

Zankiewicz, Christian; Ensthaler, Ludwig; Nottmeyer, Olga; Weizsäcker, Georg

Conference Paper

Hidden skewness: On the difficulty of multiplicative compounding under random shocks

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Information, Risk, and Uncertainty in Finance, No. B23-V4

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Zankiewicz, Christian; Ensthaler, Ludwig; Nottmeyer, Olga; Weizsäcker, Georg (2015) : Hidden skewness: On the difficulty of multiplicative compounding under random shocks, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Information, Risk, and Uncertainty in Finance, No. B23-V4, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/112815>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Hidden skewness: On the difficulty of multiplicative compounding under random shocks*

Ludwig Ensthaler Olga Nottmeyer Georg Weizsäcker
and Christian Zankiewicz

February 13, 2015

Abstract

Multiplicative growth processes that are subject to random shocks often have an asymmetric distribution of outcomes. In a series of incentivized laboratory experiments we show that a large majority of participants either strongly underestimate the asymmetry or ignore it completely. Participants misperceive the outcome distribution's spread to be too narrow-band and they estimate the median and the mode to lie too close to the distribution's center. The observed bias in expectations is irrespective to risk preferences and it appears under a variety of conditions regarding feedback, incentive size, and market contexts. The bias is largely consistent with a behavioral model in which geometric growth is confused with linear growth. This misperception is a possible explanation of investors' difficulties with real-world financial products like leveraged ETFs and is also related to issues regarding real estate investments, retirement savings plans or investments in college funds.

*We thank Erik Eyster, Dorothea Kübler, Erik Mohlin, Peter Mörters, Tobias Schmidt, Adam Szeidl and Heinrich Weizsäcker and audiences at DICE, DIW Berlin, ESMT and the IEA World Congress 2014 for helpful comments and we thank colleagues at the decision laboratories of Technical University Berlin and University College London, especially Brian Wallace and Mark Henninger, for their excellent contributions in the preparation and conduct of the experiments. Financial support by the ERC (Starting Grant 263412) is gratefully acknowledged. Ensthaler: Humboldt University Berlin and WZB Berlin, Nottmeyer: IZA Bonn, Weizsäcker: Humboldt University Berlin and DIW Berlin, Zankiewicz: DIW Berlin. Contact: weizsaecker@hu-berlin.de.

1 Introduction

Many household investors face a particular mismatch in the time frames of asset return evaluations. They acquire their most important financial assets with the intention to liquidate them in the relatively distant future but the available return information concerns much shorter time intervals. Real estate investments, retirement savings plans or investments in college funds share this feature. In all of them, the relevant outcomes are the investments' performances over several decades but the available information concerns their short-term performances, like 1-year returns. To forecast the return on the planned (or any plausible) distant selling date, an investor needs to extract the price distribution at the selling date by compounding the available short-term return distributions. This is a formidable task for the average person.

It is well documented that decision-makers both in the laboratory and in the field have difficulties when compounding deterministic growth processes—the so-called exponential growth bias. When asked to estimate the final value of an account that accumulates 7% interest for ten years a substantial fraction of respondents give an answer that is closer to 70% than to the actual 97%. The analysis of Stango and Zinman (2009) indicates that the bias is empirically relevant as it affects households' borrowing and saving decisions. Levy and Tasoff (2014) show that the exponential growth bias can lead to overconsumption, via a misperception of relative prices of consumption, and detect a statistical connection between the bias and savings behavior. Neglecting compound growth effects plausibly has adverse consequences.

But when growth occurs at a random speed, the cognitive errors in compounding may have another, even uglier face. In random processes the decision maker needs to generate a probability distribution over all possible random paths, which can lead to new misunderstandings. *All moments* of the return distribution are affected by compounding in a nonlinear way. In particular, compounding transforms a symmetric 1-period distribution into a skewed multi-period distribution. As an example consider a relatively new class of retail financial products, so-called leveraged exchange-traded funds (leveraged ETFs). These assets have a highly volatile and fairly symmetric 1-period volatility; holding them for multiple periods results in severe skew.

This paper presents a series of incentivized laboratory experiments that extend the research on the multiplicative growth bias to the stochastic domain and test the decision makers' ability to compound a sequence of multiplicative random shocks. We provide our participants with the 1-period return distribution of an asset and give incentives to compute the distribution of its selling price after a specified number of periods. As an illustration, consider a stylized experiment that we present in detail in Section 3.3. A very volatile asset either increases in value by 70% or decreases in value by 60% in every period, each growth rate realizing with a chance of one half. If the investor buys the asset she must hold it for twelve periods. With an initial investment of 10,000 EUR, what would the asset likely be worth at the end of period 12? To ask this question in an incentive-compatible way, we let the participants bet on five possible outcome ranges for the period-12 value of the asset: a) up to 6400, b) between 6400 and 12800, c) between 12800 and 19200, d) between 19200 and 25600, or e) above

25600. We then simulate the process for each of the participants and if the simulated path ends up in the outcome range that the participant has bet on, she receives a bonus of 20 Euros. If not, she receives nothing. The most popular answer is c), chosen by 43% of participants, followed by d) (28%) and b) (17%). Response options a) and e) come tied bottom with a mere 6% of responses each. However, the optimal response is a): chances that the process ends up in this interval is as high as 80%; the median of the resulting distribution is 989 EUR. The reason for a) being the optimal response is that a value increase of 70% cannot recover a value decrease by 60%, hence most trajectories have a downward trend and the distribution is highly skewed already after 12 periods. The participants fail to realize this resulting in an average expected payoff (based on their decisions) which amounts to a meager 2 Euros in this experiment, whereas the optimal response would earn them 16 Euros in expectation.

Our series of experiments examines this kind of mistake systematically, finding that participants' perception of stochastic growth deviates in predictable ways from the rational prediction. The experimental results are in line with a simple model of misperception of compounding of shocks. This model, which we label "linearity bias model", extends the hypothesis that the agent fails to do deterministic compounding to stochastic settings. It stipulates that a biased decision maker perceives a linear evolution in the sense that she perceives the distributions of absolute changes as constant over time, instead of the relative changes being constant over time. (In spirit, the model is akin to that of Levy and Tasoff, 2014, albeit developed independently.) In effect, all multiplicative growth is mistaken as additive growth with a constant distribution of increments. The model thereby predicts that the agent ignores all skewness that arises from multiplicative compounding. If the distribution of an asset's per-period returns is symmetric, then the agent perceives all longer-period returns as symmetric, too. The model also predicts larger deviations from the rational benchmark for longer investment horizons and for higher levels of return volatility. The deviations are predicted to occur in terms of misperceptions of the mode, the median, the spread, and the skewness.

The results of Section 3.2 show that, consistent with the linearity bias model, participants overestimate the median of the resulting distribution of prices significantly and strongly. The large majority of participants overestimate the median in the 70/60 example by an order of magnitude. Section 3.3 reports the above-described test showing that the perceived mode of the same distribution is also severely biased.

The experiments of Sections 4 and 5 test the predictions of the model in more realistic settings. We use both binomial-tree assets (like the above-described 70/60 example) with lower levels of volatility as well as actual historical data on the German DAX index to test for a participant's perception of the median as well as the 10th and 90th percentile of the distribution of long-run investments. All results are irrespective of the participants' risk preferences: the experimental designs of Section 4 and 5 each have two investment options involving different probabilities of winning a fixed bonus, which renders risk preferences irrelevant.

The findings in these experiments confirm the linearity bias for most participants. In all treatments with stochastic growth, the majority of participants reveal a perceived skew of the outcome distri-

bution that is below the rational prediction. The bias is highly significant in treatments where the considered asset is more volatile or the investment horizon is longer. Under these conditions, about ninety percent of participants underestimate the skew, about ninety percent underestimate the spread and about eighty percent estimate the median to lie too close to the distribution’s center. The only instance where the model falsely predicts a deviation from rationality appears in treatments with little volatility: our participants are largely immune to the exponential growth bias in settings if the level of randomness is close to zero. This is surprising given the previous evidence on exponential growth bias, but we add a control treatment that demonstrates that the surprisingly high level of rationality is possible only if the use of a calculator is allowed. Without a calculator, participants exhibit the exponential growth bias, as detailed in Section 4.4.

The rest of this paper is organized as follows. Section 2 briefly discusses related literatures. Section 3 (Study 1) introduces the linearity bias model and the main features of the experimental design by way of reporting on the stylized experiment described above, as well as on a related experiment that elicits the participants’ subjective medians. Section 4 describes Study 2, the main experiments that include systematic variations of investment horizon and asset volatility. Section 5 (Study 3) reports on the extension to a setting with a real-world asset based on the German DAX index. Section 6 concludes.

2 Review of related literature

Classic studies in cognitive psychology discuss quite extensively to what degree the human cognitive apparatus is able to account for the distinction of linear versus nonlinear relations between variables. Wagenaar and Sagaria (1975) ask participants to predict an exponential data series representing an index for pollution. They find that participants strongly underestimate exponential growth. Wagenaar and Sagaria (1978) show that underestimation of exponential growth is robust to the amount of information available to the participants and Wagenaar and Sagaria (1979) show that the effect is robust to the framing of the information. Kemp (1984) surveys perceptions of changes in the cost of living. Respondents systematically underestimate the increase in cost, which is also in line with a misperception of exponential growth. Much of the early data analysis uses responses to quiz-type questions, but a subsequent specialization of this literature more and more focuses on economic contexts, like the perception of compound growth from interest or loan payments. Eisenstein and Hoch (2005), Stango and Zinman (2009), Christandl and Fetchenhauer (2009) and Levy and Tasoff (2014) document that participants underappreciate the effects of compound interest and thereby predictably underestimate the compound effect of growth. Chen and Rao (2007) show that retailers can strategically use this bias by posting double dip price discounts (a discount of 20% followed by another 25% discount is perceived to be a 45% reduction, not the actual 40%). As described in the Introduction, our paper can be viewed as an extension of this literature to non-deterministic growth processes.

An important predecessor of our paper is the study by Benartzi and Thaler (1999) who, among other things, also study biases in the compounding of the long term distributions from a given short term

distribution. Their experimental participants choose different hypothetical retirement plans depending on whether they are given the historical return distribution of retirement plans for a one year period or a 30 year period. Benartzi and Thaler (1999) relate this bias to the effects of myopic loss aversion (see also Samuelson, 1963, Redelmeier and Tversky, 1992, Gneezy and Potters, 1997, Klos, Weber, and Weber, 2005). While we agree that myopic loss aversion likely plays a role in households' long term investment decisions, our experiments suggest that household decisions can also be misguided by a biased perception of the underlying growth processes.¹ This is also consistent with the hypothetical investment choice experiments in Stutzer and Jung Grant (2010) who find an inflated investment rate in treatments where their participants have to calculate the compound return by themselves.²

Another related literature studies whether experimental participants have a correct understanding of financial options. We refer the reader to Gneezy (1996) and Abbink and Rockenbach (2006) for previous results in this—surprisingly small—literature. We note that the assets that we construct have the same structure as the underlying asset in the well-known model of Cox, Ross, and Rubinstein (1979) of European call options. A consistent finding of misperceptions of such assets may therefore indicate a potential mispricing. This is not further studied in our paper, which focuses on investments in the underlying asset itself.

3 Study 1: Median and Mode

3.1 The Linearity Bias

In this subsection we present a simple model of biased decision making that guides our predictions for the three studies. As explained in the Introduction, our model extends the ignorance of compound interest from deterministic settings to stochastic settings. Think of a biased decision maker who perceives relative growth as absolute growth: when asked to predict the accumulated value gain of an investment that yields a per-period interest of r over T periods, she quotes a total gain of rT . That is, she wrongly perceives the absolute changes, not the relative changes, to be constant across the periods. We extend this feature—linearization bias hereafter—to the domain of stochastic growth.

Formally, let Y_0 denote the initial price of an asset with a random price series $\{Y_0, Y_1, \dots\}$ and let η_t be the random variable describing the absolute price growth in t , e.g. $Y_1 = Y_0 + \eta_1$. The linearization biased (LB) decision maker views the distribution of η_t as constant across t , with a distribution identical to that of η_1 . As a result, the LB decision maker misses out on all effects of multiplicative

¹A distinction between our study and the existing experimental work on myopic loss aversion is that the existing papers largely make use of additive growth processes.

²The experiment by Stutzer and Jung Grant (2010) uses a quite similar experimental wording as the experiment described in Section 3 and in our paper's first (2010) version despite having been developed and written independently. A separate and important experimental literature examines the preferences regarding skewness, see Deck and Schlesinger (2010), Br  nner, Levinsky and Qiu (2011), Ebert and Wiesen (2011), and Eckel and Grossman (2014). We restrict this paper to the perception of the distribution, not its valuation.

compounding, which may or may not occur in the true growth process. The absolute change in $t = 1$ has a special role in that it is assumed to be readily available to the agent. This assumption reflects our aim of capturing the neglect of compound interest, as $t = 1$ is the only period where η_t is known without compounding. The assumption is also consistent with the evidence on the exponential growth bias in deterministic settings: it prescribes that the slope of linearly perceived growth is given by the initial slope.³

To investigate the effects of the bias, we consider price series where the relative, not absolute, price growth is constant in t . Letting μ_t be the random variable describing the relative price changes occurring in t , e.g. $Y_1 = Y_0\mu_1$, we consider price series where the distribution of μ_t is i.i.d. across t . This is a natural feature of many stationary growth processes, and is often assumed in finance and economic models. In such contexts, an unbiased decision maker perceives the true distribution of period- T price as $Y_T = Y_0 \prod_{t=1}^T \mu_t$, with $\{\mu_t\}$ i.i.d., whereas the LB decision maker perceives the final price as $\tilde{Y}_T = Y_0 + \sum_{t=1}^T \tilde{\eta}_t$, with $\{\tilde{\eta}_t\}$ i.i.d. and its distribution equal to that of η_1 .

We use the model to derive qualitative (directed) predictions of biased decision making in our contexts, generated by simple numerical applications of the LB model. We also discuss the model's point predictions for completeness; but as a model of such simplicity cannot plausibly capture the precise decision process we focus the statistical analysis on the qualitative hypotheses.

3.2 Study 1 (a): Biased perception of the median

In Studies 1 and 2 we consider what is arguably the most elementary asset covered in the finance literature, among all assets with a fixed investment horizon and a non-deterministic growth path: the binomial-tree asset with fixed maturity, due to Cox, Ross and Rubinstein (1979). For this asset, the multiplicative growth μ_t is a binary random variable with a constant 50-50 chance of moving up or down for T periods, i.e. $\mu_t \in \{\mu^h, \mu^l\}$ where the percental uptick $\mu^h \geq 0$ and the percental downtick $\mu^l \geq 0$ are equiprobable in each $t = 1 \dots T$.

In the two experiments of Study 1, we set the investment horizon at $T = 12$, the initial price of the asset at $Y_0 = 10,000$ and the possible growth rates at $\{\mu^h, \mu^l\} = \{1.7, 0.4\}$. Participants are told that they can buy a given security with an current price price of 10,000 Euros and with the price transition described in the Introduction: the security's price either increases by 70% or decreases by 60% in each month, with equal probabilities. The security, if bought, has to be held for exactly 12 months and is then to be sold.

