ARTICLE

Impulse balancing versus equilibrium learning an experimental study of competitive portfolio selection

Judith Avrahami[1] · Werner Güth[2] · Yaakov Kareev[1] · Matteo Ploner[3]

© The Author(s) 2022

Abstract

The experiment lets investors interact in portfolio choices involving different risky assets, one for each state of the world. Probabilities of random states are commonly known. All assets pay the same dividend when their state is realized and becomes worthless otherwise. Whereas evolutionary stability and equilibrium behavior predict equal expected profits across assets, impulse balancing (Selten and Buchta, Games and human behavior: essays in the Honor of Amnon Rapoport, 1999) equalizes the expected regret. Thus, impulse balancing seems to capture tendencies of cyclical direction learning. In addition to analyzing whether and when behavior converges to impulse balancing or to equilibrium portfolios, we categorize portfolio adaptation by path dependence and sensitivity to state-specific probabilities. We show that portfolio choices are driven mainly by probability matching, but the effect becomes weaker over time. Furthermore, most portfolio adjustments are not compatible with directional learning.

Keywords Impulse balancing · Probability matching · Regret · Portfolio management · Experiment

JEL Classification D81 · C91 · G11

1 Introduction

Our paper experimentally investigates portfolio investment choices. We deal with this issue of vast empirical relevance within an evolutionary framework. Evolution relies on path dependence, e.g., on the average past success or fitness of the available

✉ Matteo Ploner
   matteo.ploner@unitn.it

1   The Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem, Jerusalem, Israel

2   MPI for Research on Collective Goods, Bonn, Germany

3   CEEL-Department of Economics and Management, University of Trento, Via Inama 5, 38122 Trento, Italy

options, and is usually justified by assuming an infinite population of short-lived actors who are randomly matched to interact with finitely many others. Optimal adaptation is also predicted by equilibrium behavior whose justification is, however, relying on forward-looking rational deliberations of all the interacting parties. Here, we explore whether mutual best responding could be learned. Based on the literature on phenotypical learning, one can distinguish learning in the form of best-reply dynamics, i.e., interacting agents best respond to the most recent constellation of circumstances beyond their own control like stochastic costs and others' choices, or regret-driven behavioral adaptation.

We focus essentially on the distinction of equilibrium learning and (regret) impulse balancing since, for the setup given, evolutionary stability implies equilibrium behavior. Both concepts allow for static equilibria, based on the common knowledge assumptions of strategic equilibria, but differ in what the interacting parties are supposed to optimise, namely maximize expected profits or minimize regret, across assets. Our experimental scenario modifies Blume and Easley (1992)'s evolutionary model of financial portfolio survival which, under quite strong assumptions, predicts that the share of capital, invested proportionally to the stationary positive probabilities of the various states, eventually converges to 1. Specifically, the state-specific assets pay (the same) positive dividend when their state is realized and become otherwise worthless. Thus, to preserve wealth across time, one has to engage in all assets and to invest proportionally to the state-specific probabilities, as predicted by probability matching.

Like Blume and Easley (1992), we let investors periodically determine their expenditure shares for the various assets and, due to constant supply of assets, this allows to derive the market-clearing prices. Whereas the past average profits are the driving forces of evolutionary selection, as in Blume and Easley , our experiment instead substitutes evolutionary selection by letting participants adapt their portfolio composition across time in light of feedback. Actually, the evolutionarily stable portfolio composition, namely that expenditure shares are proportional to asset probabilities, can also be justified as strategic equilibrium behavior, which requires commonly known rationality. Thus our experimental analysis confronts the joint hypothesis of *evolutionarily stable or equilibrium behavior*[1] with that of *impulse balancing*, capturing in a quantitative way phenotypical learning via the average choice probabilities of cyclic directional learning (Selten and Buchta 1999).

Both approaches provide a static and dynamic justification (see Table 1) where, of course, the dynamics generally do not have to converge. Instead, the static interpretations rely on mutual best responses by equalizing expected profits

	Blume and Easley (1992)	Our experiment
Static	Strategic equilibrium	Impulse balancing
Dymanic	Evolutionary stability	Direction learning

Table 1 The Blume and Easley 's model and our contribution

[1] Since evolutionary stability requires to be best adapted to one's habitat, it is not unusual that the results of evolutionary selection and game-theoretic equilibrium analysis coincide (see, for instance, Weibull 1997; Samuelson 2002).

(strategic equilibrium) and expected regret (impulse balancing). Thus, our study confronts static concepts, based on common knowledge (the top row of Table 1), and their dynamic justifications (the bottom row of Table 1). The former are imported from evolutionary biology; the latter predict phenotypical learning but differ in their driving forces: past average profits (evolutionary stability) or retrospective analysis of own past choices in the light of individual experiences (direction learning).

Furthermore, we will categorize portfolio adaptation in the light of feedback information on past states, past prices, and own investment success. Is portfolio adaptation shaped by both anticipating its likely consequences (the shadow of the future, as in *consequentialist logic*) and learning from the past (the shadow of the past, as in *learning and evolutionary theory*)? In view of impulse balancing, we will specifically pay attention to alternative measures of regret (e.g., Avrahami et al. 2005).

The theoretical and empirical study of financial portfolio choice and adaptation has important real-world implications, even when based on stylized setups. For instance, one could easily weaken the assumption of stationary state probabilities by allowing for rare shocks, after which probabilities become stationary again till the next shock applies. The prediction of Blume and Easley (1992) would then still predict the drift in portfolio adaptation in phases of constant probabilities. Whether and how at best boundedly rational participants anticipate such shocks seem so far only exploratively analyzed.

Regarding the comparison of equilibrium behavior and impulse balancing, the experimental evidence so far mainly rests on experimental games with (mostly unique) equilibria in completely mixed (all possible strategies are used with positive probabilities) strategies. Here, the evidence speaks quite in favor of impulse balancing as its mixing predictions are based on payoffs of the mixing players, whereas strategic equilibrium mixing is based on rendering other players indifferent and therefore depends on others' payoffs.

To assess our conclusions robustly, we vary the vector of stationary state probabilities, the number of assets, and the market size by doubling the number of interacting traders. We will describe how we disentangle between learning from the past and anticipation of changes in crucial parameters after introducing the market model with its benchmark solutions and stating hypotheses.

Altogether our data reveal strong anchoring of portfolio design on the success probabilities, in line with the intuition of probability matching. However, this behavior is less prominent in later rounds and when fewer options are available. Furthermore, reacting to own individual past success in the spirit of direction learning becomes less important over time. Among those who respond to past outcomes, different sources of regret seem to be relevant.

Section 2 describes the market model that is the basis of our experimental setting and offers the benchmark solutions for assessing actual behavior. Section 3 states some hypotheses, and Sect. 4 provides details of the experimental implementation. After the data analysis of Sect. 5, we discuss and summarize our results in the concluding Sect. 6.

2 The market model and its benchmark solution

Let $m(\geq 2)$ denote the number of assets $j = 1, \ldots, m$, whose positive success probabilities are stationary, and $n(\geq 2)$ the number of investors $i = 1, \ldots, n$. The same n investors face this stationary market environment repeatedly. In the evolutionary analysis, the focus is on the final wealth shares of stationary, e.g., genetically or phenotypically determined portfolio composition types. Reaching the end result of evolution may require infinitely many periods by any selection (see Blume and Easley 1992). However, in the experiment, the horizon of this repeated interaction, the "shadow of the future", is captured by a finite number of successive investment periods.

Assets are state-specific: for each asset $j = 1, \ldots, m$, there is a random state with positive probability $w_j (> 0)$. Evolutionary selection assumes repeated application of these stationary and state-specific probabilities, but does not require investors to know them. To allow for immediate or quick convergence to equilibrium behavior, the experiment, however, induces common knowledge of probabilities, respectively of their changes.[2] An asset j is only rewarded with return rate $R(> 0)$ when state j is realized; otherwise, asset j becomes worthless. Specifically, it will be assumed, and experimentally induced, that vector $w = (w_1, \ldots, w_m)$ is commonly known. Actually, w and m and n will be the conditions varied in our experimental analysis.

