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

Ludwig Ensthaler† Olga Nottmeyer‡ Georg Weizsäcker§
and Christian Zankiewicz¶

July 6, 2016

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 spread of the outcome distribution to be too narrow-band and they estimate the median and the mode to lie too close to the center of the distribution, failing to account for the compound nature of average growth. The observed biases are measured irrespective to risk preferences and they appear under a variety of conditions. The biases are largely consistent with a behavioral model in which geometric growth is confused with linear growth. This confusion is a possible driver of investors' difficulties with real-world financial products like leveraged ETFs and retirement savings plans.

JEL-Classification: C91, D03, D14, G02

Keywords: Behavioral Economics, Multiplicative Compounding, Skewness Neglect, Exponential Growth Bias

---

*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. The paper has much benefitted from the editorial work of the Editor, an Associate Editor and three referees. Financial support by the ERC (Starting Grant 263412) is gratefully acknowledged.

†Rocket Internet and WZB Berlin Social Science Center, ludwig@ensthaler.com

‡IZA Bonn, nottmeyer@iza.org

§Humboldt University Berlin and DIW Berlin, weizsaecker@hu-berlin.de

¶DIW Berlin, czankiewicz@diw.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.

Two biases may arise when forecasting the distribution of long-run growth. First, one may fail to compound the effects of multi-period growth, for a given growth rate. Second, one may ignore the skewness that arises over time and, e.g., confuse the mean return with the median return. The first of these biases, which we call "linearity bias" hereafter, has been studied predominantly in deterministic settings. For example, when asked to assess the total effect of accumulating 7% growth for ten periods, a substantial fraction of respondents gives an answer that is closer to 70% than to the actual 97%. The analyses of Stango and Zinman (2009) and Levy and Tasoff (2015) indicate that the bias is empirically relevant as it affects households' borrowing and saving decisions.[1] The effect is usually referred to as "exponential growth bias", but the main point of this paper is to extend the analysis of the exponential growth bias—appropriately defined—to stochastic settings and to demonstrate that it also includes the second bias. This second bias, which we refer to as "skewness neglect", is less well known in the academic literature[2] but investment practitioners and financial market regulators are aware of its effects (see Section 5). Investors apparently need to be made aware that the compounding of random growth can transform a symmetric 1-period return distribution into a skewed multi-period return distribution. An important real-world example of this is the family of leveraged exchange-traded funds (leveraged ETFs). These assets are highly volatile and have a fairly symmetric 1-period return distribution; holding them for multiple periods results in severe skew.

Our paper presents a series of incentivized laboratory experiments that extends the evidence on the perception of multiplicative growth to the stochastic domain and accounts for both of the above-described biases. As an example that demonstrates skewness neglect, consider the following stylized experiment. 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 value of 10,000, what would the asset likely be worth

[1]Both Stango and Zinman (2009) and Levy and Tasoff (2015) present survey evidence of a statistical connection between the bias and respondents' savings behaviors. Levy and Tasoff (2015) also analyse theoretical implications of the bias, e.g., an overestimation of future income that arises from too moderate time discounting of income. The effect can result in overconsumption if income is shifted to later time periods. Related effects are addressed in the experiments by McKenzie and Liersch (2011).

[2]The only other academic study that we are aware of is by Stutzer and Grant (2013), discussed in the next section.

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:[3] a) up to 6,400, b) between 6,400 and 12,800, c) between 12,800 and 19,200, d) between 19,200 and 25,600, or e) above 25,600. We then simulate the random process and if the simulated path ends up in the outcome range that a participant has bet on, she receives a prize of €20. If not, she receives nothing. The most popular answer is c), chosen by 43% of the 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); the median of the resulting distribution is 989 and the probability that the process ends up in the lowest interval is 80%. A simple reasoning for this 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 in period 12. The participants fail to realize this and instead report answers that are consistent with a confusion of mean and median. Their average expected payoff (based on their decisions) amounts to a meager €2 in this experiment, whereas the optimal response would earn them €16 in expectation.

Our series of experiments examines this kind of mistake systematically and finds that the participants' perception of stochastic growth deviates in predictable ways from the rational prediction. Both of the above biases are found to be relevant. Overall, the experimental results are in line with a simple model of misperceiving compound shocks. This model, which we label "exponential growth bias model", has the agent perceive growth as a linear process, in the sense that all multiplicative growth is mistaken as additive growth.[4] The model predicts both linearity bias and skewness neglect.

Importantly, the model also allows predictions about the strength of the two effects. It predicts that the agent has a fairly rational perception of the growth process in the case that both per-period volatility and per-period return are low. For larger volatility, skewness neglect becomes relevant and leads to an overestimation of the median. As the per-period return increases, the distribution of returns becomes more symmetric, the linearity bias becomes more dominant, and the agent underestimates the median. This somewhat intricate pattern of predictions cannot be generated by any of the biases alone but it is confirmed by the experimental data. Subsection 3.2 shows these effects in our main (novel) experiments, asking the participants to predict the most likely outcome of a growth process. In the binomial-tree processes that we use in this experiment, the most likely outcome is also the median and thus the responses can be used for assessing the subjectively perceived medians. All of our experiments are incentivized in ways that make truth-telling optimal irrespective of one's risk attitudes. The experiments of Sections 3.3, 4 and 5, and those in the Appendix, go on to test the predictions of the model in different variations of the experimental setting. We vary the incentive schemes, the level of feedback as well as the nature of both the investment strategy and the underlying asset that the participants are asked to assess. Thereby, we can investigate the robustness of the effects and we can also inquire about several other implications of the exponential growth bias model. Qualitatively, almost all predictions of the model are borne out in the data, and often with large discrepancies to

[3] Appendix A.1 contains the details of the procedure. The experiment is not one of the main treatments of the paper.

[4] The model is akin to that of Levy and Tasoff (2015) albeit developed independently.

the rational prediction. For example, in treatments with high return volatility, about 90% of the participants overestimate the median. The model is also fairly successful in predicting the participants' misperceptions of the 10th and 90th percentile of the relevant long-run distributions: the 90-10 spread is generally underestimated. This holds both for binomial-tree assets and also for a class of more realistic assets that we base on the historical returns of the German DAX index. In the latter set of experiments, we model assets that emulate leveraged ETFs where the DAX index is the underlying asset. Such leveraged products have been popular with household investors in the U.S. in recent years (though with a different underlying asset) until many investors made unexpected and seemingly unexplainable losses. Our analysis offers an explanation for the confusion related to these products: the investors appear to have been ignorant of the skew arising from high per-period volatility.