The compounding leads to a highly skewed price distribution, with a small probability of making a large compound gain and a large probability of compound losses. The median selling price after 12 months is as low as 989 Euros.⁴

³Formally, assume a constant multiplicative growth path such that $Y_T = Y_0(1+r)^T$, where r is interest. The LB decision maker perceives a constant additive growth instead, $\tilde{Y}_T = Y_0 + T\tilde{\eta}$ where $\tilde{\eta} = Y_1 - Y_0 = rY_1$. For $r > 0$ and $T > 1$, we have $\tilde{Y}_T < Y_T$, consistent with the available evidence in the cited studies.

⁴It is also true that all quantiles of Y_t approach 0 in probability as $t \rightarrow \infty$. A proof is available from the others upon request (see e.g. Mörters and v. Weizsäcker, 2009, section 3.5).

Study 1 (a) tests whether participants correctly locate the median. Through a sequence of simple choice problems we identify bounds on the median of each participant's subjectively expected distribution. The instructions differ by treatment condition only in the degree to which they show the implications of compounding over time. If a cognitive bias is at work, participants may misperceive the random price process to a greater extent in one treatment condition than in the other. The results confirm this conjecture: in the treatment where the instructions help the cognitive understanding of the random growth process, the participants' choice are close to optimal. In the treatment where no help is offered, the median is vastly overestimated.

3.2.1 Experimental Design

Choice problems: The experimental responses are bounds on the participants' subjectively perceived median of returns, irrespective of risk preferences. The monetary incentives enable this property by using only two possible payments in each choice problem—"receive a bonus" versus not—making it optimal for any participant with monotonic preferences to maximize the subjectively perceived probability of receiving the bonus.

The choice problems appear in a financial investment context: in each round of the experiments, two risky securities are on offer and the selling price of the chosen security determines whether or not the participant receives the bonus.⁵ Security A is the security described above, with a 70% / -60% price change in each period. A participant who chooses this security receives the bonus if the selling price at maturity exceeds a given threshold t_A . The alternative choice is Security B, which yields the bonus with probability one half. One can immediately see that it is subjectively optimal for a participant to choose Security A if and only if she believes that Security A yields the bonus with probability more than one half. A choice for Security A thus reveals that the median of her subjective probability distribution of Security A's selling price is above t_A .

For a balanced experimental design we describe Security B analogously to Security A, with the difference that Security B has only a single price equiprobable change of +70% or -60% during the 12 months. A participant who chooses Security B receives the bonus if the selling price of B exceeds a separate threshold t_B . This threshold is fixed at the initial price of 10,000 Euros throughout the experiment (hence Security B holds a 50-50 chance of receiving the bonus) whereas the threshold t_A varies between 10 different values. Each experimental participant makes a choice between A and B for each of the 10 possible values of t_A , allowing us to infer bounds on her subjective median of the selling price of Security A. Table 1 lists the 10 choice problems (Task 1, Task 2, etc.) as seen by the participants. Given that the true median of Security A's selling price is 989, the rational prediction is for the participants to chose A in Tasks 1 and 2 and to choose B in all subsequent tasks.

⁵The descriptions begins with the wording: "You are a manager and have to make a decision between two risky investments."

	Thresholds for Security A	Thresholds for Security B	Your decision (A or B)
<i>Task 1</i>	100	10,000	—
<i>Task 2</i>	500	10,000	—
<i>Task 3</i>	2,000	10,000	—
<i>Task 4</i>	6,000	10,000	—
<i>Task 5</i>	9,000	10,000	—
<i>Task 6</i>	12,000	10,000	—
<i>Task 7</i>	20,000	10,000	—
<i>Task 8</i>	35,000	10,000	—
<i>Task 9</i>	90,000	10,000	—
<i>Task 10</i>	250,000	10,000	—

Table 1: The 10 binary choices.

Treatments: Participants are randomly assigned to one of two treatments. Treatment NO_HELP presents only the basic explanation. To introduce Security A, the instructions first describe the price transition rule using simple language. This is followed by a statement about the independence of random draws and by the paraphrase that after month 1, the security’s price is either at 17,000 Euros or at 4,000 Euros. The instructions then repeat the random price transition, but without calculating compound effects explicitly: “At the end of month 2, the price is either 70% higher or 60% lower than at the end of month 1. At the end of month 3, the price is either 70% higher or 60% lower than at the end of month 2. And so on, ...” Security B is described next, with identical wording to that of Security A, where applicable. Finally, the thresholds t_A and t_B are explained and two examples are given.⁶

In treatment HELP we provide the participants with an additional explanation, leaving the remainder of the instructions unchanged. The additional text (about one written page) gives an explicit calculation of the distribution of compound price changes after two periods. It also points out the asymmetry in the selling price distribution and lists the implicit probabilities of receiving the bonus from choosing Security A for each value of t_A . None of the explanations in HELP adds any substantive information relative to the descriptions in NO_HELP. The only difference is that the relevant distributions are explicit in HELP and implicit in NO_HELP. Any difference in responses under the two conditions must stem from differences in the understanding of these implied truths.

Feedback and repetitions: After the participants make their 10 choices, they receive individual feedback in the form of a sample pair of selling prices of Securities A and B. This concludes the first round of the experiment. The experiment is then repeated for four additional rounds of the same nature, each including 10 choices and individual feedback. The feedback procedure and the choice format are identical for both treatments.⁷

⁶Before making their choices, participants also face an understanding test of four questions which they have to answer correctly before they may proceed. The examples and understanding test are carefully chosen to not suggest any responses to the participants.

⁷Each additional round comes with the chance to earn a new bonus (see the next paragraph in the main text), but this does not affect the simple optimality conditions for choice. Independent of other choices it remains optimal to choose A

Procedures and payments: All 128 participants (68 in NO_HELP and 60 in HELP) are students at Technical University Berlin. Six sessions, three in each treatment condition, are conducted in a paper-and-pencil format. The instructions are read aloud to the participants. Participants receive a participation fee of 5 Euros and a possible bonus of 5 Euros per round. That is, participants can earn up to five bonuses of 5 Euros each, one per round of the experiment. After completing all choices, each participant receives five random draws of integers between 1 and 10 to determine which of the 10 choice problems in each round is payoff relevant for her. She receives the bonus for a given round if the selling price of the chosen security in the payoff-relevant problem exceeds its threshold.

3.2.2 Linearity Bias Prediction

With a perceived constant distribution of absolute changes $\tilde{\eta}_t \in \{-6000, 7000\}$, the LB decision maker perceives a symmetric distribution of the period-12 selling price with mode, mean, and median at $\mathbb{E}(\tilde{Y}_T) = Y_0 + \mathbb{E}(\sum_{t=1}^T \tilde{\eta}_t) = 16000$. In treatment NO_HELP, the LB model thus predicts that participants overestimate the true median (989) by an order of magnitude and chooses a switching value in the interval $[12000, 20000]$.⁸

For notation, let $q_{0.5,i}^k$ be the *lower* bound of participant i 's median of the compound price, as elicited in treatment k : i invests in Security A for all values $t_A \leq q_{0.5,i}^k$ and invests in Security B for larger t_A . For the sake of simplicity we restrict attention to cases where participants' choices reveal such a unique switching value, a property that is true in 93% of our data.⁹ By analogy, let $q_{0.5}^k$ be the rational benchmark for $q_{0.5,i}^k$ (dropping the subscript i), i.e. the lower bound of the median that would be elicited from a rational decision maker. Here and elsewhere in the paper, we focus on revealed lower bounds when applicable; we formulate and test the LB decision maker's qualitative predictions for these lower bounds.¹⁰

Hypothesis 1 (Directed Bias in Median Perception). *The median is overestimated in treatment NO_HELP, $q_{0.5,i}^{NO_HELP} > q_{0.5}^{NO_HELP}$.*

In treatment HELP, where the arithmetic and the results of compounding are outlined explicitly, the linearity bias cannot influence the subjective beliefs without contradicting the available explanations. We therefore expect the misperception to disappear, i.e. $q_{0.5,i}^{HELP} = 500$.

iff the subjective median is above t_A , under a wide set of preferences for choice under uncertainty.

⁸Appendix B has proposition that generalizes the qualitative prediction that quantiles are misperceived to more general binomial-tree processes. The proposition also shows that for longer investment horizons, the true median and the perceived median diverge. In the main text, we confine ourselves to stating the qualitative predictions as well as the LB model's point predictions.

⁹If a participant has multiple switching points in one round, her answers in the remaining rounds are still considered in our data analysis. None of our conclusions would change if we dropped all responses by participants who switch strictly more than once in at least one round (12% of participants), or if we included all data and considered each of the 10 tasks separately.

¹⁰Focusing on the intervals' upper bound would generate qualitatively identical hypotheses and statistical results—see Appendix A.4. This alternative route introduces, however, some minor censoring pitfalls.

3.2.3 Results

Range of subjective median for Security A	Share of participants switching from A to B	
	NO_HELP	HELP
[0 – 100)	0.018	0.000
[100 – 500)	0.000	0.000
[500 – 2,000)	0.000	0.703
[2,000 – 6,000)	0.036	0.109
[6,000 – 9,000)	0.107	0.063
[9,000 – 12,000)	0.411	0.063
[12,000 – 20,000)	0.196	0.031
[20,000 – 35,000)	0.179	0.016
[35,000 – 90,000)	0.054	0.000
[90,000 – 250,000)	0.000	0.000
[250,000 – ∞)	0.000	0.016

Table 2: Subjective medians in round 1.

Table 2 lists the implied ranges for the medians of the participants' subjective distributions of Security A's selling price in round 1. Not a single NO_HELP participant reveals a subjective median between 500 and 2,000 Euros (i.e. rational switching at Task 3). Instead, 98% of NO_HELP participants in round 1 reveal that their subjective medians are above 2,000 Euros. Appendix A.1 has the results of rounds 2 to 5, showing that 86% of participants overestimate the median still in round 5. Hypothesis 1 is thus confirmed. The modal choice in round 1 (41% of participants) indicates a subjective median between 9,000 and 12,000 Euros, with the next-higher interval (12,000-20,000) attracting 20% of participants' choices. In treatment HELP, 70% of responses are at the optimal switching point of Task 3 already in round 1. Altogether, the data show strong differences between the two treatments, and parametric t-tests as well as non-parametric Wilcoxon rank-sum tests confirm that all round-by-round treatment comparisons are statistically significant at $p < 0.001$. In particular, the treatment effects are still highly significant in the last round of the experiment and exhibit a consistent pattern that the performance is poor under the NO_HELP condition and much better in HELP.¹¹

3.3 Study 1 (b): The Mode

Even though the deviations from rationality in Study 1 (a) are strong, the expected loss in payoff is rather small. Due to the random payment that results from the multiple-choice-list procedure, the participants leave only 0.46 Euros in expectation on the table, on average in round 1 of treatment NO_HELP. Reporting strongly biased medians can therefore secure near-optimal average payoffs while potentially reducing cognitive effort. To address this possibility, Study 1 (b) tests for the mode

¹¹Based on their decisions the participants can expect to earn 2.25 Euros in treatment NO_HELP and 2.61 in the HELP treatment on average in the first round of the experiment. A fully rational decision maker could expect to earn 2.71 Euros in each of the treatments.

of the distribution and thereby induces steeper incentives. The task description is also shorter and simpler than in Study 1 (a). (This is the simple design described in the paper’s Introduction.)

3.3.1 Experimental Design and Linearity Bias Prediction

The experimental asset of interest, Security A, follows the binomial-tree 70/60 process with the same parameters as in Study 1(a). $N = 69$ participants are presented with the growth process of Security A and are asked to pick one out of five investment opportunities, labeled Investment 1 through Investment 5, whose return depends on the period-12 price of Security A. Similar to Study 1(a), Study 1(b) ensures incentive compatibility under a wide set of preferences by using only two possible payments per choice problem—receive a bonus versus not. The size of the bonus is increased to 20 Euros to make the importance of the single choice between the five investment opportunities maximally salient. Participants are told that Investment 1 “makes a gain” (in effect pays the bonus of 20 Euros—see below) iff the selling price of Security A is between 0 and 6,400 Euros. Investment 2 makes a gain iff the selling price of Security A is between 6,400 and 12,800 Euros, Investment 3 makes a gain iff the selling price is between 12,800 and 19,200 Euros, Investment 4 makes a gain iff the selling price is between 19,200 and 25,600 Euros, and Investment 5 makes a gain iff the selling price is above 25,600 Euros.

Once a participant has chosen his or her investment the computer simulates the period-12 price of Security A. If the chosen investment makes a gain for that selling price, the participant receives the bonus. There are no repetitions in Study 1(b) and thus participants only answer one question.

Under any belief the decision-maker should, evidently, choose the interval that has the largest probability of containing the period-12 selling price, i.e., choose the mode of the histogram with bins $\{0 - 6400; 6400 - 12800, \dots\}$. Due to the price distribution’s large skew, the rational prediction is to choose Investment 1, which contains 80% of the probability mass. An LB decision-maker, as discussed above, perceives a symmetric price distribution \tilde{Y}_T with a mode at 16,000. This implies that the interval containing 16,000 has the highest perceived chance of yielding the bonus. An LB decision maker therefore chooses Investment 3.

3.3.2 Results

The numbers of participants (and percentages in parentheses) choosing Investments 1 through 5 are $\{Inv. 1 : 4 (6\%), Inv. 2 : 19 (28\%), Inv. 3 : 30 (43\%), Inv. 4 : 12 (17\%), Inv. 5 : 4 (6\%)\}$. The distribution is significantly different from uniform choice ($p < 0.001$, chi-square test) and indicates no tendency of participants to choose a mode near zero. While only 6% make the optimal choice of Investment 1, 43% conform with the linearity bias model and choose Investment 3. The participants give up significant amounts of money due to the bias: While the optimal choice would earn 16.12 Euros in expectation, the observed choice distribution earns only 2.07 Euros in expectation per participant.

4 Study 2: Other Quantiles and Other Growth Processes

In this section we examine the robustness of the linearity bias predictions with respect to variations in per-period volatility, investment horizon and the choice of quantiles of the participants' subjective distributions that we elicit.

The results confirm the implication of the LB model that the compound distribution is perceived as too symmetric and too narrow-band if there is substantial randomness in the growth process. We also find that the linearity bias does not appear in two treatments where volatility levels are close to zero. However, a control treatment shows that the bias—which is identical to the exponential-growth bias in these treatments—does appear if participants cannot use a calculator.

Subsection 4.1 describes the experimental designs in detail. Subsection 4.2 contains the prediction of the LB model, which we generate by applying the model introduced in Section 3.1 to the present contexts, and Subsection 4.3 has the results.

4.1 Experimental Design

Treatment conditions: There are five treatments in this study, all of them similar to the NO_HELP treatment of Study 1, i.e. without explicit explanation of the implied probabilities. In all treatments the participants can buy “Security A” at a price of £100. If they buy it they have to sell it after T^k periods, where k indexes the treatment.

The first two treatments specify that the price moves by about 20 percent in each period: In treatments *High Volatility_Short* (HVS) and *High Volatility_Long* (HVL), the parameters specifying upticks and downticks are $\mu^{h,HVS} = \mu^{h,HVL} = 1.212$ and $\mu^{l,HVS} = \mu^{l,HVL} = 0.811$. The sole difference between these two treatments is in the length of time until maturity: $T^{HVS} = 14$ and $T^{HVL} = 140$.

The next two treatments are *Low Volatility_Short* (LVS) and *Low Volatility_Long* (LVL), with a much lower price volatility at $\mu^{h,LVS} = \mu^{h,LVL} = 1.012$ and $\mu^{l,LVS} = \mu^{l,LVL} = 1.011$. The price motion is approximately deterministic in these treatments and price has positive growth with certainty. In the following we will refer to these conditions as *quasi-deterministic*. The number of time periods until maturity is analogous to the first two treatments, at $T^{LVS} = 14$ and $T^{LVL} = 140$.

