Jena Economic Research Papers

2011 – 009

Reinforcement Learning in Repeated Portfolio Decisions

by

Linan Diao
Jörg Rieskamp

www.jenecon.de

ISSN 1864-7057

The Jena Economic Research Papers is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

Reinforcement Learning in Repeated Portfolio Decisions

Linan Diao[1] & Jörg Rieskamp[2*]

[1.]Max Planck Institute of Economics, Germany;

[2] University of Basel, Switzerland

Abstract

How do people make investment decisions when they receive outcome feedback? We examined how well the standard mean-variance model and two reinforcement models predict people's portfolio decisions. The basic reinforcement model predicts a learning process that relies solely on the portfolio's overall return, whereas the proposed extended reinforcement model also takes the risk and covariance of the investments into account. The experimental results illustrate that people reacted sensitively to different correlation structures of the investment alternatives, which was best predicted by the extended reinforcement model. The results illustrate that simple reinforcement learning is sufficient to detect correlation between investments.

Keywords: repeated portfolio decisions, reinforcement learning model, correlation

JEL Classification: C91, D83, G11

[*] Authors' E-mail: diao@econ.mpg.de (Linan Diao);

joerg.rieskamp@unibas.ch (Jörg Rieskamp)

We would like to thank Werner Güth, David Hugh-Jones, Oliver Kirchkamp, Rene Levinsky, Ondrej Rydval, Eva-Maria Steiger, tobias Uske or their useful comments and suggestions, and all the audience at the ESI, IMPRS brown bag seminars, the audience at 2010 Experimental Finance Conference at Gothenburg, and the audience at the 5th Nordic Behavioral and Experimental Economics at Helsinki. We would also like to thank all the student assistants in Max Planck Institute of Economics who helped us conduct the experiments.

1. Introduction

Economic theory states that financial investments should depend on the return, the risk, and the correlation of available assets. Formally this is specified by the mean-variance (MV) model of finance (see Markowitz 1952, 1959; Tobin, 1958). We examined experimentally whether people are sensitive to these core characteristics of investment alternatives when they repeatedly make investments and are provided with feedback that allows for learning. To model the observed learning process, we tested a basic learning model that only relies on the portfolios' returns against an extended learning model that also takes the risk and the correlation between investment assets into account.

Past research has already examined whether people react sensitively to the correlation of investment alternatives. Kroll, Levy, and Rapoport (1988a, 1988b) experimentally showed that participants were not very sensitive to correlations between stocks and deviated from the MV model and separation theorem by frequently switching between available stocks and by making investments depending on the alternatives' returns in preceding investment periods. According to the separation theorem, if borrowing and lending are not constrained and the rates of borrowing and lending are the same, the efficient risky frontier is reduced to a single optimal portfolio of risky stocks. Thus, the proportions invested in the risky stocks should be fixed. However, the experimental results violated the separation theorem's prediction. Canner, Mankiw, and Weil (1997) empirically showed that even financial advisors' recommendations deviate from the separation theorem. Experimentally, Lipe (1998) and Kallir and Sonsino (2009) showed that participants could perceive different levels of covariance, but their allocations were not significantly affected by changing the correlations. Finally, Hedesstroem, Svedsaeter, and Gaerling (2006) also experimentally showed that novice investors neglect covariation when diversifying across investment alternatives. However, most of this work illustrating that people do not react sensitively to the correlation between investment alternatives did not provide participants with much learning opportunity. If learning opportunity was provided the research often did not provide a learning theory to explain the observed learning processes.

Research on decision making in general has illustrated that learning can have a strong impact on people's behavior. For instance, Erev and Roth (1998) showed that experience can often lead to quick convergence to equilibrium predictions of economic games. Likewise, Bossaerts and Plott (2002, 2004) showed in experimental repeated asset markets that investors' portfolios steadily converged toward the prediction of the capital asset pricing model. To explain these effects it is very fruitful to model the observed behavior with learning models. Learning models can provide an explanation for how learning changes behavior and under which condition learning will or will not lead to convergence with economic theory.

Reinforcement learning models have been successfully used to predict and explain repeated decisions in many situations, including financial ones. Research has shown that reinforcement learning models can describe and predict people's behavior better than the equilibrium prediction. For instance, Erev and Roth (1998) demonstrated that even very simple reinforcement learning models can explain behavior in experimental games with unique equilibrium in mixed strategies better than the equilibrium prediction. Camerer and Ho (1999) suggested an experienced-weighted attraction (EWA) learning model which additionally to reinforcement learning (i.e., learning from experienced outcomes), assumes belief learning (i.e., learning from other players' behavior). The EWA model accurately described people's decision for various tasks such as constant-sum games with unique mixed-strategy equilibria, "median-action" coordination with multiple Pareto-ranked equilibria, and a dominance-solvable "*p*-beauty contest" game with a unique equilibrium. Feltovich (2000) compared reinforcement and belief-based models against each other and found that both models predicted behavior better than Nash equilibrium predictions, and the reinforcement model predicted the observed learning processes best. Hopkins (2002) also tested reinforcement learning and belief learning models against each other and pointed out that despite their conceptual differences they often make very similar predictions. Erev and Barron (2005) developed a more high order learning model assuming that the objects of reinforcement are cognitive strategies people apply to make choices between risky gambles.

In the finance area, Rieskamp, Busemeyer, and Laine (2003) examined learning processes when participants had to allocate resources to financial assets. They found that a learning model that assumes people only slightly modify their previous allocations based on feedback described investment decisions better than a learning model that assumes people try out a large variety of allocations. Rieskamp (2006) tested participants' hypothetical retirement saving decisions in two experiments. The results showed that learning models that incorporate recency effects described the observed decisions best. Kaustia and Knuepfer (2008) tested reinforcement learning in the Finnish financial market and Choi, Laibson, Madrian, and Metrick (2009) used a naïve reinforcement learning model to explain their findings: Empirically, investors who experienced particularly rewarding outcomes from 401(k) savings increased their 401(k) savings rate more than investors who had less rewarding experiences. Shimokawa, Suzuki, Misawa, and Okano (2009) developed a modified temporal-difference reinforcement learning model to describe decision-making processes for financial investments.

In sum, the reported studies show that people often change their investment decisions when they are provided with feedback about the decision outcomes. Furthermore, reinforcement models often provide a good description of the observed learning process and their predictions are often better than standard economic equilibrium predictions. Therefore, we tested two learning models against each other to see which would be better at predicting how people make repeated portfolio decisions when they receive feedback about the portfolios' returns.

What are the core assumptions of learning models? Most of the reinforcement learning models described above assume that the probability with which an alternative is chosen is an increasing function of the previously received reinforcements specified by an investment alternative's return and a decreasing function of the reinforcement for other alternatives. Surprisingly, none of the suggested learning models described above take the correlation between the alternatives' returns explicitly into account. This is astonishing when considering that the correlation between the alternatives' returns is a core component of portfolio theory. Learning models that ignore the correlation should not be able to explain people's repeated portfolio decisions if the decisions are affected by the correlation.

We examined to what extent correlation between alternatives influences people's repeated investment decisions. If people react sensitively to the correlation between investment alternatives, this will require learning models that incorporate a mechanism for the correlation between outcomes. Therefore, we conceived of a new reinforcement learning model that takes the risk of investment alternatives and the correlations of the alternatives' returns explicitly into account. We tested this new model against a standard reinforcement learning model to describe people's repeated investments.

To test the two models rigorously against each other we conducted two experimental studies. In the experiments, participants had to make portfolio decisions and were able to change their portfolios on the basis of received feedback. First we examined whether people's portfolio decisions were qualitatively in line with the predictions of the MV model and whether the investments were sensitive to the correlations between the investment alternatives. Second, we tested which of the two learning models predicts the observed investments better. In the first experiment, the participants were given detailed information about the mean and variance of the distributions from which the returns of the alternatives (i.e., stocks) were randomly drawn. However, no information about the correlation between the alternatives was provided. In the second experiment, participants were only informed about the distributions from which the stocks' returns were drawn without being given any information about the mean or variance of the distribution. Again no information about the correlation was provided. Thus, whereas in the first experiment the participants could use the feedback to learn about the correlation, in the second experiment they also had to learn the alternatives' average returns, the involved risk, and the correlation of alternatives' returns. The results show that participants' portfolios differed for three conditions with different correlations of the investment alternatives. Even in Study 2, where the participants had to learn the various characteristics of the investment alternatives, the sensitivity to the correlation was observed.

The rest of this article is organized as follows: Section 2 describes the two reinforcement learning models we used. Section 3 describes Study 1 and its results in detail. Section 4 describes Study 2 and its results in detail. We conclude with the general discussion in Section 5.

2. The Reinforcement Learning Models

The basic idea of reinforcement learning models is that decisions are a function of the learned expectancies of alternatives. The expectancies are updated by the feedback a decision maker receives. We examined to what extent experience changes people's investment decisions by testing two reinforcement models.

2.1 Basic Reinforcement Learning Model

Past work by, for instance, Lipe (1998) indicates that people might invest by depending only on the observed returns of the available investment alternatives, thereby ignoring the risk of alternatives and the correlation between alternatives. Therefore, in the basic reinforcement (BR) learning model, the expectancies of the alternatives are only a function of the alternatives' returns. Mathematically, the proportion of the available resource invested in the alternatives is defined as:

$$PR_t(i) = \frac{\exp\,(\theta \cdot q_t(i))}{\sum_{i=1}^{N} \exp\,(\theta \cdot q_t(i))} \tag{1}$$

where $q_t(i)$ represents the decision maker's subjective expectancies for alternative i in period t and θ is a sensitive parameter restricted to positive values that determines the extent to which an investment alternative that has a higher expectancy than the others will receive a larger percentage of the available resources. The initial expectancies $q_1(i)$ are assumed to be $1/N$ for all alternatives.

The expectancies are updated in each period by the received feedback, so that

$$q_t(i) = q_{t-1}(i) + \alpha\big(r_{t-1}(i) - q_{t-1}(i)\big), \tag{2}$$

where $r_{t-1}(i)$ corresponds to the observed return of alternative i and with $0 \leq \alpha \leq 1$ as the *learning rate* that determines the impact of the observed feedback on the subjective expectancies. The difference between $r_{t-1}(i)$ and $q_t(i)$ has also been called the prediction error (Sutton & Barto, 1998) and its neurological representation has recently received much attention in the neuroscience literature (e.g., Waelti, Dickinson, & Schultz, 2001). A large learning rate of, for instance, 1 implies that the updated expectancy will be equal to the observed return of the alternative, whereas a learning rate of 0 implies that the feedback has no impact on the updated expectancies.

2.2 Reinforcement Learning Representing Return, Risk, and Correlation

If people do not make investments only on the basis of the alternatives' returns but also take the alternatives' risks and correlations into account, then the standard reinforcement model needs to be extended. We propose a new learning model with an updating rule that additionally includes the variance and the covariance of the alternatives' returns. Mathematically the risk–return–covariance (RRC) reinforcement learning model uses the following modified updating rule:

$$q_t(i) = q_{t-1}(i) + \alpha\big(R_{t-1}(i) - q_{t-1}(i)\big) \qquad (3)$$

where $R_{t-1}(i)$ is defined as

$$R_{t-1}(i) = \beta_1 \cdot r_{t-1}(i) + \beta_2 \cdot \text{abs}\big(r_{t-1}(i) - r_{t-2}(i)\big) + \ \beta_3 \cdot \sum_{j=1, j\neq i}^{N-1} (\text{abs}(r_{t-1}(i) - r_{t-1}(j)) \qquad (4)$$

for alternatives i with varying returns; for alternatives with constant returns (e.g., risk-free bonds) $R_{t-1}(i)$ is simply defined as $r_{t-1}(i)$. The parameter β_1 ($0 < \beta_1 < 1$) is the *return weight* that determines how much an alternative's expectancy is influenced by the past returns of the alternative. The *risk weight* parameter β_2 ($-1 < \beta_2 < 1$) determines how much an alternative's expectancy is influenced by the variance of the alternative's returns, that is the alternative's risk. We used the absolute differences between consecutive trials to approximate the variance of the returns. Risk-averse persons should have a negative risk weight, whereas risk-loving people should have a positive weight. Finally, the *comparison weight* parameter β_3 ($0 < \beta_3 < 1$) determines how much an alternative's expectancy is influenced by the absolute differences between an alternative's returns in comparison to the other alternatives' returns. These absolute differences are affected by the correlation between the alternatives: In a situation with negative or no correlation between alternatives the absolute differences between the alternatives will on average be larger than in situations with positive correlations, even if the mean differences between the returns of the alternatives are identical. Stocks that are negative or not correlated will have higher expectancies and will receive a larger proportion of the portfolio than positively correlated alternatives. Therefore, the risk-free alternative's expectancy in a negative or zero correlation situation is smaller than in a positive correlation situation. For the three beta weights the constraint $\beta_1 + \text{abs}(\beta_2) + \beta_3 = 1$ is set. The extension of the RRC model relative to the basic reinforcement model allows the model to predict a learning process that is sensitive not only to an alternative's average returns but also to its variance and correlation to other alternatives. To the best of our knowledge the suggestion and the experimental test of a learning model that is able to take the variance and the correlation of alternatives' returns into account represents a novel contribution of the present article to the financial literature.