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 Timmers (1978) show that this linearity bias is robust to the amount of information available to the participants and Wagenaar and Timmers (1979) demonstrate robustness of the effect 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 linearity bias. 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), McKenzie and Liersch (2011) and Levy and Tasoff (2015) 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, study biases in the compounding of long term distributions from a given short term distribution. Their experimental participants choose different hypothetical retirement plans depending on whether they observe the historical return distribution of retirement plans for a 1-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), and 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.[5] This is also consistent with the only experimental paper on skewness neglect that we found, by Stutzer and Grant (2013). Their hypothetical investment experiments find an inflated investment rate in treatments where their participants have to calculate the compound return by themselves.[6]

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 use in Sections 3 and 4 have the same structure as the underlying asset in Cox, Ross, and Rubinstein's (1979) well-known model 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 the underlying asset itself.

3 Study 1: Assessments of median and mode

Study 1 is composed of the two partial studies 1(a) and 1(b), each testing participants' perceptions of an asset's mode and/or median. We consider one of the most elementary assets covered in the finance literature: the binomial-tree asset with fixed maturity (Cox, Ross and Rubinstein (1979)). The multiplicative per-period growth μ_t of this asset is a binary random variable $\mu_t \in \{\mu^h, \mu^l\}$, where the percental uptick $\mu^h \geq 0$ and the percental downtick $\mu^l \geq 0$ are equiprobable in each period $t = 1, ..., T$.

3.1 The exponential growth bias

We start the analysis by presenting a simple model of biased decision making. Consider a decision maker who ignores compounding of interest: 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. This feature is the sole bias of our model—the exponential growth bias—and we can readily extend it to the domain of stochastic growth.

Formally, let Y_0 denote the known initial price of an asset with a random price series $\{Y_0, Y_1, ...\}$ and let μ_t be the random variable describing the relative price changes occurring in t, e.g., $Y_1 = Y_0\mu_1$. An unbiased decision maker correctly perceives the true distribution of the period-T price as $Y_T = Y_0 \prod_{t=1}^{T} \mu_t$. In contrast, an exponential growth biased (EGB) decision maker perceives the price in t

[5] A 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.

[6] The experiment by Stutzer and Grant (2013) uses a quite similar experimental wording as the experiment described in Subsection 3.3 and in the first version of this paper (2010), despite having been developed and written independently. A separate and important experimental literature examines the preferences regarding skewness, see Deck and Schlesinger (2010), Brünner, 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.

as $\tilde{Y}_T = Y_0(1 + \sum_{t=1}^{T}(\mu_t - 1))$. That is, for each t she perceives the absolute difference $Y_t - Y_{t-1}$ to be given by t's growth rate applied to the initial value Y_0. As a result, the EGB decision maker misses out on all effects of multiplicative compounding, which may or may not occur in the true growth process.

To investigate the effects of the bias, Study 1 and Study 2 consider binomial-tree price series with a constant distribution of relative price growth μ_t (but with distributions of absolute differences that vary across time, which is ignored by the EGB decision maker). Here, while the actual distribution of Y_t is skewed and approaches a lognormal distribution for large t, the EGB decision maker perceives a binomial distribution that is symmetric with its mean being equal to the median and mode: symmetry is preserved under addition of random variables.[7] The EGB model thus predicts full skewness neglect. In particular, one can check that with a strictly positive per-period average growth $\mathbb{E}[\mu] > 0$ and under the condition that $\mu^h\mu^l < 1$, the EGB decision maker overestimates the median for $t > 1$. Under the same conditions, the model also predicts a directed linearity bias: the EGB decision maker underestimates the mean, for $t > 1$. One can also check that if both μ^h and μ^l are increased by the same amount Δ, then the distribution of Y_t becomes more and more symmetric, so that the EGB decision maker's skewness neglect becomes less and less important.

In the following, we present our experiments that test these qualitative (directed) predictions of biased decision making, in each case generated by simple numerical applications of the EGB model.[8] Our main empirical focus lies on measuring the perception of the median of Y_T. We also exploit the fact that for binomial-tree processes the median is identical to the mode of Y_T, to which both optimal and EGB decision makers agree (despite disagreeing on the value). This allows formulating alternative elicitation tasks, equivalently asking for median or mode.

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

3.2.1 Experimental design

Participants in Study 1(a) are presented with a security whose price is currently at $Y_0 = 100$ and changes by a factor $\mu_t \in \{\mu^h, \mu^l\}$ with equal probabilities during each period and with all random draws being independent. The participants' task is to locate the mode of the security's outcome distribution after $T = 12$ periods. The task is made incentive compatible as follows. After a participant's response, the experimenter simulates a set of 100 values of Y_T. If at least one of these simulated values differs by less than 1 from the participant's stated value she receives a bonus of €20, otherwise not. The procedure thus prompts the participant to report the location (more precisely, an interval of length 2) where she perceives Y_T's highest likelihood. The optimal response would be to report the mode of Y_T. Notice that reporting the mode is optimal irrespective of risk preferences and of

[7] This error could also be interpreted as the decision maker wrongly computing the arithmetic mean over returns when calculating the median instead of working with log-returns.

[8] 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 our statistical analysis on the qualitative predictions.

the nature of the perceived Y_T: the incentive scheme uses only two possible payments—receive a bonus versus not—making it optimal for any participant with monotonic preferences to maximize the subjectively perceived probability of receiving the bonus by stating the price that she thinks is most likely. The procedure is also simple to understand and allows asking a straightforward question about the participants' prediction of the price evolution of the asset.[9]

Each participant is asked to report a prediction on two different securities in order to increase the number of observations and thereby the power of our statistical tests. One of the two responses is randomly picked to be payoff relevant at the conclusion of the experiment.