The fifth treatment, *Low Volatility_Long_No Calculator* (LVLNC) replicates treatment LVL but does not allow participants access to a calculator whereas hand-held calculator is available in the other four treatments. A behavioral difference between treatments LVL and LVLNC would indicate that the degree to which participants are able to do numerical calculations is a driver of choice.

Elicitation of subjective quantiles: As in the treatments of Study 1 (a), a participant of Study 2 who buys Security A receives a fixed bonus if the selling price at maturity exceeds a given threshold t_A . These thresholds differ between treatments. Table 3 lists the respective thresholds of Security A.

The alternative choice option is Security B which yields the bonus with a certain probability.¹² To

¹²Different from Study 1 (a), the instructions simply report to the participants the probability with which Security B yields the bonus, without referring to a separate threshold t_B .

elicit three different quantiles about Security A’s selling price, Security B has three different specifications. Each participant faces each specification once. Security B1 yields the bonus with 10% probability, B2 with 50% and B3 with 90%. Accordingly, each participant faces three choice lists. First, she chooses between Securities A and B1 for the different thresholds of Security A. This allows us to infer bounds on her subjective 10th percentile of Security A’s selling price. For example, suppose that participant i in treatment HVS chooses Security A over Security B1 in Tasks 1 and 2 and chooses Security B1 over Security A in Tasks 3 through 10. Inspecting Table 3 (fourth column) we see that this is subjectively optimal iff participant i ’s subjective 10th percentile for Security A’s selling price is between 30 and 45. In line with our previous notation, we would thus record the elicited lower bound of i ’s subjective quantile for Security A’s selling price as $q_{0.1,i}^{HVS} = 30$. As her second set of tasks the participant faces the analogous choices between Securities A and B2 (with the same list of thresholds for Security A). This allows us to infer a lower bound on her subjective median of the selling price, $q_{0.5,i}^{A,HVS}$. Finally, she faces the analogous list of choices between Security A and B3, allowing us to infer a lower bound on the 90th percentile of the same price, $q_{0.9,i}^{A,HVS}$ ¹³.

	Values of t_A <i>LVS</i>	Values of t_A <i>LVL</i>	Values of t_A <i>HVS</i>	Values of t_A <i>HVL</i>	Values of t_A <i>LVL_NC</i>
<i>Task 1</i>	104.0	185	15	2	185
<i>Task 2</i>	104.5	210	30	5	210
<i>Task 3</i>	105.5	240	45	15	240
<i>Task 4</i>	107.0	290	65	60	290
<i>Task 5</i>	109.0	340	95	140	340
<i>Task 6</i>	111.5	400	125	230	400
<i>Task 7</i>	114.5	460	155	350	460
<i>Task 8</i>	118.0	520	190	550	520
<i>Task 9</i>	122.0	625	225	700	625
<i>Task 10</i>	126.5	850	265	1,000	850

Table 3: The thresholds t_A by treatment condition.

Feedback and repetitions: The computer terminals report feedback to the participants in the form of a sample selling price of Security A. In each treatment, this concludes the first round of the experiment. The experiment is then repeated for four additional rounds. Hence, the experiment comprises five identical rounds for each participant.

Procedures and payments: All 148 participants are undergraduate students at University College London. Eleven sessions are conducted in a computer-based format using the software z-Tree (Fischbacher, 2007). In nine sessions, each of the first four treatment conditions (LVS, LVL, HVS, HVL) is faced by a random subset of participants in each session, without making them aware that other participants face different treatment conditions. Participants in these sessions are supplied with a

¹³After the elicitation of the subjective quantiles we also ask for the participants’ beliefs of Security A making a profit. We do not use the resulting data in the analysis but refer to the paper’s previous version (Ensthaler et al, 2013) and to the instructions for a description of the experimental details and the results.

hand-held calculator that they can use throughout the experiment. In the remaining two sessions all participants face treatment LVL_NC, without the possibility to use calculators.¹⁴

The protocol is fixed across all sessions. First, printed instructions are distributed and participants have to pass an understanding test.¹⁵ Then the computer-based experiment commences and guides the participants through the five rounds in immediate succession.

For a simpler data analysis, the participants' computer interfaces restrict responses to satisfy two constraints. First, responses must exhibit at most one switching point on a choice list between Security A and a single B-type security. That is, a participant cannot switch back and forth between Security A and the respective B-type security. Second, the elicited quantiles must be ordered in a consistent way: participants cannot switch from Security A to B1 at a threshold that exceeds the threshold at which she switches from Security A to B2, which in turn cannot exceed the threshold at which she switches from Security A to B3.¹⁶

Participants receive a participation fee of £5 and a possible bonus of £5 per round. In each round, a single choice is randomly determined to be payoff-relevant, giving an ex-ante incentive to act optimally in each task.¹⁷

4.2 Linearity Bias Prediction

The first hypothesis for Study 2 concerns the perception of skewness. From participant i 's three elicited quantiles for Security A's selling price in treatment k , we derive the empirical skewness measure

$$\gamma_i^k = \frac{1}{3} \sum_{j \in \{0.1, 0.5, 0.9\}} \left(\frac{q_{j,i}^k - \bar{q}_i^k}{sd_i^k} \right)^3,$$

with \bar{q}_i^k and sd_i^k denoting the mean and the standard deviation of participant i 's three responses $q_{0.1,i}^k$, $q_{0.5,i}^k$ and $q_{0.9,i}^k$, respectively.¹⁸

The LB decision maker perceives no skewness in the distribution of the selling price of Security A in any treatment: η_1 is symmetrically distributed, implying that \tilde{Y}_T is symmetrically distributed, too, because symmetry is preserved under addition of random variables. The true distribution of Y_T is,

¹⁴We included treatment LVLNC one year after the other treatments.

¹⁵All participants passed the understanding test, in a few cases after asking for some additional explanations.

¹⁶The instructions explain that violations of these constraints are subjectively suboptimal. Additionally, the experimental software shows an error message if a participant violates either of the two constraints. Only 2% of the participants' inputs receive one or more error messages.

¹⁷Participant in treatments HVS and HVL can earn the bonus either through the quantile elicitation task or through the profit probability elicitation task (see Footnote 13). For each round and each participant, the relevant task type (quantile or profit probability) is determined by a simulated coin flip at the end of the experiment. In treatments LVS, LVL and LVL_NC, the quantile elicitation is always payoff relevant as these treatments skip the profit probability elicitation. For the payment in the quantile elicitation, the computer randomly and equiprobably selects a single choice task between Security A and one of Securities B1, B2, or B3.

¹⁸ γ_i^k is the sample skewness of the triplet $\{q_{0.1,i}^k, q_{0.5,i}^k, q_{0.9,i}^k\}$. It is only one of many possible approximations of the underlying distribution's skewness. The literature on eliciting sets of quantiles of subjective distributions is relatively new and small; we are not aware of a standard procedure (or even any previously applied procedure) to estimate skewness from elicited quantiles. We therefore decided to stick to the sample skewness as a transparent measure.

however, severely skewed under the two treatments with high volatility and the LB model thus predicts an underestimation of skewness.

Hypothesis 2 (Directed Bias in Skewness Perception). *The elicited distribution of outcomes is too symmetric, $\gamma_i^k < \gamma^k$ for $k \in \{HVS, HVL\}$.*

The point predictions for the rational and LB decision maker in these two treatments are $\gamma^{HVS} = 0.57$, $\tilde{\gamma}^{HVS} = 0.22$, $\gamma^{HVL} = 0.71$, $\tilde{\gamma}^{HVL} = 0.20$. (In line with our previous notation, we denote rational predictions without subscript or marker but add tildes for the LB predictions.) For the quasi-deterministic treatments LVS, LVL and LVLNC, both the rational prediction and the LB prediction have all three quantiles in the same response bin of the experiment and thus the predicted skew in these treatments is zero.

For the next hypothesis, we use the observed quantiles to measure participant i 's subjectively perceived 90-10 spread of the outcome distribution in treatment k :

$$\delta_i^k = q_{0.9,i}^k - q_{0.1,i}^k.$$

An implication of linearity bias is that the decision maker fails to realize that the distribution's right tail is long. She therefore perceives the 90-10 spread to be too narrow-band in treatment HVL.

Hypothesis 3 (Directed Bias in Spread Perception). *The elicited 90-10 spread of the perceived outcome distribution is too small in treatment HVL, $\delta_i^{HVL} < \delta^{HVL}$.*

The point predictions of the rational decision maker and the LB decision maker are, respectively: $\delta^{HVL} = 700$, $\tilde{\delta}^{HVL} = 550$. In the other four treatments, both models predict identical 90-10 spreads. Inspecting the point predictions of treatments with high volatility also allows to formulate a directed hypothesis regarding comparative statics: With a longer investment horizon, the linearity bias model predicts that deviations from rationality are getting more severe.

Hypothesis 4 (Bias in Skewness and Spread Perception increases in Time Horizon). *Increasing the investment horizon increases the relative misperception of skew and spread, $\gamma_i^{HVL}/\gamma^{HVL} < \gamma_i^{HVS}/\gamma^{HVS}$ and $\delta_i^{HVL}/\delta^{HVL} < \delta_i^{HVS}/\delta^{HVS}$.*

The point predictions for the LB decision maker are $\tilde{\gamma}^{HVL}/\gamma^{HVL} = 0.28$, $\tilde{\gamma}^{HVS}/\gamma^{HVS} = 0.39$, $\tilde{\delta}^{HVL}/\delta^{HVL} = 0.79$, $\tilde{\delta}^{HVS}/\delta^{HVS} = 1$.

The final hypothesis points out that the LB decision maker's elicited quantiles are predicted to be too

low in cases of quasi-deterministic growth. This prediction coincides with the exponential growth bias, as explained in the Introduction. The LB model thus makes very different predictions in setting with versus without randomness.

Hypothesis 5 (Directed Bias in Quantile Perception for Low Volatility). *All three elicited quantiles of the distribution of outcomes are underestimated if volatility is negligible and the time horizon is long, $q_{j,i}^{LVL} < q_j^{LVL}$, for $j \in \{0.1, 0.5, 0.9\}$.*

The point prediction for all three quantiles under the rational and LB model, respectively, are 460 and 240, in treatment LVL.¹⁹

4.3 Results

We start our data analysis with a descriptive overview of the elicited lower bounds of the participants' subjective quantiles of the price distribution. Then, we run interval regressions that impose a normal decision error assumption but allow us estimating the underlying quantiles while taking into account the data's discrete nature. The final subsection reports parametric t-test and Welch tests as well as nonparametric signed-rank and rank-sum Wilcoxon tests of the above-described behavioral hypotheses.²⁰

4.3.1 Descriptive Statistics

The boxplots in Figure 1 to 5 summarize participants' behavior, separately for each treatment and for each of the five rounds. They depict the distributions of $q_{j,i}^k$ across all participants i , for the three probabilities $j \in \{0.1, 0.5, 0.9\}$ and for treatments $k \in \{HVS, HVL, LVS, LVL, LVLNC\}$. The horizontal dashed lines depict the benchmark rational predictions for the respective treatment-specific quantiles of the price distribution, q_j^k .

¹⁹In treatment LVS, the LB decision maker's elicited lower bounds of the three subjective quantiles coincide with those of the rational decision maker, at 114.5.

²⁰For reasons of simplicity, we refer only to the Wilcoxon test in the following meaning the signed-rank and rank-sum test in the respective one-sample and two-sample test environment.

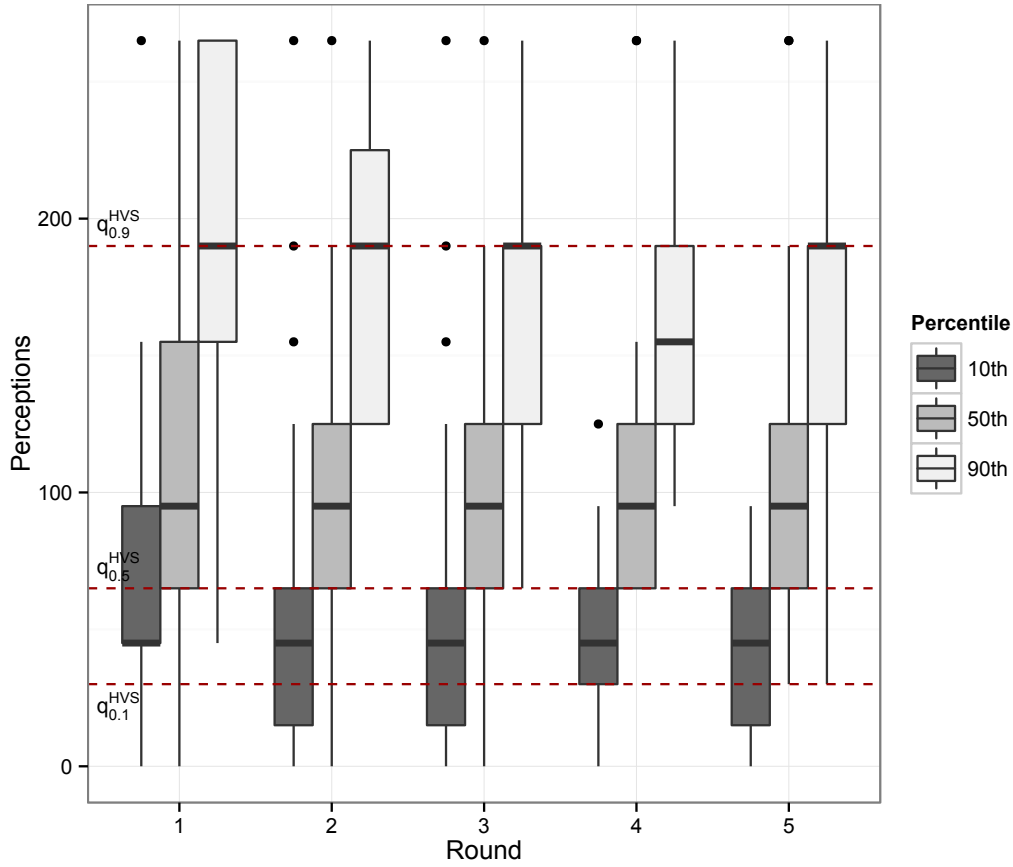


Figure 1: Elicited distributions of the participants' subjective quantiles of Security A's selling price in HVS, separately for the five rounds. Each boxplot triplet consists of (from left to right): subjective 10th percentiles (dark grey boxplot), subjective medians (medium grey boxplot), and subjective 90th percentiles (light grey boxplot). Dashed lines indicate rational benchmarks.

Figure 1 describes the switching points in treatment HVS (high volatility and short time horizon). It shows that in all five rounds the median observation of $q_{0.5,i}^{HVS}$ (the solid line in the middle boxplot within each triplet of boxplots) is strictly above the rational benchmark $q_{0.5}^{HVS}$. Thus in each round at least half of the participants strictly overestimate the median of the stochastic process in treatment HVS. Precisely, 55% of participants reveal that their subjective median is strictly too high in round 1, and 62% in round 5. The 90-10 spread, as perceived by the participants, can be gauged from the location of the two boxplots depicting the subjective 10th (first boxplot within each triplet) and 90th percentiles (third boxplot within each triplet). The perceived spread is too narrow-band compared to the rational prediction. This pattern, too, appears in each of the five rounds. With regards to perceived skewness, the arrangement of boxplots within one round tends to be more symmetric than the rational prediction. Once again, we observe that the corresponding pattern is robust over the rounds.

