis two over the experienced variance are the hard numbers from the costs of learning. Subjects might lose some money by exploring, and are less likely to explore when they experience high costs of learning. This was captured by the following hypothesis: *"Subjects explore less when they experience higher costs of learning."* The first step will be to see how this cost of learning is present in the data, after which it can be analyzed.

### Data of hypothesis 3

The cost of learning in a problem is determined by both the height of the different payoff options and the difference between the options. To see the difference in payoff per problem, one can look at Table 38 in the appendix. This shows all sixty problems used with all the differences in payoffs. From this table it can be seen that there are quite some differences between the payoffs and hence in the cost of learning within a problem. These payoffs can be compared with the choice behavior over time per problem. The results for all the problems are shown in Figure 23 and Figure 24 in the appendix, where the average choice per problem is set out over time. To zoom in on some specific examples, let us look at Figure 13 and Figure 12. In Figure 13 the risky choice over time for the problems 56 and 60 are shown. Within these problems there are small costs of learning, and it can be seen from the graphs that there is a sharp decrease in the proportion of risky choice over time. This indicates that lower costs of learning result in less exploration. But for some problems with small costs of learning, the proportion of risky choice seems to remain relatively stable over time, as can be seen from problem 5 and 57 in Figure 12.

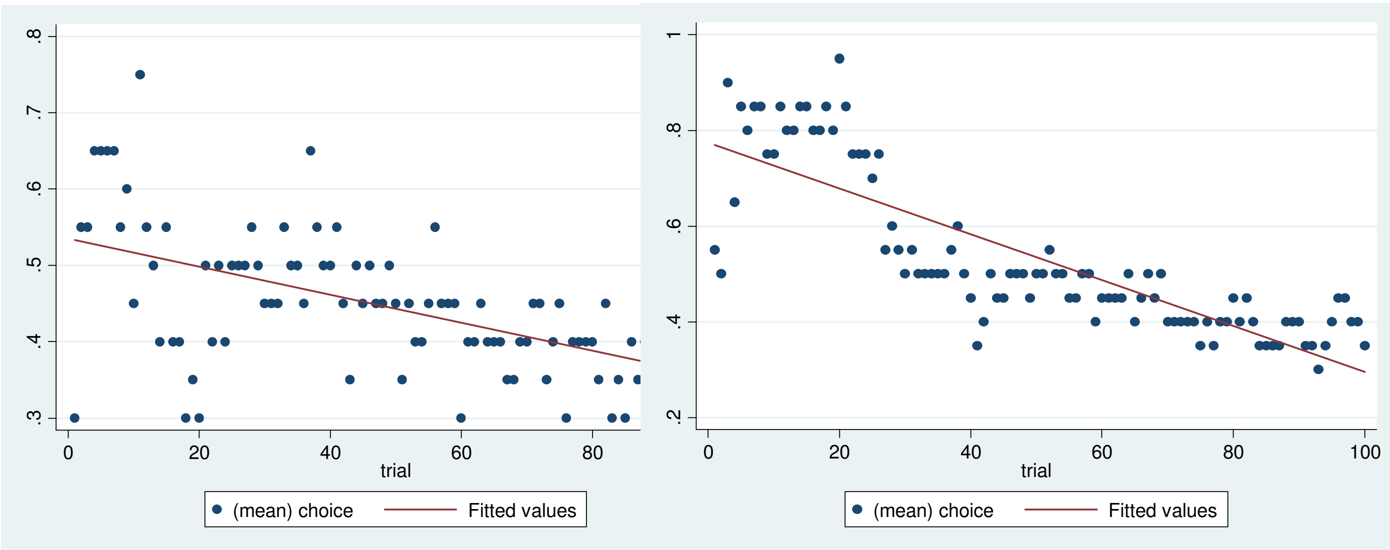



Figure 13: choice over time for problems 56 (left) and 60 (right)

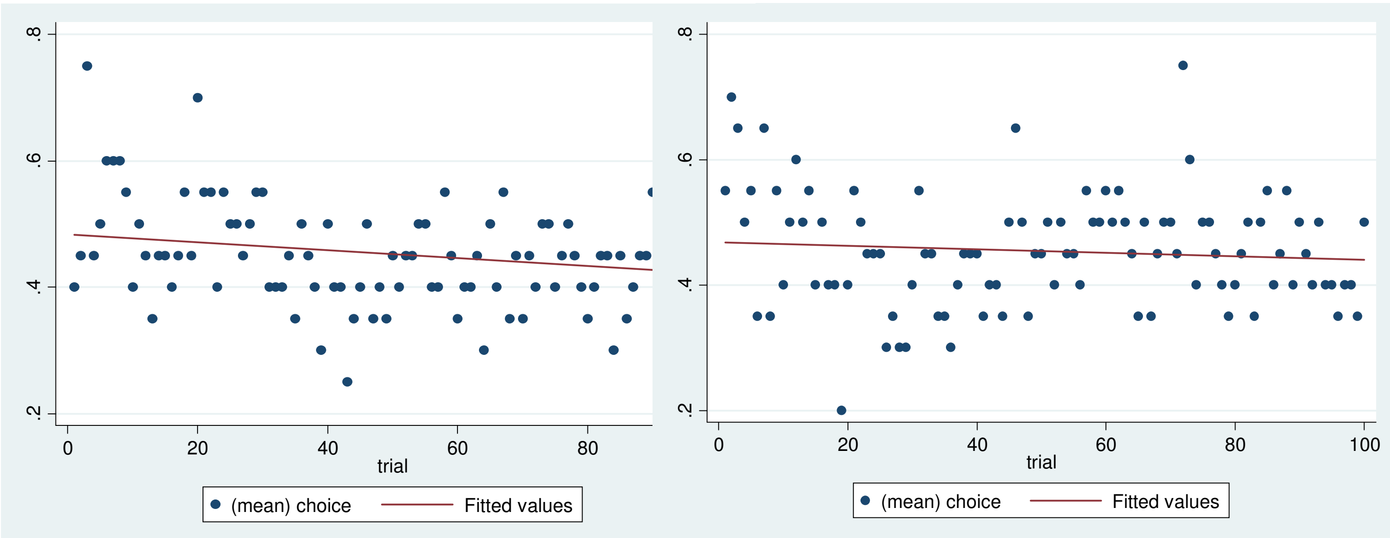



Figure 12: choice over time for problems 5 (left) and 57 (right)

To see whether the opposite is true, we look at problems with relative large costs of learning in the payoffs. In Figure 14 we have problem 6 and 32, and it can be seen that for problems where the costs of learning are relatively high, the proportion of risky choice seems to be more stable This means that a higher cost of learning results in more exploration. However, the mean risky choice is occasionally also decreasing when the variance is high. This can be seen in Figure 15, showing the proportion of risky choice over time for the problems 9 and 20. This thus indicates that there is mixed evidence on the effect of the variance of payoffs on the problems. Both high and low variance indicate that there is mixed evidence on the behavior over time. No clear relation can be from the graphs. Statistical tests and analyses are needed to see what is true, and whether there actually exist such a relation

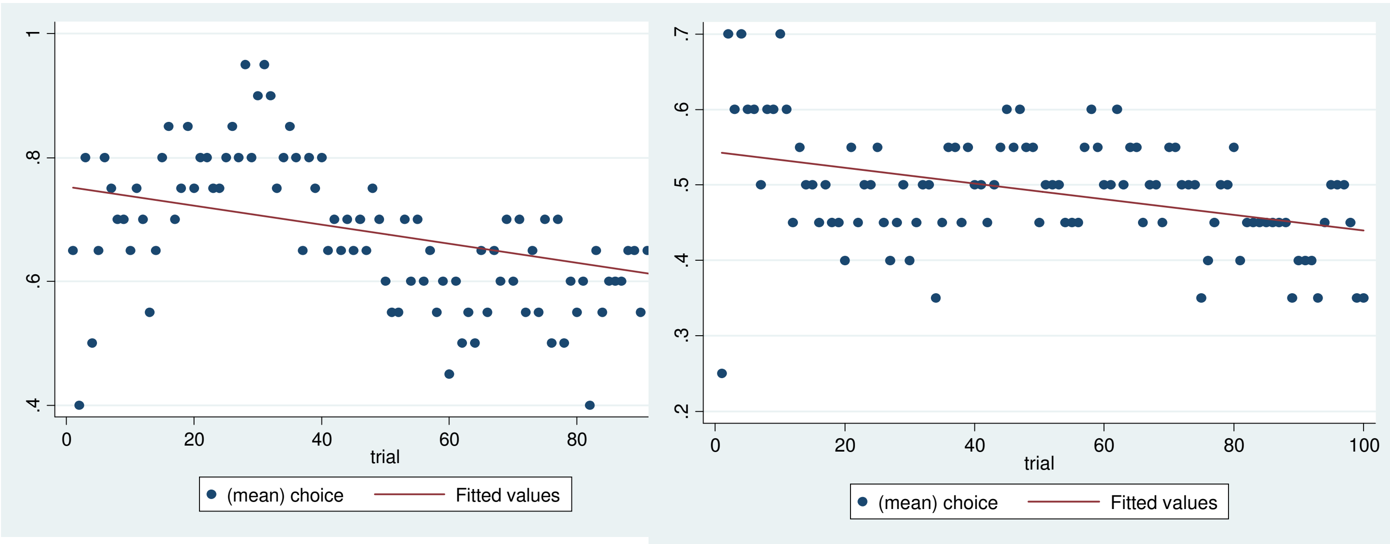



Figure 14: mean choice over time for problems 6 (left) and 32 (right)

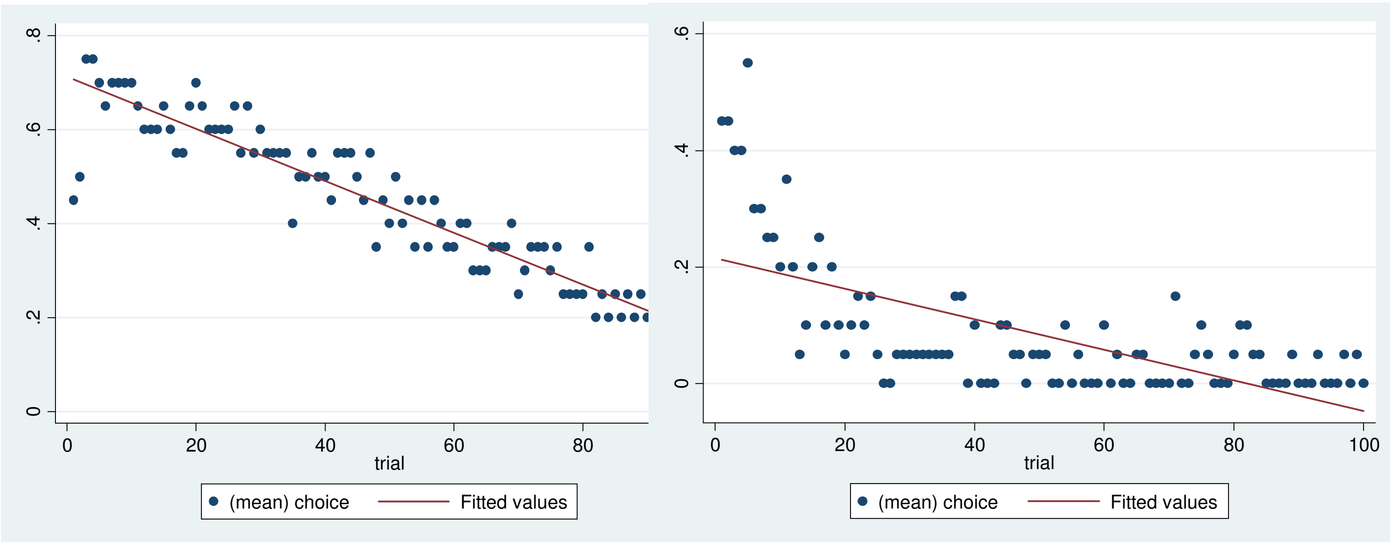



Figure 15: mean choice over time for problems 9 (left) and 20 (right)

## Testing hypothesis 3

To test the cost of learning, a logit model will be estimated where the difference between the high and the safe option and the difference between the safe and the low option will be regressed on the variable choice. This means that the explaining variables are the cost of learning, measured by the difference between the safe option and the relevant high and low option. These differences are labeled as the variables gain and lose, as the exploration on the payoffs can either be a gain or a loss relative to the certain amount the subjects can receive. The results for this restricted logit model are shown in Table 13. From this table it can be concluded that there is a negative significant relation between the variable gain and the choice. This means that when the difference between the safe and the high option gets larger (the gains of learning), the subjects are less likely to select the risky choice.

Looking at the variable loss, it can be seen that there is a significant positive relation. This means that if the difference between the safe option and the low option from the risky choice gets larger (the cost of learning increase), the subject is more likely to select the risky choice. It is expected that this negative relation can be explained by the associated probabilities, as a higher payoff in the risky choice is often followed by a low chance of happening. Therefore the main model needs to be expanded with several control variables.

| Dependant variable: Choice | Coefficient | Standard Deviation |
|---|---|---|
| gain | -0.072*** | 0.015 |
| lose | 0.069*** | 0.018 |
| _cons | -0.735*** | 0.169 |

**Table 13: Logit model costs of learning**

In the unrestricted logit model over the cost of learning, the trial, order, and the probabilities will be added to the model. The safe option, measured by the variable medium, will also be added to the model to show whether the size of the safe option matters. Table 13 shows the results for this model. Looking at the control variables it can be seen that the variable trial is negative and significant, similar as before. The variables *phigh* and medium are positive and significant. For the variable medium, this is quite unexpected, because an increase in the safe option would make the subject more likely to select the risky option, which seems highly irrational. This might be explained by the fact that the difference between the safe and high option could be smaller when the safe option is larger. Looking at the variables of interest it can be seen that the variable for gains (the gains of learning) is significant and positive at a 1% significance level. This means that when the subject can gain from learning, the subject is more likely to select the risky choice. This because they are more likely to select the risky option when the risky choice can yield more. Opposite to this it can be seen that the variable for losses (the costs of learning) is significant and positive at a 5% significance level. This means that when the cost of learning increase, the subjects are less likely to select the risky choice. Intuitively this means that subjects are less likely to select a risky choice if the difference between the safe option and the low option of the risky choice is larger: subjects can lose more when they select the risky choice. This means that subjects tend to explore less (more) when the cost of learning are high(low).

As seen in hypothesis 2, the experienced variance seems to have an effect on the risky choice in the previous trials. Therefore in the second column of Table 14, the lagged effects are included.[9] The difference between the two models is determined by whether we did or did not include the experienced variance. As can be seen in the table the relations remain the same. There are minor changes in the size of the significance: the variables gain is now significant at a 5% significance level. This finding proves that the effects of the cost of learning are not biased by the experienced variance.

| Dependant variable: Choice | Coefficient | Standard Deviation | Coefficier | Standard Deviation |
|---|---|---|---|---|
| gain | 0.063*** | 0.022 | 0.052** | 0.023 |
| lose | -0.060** | 0.024 | -0.054** | 0.023 |
| medium | 0.015** | 0.007 | 0.018** | 0.007 |
| trial | -0.011*** | 0.000 | -0.011*** | 0.000 |
| order | 0.005 | 0.018 | 0.004 | 0.018 |
| phigh | 3.101*** | 0.378 | 3.166*** | 0.383 |
| _cons | -1.905*** | 0.292 | -2.098*** | -7.080 |
| lagged rare outcome included? | No | | Yes | |

**Table 14: Logit model gain or lose (unrestricted)**

To give more insight on the costs of learning, let us consider two separate cases. The first one is where the good rare outcome is never observed. This makes the cost of learning the difference between the medium and the low option. The second case deals with subjects who never observe the bad rare outcome. This makes the cost of learning the difference between the high and the medium option. To test this model in Table 14 has been estimated separately for subjects who either did not experience the good or the bad rare outcome. The results for this are shown in Table 15. The left side deals with subjects who did not experience the good rare outcome. The cost of learning are captured here by the variable *lose*. As can be seen from the table the cost of learning are negative but not significant. The other variables dealing with payoffs are also not significant. The right side of the table deals with subjects who never experienced a bad rare outcome. Here the cost of learning are captured here by the variable *gain*. It can be seen that when the cost of learning get higher there is  a significant increase in the proportion of risky choice. Interesting to note is that the order of the problems is significant in this model. The main conclusion from this table is that there are  significant costs of learning for subjects who never experience the bad rare outcome, but not for those who never experienced the good rare outcome.