Overall, Study 1(a) covers four securities that only differ in the values $\{\mu^h, \mu^l\}$, appearing pairwise in two treatments. Participants in treatment 1 assess the modal values of Security 1 ($\mu^h = 1.7, \mu^l = 0.4$) and Security 2 ($\mu^h = 1.075, \mu^l = 1.025$), whereas participants in treatment 2 assess the modal values of Security 3 ($\mu^h = 1.4, \mu^l = 0.7$) and Security 4 ($\mu^h = 1.8, \mu^l = 1.1$). Each participant thus faces one security which can depreciate as well as appreciate, and one security which can only appreciate. In treatment 1, the two securities have identical means but different per-period volatilities (as measured by the spread $(\mu^h - \mu^l)$) and in treatment 2, the two securities have different means but identical per-period volatilities. Moreover, the mean of Security 3 is identical to that of Security 1 and 2.

Participants are randomly assigned to treatments 1 or 2. To account for possible learning effects, the order of the two securities randomly varies between the participants within a treatment. All 127 participants (63 in treatment 1 and 64 in treatment 2) are students at Technical University Berlin. Six sessions, three in each treatment, are conducted in a computer-based format using the software z-Tree (Fischbacher (2007)). Participants receive a participation fee of €5 in addition to their possible bonus of €20.

3.2.2 Exponential growth bias prediction

The securities in Study 1(a) are specified such that they allow predictions about the relative strengths of the above-described two effects, linearity bias and skewness neglect. To examine the effect of a large per-period volatility, we consider Security 1 ($\mu^h = 1.7, \mu^l = 0.4$) with +70% and -60% as possible percentage changes and predict a strong effect of skewness neglect. With a perceived constant distribution of absolute changes that lie in $\{-60; +70\}$, the EGB decision maker perceives a symmetric distribution of the period-12 price that locates mode, mean, and median at $\mathbb{E}(\tilde{Y}_T) = 160$. Thus, skewness neglect leads to a strong overestimation of the true mode of Security 1, which is at 9.89. This effect stems entirely from skewness neglect, whereas linearity bias has only a mild effect: the EGB decision maker's belief about the mean is close to the true mean $\mathbb{E}(Y_T) = 179.59$.

For a relatively lower per-period volatility, as captured by Security 3 ($\mu^h = 1.4, \mu^l = 0.7$) with +40% and -30% as possible percentage changes, skewness neglect becomes less extreme and results in a weaker, but still sizable, overestimation of the mode: the EGB model predicts a response of 160 in-

[9] The procedure is novel to the experimental literature, to our knowledge. All instructions are in the online appendix, including the instructions for an understanding test that participants had to pass.

stead of the correct 88.58.

Further decreasing per-period volatility, as for Security 2 ($\mu^h = 1.075, \mu^l = 1.025$), results in the predictions that the EGB decision maker has a fairly rational perception of the growth process: she perceives the most likely period-12 price of Security 2 at 160 instead of the correct 178.97. That is, even a model allowing for both skewness neglect and linearity bias is fairly ineffective here and predicts a mild underestimation of Security 2's mode.

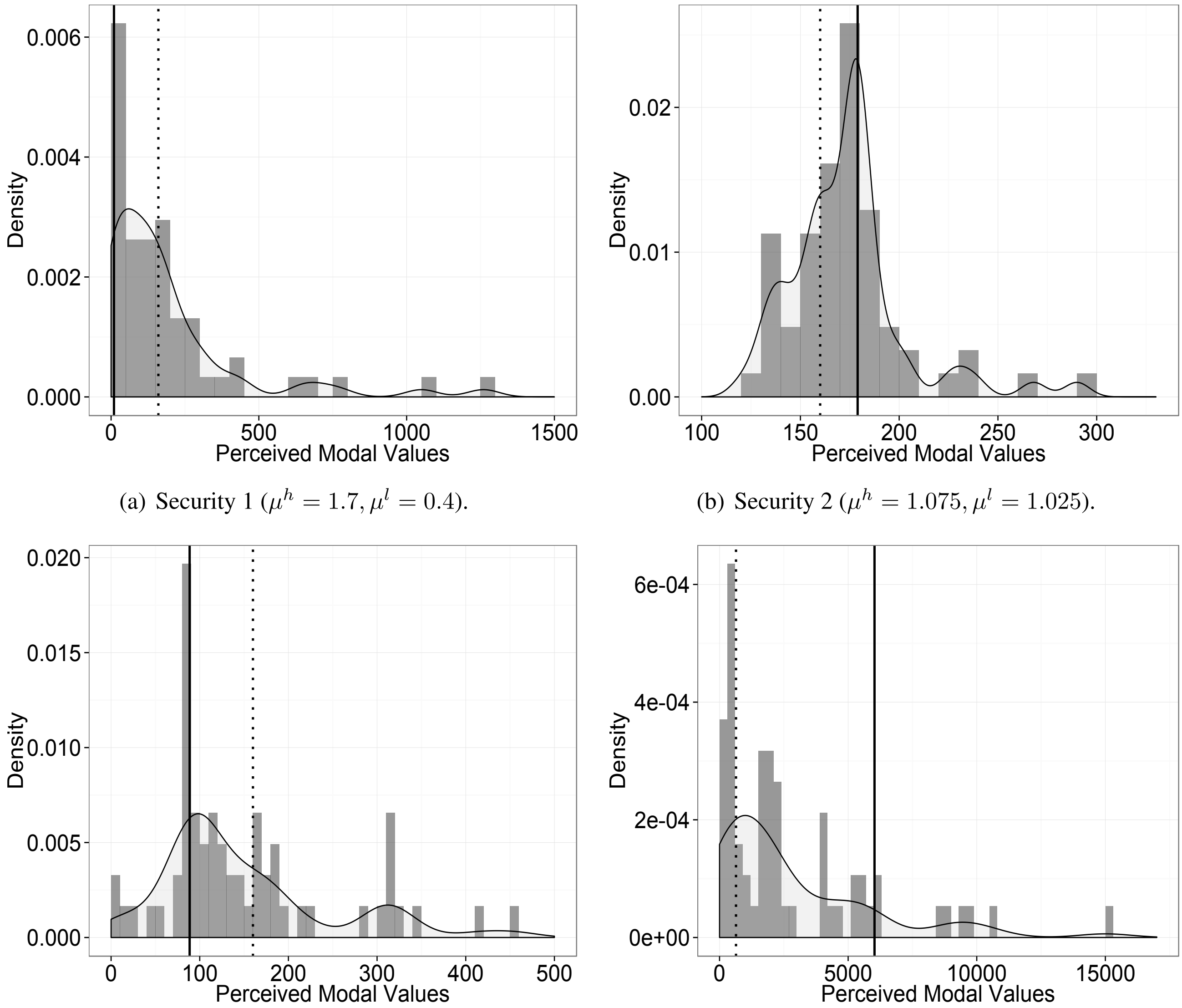
Security 4 ($\mu^h = 1.8, \mu^l = 1.1$) allows for a much stronger effect towards underestimation, which is due to linearity bias. Here, the EGB decision maker perceives the most likely period-12 price at 640, while the true mode of the price distribution lies at 6,025.47. With such high per-period mean growth, skewness considerations become less important than the effect of linearity bias.