3. Study 1

In the first experiment the participants faced a portfolio task in which they repeatedly had to make portfolio decisions. Additionally at the end of the experiment, the participants performed a risk preference task that assessed their risk attitudes. In the portfolio task, participants made portfolio decisions by allocating endowments among three investment assets. For the first two investment alternatives the returns varied from period to period. In contrast, the third investment alternative was risk free and the returns were constant across all periods. In the experiment's instructions (see the Appendix), we used the neutral word "assets" for all three investment alternatives, but for unambiguous presentation in the following we will refer to the first two alternatives as "the risky stocks" and to the risk-free alternative as the "risk-free bond." The participants were told that the returns of the stocks were drawn from a normal distribution with a specific mean and variance. Thus, participants were informed about the average return of the stocks and the risk involved. However, no information about the correlation between the returns of the two stocks was provided.

The goal of Study 1 was to examine whether people would make different investments when the correlation between the two stocks differed, as predicted by economic theory. Because no information about the correlation was given, participants had to learn this on the basis of the received feedback. To examine people's sensitivity to the correlation of returns, for one group of participants, the returns of the two stocks were negatively correlated, whereas in the second group they were positively correlated, and in a third group they were independent. Furthermore, we explored overall learning effects by repeating the portfolio task three times.

With this experimental design, we tested whether participants' investments differed in the three conditions and to what extent they were in line with the MV model. Furthermore, we tested which of the two learning models did best in predicting participants' investments. Finally, we tested the prediction that participants who were assessed as more risk averse would invest a larger proportion of their endowment in the risk-free investment alternative.

3.1 Method

Participants

In the experiment 72 students (32 male and 40 female) with an average age of 24 years (*SD* = 2.4) participated. The whole experiment lasted around 45 min. Participants were students from the Friedrich Schiller University in Jena, Germany. The majority (66) studied economics, including business administration (42), economics (16), and economic mathematics (8). The remaining six students had other majors. For their participation, they received a show-up payment of 4 euros (ca. US $5). All additional payments depended on participants' performance: They were paid a percentage of the final portfolio value, resulting in average payments of 6.57 euros (*SD* = 3.48). In addition, participants received another payment for the risk preference task (0.10, 1.60, 2.00, or 3.85 euros).

Procedure

Participants were invited to the experimental laboratory of the Max Planck Institute of Economics in Jena, Germany. Participants sat in separate cabins and were not disturbed by others. They were paid and left the cabins at the end of the experiment. In the beginning of the experiment, participants had to fill out a pre-experiment questionnaire asking for age, gender, major, if they had taken finance or investment-related courses, if yes, how many, and if they had real investment experience. Then the main part of the experiment consisted of the portfolio decision task followed by the risk preference task.

Participants were randomly assigned to one of the three between-subjects conditions, with 24 participants each. For the portfolio task the participants had to invest their endowment in 30 periods in two stocks (A, B) and one bond (C). Stock A's returns were drawn from an i.i.d. normal distribution of $N(6\%, 6\%)$ and stock B's returns were drawn from an i.i.d. normal distribution of $N(9\%, 14\%)$. Thus, stock B offered on average a higher return than stock A but also involved greater risk than A. The risk-free bond C had a constant return of 2%. To prevent very extreme returns, we truncated the normal distribution at the positive and negative end by three times the standard deviations of the distributions. This implies that stock A's and stock B's returns were restricted to between -12% and 24%, and -33% and 51%, respectively. At the beginning of each period, participants invested all their accumulated wealth in three investment alternatives. Then the actual returns of the two stocks were revealed. At the end of each period, the gain/loss and the final wealth of this period were calculated. The portfolio task was repeated three times, to explore whether due to learning, participants would change their investments in the later tasks. In the three experimental between-subjects conditions the correlation between the returns of stock A and B differed: In the positive, independent, and negative correlation condition the correlation between the two stocks was 0.80, zero, and -.80, respectively.

At the beginning of the portfolio task, participants got an initial endowment of 10,000 experimental currency units (ECUs) for investment. In each of the 30 periods of the task participants were able to invest their resource defined as their endowment plus any potential gain or loss they made in the previous periods. They decided what percentage of their resource they wanted to invest in each of the three investment alternatives. In the first period they had to make a decision about their endowment of 10,000 ECUs, whereas in the following periods they made decisions about the accumulated resource. To make a decision they had to move a slide rule (see Figure 1 for a screenshot) to indicate their investment in stock A and stock B. The remaining part of the resource was invested in bond C. To finalize their decision, which was represented by a pie chart, they pressed another button. After they made their investments they were informed about the investment returns by means of a table and a graphic. The table was visible during the whole portfolio task and provided information about the rate of return of each investment alternative, the gain or loss the participants made with this alternative, the total gain or loss for the whole portfolio, and the value of the overall resource. This information was given for the present period, but also for all previous periods (see Figure 1). The total value of the portfolio changed as a function of the three alternatives' returns and the percentage invested in the alternatives. At the end of the task the participants were shown the accumulated resource in ECUs and additionally the transformed value in euros using an exchange rate of 10,000 ECUs = 1 euro. After the end of the first portfolio task a new task started. Importantly, the participants were informed at the very beginning of the experiment that they would only receive real payment for one of the three tasks, which was randomly selected at the end of the experiment.

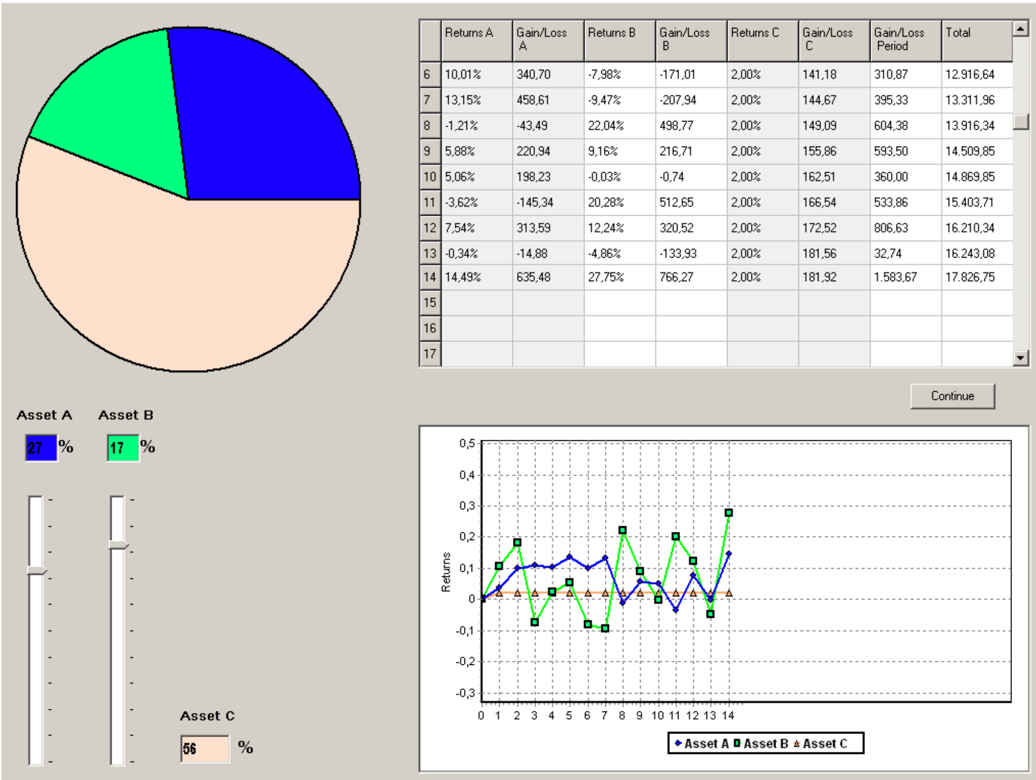


Figure 1. Screenshot of the experimental software used in Studies 1 and 2. Participants made their decisions by moving the tab up and down a slide rule below each investment alternative. The pie chart reflected their investment decisions.

After the participants finished the third portfolio task they proceeded with the risk preference task. This task was almost identical to that used by Holt and Laury (2002) to assess people's risk attitudes. The task is based on 10 choices between two pairs of gambles, presented in Table 1. Gamble A always led to an outcome of either 2.00 euros or 1.60 euros, whereas gamble B always led to a payoff of 3.85 euros or 0.10 euros. For the 10 choices the probability of receiving the two outcomes was varied, starting with a probability of 0.10 of receiving the larger of the two payoffs of each gamble. This probability was increased by 0.10 for each consecutive pair of gamble, so that for the last choice the larger payoff was guaranteed with a probability of 1.

Because the payoffs for gamble A are less variable than the payoffs for gamble B, gamble B is in general riskier than gamble A. However, by increasing the probability of obtaining the larger payoff for both gambles the expected value of gamble B increases in comparison to gamble A, so that it becomes increasingly attractive. Therefore a risk-neural decision maker would choose gamble A for the first four pairs before switching to gamble B for the remaining six pairs. Participants were told that after they completed the risk preference task, one pair of gambles would be selected randomly, and using the gamble they had chosen, the gamble would be played out in front of them and the realized payoff would be paid to them. The Appendix provides the full instructions for the portfolio task and the risk preference task.

Table 1. Risk preference task

	Pairs of gambles[a]		Participants choosing gamble B (%)	
Pair	Gamble A	Gamble B	Study 1	Study 2
1	0.10 of €2.00, 0.90 of €1.60	0.10 of €3.85, 0.90 of €0.10	4.2	2.8
2	0.20 of €2.00, 0.80 of €1.60	0.20 of €3.85, 0.80 of €0.10	6.9	5.6
3	0.30 of €2.00, 0.70 of €1.60	0.30 of €3.85, 0.70 of €0.10	6.9	4.2
4	0.40 of €2.00, 0.60 of €1.60	0.40 of €3.85, 0.60 of €0.10	13.9	5.6
5	0.50 of €2.00, 0.50 of €1.60	0.50 of €3.85, 0.50 of €0.10	26.4	22.2
6	0.60 of €2.00, 0.40 of €1.60	0.60 of €3.85, 0.40 of €0.10	33.3	37.5
7	0.70 of €2.00, 0.30 of €1.60	0.70 of €3.85, 0.30 of €0.10	62.5	69.4
8	0.80 of €2.00, 0.20 of €1.60	0.80 of €3.85, 0.20 of €0.10	84.7	87.5
9	0.90 of €2.00, 0.10 of €1.60	0.90 of €3.85, 0.10 of €0.10	88.9	95.8
10	1.00 of €2.00, 0.00 of €1.60	1.00 of €3.85, 0.00 of €0.10	94.4	95.8

[a] The first and third value for each alternative represent the probability with which the two payoffs occur that are represented by the third and fourth value.