[9] The lags are not reported to keep the table simple and clear. Significance and relation similar as findings reported in the chapter on hypothesis 2.

| Dependant variable: Choice | Coefficient | Standard Deviation | Coefficier | Standard Deviation |
|---|---|---|---|---|
| gain | 0.042 | 0.026 | 0.061** | 0.026 |
| lose | -0.016 | 0.021 | -0.462*** | 0.088 |
| medium | 0.007 | 0.007 | 0.008 | 0.009 |
| trial | -0.015*** | 0.000 | -0.008*** | 0.000 |
| order | -0.046*** | 0.017 | 0.047* | 0.028 |
| phigh | 4.375*** | 0.397 | 6.332*** | 1.008 |
| _cons | -2.852*** | 0.331 | -2.227*** | 0.350 |
| Specification: of subjects | Never experienced the good rare outcome | | Never experienced the bad rare outcome | |

**Table 15: Logit model gain or lose for subjects who did not experience type of rare outcome[10]**

## Conclusions on hypothesis 3

Looking at the results from this analysis it can be seen that the first important finding is that the probability of the risky choice is important to take into account. If this is not included reversed relations are found. The main finding with respect to the cost of learning is that subjects tend to explore less when the cost of learning are high. This is measured trough the difference with the low option and the safe option; if this difference gets larger the cost of learning get larger, and it can be seen that the risky choice is decreasing. The opposite holds for the gains of learning: if the difference between the high option and the safe option gets larger, the gains of learning are increasing and the subject is more likely to explore. Correcting for the experienced variances shows that the findings seem to hold. The findings for the gains of learning hold when the results are split for subjects who never experienced the bad rare outcome. The findings for the cost of learning are no longer significant for subjects who never experienced the good rare outcome. Therefore it can be concluded that hypothesis 3 seems to be true, but is affected dependant on whether subjects experienced a type of rare outcome or not.

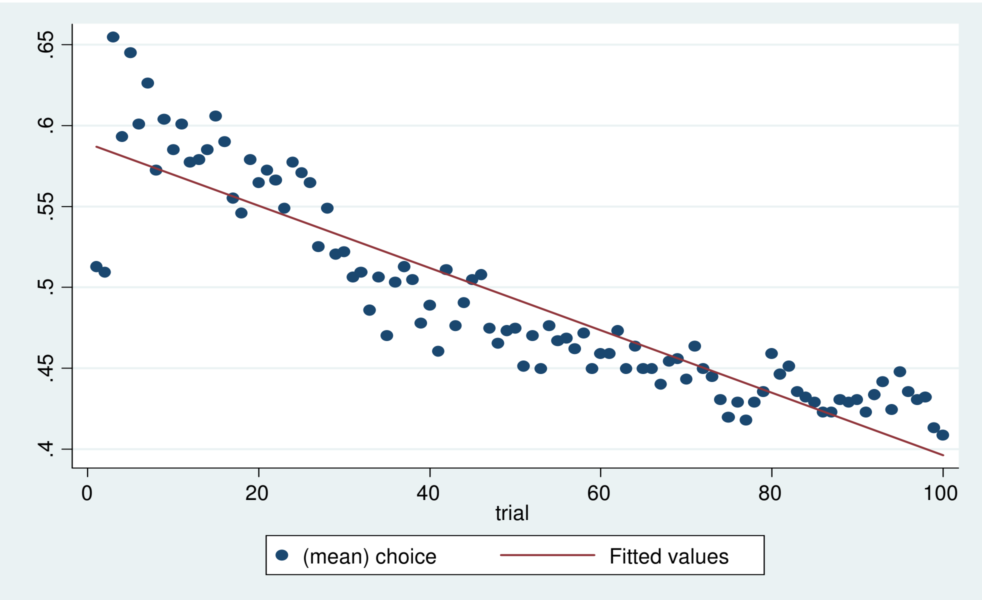
[10] Note: if we were to remove the variable for the difference not of interest (so gain for subjects who never experience the good rare outcome, and lose for subjects who never experience the bad rare outcome), the findings on the cost of learning do not change.

## Analysis of hypothesis 4 - gambler's fallacy & hot stove effect

The final main point of analysis deals with the gamblers fallacy and the hot stove effect. It is expected that wins and gains affect the settling down making a subject either loving or averse to the risky choice. This was depicted in the following hypothesis: "*The hot stove effect and the gamblers fallacy affect the exploration behavior, as reflected in the settling down behavior of the subjects. "*. The first step will be to see whether and where these biases are present in the data.

### Data on hypothesis 4

To look for the presence of the biases it is important to see what happens to the choice behavior directly after a shock. Hence the first step is to look at the difference in choice behavior for subject who experienced a bad or a good rare outcome and those who observed a good rare outcome. This is shown in Figure 16 and Figure 17. From the figures it can be seen that there is a decreasing average risky choice over trials for subjects who experienced a bad rare outcome. This indicates that there is some sort of hot stove effect, making them averse to the risky choice. The opposite is found for subjects who experienced a good rare outcome. The fitted curve tells us that there is an increase in the average choice over time. This means that subjects who experienced a good rare outcome are more likely to select the rare choice in the following trials.



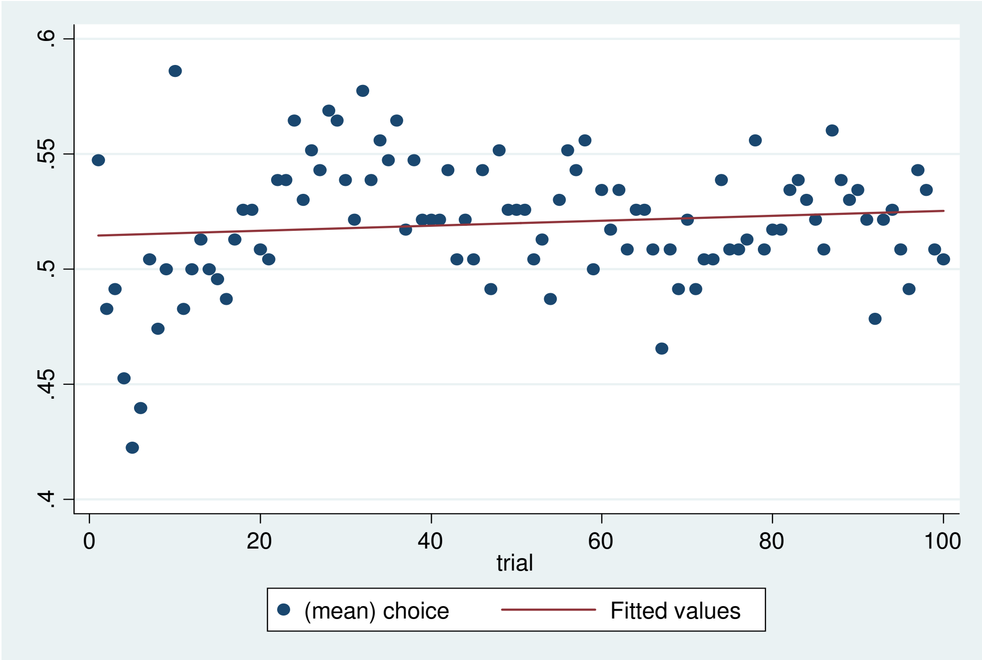



Figure 16: Mean choice for subjects with any bad rare outcome observed

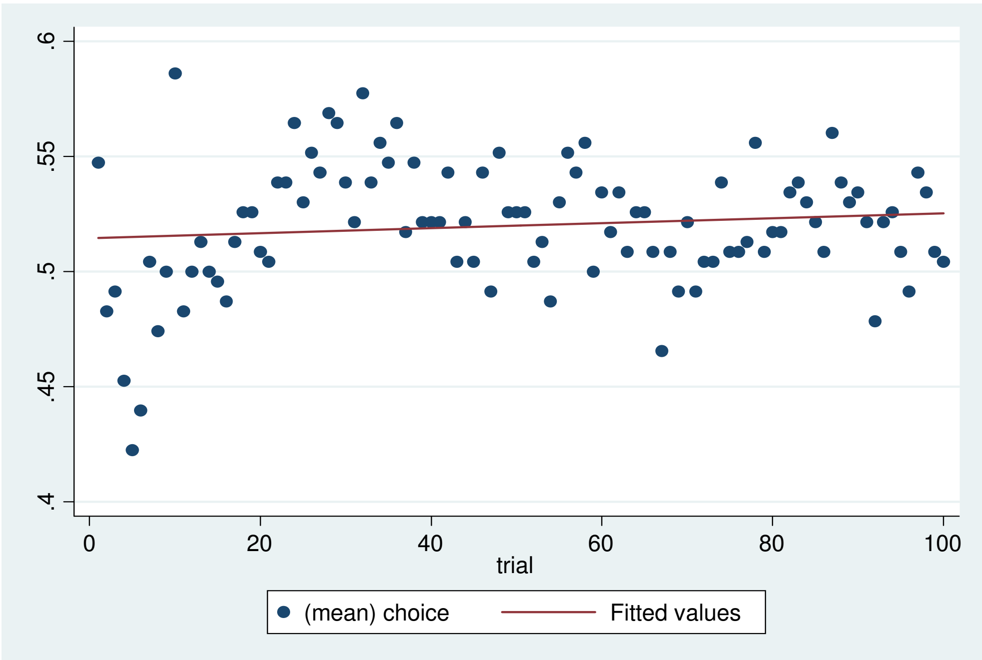



Figure 17: Mean choice for subjects with any good rare outcome observed

The downside from the above figures is that they do not tell us anything about when the rare outcome occurred. To see whether it matters when the rare outcome occurred, the mean choice is set out for subjects who experienced the rare outcome in the first 10, 20, 30, 40, or 50 trials. If the rare outcome occurred in the first 10 trials it is thus also included in the graph for the first 20 trials. Figure 18 shows the results for when the bad outcome was rare and occurred. It can be seen that there is a decrease in the risky choice for all the trials. This gives an indication that the hot stove effect is present in the dataset, subjects are moving away from the risky choice in all the cases. In all cases when the bad rare outcome occurred the risky choice is decreasing . An interesting finding is that this decrease seems to be the steepest for the first 20 and 30 trials. Combining this with the results for the good outcome, it seems that the trials between 10 and 30 seem to have the most effect on the choice. To see whether this is true, statistical tests need to be done.

In Figure 19 this is graphed for when the good outcome was rare. It can be seen that there is an increase in the average risky choice when the good rare outcome occurred in the first 30 or 40 trials. If it occurred in the first 10 trials, subjects seem to become averse to the risky choice. In the first 20 trials, there seems to be no change in the choice. Statistical testing needs to be done on this to see whether the differences are significant, as no conclusions regarding the timing of these effects can be drawn from these figures.

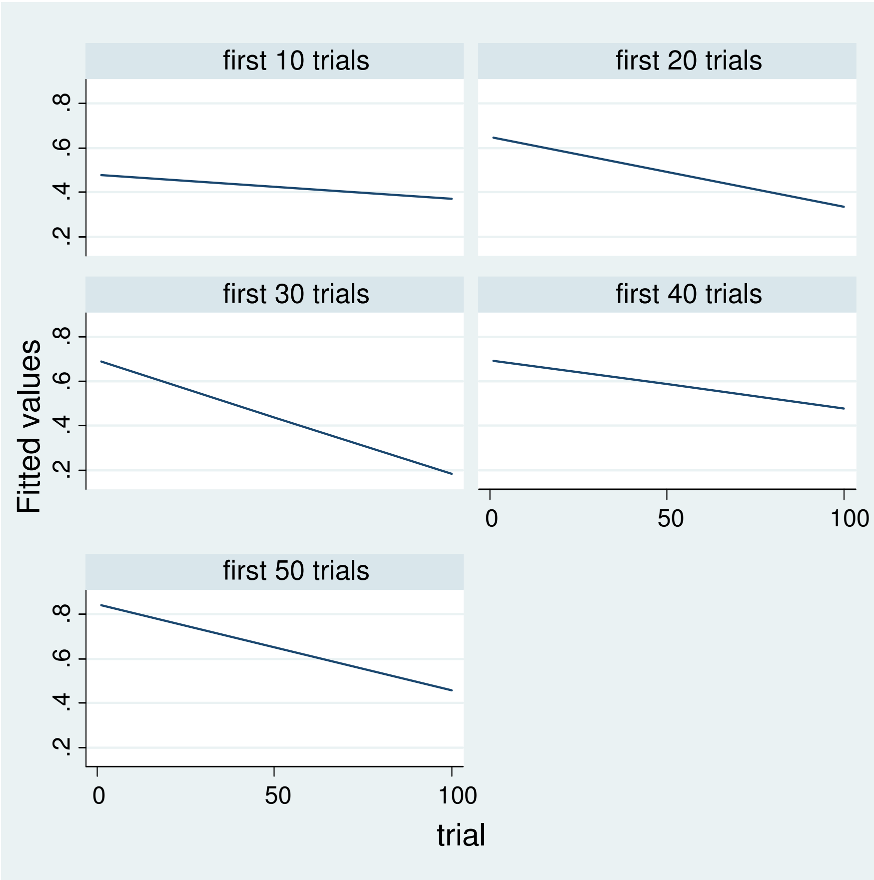



Figure 18: average choice dependant on when bad rare outcome occurred

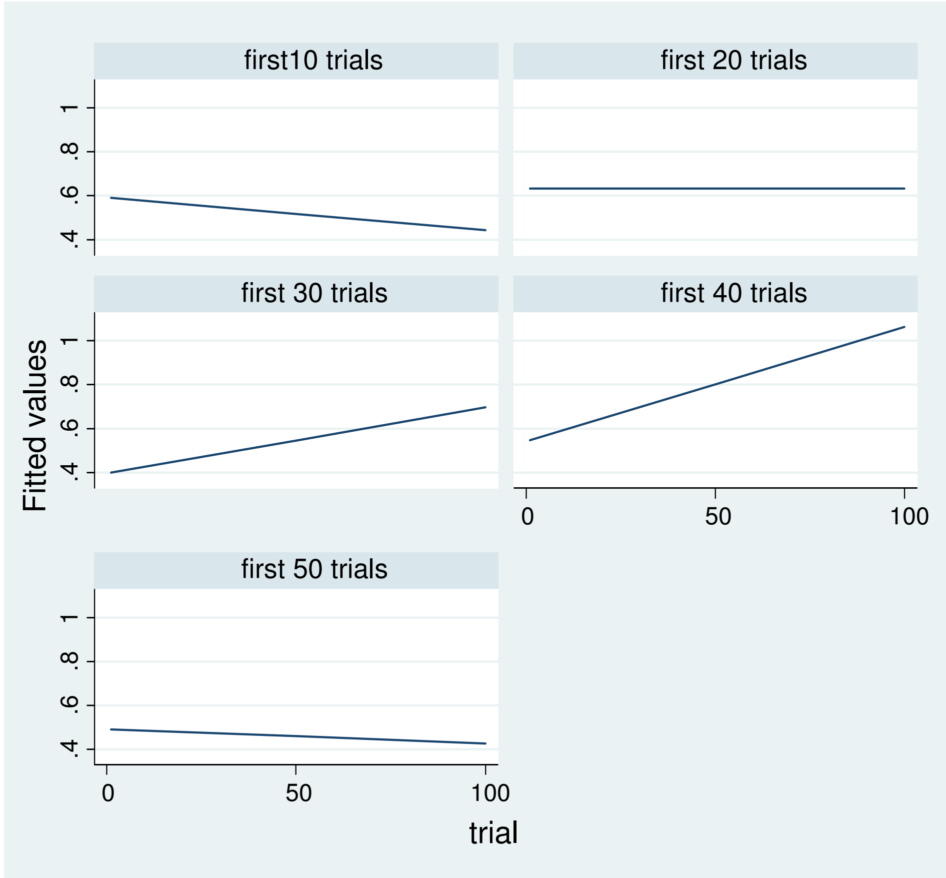



Figure 19: average choice dependant on when good rare outcome occurred

## Testing of hypothesis 4

To see how the choice behavior over time differs dependant on the rare outcomes, the difference in mean choice over time are tested for both the good and the bad rare outcome. This is done in two steps. The first step is to compare the difference in the proportion of risky choice at any point in time between subjects who experienced a good rare outcome and a bad rare outcome. The second step is to see how the choice behavior of subjects who experience a good or a bad rare outcome differs over time.