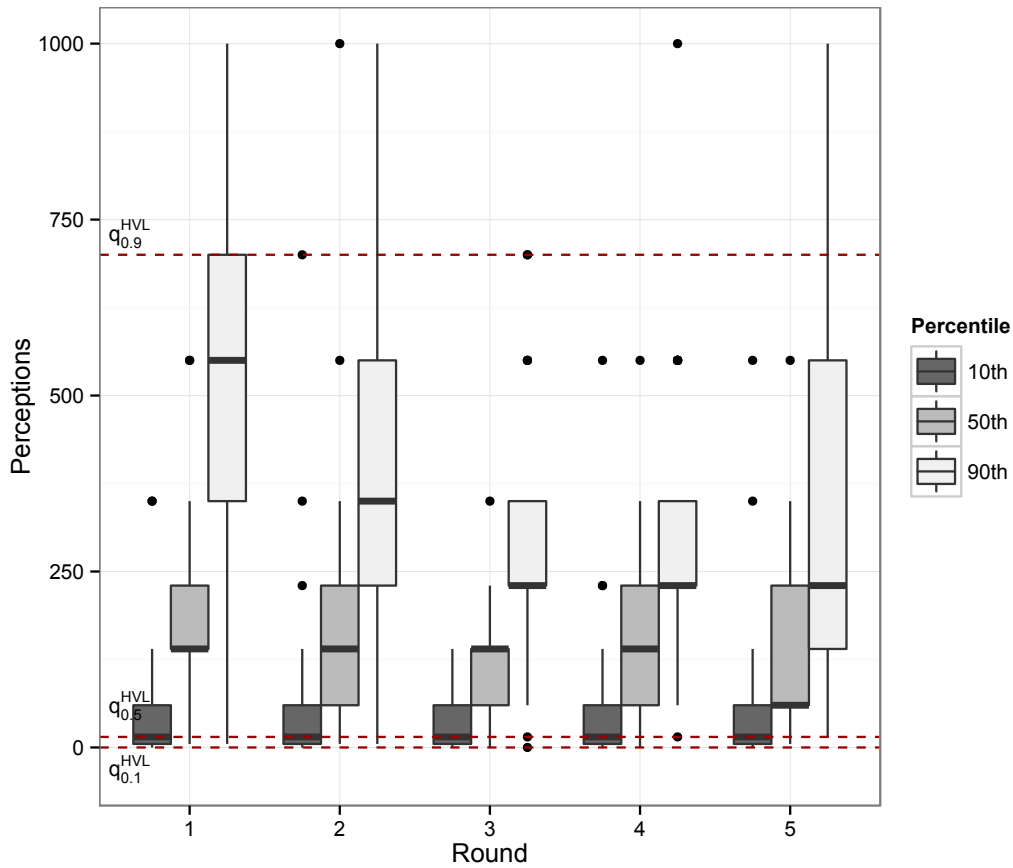


Figure 2: Elicited distributions of the participants' subjective quantiles of Security A's selling price in HVL, separately for the five rounds. Each boxplot triplet consists of (from left to right): subjective 10th percentiles (dark grey boxplot), subjective medians (medium grey boxplot), and subjective 90th percentiles (light grey boxplot). Dashed lines indicate rational benchmarks.

We get qualitatively the same results for treatment HVL (Figure 2), but they are even stronger. Here, the interquartile ranges of the median-perception boxplots are located strictly above the rational level in all rounds. Hence, more than 75% of participants overestimate the median in each round of this treatment. The precise numbers are 93% in round 1 and 79% in round 5. Likewise, the perceived price spread in this condition is underestimated by almost all participants over the rounds and the arrangement of boxplots within each round is much more symmetric than the rational benchmark. If we relax our rational benchmark a bit and count a participant as "overestimating the median" only if she switches more than one step (*task*) later than rational, we still count proportions of participants overestimating the median of 79% in the first round and 45% in the last round of treatment HVL. In HVS, these proportions are 34% and 37% for the first and last round, respectively.

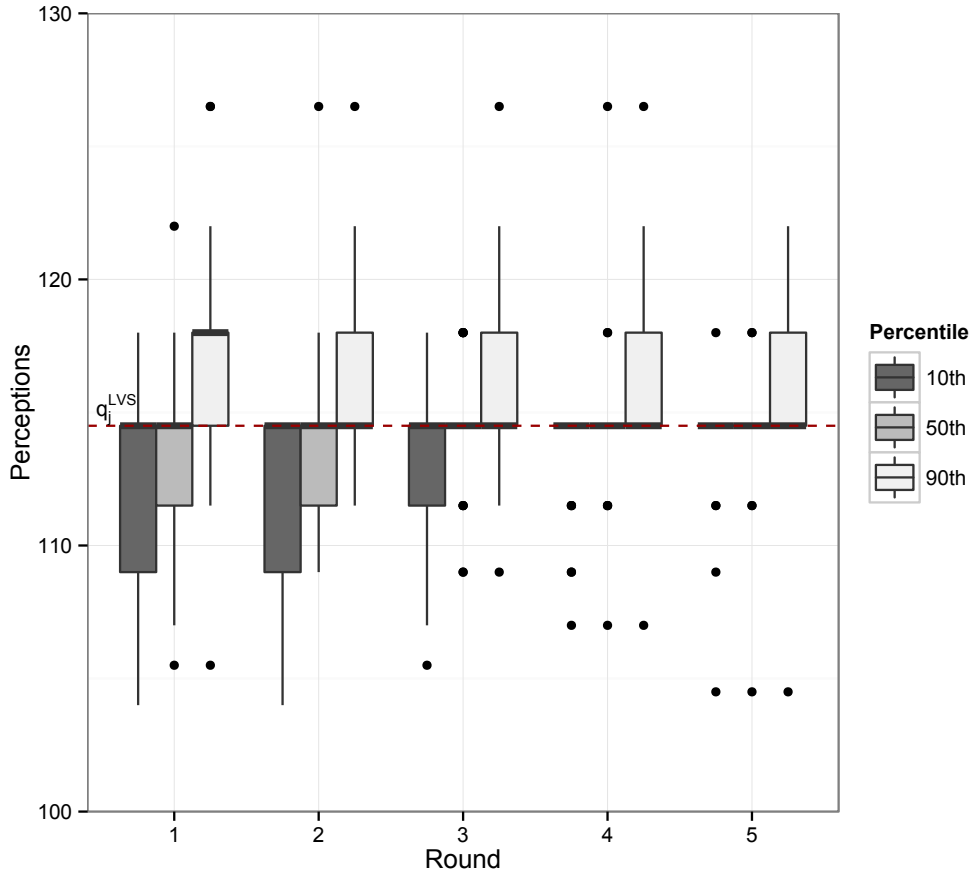


Figure 3: Elicited distributions of the participants' subjective quantiles of Security A's selling price in LVS, separately for the five rounds. Each boxplot triplet consists of (from left to right): subjective 10th percentiles (dark grey boxplot), subjective medians (medium grey boxplot), and subjective 90th percentiles (light grey boxplot). Dashed line indicates rational benchmarks of q_j^{LVS} , for $j \in \{0.1, 0.5, 0.9\}$.

Figure 3 depicts perceptions in treatment LVS, with low per-period volatility and a short investment horizon. The results are very different under these conditions: we observe perceptions of the median that are exactly at, or very close to, the optimal level already in round 1 of the experiment.²¹ Over the rounds, the distributions of price median perceptions quickly collapse towards a single point, which is located at the optimal level. Inspecting the arrangement of boxplots within one round, one can also see that the participants perceive skewness and price spread close to optimally (at zero). There is little dispersion in the subjective quantiles, and most of it vanishes in the course of five rounds.

²¹9 extreme downward outliers over the five rounds and for all elicited percentiles (2% of all treatment specific decisions) in LVS are not pictured in Figure 3 as they would stretch the figure's ordinate too much towards zero and render it uninformative.

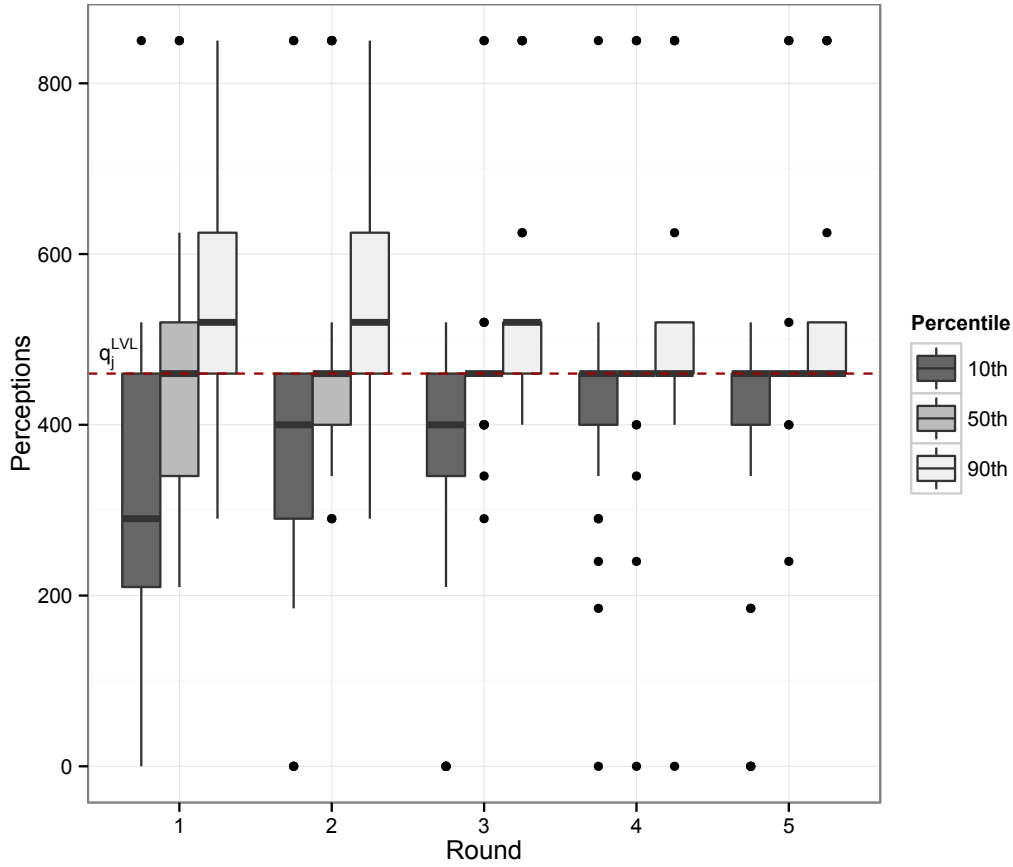


Figure 4: Elicited distributions of the participants' subjective quantiles of Security A's selling price in LVL, separately for the five rounds. Each boxplot triplet consists of (from left to right): subjective 10th percentiles (dark grey boxplot), subjective medians (medium grey boxplot), and subjective 90th percentiles (light grey boxplot). Dashed line indicates rational benchmarks of q_j^{LVL} , for $j \in \{0.1, 0.5, 0.9\}$.

Figure 4 shows a very similar picture for treatment LVL. Again, elicited medians show low levels of misperception and become quite accurate from the third round on. In the first two rounds, spread and skewness perceptions are too large but they approach zero fast.

Again relaxing our rational benchmark and counting a participant as "underestimating the median" not before she switches more than one task earlier than rational yields a similar picture: we count proportions of participants underestimating the median of 31% in the first round and only 3% in the last round of treatment LVL. In LVS, these proportions are 20% and 7% for the first and last round, respectively.

The result that the participants almost perfectly understand the quasi-deterministic growth in LVS and LVL is surprising. There is no significant exponential growth bias in our experiments if the process is quasi-deterministic. A comparison with the existing literature on the exponential growth bias suggests that the difference to the previous evidence may be driven by the fact that we equip our participants with calculators. In order to test if withdrawing calculators would bring the exponential growth bias back to life, we ran the additional treatment LVLNC where calculators are not allowed.

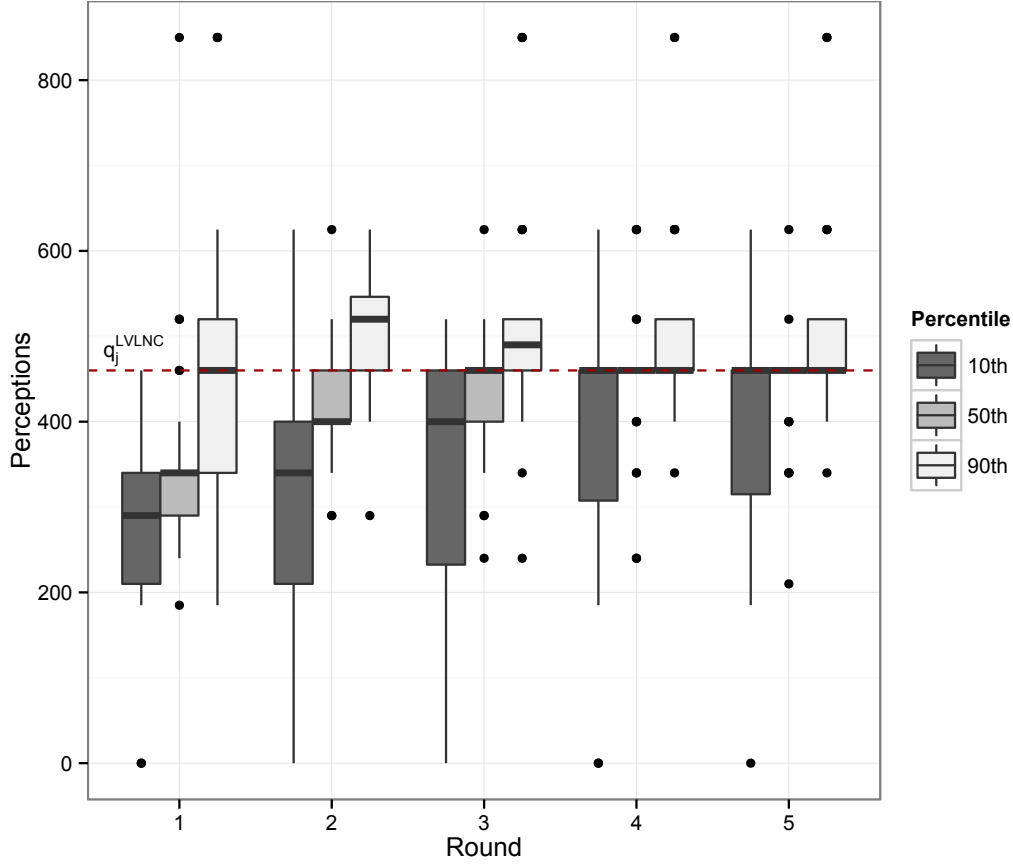


Figure 5: Elicited distributions of the participants' subjective quantiles of Security A's selling price in LVLNC, separately for the five rounds. Each boxplot triplet consists of (from left to right): subjective 10th percentiles (dark grey boxplot), subjective medians (medium grey boxplot), and subjective 90th percentiles (light grey boxplot). Dashed line indicates rational benchmarks of q_j^{LVLNC} , for $j \in \{0.1, 0.5, 0.9\}$.