3.2 Results

We analyzed the results in three steps. First, we examined people's risk preferences and related them to their investment decisions. Second, we examined whether participants' investments differed for the three conditions with different correlations between the stocks' returns and compared the portfolios of the participants with the predictions of the MV model. Finally, we examined the learning effects and tested the two learning models against each other.

Risk attitude

Table 1 shows the results of the risk preference task. Consistent with the logic of the task, only a few participants chose the risky gamble for the first pair of gambles and the safer gamble for the last pair of gambles. On average, the choice percentages of the risky gamble increased monotonically across the 10 pairs in Study 1. We first analyzed whether the participants from the three correlation conditions had different risk attitudes. Since we are not interested in the absolute risk preference, we simply used the frequency with which the safer gamble was chosen by the participants to measure their risk preferences. The more risk averse, the higher the frequency of choosing the safer gamble. The average number choosing the safer gamble A in the positive correlation condition was 5.92 (*SD* = 1.97), in the negative correlation condition it was 5.29 (*SD* = 1.33) and in the independent condition 6.13 (*SD* = 1.85). These differences were not significant according to an analysis of variance (ANOVA) with the three correlation conditions as a between-subjects factor ($F = 1.49$, $p = 0.233$).

Table 2. Kendall's tau rank correlation between risk preference and investment in the risk-free bond (with *p*-value)

	Correlation Condition			Average
	Positive	Negative	Zero	
	0.09	0.18	0.27	0.20
Study 1	(0.554)	(0.26)	(0.085)	(0.027)
	0.44	0.33	0.13	0. 26
Study 2	(0.005)	(0.039)	(0.413)	(0.003)

If people's investments are correlated with their risk attitudes, then the more risk averse a person is, the more that person should invest in the risk-free bond. Table 2 lists the Kendall's tau rank correlation between risk preference and the percentage of the portfolio that was invested in the risk-free bond. Although we observed on average a positive correlation of 0.20 between the investment in the risk-free bond and the risk attitude, the correlation is not very strong and differs for the three experimental conditions.

The MV model's prediction and investment decisions

Table 3 lists the average portfolios of the participants for the three correlation conditions of Study 1. Consistent with the core principle of portfolio theory participants invested the least in the risk-free bond and the most in the riskier stock (stock B) when the two stocks were negatively correlated. Likewise, participants invested the most in the risk-free bond when the two stocks were positively correlated, although their investment in the risky stock (stock B) was also quite high. In all treatments, participants increased their investment in the riskier stock (stock B) that offered on average a higher return and decreased their investment in the risk-free bond (bond C) over the three rounds of the portfolio task.

Table 3. Investments in the three assets across the three correlation conditions of Study 1

	Positive			Zero			Negative		
	Stock A	Stock B	Bond C	Stock A	Stock B	Bond C	Stock A	Stock B	Bond C
Round 1	0.25	0.45	0.30	0.42	0.36	0.22	0.34	0.45	0.21
Round 2	0.22	0.48	0.30	0.40	0.40	0.19	0.36	0.50	0.14
Round 3	0.24	0.50	0.26	0.39	0.43	0.18	0.34	0.52	0.14
Overall	0.24	0.47	0.29	0.40	0.40	0.20	0.35	0.49	0.16

Note. Stock A led to an average return of 6% (*SD* = 6%), B led to an average return of 9% (*SD* = 14%), and C led to a fixed return of 2%.

We performed two ANOVAs with the proportion invested in the risk-free bond C and the risky stock B, respectively, as the dependent variable and the correlation as a between-subjects factor and the repetition as a within-subject factor (Table 4). The results show that the investment in the risk-free bond C differed among the three correlation conditions and among the three repetitions of the portfolio task. Post hoc contrasts revealed that investment in bond C differed among all three correlation conditions; investment in bond C differed between the positive and negative correlation conditions ($t = 22.53$, $p < .001$); between positive and zero correlation conditions ($t = 21.70$, $p < .001$), and between the negative and zero correlation conditions ($t = 7.72$, $p < .001$). Likewise investment in bond C decreased significantly from the first to the second ($t = 3.67$, $p < .001$) and from the second to the third ($t = 1.98$, $p = .024$) repetition of the portfolio task. The observed interaction is presumably due to no decrease of investment in bond C between the first and second repetition in the positive correlation condition and between the second and third repetition in the negative correlation condition.

Table 4. Two-factor repeated measurement ANOVA of Study 1

	Source of variation	*SS*	*df*	*MS*	*F*	*p*-value
Stock B	Rounds	0.20	2	0.10	107.41	< .001
	Conditions	0.44	2	0.22	242.62	< .001
	Interaction	0.00	4	0.00	0.96	.430
	Within	0.24	261	0.00		
	Total	0.88	269			
Bond C	Rounds	0.13	2	0.06	80.97	< .001
	Conditions	0.72	2	0.36	455.76	< .001
	Interaction	0.04	4	0.01	13.49	< .001
	Within	0.21	261	0.00		
	Total	1.10	269			

Investment in the riskier stock B also differed among the three conditions. Post hoc contrasts revealed that investment differed in the positive and zero correlation conditions ($t = 19.16$, $p < .001$); the positive and negative correlation conditions ($t = -4.08$, $p < .001$), and the negative and zero correlation conditions ($t = 18.63$, $p < .001$). Likewise investment in stock B increased significantly from the first to the second ($t = 5.49$, $p < .001$), and from the second to the third ($t = 3.18$, $p < 0.001$) repetition of the portfolio task. We did not observe a significant interaction between conditions and repetitions. To compare the average results with the prediction of the MV model more specifically, we first determined the efficient frontier for all three conditions. In addition, Figure 2 shows the expected returns and standard deviations for the average portfolios in the three correlation conditions and the three repetitions of the portfolio task. Contrary to the MV model prediction, it becomes clear that participants' portfolios are on average not at the frontier lines; that is, participants could have constructed portfolios with a higher average return without increasing their risk. This situation was also not improved by learning, because the results also hold for the third repetition of the portfolio task. However, in line with the MV model we observed that investment in the risk-free bond was the largest in the condition with positively correlated stocks and smallest in the condition with negatively correlated stocks. Likewise, consistent with the MV model, investment in the riskier stock B was largest in the condition with negatively correlated stocks. However, contrary to the MV model, investment in the riskier stock was not smallest in the condition with positively correlated stocks.

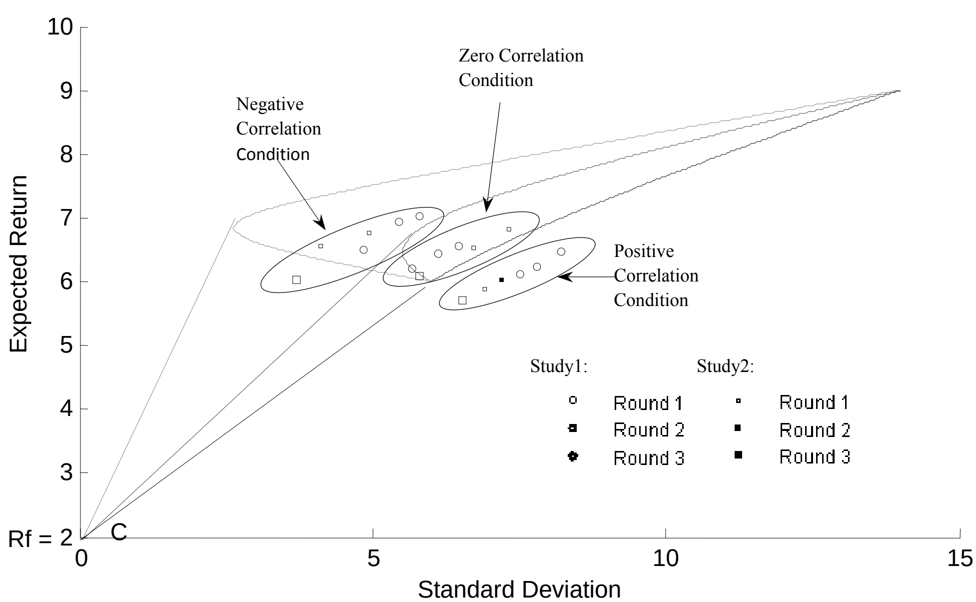


Figure 2. Efficient frontier for the three correlation conditions and average means and standard deviations of the actual portfolios by condition and round for the two studies. Rf is the return of risk-free bond.

The learning models

We tested the two learning models against each other and against a simple baseline model. According to the baseline model each person invests an equal share of his or her resource in the three investment alternatives. The baseline model implements the idea of the so-called 1/*N* investment strategy (see Benartzi & Thaler, 2001), according to which people diversify their portfolio in a naïve way. The free parameters of the learning models were estimated for each individual using the data of the first two repetitions of the portfolio task. To estimate the parameters we minimized the average Euclidean distance between the predicted and the observed portfolios of the participants as a goodness-of-fit criterion defined as:

$$D_{\text{Fit}} = \min\left[1/N\sum_{t=1}^{N}\sqrt{\sum_{i=1}^{3}(PR_{i,t}^{RL} - PR_{i,t})^2}\right] \quad (5)$$

where $PR_{i,t}$. are the proportions of the resource the participants invested in the three investment alternatives *i* for period *t* for all periods *N* and PR^{RL} are the proportions invested in the three investment alternatives predicted by the learning models. To optimize the parameters for each participant, reasonable parameter values were first selected by a grid-search technique, and thereafter the best-fitting grid values were used as a starting point for a subsequent optimization using the simplex algorithm (Nelder & Mead, 1965) as implemented in Matlab.

The learning models were tested against each other and against the baseline model by their predictions for the third repetition of the portfolio task. The predictions were generated using the estimated parameters from the first and second repetition of the portfolio task, so that no data fitting was involved for the critical third repetition. With this cross-validation generalization test of the models the complexities of the models (i.e., their free parameters) are taken into account.

Table 5 shows the goodness of fit of the three models for the first two repetitions of the portfolio task and for the crucial last repetition for which the models' independent predictions were tested. The table clearly illustrates that both learning models predict the behavior much better than the simple baseline model. Thus, the participants made investments that did depend on the different returns and risks of the investment alternatives. Likewise, the participants changed their investments over the course of the 30 periods of the task, which cannot be explained by any model that predicts a constant allocation.

Table 5. Goodness-of-fit and predicted values for different reinforcement learning models and the 1/*N* model of Study 1

Model	Correlation condition					
	Positive		Zero		Negative	
	Fit first 60 periods	Predict last 30 periods	Fit first 60 periods	Predict last 30 periods	Fit first 60 periods	Predict last 30 periods
1/*N*	635.74	327.90	598.51	304.71	565.23	312.24
BR (α,β_1,θ)	404.56	210.29	468.80	231.38	396.59	189.52
RL1 (α, β_1,β_2,θ)	303.79	193.90	362.06	225.18	307.30	186.77
RL2 (α, β_1,β_3,θ)	389.36	200.83	395.11	196.30	337.88	169.97
RRC (α, $\beta_1,\beta_2,\beta_3,\theta$)	292.47	171.82	256.29	174.23	244.33	159.85

Note. BR, basic reinforcement learning model, with parameters α,β_1,θ; RL1, reinforcement learning model with parameters α, β_1,β_2,θ ; RL2, reinforcement learning model with parameters α, β_1,β_3,θ; RRC, risk–return–covariance reinforcement learning model with all parameters.