Over the same trial groups as in the previous analyses, the proportion of risky choice per ten trials has been tested trough a mean comparison test. The results for this are shown in Table 16. As can be seen there are significant differences for most of the cases between the subjects dependant on the type of rare outcome they observed. As expected, the subjects who experienced a good rare outcome select the risky choice more often compared to subjects who experienced a bad rare outcome. The exception to this lies within the first ten trials, where subjects who experience a bad rare outcome are more likely to pick the risky choice. After the initial ten trials, the difference in the trials 11-20 and 21-30 are not significant. This might be explained by the fact that the rare outcome did not occur yet in that phase of the model, and therefore no significant differences caused by the rare outcome are found. The fact that subjects who have experienced the good rare outcome are more likely to select the risky choice and subjects who experienced the bad rare outcome are less likely to select the risky findings is in line with the expectations, it indicates that the hot stove effect and the gamblers fallacy are present. But more importantly, it indicates that the fact that subjects settle down more easily after they have experienced a rare outcome does not seem to differ related to whether it was a good or a bad rare outcome. This because the difference in the mean risky choice seems to remain relatively stable over time, the difference between them does not seem to be increasing or decreasing after about 40 trials.

From this analysis it can be concluded that there are significant difference in choice behavior dependant on the good or the bad rare outcome, as we see that subjects who experience a good rare outcome are more likely to select the risky choice, opposite to subjects who experience the bad rare outcome. No conclusions with respect to the settling down behavior can be drawn based on this table.

| Trials | Over | Difference | Standard Deviation |
|---|---|---|---|
| 1-10 | good vs. bad | -0.100*** | 0.012 |
| 11-20 | good vs. bad | -0.006 | 0.024 |
| 21-30 | good vs. bad | 0.016 | 0.021 |
| 31-40 | good vs. bad | 0.042** | 0.018 |
| 41-50 | good vs. bad | 0.037** | 0.017 |
| 51-60 | good vs. bad | 0.089*** | 0.016 |
| 61-70 | good vs. bad | 0.061*** | 0.015 |
| 71-80 | good vs. bad | 0.048*** | 0.014 |
| 81-90 | good vs. bad | 0.086*** | 0.013 |
| 91-100 | good vs. bad | 0.085*** | 0.012 |

**Table 16: Difference in proportion of risky choice dependant on either good or bad rare outcome observed**

To see whether there are differences in the settling down behavior over time, the analysis in Table 8 is ran separately for the good and bad rare outcome. This will show whether there are differences in the size of the effects over time, dependant on the type of rare outcome. The results for this are shown in Table 17. From the table it can be seen that the differences over time seem to be small and not consistently significant for both the timing of the good and the bad rare outcome. The largest changes in the risky choice seems to occur in the first 50 trials. An interesting case is the switch between the trials 31-40 and 41-50, where there is a significant *decrease* of choice for both types of rare outcomes. It can be seen that after the first ten trials, subjects who experienced a good rare outcome significantly increase their proportion of risky choice, whereas subjects who experience a bad rare outcome significantly decrease their proportion of risky choice. This implies that the hot stove effect and the gamblers fallacy are both present. Interesting to note is that once the initial settling down occurred, the effects seems to disappear. There is little to no significant change in the proportion of risky choice after the first trials.

| Trials | Good Rare outcome Observed | | Bad Rare outcome Observed | |
|---|---|---|---|---|
| | *Difference* | *Standard Deviation* | *Difference* | *Standard Deviation* |
| 11-20 to 1-10 | 0.054** | 0.023 | -0.041*** | 0.015 |
| 21-30 to 11-20 | 0.026 | 0.027 | 0.005 | 0.017 |
| 31-40 to 21-30 | -0.073*** | 0.024 | -0.099*** | 0.015 |
| 41-50 to 31-40 | -0.003 | 0.022 | 0.001 | 0.013 |
| 51-60 to 41-50 | 0.047** | 0.021 | -0.005 | 0.012 |
| 61-70 to 51-60 | -0.019 | 0.019 | 0.010 | 0.011 |
| 71-80 to 61-70 | -0.017 | 0.018 | -0.005 | 0.010 |
| 81-90 to 71-80 | 0.026 | 0.017 | -0.012 | 0.010 |
| 91-100 to 81-90 | -0.017 | 0.015 | -0.015* | 0.009 |

**Table 17: Difference in choice behavior over time dependant on type of rare outcome**

To investigate this table in more detail it is important to test when the rare outcome occurred, which will help to conclude whether the timing of the rare outcome matters. To test the effect of when the rare outcome occurred, a pair wise comparison of the risky choice dependant on when the rare outcome occurred will be used. This means that the differences in choice behavior over time for all the different moments will be tested against each other. The results for this are shown in Table 18. This table thus shows the differences in the mean risky choice dependant on when the rare outcome occurred. The differences are between the left column and the top row.[11] From the table it can be concluded that the differences dependant on when the rare outcome occurred is significant in all but two cases. Furthermore the lowest average choice is for subjects who experienced the rare outcome only in the last ten trials, followed by the subjects who experienced it in the first ten trials. The highest risky choice is for the subjects who experienced the risky choice in the trials 51-60, followed by the subjects who experienced the rare outcome in trials 61-70 and 71-80.

The main finding here is that in general subjects who experience the rare outcome in a later part of the trials are more eligible to keep selecting the risky choice. This means that subjects stay with the risky choice, but move towards the safe option when a rare outcome occurs. The fact that subjects who experienced the rare outcome later settle down later is caused by the timing of the rare outcome. The exception to this are subjects who experience the rare outcome in the final ten trials. This might be explained by differences in risk aversion, some only select the risky subject in the final part of the trials.

[11] How to read this table: average risky choice is 0.082 higher when rare outcome occurred in the trials 11-20 compared to when it occurred in trials 1-10, significant at 1%. Reversed differences hold for the white parts of the table, because difference between the groups 1-10 with 11-20 is the negative difference of 11-20 with 1-10

| Difference with mean choice | 1 to 10 | 11 to 20 | 21 to 30 | 31 to 40 | 41 to 50 | 51 to 60 | 61 to 70 | 71 to 80 | 81 to 90 | 91 to 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 to 10 | xxx | | | | | | | | | |
| *(standard deviation)* | *(xxx)* | | | | | | | | | |
| 11 to 20 | 0.082*** | xxx | | | | | | | | |
| *(standard deviation)* | *(0.008)* | *(xxx)* | | | | | | | | |
| 21 to 30 | 0.010 | -0.072*** | xxx | | | | | | | |
| *(standard deviation)* | *(0.008)* | *(0.011)* | *(xxx)* | | | | | | | |
| 31 to 40 | 0.187*** | 0.104*** | 0.177*** | xxx | | | | | | |
| *(standard deviation)* | *(0.008)* | *(0.011)* | *(0.011)* | *(xxx)* | | | | | | |
| 41 to 50 | 0.113*** | 0.104** | 0.103**** | -0.073*** | xxx | | | | | |
| *(standard deviation)* | *(0.010)* | *(0.011)* | *(0.013)* | *(0.012)* | *(xxx)* | | | | | |
| 51 to 60 | 0.288*** | 0.206*** | 0.278*** | 0.101*** | 0.175 | xxx | | | | |
| *(standard deviation)* | *(0.008)* | *(0.011)* | *(0.011)* | *(0.011)* | *(0.013)* | *(xxx)* | | | | |
| 61 to 70 | 0.240*** | 0.157*** | 0.230*** | 0.000*** | 0.126 | -0.048*** | xxx | | | |
| *(standard deviation)* | *(0.008)* | *(0.011)* | *(0.012)* | *(0.000)* | *(0.013)* | *(0.011)* | *(xxx)* | | | |
| 71 to 80 | 0.223*** | 0.141*** | 0.213*** | 0.036*** | 0.110*** | -0.065*** | -0.017 | xxx | | |
| *(standard deviation)* | *(0.008)* | *(0.011)* | *(0.011)* | *(0.011)* | *(0.013)* | *(0.011)* | *(0.011)* | *(xxx)* | | |
| 81 to 90 | 0.145*** | 0.062*** | 0.134*** | -0.042*** | 0.031** | -0.144*** | -0.095*** | -0.078*** | xxx | |
| *(standard deviation)* | *(0.009)* | *(0.011)* | *(0.012)* | *(0.011)* | *(0.013)* | *(0.012)* | *(0.012)* | *(0.011)* | *(xxx)* | |
| 91 to 100 | -0.214*** | -0.297*** | -0.224*** | -0.401*** | -0.327*** | -0.502*** | -0.454*** | -0.437*** | -0.359*** | xxx |
| *(standard deviation)* | *(0.003)* | *(0.008)* | *(0.008)* | *(0.008)* | *(0.010)* | *(0.008)* | *(0.008)* | *(0.008)* | *(0.009)* | *(xxx)* |

**Table 18: differences in mean choice dependant on when rare outcome occurred**

Since it is likely that subjects behave different dependant on the type of rare outcome, these results will be repeated for good and bad rare outcomes.  To show the findings for the different types of rare outcomes, the same analysis as in Table 18 will be used, but split out for subjects who experienced a good or a bad rare outcome.

The outcome for the bad rare outcome is shown in Table 19 to conclude on the presence of the hot stove effect. From the table it can be seen that the final ten trials have the lowest mean choice. For the rest of the observations there seems to be a clear trend in the proportion of risky choice. If the bad outcome occurred in the first part of the trials, the proportion of risky choice is lower compared to when it occurred in the final part of the trials. This implies that subjects who experience the bad outcome in the first part become averse to the risky choice, and are more likely to select the safe option. The proportion of risky choice is the highest for subjects who experience the rare outcome in the trials 61-70, followed by 51-60 and 71-80. This leads to the conclusion that subjects who experience a bad rare outcome in the first part of the trials will pick the risky option less often than subjects who experience the bad rare outcome in the final part of the trials. Relating this to the hot stove effect, it can be concluded that there is clear settling down behavior visible here: Subjects who experience the bad outcome early are more likely to pick the safe option opposite to subjects who experience it later on. This means that the hot stove effect affects settling down behavior, dependant on when the hot stove effect occurred.

This analysis is repeated for when the rare good outcome occurred to conclude on the gambler's fallacy. In Table 20 the results for the mean choice comparison of when the good rare outcome occurred are shown, which seem to be significant for most of the cases. The same finding as the previous tables seems to hold: the proportion of risky choice is the lowest for subjects who experienced the rare outcome in the final ten trials. This is followed by the trials 41-50 and 61-70. The subjects who experienced the good outcome in the trials 31-40 have the highest average risky choice.. This is followed by the trials 51-60 and the trials 11-20. There seems to be no clear order of the average risky choice here, which implies that *when* the good rare outcome occurred does *not* affect the risky choice. Remember that it *does* matter *whether* the good rare outcome occurred. Related to the gambler's fallacy it can be concluded that it does not seem to matter when a good rare outcome occurred, as the proportion of risky choice is not significantly higher for subjects who experienced the good rare outcome early on. This means that the settling down behavior (on the risky choice) does not seem to be affected by the occurrence of the positive shock.

| Difference with mean choice | 1 to 10 | 11 to 20 | 21 to 30 | 31 to 40 | 41 to 50 | 51 to 60 | 61 to 70 | 71 to 80 | 81 to 90 | 91 to 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 to 10 | xxx | | | | | | | | | |
| *(standard deviation)* | *xxx* | | | | | | | | | |
| 11 to 20 | 0.067*** | xxx | | | | | | | | |
| *(standard deviation)* | *(0.010)* | *xxx* | | | | | | | | |
| 21 to 30 | 0.010 | -0.057*** | xxx | | | | | | | |
| *(standard deviation)* | *(0.010)* | *(0.013)* | *xxx* | | | | | | | |
| 31 to 40 | 0.160*** | 0.093*** | 0.150*** | xxx | | | | | | |
| *(standard deviation)* | *(0.009)* | *(0.013)* | *(0.013)* | *xxx* | | | | | | |
| 41 to 50 | 0.223*** | 0.157*** | 0.213*** | 0.064*** | xxx | | | | | |
| *(standard deviation)* | *(0.014)* | *(0.017)* | *(0.017)* | *(0.016)* | *xxx* | | | | | |
| 51 to 60 | 0.327*** | 0.261*** | 0.317*** | 0.168*** | 0.104*** | xxx | | | | |
| *(standard deviation)* | *(0.009)* | *(0.013)* | *(0.013)* | *(0.012)* | *(0.016)* | *xxx* | | | | |
| 61 to 70 | 0.336*** | 0.270*** | 0.326*** | 0.176*** | 0.112*** | 0.008 | xxx | | | |
| *(standard deviation)* | *(0.010)* | *(0.014)* | *(0.014)* | *(0.013)* | *(0.017)* | *(0.013)* | *xxx* | | | |
| 71 to 80 | 0.300*** | 0.233*** | 0.290*** | 0.140*** | 0.076*** | -0.028 | -0.036*** | xxx | | |
| *(standard deviation)* | *(0.010)* | *(0.013)* | *(0.013)* | *(0.013)* | *(0.017)* | *(0.013)* | *(0.014)* | *xxx* | | |
| 81 to 90 | 0.238*** | 0.171*** | 0.228*** | 0.078*** | 0.01433 | -0.090*** | -0.098*** | -0.062*** | xxx | |
| *(standard deviation)* | *(0.013)* | *(0.015)* | *(0.015)* | *(0.015)* | *(0.018)* | *(0.015)* | *(0.016)* | *(0.016)* | *xxx* | |
| 91 to 100 | -0.106*** | -0.172*** | -0.116*** | -0.265*** | -0.329*** | -0.433*** | -0.441*** | -0.405*** | -0.343*** | xxx |
| *(standard deviation)* | *(0.003)* | *(0.009)* | *(0.009)* | *(0.009)* | *(0.014)* | *(0.009)* | *(0.010)* | *(0.010)* | *(0.012)* | *xxx* |