The only possibility to save wealth from one period to the next is by investing it, since what is not invested is lost. For all investors $i = 1, \ldots, n$, let $E(i)$ denote the wealth of investor i which i can invest in the m assets by choosing portfolio shares $s(i) = (s_1(i), \ldots, s_m(i))$, with $0 \leq s_j(i) \leq 1$ for $j = 1, \ldots, m$ and $s_1(i) + \cdots + s_m(i) = 1$, i.e., $s_j(i)E(i)$ is the amount which i invests in asset $j = 1, \ldots, m$.[3]

For each of the m assets, there exists a constant supply of $S = 100$ units. Thus, the market-clearing price p_j for asset $j = 1, \ldots, m$ is determined by $p_j S = \sum_{i=1}^{n} s_j(i)E(i)$. At this price, p_j investor i buys according to i's hyperbolic demand function $\frac{s_j(i)E(i)}{p_j}$ units of asset j which pay him, in case of state j being selected, $\pi_j(i) = R \cdot \frac{s_j(i)E(i)}{p_j}$ and become worthless otherwise.[4] The expected payoff of choice $s(i)$ by investor i is $\sum_{j=1}^{m} w_j \pi_j(i)$ for $j = 1, \ldots m$.

The equilibrium benchmark s for this interactive portfolio choice is portfolio $s^* = (s_1^*, \ldots, s_m^*)$ with $s_j^* = w_j$ for $j = 1, \ldots, m$ and $s^*(i) = s^*$ for $i = 1, \ldots, n$. In addi-

---

[2] When different predictions like evolutionary stability and equilibrium behavior rely on differences in (common) knowledge, it may be possible to discriminate between them experimentally, even when they predict the same benchmark outcome. However, since we want to compare equilibrium and learning concepts, this is not exploited here, and we chose to implement a scenario to which all concepts may apply.

[3] In the experiment, only integer percentages are allowed for portfolio shares so that $s_k(i) \in \{0\%, 1\%, \ldots, 99\%, 100\%\}$ for $k = 1, \ldots, m$ and $i = 1, \ldots, n$.

[4] Such demand functions can be rationalized by suitable utility functions and have been frequently estimated econometrically in the past (for an earlier application of so-called linear expenditure systems, see Stone (1954).)

tion to its obvious prominence as the only possible anchor for participants in the experiment, the equilibrium benchmark can be justified by evolutionary stability.[5]

If $E(i) = E$ for all interacting investors $i = 1, \ldots, n$, as experimentally implemented, common reliance on benchmark behavior $s(i) = s^*$ for $i = 1, \ldots, n$ implies the equilibrium prices $p_j = \frac{nE}{S} w_j$ for $j = 1, \ldots, n$ and expected payoffs RS/n. Thus, investors would earn in expectation less when only n increases, but would earn the same when both, s and n, increase, such that S/n remains constant. The effect of increasing R is obvious.

In their innovative evolutionary justification, Blume and Easley assume each investor $i = 1, \ldots, n$ to be only once positively endowed. Therefore, in later periods, one's endowment is positive only when having positively invested in all previous winning states. In the experiment, we deviate from this to keep all participants "alive" in the sense of maintaining them as active investors. Specifically, we endow participants repeatedly in each period and also allow them to change their portfolio composition, i.e., to adapt their portfolio choice from period to period. Nevertheless, the intuition of the evolutionarily stable portfolios $s^* = w$ provides a useful benchmark prediction about how portfolio choices might converge with more and more experience. We predict learning to let average portfolio choices $s(i)$ converge to $s^* = w$, eventually.

Evolutionarily stable choices often coincide with equilibrium behavior in a game-theoretic sense, i.e., when the game and rationality of all investors are commonly known.[6] The normative justification, based on common(ly known) rationality and risk neutrality, does not require learning or evolution, but is purely based on consequentialist forward-looking deliberation (the shadow of the future).[7] Crucially, observing convergence to s^* choices only after learning would support the evolutionary or learning rather than the normative s^* justification.

Even when participants frequently rely on s^* already initially, this may not necessarily support the equilibrium justification of s^*, as initial reliance on s^* is predicted by anchoring in the form of *probability matching*. Actually, $w = (w_1, \ldots, w_m)$ is the only numerical anchor provided by the instructions. If at all, the only alternative anchor, namely $s_j(i) = 1/m$ for $j = 1, \ldots, m$ and $i = 1, \ldots, n$, the so-called 1/m-heuristic (Benartzi and Thaler 2001) (often referred to as the Golden Mean, $1/m$), is only implicitly provided.

However, neither the static equilibrium concept nor anchoring predict whether and how participants will question its recommendation after receiving feedback information. If, by chance, the first randomly selected states are rather unexpected, e.g., by being very unlikely, one might question anchoring on w. Similarly, when frequently experiencing a state j with w_j exceeding $1/m$ considerably, the 1/m-heuristic

---

[5] For a derivation under partly different assumptions, see Blume and Easley (1992).

[6] It is one advantage of the evolutionary justification that it does not require common knowledge assumption, as only nature has to be aware of what is fitness-enhancing.

[7] Risk neutrality for the evolutionary analysis follows from randomly forming oligopoly markets with n investors each in each period, based on an infinite population, i.e., on infinitely many such markets, so that fitness is measured by average success.

might lose its appeal. Furthermore, when allowing subjective probability adaptation to past realizations of states in spite of the independent repeated application of w, one may engage in best-reply dynamics to past results, for instance, due to the gambler's fallacy.

An alternative equilibrium concept presupposes regret balancing, for instance, impulse balance equilibria (Selten and Buchta 1999; Ockenfels and Selten 2005) capturing the average choice probabilities of cyclical direction learning. In our context, this means to balance the possible regret when investing in the m different assets. Here, we measure regret via investment shares rather than by expected monetary returns, which depend on market prices and others' behavior. Specifically, when asset j was drawn, agent i regrets everything not invested in asset j. This occurs with probability w_j and amounts to the expected regret of $w_j(1 - s_j(i))$. When asset j was not drawn, i regrets everything invested in j, i.e., the portfolio share $1 - s_j(i)$. This occurs with probability $(1 - w_j)$, amounting to $(1 - w_j)s_j(i)$. Balancing these two amounts results in $s_j(i) = w_j$, so that, in the case of monomorphic $s^* = w$-behavior, no investor would suffer from unbalanced regret. In the analysis of our experimental data, we also present other, more behaviorally motivated measures of regret and explore their relationship to dynamic portfolio adjustment (see Sect. 5.3.3).

When considering only their final predictions, all four approaches[8]—*evolution, equilibrium, w-anchoring, and impulse balancing*—predict the same portfolio compositions, i.e., $s^* = w$, whereas the 1/m-heuristic predicts, at least initially, w-insensitivity. For the experimental analysis, the research questions, therefore, concern more the dynamics, i.e., the individual adaptation of portfolio composition, than the final choices.

- Will the same final s^*-predictions be confirmed by monotonic adaptations, possibly by moving away from $s_j(i) = 1/m$ choices, towards s^*

  - already initially,
  - only after learning,
  - not at all?

- Will portfolio adaptation be cyclic as suggested by direction learning (Selten 2004)?
- Will the experimentally observed adaptation steps allow to distinguish between the static (anchoring, equilibrium, and impulse balancing) or dynamic (evolutionary, learning, e.g., qualitative direction learning) s^*-justification?