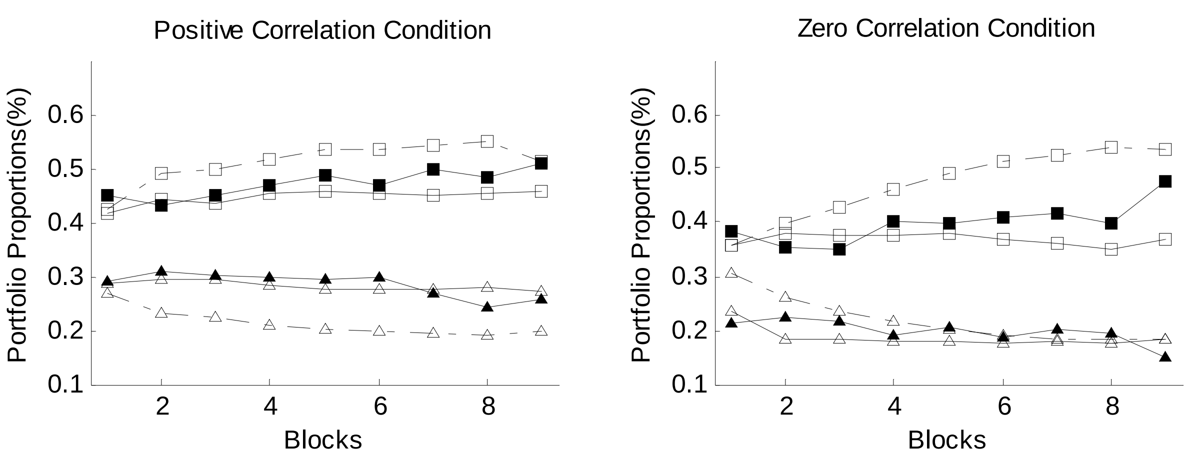
The BR learning model is nested within the RRC model, so by necessity the more complex RRC model will be better than or as good as the basic model in describing the data of the first two repetitions of the task, which can be seen in Table 5. What is more crucial is how good the models are in predicting independent data. Here the suggested RRC model was better than the simpler learning model and the baseline model in predicting the behavior in all three experimental conditions. Table 5 also shows the results of learning models that only include the risk (RL1) or the covariance component (RL2) of the RRC model (by setting the corresponding parameter in the RRC model to a value of zero). The results show that both components improve the predictive accuracy of the RRC model, where the covariance component seems to have the biggest positive impact on improvement.

Table 6 lists the average estimated parameter values for the two reinforcement models. These parameters have to be interpreted cautiously, because the different components of the models might interact with each other, and the absolute values of the parameters might also depend on the scale of the returns, the variance, and the covariance of the investment alternatives. Nevertheless, the learning rate, α, is given a moderate weight in line with previous modeling results of learning. Interestingly the RRC model shows that participants seemed to consider the return, the variance, and the covariance of the alternatives. Whereas the return is given a similar weight across all three experimental conditions, the weight given to the variance of the returns is lower in the condition with no correlation between the alternatives. The weight given to the variance substantially indicated the risk-seeking and risk-averse behavior of the participants. The weight given to the covariance was relatively high and was estimated to be lower in the condition with positive correlation between the alternatives.

Table 6. Optimal Estimated Parameters of the Two Learning Models of Study2

			α	β_1	β_2	abs(β_2)	β_3	θ
BR	Positive correlation	Mean	0.44	1.00	-	-	-	8.65
		SD	0.48	0.00	-	-	-	25.78
	Zero correlation	Mean	0.35	1.00	-	-	-	5.17
		SD	0.47	0.00	-	-	-	1.66
	Negative correlation	Mean	0.25	1.00	-	-	-	1.66
		SD	0.41	0.00	-	-	-	5.67
RRC	Positive correlation	Mean	0.20	0.16	0.01	0.64	0.20	4.04
		SD	0.27	0.25	0.76	0.39	0.30	9.34
	Zero correlation	Mean	0.19	0.13	-0.26	0.38	0.49	4.26
		SD	0.24	0.20	0.46	0.36	0.42	9.44
	Negative correlation	Mean	0.14	0.14	0.05	0.44	0.42	2.29
		SD	0.19	0.16	0.58	0.38	0.37	5.45

Figure 3 shows the mean portfolios across the 90 periods of the experiment partitioned into nine blocks with 10 periods each. The figure illustrates that in all three conditions and especially in the negative correlation condition, participants change their portfolios gradually. In the condition with positively correlated stocks, participants apparently realized the return disadvantage of stock A compared to stock B and no diversification advantage, so that they increased their investment in stock B. When the two stocks were not correlated participants invested an equal share in both stocks, which slightly increased over the course of the task. In the condition with the negatively correlated stocks, the participants apparently realized that diversification would reduce the risk of their investment, so that their investment in the risk-free bond decreased and their investment in the risky stock B increased over time. These dynamics were better predicted by the RRC model. The BR model overestimates the proportion invested in the riskier stock B and underestimates the investment in the safer stock A, because it solely focuses on the return advantage of stock B. In contrast, the prediction of the RRC model seems to be more in line with the observed behavior as it takes the risk of the stocks into account.



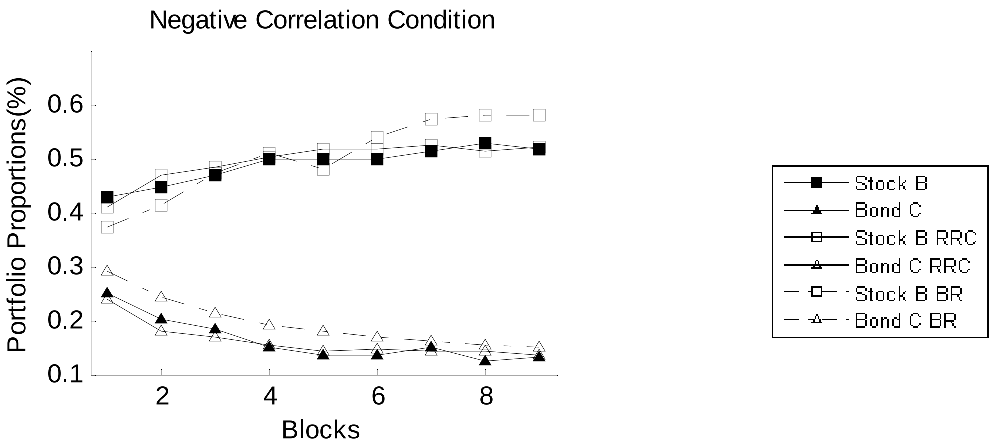


Figure 3. Participants' average portfolios represented by the investments in stock B and bond C of the three conditions of Study 1, compared with the average predictions of the basic reinforcement (BR) learning model and the risk–return–covariance (RRC) model. The *x*-axes represent the 90 periods of the whole experiment partitioned into nine blocks of 10 periods each.

3.3 Summary of Study 1

Study 1 illustrates that the participants made portfolio decisions that took the return and risk of the investment alternatives into account. Interestingly and in contrast to the results of past studies (see, e.g., Canner et al., 1997; Hedesstroem et al., 2006; Kallir & Sonsino, 2009; Kroll et al., 1988a, 1988b; Lipe, 1998; Svedsaeter & Gaerling, 2006), in our experiment the participants were influenced by the correlation between the stocks' returns, which were unknown to the participants at the beginning of the experiment. In line with the general principles of portfolio theory, investment in the risk-free bond was highest when the two stocks were positively correlated and did not offer a reduction of risk by diversification. In contrast, when the stocks were negatively correlated investment in the risk-free bond was lowest, exploiting the risk reduction of diversification. Likewise investment in the riskier stock B was highest in the condition with the negative correlation between stocks. In sum, these results are all in line with the MV model. However, the results did not match the predictions of the MV model completely: Investment in stock B was relatively high in the condition with positive correlation between the stocks and the average portfolios were not at the efficient frontier.

The results also illustrate that the participants changed their portfolios during the 30 periods of each portfolio task. These dynamics cannot be explained by a model that predicts a constant portfolio over the whole investment period such as the baseline model implementing the 1/*N* strategy. In contrast, the two learning models did better in predicting the observed dynamics. In particular, the RRC model predicted the behavior best. This is an important finding as it illustrates that people do not only make their investments on the basis of the alternatives' returns, as the BR model predicts, but take the risk of the alternatives into account. It appears that the two additional components of the RRC model might successfully describe how people are influenced by the risk of the investment alternatives.

4. Study 2

One might argue that the surprising finding of Study 1 that participants reacted very sensitively to the correlation between the stocks' returns might be due to the detailed information they got about the alternatives. Because the participants were told the average and the variance of the alternatives' returns, they could focus on detecting the correlation between the stocks. In Study 2 we tested whether people are also sensitive to the correlation between stocks' returns when they are not provided with detailed information about the investment alternatives.

In Study 2 the portfolio task was the same as in Study 1. The participants made portfolio decisions by allocating endowments for a fixed period of time among three investment alternatives. Participants were told that the three investment alternatives represented financial assets. Furthermore, the participants were told that the returns of two assets were drawn from a normal distribution with a specific mean and variance. However, participants were not informed about the mean and the variance of the distribution, which were identical to Study 1. Again no information about the correlation between the two stocks was provided. Participants were informed that the third asset provided a constant return of 2%. To examine participants' sensitivity to the correlation of returns, again there were three experimental conditions with positive, negative, and zero correlation between the stocks' returns, identical to Study 1. After the participants completed the portfolio task they continued the experiment with the risk preference task as in Study 1. Additionally, at the very end of the experiment the participants were asked to estimate the average returns, the 95% intervals of the returns of the two stocks, and the correlation between the two stocks.

Again we tested with Study 2 whether participants' investments differed in the three correlation conditions and to what extent their investments were in line with the MV model under the condition of reduced a priori information about the investment situation. Furthermore, we tested which of the two learning models predicts people's repeated investments best.

4.1 Methods

Participants

In Study 2, 72 students (42 male and 30 female) with an average age of 23 years (*SD*=2.5) participated in the experiment. The whole experiment lasted around 45 min. Participants were students from Friedrich Schiller University, Jena. The majority (64) studied economics, including business administration, economics, and economic mathematics, five were from mathematics, and the remaining three students had other majors. For their participation, they received a show-up payment of 4 euros (ca. US $5). All additional payments depended on the participants' performance: They were paid a percentage of the final portfolio value, resulting in an average payment of 6.17 euros (*SD* = 2.59), not including the show-up payment and payment for the risk preference task. The payment for the risk preference task was 0.10, 1.60, 2.00, or 3.85 euros.

Procedure

Participants were invited to the experimental laboratory of the Max Planck Institute of Economics in Jena, Germany. Participants sat in separate cabins and were not disturbed by others. They were paid and left the cabins at the end of the experiment. In the beginning of the experiment, participants had to fill out a pre-experiment questionnaire asking for age, gender, major, if they had taken finance or investment-related courses, if yes, how many, and if they had real investment experience. Then the portfolio task started, followed by the risk preference task. Study 2 was conducted identically to Study 1, with the only difference being that no information about the mean and variance of the returns for the risky stocks were provided. Furthermore, to see whether the participants had a good understanding of the distribution of the two stocks, we asked participants to estimate the average returns, the 95% intervals of returns, and the correlations of the two stocks at the very end of the experiment.

4.2 Results

We analyzed our results in four steps. First, we analyzed participants' risk preferences and related them to their investment decisions. Second, we examined whether participants' investments differed for the three correlation conditions and compared the investments with the MV model. Third, we compared the investments in the two experiments. Finally, we examined the learning effects and tested the two learning models against each other.

Risk attitude

The results of the risk preference task are shown in Table 1. Consistent with the logic of the task, in Study 2, only a few participants chose the risky gamble (Gamble B) for the first pair of gambles and the safer gamble (Gamble A) for the last pair of gambles. With the exception of pair 2, the choice percentages of the risky gamble increased monotonically across the 10 pairs. We tested whether the participants from the three correlation conditions had different risk attitudes. Table 1 shows the average frequency and standard deviation of how often the safer gamble was chosen. We used this measurement as the approximation of the risk preference. The more risk averse, the higher the frequency of choosing the safer gamble should be. The average number choosing the safer gamble was 5.42 (SD = 2.04) in the condition with positive correlation, 6.21 (SD = 1.50) in the negative correlation condition, and 5.58 (SD = 1.32) in the zero correlation condition, which did not differ significantly according to an ANOVA test (F = 1.54, p = 0.222).