**Table 19: mean choice when bad rare occurred**

| Difference with mean choice | 1 to 10 | 11 to 20 | 21 to 30 | 31 to 40 | 41 to 50 | 51 to 60 | 61 to 70 | 71 to 80 | 81 to 90 | 91 to 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 to 10 | xxx | | | | | | | | | |
| *(standard deviation)* | *xxx* | | | | | | | | | |
| 11 to 20 | 0.118*** | xxx | | | | | | | | |
| *(standard deviation)* | *(0.016)* | *xxx* | | | | | | | | |
| 21 to 30 | 0.031* | -0.086**** | xxx | | | | | | | |
| *(standard deviation)* | *(0.019)* | *(0.024)* | *xxx* | | | | | | | |
| 31 to 40 | 0.288*** | 0.171*** | 0.257*** | xxx | | | | | | |
| *(standard deviation)* | *(0.017)* | *(0.022)* | *(0.024)* | *xxx* | | | | | | |
| 41 to 50 | -0.058*** | -0.176*** | -0.090*** | -0.347*** | xxx | | | | | |
| *(standard deviation)* | *(0.016)* | *(0.022)* | *(0.024)* | *(0.022)* | *xxx* | | | | | |
| 51 to 60 | 0.131*** | 0.013 | 0.099*** | -0.158*** | 0.189*** | xxx | | | | |
| *(standard deviation)* | *(0.022)* | *(0.027)* | *(0.028)* | *(0.027)* | *(0.027)* | *xxx* | | | | |
| 61 to 70 | -0.013 | -0.131*** | -0.044 | -0.301*** | 0.045 | -0.144*** | xxx | | | |
| *(standard deviation)* | *(0.017)* | *(0.022)* | *(0.024)* | *(0.023)* | *(0.022)* | *(0.027)* | *xxx* | | | |
| 71 to 80 | 0.041*** | -0.078*** | 0.009 | -0.248*** | 0.098*** | -0.091*** | 0.0529** | xxx | | |
| *(standard deviation)* | *(0.015)* | *(0.021)* | *(0.023)* | *(0.022)* | *(0.021)* | *(0.026)* | *(0.022)* | *xxx* | | |
| 81 to 90 | 0.012 | -0.106*** | -0.020 | -0.277*** | 0.070*** | -0.119*** | 0.024 | -0.029 | xxx | |
| *(standard deviation)* | *(0.013)* | *(0.020)* | *(0.022)* | *(0.020)* | *(0.020)* | *(0.025)* | *(0.020)* | *(0.019)* | *xxx* | |
| 91 to 100 | -0.149*** | -0.267*** | -0.181*** | -0.438*** | -0.091*** | -0.280*** | -0.137*** | -0.189*** | -0.161*** | xxx |
| *(standard deviation)* | *(0.005)* | *(0.015)* | *(0.018)* | *(0.016)* | *(0.015)* | *(0.022)* | *(0.016)* | *(0.015)* | *(0.012)* | *xxx* |

**Table 20: mean choice when good rare outcome occurred**

## Conclusions on hypothesis 4

The findings on hypothesis four indicate there are no clear significant pattern differences in the proportion of risky choice between subjects who did and who did not experience a good or a bad rare outcome. Subjects who experience a rare outcome in the final ten trials have the lowest proportion of risky choice for both the types of rare outcomes. What can be concluded with respect to the settling down is that the settling down behavior is clearly present over the timing of the rare bad outcomes, but not so much over the timing of the rare good outcomes.

These findings tell us that the hot stove effect seems to be present here when we combine this finding with the previous findings. This because from the analysis of hypothesis two it became clear that subjects who experience the bad rare outcome are less likely to select the risky choice. Opposite to this we know that subjects who experience the good rare outcome are more likely to select the risky choice. When the rare outcome occurred does not seem to matter for the good outcome, but does matter for the bad rare outcome. This is in line with the findings from the experienced variance from hypothesis two. Therefore this lead us to conclude that both the biases are present in the choice behavior of the subjects, but the settling down behavior seems only relevant for the subjects who experienced the bad rare outcome. For these subjects the settling down behavior is stronger for subjects who experience the bad rare outcome early on. This means that hypothesis four is partly true: the biases are present but the settling down behavior is only present in the first ten trials. There is stronger evidence for the hot stove effect over the gamblers fallacy.

## Analysis of hypothesis 5: the type of problem?

The final hypothesis of interest states that the exploration is also affected by the type of problem a subject is facing, as there might be differences in behavior when a subject is either facing losses or gains. This was reflected in the hypothesis: "*the choice behavior of subjects differs when facing losses compared to gains*". This hypothesis will be tested and analyzed in this chapter. The first step is to look at the difference in choice behavior in the data.

### Data related to hypothesis 5

There are three types of problems a subject can be facing: losses only, gains only, or over a combination of these two. The used problems per session are shown in Table 39 in the appendix. In all sessions, subjects faced all three types of problems: over gains (low and high and safe are all positive), over losses (low and high and safe are all negative) and over mixed (high is positive, low is negative, and safe is either negative or positive).

To see how the choice behavior between these three type of problems differs, the proportion of risky choice over time will be plotted over the different types of problems .These are shown respectively in Figure 20 (gains) Figure 21 (losses) Figure 22 (mixed). As can be seen from these figures, there is settling down on the safe choice in all the types of problems. The proportion of risky choice is the highest for subjects facing gains, and about equal between subjects facing losses and mixed problems. From the graphs it is not clear whether there are differences in searching behavior over time dependant on the type of problems. What can be seen is that there is a drop in the proportion of risky choice in the first ten trials for all type of problems.

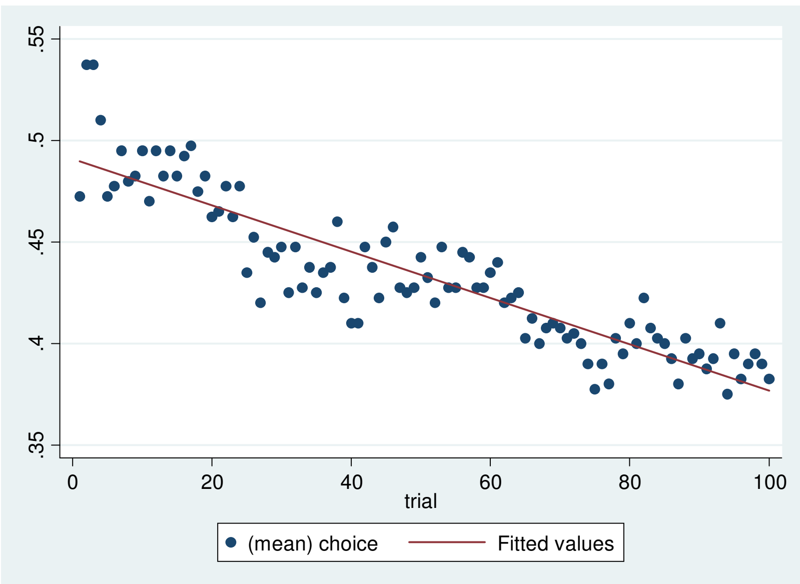



Figure 20: proportion of risky choice over time facing gains

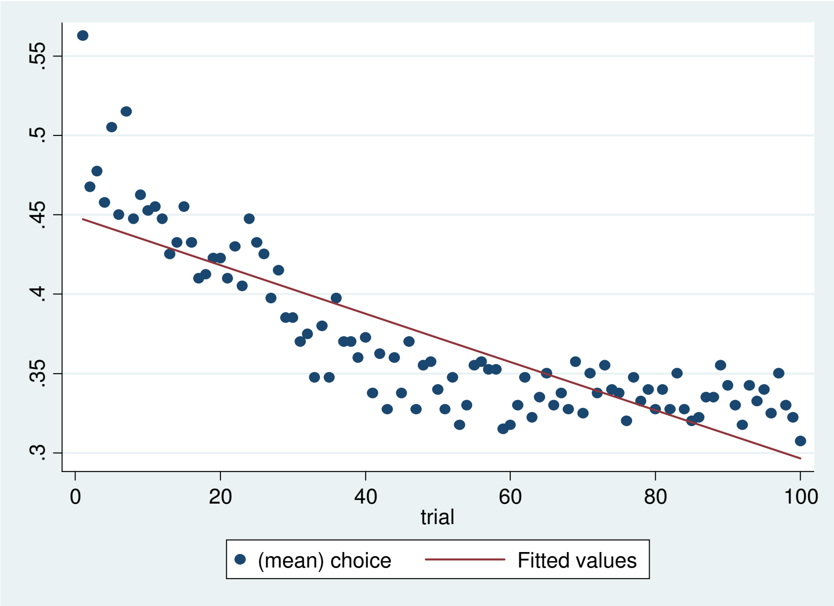



Figure 21: proportion of risky choice over time facing losses

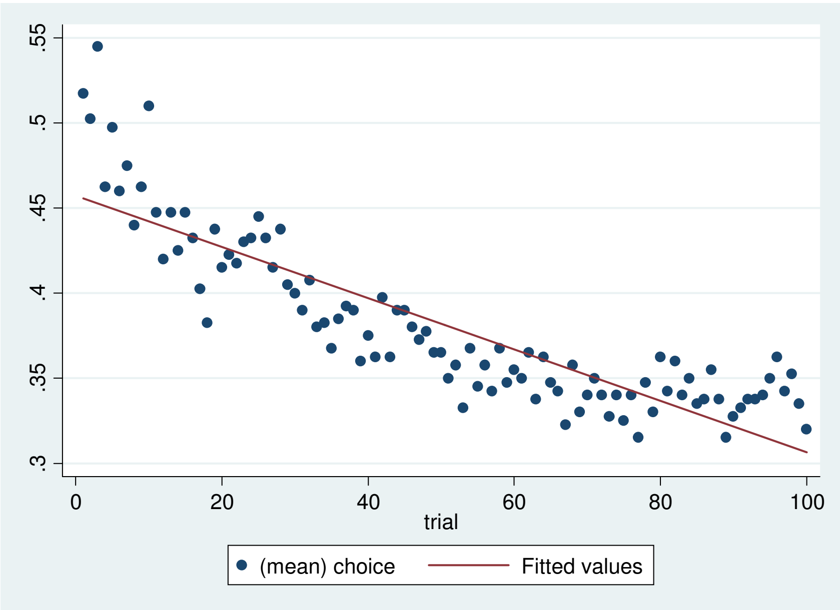



Figure 22: proportion of risky choice over time facing mixed problems

## Testing hypothesis 5

To test whether and how the type of problem affects the risky choice, dummies for the type of problems can be included in a panel logit model with the risky choice as the dependant variable. Since the three types are mutually exclusive, the dummy variable for mixed problems is used as a reference category. The results for the restricted model are shown in Table 21, where the dummy variables for problems over gains and problems over losses are regressed on the risky choice.

From this table it can be seen that subjects are more likely to select the risky choice when facing gains compared to facing a mixed problem, as there is a positive relation between the dummy variable for gain problems and the choice. This is significant at  a 5% significance level and implies that subjects are risk seeking over gains. Another interesting implication is that subjects seem to treat mixed problems the same loss problems. This because the coefficient for the dummy variable gains is almost equal to zero, implying that it should be equal to the reference category of the mixed problems. It should be noted that this is not  significant in this specification. These findings confirm the initial findings from the graphs.

| Dependant variable: Choice | Coefficient | Standard Deviation |
|---|---|---|
| lossproblem | -0.001 | 0.159 |
| gainproblem | 0.359** | 0.159 |
| _cons | -0.835*** | 0.113 |

Table 21: Restricted model on effect of type of problems on choice

In the above model, some noise will be present. Therefore the usual control variables are added, including the payoffs, probabilities, trial, and order of the problems. Again, mixed problems are selected as the reference category. The results for this model are shown in Table 22. The same relations seem to hold for the problem dummy's, subjects are more likely to select the risky choice when facing a problem over gains compared to a mixed problem. The loss problems remains insignificant and about equal to the behavior in mixed problem. The rest of the relations and significances seem to be the same as in the model from Table 5.

| Dependant Variable Choice | Coefficient | Standard Deviation |
|---|---|---|
| gainproblem | 0.521** | 0.224 |
| lossproblem | 0.001 | 0.222 |
| phigh | 3.243*** | 0.403 |
| high | 0.053** | 0.024 |
| low | 0.045* | 0.026 |
| medium | -0.165*** | 0.036 |
| trial | -0.011*** | 0.000 |
| order | 0.005 | 0.019 |
| lagged payoff | 0.061*** | 0.004 |
| _cons | -2.224*** | 0.336 |

Table 22: Unrestricted model on effect of type of problems on choice

To see whether the probability has a different effect on the risky choice in the different type of problems, interaction terms are created. These should show whether subjects take more or less risk dependant on *both* the type of problem and the probabilities. The dummy variables for the type of problems are thus linked with the probability of the high outcome. The results for this are shown in Table 23. The interaction effect shows how the choice between the different type of problems differs in probability. Again, the mixed problems remain the reference category. The regression model shows that if the probability of the high outcome is close to 1 (hence very likely), the subjects are more likely to select the risky choice in a mixed problem than in a problem over gains. To see this, one can see the following reasoning: if the high outcome has a probability of 95% of happening, this means that the difference in proportion of risky choice between gain problems and mixed problems is 0.95 * -0.86 = -0.817, ceteris paribus. Opposite to this, if the probability of the high outcome is 5%, the difference in the proportion of risky choice between gain problems and mixed problems becomes 0.05 * -0.86 = -0.043, ceteris paribus. This means that the difference in the proportion of risky choice between gain and mixed problems increases when the likelihood of the high outcome increases. The reverse holds with respect to the probability of the low outcome. This relation is significant at a 5 % significance level.

Looking at the difference between mixed problems and problems over losses, it can be seen that there is a positive relation. This implies that when the “high” outcome in a losses problem (which is still a loss, but a smaller loss than the safe option) has a large probability of happening (95%), the difference in risky choice between loss problems and mixed problems becomes 0.95 * 0.70 = 0.665, ceteris paribus. Opposite to this, if the high outcome has a 5% chance of happening, the difference becomes 0.05 * 0.70 = 0.035, ceteris paribus. This means that the difference in risky choice between losses and mixed problems increases in line with the probability for the high outcome. This relation is significant at a 1% significance level.

These two findings let us to believe that there exist a relation between the type of problem and the risky choice, and the magnitude of this difference is dependent on the probability of the options. This difference exist between the mixed problems against either a loss or a gain problem. The difference between loss and gain problems has also been tested, but this does not result in different findings and is therefore not reported. The same holds for when linked with the probabilities.

Combining these findings thus leads to the conclusion that there are significant differences in choice behavior dependant on the type of problem facing. Subjects seem to be more willing to select the risky choice if they are facing gains, and seem indifferent between problems over losses and mixed payoffs. This implies that subjects are risk seeking over gains and risk averse over losses. However, when the findings are combined with the probabilities, it can be seen that subjects facing gains take more risk when the probabilities are high, and less risk when the probabilities are small. This under the assumption that selecting the risky choice is seen as risky. For losses, it can be seen that subjects take more risk when the probabilities are high, and less risk when the probabilities are small. This means that the findings based on risk seeking over gains and risk averse over losses do not seem to hold.

| Dependant Variable Choice | Coefficient | Standard Deviation |
|---|---|---|
| gainproblem * phigh | -0.866*** | 0.406 |
| lossproblem * phigh | -1.22** | 0.424 |
| | | |
| gainproblem | 0.985** | 0.335 |
| lossproblem | 0.695*** | 0.330 |
| | | |
| phigh | 4.175*** | 0.510 |
| high | 0.064** | 0.025 |
| low | 0.065** | 0.027 |
| medium | -0.191*** | 0.036 |
| trial | -0.011*** | 0.000 |
| order | 0.005 | 0.019 |
| lagged payoff | 0.061*** | 0.004 |
| _cons | -2.670*** | 0.377 |

Table 23: effect of type of problems on choice with interaction effects

## Conclusions on hypothesis 5

In this chapter the difference in choice behavior regarding the type of problem have been analyzed. It can be concluded that the subjects seem to treat problems over losses and over mixed payoffs equal. There seem to be slight differences in choice behavior dependant on the type of problems faced. Subjects are more likely to pick the risky choice in gains problems compared to subjects facing a mixed or a loss problem. If this is combined with the probabilities of the problems it can be seen that subjects are more likely to pick a risky choice over gains and losses compared to a mixed problem when this is very likely. This is interesting, because it implies that subjects always seem to take more risk over high probabilities, which implies that there might be overweighting of large probabilities for the best payoff dependant on the type of problem faced. The difference between the type of problems disappear when the subjects are facing small probabilities. This implies that the type of problem does not matter for small probabilities, but does matter for large probabilities. Naturally, the reverse holds for the worst payoff a subject is facing. In general it can be concluded that hypothesis 5 holds, subjects are more likely to select the risky choice when facing gains, but subjects treat a mixed problem equal to a loss problem.