3.2.3 Results

Figure 1 illustrates the distributions of subjective mode perceptions for all four securities estimated from our experimental data. The participants' predictions for Security 1 ($\mu^h = 1.7, \mu^l = 0.4$) are displayed in Figure 1(a) and show a substantial degree of overestimation (i.e., most of the probability mass is located to the right of the vertical solid line, indicating the optimal response). Consistent with skewness neglect, 87% of the participants overestimate the true mode and this frequency of overestimation lies significantly above 50% (p-value<0.001, one-sided binomial test). Although the data show a peak in the neighborhood around the optimal value, most participants' degree of overestimation is substantial. Half of them predict the mode of the distribution to lie above 120—more than 12 times the true value.

Figure 1(c) illustrates the subjective mode perceptions for Security 3 ($\mu^h = 1.4, \mu^l = 0.7$). As for Security 1, the data show a notable proportion of participants, 70%, overestimating the mode. While the frequency of overestimation lies significantly below that of Security 1 (p-value<0.001, two-sample binomial test), it is still significantly greater than 50% (p-value<0.001, one-sided binomial test). These observations are consistent with the EGB model in the sense that the model predicts an overestimation for both securities and a larger overestimation for Security 1 than for Security 3. But also apart from the model, the comparison between Security 1 and Security 3 is relevant as it shows the effect of skewness neglect in isolation: the mean is constant between them whereas the higher volatility in Security 1 changes median and mode. The higher frequency of overestimation in Security 1 illustrates that participants do not fully appreciate this difference.

The participants' assessments of Security 2 ($\mu^h = 1.075, \mu^l = 1.025$) are depicted in Figure 1(b) and (again consistent with the EGB model) show a different picture. Only a minor underestimation for Security 2 appears: 57% of participants state modal values below the true mode, a proportion that is not significantly greater than 50% (p-value>0.15, one-sided binomial test). Moreover, the median response is not significantly different from the optimal value (p-value>0.05, Wilcoxon signed-rank test). Neither skewness neglect nor linearity bias show to be relevant for this security.



(a) Security 1 ($\mu^h = 1.7, \mu^l = 0.4$).

(b) Security 2 ($\mu^h = 1.075, \mu^l = 1.025$).

(c) Security 3 ($\mu^h = 1.4, \mu^l = 0.7$).

(d) Security 4 ($\mu^h = 1.8, \mu^l = 1.1$).

Figure 1: Densities of the subjectively perceived modal values for securities 1 through 4. Solid lines indicate rational benchmarks at 9.89 (Security 1), 178.97 (Security 2), 88.58 (Security 3), and 6,025.47 (Security 4). Dotted lines illustrate EGB predictions for mean, mode and median.

The perceptions for Security 4 ($\mu^h = 1.8, \mu^l = 1.1$), with a higher average per-period return, are illustrated in Figure 1(d) and show a substantial degree of underestimation. Here, 89% of the participants state responses that lie below the true mode. This share is significantly larger than 50% (p-value<0.001, one-sided binomial test) and also significantly larger than the share of participants who underestimate Security 2's modal value (p-value<0.001, two-sample binomial test). Once again, these observations are consistent with the much stronger prediction of the EGB model for Security 4 than for Security 2. Moreover, it is notable that the data confirm the EGB model's prediction that linearity bias is more relevant than skewness neglect in Security 4.

3.3 Study 1(b), Robustness: Choice list format and repetitions

Study 1(b) focuses on eliciting the median. We use a novel choice list mechanism to identify bounds on the median of each participant's subjectively expected price distribution of a binomial-tree asset. As for 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. We also let the participants repeat this task over five rounds.[10]

3.3.1 Experimental design

In each round of the experiment, two risky securities are on offer and the selling price of the chosen security determines whether or not the participant receives the bonus. Security A follows a binomial-tree -60%/+70% process over 12 periods that is identical to Security 1 in Study 1(a) with the sole exception that its initial price now is 10,000. 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 a 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 equiprobable price change of -60% or +70% during the 12 periods. A participant who chooses Security B receives the bonus if the selling price of Security B exceeds a separate threshold t_B. This threshold is fixed at the initial price of 10,000 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 participant makes a choice between Security A and B for each of the 10 values of t_A, allowing us to infer bounds on her subjective median of Security A's selling price distribution. 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 distribution is 989, the rational prediction is for the participants to choose Security A in Task 1 and Task 2 and to choose Security B in all subsequent tasks.

[10] In a further treatment variation of Study 1(b), we additionally provide the participants with an explicit calculation of the distribution of compound price changes after two periods for the respective security and we point out the asymmetry in the price distribution. The observed choice bias decreases strongly in this treatment, consistent with the presumption that the bias stems from a cognitive problem and is not driven by the particular choice format. A detailed description of this treatment is in Appendix A.2.

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

Table 1: The 10 binary choices.

After the participants make their 10 choices, they receive individual feedback in the form of a sample pair of selling prices of Security 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.[11] Three sessions are conducted in a paper-and-pencil format, with 68 student participants at Technical University Berlin. Participants receive a participation fee of €5 and a possible bonus of €5 per round. That is, participants can earn up to five bonuses of €5 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.

3.3.2 Exponential growth bias prediction

The EGB model predicts that a biased decision maker perceives a binomial distribution with mean equal to mode equal to median at 16,000. She would therefore overestimate the true median (989) by an order of magnitude and choose a switching value in the interval [12,000; 20,000).