---

[8] Confounded justifications of the same benchmark behavior is rare in experimental research (one usually employs a design separating predictions of different "solution" ideas), but seems important and interesting as off-benchmark behavior is the rule rather than the exception. Our setup illustrates that off-benchmark behavior can be the only way to distinguish between "solution" ideas.

3 Hypotheses

From a behavioral perspective, (expected) regret may not be purely restricted to investment, but is probably also affected by wealth effects that depend on market-clearing prices. Therefore, for example, impulse balancing, the underlying intuition of directional learning, provides a clear hypothesis of adaptive behavior: when i has invested $s_j(i)E(i)$ in asset j whose state j is not selected, we expect a likely $s_j(i)$ -decrease in the next period. The effect should be regret related via $(1 - s_j)\frac{s_j(i)E(i)}{p_j}$ or via $s_k\frac{s_j(i)E(i)}{p_k}$, where k denotes the actually selected state in the past period. The first specification does not presuppose that one could have guessed correctly the actually realized state j; the second assumes that one could have guessed right. Of course, in the case of two states only, $m = 2$, which we also explore, the two specifications coincide.

The evolutionary analysis by Blume and Easley does not rely on the individual adaptation of portfolios but on individual wealth. To capture this intuition and test it by experimental data, one could explore whether, across time those investors with past choices closer to s^* have on average also earned more. This would, at least qualitatively, confirm the intuition of the evolutionary selection driving the result of Blume and Easley . For the static justifications of s^*, one may at best observe some initial effects, e.g., when by chance the first random realizations differ considerably from w. Will this induce adaptation of subjective beliefs away from w, or will anchoring at w persist and be revealed by constant portfolio choices close to w?

In our view, the multi-dimensional action space for choosing $s(i)$ and the rather fine grid, imposed by integer percentages from 0% to 100% when selecting $s_j(i)$, provide an interesting setup for studying and comparing learning in a rather rich setup.

Behaviorally, we cannot hope for the same portfolio selection by all individuals, which renders the anticipation of market-clearing prices questionable. When analyzing how portfolios are adapted, one can use past instead of anticipated prices. From an evolutionary perspective, one may, for instance, rely on w_j/p_j, with p_j denoting the past price of asset $j = 1, \dots, m$ to compare what a marginal token would achieve in different states j.

Concerning impulse balancing, we rely on the dynamic intuition of direction learning (see Selten and Buchta 1999) to investigate whether and how participants react to feedback information after each period. This implies conditioning on just the past state k, so that $s_k(i)$ will not be increased when state k had not been drawn and that $s_k(i)$ should not be decreased when state k was realized. Of course, how one reacts to (non) realized states may also depend on past prices due to impulses $s_j(i)E(i)/p_j$ and $s_k(i)E(i)/p_k$, a quantitative aspect of impulse balancing but not of purely qualitative direction learning.[9]

Participants may over-or under-estimate the probability of observing the same state in successive rounds, resulting in frequent restructuring of portfolios, which

---

[9] Impulses $s_j(i)E(i)/p_j$ and $s_k(i)E(i)/p_k$ have a broader meaning that in our justification of the s^*–impulse balance equilibrium in Sect. 2.

Table 2 Probability of success for each prospect (w_j)

Prob (%)	Prospect 1	Prospect 2	Prospect 3	Prospect 4
Set 1	43.8	56.3	–	–
Set 2	25.0	75.0	–	–
Set 3	12.5	21.9	28.1	37.5
Set 4	12.5	15.6	34.4	37.5
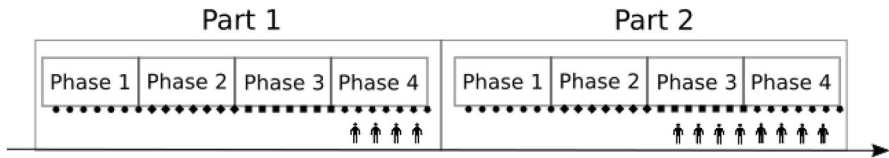
Fig. 1 Experimental structure (Part 1 (2): 4 (8) investors)

may question (convergence to) s^*-behavior. Behaviorally, participants may be influenced by past prices p_i and state-specific w_i in other ways than predicted by impulse balancing or equilibrium (evolution). As postulated by reinforcement learning, path dependence may also be based on longer memory length than just one round as supposed by direction learning.

The theoretical predictions do not depend on endowments and differences in individual endowments. Furthermore, probabilities of the m states can be rather similar and very different, which will be explored experimentally by varying the probabilities of the m states. If, for instance, the probabilities of the m different states do not differ much, this should enhance the support for the so-called "1/m" heuristic (Benartzi and Thaler 2001) predicting equal portfolio shares for all m assets. In the case of substantially different probabilities, even investors, fearing the risk of neglecting low probability assets, should learn to invest significantly less in substantially less likely states.

4 Experimental design and procedures

Participants i are asked to allocate 100 tokens to the available assets or prospects by choosing $s(i)$, repeatedly. When a prospect is selected, its returns are positive; otherwise, its returns are null. Specifically, the returns from the prospect are given by €0.10 × the number of units purchased of the selected prospect.[10]

Overall, four different sets of prospects are considered, with either m=2 or m=4 states (see Table 2).

The experiment consists of two parts (see Fig. 1): *Part 1* and *Part 2*. In *Part 1*, participants are matched together in groups of 4, while in *Part 2*, they are matched

---

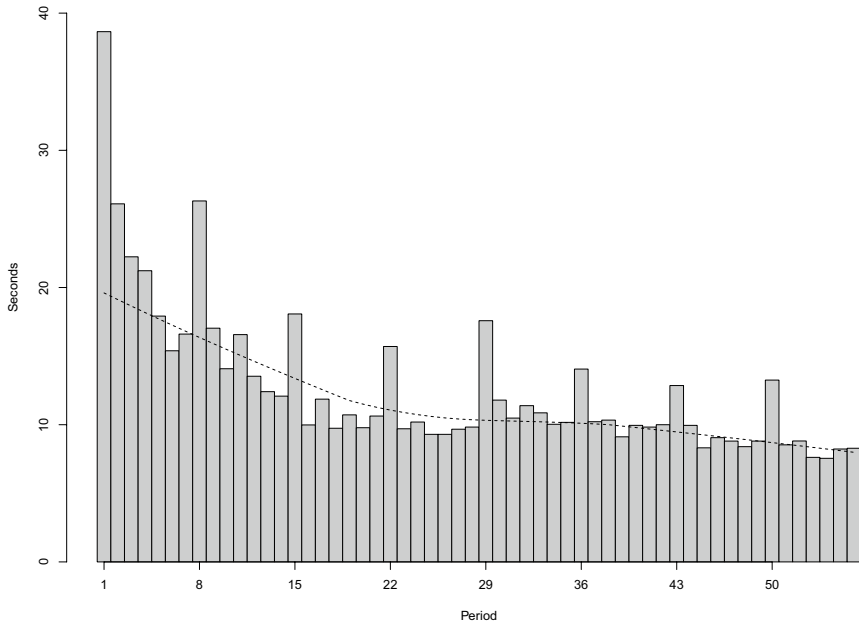[10] See the translated instructions in Appendix A for more details

Fig. 2 Decision time

together in groups of 8. Each part has four phases: *Phase 1*, *Phase 2*, *Phase 3*, and *Phase 4*. In each phase, participants choose how to invest in the available prospects, facing one of the four sets in Table 2 over seven successive rounds. Which set is displayed in which phase is randomly defined at the matching group level, but all subjects see all sets, and within a phase, the set is kept constant. In Part 1, when groups are made of two participants, individuals interact in matching groups of size 8. In part 2, the matching groups include 16 subjects each. The composition of the group does not change within a phase, but is randomly redefined from phase to phase.

Overall, participants face 56 successive decision tasks. The final payment in the experiment is given by randomly drawing one round from each of the four phases in Part 1 and one round from each of the four phases in Part 2.