## Analysis with competition dataset

In the main analysis we made use of the estimation dataset from Erev et al (2010). Next to this they also created the competition dataset. This covers the same amount of observations; 100 students in 5 sessions of 20 students, with 12 problems and 100 trials per student, yielding another 120000 observations. To test the robustness of the previous findings, the main findings can thus be checked by using the other dataset. For each hypothesis and check, the main findings will be checked against this dataset. Note that the same transformations have been used on the competition dataset as on the estimation dataset.

### Robustness data description

The first step will be to see whether the competition and the estimation data set are actually comparable. To see whether this is the case, summary statistics are given. Table 24 shows both the robustness (competition) and the original (estimation) summary statistics. The main difference in the prospects seems to lie in the safe option. This is on average negative in the robustness dataset, whereas it was positive in the estimation dataset. This is reflected in the achieved payoff of the subject, which is now also negative. Another difference lies in the probabilities: the high outcome was on average less likely than the low outcome in the robustness dataset, whereas the reverse holds in our original estimation set.

Looking at the difference in choice behavior based on the available subjects, it can be seen that the risky choice seems to be about the same. In the robustness check, the subjects picked the risky option in 38.10% of all the cases, slightly lower than the original 39.54%. Looking at the occurrence of the rare outcome it can be seen that this occurred in 4.04% of all cases, which translates to 10.63% of the risky choice resulting in the rare outcome. Again, this is slightly lower than the original 10.84%. Another interesting difference lies in the type of rare outcome which occurred: the rare high outcome occurred more often than the rare low outcome in the robustness check, opposite to the outcomes of the original analysis. It will be interesting to see whether and how this affected the subjects' choice behavior.

| Variable | Observations | Mean | Standard Error | Min | Max | | Mean | Standard I | Min | Max |
|---|---|---|---|---|---|---|---|---|---|---|
| *choice* | 120000 | 0.3810 | 0.4856 | 0 | 1 | | 0.3954 | 0.4889 | 0 | 1 |
| *risky* | 120000 | 0.3810 | 0.4856 | 0 | 1 | | 0.3954 | 0.4889 | 0 | 1 |
| *safe* | 120000 | 0.6190 | 0.4856 | 0 | 1 | | 0.6046 | 0.4889 | 0 | 1 |
| *high (risky)* | 120000 | 5.1783 | 9.4260 | -9.7 | 27.5 | | 5.2317 | 9.3011 | -10 | 26.5 |
| *low (risky)* | 120000 | -5.2183 | 9.4623 | -26.5 | 9.2 | | -5.0767 | 9.4692 | -29.2 | 9.7 |
| *medium (safe)* | 120000 | -0.2417 | 11.2293 | -26.3 | 22.1 | | 0.2933 | 10.7588 | -25.6 | 25.2 |
| *phigh* | 120000 | 0.4852 | 0.3925 | 0.02 | 0.99 | | 0.5522 | 0.3953 | 0.01 | 0.99 |
| *plow* | 120000 | 0.5148 | 0.3925 | 0.01 | 0.98 | | 0.4478 | 0.3953 | 0.01 | 0.99 |
| *achieved payoff* | 120000 | -0.1188 | 11.3993 | -26.5 | 27.5 | | 0.3775 | 10.8851 | -29.2 | 26.5 |
| *expected payoff* | 120000 | -0.2408 | 11.1936 | -24.71 | 22.294 | | 0.2902 | 10.6472 | -25.36 | 25.408 |
| *low_rare* | 120000 | 0.4500 | 0.4975 | 0 | 1 | | 0.5667 | 0.4955 | 0 | 1 |
| *high_rare* | 120000 | 0.5000 | 0.5000 | 0 | 1 | | 0.4167 | 0.4930 | 0 | 1 |
| *gain* | 120000 | 5.4200 | 5.2208 | -0.1 | 17.7 | | 4.9383 | 5.3879 | 0 | 16.7 |
| *lose* | 120000 | 4.9767 | 4.9751 | -0.3 | 16.9 | | 5.3700 | 4.5859 | -0.4 | 16.9 |
| *rare outcome occurred* | 120000 | 0.0404 | 0.1969 | 0 | 1 | | 0.0429 | 0.2027 | 0 | 1 |
| *rare high outcome occurred* | 120000 | 0.0221 | 0.1471 | 0 | 1 | | 0.0139 | 0.1169 | 0 | 1 |
| *rare low outcome occurred* | 120000 | 0.0183 | 0.1340 | 0 | 1 | | 0.0291 | 0.1680 | 0 | 1 |
| *Dataset:* | | Competiton (robustness) | | | | | Estimation (main analysis) | | | |

Table 24: summary statistics robustness check

## Robustness check hypothesis 1

The first step in testing the robustness of hypothesis one will be to see whether there is a significant difference between the occurrence of the rare outcomes and the timing of the rare outcomes. Table 25 shows the results for this over the robustness dataset. As can be seen the same relations hold as before: the proportion of risky choice is significantly higher for subjects who did experience the rare outcome. The only difference lies in the height of the differences, as these seem larger than the difference in the original analysis. This might be explained by the mean negative payoff in the robustness dataset. This means that subjects might make their choice whilst being more risk averse in this dataset because they are facing losses.

| Trials | Over | Difference | Standard Deviation |
|---|---|---|---|
| 1-10 | rare vs. no rare | 0.143*** | 0.008 |
| 11-20 | rare vs. no rare | 0.319*** | 0.021 |
| 21-30 | rare vs. no rare | 0.365*** | 0.017 |
| 31-40 | rare vs. no rare | 0.403*** | 0.016 |
| 41-50 | rare vs. no rare | 0.541*** | 0.014 |
| 51-60 | rare vs. no rare | 0.545*** | 0.012 |
| 61-70 | rare vs. no rare | 0.606*** | 0.011 |
| 71-80 | rare vs. no rare | 0.603*** | 0.011 |
| 81-90 | rare vs. no rare | 0.575*** | 0.010 |
| 91-100 | rare vs. no rare | 0.598*** | 0.010 |

**Table 25: Robustness check difference rare outcome observed vs. Not observed**

To give more insight on the robustness test in this, it is also useful to look at choice behavior over time in the robustness dataset. This means that the proportion of risky choice will be tested over time separately for subjects who did and who did not experience the rare outcome. This is shown in Table 26. The risky choice seems to be increasing over time for subjects who experience the rare outcome, similar to the main analysis. The negative relation for subjects who did not experience a rare outcome is found here as well. The results for this are more often significant in the robustness check compared to the original analysis. This means that the robustness check on this part confirms the original findings in this.

With respect to hypothesis 1 it can be concluded that the robustness check confirms the findings from the original analysis. The proportion of risky choice is found to be decreasing over time for subjects who did not experience the rare outcome, opposite to increasing for subjects who did experience the rare outcome.  This settling down is confirmed to differ dependant on the type of rare outcome. Bad rare outcome results in settling down on the safe choice, whereas the subjects who experience the good rare outcome settle down on the risky choice.

| Trials | Over | Difference | Standard Deviation |
|---|---|---|---|
| 1 to 2 | rare observed | 0.031 | 0.020 |
| | rare not observed | -0.145*** | 0.011 |
| 2 to 3 | rare observed | 0.013 | 0.024 |
| | rare not observed | -0.034*** | 0.013 |
| 3 to 4 | rare observed | 0.075*** | 0.021 |
| | rare not observed | 0.036*** | 0.011 |
| 4 to 5 | rare observed | 0.105*** | 0.019 |
| | rare not observed | -0.033*** | 0.009 |
| 5 to 6 | rare observed | -0.007 | 0.016 |
| | rare not observed | -0.011 | 0.009 |
| 6 to 7 | rare observed | 0.043*** | 0.015 |
| | rare not observed | -0.019** | 0.008 |
| 7 to 8 | rare observed | -0.013 | 0.014 |
| | rare not observed | -0.009 | 0.007 |
| 8 to 9 | rare observed | -0.011 | 0.013 |
| | rare not observed | 0.016** | 0.007 |
| 9 to 10 | rare observed | 0.025** | 0.013 |
| | rare not observed | 0.002 | 0.006 |

**Table 26: Robustness check mean choice over time rare vs. not rare**

## Robustness check hypothesis 2

In testing the robustness of hypothesis 2 the occurrence of a rare outcome in the previous trial is of interest. The analysis of lags is thus repeated here. In Table 27 the same logit model as in the original analysis has been estimated. From this it can be seen that the same relation is found for the lagged rare high outcome. The opposite is found for the lagged bad outcome. This is now significantly negative. In the original analysis the first lag was found to be insignificant, regardless of the amount of lags. This implies that subjects immediately respond to a negative shock. To see whether the first lag remains significant in the robustness dataset, more lags need to be added. The only other difference is the control variable *low*, which is not significant.

| Dependant Variable Choice | Coefficient | Standard Deviation |
|---|---|---|
| 1 lag (high) | 1.333*** | 0.066 |
| 1 lag (low) | -0.110* | 0.061 |
| phigh | 2.753*** | 0.444 |
| high | 0.087*** | 0.025 |
| low | 0.013 | 0.027 |
| medium | -0.100*** | 0.038 |
| trial | -0.010*** | 0.000 |
| order | 0.007 | 0.020 |
| _cons | -2.164*** | 0.314 |

**Table 27: Robustness check lagged rare outcomes**

In Table 28 the model has been expanded with 20 lags. There are some interesting findings here. The first one is that the first lag for the bad rare outcome remains significant and negative. This is opposite to the findings from the initial analysis, implying that subjects might actually take a rare outcome into account if it happened in the previous trial, and move away from the risky choice. Another interesting finding is that the same lag length as the initial analysis seem to hold. All the lags are significant and positive for the good rare outcome, and only the first seven lags are significant for the bad rare outcome. This is one lag less than the initial findings, which found that the first eight lags are significant for the bad rare outcome. The effects are decreasing over time which is in line with the initial findings.

This robustness check thus confirms the findings from the main analysis; there is a difference in behavior dependant on the experienced variance. A positive variance is always taken into account when making a decision, whereas bad variance is no longer used after 7-8 lags. The difference lies in the occurrence of a bad outcome in the previous sample: if subjects experience a bad outcome in the previous sample they react immediately  by moving away from that choice. This was not found in the initial analysis. This can be explained intuitively: subjects respond immediately to a bad shock, moving away from the risky choice. But after a while they might be interested in the risky choice wondering what the other options yields. Therefore the proportion of risky choice increases when the rare outcome occurred recently.

| model with 20 lags | | |
|---|---|---|
| Dependant variable: Choice | Coefficient | Standard Deviation |
| trial | -0.010*** | 0.000 |
| order | 0.005 | 0.019 |
| phigh | 2.586*** | 0.420 |
| high | 0.063*** | 0.024 |
| low | -0.007 | 0.025 |
| medium | -0.062* | 0.036 |
| 1 lag (high) | 1.293*** | 0.068 |
| 2 lag (high) | 1.035*** | 0.066 |
| 3 lag (high) | 0.712*** | 0.064 |
| 4 lag (high) | 0.702*** | 0.064 |
| 5 lag (high) | 0.480*** | 0.063 |
| 6 lag (high) | 0.545*** | 0.063 |
| 7 lag (high) | 0.528*** | 0.063 |
| 8 lag (high) | 0.318*** | 0.062 |
| 9 lag (high) | 0.407*** | 0.062 |
| 10 lag (high) | 0.367*** | 0.062 |
| 11 lag (high) | 0.311*** | 0.062 |
| 12 lag (high) | 0.281*** | 0.061 |
| 13 lag (high) | 0.248*** | 0.061 |
| 14 lag (high) | 0.289*** | 0.062 |
| 15 lag (high) | 0.341*** | 0.061 |
| 16 lag (high) | 0.293*** | 0.061 |
| 17 lag (high) | 0.357*** | 0.062 |
| 18 lag (high) | 0.331*** | 0.061 |
| 19 lag (high) | 0.258*** | 0.061 |
| 20 lag (high) | 0.290*** | 0.061 |
| 1 lag (low) | -0.185**** | 0.061 |
| 2 lag (low) | 0.379*** | 0.063 |
| 3 lag (low) | 0.308*** | 0.063 |
| 4 lag (low) | 0.337*** | 0.063 |
| 5 lag (low) | 0.171*** | 0.062 |
| 6 lag (low) | 0.149** | 0.062 |
| 7 lag (low) | 0.107* | 0.062 |
| 8 lag (low) | 0.041 | 0.062 |
| 9 lag (low) | 0.028 | 0.062 |
| 10 lag (low) | -0.021 | 0.062 |
| 11 lag (low) | -0.093 | 0.062 |
| 12 lag (low) | -0.051 | 0.062 |
| 13 lag (low) | -0.094 | 0.062 |
| 14 lag (low) | -0.094 | 0.062 |
| 15 lag (low) | -0.126 | 0.062 |
| 16 lag (low) | -0.103 | 0.062 |
| 17 lag (low) | -0.138 | 0.062 |
| 18 lag (low) | 0.009 | 0.062 |
| 19 lag (low) | 0.009 | 0.062 |
| 20 lag (low) | 0.038 | 0.062 |
| constant | -2.222*** | 0.297 |



**Table 28: Robustness check lagged rare outcome (20 lags)**

## Robustness check hypothesis 3

To test the robustness of hypothesis 3, the difference between the payoffs of an option is of interest. This means that the same logit model as in the original analysis will be estimated, with the differences between the options as the main independent variables of interest. The results for this test are shown in Table 29. The left side of this table deals with the main model, whereas the right side includes the experienced variance, measured by the lagged rare outcomes from hypothesis two.[12]

Looking at the left side of the table it can be seen that the same significant relation is found for the variable *gain*. This implies that when a subject can gain from learning, he is more likely to select the risky choice, the same finding as in the main analysis. The same relation as before is also found when looking at the cost of learning, measured by the variable *lose.* Opposite to the main analysis however, this is not significant. This result weakens the previous findings with respect to the cost of learning, but strengthens the findings regarding the gains of learning.

The next step is to look at the model which includes the experienced variance, in the right side of the table. The main difference lies in the variable measuring the cost of learning, *lose*. This became positive, but remains insignificant. This finding is in both ways opposite to the initial findings. This means that the experienced variance does not seem to affect the relation regarding the gains of learning, but does affect the relation regarding the cost of learning. For the lags the same findings as the robustness check for hypothesis two hold.

| Dependant variable: Choice | Coefficient | Standard Deviation | Coefficier | Standard Deviation |
|---|---|---|---|---|
| gain | 0.085*** | 0.024 | 0.063*** | 0.024 |
| lose | -0.015 | 0.025 | 0.007 | 0.025 |
| medium | 0.001 | 0.007 | -0.007 | 0.007 |
| trial | -0.010*** | 0.000 | -0.010*** | 0.000 |
| order | 0.007 | 0.019 | 0.005 | 0.019 |
| phigh | 2.601*** | 0.419 | 2.586*** | 0.420 |
| _cons | -1.997*** | 0.296 | -2.222*** | 0.297 |
| lagged rare outcome included? | No | | Yes | |

**Table 29: Robustness check cost of learning**

[12] Again, the coefficients per lag are not included, as this will overcomplicate the table.