According to economic principles, more risk-averse participants should invest more in the risk-free bond C. The Kendall's tau rank correlations between the risk preference and the percentage of the portfolio invested in the risk-free bond C are listed in Table 2. Overall, we observed a significant positive correlation 0.26 (p = .003). However, the correlation differed among the three experimental conditions and was not very strong in the condition with no correlation between the risky alternatives.

Comparison with the MV model's prediction

The average portfolios across participants for each condition and round are shown in Table 7. Consistent with the MV model, participants invested the most in the risk-free bond C and the least in the riskier stock B when the two stocks were positively correlated. Likewise, participants invested the least in bond C when the two stocks were negatively correlated, although their investment in the riskier stock (stock B) was lower than when the two stocks were independent.

Table 7. Investments in the three assets across the three correlation conditions of Study 2

	Correlation condition								
	Positive			Zero			Negative		
	Stock A	Stock B	Bond C	Stock A	Stock B	Bond C	Stock A	Stock B	Bond C
Round 1	0.31	0.35	0.33	0.35	0.39	0.27	0.36	0.37	0.27
Round 2	0.29	0.39	0.32	0.33	0.46	0.21	0.41	0.42	0.18
Round 3	0.29	0.41	0.30	0.32	0.51	0.18	0.37	0.47	0.16
Overall	0.30	0.38	0.32	0.33	0.45	0.22	0.38	0.42	0.20

Note. Stock A led to an average return of 6% (*SD* = 6%), B led to an average return of 9% (*SD* = 14%), and C led to a fixed return of 2%.

We performed two ANOVAs with the proportion invested in the risk-free bond C and the riskier stock B, respectively, as the dependent variable and the correlations as a between-subjects factor and the repetition as the second, within-subject factor (Table 8). The results show that investment in the risk-free bond differed among the three correlation conditions and among the three repetitions of the portfolio task. Post hoc contrasts revealed that the investment in bond C differed among all three correlation conditions. Paired two-sample *t*-tests showed that the investment in bond C differed between the positive and negative conditions ($t = 15.89$, $p < .001$); the positive and independent conditions ($t = 15.37$, $p < .001$); and the negative and independent conditions ($t = -2.64$, $p = .010$). Likewise, the investment in bond C decreased significantly from the first to the second ($t = -5.79$, $p < 0.001$), and from the second to the third ($t = -2.10$, $p < 0.02$) repetition of the portfolio task. The observed interaction is presumably due to no substantial decrease of investment in bond C between the first and second repetition in the positive correlation condition and a similar investment in bond C in the first repetition of the zero and negative correlation conditions.

Table 8. Two factor repeated measurement ANOVA of Study 2

	Source of variation	*SS*	*df*	*MS*	*F*	*P*-value
Stock B	Rounds	0.38	2	0.19	126.18	< .001
	Conditions	0.20	2	0.10	65.71	< .001
	Interaction	0.03	4	0.01	4.934	< .001
	Within	0.39	261	0.00		
	Total	1.00	269			
Bond C	Rounds	0.28	2	0.14	94.89	< .001
	Conditions	0.73	2	0.36	250.19	< .001
	Interaction	0.07	4	0.02	11.21	< .001
	Within	0.38	261	0.00		
	Total	1.45	269			

The investment in the riskier stock B also differed among the three conditions. Post hoc contrasts (paired two-sample *t*-tests) revealed a difference between the positive and negative conditions ($t = -5.81$; $p < .001$), the positive and independent conditions ($t = -9.62$; $p < .001$), and the negative and independent conditions ($t = -6.00$; $p < .001$). Likewise, the investment in stock B increased significantly from the first to the second ($t = -7.69$, $p < 0.001$) and from the second to the third ($t = -5.26$, $p < 0.001$) repetition of the portfolio task. The observed interaction was presumably due to no substantial increase of investment in stock B between the second and third repetition in the positive correlation situation.

To compare the average results with the prediction of the MV model more specifically, Figure 2 shows the expected return and standard deviation for the average portfolios in the three experimental conditions and the three repetitions of the portfolio task in Study 2. Contrary to the MV model prediction, it is clear that participants' portfolios are not at the frontier lines. That is, participants could have constructed portfolios with a higher expected return without increasing their risk. The situation was not improved by learning, because the results also hold for the third repetition. However, in line with the MV model, we observed that investment in the risk-free bond was largest in the condition with positively correlated stocks and smallest in the condition with negatively correlated stocks. Likewise, consistent with the MV model, investment in the riskier stock B was smallest in the condition with the positively correlated stocks. However, contrary to the MV model, investment in the riskier stock B in the condition with the negatively correlated stocks was not the largest.

Comparison of the two studies

Although the two studies were conducted at different times and the participants were not randomly assigned to the two experiments, a comparison across the two experiments is interesting. Because both experiments were conducted almost identically, with the only difference being the information about the stocks' average returns and variances, differences in participants' behavior are most likely due to the different information provided.

We compared the average portfolios across the two experiments, to examine whether providing information about the average returns and risk had an impact on the participants' behavior. Table 9 shows the average portfolios of the two experiments, as well as the first and last 10 periods' average portfolios.

Table 9. Average portfolios in Studies 1 and 2

		Correlation condition								
		Positive			Zero			Negative		
		Stock A	Stock B	Bond C	Stock A	Stock B	Bond C	Stock A	Stock B	Bond C
First 10 periods	Study1	0.26	0.45	0.29	0.40	0.38	0.21	0.32	0.43	0.25
	Study2	0.32	0.36	0.32	0.37	0.35	0.28	0.33	0.36	0.31
Last 10 periods	Study1	0.23	0.51	0.26	0.37	0.47	0.15	0.35	0.52	0.13
	Study2	0.25	0.44	0.31	0.32	0.52	0.16	0.36	0.47	0.17
All	Study1	0.24	0.47	0.29	0.40	0.40	0.20	0.35	0.49	0.16
	Study2	0.30	0.38	0.32	0.33	0.45	0.22	0.38	0.42	0.20

When comparing the average portfolios of two experiments, first, we looked at investments in bond C. In all three experimental conditions, the investment in bond C was larger in Study 2 than in Study 1. In the beginning of Study 2, when the participants had no information about the returns of the two stocks, they invested a larger proportion of their resource in the safe bond C compared to Study 1. However, participants obviously learned through the experiment that the other alternatives offered on average a higher return and increased their investment in the stocks. However, even in the last 10 periods investment in bond C was still larger than in Study 1.

We performed a single factor ANOVA analysis of Study 2 with participants' investments in the first periods in stock A as dependent variable, the three correlation conditions and 1/*N* as a between-group factor (1/N is a hypothetical group with 24 subjects, which are assumed to use the 1/N strategy, i.e., the investment proportion is 1/3), which revealed no significant difference of investment in the three conditions ($F = 0.22$, $p = 0.88$) and they had no significant difference with the 1/*N* strategy. We got the same results for stock B ($F = 1.08$, $p = 0.36$). Not surprisingly without any information about the stocks' returns the investments did not differ among the three conditions. Apparently participants commenced by investing an equal proportion of their resource in the three investment alternatives following a 1/*N* investment strategy. In contrast, in Study 1 in which the participants were provided with information about the returns of the alternatives, a higher proportion of the resource was invested in stock B with the highest average return (despite its higher risk). In Study 2 participants increased their investment in the high-return stock B due to learning (although the final investment in stock B was still lower compared to Study 1).

At the very end of Study 2 we asked participants to estimate the average returns, the 95% return intervals, and the correlations between the two stocks. Table 10 lists participants' estimations.[1] Surprisingly, the estimations of returns were all higher than the real average returns of the stocks. The average return of the risky stock B, for instance, was estimated to be about 14% compared to the real average return of only 9%. Likewise the return for the safer stock A was estimated to be around 11% compared to the real 6% return.

[1] We excluded two participants with very extreme estimates from this analysis as outliers (one participant from the positive correlation condition and one participant from the zero correlation condition).

Table 10. Participants' estimations of the two stock returns and variability in Study 2

Correlation	Return stock A	Return stock B	95% intervals stock A		95% intervals stock B		Real Correlation	Estimated Correlation
Real	0.06	0.09	-0.12	0.24	-0.33	0.51	-	-
Positive	0.10	0.14	0.00	0.20	-0.04	0.25	0.80	0.25
(*SD*)	(0.07)	(0.09)	(0.06)	(0.13)	(0.10)	(0.14)		(0.35)
Negative	0.14	0.14	0.02	0.23	-0.07	0.31	-0.80	-0.07
(*SD*)	(0.19)	(0.12)	(0.16)	(0.26)	(0.20)	(0.23)		(0.40)
Zero	0.10	0.14	0.00	0.25	-0.02	0.31	0	-0.03
(*SD*)	(0.11)	(0.16)	(0.19)	(0.21)	(0.20)	(0.22)		(0.44)

Despite these overestimates of the average returns, the variability of the returns was underestimated. Finally the participants were asked to estimate the correlation between the returns of the two stocks. In fact, the participants were on average aware of the differences among the three conditions, with the highest positive correlation estimate being given in the positive correlation condition and the lowest negative correlation estimate being given in the negative correlation condition. In the positive correlation condition, 15 participants (out of 24) estimated that the two stocks were positively correlated, whereas in the negative correlation condition, 10 participants (out of 24) estimated the two stocks being negatively correlated. In the zero correlation condition, only 5 participants judged the returns of the two stocks to be independent. Furthermore, the average range estimates in the positive and negative correlation conditions were largely underestimated.

The learning models

Again we tested the BR learning model and the extended RRC learning model against each other in predicting participants' investments. Likewise we estimated the parameters using the first two repetitions of the portfolio task and predicted the observed learning process in the last repetition. The two learning models were compared against each other and against the baseline model (1/*N* strategy). Table 11 lists the goodness-of-fit for the first two repetitions of the portfolio task and the models' independent predictions for the last repetition. Table 11 illustrates that all reinforcement learning models predicted participants' portfolio behavior much better than the simple baseline model.

Table 11. Goodness-of-fit and predicted values for different reinforcement learning models and the 1/*N* model of Study 2

Model	Correlation condition					
	Positive		Zero		Negative	
	Fit first 60 periods	Predict last 30 periods	Fit first 60 periods	Predict last 30 periods	Fit first 60 periods	Predict last 30 periods
1/*N*	553.29	306.06	558.58	313.22	535.77	298.64
RL (α,β_1,θ)	442.10	253.42	445.14	220.20	416.86	252.40
RL1 ($\alpha, \beta_1,\beta_3,\theta$)	421.29	247.37	387.55	200.23	313.80	162.13
RL2 ($\alpha, \beta_1,\beta_2,\theta$)	359.34	245.55	394.62	217.28	338.54	244.48
RRC ($\alpha, \beta_1,\beta_2,\beta_3,\theta$)	341.04	224.61	352.10	200.35	234.15	153.11

Note. BR, basic reinforcement learning model, with parameters α,β_1,θ; RL1, reinforcement learning model with parameters $\alpha, \beta_1,\beta_2,\theta$; RL2, reinforcement learning model with parameters $\alpha, \beta_1,\beta_3,\theta$; RRC, risk–return–covariance reinforcement learning model.

The suggested RRC learning model was better than the simple BR learning model and the baseline model in predicting the behavior in all three experimental conditions. Table 5 also shows the results of learning models that only include the risk (RL1) or the covariance (RL2) component of the RRC learning model. Here the results illustrate that both components are important, especially in the positive and negative correlation conditions. In the condition with a zero correlation between the stocks, however, the simplified RRC model (RL1) that did not include the risk component did as well as the RRC model. In general, the results show that the participants changed their investment over the course of the experiment, which can be explained by the RRC model, but not by any model that predicts a constant allocation. Table 12 shows the average estimated parameters.