The experiment was run at the laboratory of the Max Planck Institute Jena, Germany. The experiment was programmed and conducted using zTree software (Fischbacher 2007). A total of 128 participants took part in the experiment over four experimental sessions. In total, we collected 7168 individual portfolio choices, $s(i)$.

5 Results

First, we analyze decision times for completing the task to understand how participants familiarize themselves with the task. Then, we present investment choices and their deviation from probability matching behavior. In Sect. 5.3, we investigate

investment dynamics with specific attention to alternative regret measures. Finally, regression analysis provides us with a general assessment of our main results.

5.1 Decision times

The graph (Fig. 2) represents the distribution of median decision times in seconds in each round of the experiment.

Participants become much faster in choosing as they familiarize themselves with the task, testified by the polynomial fitting of the data (dashed line). Spikes correspond to periods $1 + 7x$ with $x \in [0, 7]$ and highlight the adjustment needed when a new experimental condition, a new probability vector w and/or a larger number of investors, is announced and portfolios are adjusted what, as such, indicates forward-looking inclinations. A Wilcoxon signed-rank test comparing choice times in the first and in the last period shows that choice times significantly decrease over time (p value< 0.001).

Thus, our data show that proficiency in the task increases over rounds as measured by time to choose. Time spent in the choice is generally higher in the first round after announcing a new experimental condition.

5.2 Investment choices

Figure 3 displays how deviations from probability matching are distributed for each prospect whose probability is given below. The symbol "x" denotes the average deviations and medians are indicated by bold lines.

The graphs show that median choices are generally close to probability matching. However, the distribution of tokens is widely dispersed, and several outliers are present in each condition of choice. Choices clearly reveal that commonly known numerical probabilities provide a much stronger cognitive anchor than potential alternatives, like the Golden Mean or the 1/m-heuristic, especially in the case of the prospect sets with probabilities differing considerably from 1/m. From this, one may infer that the 1/m-heuristic may require ambiguity of success probabilities to emerge.

A series of non-parametric tests shows that the central tendency of the token distribution in the prospect with the lowest probability is significantly different from probability matching, at the conventional 5% level, for Set 2 and 3 when the group size is equal to 4 and for Set 1 when the group size is equal to 8. When pooling data

Group size=4
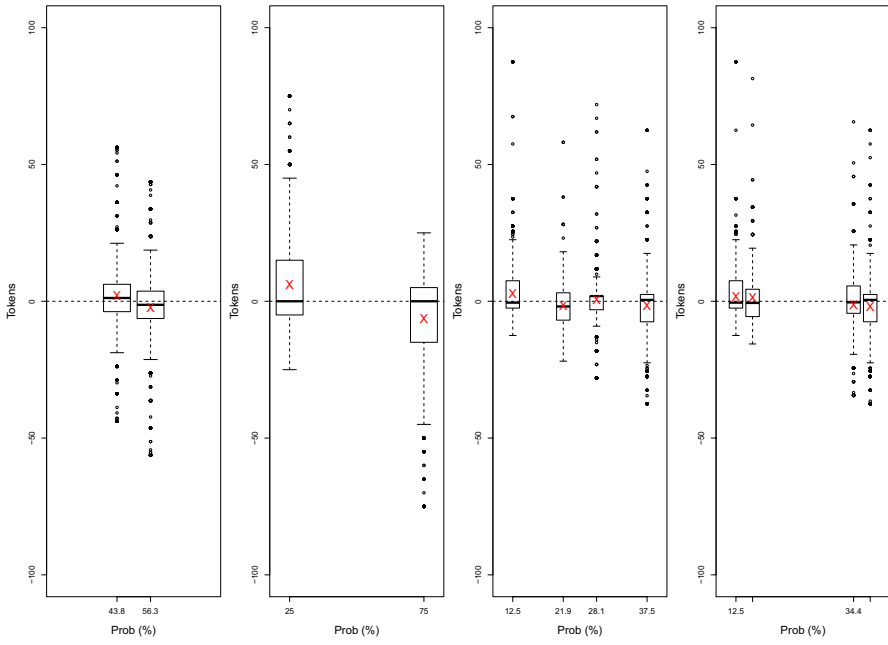
Group size=8
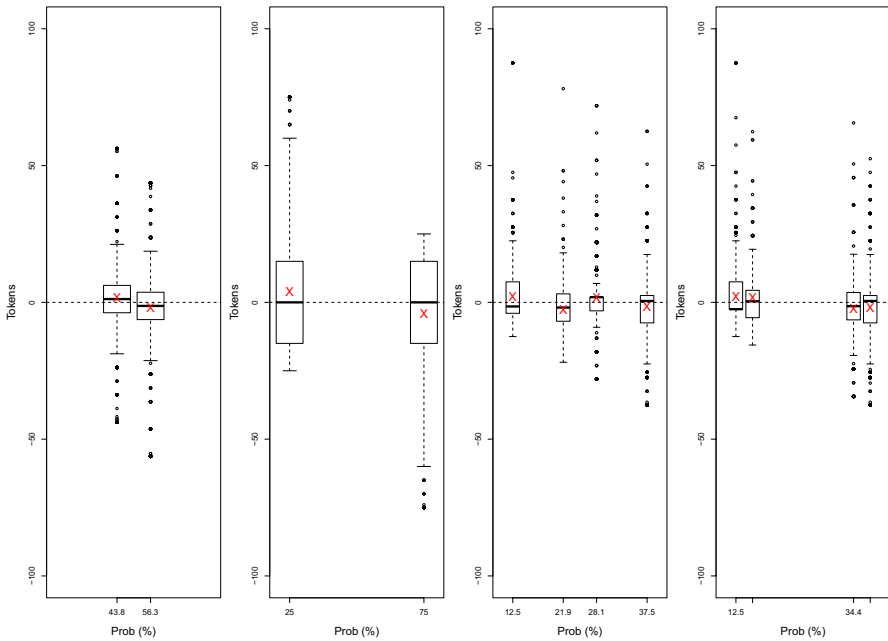
irrespective of set and group size, a (positive) statistically significant difference is registered (p value<0.001).[11] In terms of investment dynamics, no significant differences are observed in token allocations, when comparing the first and the last round in each block with seven successive choices (all p values ≥ 0.152). Table 3 presents a few summary statistics of absolute deviations from equilibrium prices in each set and for each group size.

Individual heterogeneity in token allocation does not average out in terms of market prices. Actually, prices, a measure of group behavior, are far from their equilibrium levels in each set and for each group size. Furthermore, standard deviations are quite sustained in all experimental conditions. Deviations are generally larger for larger group sizes, as confirmed by a series of Wilcoxon rank-sum tests comparing averages at the matching group levels across the two group sizes (all p values ≤ 0.019).

As a complement of the "static" analysis presented above, Fig. 4 provides us with a representation of the evolution of probability matching over rounds. The graph reports the average absolute difference between the share of tokens invested in a prospect by a participant ($s_j(i)$) and the probability of the prospect ($w_j(i)$), over the seven rounds.

According to Fig. 4, the average distance from probability matching tends to increase over time. A Wilcoxon signed-rank test confirmed this, showing that choices in the first and the last round are significantly different (p value< 0.001, test on averages at the matching group level).

5.3 Investment dynamics

5.3.1 Individual-level changes in drawn asset

Figure 5 shows the average variations at the individual level in the share of tokens allocated to the prospect drawn in the previous period ($\Delta s_{\hat{k}}$). A positive (negative) value denotes individuals who, on average, increase (decrease) the amount allocated to the winning asset of the previous round. The vertical dashed line captures the average change.