For notation, let $q_{0.5,i}$ be the elicited *lower* bound of participant i's assessment of the median: i invests in Security A for all values $t_A \leq q_{0.5,i}$ and invests in Security B for all strictly 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.[12] By analogy, let $q_{0.5}$ be the rational benchmark for $q_{0.5,i}$ (dropping the subscript i), i.e., the lower bound of the median that would be elicited from a decision maker who behaves optimally. Here and elsewhere in the paper, we focus on revealed lower bounds when applicable.

[11] Each additional round comes with the chance to earn a new bonus but this does not affect the simple optimality conditions for choice. Independent of other choices it remains optimal to choose A iff the subjective median is above t_A.

[12] If 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 switching 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.

3.3.3 Results

Range of subjective median	Share of participants switching from A to B				
	Round 1	Round 2	Round 3	Round 4	Round 5
$[0 \; ; \; 100)$	0.018	0.000	0.000	0.018	0.000
$[100 \; ; \; 500)$	0.000	0.000	0.000	0.000	0.035
$[500 \; ; \; 2,000)$	0.000	0.054	0.072	0.072	0.107
$[2,000 \; ; \; 6,000)$	0.036	0.145	0.127	0.200	0.303
$[6,000 \; ; \; 9,000)$	0.107	0.090	0.254	0.309	0.142
$[9,000 \; ; \; 12,000)$	0.411	0.381	0.309	0.236	0.196
$[12,000 \; ; \; 20,000)$	0.196	0.181	0.109	0.127	0.142
$[20,000 \; ; \; 35,000)$	0.179	0.090	0.109	0.036	0.053
$[35,000 \; ; \; 90,000)$	0.054	0.054	0.000	0.000	0.017
$[90,000 \; ; \; 250,000)$	0.000	0.000	0.000	0.000	0.000
$[250,000 \; ; \; \infty)$	0.000	0.000	0.018	0.000	0.000

Table 2: Subjective median ranges over the five rounds.

Table 2 lists the frequencies with which the participants' subjective medians lie in relevant ranges of Security A's selling price distribution, over the five rounds. In round 1, not a single participant gives the optimal response of $q_{0.5} = 500$ (i.e., optimal switching at Task 3). Instead, 98% of the participants reveal that their subjective medians are above 2,000. The results of rounds 2 to 5 show that a proportion of 86% of participants overestimate the median still in round 5. (The proportion lies significantly above 50%, with p-value<0.001 in a one-sided binomial test). The modal choice in round 1 (41% of participants) indicates a subjective median between 9,000 and 12,000, with the next-higher interval [12,000; 20,000) attracting 20% of participants' choices.

4 Study 2: Other quantiles and other growth processes

In this section we examine the robustness of the exponential growth bias predictions with respect to variations of the investment horizon and the choice of quantiles of the subjective distributions that we elicit. The results confirm the EGB model's implication that the compound distribution is perceived as too symmetric and too narrow-band if there is substantial randomness in the growth process.

4.1 Experimental design

There are two treatments in Study 2, both involving assets similar to those in Study 1. Participants can buy a 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 price of Security A moves by about +/-20% in each period: in both 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 $T^{HVL} = 140$.[13]

As in Study 1(b), 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 and are listed in Table 3. The alternative choice option is Security B which yields the bonus with a certain probability.[14] To elicit three different subjective quantiles of Security A's selling price distribution, Security B has three different specifications. Security B1 yields the bonus with 90% probability, B2 with 50% and B3 with 10%. Accordingly, each participant faces three choice lists. First, she chooses between Security 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 distribution. For example, suppose that participant i in treatment HVS chooses Security A over Security B1 in Task 1 and Task 2 and chooses Security B1 over Security A in tasks 3 through 10. Inspecting Table 3 (first column) we see that this is subjectively optimal iff participant i's subjective 10th percentile for Security A's selling price distribution is between 30 and 45. In line with our previous notation, we would thus record the elicited lower bound of i's subjective 10th percentile for Security A's selling price distribution as $q^{HVS}_{0.1,i} = 30$.

As her second set of tasks a participant faces the analogous choices between Security 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 distribution, $q^{HVS}_{0.5,i}$. Finally, she faces the analogous list of choices between Security A and B3, allowing us to infer a lower bound on her subjective 90th percentile of the same price distribution, $q^{HVS}_{0.9,i}$.[15]

After the participants make their 30 choices, the computer terminals report feedback to them 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.[16] All 58 participants are undergraduate students at University College London. Each of the two treatment conditions is faced by a random subset of participants in each session, without making them aware that other participants face different treatment conditions (N=29 in both HVS and HVL). 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] 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

[13] In three further treatments, the price motion is approximately deterministic (i.e., the price volatility is very low) and the price has positive growth with certainty. A detailed description and the results can be found in the online appendix.

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

[15] After 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.

[16] All participants passed an understanding test, in a few cases after asking for some additional explanations. In contrast to Study 1, participants in Study 2 are supplied with a hand-held calculator that they can use throughout the experiment.

[17] Participants can earn the bonus either through the quantile elicitation task or through the profit probability elicitation task (see Footnote 15). 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.

Security A and the respective B-type security. Second, the elicited quantiles must be ordered in a consistent way: a participant 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.[18]

	Values of t_A *HVS*	Values of t_A *HVL*
Task 1	15	2
Task 2	30	5
Task 3	45	15
Task 4	65	60
Task 5	95	140
Task 6	125	230
Task 7	155	350
Task 8	190	550
Task 9	225	700
Task 10	265	1,000

Table 3: The thresholds t_A by treatment condition.

4.2 Exponential growth bias prediction

The treatment comparison in Study 2 focuses on the effects of investment horizon variations. With a subjective constant distribution of absolute changes in $\{-18.9; 21.2\}$, the EGB decision maker perceives a perfectly symmetric selling price distribution with its median at 116.10 in treatment HVS and at 261.00 in treatment HVL. The EGB model thus predicts that participants overestimate the true median (88.64 in HVS and 29.96 in HVL). This effect is rather mild in HVS and much more pronounced in HVL.