Table 12. Optimal Estimated Parameters of the Two Learning Models of Study2

			α	β_1	β_2	abs(β_2)	β_3	θ
BR	Positive correlation	Mean	0.51	1.00	-	-	-	1.10
		SD	0.46	0.00	-	-	-	3.72
	Zero correlation	Mean	0.32	1.00	-	-	-	0.35
		SD	0.45	0.00	-	-	-	0.45
	Negative correlation	Mean	0.22	1.00	-	-	-	0.44
		SD	0.41	0.00	-	-	-	0.37
RRC	Positive correlation	Mean	0.23	0.24	-0.34	0.52	0.24	2.66
		SD	0.30	0.33	0.58	0.42	0.37	7.45
	Zero correlation	Mean	0.19	0.17	-0.10	0.33	0.50	1.84
		SD	0.26	0.29	0.51	0.39	0.45	4.58
	Negative correlation	Mean	0.20	0.13	-0.22	0.24	0.63	2.99
		SD	0.33	0.25	0.36	0.34	0.41	5.79

Figure 4 shows the average portfolios and the learning models' predictions across all 90 periods averaged for blocks of 10 periods each. In the first block of all three correlation conditions the average investments were very close to an equal allocation to all three investment alternatives as predicted by a 1/*N* strategy (see also Table 9). Through the experiment, in all three conditions, participants increased their investments in the riskier stock B and decreased their investment in the risk-free bond C, especially in the negative and zero correlation conditions. These dynamics were better predicted by the RRC model. As in Study 1, the BR learning model overestimated the proportion invested in the riskier stock B and underestimated the proportion invested in the risk-free bond C, because it focused solely on the return advantage of stock B and ignored the risk component. In contrast, the prediction of the RRC model seems to be more in line with the observed behavior as it took the risk of the stocks into account.

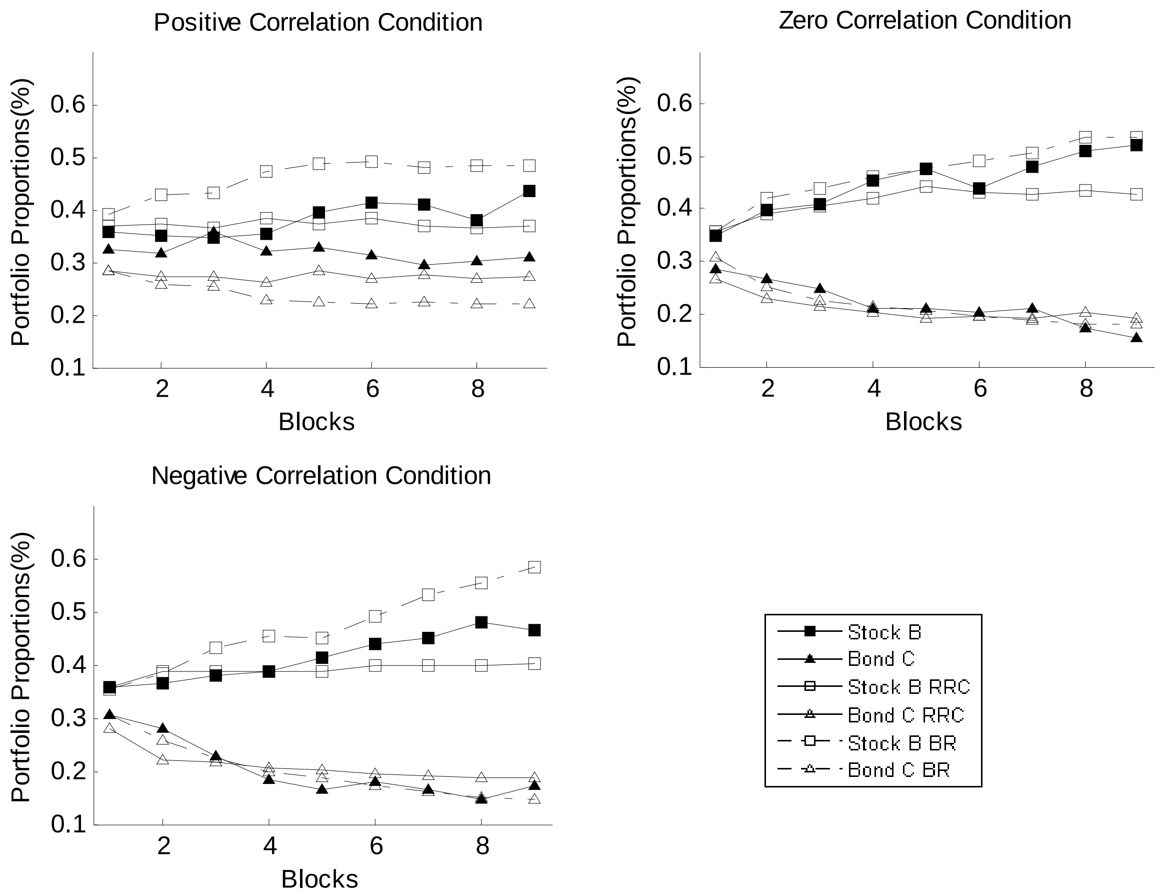


Figure 4. Average portfolios for the three correlation conditions in Study 2, compared with the average predicted portfolios of the BR learning model and the RRC learning model. The *x*-axes represent the 9 blocks with 10 investment periods each. Since the investment in A results from the investment in B and C, only the latter is presented.

4.2 Summary of Study 2

Study 2 demonstrated that people's investments can be influenced by investment alternatives' returns and variance, and the correlation of the returns of the investment alternatives. Surprisingly, Study 2 demonstrated this sensitivity in an investment situation where the participants were provided with only little information about the investment alternatives. In line with the principle of portfolio theory and the results from Study 1, investments in the risk-free bond were highest when the two stocks were positively correlated. Participants apparently realized that they could not reduce the risk by diversifying. Likewise, investment in the riskier stock B was lowest when the two stocks were positively correlated. In contrast, when the two stocks were negatively correlated, participants invested the least in the risk-free bond C and exploited the risk reduction of diversification. These findings are surprising when considering past research illustrating that people are insensitive to the correlation between investment alternatives (e.g., Hedesstroem et al., 2006; Kallir & Sonsino, 2009; Kroll et al., 1988a, 1988b; Lipe, 1998). However, the investments did not match the prediction of portfolio theory completely: The investment in the riskier stock B was higher in the negative correlation condition than in the positive correlation condition, although it was still lower than when the two stocks were independent. As in Study 1, the average portfolios for all three conditions followed the qualitative differences of the MV model. However, we did not observe efficient portfolios at the frontier lines, illustrating that the participants could have constructed better portfolios, that is, portfolios with similar average returns but lower risks.

The results also illustrate that the participants changed their portfolios during the 30 periods of each portfolio task. These dynamics cannot be explained by a model that predicts a constant portfolio over the whole investment period such as the baseline model implementing the $1/N$ strategy. In contrast, the two learning models did better in predicting the observed dynamics. In particular the RRC model predicted the behavior best. This is an important finding as it illustrates that people do not make their investments only on the basis of the alternatives' returns, as the BR model predicts, but take the risks of the alternatives into account. The two additional components of the RRC model appear as successful mechanisms to describe how people are influenced by the risk of investment alternatives.

5. General Discussion

How do people make portfolio decisions when they receive feedback about their decisions? This is the central question of the present article. To address it, we tested the standard financial MV model against two reinforcement models. The BR model predicts a learning process that relies solely on the return of the investment alternatives. In contrast, the extended reinforcement model (RRC model) also takes the variance and the correlation of investment alternatives' returns into account. The three models were tested with two experimental studies, in which participants repeatedly had to select a portfolio and received feedback about the portfolio's returns. In both experiments the participants received no information about the correlation of the risky assets, but they could detect it on the basis of received feedback. Past research examining our research question has led to a rather negative view of people's ability to detect the correlation between assets and to take that correlation into account for their investments (e.g., Hedesstroem et al., 2006; Kallir & Sonsino, 2009; Kroll et al., 1988a, 1988b; Lipe, 1998). We also found that the portfolios of the participants were not efficient and did not match the efficient frontier of portfolio theory. However, we observed that the participants reacted to the different correlation structures by making different investment decisions. In the cases of negative correlation between the risky assets, investment in the risk-free alternative was lowest and the opposite held for the positive correlation condition, which is exactly what is expected from portfolio theory.

In Study 1 the participants received quite a lot of information about the investment alternatives, so that they knew the average returns of the risky alternatives, the maximum and minimum returns, the distribution from which the returns were drawn, and the variability of the distribution. Only information about the correlation between the risky alternatives was withheld. With the provided information the participants could concentrate on using the received feedback to learn the correlation between the risky alternatives. This is exactly what happened: The participants changed their portfolio as a result of the feedback they received. In Study 2 the participants received only a little information. In contrast to Study 1 no information about the average returns or the variance of the returns was provided, making the portfolio task rather complicated. We expected that the participants with this little information would have had a hard time taking the correlation between the risky alternatives into account, because they also had to learn which alternative provided on average the higher return. Nevertheless, even with the little information provided we observed similar results to those in Study 1; that is, investment in the risk-free alternative differed among the three correlation conditions. Because they had little information, the participants preferred to allocate their resources equally among the investment alternatives at the beginning of the experiment, as one would predict with a $1/N$ investment strategy. However, this allocation quickly changed to one that increased investment in the alternatives with higher returns. However, even after substantial learning the participants in Study 2 invested more in the risk-free alternative compared to the participants in Study 1, who had explicit information about the stocks' expected returns. At the end of Study 2 the average portfolios in the three correlation conditions differed on a qualitative level as predicted by the MV model, similar to in Study 1.

When we compared the average investments with the MV model the results were only partially in line with its predictions. According to the MV model, participants should invest the most in the risk-free bond, when the stocks are positively correlated, and the least when stocks are negatively correlated. Our results are consistent with this for both experiments. However, the MV model also predicts that the investment in the riskier stocks should be highest when the stocks are negatively correlated, in contrast to our results. Even though investment in the riskier stock was highest in Study 1, this was not the case in Study 2. Furthermore, the MV model predicts that investment in the riskier stock should be lowest when stocks are positively correlated. Indeed, in Study 2 the participants invested the lowest proportion of their portfolio in the riskier stock. However, in Study 1 this was not the case. One reason for this high investment in the riskier stock B in Study 1 was its large return advantage compared to the risk-free bond. In Study 1 the participants were informed about stock B's return. Apparently, the return advantage of the riskiest stock was so high that the participants were willing to take the risk. In contrast, in the conditions with negative and zero correlation between the risky stocks, the participants realized the opportunity to diversify and invested less in the riskiest stock and more in the safer stock to reduce their overall investment risk.

In contrast, in Study 2 participants were not told about the large return advantage of stock B, so at the beginning they invested an equal share in all three investment alternatives. After some time they realized the return advantage of stock B, so they increased their investment in this stock. The increase was rather small as predicted by the RRC model, so even at the end of the experiment the investment was not as large as in Study 1. As predicted by the RRC model, the participants were also sensitive to the correlation and risk. They realized in the positive correlation condition that the returns of the two stocks went up and down almost together, so that there was no way to reduce the risk by diversifying. Therefore, they did not increase their investment in the riskier stocks as much as in the zero and negative correlation conditions, where the risk could be reduced by diversification. However, people's explicit knowledge of the correlation structure was limited: When they were asked at the end of the experiment to estimate the correlation their estimates were far from the real correlation, although they differed qualitatively in the right direction (see Table 10). The low accuracy of the correlation estimates might also be due to people's limited understanding of the correlation scale, in general.

With the risk preference task we estimated participants' risk preferences. According to investment theory the more risk averse participants were, the more they should have invested in the risk-free bond. This prediction was confirmed in both of our experiments. The correlation between risk preference and investment in the risk-free bond was positive in both experiments. However, the correlations were not very high and on closer examination, some were rather low. Therefore the investments were more a function of the received feedback and the observed learning process than of the participants' stable risk preferences. To explain participants' changes in their portfolios the learning theory comes into play.