According to Fig. 5, the average change is negative but small: most of the changes (78.1%) lie in the interval -0.05/+0.05. Non-parametric tests show that the central tendency of the distribution is negative, either when conditioning the tests at the

---

[11] The tests take into account differences between tokens allocated to and probability of the asset with the lowest probability of realization ($s_1(i) - w_1$). This approach was chosen in light of the nature of the choices at hand. Choosing only one allocation allows us to compare prospects with $m = 2$ and $m = 4$, using a common reference value. Furthermore, considering deviations for all assets would just make the test irrelevant as deviations cancel out ($\sum_j^m (s_j(i) - w_j) = 0$). At the same time, absolute differences ($|(s_1(i) - w_1)|$) would bias the test by rendering it heavily asymmetric around the null hypothesis. Finally, to preserve statistical independence, the Wilcoxon signed-rank tests whose p values are reported in the text are run on median choices at the matching group level.

Table 3 Deviations from equilibrium prices

Set	Mean	Median	SD
Size = 4			
1	0.621	0.500	0.453
2	0.699	0.500	0.577
3	0.356	0.276	0.268
4	0.324	0.256	0.237
Size = 8			
1	0.811	0.620	0.698
2	1.284	1.050	1.018
3	0.610	0.497	0.358
4	0.607	0.498	0.374

matching group level or at the individual level (p value= 0.001 and p value< 0.001, respectively).

The pattern emerging from Fig. 5 suggests that more investors tend to be discouraged from investing in the previously drawn prospect, whereas only the opposite adjustment (or no adjustment at all) is compatible with the notion of directional learning. The observed adjustment may relate to a well-known bias in decision-making under uncertainty, i.e., the gambler's fallacy, which denies the independence of successive iid-chance moves.

5.3.2 Directional learning

Figure 6 presents the change in the share invested in the winning prospect $\hat{k}$ of the previous round: $\Delta s_{\hat{k}}(i) = s_{\hat{k}}^t(i) - s_{\hat{k}}^{t-1}(i)$, where $s_{\hat{k}}^t(i)$ is the share of tokens invested by subject i at time t in the winning asset $\hat{k}$ at time $t-1$ and $s_{\hat{k}}^{t-1}(i)$ is the share of tokens at time $t-1$.[12] A change in the share invested in the previously drawn prospects is compatible with directional learning if the share allocated at time t is increased relative to the share at time $t-1$. White bars capture adjustments compatible with directional learning,[13] i.e., when the change is positive ($\Delta s_{\hat{k}}(i) > 0$), light gray bars when there is no change ($\Delta s_{\hat{k}}(i) = 0$), and dark gray bars capture negative changes ($\Delta s_{\hat{k}}(i) < 0$).

Overall, the majority of changes from round to round are not compatible with directional learning. For the groups of size 4, the average frequency of positive changes is 26.43%; the average frequency of negative changes is 29.82%; the remainder 43.75% is due to inertia. A similar pattern emerges for group size 8, with

---

[12] All measures are computed separately for experimental treatments differing in terms of prospect sets and group size. Accordingly, for each condition, six observations are available at the individual level.

[13] Unlike Selten and Buchta, we do not view inertia, $\Delta s_{\hat{k}}(i) = 0$ as support of directional learning, but rather compare only the changes in the predicted direction with those in the opposite direction.

positive changes covering 22.59% of the observations, negative changes 27.08%, and absence of change 50.33%. Altogether, no adjustment dominates and increases its share as rounds progress. It seems that over time participants appreciate the true randomness of winning chances and become less sensitive to the last random event.

Concerning the evolution of changes over rounds, a series of non-parametric tests find a significant decrease in positive variations from the first available round to the last one, both for group size 4 and 8, as shown by Wilcoxon signed-rank tests at the matching group level (p value=0.021 and p value=0.014, respectively). At the same time, inertia significantly increases across rounds, both for group size 4 and 8 (p value=0.017 and p value=0.014, respectively). Furthermore, the difference between the share of negative changes in the first and last round is statistically significant for group size 8 (p value=0.014) but not for group size 4 (p value=0.417). Altogether, the general tendency is a progressive shift from active change, in both directions, to the invariance of choices.

5.3.3 Regret analysis

To better understand the adaptation dynamics of token allocation, we distinguish four alternative sources of regret which may correlate with a positive change in the token allocation to the winning asset $\widehat{k}$ in the previous round ($\Delta s_{\widehat{k}} > 0$ or $\Delta^+ s_{\widehat{k}}$, see above). The analysis relies on 1506 out of the 6144 available observations (24.5%).

The first measure we adopt is $regret\_payoff_t(i) = R\frac{1}{1+\sum_{j\neq i} s_{\widehat{k}}^{t-1}(j)} - R\frac{s_{\widehat{k}}^{t-1}(i)}{\sum_{j=1}^{n} s_{\widehat{k}}^{t-1}(j)}$, the difference between the payoff individual i would have earned had she invested all her resources in the previously drawn asset $\widehat{k}^{t-1}$ and the amount actually earned by i in period $t-1$. Second and third we adopt $regret\_tokens_t(i) = 1 - s_{\widehat{k}}^{t-1}(i)$, capturing the share not invested in the drawn asset, and $regret\_others_t(i) = \frac{1}{n-1}\sum_{j\neq i}^{n} s_{\widehat{k}}^{t-1}(j) - s_{\widehat{k}}^{t-1}(i)$, the difference between the average amount invested by the others in one's (market) and share invested by individual i in the drawn asset.[14] The fourth measure is $regret\_probabilities_t = w_{\widehat{k}}^{t-1} - s_{\widehat{k}}^{t-1}(i)$, the difference between the exogenously given probability $w_{\widehat{k}}$ and share invested in the drawn asset.

According to Table 4, all regret measures are positively and significantly correlated to changes in token allocation to the last drawn assets $\Delta^+ s_{\widehat{k}}$, except for $regret\_payoff$. Table 4 also highlights strong correlations among alternative regret measures. In particular, $regret\_others$ and $regret\_probabilities$ are strongly correlated (rho=0.922). This is most likely due to the fact that the mean investment in a prospect is very close to the probability of the winning prospect (see Fig. 3).

---

[14] In principle, past asset price information allows to infer which token amounts were invested in total by others, but inverting the price equations is cognitively demanding and very unlikely. Nevertheless, participants may still qualitatively infer others' average behavior from price vector information.
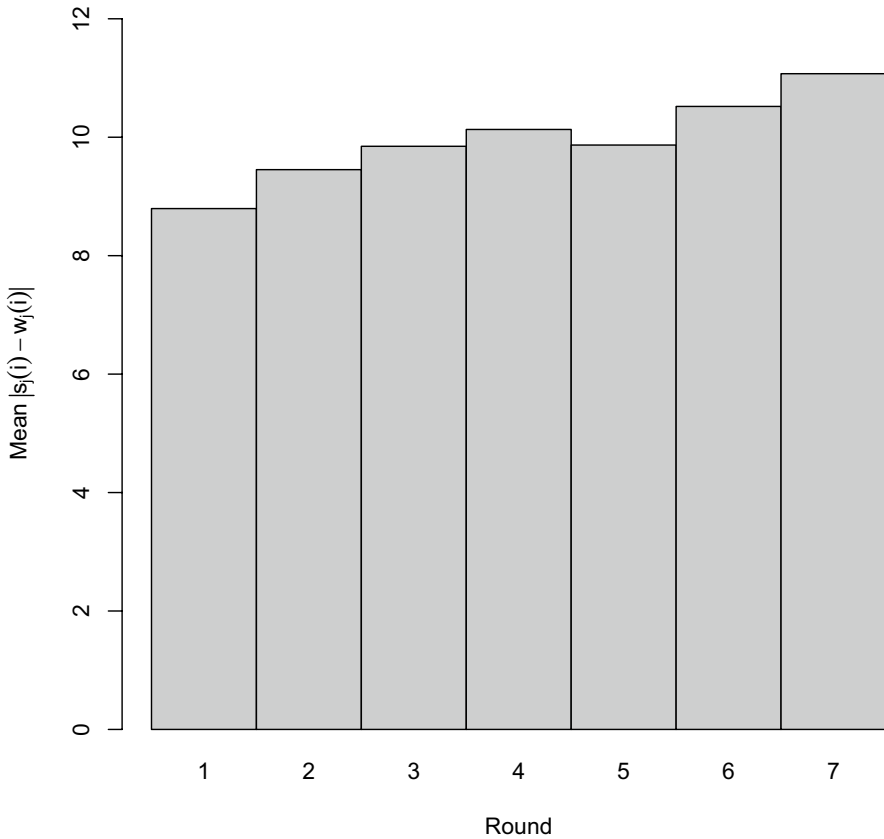
Fig. 4 Average distance from probability matching over rounds