Finally it might be true that there are differences based on the subjects who either did not experience the good or the bad rare outcome. The results for this regression are shown in Table 30. The left side of the table shows the results for subjects who did not experience the good rare outcome. The variable *lose* measures the cost of learning, which is not significant, in line with previous findings. For the subjects who did not experience the bad rare outcome, *gain* measures the cost of learning. This is significant and positive, similar as before. The robustness findings thus hold for this hypothesis . Interesting is that the order of the problems is not significant in this dataset.

With respect to hypothesis three it can be concluded that the findings on the cost of learning are weakened as there is no significant relation found here. This could be explained by differences in risk taking and losses in the dataset. The robustness check confirms the findings on gains of learning. When this is expanded with the experienced variance, the relation for the cost of learning goes into the opposite direction. The same findings as before hold when we correct for subject who did not experience the good or the bad rare outcome. Therefore the initial findings are concluded to hold.

| Dependant variable: Choice | Coefficient | Standard Deviation | Coefficier | Standard Deviation |
|---|---|---|---|---|
| gain | 0.117*** | 0.029 | 0.059** | 0.025 |
| lose | 0.001 | 0.024 | -0.121** | 0.056 |
| medium | 0.009 | 0.008 | -0.004 | 0.009 |
| trial | -0.018*** | 0.000 | -0.009*** | 0.000 |
| order | -0.001 | 0.020 | 0.009 | 0.025 |
| phigh | 4.768*** | 0.456 | 2.013** | 0.813 |
| _cons | -3.767*** | 0.357 | -1.628*** | 0.330 |
| Specification: of subjects | Never experienced the good rare outcome | | Never experienced the bad rare outcome | |

**Table 30: Robustness check cost of learning for subjects who did not experience one of the rare outcome**

## Robustness check hypothesis 4

The first robustness check in hypothesis four deals with the difference in choice behavior over time  between subjects who experienced a good or a bad rare outcome. The results for this are shown in Table 31. From this table it can be seen that the same relation is found in the final 50 trials: subjects who experienced the good rare outcome have a higher proportion of risky choice compared to the bad rare outcome. This confirms the findings regarding the settling down behavior. Looking at the first 50 trials it can be seen that  the only difference lies in the trials 31-40, where the reverse relation is found. Since this is the only deviation it is concluded that the robustness check confirms the previous findings.

| Trials | Over | Difference | Standard Deviation |
|---|---|---|---|
| 1-10 | good vs. bad | -0.106*** | 0.011 |
| 11-20 | good vs. bad | 0.006 | 0.028 |
| 21-30 | good vs. bad | 0.024 | 0.016 |
| 31-40 | good vs. bad | -0.083*** | 0.018 |
| 41-50 | good vs. bad | 0.006 | 0.016 |
| 51-60 | good vs. bad | 0.017 | 0.014 |
| 61-70 | good vs. bad | 0.015 | 0.013 |
| 71-80 | good vs. bad | 0.040*** | 0.012 |
| 81-90 | good vs. bad | 0.065*** | 0.012 |
| 91-100 | good vs. bad | 0.098*** | 0.011 |

**Table 31: Robustness check difference good and bad rare outcome**

The next step is be to test the choice behavior over time separately for subject who experienced the bad rare outcome and the good rare outcome. The results for this are shown in Table 32. The main difference with respect to the original analysis is that there are more significant differences over time. But again, there does not seem to be any clear pattern. Interesting to note is that almost all of the observations in the first 50 trials are significant. This indicates that the first 50 trials are the most important, the most variation seems to occur in these trials, implying that the hot stove effect and the gamblers fallacy are most likely to be found in these trials.

| Trials | Over | Difference | Standard Deviation |
|---|---|---|---|
| 1 to 2 | good rare observed | -0.067*** | 0.007 |
| | bad rare observed | -0.051*** | 0.007 |
| 2 to 3 | good rare observed | 0.017** | 0.008 |
| | bad rare observed | -0.004 | 0.008 |
| 3 to 4 | good rare observed | -0.015** | 0.007 |
| | bad rare observed | -0.017** | 0.007 |
| 4 to 5 | good rare observed | 0.018*** | 0.006 |
| | bad rare observed | 0.005 | 0.006 |
| 5 to 6 | good rare observed | 0.018*** | 0.006 |
| | bad rare observed | -0.012** | 0.006 |
| 6 to 7 | good rare observed | -0.002 | 0.005 |
| | bad rare observed | -0.005 | 0.005 |
| 7 to 8 | good rare observed | -0.008 | 0.005 |
| | bad rare observed | -0.001 | 0.005 |
| 8 to 9 | good rare observed | 0.007 | 0.004 |
| | bad rare observed | -0.003 | 0.004 |
| 9 to 10 | good rare observed | -0.010** | 0.004 |
| | bad rare observed | 0.004 | 0.004 |

Table 32: Robustness check hot stove and gambler's fallacy over time

The next step is to  differentiate with respect to the timing outcome occurred. This is done in the same way as before, and the results for the bad rare outcome are shown in Table 33. Similar to before most of the timings are significant when we compare the proportion of risky choice dependant on where the rare bad outcome occurred. The lowest proportion of risky choice lies for subjects who only experience the rare outcome in the final 10 trials. Furthermore it can be concluded who experience the rare outcome in the later part of the trials have a higher proportion of risky choice; they will switch towards the safe option after they experienced the bad rare outcome, as the proportion of risky choice is significantly increasing. This is in line with the expectations as it was expected that the risky choice should be the lowest for subjects who experience the bad rare outcome in the first part of the trials. Interesting is that the findings of the robustness check with respect to the setting down are more significant compared to the original analysis.

This analysis is of course repeated for when the good outcome is rare. The results for this are shown in Table 34. Again most of the differences between the timing of the good rare outcome are significant. Looking at the proportion of risky choice it can be seen that again, there is no clear pattern visible in the difference between the proportion of risky choice for the timing of the rare effect for subjects who experience the good rare outcome. The proportion of risky choice is the highest for subjects who experience the good rare outcome in the trials 21-30, followed by the trials 41-50 and 51-60. It is the lowest for the subjects who experienced a good rare outcome in the final twenty trials.

With respect to hypothesis four it can be concluded that the robustness check confirms the findings. There is  slightly clearer settling down behavior present for subjects who experienced the bad rare outcome compared to the original analysis. Again, no clear setting down is found for subject who experienced the good rate outcome. This robustness check confirms that the biases are present, but does not clarify the settling down behavior.

| Difference with mean choice | 1 to 10 | 11 to 20 | 21 to 30 | 31 to 40 | 41 to 50 | 51 to 60 | 61 to 70 | 71 to 80 | 81 to 90 | 91 to 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 to 10 | xxx | | | | | | | | | |
| *(standard deviation)* | *xxx* | | | | | | | | | |
| 11 to 20 | 0.029** | xxx | | | | | | | | |
| *(standard deviation)* | *(0.012)* | *xxx* | | | | | | | | |
| 21 to 30 | 0.128*** | 0.099*** | xxx | | | | | | | |
| *(standard deviation)* | *(0.011)* | *(0.015)* | *xxx* | | | | | | | |
| 31 to 40 | 0.239*** | 0.210** | 0.111*** | xxx | | | | | | |
| *(standard deviation)* | *(0.011)* | *(0.015)* | *(0.015)* | *xxx* | | | | | | |
| 41 to 50 | 0.285*** | 0.256*** | 0.157*** | 0.046*** | xxx | | | | | |
| *(standard deviation)* | *(0.010)* | *(0.015)* | *(0.014)* | *(0.014)* | *xxx* | | | | | |
| 51 to 60 | 0.270*** | 0.241*** | 0.142*** | 0.031** | -0.014 | xxx | | | | |
| *(standard deviation)* | *(0.012)* | *(0.016)* | *(0.015)* | *(0.015)* | *(0.015)* | *xxx* | | | | |
| 61 to 70 | 0.252*** | 0.223*** | 0.124*** | 0.013 | -0.033** | -0.018 | xxx | | | |
| *(standard deviation)* | *(0.012)* | *(0.016)* | *(0.015)* | *(0.015)* | *(0.015)* | *(0.016)* | *xxx* | | | |
| 71 to 80 | 0.288*** | 0.259*** | 0.160*** | 0.049*** | 0.003 | 0.018 | 0.036** | xxx | | |
| *(standard deviation)* | *(0.014)* | *(0.018)* | *(0.017)* | *(0.017)* | *(0.017)* | *(0.018)* | *(0.018)* | *xxx* | | |
| 81 to 90 | 0.314*** | 0.285*** | 0.186*** | 0.075*** | 0.029** | 0.044*** | 0.062*** | 0.026 | xxx | |
| *(standard deviation)* | *(0.011)* | *(0.015)* | *(0.015)* | *(0.015)* | *(0.014)* | *(0.015)* | *(0.015)* | *(0.017)* | *xxx* | |
| 91 to 100 | -0.126*** | -0.155*** | -0.254*** | -0.365*** | -0.411*** | -0.396*** | -0.378*** | -0.414*** | -0.440*** | xxx |
| *(standard deviation)* | *(0.004)* | *(0.011)* | *(0.011)* | *(0.010)* | *(0.010)* | *(0.011)* | *(0.011)* | *(0.014)* | *(0.011)* | *xxx* |

**Table 33: Robustness check timing of bad rare outcome**

| Difference with mean choice | 1 to 10 | 11 to 20 | 21 to 30 | 31 to 40 | 41 to 50 | 51 to 60 | 61 to 70 | 71 to 80 | 81 to 90 | 91 to 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 to 10 | xxx | | | | | | | | | |
| *(standard deviation)* | *xxx* | | | | | | | | | |
| 11 to 20 | -0.073*** | xxx | | | | | | | | |
| *(standard deviation)* | *(0.015)* | *xxx* | | | | | | | | |
| 21 to 30 | 0.041*** | 0.113*** | xxx | | | | | | | |
| *(standard deviation)* | *(0.013)* | *(0.020)* | *xxx* | | | | | | | |
| 31 to 40 | -0.111*** | -0.038** | -0.152*** | xxx | | | | | | |
| *(standard deviation)* | *(0.012)* | *(0.019)* | *(0.017)* | *xxx* | | | | | | |
| 41 to 50 | 0.001 | 0.073*** | -0.040** | 0.112*** | xxx | | | | | |
| *(standard deviation)* | *(0.013)* | *(0.020)* | *(0.018)* | *(0.017)* | *xxx* | | | | | |
| 51 to 60 | 0.022 | 0.094*** | -0.019 | 0.133*** | 0.021 | xxx | | | | |
| *(standard deviation)* | *(0.013)* | *(0.020)* | *(0.018)* | *(0.017)* | *(0.018)* | *xxx* | | | | |
| 61 to 70 | -0.121*** | -0.049** | -0.162*** | -0.010 | -0.122*** | -0.143*** | xxx | | | |
| *(standard deviation)* | *(0.016)* | *(0.022)* | *(0.020)* | *(0.020)* | *(0.020)* | *(0.020)* | *xxx* | | | |
| 71 to 80 | -0.066*** | 0.006 | -0.107*** | 0.045** | -0.067*** | -0.088*** | 0.055*** | xxx | | |
| *(standard deviation)* | *(0.014)* | *(0.020)* | *(0.019)* | *(0.018)* | *(0.019)* | *(0.019)* | *(0.021)* | *xxx* | | |
| 81 to 90 | -0.123*** | -0.051** | -0.164*** | -0.012 | -0.124*** | -0.145*** | -0.002 | -0.057*** | xxx | |
| *(standard deviation)* | *(0.013)* | *(0.020)* | *(0.018)* | *(0.017)* | *(0.018)* | *(0.018)* | *(0.020)* | *(0.019)* | *xxx* | |
| 91 to 100 | -0.229*** | -0.157*** | -0.270*** | -0.118*** | -0.230*** | -0.251*** | -0.108*** | -0.163*** | -0.106*** | xxx |
| *(standard deviation)* | *(0.004)* | *(0.015)* | *(0.013)* | *(0.012)* | *(0.013)* | *(0.013)* | *(0.016)* | *(0.014)* | *(0.013)* | *xxx* |

**Table 34: Robustness check timing of good rare outcome.**

## Robustness check hypothesis 5

The final robustness check deals with the type of problem. The first step will be to look at the initial effect of the type of problem on the choice. Again, mixed problems are selected as the reference category. The results for this regression are shown in Table 35. As can be seen, there seem to be no significant differences with respect to mixed problems for both gain and loss problems. It can be concluded that the proportion of risky choice with respect to the mixed problems are quite different to the original findings. The proportion of risky choice is the lowest for subjects facing losses, and the highest for subjects facing mixed problems. Again, note that this is not significant. Furthermore, it might be caused by the height of the payoffs as we know that the mean payoff is negative.

| Dependant variable: Choice | Coefficient | Standard Deviation |
|---|---|---|
| lossproblem | -0.111 | 0.164 |
| gainproblem | -0.074 | 0.164 |
| _cons | -0.727*** | 0.116 |

**Table 35: Robustness check type of problem (restricted)**

To be able to give definitive proof of the findings from the restricted model on the type of problems the complete model will be analyzed, with both interaction terms and all control variables. The results for this are shown in Table 36. Interesting to note is that a different relation as before is found for the interaction term between problems over gains and the probability of the high outcome. This is now positive and significant, opposite to negative and significant. The interaction term over problems facing losses does have the same relation, but is no longer significant. Another important finding compared to the initial analysis is that the dummy variables for the type of problems are not significant. This implies that there are no significant differences between the type of problems, which means that the difference in behavior is only determined by the probability of the risky choice; the more likely the best option becomes, the more likely a subject will be to select the risky choice. This means that we cannot state anything about over or underweighting, and it contradicts the initial findings. For the rest of the control variables, the same relations seem to hold.

With respect to hypothesis five in the robustness check it can be concluded that the findings here weaken the initial findings, as opposite results are found. The explanation for this is likely to be the negative payoff, which affects the choice behavior of the subject in the competition data set. This implies that there might be some sort of underlying behavior present. One suggestion might be that subjects relate their payoffs to their total wealth, which is something for future research.

| Dependant Variable Choice | Coefficient | Standard Deviation |
|---|---|---|
| gainproblem * phigh | 0.915** | 0.457 |
| lossproblem * phigh | -0.056 | 0.437 |
| | | |
| gainproblem | -0.472 | 0.325 |
| lossproblem | 0.080 | 0.338 |
| | | |
| phigh | 2.427*** | 0.542 |
| high | 0.083*** | 0.027 |
| low | 0.017 | 0.029 |
| medium | -0.161*** | 0.039 |
| trial | -0.010*** | 0.000 |
| order | 0.008 | 0.020 |
| lagged payoff | 0.066*** | 0.004 |
| _cons | -1.954 | 0.359 |

**Table 36: Robustness check type of problems (unrestricted) with interaction terms**

## Conclusions on the robustness checks

This chapter tested the robustness of the main findings trough the use of another dataset on the same experiment; the competition set. From the summary statistics it became clear that the experiments are comparable. The main difference lies in the payoff of the options. The mean payoff was negative in this experiment, opposite to the original experiment yielding a positive payoff. This might explain differences in choice behavior between the experiments. For all the four hypotheses of interest in this paper, the analysis is repeated using the robustness dataset. For hypothesis one, the findings are confirmed. For hypothesis two, one small difference was found with respect to the experienced variance, namely the effect of a bad rare outcome in the previous sample. This is found to be significant and negative, opposite to the initial findings. The rest of the findings regarding hypothesis two seem to hold. With respect to hypothesis three, it is found that there is no relation found regarding the cost of learning, but the relation for the gains of learning is confirmed. This might be caused by the differences in payoffs between the two experiments. Finally in hypothesis four, the same findings as before have been found. However, this did not yield clarification. In hypothesis five, the opposite of the initial findings are found. This might be caused by the difference in the mean payoffs, which is something future research could analyze.