Figure 5 shows the results of the LVLNC treatment. In the first two rounds a substantial proportion of participants underestimate the median. The interquartile range of the median-perception boxplot of the first round is located strictly below the rational level, indicating that at least 75% of the participants underestimate the median in this round. The exponential growth bias indeed shows up in the absence of calculators. But in quasi-deterministic treatments, the feedback is very informative and learning is fast. Over the last three rounds, distributions of price median perceptions collapse towards a single point at the optimal level. Price spread and skewness perceptions remain almost unchanged compared to LVL. Relaxing the rational benchmark just like for LVL before, we observe proportions of participants underestimating the median of 78% in the first round and 13% in the last round of treatment LVLNC.

In sum, the descriptive statistics indicate that in treatments HVS and HVL the linearity bias model makes qualitatively correct predictions. For the quasi-deterministic treatments LVL and LVS, a surprising accuracy in perceptions appears in the data. However, this accuracy vanishes in LVLNC

where no calculator is available. Here, the direction of deviation once again confirms the linearity bias model.²²

4.3.2 Interval Regressions

The above descriptive analysis is based on the elicited lower bounds of our participants' perceived quantiles of Security A's price distribution. This subsection investigates point estimates instead of lower bounds and employs interval regressions—a modified version of the ordered probit regressions (see e.g. Wooldridge, 2002). The analysis takes into account the interval nature of the data and assumes that the subjectively perceived quantiles are subject to normally distributed disturbances. Under this assumption, the mean of the participants' subjective quantiles can be estimated via maximum likelihood. Figures 6 to 8 report the corresponding estimates of the population means for the subjectively perceived quantiles about Security A's selling price, separately for each of the treatments and for each round.

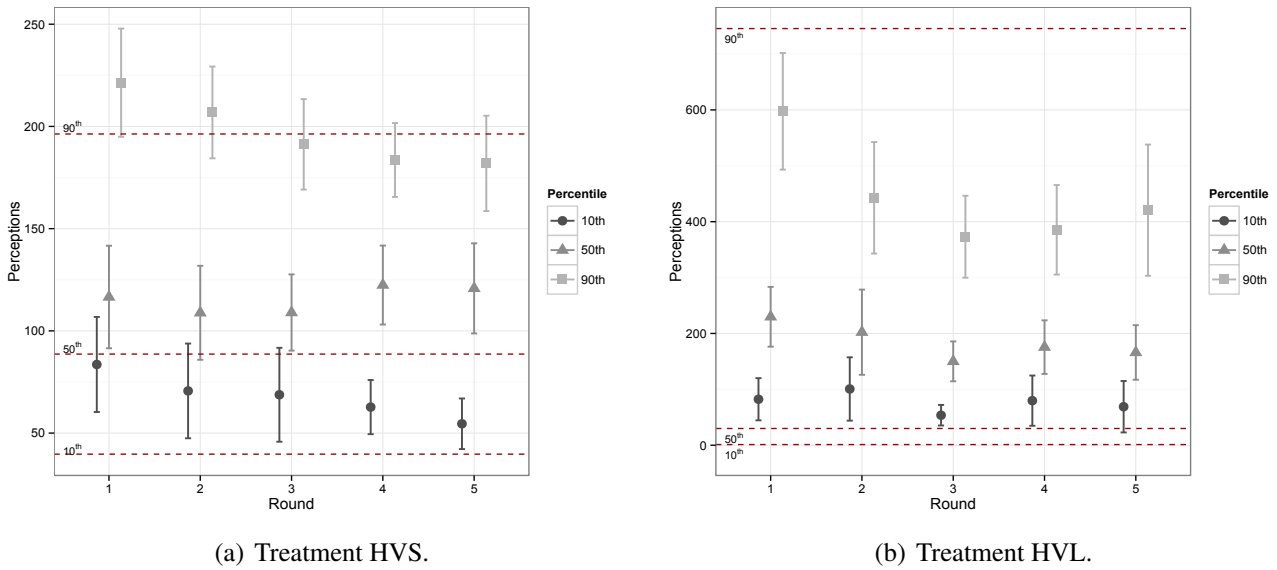


Figure 6: Point estimates enclosed by 95% confidence intervals of the participants' subjective quantiles of Security A's selling price. For each of the five rounds, separate estimates refer to the subjective 10th percentiles (circle), subjective medians (triangle), and subjective 90th percentiles (square). Dashed lines indicate rational benchmarks.

Figures 6, 7 and 8 illustrate the average perceptions of the respective outcome distributions, separately for each treatment and for each of the five rounds under the above functional form assumptions.²³ The

²² Average expected earnings (based on decisions) in the first round of the experiment compare to the expected earnings of a rational decision maker as follows: HVS £3.14 (rational £3.35), HVL £2.96 (£3.14), LVS £4.09 (£4.25), LVL £3.99 (£4.25), LVLNC £3.83 (£4.25).

²³ For a detailed listing of interval regression estimates, see Appendix A.3.

horizontal dashed lines depict the benchmark rational predictions for the respective treatment-specific quantiles. The point estimates of participants' average perceptions are enclosed by the respective 95% confidence intervals.

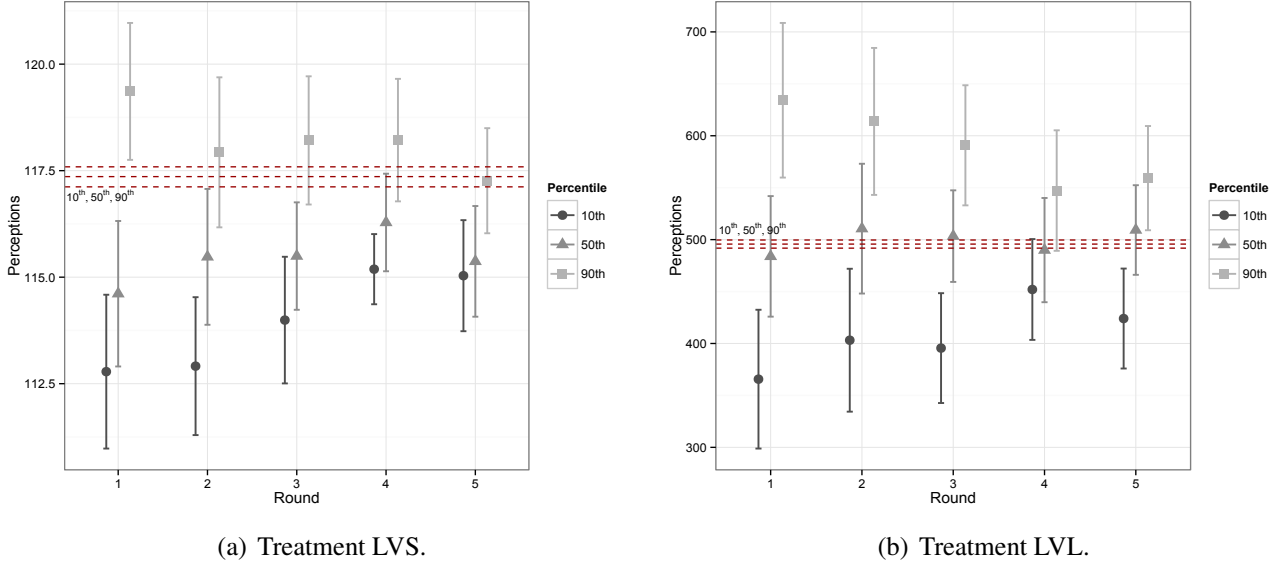


Figure 7: Point estimates enclosed by 95% confidence intervals of the participants' subjective quantiles over the five rounds, with subjective 10th percentiles (circle), subjective medians (triangle), and subjective 90th percentiles (square). Dashed lines indicate rational benchmarks.

Figure 6, 7 and 8 all confirm the previous subsection's descriptive analysis. Based on the point estimates of perceived quantiles, the skewness and spread of the respective outcome distribution are mostly underestimated in treatment HVS (Figure 6(a)) and this effect exacerbates considerably in treatment HVL (Figure 6(b)). Analogously, overestimation of the median appears in both treatments HVS and HVL, but much more severely in HVL. Average perceptions in treatments LVS and LVL are much closer to rationality (Figure 7). In LVS, we observe average perceptions of the median that are significantly below the rational level in some of the rounds (Figure 7(a)). This effect is, however, small in absolute values and it is only marginally significant especially in later rounds. In LVL, average perceptions of the median are quite accurate from round 1 onwards (Figure 7(b)). The estimated 90-10 spread in LVS and LVL diminishes over the rounds. Figure 8 confirms that the exponential growth bias is much stronger when calculators are unavailable, as participants in this treatment strongly and significantly underestimate the median in the first two rounds of treatment LVLNC.

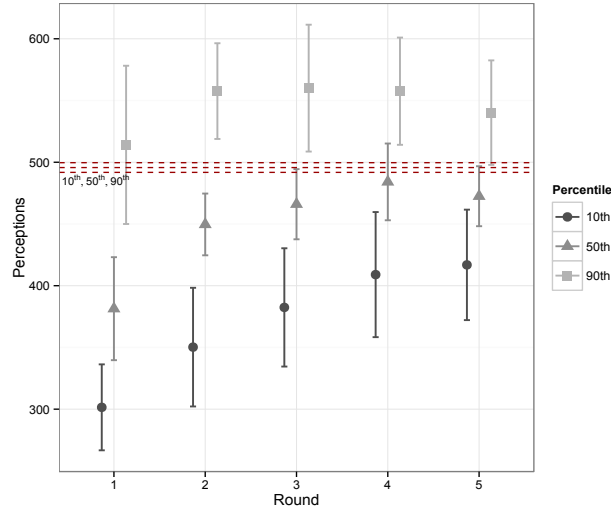


Figure 8: Point estimates in LVLNC enclosed by 95% confidence intervals of the participants' subjective quantiles over the five rounds, with subjective 10th percentiles (circle), subjective medians (triangle), and subjective 90th percentiles (square). Dashed lines indicate rational benchmarks.

4.3.3 Hypothesis Testing

To test the hypotheses of Subsection 4.2, we calculate the described statistical measures of individually perceived skewness (γ_i^k) and spread (δ_i^k), for each participant i and each treatment k . Based on these measures, we analyse deviations from rationality by means of standard parametric and nonparametric tests. Additionally, we conduct tests of the frequencies with which the measures deviate from rationality in the same direction as predicted by the hypotheses.

Regarding **Hypothesis 2 (Directed Bias In Skewness Perception)**, we observe from Figure 1, 2 and 6 that the elicited subjective 50th percentiles tend to lie too close to the midpoint between the 10th percentile and the 90th percentile, in the relevant treatments HVS and HVL. Calculating the individual participants' skewness measures γ_i^k leads to analogous conclusions. In the first round, the average values for γ_i^k are 0.40 in HVS and 0.34 in HVL, compared to the rational benchmarks of 0.57 and 0.71, respectively. Thus, participants underestimate the skew on average by 30% in HVS and by 51% in HVL. The statistical significance of these downward deviations in HVS and HVL is confirmed by parametric t-tests ($p < 0.01$, one-sided) and non parametric Wilcoxon tests ($p < 0.05$, one-sided). Table 4 shows that this misperception becomes even stronger for the following four rounds, with average values for γ_i^k as low as 0.109 in HVS and 0.239 in HVL. Hypothesis 2 is thus confirmed.

The proportions of participants who underestimate skewness are high. While in round 1 of treatment HVS, a moderate 55% of the participants underestimate the skew, 82% do so in round 5. The corresponding proportions are even larger in HVL with 93% in round 1 as well as in round 5. In HVS, 4 out of 5 rounds exhibit proportions of participants underestimating the skew which are significantly greater than 50% ($p < 0.01$, binomial tests). For HVL, in all rounds these proportions are larger than

50% ($p < 0.001$, binomial tests).²⁴

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
γ_i^{HVS}	0.398***	(0.06)	0.341***	(0.06)	0.244***	(0.07)	0.109***	(0.08)	0.129***	(0.08)
γ_i^{HVL}	0.344***	(0.06)	0.347***	(0.07)	0.382***	(0.06)	0.239***	(0.08)	0.320***	(0.07)

Significance levels: * : <10% ** : <5% *** : <1%

Table 4: Average values for γ_i^k over the five rounds, with $k \in \{HVS, HVL\}$. The indicated significance values refer to one-sided parametric t-tests for downward deviations from the rational values.

Hypothesis 3 (Directed Bias In Spread Perception) relates only to treatment HVL. From Figure 2 and 6(b) we observe that the 90-10 spread tends to be too small, as predicted in Hypothesis 3. At the individual level, we observe in round 1 an average value for δ_i^{HVL} of 449.9, compared to the rational prediction of 700, a large and significant underestimation ($p < 0.001$, one-sided t and Wilcoxon test). Table 5 shows that the average values of the perceived spread shrink even more for the following four rounds, with a minimum value of 259.8 in round 4. In sum, the evidence confirm Hypothesis 3. In terms of proportions, 89% of the participants in treatment HVL underestimate the 90-10 spread in round 1 and even 93% do so in round 5. For all rounds, the proportions of participants underestimating the 90-10 spread are significantly larger than 50% ($p < 0.01$, binomial tests).

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
δ_i^{HVL}	449.93***	(48.6)	289.44***	(40.2)	270.44***	(32.6)	259.75***	(39.5)	304.48***	(52.6)

Significance levels: * : <10% ** : <5% *** : <1%

Table 5: Average values for δ_i^{HVL} over the five rounds. The indicated significance values refer to one-sided parametric t-tests for downward deviations from the rational value.

Next, we document the extent to which statistical tests support **Hypothesis 4 (Bias in Skewness and Spread Perception increases in Time Horizon)**, which compares short and long investment horizons in treatments with large per-period volatility. In order to test the hypothesis, we calculate the relative measures $\gamma_i^{HVS}/\gamma_i^{HVS}$, $\gamma_i^{HVL}/\gamma_i^{HVL}$, $\delta_i^{HVS}/\delta_i^{HVS}$ and $\delta_i^{HVL}/\delta_i^{HVL}$ for each participant, summarized in Table 6.

In round 1, skewness is indeed underestimated to a greater extent in HVL than in HVS, with average values for $\gamma_i^{HVL}/\gamma_i^{HVL}$ and $\gamma_i^{HVS}/\gamma_i^{HVS}$ of 0.49 and 0.70, respectively. This difference, $\Delta_\gamma =$

²⁴For a detailed listing of binomial test results, see Appendix A.2.

$\gamma_i^{HVL}/\gamma^{HVL} - \gamma_i^{HVS}/\gamma^{HVS}$, is significantly smaller than zero ($p < 0.1$, one-sided Welch and Wilcoxon test). However, Table 6 shows that both the negative sign and the significance of Δ_γ vanish quickly in later rounds, giving overall only very partial support for the first part of Hypothesis 4.

With regard to the 90-10 spread, Table 6 exhibits that in all 5 rounds the degree of relative underestimation is larger in HVL than in HVS. That is, $\Delta_\delta = \delta_i^{HVL}/\delta^{HVL} - \delta_i^{HVS}/\delta^{HVS}$ is on average significantly smaller than zero for each round of the experiment ($p < 0.1$, one-sided Welch and Wilcoxon test) and thereby supports the second part of Hypothesis 4.