5.4 Regression analysis

The multivariate regression analysis of Table 5 provides a more refined picture of the relative impact of the alternative regret sources. To corroborate results emerging from the analysis reported above, we present four regression estimates with different model specifications, both in terms of dependent and independent variables as well as in terms of sample size. Regressions (1), (2), and (4) of Table 5 report the outcome of linear mixed models (LMM). Regression (3) is a logit generalized linear mixed model (GLMM Logit). All models are estimated with random effect intercepts at individual and matching group levels.

In Regression (1), the dependent variable is the absolute difference between the share of tokens invested in a prospect by a participant ($s_j(i)$) and the probability of the prospect ($w_j(i)$), which we interpret as a direct measure of probability matching.
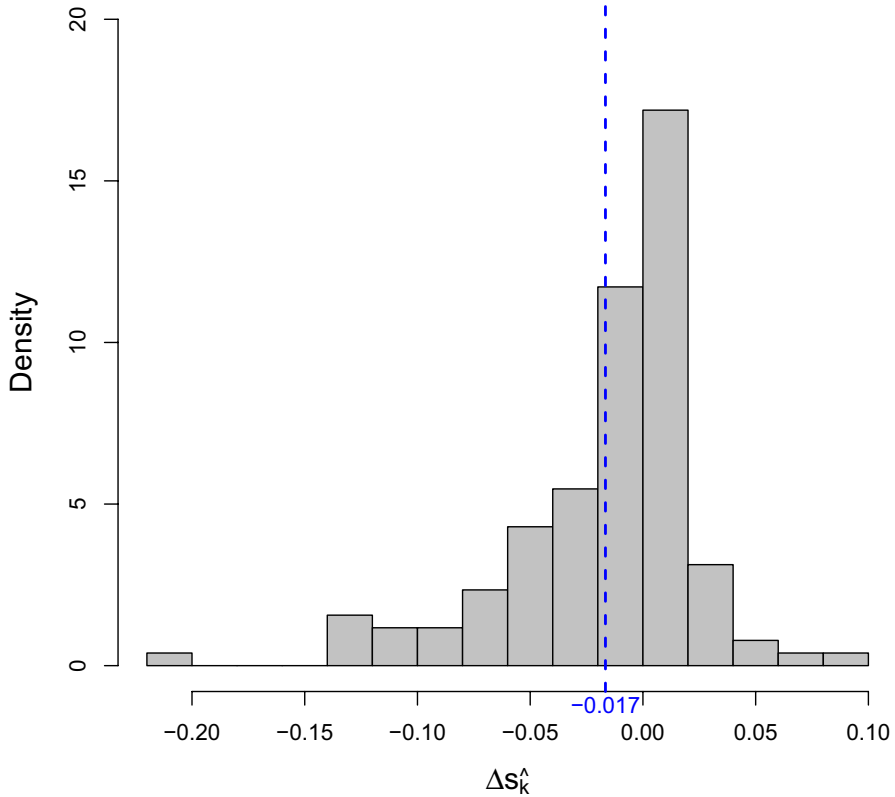
Fig. 5 Changes in allocation to drawn prospect

In Regression (2), the dependent variable is the expected payoff of a participant in each round of choice, given the actual choices of the others in the group ($\xi[\Pi] = \sum_1^m w_j \pi_j$, where m is the number of prospects in a set and π_j are the earnings in prospect j). This measures the expected profitability of one's allocation strategy.

In Regression (3), the dependent variable $I(\Delta^+ s_{\hat{k}}(i)$ is equal to 1 if the difference between the amount invested at time t and at time $t-1$ in the last winning prospect $\hat{k}$ ($\Delta s_{\hat{k}}(i) = s_{\hat{k}}^t(i) - s_{\hat{k}}^{t-1}(i)$) is strictly positive and thus compatible with directional learning. Otherwise, the dependent variable is 0, i.e., if $\Delta s_{\hat{k}}(i) = s_{\hat{k}}^t(i) - s_{\hat{k}}^{t-1}(i) \leq 0$.

In Regression (4), the dependent variable is the difference between the amount invested at time t and at time $t-1$ in winning asset $\hat{k}$ at round $t-1$. ($\Delta s_{\hat{k}}(i) = s_{\hat{k}}^t(i) - s_{\hat{k}}^{t-1}(i)$). The dependent variable captures the investment reaction to the previous random event, $\hat{k}$ in t-1. Only positive adjustments are considered, to shed light on regret balancing.

As explanatory variables, experimentally controlled manipulations are employed across all three model specifications: *Round* captures the period of choice; $Size_8$ is a dummy variable taking value 1 when the group size is 8 and value 0 when it is 4; Set_i with $i \in \{1, 2, 3, 4\}$ is a dummy variable equal to 1 when choices are from set i and equal to 0 otherwise.

In Regression (2), a measure of absolute relative deviation from probability matching for each portfolio of choices is computed ($Abs\_Rel\_Diff = \frac{|s_j(i) - w_j(i)|}{w_j(i)}$). In Regression (4), a few additional controls are added to capture alternative sources of potential regret related to choices in the previous period. A detailed description of these measures is provided above (see Sect. 5.3).

Regression (1) reveals that the distance from probability matching tends to increase over rounds. Interestingly, in Set 2, for which the 1/m-heuristic is much closer to probability matching, the shares are further away from probability matching than in Set 1. Furthermore, a Linear Hypothesis test shows that deviations are statistically larger in Set 3 than in Set 4 (Chi-square test, *p* value= 0.037).

Regression (2) shows that payoffs in Set 1 are lower than those in other sets. Furthermore, a larger distance from probability matching implies a considerable decrease in expected payoffs.

Our findings of the relevance of *probability matching* are summarized in the following result:

Result 1 Overall, probability matching seems to provide a strong anchoring for investments decisions. However, deviations from probability matching increase over rounds and are larger for smaller sets of options. This has an impact on outcomes, as larger deviations from probability matching impact expected payoffs negatively.

Regression (3) confirms that the likelihood of adjusting in line with directional learning decreases over rounds. The tendency towards directional learning is weaker in larger groups than in smaller ones ($Size_8$). Furthermore, sets with more investment options are less likely to induce behavior compatible with directional learning (Set_3 and Set_4).

Our findings of the relevance of *directional learning* are summarized in the following result:

Result 2 Overall, directional learning is not a main behavioral driver, and its strength is even weaker in later rounds, for larger groups, and larger sets of options.

Regression (4) highlights the impact of alternative regret measures on changing one's allocation: both the amount of tokens not invested in the drawn prospect (*regret_tokens*) and the distance between the probability of the drawn asset and the share invested in it (*regret_probabilities*) positively affect how much is invested in the drawn asset. Interestingly, the distance from the behavior of others, which has to be inferred from past market-clearing prices, has no significant impact, what had to be expected, since such inference is cognitively demanding. The difference between the payoff that one would have obtained by investing in the drawn asset

Fig. 6 Round-to-round changes in allocation to drawn prospect ▶

and the actual payoff (*regret_payoff*) negatively impacts on behavior what seems compatible with the intuition of directional learning. However, these results must be taken with caution because of the strong correlation between *regret_probabilities* and *regret_others* (see Table 4) and the potential issue of multicollinearity.