When the findings with respect to the robustness analysis and the original findings are combined, conclusions can be drawn with respect to the exploration-exploitation tradeoff. This is described in the conclusion in the next chapter.

## Conclusions

In this research  decision making has been analyzed. Specifically, decision making from experience where the underlying probabilities must become known trough experiencing and exploring the outcomes. Within this decision making from experience, a tradeoff is found to be present between this exploring and the alternative, setting down behavior. In literature, this tradeoff has become known as the exploration-exploitation tradeoff. Previous research analyzed this tradeoff mainly with respect to the small-sample problems, (Erev & Barron, 2005) the searching behavior, (Lejarraga, Hertwig, & Gonzalez, 2012) and the one-armed bandit problems (Berry & Fristedt, 1985). The goal of this research was to shed more light on the tradeoff and the choice behavior when subjects are facing an unknown risky and an unknown safe choice. Here, subjects received feedback on the payoff of their previous choice. The previous theory and literature gave rise to five main hypotheses:

*Hypothesis 1: The lack of exploration is the main reason of the description-experience gap in the feedback paradigm.*

*Hypothesis 2: Subjects explore more when they experience more variance*.

*Hypothesis 3: Subjects explore less when they experience higher costs of learning*

*Hypothesis 4: The hot stove effect and the gamblers fallacy affect the exploration behavior, as reflected in the settling down behavior of the subjects.*

*Hypothesis 5: the choice behavior of subjects differ  when facing losses compared to gains.*

These were analyzed trough the estimation dataset from Erev. et al (2010), which consists of five sessions of twenty subjects, each facing twelve problems of 100 samples over decision making from experience. The main variable of interest here is the variable *choice* which measures whether subjects picked the safe option (choice = 0) or the risky option (choice = 1). As a robustness check, the analysis is repeated over the competition dataset from Erev. et al (2010), which has similar characteristics.

The general findings are that differences in the choice of subjects can be explained differences in settling down behavior. All the subjects have a decreasing proportion of risky choice over time, which shows settling down behavior. This is found to be stronger when subjects experience a rare outcome, resulting in small-sample problems. Furthermore, experienced variance affects this settling down behavior as well, but this differs dependant on the type of experienced variance. Negative variance results in smaller samples compared to positive variance. It  is also found that higher gains of learning make the subject more likely to explore the risky choice, whereas a higher cost of learning makes the subject less likely to explore.

Both the  gamblers fallacy and the hot stove effect are found to be present in the behavior of the subjects. The settling down behavior is found to be the strongest for subjects for whom the hot stove effect is of importance. Furthermore, subjects behave different regarding the types of payoffs they are facing. They are more likely to select the risky choice when they are facing gains compared to problems over mixed payoff or over losses, but does not seem to hold when this is tested over the competition dataset. In this, the opposite is found for the behavior in different problems.

From this research we have thus learned that regarding the exploration and exploitation tradeoff in decision making from experience, there are many factors affecting the choice behavior of subjects. This is mainly caused by the small-sample problems, but also by the hot stove effect and the gamblers fallacy. Furthermore, it is found that the experienced variance and the cost of learning also affect this tradeoff. This means that in practice, investors should avoid small-sample problems in their decision making. One sudden shock in the stock returns, either positive or negative, should not change the investors' beliefs  as they might fall for the hot stove effect, or the gamblers fallacy. This is reflected in works of Barbaris, Sheifler, and Vishny (1998), stating that investors should update their beliefs dependant on their risk aversion: monthly if risk loving, yearly if risk averse.

There are some restrictions to this research. For instance, we did not create our own experiment, which means that we lost control over the reliance of the data. Furthermore, the optimal lag length could not be determined because the lags for the rare outcome have to be specified individually per lag. Therefore there is no definitive proof on the optimal lag length. Finally, there are some restrictions in our conclusions regarding the robustness check. These do not seem to give definitive evidence on some conclusions, e.g. the effect of the lagged bad rare outcome, or the settling down behavior of the gamblers fallacy. The main restrictions are the findings with respect to hypothesis five, the type of problem faced.

This research gives some interesting findings for future research. For example, it might be interesting to focus specific on the gamblers fallacy, as there is no clear evidence on the choice behavior over time of subjects within this. My suggestion would be to combine this with a one-armed bandit problem, as there are currently few experimental researches dealing with one-armed bandits in this field. The main suggestion would be to focus on the risk preferences of subjects when facing decisions from experience. The findings with respect to the type of problem a subject was facing showed that there seem to be differences in behavior regarding the likelihood and the payoffs of the different problems, and this research did not capture the causes for this. It is hypothesized that the mean payoff may be causing this, affecting the risk behavior. Could there be some sort of Prospect Theory underlying the beliefs of subjects here?

## Used References


Aghion, P., Bolton, P., Harris, C., & Jullien, B. (1991). Optimal learning by experimentation. *Review of economic studies, volume 58:4* , 621-654.

Audibert, J.-Y., Munos, R., & Szepesvari, C. (2009). Exploration-exploitation tradeoff using variance estimates in multi-armed bandits. *Theoretical computer science, volume 410:29* , 1876-1902.

Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: two faces of subjective randomness. *Memory & Cognition, volume 32:8* , 1369-1378.

Barbaris, N., Sheifler, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, Volume 49:3* , 307-343.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description based decisions. *Journal of Behavioral decision making, Volume 16* , 215-233.

Berry, D., & Fristedt, B. (1985). Bandit problems: sequential allocation of experiments. In *Monographs on Statistics and Applied Probability.* London: Chapman & Hall.

Bubeck, S., Munos, R., & Stoltz, G. (2011). Pure exploration in finite-armed and continous-armed bandits. *Theoretical computer science, volume 412:19* , 1832-1852.

Camerer, C., & Ho, T. (1999). Experience-weighted Attraction Learning in Normal Form Game. *Econometrica, Volume 67:4* , 827-874.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review, volume 112* , 912-931.

Erev, I., & Roth, A. (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *American Economic Review, Volume 88:4* , 848-881.

Erev, I., Ert, E., & Roth, A. (2010). A choice prediction competition for market entry games: an introduction. *Games, volume 1:3* , 221-225.

Friedman, D. (1998). Evolutionary economics goes mainstream: a review of the theory of learning in games . *Journal of Evolutionairy Economics, Volume 8:4* , 423-432.

Fujikawa, T. (2009). On the importance of the hot stove effect and the tendency to rely on small samples. *Judgement and decision making, volume 4:5* , 429-435.

Gittins, J., & Jones, D. (1979). A dynamic allocation index for the discounted multiarmed bandit problem. *Biometrica, volume 66:3* , 561-565.

Gureckis, T., & Love, B. (2009). Learning in Noise: Dynamic Decision-Making in a Variable Environment. *J Math Psychology, volume 53:3* , 180-193.

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in cognitive science, volume 13:12* , 517-523.

Hertwig, R., & Pleskac, T. (2010). Decisions from experience, why small samples? *Cognition, volume 115* , 225-237.

Hertwig, R., Barron, G., & Weber, E. E. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, Volume 15:8* , 534-539.

Hills, T., & Hertwig, R. (2010). Information search in decision from experience: do our patterns of sampling foreshadow our decisions. *Psychological science, volume 21:12* , 1787-1792.

Hogarth, R., & Einhorn, H. (1992). Order effects in belief updating: the belief-adjustment model. *Cognitive Psychology, Volume 24:1* , 1-55.

Jessup, R., Bishara, A., & Busemeyer, J. (2008). Feedback produces divergence from prospect theory in descriptive choice. *Psychology Science, volume 19* , 1015-1022.

Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica Volume 47:2* , 263-291.

Lejarraga, T., Hertwig, R., & Gonzalez, C. (2012). How choice ecology influences search in decisions from experience. *Cognition, volume 124* , 334-342.

Ludvig, E., & Spetch, M. (2011). Of black swans and tossed coins: is the desciption-experience gap in risky choice limited to rare events? *PLoS one, volume 6:6* , e20262.

March, J. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science, volume 2:1* , 71-87.

Shafran, A. (2011). Self-protection against repeated low probability risks. *Journal of Risk and Uncertainty, volume 42:3* , 263-285.

Shanks, D., Tunney, R., & McCarthy, J. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making, volume 15:3* , 233-250.

Yechiam, E., & Busemeyer, J. (2006). The effect of forgone payoffs on underweighing small probabiliity events. *Journal of behavioral decision making, Volume 19:1* , 1-16.

# Appendix

| Problem: | Payoff high: | Probability High: | Payoff Low: | Probability Low: | Expected payoff risky option: | Payoff Safe option: | Highest payoff? |
|---|---|---|---|---|---|---|---|
| 1 | -8.70 | 6% | -22.80 | 94% | -21.95 | -21.40 | safe |
| 2 | -2.20 | 9% | -9.60 | 91% | -8.93 | -8.70 | safe |
| 3 | -2.00 | 10% | -11.20 | 90% | -10.28 | -9.50 | safe |
| 4 | -1.40 | 2% | -9.10 | 98% | -8.95 | -9.00 | risky |
| 5 | -0.90 | 7% | -4.80 | 93% | -4.53 | -4.70 | risky |
| 6 | -4.70 | 91% | -18.10 | 9% | -5.91 | -6.80 | risky |
| 7 | -9.70 | 6% | -24.80 | 94% | -23.89 | -24.20 | risky |
| 8 | -5.70 | 96% | -20.60 | 4% | -6.30 | -6.40 | risky |
| 9 | -5.60 | 10% | -19.40 | 90% | -18.02 | -18.10 | risky |
| 10 | -2.50 | 60% | -5.50 | 40% | -3.70 | -3.60 | safe |
| 11 | -5.80 | 97% | -16.40 | 3% | -6.12 | -6.60 | risky |
| 12 | -7.20 | 5% | -16.10 | 95% | -15.66 | -15.60 | safe |
| 13 | -1.80 | 93% | -6.70 | 7% | -2.14 | -2.00 | safe |
| 14 | -6.40 | 20% | -22.40 | 80% | -19.20 | -18.00 | safe |
| 15 | -3.30 | 97% | -10.50 | 3% | -3.52 | -3.20 | safe |
| 16 | -9.50 | 10% | -24.50 | 90% | -23.00 | -23.50 | risky |
| 17 | -2.20 | 92% | -11.50 | 8% | -2.94 | -3.40 | risky |
| 18 | -1.40 | 93% | -4.70 | 7% | -1.63 | -1.70 | risky |
| 19 | -8.60 | 10% | -26.50 | 90% | -24.71 | -26.30 | risky |
| 20 | -6.90 | 6% | -20.50 | 94% | -19.68 | -20.30 | risky |
| 21 | 1.80 | 60% | -4.10 | 40% | -0.56 | 1.70 | safe |
| 22 | 9.00 | 97% | -6.70 | 3% | 8.53 | 9.10 | safe |
| 23 | 5.50 | 6% | -3.40 | 94% | -2.87 | -2.60 | safe |
| 24 | 1.00 | 93% | -7.10 | 7% | 0.43 | 0.60 | safe |
| 25 | 3.00 | 20% | -1.30 | 80% | -0.44 | -0.10 | safe |
| 26 | 8.90 | 10% | -1.40 | 90% | -0.37 | -0.90 | risky |
| 27 | 9.40 | 95% | -6.30 | 5% | 8.62 | 8.50 | risky |
| 28 | 3.30 | 91% | -3.50 | 9% | 2.69 | 2.70 | safe |
| 29 | 5.00 | 40% | -6.90 | 60% | -2.14 | -3.80 | risky |
| 30 | 2.10 | 6% | -9.40 | 94% | -8.71 | -8.40 | safe |
| 31 | 0.90 | 20% | -5.00 | 80% | -3.82 | -5.30 | risky |
| 32 | 9.90 | 5% | -8.70 | 95% | -7.77 | -7.60 | safe |
| 33 | 7.70 | 2% | -3.10 | 98% | -2.88 | -3.00 | risky |
| 34 | 2.50 | 96% | -2.00 | 4% | 2.32 | 2.30 | risky |
| 35 | 9.20 | 91% | -0.70 | 9% | 8.31 | 8.20 | risky |
| 36 | 2.90 | 98% | -9.40 | 2% | 2.65 | 2.90 | safe |
| 37 | 2.90 | 5% | -6.50 | 95% | -6.03 | -5.70 | safe |
| 38 | 7.80 | 99% | -9.30 | 1% | 7.63 | 7.60 | risky |
| 39 | 6.50 | 80% | -4.80 | 20% | 4.24 | 6.20 | safe |
| 40 | 5.00 | 90% | -3.80 | 10% | 4.12 | 4.10 | risky |
| 41 | 20.10 | 95% | 6.50 | 5% | 19.42 | 19.60 | safe |
| 42 | 5.20 | 50% | 1.40 | 50% | 3.30 | 5.10 | safe |
| 43 | 12.00 | 50% | 2.40 | 50% | 7.20 | 9.00 | safe |
| 44 | 20.70 | 90% | 9.10 | 10% | 19.54 | 19.80 | safe |
| 45 | 8.40 | 7% | 1.20 | 93% | 1.70 | 1.60 | risky |
| 46 | 22.60 | 40% | 7.20 | 60% | 13.36 | 12.40 | risky |
| 47 | 23.40 | 93% | 7.60 | 7% | 22.29 | 22.10 | risky |
| 48 | 17.20 | 9% | 5.00 | 91% | 6.10 | 5.90 | risky |
| 49 | 18.90 | 90% | 6.70 | 10% | 17.68 | 17.70 | safe |
| 50 | 12.80 | 4% | 4.70 | 96% | 5.02 | 4.90 | risky |
| 51 | 19.10 | 3% | 4.80 | 97% | 5.23 | 5.20 | risky |
| 52 | 12.30 | 91% | 1.30 | 9% | 11.31 | 12.10 | safe |
| 53 | 6.80 | 90% | 3.00 | 10% | 6.42 | 6.70 | safe |
| 54 | 22.60 | 30% | 9.20 | 70% | 13.22 | 11.00 | risky |
| 55 | 6.40 | 9% | 0.50 | 91% | 1.03 | 1.50 | safe |
| 56 | 15.30 | 6% | 5.90 | 94% | 6.46 | 7.10 | safe |
| 57 | 5.30 | 90% | 1.50 | 10% | 4.92 | 4.70 | risky |
| 58 | 21.90 | 50% | 8.10 | 50% | 15.00 | 12.60 | risky |
| 59 | 27.50 | 70% | 9.20 | 30% | 22.01 | 21.90 | risky |
| 60 | 4.40 | 20% | 0.70 | 80% | 1.44 | 1.10 | risky |