In both experiments the participants changed their investment over the course of the investment period. The ANOVA showed that the average investments in the three repetitions of the portfolio task differed substantially. Participants increased their investment in the riskier stock B and decreased their investment in the risk-free bond C. To explain these results we tested two learning theories against each other. The first learning model we used was the BR learning model, where investment alternatives' expectancies are only updated by the investments' returns. This BR learning model essentially corresponds to the standard reinforcement model widely used in the literature (e.g., Sutton & Barto, 1998). We interpret the choice rule (Equation 1) such that it predicts the proportion of the resource invested in the various assets. Our results are consistent with those of previous studies (J. J. Choi et al., 2009, M. Kaustia and S. Knupfer, 2008, J. Rieskamp, 2006, J. Rieskamp et al., 2003, T. Shimokawa et al., 2009) showing that investors change their investments on the basis of experience. These feedback effects beg learning theories that can predict people's decisions as a function of experience. Yet surprisingly, all the reinforcement learning models developed so far focus solely on the direct outcomes (i.e., returns) of investment alternatives and ignore the variance and covariance of returns. Our results show that participants did not take only the returns into account, but also the returns' variance and the correlation between the stocks. Therefore, the BR model consistently did not describe the observed data very well. Instead, we suggested a new learning model in which the investments' expectancies were updated by three parts: the investments' returns, the returns' variance, and the covariance between the investment alternatives. To approximate the returns' variance with a simplified updating mechanism we used the absolute difference between two consecutive periods' returns. We suggest this approximation to illustrate how the variance of returns could be continuous estimated by the observed feedback after each investment period. This approximation works because the alternative with the higher (lower) return variance will also have higher (lower) absolute differences between consecutive returns. Likewise, we suggested a simple updating mechanism to approximate the covariance between the investment alternatives, by using only the absolute difference between the alternatives' returns. The absolute difference is lower (higher) if stocks have positive (negative) correlation. This new model did very well in describing the observed learning process and predicted the investments in the last repetition of the investment task better than the baseline model and the BR model. In sum, the experimental results illustrate that people react sensitively to the return, variance, and correlation of investment alternatives when they make repeated investments and receive feedback about their decisions. The feedback leads to a learning process that is best described by the RRC model.

However, our conclusions have to be limited. We have examined a rather simple investment situation with only three investment alternatives. Only two of the offered alternatives had varying returns, so that only the correlation between these two alternatives had to be considered. With this limited setting we observed people's sensitivity to the correlation, although the investments were not totally in line with the MV model. With more investment alternatives the number of correlations would increase exponentially, making their detection by an individual decision maker very hard if not even impossible. However, we do not aim to generalize our results to an investment situation with many investment alternatives. Rather, we were only interested whether people in general can notice the correlation between assets on the basis of trial-by-trial feedback and whether they recognize the diversification effect that results from correlation that can reduce the risk of an investment. With these positive results it might be beneficial to develop training tools where people can experience the advantage of diversification in hypothetical investment scenarios. These scenarios might illustrate the risk reduction advantage of investments in different types of assets that are not correlated with each other (or are even negatively correlated) compared to investments in similar types of assets that are positively correlated.

In summary, our studies show with a repeated portfolio task that people are sensitive to the correlation between stocks and their investments are partially consistent with the MV model. Participants showed a learning effect through the investment periods. The results show that people take information about stocks' returns, risks, and correlations into account. The RRC reinforcement model is able to describe this process and predicts people's behavior better than a simple, conventional learning model that focuses only on the investments' returns.

References

Benartzi, Shlomo and Richard H. Thaler. 2001. "Naive Diversification Strategies in Defined Contribution Saving Plans." American Economic Review, 91(1), 79-98.

Bossaerts, Peter and Charlie Plott. 2002. "The Capm in Thin Experimental Financial Markets." Journal of Economic Dynamics & Control, 26(7-8), 1093-112.

Bossaerts, Peter and Charlie Plott. 2004. "Basic Principles of Asset Pricing Theory: Evidence from Large-Scale Experimental Financial Markets." Review of Finance, 8(2), 135-69.

Camerer, Colin and Teck-Hua Ho. 1999. "Experience-Weighted Attraction Learning in Normal Form Games." Econometrica, 67(4), 827-74.

Canner, Niko; N. Gregory Mankiw and David N. Weil. 1997. "An Asset Allocation Puzzle." American Economic Review, 87(1), 181-91.

Choi, James J.; David Laibson; Brigitte C. Madrian and Andrew Metrick. 2009. "Reinforcement Learning and Savings Behavior." Journal of Finance, 64(6), 2515-34.

Erev, Ido and Greg Barron. 2005. "On Adaptation, Maximization, and Reinforcement Learning among Cognitive Strategies." Psychological Review, 112(4), 912-31.

Erev, Ido and Alvin E. Roth. 1998. "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria." American Economic Review, 88(4), 848-81.

Feltovich, Nick. 2000. "Reinforcement-Based Vs. Belief-Based Learning Models in Experimental Asymmetric-Information Games." Econometrica, 68(3), 605-41.

Hedesstrom, T. Martin; Henrik Svedsater and Tommy Garling. 2006. "Covariation Neglect among Novice Investors." Journal of Experimental Psychology-Applied, 12(3), 155-65.

Hopkins, Ed. 2002. "Two Competing Models of How People Learn in Games." Econometrica, 70(6), 2141-66.

Kallir, Ido and Doron Sonsino. 2009. "The Neglect of Correlation in Allocation Decisions." Southern Economic Journal, 75(4), 1045-66.

Kaustia, Makku and Samuli Knupfer. 2008. "Do Investors Overweight Personal Experience ? Evidence from Ipo Subscriptions." Journal of Finance, 63(6), 2679-702.

Kroll, Yoram; Haim Levy and Amnon Rapoport. 1988. "Experimental Tests of the Mean Variance Model for Portfolio Selection." Organizational Behavior and Human Decision Processes, 42(3), 388-410.

Kroll, Yoram; Haim Levy and Amnon Rapoport. 1988. "Experimental Tests of the Separation Theorem and the Capital-Asset Pricing Model." American Economic Review, 78(3), 500-19.

Lipe, M. G. 1998. "Individual Investors' Risk Judgments and Investment Decisions: The Impact of Accounting and Market Data." Accounting Organizations and Society, 23(7), 625-40.

Markowitz, Harry. 1952. "Portfolio Selection." The Journal of Finance, 7(1), 77-91.

Markowitz, Harry. 1959. Portfolio Selection: Efficient Diversification of Investments. New York: Wiley.

Nelder, J. A. and R. Mead. 1965. "A Simplex-Method for Function Minimization." Computer Journal, 7(4), 308-13.

Rieskamp, Jorg. 2006. "Positive and Negative Recency Effects in Retirement Savings Decisions." Journal of Experimental Psychology-Applied, 12(4), 233-50.

Rieskamp, Jorg; Jerome R. Busemeyer and Tei Laine. 2003. "How Do People Learn to Allocate Resources? Comparing Two Learning Theories." Journal of Experimental Psychology-Learning Memory and Cognition, 29(6), 1066-81.

Shimokawa, Tetsuya; Kyoko Suzuki; Tadanobu Misawa and Yoshitaka Okano. 2009. "Predicting Investment Behavior: An Augmented Reinforcement Learning Model." Neurocomputing, 72(16-18), 3447-61.

Sutton, Richard S. and Andrew G. Barto. 1998. “Reinforcement Learning: An Introduction.” Cambridge: MIT Press.

Tobin, James. 1958. "Liquidity Preference as Behavior Towards Risk." The Review of Economic Studies, 25(2), 65-86.

Waelti, Pascale; Anthony Dickinson and Wolfram Schultz. 2001. "Dopamine Responses Comply with Basic Assumptions of Formal Learning Theory." Nature, 412, 43-48.

Appendix (Not for publication)

Instructions for Study 1

Dear Participants:

Thank you for participating in our experiment. There are two parts to this experiment.

Part I: Portfolio decision making

Part II: Gamble choice

Part I. Portfolio decision making

First, please answer a few questions on the first page, then press the <next> button. After that, please read the following part of these instructions.

In this experiment, you have to make repeated investment decisions. There are three rounds in this experiment. In each round, you will start with a hypothetical endowment of 10,000 experimental currency units (ECUs) that you can then invest over 30 periods. In each period you have to invest all your gains from the previous period.

Different financial assets will be available, among which you have to allocate your money. You can freely choose how to split a specific amount among the available investment alternatives. The value of your portfolio in every period depends on the change in value of the financial assets in which you invested. You can change your investment choices in every period.

You can invest in three different assets. The returns of the first two assets will vary from period to period. In contrast, the third assets return is risk-free and constant across all periods. The available assets in this experiment can be described as follows:

Asset A:

This asset will provide a return that varies as illustrated in Figure 1. The figure shows the distribution of possible returns and how likely they are to occur. The figure shows that in around 68% of all cases the returns of the fund will lay between 0% and 12%. Returns larger than 12% will occur in around 16% of all cases. Returns lower than 0% will occur in 16% of the cases. The distribution of returns illustrated in Figure 1 has an average return of 6%.

In statistical terms, the returns of asset A will be drawn from a so-called *normal distribution* with a mean of 6% and a standard deviation of 6%. The standard deviation is a statistic that tells you how tightly all possible returns are clustered around the mean. In every period of the experiment, the return of the investment will be randomly drawn from this distribution to determine the return for your investment in asset A. Each draw in each period is independent from previous returns. In principle the distribution also allows for very large positive and negative returns. However, in the present experiment the possible returns are restricted to a minimum return of -12% and a maximum return of +24%.

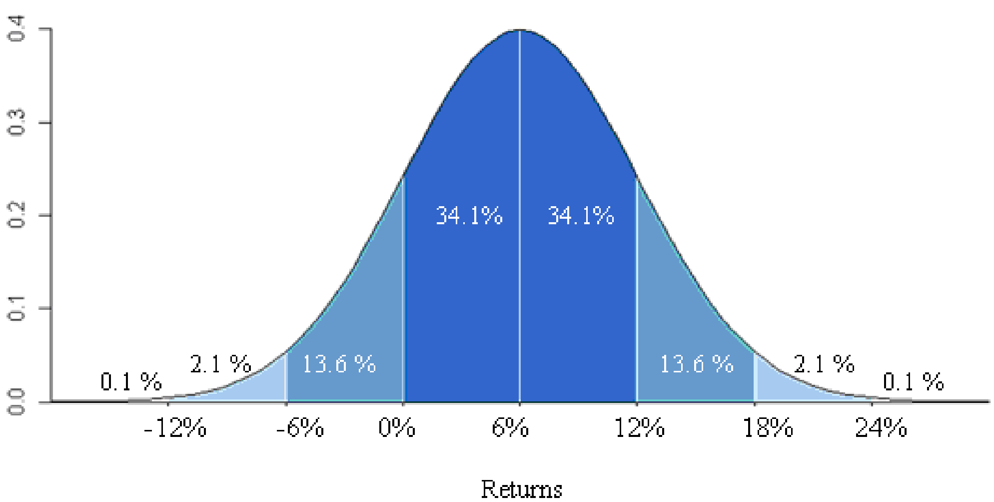


Figure 1. Theoretical distribution of the possible returns for asset A. Around 68% of the returns will lie between 0% and 12%. Around 95% of the returns will be between -6% and 18%.

Asset B:

This asset will provide a return that varies as illustrated in Figure 2. The figure shows the distribution of possible returns and how likely are to occur. The figure shows that in around 68% of all cases the returns of the fund will lay between -5% and 23%. Returns larger than 23% will occur in around 16% of all cases. Returns lower than -5% will occur in 16% of all cases. The distribution of returns illustrated in Figure 2 has an average return of 9%.

In statistical terms the returns of asset B will be drawn from a normal distribution with a mean of 9% and a standard deviation of 14%. In every period of the experiment, the return of the investment will be randomly drawn from this distribution to determine the return for your investment. Each draw in each period is independent from previous returns. In principle the distribution also allows for very large positive and negative returns. However, in the present experiment the possible returns are restricted to a minimum return of –34% and a maximum return of +51%.