Our findings of the relevance of alternative *regret measures* are summarized in the following result:

Result 3 Regret related to foregone payoffs impacts portfolio adjustments, in line with directional learning. Furthermore, the distance of the investment from the probability of the drawn asset and the distance from entirely investing in it also impact portfolio adjustments.

6 Discussion

Our analysis investigates the choice behavior of enormous field relevance, namely how to invest in different financial assets. But then, we seem to lose all field relevance, partly in line with the theoretical finance literature, by presupposing and experimentally inducing numerically specified and commonly known winning chances. If at all, winning chances in the field are ambiguous, and many professional investors claim individual superiority in knowing them. Interestingly, the evolutionary justification of the benchmark $s(i) = s^*$ for $i = 1, \dots, n$ only does require sufficient stationarity of winning chances to allow for evolutionary selection.

In our view, it is surprising that investors' awareness of numerical winning chances w_j for $j = 1, \dots, m$ seems to crowd out heuristics like, for instance, the Golden Mean (1/m). Although this appears like an obvious default in case of considerably blurred winning chances, it is even more astonishing that this crowding out is immediate. The 1/m anchoring is substituted by w-anchoring, at least when w is commonly known. But then, the common knowledge and cognitive demands of the two equilibrium justifications of $s(i) = s^*$-behavior for all $i = 1, \dots, n$ question seriously their behavioral relevance, especially its psychological validity. Could the average choice tendencies in our data, which are quite in line with this behavior, be due to reaction to myopic regret? This means to retrospectively analyze what would have been best, or at least better, choice behavior in the last period and to assess how much one has actually lost compared to this in the last period. Obviously, the latter can never exceed the former, and a positive difference measures a loss, assessed via a retrospective choice analysis when knowing the last winning asset.

Compared to other investment scenarios, our setup is simple via its block-wise stationary vector of state-specific probabilities across all seven periods of a phase but also much more complex than usual portfolio-choice experiments, since investors endogenously determine the market-clearing prices for markets with 4, respectively 8, interacting agents. The initial steep decline in decision times reveals that the task is at least initially cognitively demanding. The prominent role of probability
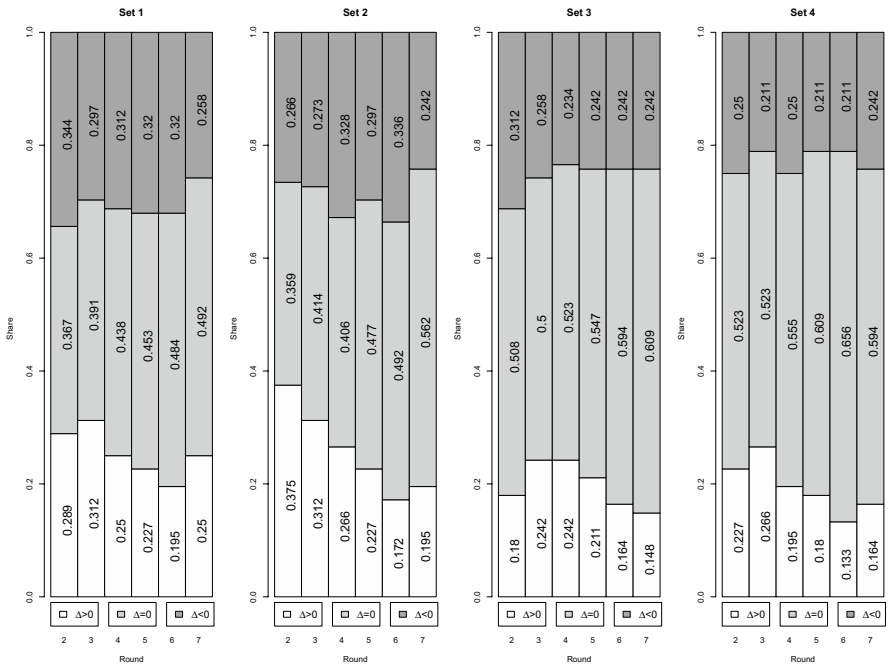
Group size=4
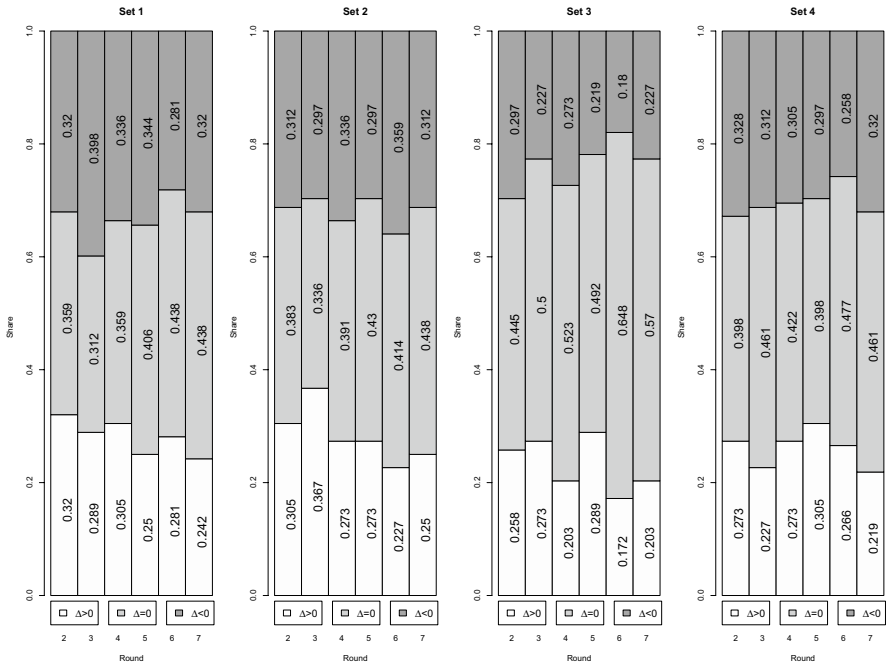
Group size=8

Table 4 Potential regret sources of behavioral adaptation (Pearson's product–moment)

	$\Delta^+ s_{\hat{k}}$	regret_payoff	regret_others	regret_probabilities
regret_payoff	0.018			
regret_others	0.517***	0.242***		
regret_probabilities	0.530***	0.222***	0.922***	
regret_tokens	0.232***	0.782***	0.585***	0.537***

Table 5 Mixed models estimations (random effect intercepts at individual and matching group levels)

	(1) $|s_j(i) - w_j(i)|$	(2) $\xi[\Pi] = \sum_j^m w_j \pi_j$	(3) $I(\Delta^+ s_{\hat{k}}(i))$	(4) $\Delta^+ s_{\hat{k}}(i)$
(Intercept)	0.117 (0.007)***	2.577 (0.008)***	−0.591 (0.141)***	0.074 (0.018)***
Round	0.003 (0.000)***	0.004 (0.001)***	−0.105 (0.018)***	0.000 (0.002)
$Size_8$	0.001 (0.004)	−1.247 (0.008)***	−0.233 (0.110)*	0.001 (0.006)
Set_2	0.030 (0.002)***	0.056 (0.005)***	0.015 (0.086)	0.007 (0.008)
Set_3	−0.054 (0.002)***	0.030 (0.005)***	−0.318 (0.089)***	−0.072 (0.010)***
Set_4	−0.051 (0.002)***	0.042 (0.005)***	−0.243 (0.088)**	−0.080 (0.009)***
Abs_Rel_Diff		−35.313 (0.626)***		
regret_payoff				−1.936 (0.552)***
regret_others				0.032 (0.049)
regret_probabilities				0.251 (0.046)***
regret_tokens				0.160 (0.037)***
Model	LMM	LMM	GLMM Logit	LMM
Num. obs.	21504	7168	6144	1506
Num. indiv.	128	128	128	123
Num. groups	24	24	24	24

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, ·$p < 0.1$

matching in explaining investment choices, for instance, vis-a-vis the 1/m or Golden mean heuristic, via allocating wealth shares questions that one changes one's portfolio in the light of myopic feedback information about the most recent results as, for instance, postulated by direction learning. Apparently participants develop more stable investment behavior after more experiences with random success, revealed by increasing inertia across time. Less myopic path dependence, as in reinforcement learning, could account for such increasing inertia share. In our regret analysis, we restricted ourselves to myopic path dependence but, as in direction learning, distinguished several obvious and direct regret measures whose effects on portfolio adjustments in the light of only the most recent results could be confirmed.