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
$\frac{\gamma_i^{HVS}}{\gamma^{HVS}}$	0.696	(0.11)	0.597	(0.09)	0.427	(0.12)	0.190	(0.13)	0.227	(0.15)
$\frac{\gamma_i^{HVL}}{\gamma^{HVL}}$	0.487	(0.08)	0.491	(0.09)	0.541	(0.08)	0.339	(0.12)	0.453	(0.10)
Δ_γ	-0.209*	(0.14)	-0.105	(0.13)	0.113	(0.14)	0.148	(0.17)	0.226	(0.18)
$\frac{\delta_i^{HVS}}{\delta^{HVS}}$	0.789	(0.08)	0.790	(0.08)	0.715	(0.08)	0.713	(0.07)	0.752	(0.09)
$\frac{\delta_i^{HVL}}{\delta^{HVL}}$	0.642	(0.07)	0.413	(0.06)	0.386	(0.05)	0.371	(0.06)	0.434	(0.08)
Δ_δ	-0.147*	(0.11)	-0.377***	(0.09)	-0.329***	(0.09)	-0.342***	(0.09)	-0.317***	(0.11)

Significance levels: * : <10% ** : <5% *** : <1%

Table 6: Average values of relative deviation measures with the significance levels referring to the one-sided parametric Welch test of Δ_δ and Δ_γ being significantly smaller than zero.

Finally, we examine **Hypothesis 5 (Directed Bias in Median Perception for Low Volatility)**, predicting the exponential growth bias in treatment LVL: relative to the rational benchmarks, all three subjective quantiles are predicted to be too small. Like the interval regressions, the hypothesis tests do not support Hypothesis 5. Table 7 shows the averages of all lower bounds for the elicited quantiles of the subjective price distributions. Inspecting the second row, we see that in round 1 the average perceived lower bound for the median is 448.4. The hypothesized downward deviation from the rational value of 460 can neither be supported by a t-test ($p=0.34$, one-sided) nor by a Wilcoxon test ($p=0.17$, one-sided). Regarding the averages of elicited 10th and 90th percentiles in round 1, the value of 324.6 for the 10th percentile confirms a significant underestimation of the rational value ($p < 0.01$, one-sided t and Wilcoxon test) while 577.8 for the 90th percentile does not ($p=0.99$, one-sided t and Wilcoxon test). Table 7 also shows that the deviations from rationality decrease over the course of the subsequent rounds. Altogether, Hypothesis 5 is not supported.

The proportion of participants who strictly underestimate the median is 41% in round 1, whereas 28% report the rational lower bound and 31% report a too-high lower bound. But learning is fast and by

the last round, 79% of the participants report the rational lower bound for the 50th percentile, which is significantly more than 50% ($p < 0.01$, binomial tests).

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
$q_{0.1,i}^{LVL}$	324.6***	(37.0)	365.3***	(37.0)	349.3***	(32.3)	418.4*	(25.7)	385.1**	(28.5)
$q_{0.5,i}^{LVL}$	448.4	(27.5)	477.2	(29.9)	470.7	(21.3)	455.1	(26.8)	477.2	(20.9)
$q_{0.9,i}^{LVL}$	577.8	(31.7)	561.0	(30.8)	541.7	(25.3)	502.6	(28.7)	518.4	(22.3)

Significance levels: * : <10% ** : <5% *** : <1%

Table 7: Average values for $q_{j,i}^{LVL}$ over the five rounds, with $j \in \{0.1, 0.5, 0.9\}$. The listed significance values refer one-sided parametric t-tests for downward deviations from the rational value ($q_j^{LVL} = 460$, for $j \in \{0.1, 0.5, 0.9\}$).

In treatment LVLNC, where we do not allow calculators, the exponential growth bias shows up significantly. There, participants significantly underestimate the median of the relevant price at maturity, as shown in Table 8. In round 1, the average lower bound of the perceived median is 353.3, which is a significant downward deviation from the rational value of 460 ($p < 0.001$, one-sided t and Wilcoxon test). Table 8 also shows that this misperception almost completely vanishes by round 4.

The bias is also strong in terms of frequencies: 84% of the participants underestimated the median in round 1 of treatment LVLNC which is a proportion significantly larger than 50% ($p < 0.001$, binomial tests). In the final round, only 22% still do so.

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
$q_{0.1,i}^{LVLNC}$	272.0***	(19.5)	314.5***	(26.6)	345.6***	(27.1)	374.7***	(27.1)	386.3***	(23.3)
$q_{0.5,i}^{LVLNC}$	353.3***	(20.7)	418.9***	(12.3)	434.5**	(13.7)	452.3	(15.0)	442.3*	(11.7)
$q_{0.9,i}^{LVLNC}$	474.2	(29.4)	507.2	(15.6)	516.4	(23.4)	514.4	(19.2)	501.7	(19.1)

Significance levels: * : <10% ** : <5% *** : <1%

Table 8: Average values for $q_{j,i}^{LVLNC}$ over the five rounds, with $j \in \{0.1, 0.5, 0.9\}$. The listed significance values refer one-sided parametric t-tests for downward deviations from the rational value ($q_j^{LVLNC} = 460$, for $j \in \{0.1, 0.5, 0.9\}$).

5 Study 3: Linearity Bias in the Perception of Exchange Traded Funds

In this study we test the robustness of the LB model in a setting where the asset price depends on real-world historical data. We simulate leveraged and unleveraged exchange-traded funds (ETFs) on past data of the German stock market index DAX30 to examine how changes in volatility affect the participants' perceptions of real-life growth processes.

Leveraged ETFs move by a given multiple relative to an underlying asset, compounded at the end of each trading day. A triple leveraged ETF on the DAX30 index increases by three per cent on a trading day if the DAX30 increases by one per cent on that day and it falls by three per cent if the DAX30 falls by one per cent. Leveraged ETFs are a popular asset class amongst household investors but have come under severe scrutiny as many investors were perplexed when the products made a loss in a period where the underlying index made a gain.²⁵ Our experiment confirms the hypothesis derived from the LB model that skewness and spread are strongly underestimated if the volatility is high.

Section 5.1 describes the experimental design. Section 5.2 presents the behavioral hypothesis derived from the model and section 5.3 discusses the results.

5.1 Experimental Design

Treatment conditions: There are two treatments in this study. In both treatments, Security A is an ETF based on the DAX30. The two treatments differ only in that their respective versions of Security A differ in per-period volatility. In treatment ETF_3, the relevant security is a triple-leveraged ETF based on the DAX30. Its price changes, on each trading day, by three times the daily percentage changes of the underlying index DAX30. In treatment ETF_1, in contrast, Security A is simply the DAX30 ETF itself.

The time horizon until maturity of the ETF is 2000 trading days both for ETF_1 and ETF_3. To generate realized price paths for the two assets, we sample 2000 consecutive DAX30 closing values, drawn at random from the time period 1964 to 2012.²⁶ As in Study 2, participants can buy the ETF at a price of £100 and have to hold it until maturity.

Elicitation of subjective quantiles: Also as in Study 2, in both treatments, a participant who chooses Security A receives a fixed bonus if the selling price at maturity exceeds a given threshold t_A . These

²⁵Regulatory units and the financial media issued extensive warnings that involve explanations of these counter-intuitive possibilities. The U.S. securities regulator FINRA issued a note in 2009 (FINRA Regulatory Note 09-31) saying that "...while such products may be useful in some sophisticated trading strategies, they are highly complex financial instruments that are typically designed to achieve their stated objectives on a daily basis. Due to the effects of compounding, their performance over longer periods of time can differ significantly from their stated daily objective..."

²⁶The instruction in the ETF treatments are analogous to the other treatments in study 2. The participants receive general information about the DAX30 and a data summary of daily DAX30 movements in the relevant time period. The information is given in the form of a histogram as well as statements specifying the 90% confidence interval $([-1.8\%, 1.8\%])$ and the overall average of daily percentage changes (0.03%). Note that the participants are UK-based students who typically have little knowledge about German stock markets.

thresholds do not differ between ETF_3 and ETF_1. They are listed in Table 5.

To elicit three different quantiles, the alternative choice option Security B has three different specifications which are equal to those in Study 2, i.e., Security B1 yields the bonus with 10% probability, B2 with 50% and B3 with 90%. Like in Study 2, each participant faces three choice lists and the choices allow us to infer the subjectively perceived quantiles.

Thresholds for Security A in ETF_3 and ETF_1	
<i>Task 1</i>	30
<i>Task 2</i>	60
<i>Task 3</i>	90
<i>Task 4</i>	140
<i>Task 5</i>	200
<i>Task 6</i>	260
<i>Task 7</i>	330
<i>Task 8</i>	450
<i>Task 9</i>	650
<i>Task 10</i>	1000
<i>Task 11</i>	1600

Table 9: The 11 thresholds.

Procedures: The computer terminals report feedback to the participants in the form of a sample selling price of Security A. That is, the computer randomly samples a sequence of 2000 consecutive trading days from the set of all available 2000-day histories of the DAX30 and uses it to simulate the asset price at maturity. The participants learn the result of the simulation and their payoff.

All other aspects of the protocol are identical to Study 2. In both treatments of Study 3, the basic procedure is repeated four times, making for five identical rounds for each participant. 59 participants are in one of the treatments of Study 3, all of them undergraduate students at University College London. The incentivisation structure contains a participation fee and a possible bonus of £ 5.00 for each round.

5.2 Linearity Bias Prediction

While Study 2 analysed the effects of an increased time horizon, the treatment comparison in Study 3 focuses on the effects of increasing the per-period volatility. Again, we use the LB decision maker, who perceives the distribution of absolute returns as constant over time, in order to derive the hypothesis for our statistical tests.

We generate the perceived distribution of an LB decision maker by means of simulation: We randomly sample 500 price paths of ETF_3 and ETF_1 as perceived by an LB decision maker to generate the perceived distribution. By analogy to the previous discussion, the LB decision maker is assumed to correctly perceive the distribution of daily changes in the relevant asset price, but views them as an absolute changes and perceives their distribution to be constant over time. She therefore neglects all

compounding. In detail, we simulate the perceived selling price of an ETF by first randomly selecting a start date $t = 0$ at which the price is fixed (by design of the experiment) at $Y_0 = 100$. For each of the ensuing 2000 trading days $s > t$ we consider the relative change in value on that day, μ_s , and add $\tilde{\eta}_s = (\mu_s - 1)Y_0$ to the current perceived price of the asset: $\tilde{Y}_s = \tilde{Y}_{s-1} + \tilde{\eta}_s$ (with correct perception of the starting value, $\tilde{Y}_0 = Y_0$). For example, suppose that $t = \text{Jan 2, 1978}$ was randomly chosen as the starting date and that the DAX30 increased in value by 1.4% on $s = \text{Feb 15, 1979}$. As the perceived absolute increase on the latter date, the simulation simply adds $\tilde{\eta}_s = \pounds 1.4$ to the price of the asset, as perceived by the LB decision maker. This way the simulation arrives at a perceived selling price at maturity. Repeating this procedure 500 times for random starting dates generates the perceived distribution of selling prices.

The simulation of perceived growth leads us to the following main hypothesis of this section, where γ and δ are constructed as in Section 4. The hypothesis describes that the increased per-period volatility of leveraged ETFs leads to a larger bias. This is intuitive as the neglect of compounding is worse in situations where the proportional growth rates are further away from 1 and their multiplication is thus more different from a perceived linear growth.

Hypothesis 6 (Bias in Skewness Perception increases in Volatility). *Increasing the per-period volatility increases participant i 's misperception of the skewness and the spread of the outcome distribution, $\gamma_i^{ETF-3}/\gamma^{ETF-3} < \gamma_i^{ETF-1}/\gamma^{ETF-1}$, $\delta_i^{ETF-3}/\delta^{ETF-3} < \delta_i^{ETF-1}/\delta^{ETF-1}$.*

The point predictions for the LB decision maker are $\tilde{\gamma}^{ETF-1}/\gamma^{ETF-1} = 0.24$, $\tilde{\gamma}^{ETF-3}/\gamma^{ETF-3} = 0.10$, $\tilde{\delta}^{ETF-1}/\delta^{ETF-1} = 0.65$, $\tilde{\delta}^{ETF-3}/\delta^{ETF-3} = 0.37$.

5.3 Results

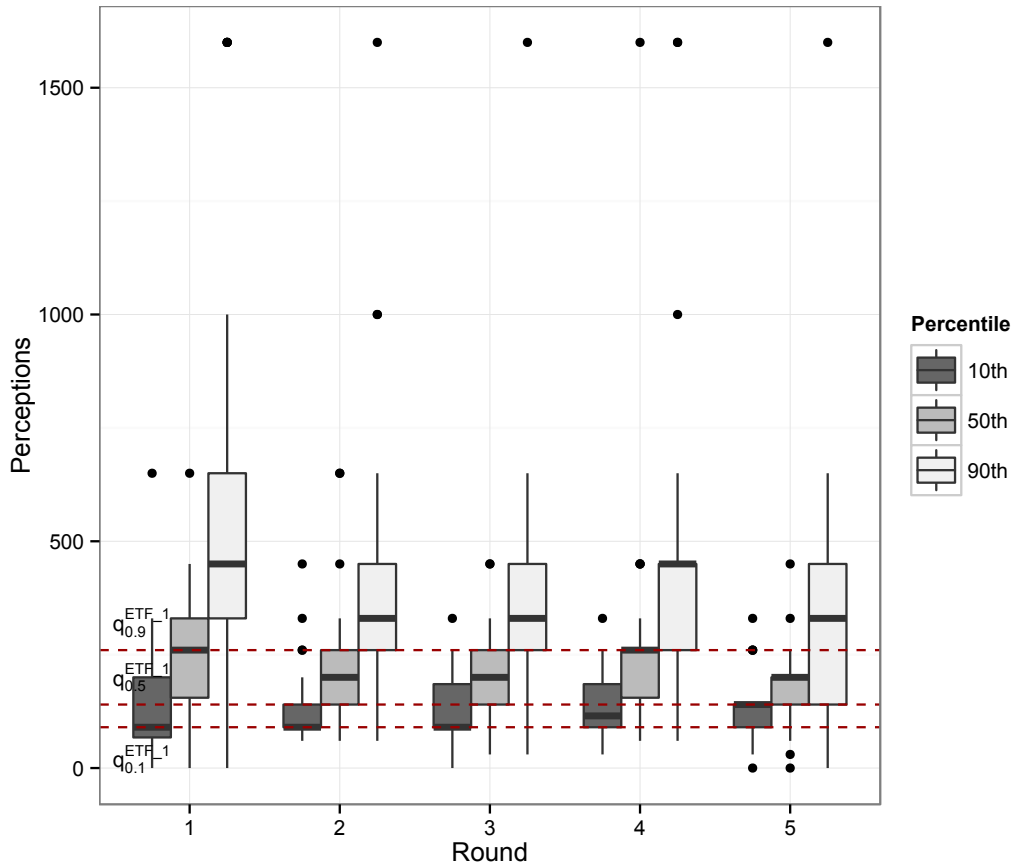


Figure 9: Elicited distributions of the participants' subjective quantiles of Security A's selling price in ETF_1, separately for the five rounds. Each boxplot triplet consists of (from left to right): subjective 10th percentiles (dark grey boxplot), subjective medians (medium grey boxplot), and subjective 90th percentiles (light grey boxplot). Dashed lines indicate rational benchmarks.

Figure 9 illustrates the distributions of participants' perceptions for condition ETF_1. The figure shows that all medians of the median-perception distributions (the solid line in the middle boxplot within each triplet of boxplots) are above the rational level. This indicates a substantial degree of overestimation of the typical price growth by the participants. Participants in our experiment are overly optimistic about the simple index ETF. The arrangement of boxplots also shows that the perceived distributions are quite symmetric. But we note that with a simple ETF, the true distribution is relatively symmetric as well.