Moreover, as discussed in Subsection 3.1, the EGB decision maker perceives no skewness in the distribution of the selling price of Security A. The true distribution Y_T is skewed, however, and the EGB model thus predicts a false assessment of the 10th and 90th percentiles. In particular, the EGB decision maker fails to realize that the right tail of the distribution is long, especially with a long investment horizon. We chose the experimental parameters such that under the EGB model this bias would have no discernible effect in treatment HVS, but predicts an effect in HVL. That is, the model predicts that the 90-10 spread is too narrow-band in treatment HVL.[19]

[18]The instructions explain that violations of these constraints are subjectively suboptimal and 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.

[19]The point predictions of the EGB model for the 10th and 90th percentile for treatment HVS are 35.90 and 196.30, respectively. The rational quantile assessments are only marginally different at 39.69 and 197.98. In treatment HVL, the EGB model predicts the 10th percentile at 0 and the 90th percentile at 581.80, with the corresponding rational values at 1.20 and 745.58.

4.3 Results

We start the analysis of the results with a descriptive overview. We then present interval regressions that impose a normal decision error and allow estimating the underlying quantiles while taking into account the discrete nature of the experimental data.

4.3.1 Descriptive overview

In treatment HVS (with the 14-period time horizon), at least half of the participants within each round strictly overestimate the median of the stochastic process, i.e., switch at least one task later than rational. Precisely, the shares of participants that reveal such a misperception of the median are: $\{$Round 1: 55%, Round 2: 65%, Round 3: 55%, Round 4: 69%, Round 5: $62\%\}$. However, only the share in Round 4 is significantly greater than 50% (p-value<0.05, one-sided binomial test), which is consistent with the EGB model's prediction of a mild median misperception for HVS.

As also predicted by the EGB model, the 90-10 spread perceptions in HVS, too, deviate only slightly from the rational prediction. This pattern appears in each of the five rounds. The respective shares of participants that underestimate the spread are as follows: $\{$Round 1: 65%, Round 2: 65%, Round 3: 72%, Round 4: 72%, Round 5: $65\%\}$. These shares are significantly greater than 50% (p-values<0.05, one-sided binomial test) in Round 3 and Round 4 only.

We get the same qualitative results for treatment HVL (with the 140-period investment horizon), but they are much stronger. This is further support for the EGB model. More than 75% of participants overestimate the median in each round of this treatment: $\{$Round 1: 93%, Round 2: 79%, Round 3: 79%, Round 4: 79%, Round 5: $79\%\}$ with all shares exceeding 50% significantly (p-values<0.001, one-sided binomial test). Moreover, for all but two rounds, these shares are significantly greater in HVL than in HVS (p-values<0.05, two-sample binomial test).

Finally, as the EGB model predicts, we observe that the perceived 90-10 spread in HVL is underestimated by almost all participants in almost all rounds: $\{$Round 1: 89%, Round 2: 96%, Round 3: 100%, Round 4: 96%, Round 5: $93\%\}$ (p-values<0.001, one-sided binomial test for >50%).[20] For each round these shares are significantly greater in HVL than in HVS (p-values<0.05, two-sample binomial test).

4.3.2 Interval regressions

The above descriptive analysis is based on the elicited lower bounds of the 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 regression

[20] EGB predictions and rational benchmarks for the 10th and 90th percentiles differ only for the latter in HVL. There, the shares of those underestimating the 90th percentile range between 72% and 96% over the rounds and are always significant, supporting the EGB model (p-values<0.05, one-sided binomial test for >50%). In HVS, where EGB and rational predictions coincide for the 90th percentiles, the respective shares are much lower and not significantly different from 50% (ranging from 31% to 58%).

(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, and standard hypothesis testing applies. Figures 2 and 3 report the corresponding estimates of the population means for the subjectively perceived quantiles of Security A's selling price distribution, separately for each of the two treatments and for each round. The horizontal dashed lines depict the benchmark rational predictions for the respective treatment-specific quantiles. The point estimates of the participants' average perceptions are enclosed by 95% confidence intervals.[21]

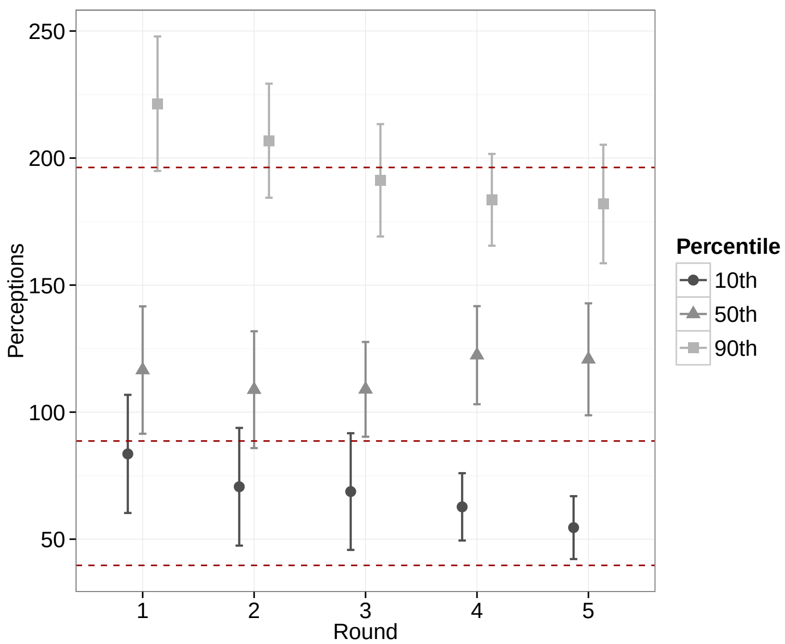


Figure 2: Point estimates of the participants' subjective quantiles of Security A's selling price distribution in HVS enclosed by 95% confidence intervals. Dashed lines indicate rational benchmarks for the 10th percentile (lowest), median (middle) and 90th percentile (uppermost).

Figure 2 confirms the previous subsection's descriptive analysis. In HVS, the round-wise 95% confidence intervals of the point estimates for the average median perceptions (triangle) are only marginally above the rational predictions (middle dashed line). In contrast to that and in accordance with the EGB model, the HVL median perception confidence intervals (Figure 3) differ significantly from the rational benchmarks for all five rounds.