Altogether, our results support the prominent and outstanding theoretical prediction of *w*-homogeneity in investing—a prediction supported by evolutionary

stability, strategic equilibrium, anchoring, and regret balancing. It captures the main tendencies even after block-wise changing probability vectors, but does not account for the persistent heterogeneity, which we could at least partly explain by myopic regret effects.

In conclusion, our stylized setup offers insights into policy interventions in real financial markets. Therefore, one might pool all available information and provide it publicly and enhance thereby block-wise stationarity of winning chances. It would also render the assumption that all investors rely on the same winning chances, a crucial assumption of our setup, more realistic and could render financial markets less perilous by limiting exploitation by privately informed traders.

Appendix

A. Translated instructions

General instructions

[HERE SOME STANDARD ABOUT THE EXPERIMENT AND THE SHOW-UP FEE]

If you have any question, please raise your hand and we will answer to it privately. Before the experiment starts, you are asked to answer a few questions checking the understanding of these instructions. The experiment will start only when all participants have correctly answered the control questions.
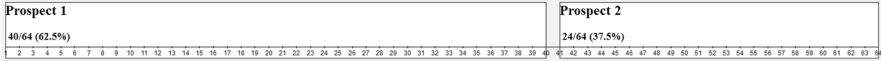
Your choice task

Prospects

During the experiment, you are going to make choices involving prospects.

A prospect has a certain probability of being selected by the computer and when a prospect is selected by the computer, it generates some positive returns. Otherwise, it generates nothing.

Distinct prospects are represented on the screen of your computer via bars of varying length, indicating their likelihood of being selected. The following figure provides you with an example of a situation in which there are two prospects, Prospect 1 and Prospect 2. The number of prospects can change during the experiment and can be either equal to 2 or to 4.

Prospect 1	Prospect 2
40/64 (62.5%)	24/64 (37.5%)
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40	41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64

The length of the bar graphically represents the numerical probability that a prospect is selected by the computer and generates positive returns. The longer the bar, the higher the probability that the corresponding prospect is selected. The numerical probability is also provided in correspondence to each bar.

The procedure adopted by the computer to select a prospect is as follows:

- an integer number in the range from 1 to 64 is randomly drawn with all numbers from 1 to 64 being equally likely. The same number is drawn for all participants.
- when the number "falls" in a bar representing prospect probabilities, the prospect associated with the bar is selected and generates positive returns.

For the figure above, the numbers from 1 to 40 select Prospect 1, whereas all numbers from 41 to 64 select Prospect 2. Thus, exactly one prospect can be selected to generate positive returns.

Tokens

During the experiment, you are matched with other participants to form a group. The size of the group can change during the experiment and can be either equal to 4 or to 8.

You and every other member of your group are endowed with 100 tokens.

Your task is to allocate the 100 tokens to the available prospects. You and the other in your group must allocate all 100 tokens to the available prospects.

Prospects bought

For each prospect i, there are 100 units to be bought by you and the others in your group.

The number of units of prospect i you buy (u_i) is given by the proportion between the tokens allocated to the prospect by you (t_i) and the total tokens allocated to the prospect by your group (T_i).

In formal terms, the number of units of prospect i you buy is given by $u_i = 100\frac{t_i}{T_i}$. It will be possible to buy non-integer units of prospects.

Returns

When a prospect is selected by the computer (i^*), it generates positive returns.

Your returns R from the selected prospect i^* are equal to €0.10 multiplied by the number of units of the selected prospect you bought (u_{i^*}).

In formal terms, the returns from the selected prospect are given by $R = u_{i*} \times$ €0.10. Units of non-selected prospects become worthless.

Parts, phases, and rounds

The experiment is made of 2 parts: Part 1 and Part 2.

Each part is made of 4 phases: Phase 1, Phase 2, Phase 3, and Phase 4. Prospects may change between phases, but not within a given phase.

Each phase consists of seven rounds. In each round, you will be endowed with 100 tokens and asked to allocate your tokens to available prospects.

Overall, you will face 56 rounds.

Throughout a phase, you are matched with the same participants, but between phases the composition of the group is randomly changing.

Detailed information about the prospects and the size of your group are given to you on the computer screen. You are also constantly informed about the Part, Phase, and Round you are in.

Final payment

The final payment is determined by randomly picking one round from each of the four phases in Part 1 and one round from each of the four phases in Part 2.

Thus, your final earnings in the experiment are the sum of your earnings in the 8 rounds randomly chosen by the computer. You will learn which rounds have been chosen only at the end of the experiment.

Example

This example illustrates the procedure of the experiment, but is not meant to give you any indication about how to behave in the experiment.

Consider the prospects in the figure above.

Assume you allocate 50 tokens to Prospect 1 ($t_1 = 50$) and 50 tokens to Prospect 2 ($t_2 = 50$).

Now, assume that the total number of tokens allocated by the four in your group (you included) to Prospect 1 and Prospect 2 is $T_1 = 250$ and $T_2 = 150$, respectively. Accordingly, you succeed in buying $100 * \frac{50}{250} = 20$ units of Prospect 1 and $100 * \frac{50}{150} = 33.3$ units of Prospect 2.

If a number smaller than or equal to 40 is randomly drawn, Prospect 1 is selected by the computer and you earn €.10 × 20 = €2.0. If a number larger than 40 is selected, Prospect 2 would be chosen and you would earn €0.10 × 33.3 = €3.3.

Declarations

References

Avrahami J, Güth W, Kareev Y (2005) Games of competition in a stochastic environment. Theor Decis 59(4):255–294

Benartzi S, Thaler RH (2001). Naive diversification strategies in defined contribution saving plans. Am Econ Rev 79–98

Blume LE, Easley D (1992) Evolution and market behavior. J Econ Theory 58(1):9–40

Fama EF (1970) Efficient capital markets: a review of theory and empirical work. J Financ 25(2):383–417

Fischbacher U (2007). z-Tree: Zurich toolbox for ready-made economic experiments Exp Econ (2):171–178

Kroll Y, Levy H, Rapoport A (1988) Experimental tests of the mean-variance model for portfolio selection. Organ Behav Hum Decis Process 42(3):388–410

Ockenfels A, Selten R (2005) Impulse balance equilibrium and feedback in first price auctions. Games Econom Behav 51(1):155–170

Samuelson L (2002) Evolution and game theory. J Econ Perspect 16(2):47–66

Selten R (2004) Learning direction theory and impulse balance equilibrium. Econ Lab 133

Selten R, Buchta J (1999) Experimental sealed bid first price auctions with directly observed bid functions. In: Budescu D, Erev I, Zwick R (eds) games and human behavior: essays in the honor of Amnon Rapoport. Lawrenz Associates, Mahwah, NJ

Stone R (1954) Linear expenditure systems and demand analysis: an application to the pattern of British demand. Econ J 64(255):511–527

Weibull JW (1997) Evolutionary game theory. MIT Press, Boca Raton