Table 37 :Distribution of problem, payoffs, and probabilities

| Problem: | High - Low | High - Medium | Medium - low | | Difference (H-L) & (H-M) | Difference (H-M) & (M-L) | Difference (H-L) & (M-L) |
|---|---|---|---|---|---|---|---|
| 1 | 14.1 | 12.7 | 1.40 | | 1.4 | 11.3 | 12.7 |
| 2 | 7.4 | 6.5 | 0.90 | | 0.9 | 5.6 | 6.5 |
| 3 | 9.2 | 7.5 | 1.70 | | 1.7 | 5.8 | 7.5 |
| 4 | 7.7 | 7.6 | 0.10 | | 0.1 | 7.5 | 7.6 |
| 5 | 3.9 | 3.8 | 0.10 | | 0.1 | 3.7 | 3.8 |
| 6 | 13.4 | 2.1 | 11.30 | | 11.3 | -9.2 | 2.1 |
| 7 | 15.1 | 14.5 | 0.60 | | 0.6 | 13.9 | 14.5 |
| 8 | 14.9 | 0.7 | 14.20 | | 14.2 | -13.5 | 0.7 |
| 9 | 13.8 | 12.5 | 1.30 | | 1.3 | 11.2 | 12.5 |
| 10 | 3.0 | 1.1 | 1.90 | | 1.9 | -0.8 | 1.1 |
| 11 | 10.6 | 0.8 | 9.80 | | 9.8 | -9.0 | 0.8 |
| 12 | 8.9 | 8.4 | 0.50 | | 0.5 | 7.9 | 8.4 |
| 13 | 4.9 | 0.2 | 4.70 | | 4.7 | -4.5 | 0.2 |
| 14 | 16.0 | 11.6 | 4.40 | | 4.4 | 7.2 | 11.6 |
| 15 | 7.2 | -0.1 | 7.30 | | 7.3 | -7.4 | -0.1 |
| 16 | 15.0 | 14.0 | 1.00 | | 1.0 | 13.0 | 14.0 |
| 17 | 9.3 | 1.2 | 8.10 | | 8.1 | -6.9 | 1.2 |
| 18 | 3.3 | 0.3 | 3.00 | | 3.0 | -2.7 | 0.3 |
| 19 | 17.9 | 17.7 | 0.20 | | 0.2 | 17.5 | 17.7 |
| 20 | 13.6 | 13.4 | 0.20 | | 0.2 | 13.2 | 13.4 |
| 21 | 5.9 | 0.1 | 5.80 | | 5.8 | -5.7 | 0.1 |
| 22 | 15.7 | -0.1 | 15.80 | | 15.8 | -15.9 | -0.1 |
| 23 | 8.9 | 8.1 | 0.80 | | 0.8 | 7.3 | 8.1 |
| 24 | 8.1 | 0.4 | 7.70 | | 7.7 | -7.3 | 0.4 |
| 25 | 4.3 | 3.1 | 1.20 | | 1.2 | 1.9 | 3.1 |
| 26 | 10.3 | 9.8 | 0.50 | | 0.5 | 9.3 | 9.8 |
| 27 | 15.7 | 0.9 | 14.80 | | 14.8 | -13.9 | 0.9 |
| 28 | 6.8 | 0.6 | 6.20 | | 6.2 | -5.6 | 0.6 |
| 29 | 11.9 | 8.8 | 3.10 | | 3.1 | 5.7 | 8.8 |
| 30 | 11.5 | 10.5 | 1.00 | | 1.0 | 9.5 | 10.5 |
| 31 | 5.9 | 6.2 | -0.30 | | -0.3 | 6.5 | 6.2 |
| 32 | 18.6 | 17.5 | 1.10 | | 1.1 | 16.4 | 17.5 |
| 33 | 10.8 | 10.7 | 0.10 | | 0.1 | 10.6 | 10.7 |
| 34 | 4.5 | 0.2 | 4.30 | | 4.3 | -4.1 | 0.2 |
| 35 | 9.9 | 1.0 | 8.90 | | 8.9 | -7.9 | 1.0 |
| 36 | 12.3 | 0.0 | 12.30 | | 12.3 | -12.3 | 0.0 |
| 37 | 9.4 | 8.6 | 0.80 | | 0.8 | 7.8 | 8.6 |
| 38 | 17.1 | 0.2 | 16.90 | | 16.9 | -16.7 | 0.2 |
| 39 | 11.3 | 0.3 | 11.00 | | 11.0 | -10.7 | 0.3 |
| 40 | 8.8 | 0.9 | 7.90 | | 7.9 | -7.0 | 0.9 |
| 41 | 13.6 | 0.5 | 13.10 | | 13.1 | -12.6 | 0.5 |
| 42 | 3.8 | 0.1 | 3.70 | | 3.7 | -3.6 | 0.1 |
| 43 | 9.6 | 3.0 | 6.60 | | 6.6 | -3.6 | 3.0 |
| 44 | 11.6 | 0.9 | 10.70 | | 10.7 | -9.8 | 0.9 |
| 45 | 7.2 | 6.8 | 0.40 | | 0.4 | 6.4 | 6.8 |
| 46 | 15.4 | 10.2 | 5.20 | | 5.2 | 5.0 | 10.2 |
| 47 | 15.8 | 1.3 | 14.50 | | 14.5 | -13.2 | 1.3 |
| 48 | 12.2 | 11.3 | 0.90 | | 0.9 | 10.4 | 11.3 |
| 49 | 12.2 | 1.2 | 11.00 | | 11.0 | -9.8 | 1.2 |
| 50 | 8.1 | 7.9 | 0.20 | | 0.2 | 7.7 | 7.9 |
| 51 | 14.3 | 13.9 | 0.40 | | 0.4 | 13.5 | 13.9 |
| 52 | 11.0 | 0.2 | 10.80 | | 10.8 | -10.6 | 0.2 |
| 53 | 3.8 | 0.1 | 3.70 | | 3.7 | -3.6 | 0.1 |
| 54 | 13.4 | 11.6 | 1.80 | | 1.8 | 9.8 | 11.6 |
| 55 | 5.9 | 4.9 | 1.00 | | 1.0 | 3.9 | 4.9 |
| 56 | 9.4 | 8.2 | 1.20 | | 1.2 | 7.0 | 8.2 |
| 57 | 3.8 | 0.6 | 3.20 | | 3.2 | -2.6 | 0.6 |
| 58 | 13.8 | 9.3 | 4.50 | | 4.5 | 4.8 | 9.3 |
| 59 | 18.3 | 5.6 | 12.70 | | 12.7 | -7.1 | 5.6 |
| 60 | 3.7 | 3.3 | 0.40 | | 0.4 | 2.9 | 3.3 |

**Table 38: Distribution of variance within problems**

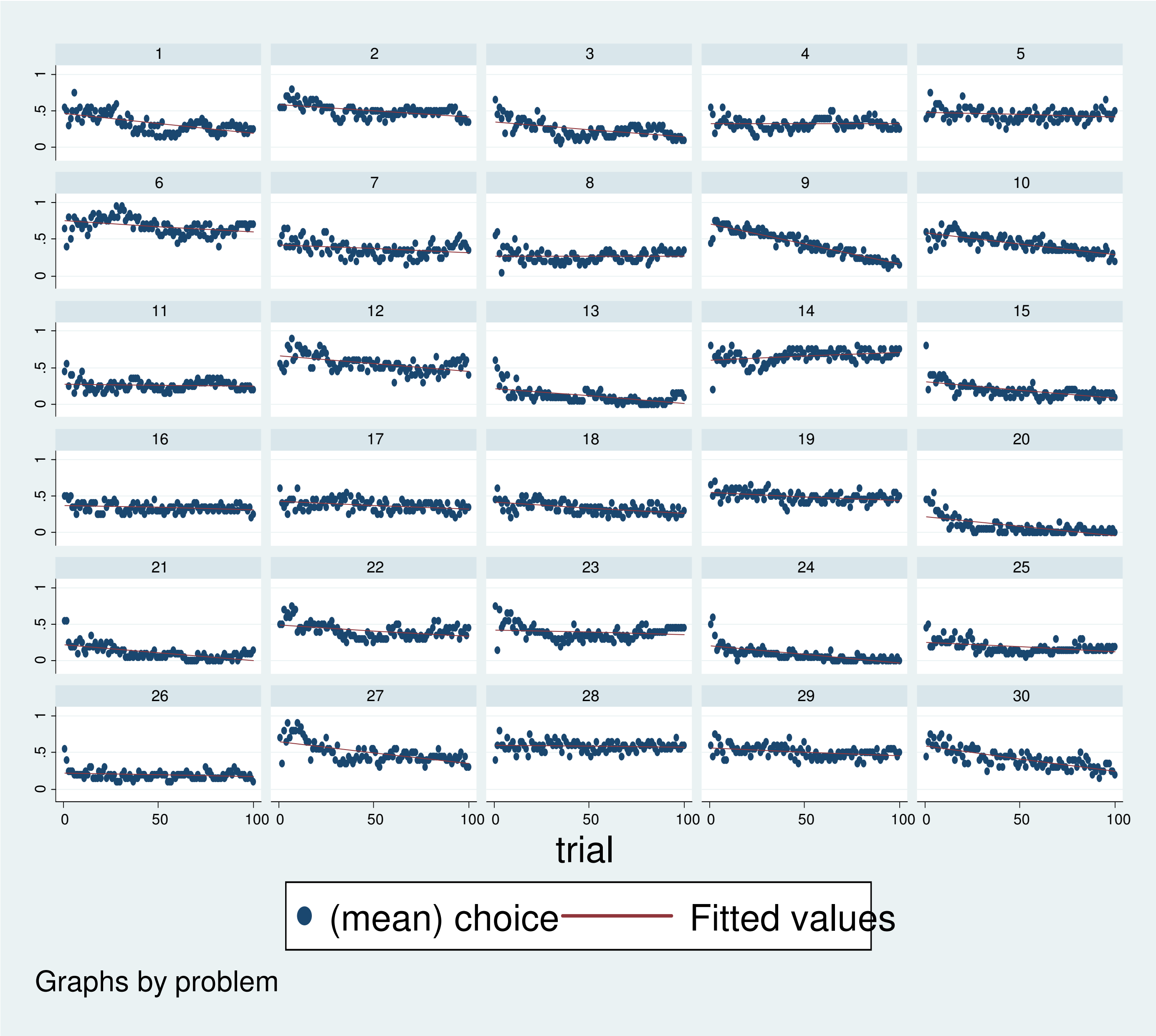



**Figure 23: average choice per trial for problem 1-30**

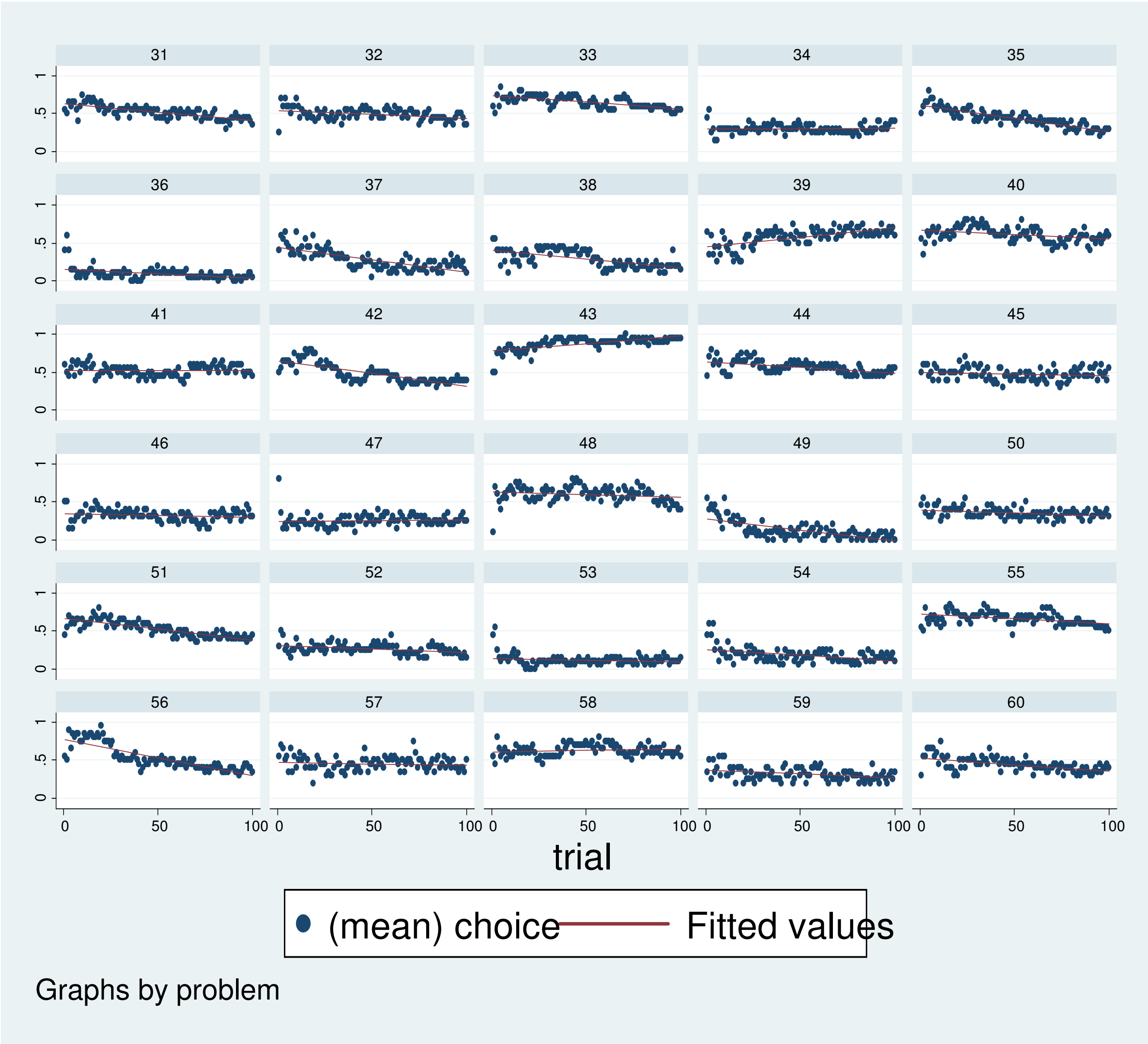



Figure 24: average choice per trial for problem 31-60

| | session | | | | | | session | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| problem | 1 | 2 | 3 | 4 | 5 | problem | 1 | 2 | 3 | 4 | 5 |
| 1 | X | | | | | 31 | | | X | | |
| 2 | X | | | | | 32 | | | X | | |
| 3 | X | | | | | 33 | | | | X | |
| 4 | X | | | | | 34 | | | | X | |
| 5 | | X | | | | 35 | | | | X | |
| 6 | | X | | | | 36 | | | | X | |
| 7 | | X | | | | 37 | | | | | X |
| 8 | | X | | | | 38 | | | | | X |
| 9 | | | X | | | 39 | | | | | X |
| 10 | | | X | | | 40 | | | | | X |
| 11 | | | X | | | 41 | X | | | | |
| 12 | | | X | | | 42 | X | | | | |
| 13 | | | | X | | 43 | X | | | | |
| 14 | | | | X | | 44 | X | | | | |
| 15 | | | | X | | 45 | | X | | | |
| 16 | | | | X | | 46 | | X | | | |
| 17 | | | | | X | 47 | | X | | | |
| 18 | | | | | X | 48 | | X | | | |
| 19 | | | | | X | 49 | | | X | | |
| 20 | | | | | X | 50 | | | X | | |
| 21 | X | | | | | 51 | | | X | | |
| 22 | X | | | | | 52 | | | X | | |
| 23 | X | | | | | 53 | | | | X | |
| 24 | X | | | | | 54 | | | | X | |
| 25 | | X | | | | 55 | | | | X | |
| 26 | | X | | | | 56 | | | | X | |
| 27 | | X | | | | 57 | | | | | X |
| 28 | | X | | | | 58 | | | | | X |
| 29 | | | X | | | 59 | | | | | X |
| 30 | | | X | | | 60 | | | | | X |

Table 39: Problem distribution per session