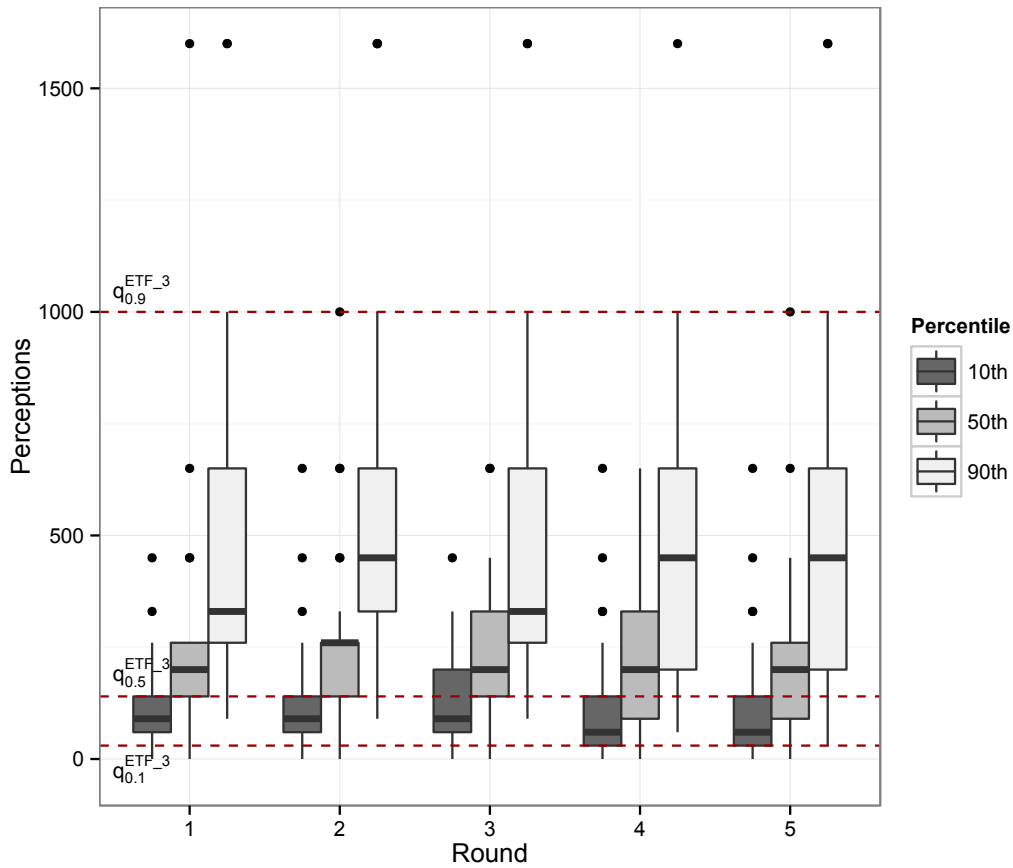
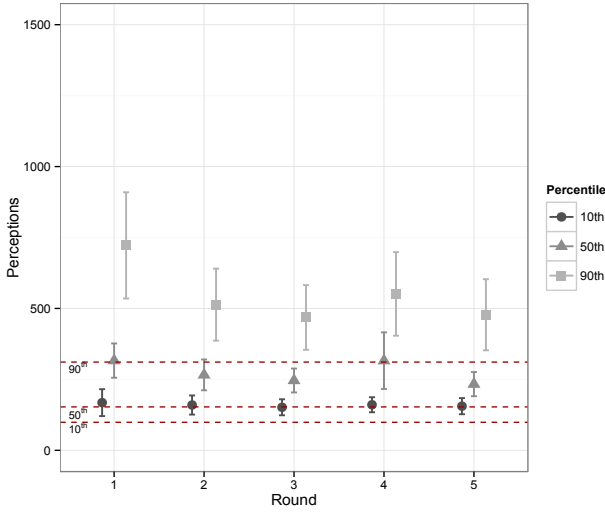


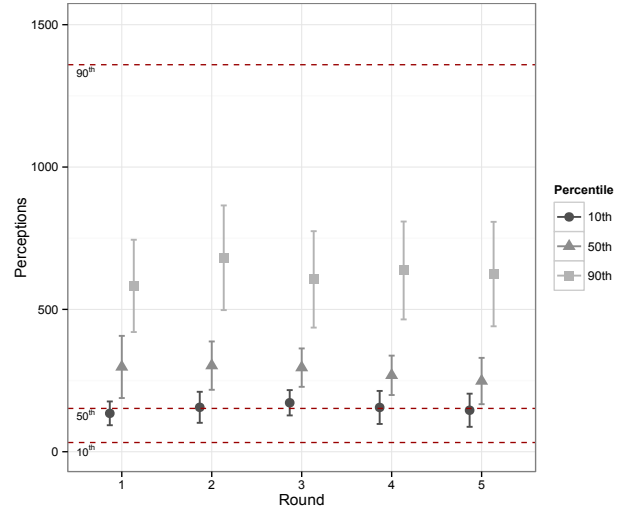
Figure 10: Elicited distributions of the participants' subjective quantiles of Security A's selling price in ETF_3, separately for the five rounds. Each boxplot triplet consists of (from left to right): subjective 10th percentiles (dark grey boxplot), subjective medians (medium grey boxplot), and subjective 90th percentiles (light grey boxplot). Dashed lines indicate rational benchmarks.

Figure 10 captures the participants' perceptions in treatment ETF_3. Again, perceived medians show a notable level of overestimation. That is, the median of the perceived median distributions lies strictly above the optimal level. The arrangement of boxplots within a round also shows that the perceived spread and the perceived skewness are too small (in absolute terms). The participants do show a tendency to report skewed distributions but they far underappreciate the actual level of skewness.²⁷

²⁷Fully rational subjects can expect to earn £3.30 in ETF_1 and £3.20 in ETF_3. On average (based on their decisions) the participants can expect to earn £3.07 in ETF_1 and £3.01 in treatment ETF_3 in round 1 of the experiment.



(a) Treatment ETF_1.



(b) Treatment ETF_3.

Figure 11: Point estimates enclosed by 95% confidence intervals of the participants' subjective quantiles of Security A's selling price. For each of the five rounds, separate estimates refer to the subjective 10th percentiles (circle), subjective medians (triangle), and subjective 90th percentiles (square). Dashed lines indicate rational benchmarks.

We now consider interval regressions. Figure 11 exhibits that these regression, under the assumptions of normally distributed decision errors, confirm the main conclusions drawn from Figure 9 and 10. That is, in both treatments, ETF_1 and ETF_3, the outcome distributions' median is significantly overestimated within each of the five rounds. In treatment ETF_3, which has a much more skewed and spread out outcome distribution than ETF_1, participants far underappreciate the skewness and spread over all rounds.

Finally, we test **Hypothesis 6 (Bias in Skewness and Spread Perception increases in Volatility)** by calculating for each participant the measures of relative deviation for which the hypothesis makes a directed statement. In round 1, skewness is underestimated to a greater extent in ETF_3 than in ETF_1, with average values for $\gamma_i^{ETF_3}/\gamma^{ETF_3}$ and $\gamma_i^{ETF_1}/\gamma^{ETF_1}$ of 0.38 and 0.59, respectively. The significance of this statement is supported by a nonparametric Wilcoxon test ($p < 0.05$, one-sided) but not by a parametric Welch test for $\Delta_\gamma = \gamma_i^{ETF_3}/\gamma^{ETF_3} - \gamma_i^{ETF_1}/\gamma^{ETF_1}$ ($p = 0.14$, one-sided). Table 10 points out that the negative difference of the first round quickly vanishes in the course of the experiment. Just like for the comparative statics of γ in Hypothesis 4, we thus find support for the first part of Hypothesis 6 only weakly and in early rounds.

With regard to the 90-10 spread, Table 10 exhibits that in all 5 rounds the degree of relative underestimation is larger in ETF_3 than in ETF_1. That is, $\Delta_\delta = \delta_i^{ETF_3}/\delta^{ETF_3} - \delta_i^{ETF_1}/\delta^{ETF_1}$ is on average significantly smaller than zero for each round of the experiment ($p < 0.001$, one-sided Welch and Wilcoxon tests) and the measures thereby support the second part of Hypothesis 6.

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
$\frac{\gamma_i^{ETF_1}}{\gamma^{ETF_1}}$	0.587	(0.16)	0.370	(0.20)	0.457	(0.18)	0.367	(0.16)	0.535	(0.15)
$\frac{\gamma_i^{ETF_3}}{\gamma^{ETF_3}}$	0.383	(0.11)	0.436	(0.09)	0.244	(0.14)	0.490	(0.10)	0.553	(0.10)
Δ_γ	-0.203	(0.19)	0.066	(0.22)	-0.213	(0.23)	0.122	(0.19)	0.017	(0.18)
$\frac{\delta_i^{ETF_1}}{\delta^{ETF_1}}$	2.758	(0.45)	1.768	(0.29)	1.580	(0.28)	1.964	(0.41)	1.590	(0.29)
$\frac{\delta_i^{ETF_3}}{\delta^{ETF_3}}$	0.392	(0.07)	0.455	(0.08)	0.374	(0.07)	0.407	(0.07)	0.410	(0.07)
Δ_δ	-2.366***	(0.46)	-1.312***	(0.29)	-1.205***	(0.29)	-1.557***	(0.41)	-1.179***	(0.30)

Significance levels: * : <10% ** : <5% *** : <1%

Table 10: Average values of relative deviation measures with the significance levels referring to the one-sided parametric Welch test of Δ_δ and Δ_γ being significantly smaller than zero.

6 Conclusion

This paper investigates and finds support for the hypothesis that people underestimate the level of asymmetry in growth processes where nothing in the description of the environment explicitly raises the prevalence of asymmetry. A wide set of stochastic growth processes have this feature—skewness is "hidden". The paper thereby adds to the list of known facts on cognitive biases in financial decision problems. In realistic (stochastic) settings, a failure to correctly compound can have more consequences than just an underestimation of growth. This may be especially relevant in the context of preferences that consider quantiles of the outcome distribution, like value-at-risk or expected utility with highly convex utility functions. The experimental tests in this paper are, however, independent of risk attitudes.

Questions about compound interest are, by now, standard procedure in surveys about financial literacy—see e.g. the relevant module in the Health and Retirement Survey documented in Lusardi and Mitchell (2011). The typical evidence is that calculations of multiplicative growth effects show a strong downward bias, often to the extent that all compounding is ignored. The bias seems robust and economically important. Our experiments arguably give the respondents a very good shot at correctly detecting the speed of exponential growth, especially since we use highly selected and quantitatively skilled students and in most of our treatments we provide them with calculators. Indeed we do not find evidence of the exponential growth bias in settings where growth is quasi-deterministic and calculators are available. It is perhaps all the more notable that we nevertheless find a strong bias in analogous stochastic settings. With random growth, the effects of compounding are reliably, and in many cases strongly, under-appreciated.

References

- [1] Abbink, K., and Rockenbach, B. (2006). *Option pricing by students and professional traders: a behavioral investigation*. *Managerial and Decision Economics* 27 (6): 497-510.
- [2] Benartzi, S., and Thaler, R.H. (1999). *Risk aversion or myopia? Choices in repeated gambles and retirement investments*. *Management Science* 45: 364-381.
- [3] Br  nner, T., Levinsky, R., and Qiu, J. (2011). *Preference for skewness: evidence from a binary choice experiment* *European Journal of Finance*, 17: 525-538.
- [4] Chen, H., and Rao, A.R. (2007). *When two plus two is not equal to four: Errors in processing multiple percentage changes*. *Journal of Consumer Research* 34: 327-340.
- [5] Christandl, F., and Fetchenhauer, D. (2009). *How laypeople and experts misperceive the effect of economic growth*. *Journal of Economic Psychology* 30: 381-92.
- [6] Cox, J.C., Ross, S.A., and Rubinstein, M. (1979). *Option pricing: a simplified approach*. *Journal of Financial Economics* 7: 229-263.
- [7] Deck, C., and Schlesinger, H. (2010). *Exploring higher order risk effects* *Review of Economic Studies* 77: 1403-1420.
- [8] Ebert, S., and Wiesen, D. (2011). *Testing for prudence and skewness-seeking* *Management Science* 57: 1334-1349.
- [9] Eckel, C.C., and Grossman, P.J. (2014) *Loving the long shot: risk taking with skewed lotteries*, Working paper, Texas A&M University.
- [10] Eisenstein, E.M., and Hoch, S.J. (2005). *Intuitive compounding: Framing, temporal perspective, and expertise*. Working paper, Johnson Graduate School of Management, Cornell University.
- [11] Ensthaler, L., Nottmeyer, O., Weizs  cker, G., and Zankiewicz, C. (2013) *Hidden skewness: on the difficulty of multiplicative compounding under random shocks*, Working paper, Humboldt University Berlin.
- [12] Financial Industry Regulatory Authority (2009), FINRA Regulatory Note 09-31:1-5
- [13] Fischbacher, U. (2007). *z-Tree: Zurich Toolbox for Ready-made Economic Experiments*, *Experimental Economics* 10(2): 171-178.
- [14] Gneezy, U. (1996). *Probability judgments in multi-stage problems: experimental evidence of systematic biases*. *Acta Psychologica* 93: 59-68.
- [15] Grether, D. (1981). *Financial Incentive Effects and Individual Decision Making* Working Paper 401, California Institute of Technology.

- [16] Holt, C.A. (2007). *Markets, Games & Strategic Behavior*.
- [17] Kemp, S. (1984). *Perception of changes in the cost of living*. Journal of Economic Psychology 5(4): 313-323.
- [18] Klos, A., Weber, E.U., and Weber, M. (2005). *Investment decisions and time horizon: risk perception and risk behavior in repeated gambles*. Management Science 51: 1777-1790.
- [19] Levy, M., and Tasoff, J. (2014). *Exponential-growth bias*. Working paper, London School of Economics and Political Science.
- [20] Lusardi, A., and Mitchell, O.S. (2011). *Financial literacy and planning: Implications for retirement wellbeing*. NBER Working Paper 17078.
- [21] Mörters, P., and v. Weizsäcker, H. (2009). *Stochastische Methoden*. (3rd edition) Universität Kaiserslautern.
- [22] Redelmeier, D.A., and Tversky, A. (1992). *On the framing of multiple prospects*. Psychological Science 3(3): 191-93.
- [23] Samuelson, P. A. (1963). *Risk and uncertainty: a fallacy of large numbers*. Scientia 98: 108-113.
- [24] Stango, V., and Zinman, J. (2009). *Exponential growth bias and household finance*. Journal of Finance 64(6): 2807-2849.
- [25] Stutzer, M., and Jung Grant, S. (2010). *Expected return or growth rate? Choices in repeated gambles that model investments*. Working paper, Leeds School of Business, University of Colorado.
- [26] Wagenaar, W.A., and Sagaria, S.D. (1975). *Misperception of exponential growth*., Perception & Psychophysics 18(6): 416-422.
- [27] Wagenaar, W.A., and Timmers, H. (1978). *Extrapolation of exponential time series is not enhanced by having more data points*., Perception & Psychophysics 24(2): 182-184.
- [28] Wagenaar, W.A., and Timmers, H. (1979). *The pond-and-duckweed problem: Three experiments on the misperception of exponential growth*., Acta Psychologica 43(3): 239-251.
- [29] Wooldridge, J.M. (2002). *Econometric analysis of cross section and panel data*., Cambridge, MA: MIT Press.