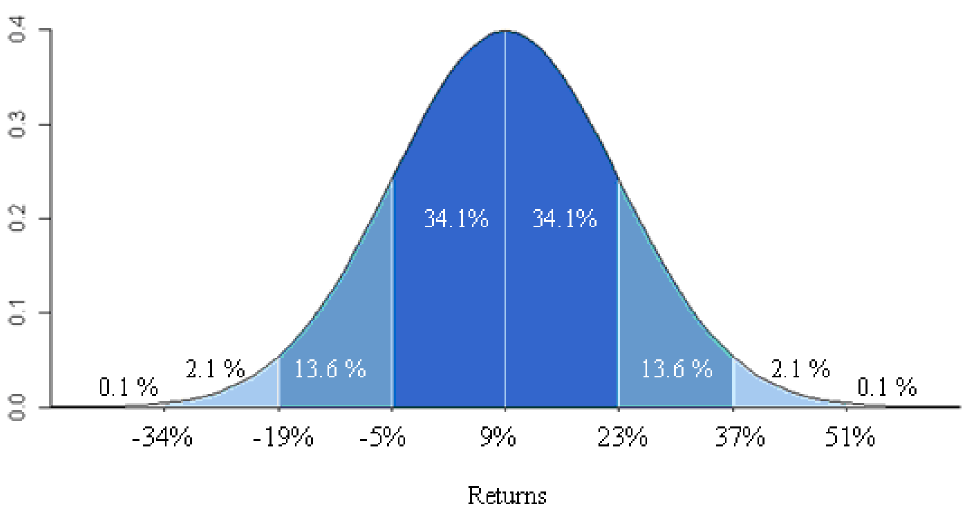


Figure 2. Distribution of the possible returns for asset B. Approximately 68% of the returns will lie between -5% and 23%. Around 95% of the returns will be between -19% and 37%.

Asset C:

This asset will provide a fixed return of 2%. Thus, in every period of the experiment any amount invested in this fund will have a return of 2%. Any amount of money not invested in assets A and B in each period will automatically be invested in asset C.

To illustrate this situation, Figure 3 shows for 100 cases the frequency with which you can expect specific returns for the three investments. For asset C the situation is easy: This fund will always lead to a return of 2%. For asset A you can see that a return of around 6% will occur in most of the cases, in around 40 of 100 cases. In most cases the returns will lie between -2% and 14%. For asset B a return of around 9% is most likely to occur, in around 30 of 100 cases. In most cases the returns will lie between -1% and 19%.

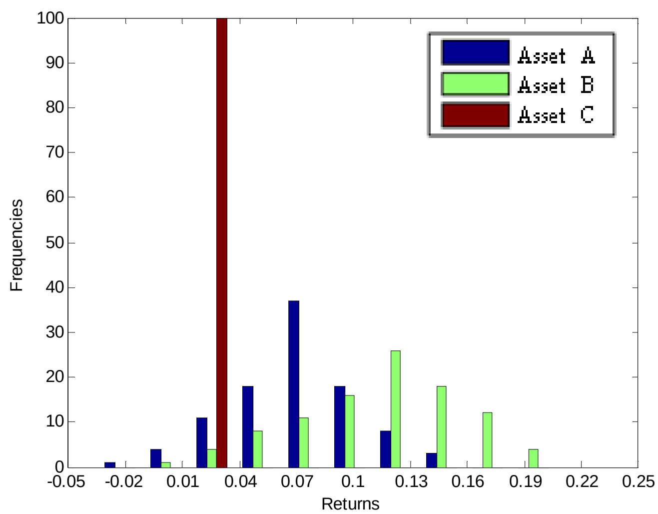


Figure 3. The expected frequencies with which you can expect specific returns for the three investments, over 100 cases.

How you make an investment

You need to decide how to allocate your endowments by stating the percentage you want to invest in the three assets. For simplification, you can only use integers (for example: 35% instead of 35.2%).

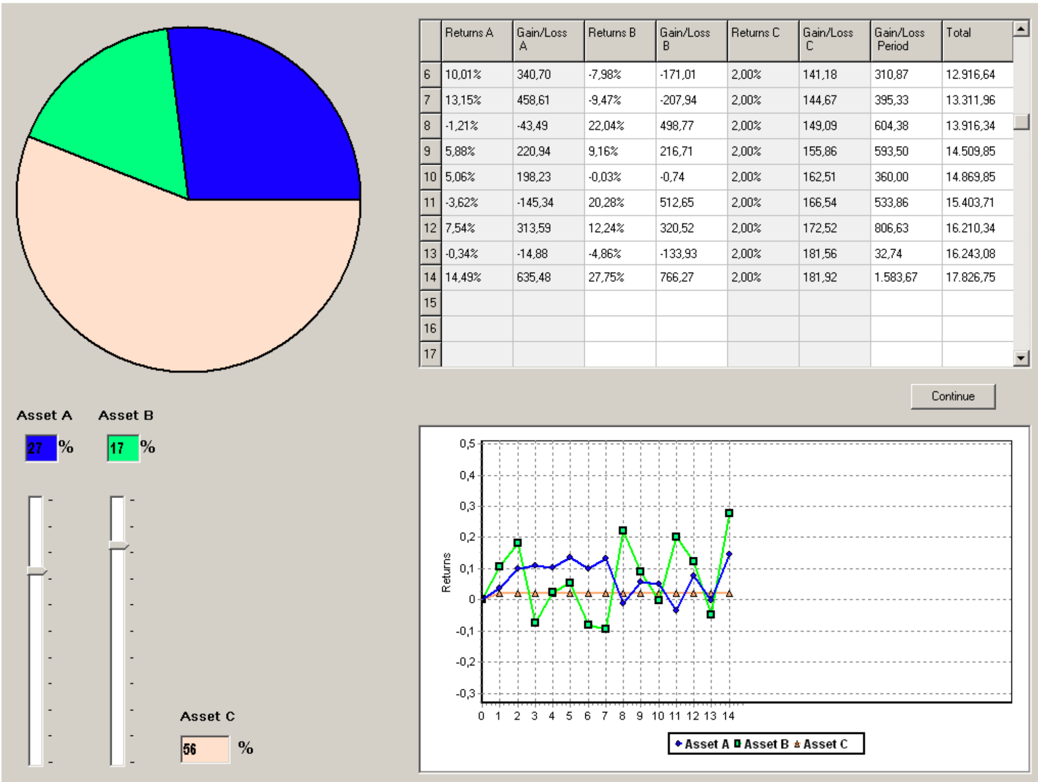
On the next page, you will find an example of what the computer screen looks like in the experiment.

To allocate your investments, click and hold the left mouse button on the two slide rules and tabs beneath the pie diagram, shown on the computer screen. The first slider refers to the proportion of your resource that you want to invest in asset A. The second slider refers to the proportion of your resource that you want to invest in asset B. The proportion of your resource that you do not invest in asset A or asset B will be automatically invested in asset C, which is represented by a box next to the two sliders.

To make or change your investment, slide the tabs on the slide rules up or down. Sliding a tab down means you want to increase the proportion of your investment in this asset and decrease the proportion of your investment in the other assets. Sliding a tab up will decrease the proportion of your investment in this asset and increase the proportion of investment in the other assets. Once you release the mouse button, the pie chart in the upper left will be updated to reflect your investment choices. You can adjust your investment as many times as you like. When you are finished adjusting, please click the <next> button below the history window.

After you have made your decision, the history window and the return window will be updated. In the history window, you will see the percentage of your portfolio you invested in one of the three assets for each period. Additionally, you will see the return rate of each asset in percentages. Finally, the gain (or loss) you made with your investment in each period and the total value of your portfolio will be shown in the table. Beneath the history window you will see a figure that presents the returns of the three assets graphically.

Example Screen



	Returns A	Gain/Loss A	Returns B	Gain/Loss B	Returns C	Gain/Loss C	Gain/Loss Period	Total
6	10,01%	340,70	-7,98%	-171,01	2,00%	141,18	310,87	12.916,64
7	13,15%	458,61	-9,47%	-207,94	2,00%	144,67	395,33	13.311,96
8	-1,21%	-43,49	22,04%	498,77	2,00%	149,09	604,38	13.916,34
9	5,88%	220,94	9,16%	216,71	2,00%	155,86	593,50	14.509,85
10	5,06%	198,23	-0,03%	-0,74	2,00%	162,51	360,00	14.869,85
11	-3,62%	-145,34	20,28%	512,65	2,00%	166,54	533,86	15.403,71
12	7,54%	313,59	12,24%	320,52	2,00%	172,52	806,63	16.210,34
13	-0,34%	-14,88	-4,86%	-133,93	2,00%	181,56	32,74	16.243,08
14	14,49%	635,48	27,75%	766,27	2,00%	181,92	1.583,67	17.826,75
15								
16								
17								

After the first 30 periods, you will have the chance to repeat your investment choices another two times. In rounds 2 and 3, you will see the same screen again and make another 30 periods' investment choices. The distributions of possible returns are the same as in round 1.

After completing all three sessions you will get further instructions and will have to fill out a questionnaire.

Payment for part I:

You will receive a bonus payment that will depend on the success of your investments. Only one round will be randomly chosen to determine your payment. For every 10,000 ECUs you earn with your investment you will receive 1 euro. For example, if you earn 46,300 ECUs, you will get 4.60 euros cash.

You are NOT being videotaped. Please do not hesitate to ask if you have any questions. Once you have finished the portfolio decision-making part, please read the following instructions for the second part.

Part II. Gamble choices

In this part of the experiment you have to make 10 decisions. With each decision you can choose between two gambles: Gamble A and Gamble B. Both lotteries are based on a certain probability and will lead to two different pay-off alternatives. After you make all 10 decisions, the computer will randomly choose one of the 10 pairs. The pair picked will be played and the resulting pay-off will be paid to you.

Total payment:

For your participation in this experiment, you will receive a fixed payment of 4 euros. Additionally you will receive a payment based on the outcome of the portfolio decision and the gamble choice parts.

Thank you for your participation in the experiment. When you finish the experiment, please approach the experimenter and collect your payment.

Instructions for Study 2:

The instructions for Study 2 are the same as for Study 1, except that we eliminated the distribution parameters of the two stocks. Therefore, the explanation for the two stocks was as follows:

Asset A (B):

This asset will provide a return that varies as illustrated in Figure 1. The figure shows the distribution of possible returns and how likely they are to occur. The exact average return is unknown. However, you can see that the returns will be below (above) the average return in 50% of the cases.

The returns of asset A (B) were drawn from a so-called *normal distribution*. The standard deviation is a statistic that tells you how tightly all possible returns are clustered around the mean. In every period of the experiment, the return of the investment will be randomly drawn from this distribution to determine the return for your investment in asset A. Each draw in each period is independent from previous returns. In principle the distribution also allows for very large positive and negative returns. However, in the present experiment the possible returns are restricted.

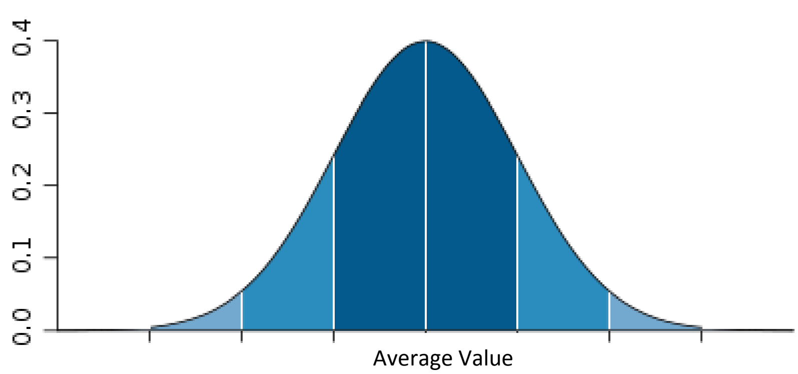


Figure 1. Distribution of the possible returns for assets A and B.