Regarding the perceived skewness, the estimates also confirm the EGB model, as they indicate that on average the underlying subjective distributions are only marginally more symmetric than the rational prediction in HVS (Figure 2) but significantly more symmetric compared to the rational predictions in HVL (Figure 3). Once again, we observe that this pattern is robust over the rounds. The same

[21] For a detailed listing of interval regression estimates, see the online appendix.

is true for average perceptions of spread. The difference between the estimates for 10th percentile perceptions (circle) and 90th percentile perceptions (square) indicate subjective distributions that are much more narrow-band compared to the rational predictions (lowest and uppermost dashed lines) in HVL than in HVS.

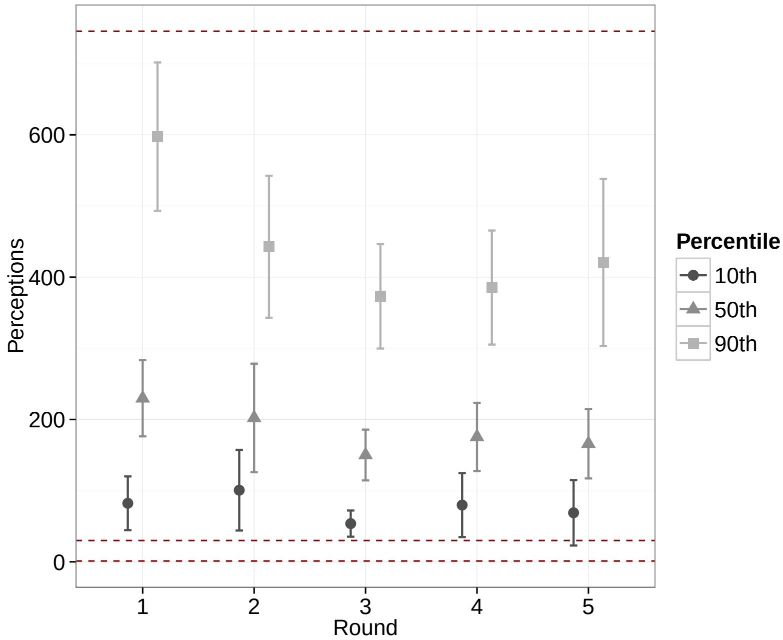


Figure 3: Point estimates of the participants' subjective quantiles of Security A's selling price distribution in HVL enclosed by 95% confidence intervals. Dashed lines indicate rational benchmarks for the 10th percentile (lowest), median (middle) and 90th percentile (uppermost).

5 Study 3: Exponential growth bias in the perception of exchange traded funds

In this study we test the robustness of the EGB model in a setting where the asset price depends on real-world 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 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, for example, changes by +3% on a trading day if the DAX30 increases by 1% on that day and it changes by -3% if the DAX30 falls by 1%. 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.[22] Our experiment confirms the prediction derived from the EGB model that skewness and spread are strongly underestimated if the volatility is high.

5.1 Experimental design

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 2,000 trading days both for ETF_1 and ETF_3. To generate realized price paths for the two assets, we sample 2,000 consecutive DAX30 closing values, drawn at random from the time period 1964 to 2012.[23] Participants can buy the ETF at a price of 100; if they buy it, they have to hold it for 2,000 trading days. As in Study 2, in both treatments, a participant who chooses Security A receives a fixed bonus if the selling price exceeds a given threshold t_A. These thresholds do not differ between ETF_3 and ETF_1. They are listed in Table 4.

	Thresholds for Security A in ETF_3 and ETF_1
Task 1	30
Task 2	60
Task 3	90
Task 4	140
Task 5	200
Task 6	260
Task 7	330
Task 8	450
Task 9	650
Task 10	1,000
Task 11	1,600

Table 4: The 11 thresholds.

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 90% probability,

[22] Regulatory 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..."

[23] The instructions in the ETF treatments are analogous to the treatments in Study 2. Additionally, 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.

B2 with 50% and B3 with 10%. Like in Study 2, each participant faces three choice lists and the choices allow us to infer bounds on the subjectively perceived quantiles.

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 2,000 consecutive trading days from the set of all available 2,000-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 (29 in ETF_3 and 30 in ETF_1), all of them undergraduate students at University College London. The incentivisation structure contains a $\pounds$ 5 participation fee and a possible bonus of $\pounds$ 5 for each round.

5.2 Exponential growth bias prediction

While Study 2 analysed the effects of investment horizon variations, the treatment comparison in Study 3 focuses on the effects of increasing the per-period volatility. We generate the perceived distribution of an EGB decision maker by means of simulations: we randomly sample 500 price paths of ETF_3 and ETF_1 as perceived by an EGB decision maker. By analogy to the previous discussion, the EGB decision maker is assumed to correctly perceive the distribution of daily changes in the relevant asset price, but views them as 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 at $Y_0 = 100$. For each of the ensuing 2,000 trading days $s > t$ we consider the relative change in value on that day, μ_s, and add $(\mu_s - 1)Y_0$ to the current perceived price of the asset: $\tilde{Y}_s = \tilde{Y}_{s-1} + Y_0(\mu_s - 1)$ (with $\tilde{Y}_0 = Y_0$). For example, suppose that $t =$ Jan 2, 1978 was randomly chosen as the starting date and that the DAX30 increased in value by 1.4% on $s =$ Feb 15, 1979. As an EGB decision maker's perceived absolute increase on the latter date, the simulation simply adds $Y_0(\mu_s - 1) = 1.4$ to the price of the asset in the ETF_1 condition (and 3×1.4 in the ETF_3 condition). The simulation thus arrives at a perceived selling price at maturity. Repeating this procedure 500 times for random starting dates generates the perceived distribution of selling prices.

Consistent with the results of previous sections, the simulations generate the predictions that an increased per-period volatility leads to (i) a larger overestimation of the median return, and (ii) a larger underestimation of the outcome distribution's skew and spread.