A Tables

A.1 Median perceptions in round 2 to 5 of Study 1(a)

Share of participants switching from A to B									
Range of subjective median for Security A	Round 2		Round 3		Round 4		Round 5		
	NO_HELP	HELP	NO_HELP	HELP	NO_HELP	HELP	NO_HELP	HELP	
[0 – 100)	0.000	0.032	0.000	0.047	0.018	0.046	0.000	0.092	
[100 – 500)	0.000	0.016	0.000	0.000	0.000	0.046	0.035	0.046	
[500 – 2,000)	0.054	0.612	0.072	0.666	0.072	0.676	0.107	0.661	
[2,000 – 6,000)	0.145	0.145	0.127	0.095	0.200	0.138	0.303	0.046	
[6,000 – 9,000)	0.090	0.048	0.254	0.063	0.309	0.046	0.142	0.061	
[9,000 – 12,000)	0.381	0.064	0.309	0.063	0.236	0.000	0.196	0.030	
[12,000 – 20,000)	0.181	0.064	0.109	0.031	0.127	0.462	0.142	0.046	
[20,000 – 35,000)	0.090	0.000	0.109	0.015	0.036	0.000	0.053	0.015	
[35,000 – 90,000)	0.054	0.000	0.000	0.015	0.000	0.000	0.017	0.000	
[90,000 – 250,000)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
[250,000 – ∞)	0.000	0.016	0.018	0.000	0.000	0.000	0.000	0.000	

Table 11: Subjective medians for rounds 2-5.

A.2 Fractions of decisions supporting the hypotheses.

Treatment	Proportion Round 1		Proportion Round 2		Proportion Round 3		Proportion Round 4		Proportion Round 5	
<i>HVS</i>	0.552	(0.09)	0.758***	(0.07)	0.827***	(0.07)	0.896***	(0.06)	0.827***	(0.07)
<i>HVL</i>	0.931***	(0.05)	0.827***	(0.07)	0.827***	(0.07)	0.896***	(0.06)	0.931***	(0.05)

Significance levels: * : <10% ** : <5% *** : <1%

Table 12: Proportions of participants strictly underestimating skewness over the rounds. Significance values referring to a binomial test of the proportions being significantly larger than 0.5.

Treatment	Proportion Round 1		Proportion Round 2		Proportion Round 3		Proportion Round 4		Proportion Round 5	
<i>HVL</i>	0.896***	(0.06)	0.965***	(0.03)	1.000***	(0.00)	0.965***	(0.03)	0.931***	(0.05)

Significance levels: * : <10% ** : <5% *** : <1%

Table 13: Proportions of participants strictly underestimating the price spread over the rounds. Significance values referring to a binomial test of the proportions being significantly larger than 0.5.

Treatment	Proportion Round 1		Proportion Round 2		Proportion Round 3		Proportion Round 4		Proportion Round 5	
<i>LVL</i>	0.413	(0.09)	0.413	(0.09)	0.241	(0.08)	0.172	(0.07)	0.103	(0.06)
<i>LVLNC</i>	0.844***	(0.06)	0.563	(0.09)	0.344	(0.09)	0.219	(0.07)	0.219	(0.07)

Significance levels: * : <10% ** : <5% *** : <1%

Table 14: Proportions of participants strictly underestimating the median over the rounds. Significance values referring to a binomial test of the proportions being significantly larger than 0.5.

A.3 Interval regressions

	10th Percentile		50th Percentile		90th Percentile	
	Mean	Std.Err.	Mean	Std.Err.	Mean	Std.Err.
<i>Rational Model</i>	39.69		88.64		197.98	
<i>LB Model</i>	35.90		116.10		196.30	
<i>Round 1</i>	83.56	(11.8)	116.56	(12.7)	221.40	(13.5)
<i>Round 2</i>	70.63	(11.8)	108.84	(11.7)	206.84	(11.4)
<i>Round 3</i>	68.74	(11.7)	108.97	(9.5)	191.24	(11.2)
<i>Round 4</i>	62.72	(6.7)	122.41	(9.8)	183.57	(9.2)
<i>Round 5</i>	54.55	(6.3)	120.79	(11.2)	181.92	(11.9)

Table 15: Interval regression estimates for the mean perceptions of the three elicited percentiles in HVS complemented by LB and rational predictions at the top of each column.

	10th Percentile		50th Percentile		90th Percentile	
	Mean	Std.Err.	Mean	Std.Err.	Mean	Std.Err.
<i>Rational Model</i>	1.20		29.96		745.58	
<i>LB Model</i>	-59.80		261.00		581.80	
<i>Round 1</i>	82.27	(19.2)	229.80	(27.3)	597.49	(53.2)
<i>Round 2</i>	100.65	(28.9)	202.24	(38.9)	442.75	(50.8)
<i>Round 3</i>	53.68	(9.3)	150.07	(18.2)	373.06	(37.4)
<i>Round 4</i>	79.75	(22.9)	175.52	(24.4)	385.44	(40.9)
<i>Round 5</i>	68.90	(23.4)	165.97	(24.8)	420.60	(59.9)

Table 16: Interval regression estimates for the mean perceptions of the three elicited percentiles in HVL complemented by LB and rational predictions at the top of each column.

	10th Percentile		50th Percentile		90th Percentile	
	Mean	Std.Err.	Mean	Std.Err.	Mean	Std.Err.
<i>Rational Model</i>	117.12		117.36		117.59	
<i>LB Model</i>	115.90		116.10		116.30	
<i>Round 1</i>	112.78	(0.92)	114.61	(0.87)	119.36	(0.82)
<i>Round 2</i>	112.91	(0.83)	115.47	(0.81)	117.93	(0.90)
<i>Round 3</i>	113.99	(0.76)	115.49	(0.64)	118.21	(0.77)
<i>Round 4</i>	115.18	(0.42)	116.28	(0.58)	118.21	(0.73)
<i>Round 5</i>	115.03	(0.67)	115.37	(0.66)	117.26	(0.63)

Table 17: Interval regression estimates for the mean perceptions of the three elicited percentiles in LVS complemented by LB and rational predictions at the top of each column.

	10th Percentile		50th Percentile		90th Percentile	
	Mean	Std.Err.	Mean	Std.Err.	Mean	Std.Err.
<i>Rational Model</i>	491.79		495.69		499.63	
<i>LB Model</i>	260.20		261.00		261.80	
<i>Round 1</i>	365.65	(34.9)	483.87	(29.6)	634.18	(37.9)
<i>Round 2</i>	403.14	(35.1)	510.56	(31.8)	613.80	(36.1)
<i>Round 3</i>	395.55	(26.9)	503.36	(22.4)	590.81	(29.4)
<i>Round 4</i>	451.99	(24.7)	489.91	(25.6)	547.23	(29.5)
<i>Round 5</i>	424.01	(24.5)	509.23	(22.0)	559.20	(25.6)

Table 18: Interval regression estimates for the mean perceptions of the three elicited percentiles in LVL complemented by LB and rational predictions at the top of each column.

	10th Percentile		50th Percentile		90th Percentile	
	Mean	Std.Err.	Mean	Std.Err.	Mean	Std.Err.
<i>Rational Model</i>	491.79		495.69		499.63	
<i>LB Model</i>	260.20		261.00		261.80	
<i>Round 1</i>	301.48	(17.8)	381.39	(21.28)	514.08	(32.7)
<i>Round 2</i>	350.25	(24.5)	449.63	(12.76)	557.62	(19.7)
<i>Round 3</i>	382.42	(24.4)	466.01	(14.51)	560.03	(26.2)
<i>Round 4</i>	408.99	(25.8)	484.08	(15.86)	557.55	(22.1)
<i>Round 5</i>	416.87	(22.8)	472.47	(12.39)	540.17	(21.6)

Table 19: Interval regression estimates for the mean perceptions of the three elicited percentiles in LVLNC complemented by LB and rational predictions at the top of each column.

	10th Percentile		50th Percentile		90th Percentile	
	Sim.Value	Bootstr.Conf.Inter.	Sim.Value	Bootstr.Conf.Inter.	Sim.Value	Bootstr.Conf.Inter.
<i>Rational Model</i>	98.80	[94.95; 101.25]	152.98	[146.18; 162.40]	310.98	[292.75; 322.40]
<i>LB Model</i>	109.36	[104.21; 113.83]	162.72	[156.63; 169.16]	227.10	[221.59; 230.66]
	Mean	Std.Err.	Mean	Std.Err.	Mean	Std.Err.
<i>Round 1</i>	168.02	(24.2)	316.25	(30.8)	722.22	(95.5)
<i>Round 2</i>	159.65	(17.2)	265.84	(27.7)	513.36	(64.8)
<i>Round 3</i>	151.62	(14.4)	246.15	(21.5)	468.22	(58.2)
<i>Round 4</i>	160.56	(13.5)	315.93	(50.9)	551.12	(75.1)
<i>Round 5</i>	155.61	(14.5)	233.18	(21.7)	477.73	(63.9)

Table 20: Interval regression estimates for the mean perceptions of the three elicited percentiles in ETF_1 complemented by LB and rational predictions at the top of each column.

	10th Percentile		50th Percentile		90th Percentile	
	Sim.Value	Bootstr.Conf.Inter.	Sim.Value	Bootstr.Conf.Inter.	Sim.Value	Bootstr.Conf.Inter.
<i>Rational Model</i>	32.39	[29.24; 36.31]	152.14	[123.17; 178.88]	1,359.42	[1,163.44; 1,555.64]
<i>LB Model</i>	128.09	[112.64; 141.51]	288.17	[269.91; 307.49]	481.30	[464.79; 491.98]
	Mean	Std.Err.	Mean	Std.Err.	Mean	Std.Err.
<i>Round 1</i>	134.87	(21.3)	297.83	(55.7)	582.54	(82.7)
<i>Round 2</i>	156.07	(27.8)	302.48	(43.3)	681.32	(93.8)
<i>Round 3</i>	172.03	(22.7)	295.44	(34.4)	605.39	(86.4)
<i>Round 4</i>	155.59	(29.6)	268.41	(35.3)	636.71	(87.7)
<i>Round 5</i>	145.72	(29.7)	248.39	(41.5)	624.07	(93.6)

Table 21: Interval regression estimates for the mean perceptions of the three elicited percentiles in ETF_3 complemented by LB and rational predictions at the top of each column.

A.4 Average Perceptions for *Upper Bounds*

We replicate Tables 4, 5, 6, 7, 8 and 10 for the elicited upper bounds of the subjective quantiles of the selling price.

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
γ_i^{HVS}	0.388*	(0.06)	0.315***	(0.06)	0.210***	(0.07)	0.067***	(0.07)	0.100***	(0.09)
γ_i^{HVL}	0.336***	(0.06)	0.309***	(0.07)	0.353***	(0.06)	0.183***	(0.08)	0.226***	(0.07)

Significance levels: * : <10% ** : <5% *** : <1%

Table 22: Average values for γ_i^k over the five rounds, with $k \in \{HVS, HVL\}$. The indicated significance values refer to one-sided parametric t-tests for downward deviations from the rational values.

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
δ_i^{HVL}	584.34***	(62.4)	393.20***	(55.2)	375.44***	(43.5)	351.65***	(49.5)	406.65***	(66.5)

Significance levels: * : <10% ** : <5% *** : <1%

Table 23: Average values for δ_i^{HVL} over the five rounds. The indicated significance values refer to one-sided parametric t-tests for downward deviations from the rational values.

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
$\frac{\gamma_i^{HVS}}{\gamma^{HVS}}$	0.802	(0.12)	0.651	(0.12)	0.435	(0.14)	0.138	(0.15)	0.206	(0.17)
$\frac{\gamma_i^{HVL}}{\gamma^{HVL}}$	0.481	(0.08)	0.442	(0.09)	0.506	(0.08)	0.262	(0.11)	0.324	(0.10)
Δ_γ	-0.320**	(0.15)	-0.208*	(0.15)	0.070	(0.16)	0.123	(0.19)	0.117	(0.20)
$\frac{\delta_i^{HVS}}{\delta^{HVS}}$	0.766	(0.08)	0.773	(0.08)	0.703	(0.08)	0.699	(0.07)	0.739	(0.09)
$\frac{\delta_i^{HVL}}{\delta^{HVL}}$	0.585	(0.06)	0.394	(0.05)	0.376	(0.03)	0.352	(0.05)	0.407	(0.08)
Δ_δ	-0.180*	(0.10)	-0.379***	(0.09)	-0.326***	(0.09)	-0.346***	(0.09)	-0.332***	(0.11)

Significance levels: * : <10% ** : <5% *** : <1%

Table 24: Average values of relative deviation measures with the significance levels referring to the one-sided parametric Welch test of Δ_δ and Δ_γ being significantly smaller than zero.

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
$q_{0.1,i}^{LVL}$	406.0***	(38.0)	443.9**	(40.0)	432.9***	(25.2)	487.8	(28.4)	457.3***	(28.5)
$q_{0.5,i}^{LVL}$	532.4	(36.5)	560.8	(40.4)	544.8	(29.0)	530.5	(31.3)	549.8	(28.5)
$q_{0.9,i}^{LVL}$	708.3	(44.3)	682.9	(42.7)	652.9	(36.3)	600.5	(35.7)	610.5	(32.5)

Significance levels: * : <10% ** : <5% *** : <1%

Table 25: Average values for $q_{j,i}^{LVL}$ over the five rounds, with $j \in \{0.1, 0.5, 0.9\}$. The listed significance values refer one-sided parametric t-tests for downward deviations from the rational value of 520 for the upper bounds in LVL.

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
$q_{0.1,i}^{LVLNC}$	328.1***	(18.0)	385.0***	(26.9)	414.2***	(24.8)	446.6***	(29.5)	449.53***	(25.3)
$q_{0.5,i}^{LVLNC}$	417.0***	(26.2)	486.3**	(16.5)	504.4	(18.1)	530.0	(22.3)	507.9	(15.8)
$q_{0.9,i}^{LVLNC}$	576.7	(40.5)	625.0	(26.7)	621.1	(33.4)	616.6	(29.4)	593.1	(28.9)

Significance levels: * : <10% ** : <5% *** : <1%

Table 26: Average values for $q_{j,i}^{LVLNC}$ over the five rounds, with $j \in \{0.1, 0.5, 0.9\}$. The listed significance values refer one-sided parametric t-tests for downward deviations from the rational value of 520 for the upper bounds in LVLNC.

Measure	Round 1		Round 2		Round 3		Round 4		Round 5	
$\frac{\gamma_i^{ETF_1}}{\gamma^{ETF_1}}$	0.751	(0.17)	0.511	(0.21)	0.587	(0.20)	0.552	(0.18)	0.774	(0.15)
$\frac{\gamma_i^{ETF_3}}{\gamma^{ETF_3}}$	0.423	(0.11)	0.478	(0.09)	0.285	(0.13)	0.515	(0.10)	0.575	(0.10)
Δ_γ	-0.328*	(0.20)	-0.032	(0.23)	-0.301	(0.24)	-0.036	(0.20)	-0.197	(0.18)
$\frac{\delta_i^{ETF_1}}{\delta^{ETF_1}}$	3.603	(0.57)	2.279	(0.39)	2.010	(0.37)	2.533	(0.52)	2.075	(0.39)
$\frac{\delta_i^{ETF_3}}{\delta^{ETF_3}}$	0.353	(0.07)	0.413	(0.07)	0.348	(0.07)	0.388	(0.07)	0.382	(0.07)
Δ_δ	-3.250***	(0.46)	-1.865***	(0.39)	-1.662***	(0.37)	-2.144***	(0.52)	-1.693***	(0.39)

Significance levels: * : <10% ** : <5% *** : <1%

Table 27: Average values of relative deviation measures with the significance levels referring to the one-sided parametric Welch test of Δ_δ and Δ_γ being significantly smaller than zero.