5.3 Results

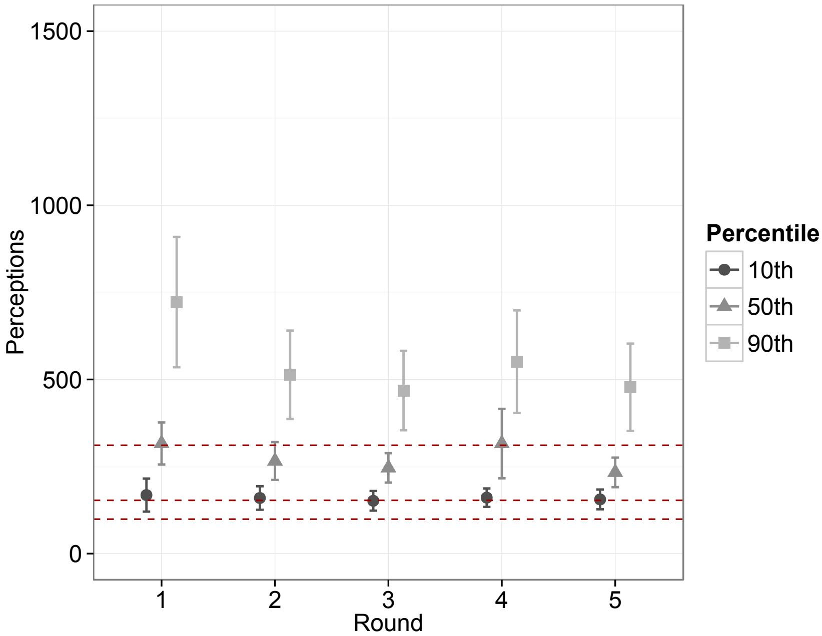


Figure 4: Point estimates of the participants' subjective quantiles of Security A's selling price distribution in ETF_1 enclosed by 95% confidence intervals. Dashed lines indicate rational benchmarks for the 10th percentile (lowest), median (middle) and 90th percentile (uppermost).

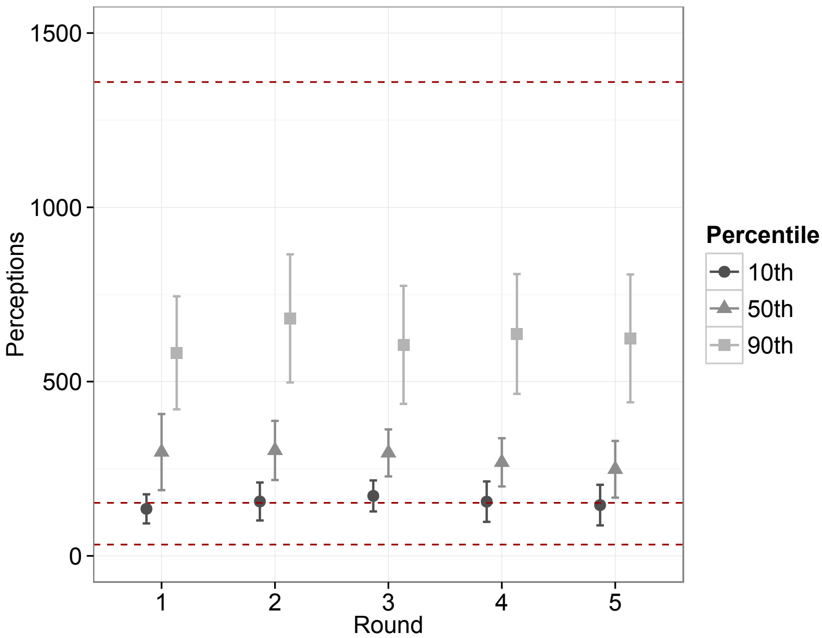


Figure 5: Point estimates of the participants' subjective quantiles of Security A's selling price distribution in ETF_3 enclosed by 95% confidence intervals. Dashed lines indicate rational benchmarks for the 10th percentile (lowest), median (middle) and 90th percentile (uppermost).

Figure 4 illustrates interval regression estimates of the participants' average perceptions in ETF_1. It shows that for all five rounds average median perceptions (triangle) are significantly above the rational level (middle dashed line). Participants in the experiment are overly optimistic about the simple index ETF. The round-wise arrangement of the three average percentile perception estimates 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 5 captures participants' average perceptions in treatment ETF_3. Again, perceived medians show on average a notable level of overestimation. Moreover, supporting the EGB model, the underlying distributions of subjective percentiles within each round also indicate that the perceived spread and the perceived skewness are on average too small compared to the rational benchmarks. The participants show at least a mild tendency to report skewed distributions but they far underappreciate the actual level of skewness. This misperception is much stronger in ETF_3 than it is in its low-volatility version ETF_1.

6 Conclusion

This paper investigates how people perceive two important implications of compounding random growth. First, as has been established in the literature, decision makers have a tendency to neglect nonlinear growth. Our experiments add to the evidence of this effect by measuring it in a context with random growth rates. Second, people underestimate the level of asymmetry in growth processes—skewness is "hidden". This is a novel effect in the academic literature (modulo the independent description in Stutzer and Grant (2013)) and may be especially relevant in the context of preferences that consider quantiles of the outcome distribution, like value-at-risk or expected utility with convex utility functions. However, it is important to note that this paper is about misperception, not preferences, and that we measure the effects irrespective 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. One may speculate that if decision makers were trained in the use of log returns, the bias could be reduced. At least some of our experiments give the respondents a very good shot at detecting the nonlinear effects of growth, since we use highly selected and quantitatively skilled students and partially provide them with calculators. It is perhaps all the more notable that we, too, find a strong bias towards linear perceptions.

In Study 3 we also indicate that the effect of skewness is relevant beyond abstract settings. We chose a setting where leveraged ETFs are on offer, as these assets were primarily bought and held by household investors who do not usually have sophisticated risk management tools available. Just as many of these household investors were surprised by the losses they incurred, our participants show a misperception of the effects of leverage, in line with exponential growth bias.

The results may well extend to other contexts of household finance. As we describe in the Introduction, compounding of stochastic growth is often required in contexts of retirement savings. It will be difficult to quantify the effects of the many established anomalies in savings behavior. But progress is made step by step. Our stylized experiments, with full control over information flow and incentives, can at least establish that the expectation of compound distributions deviates predictably from the rational benchmark. Our experiments also show that long investment horizons increase the strength of the misperception.

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] Brünner, 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., Weizsäcker, 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] Gneezy, U., and Potters, J. (1997). *An Experiment on Risk Taking and Evaluation Periods.* Quarterly Journal of Economics, 112(2): 631-645.

[16] Kemp, S. (1984). *Perception of changes in the cost of living.* Journal of Economic Psychology, 5(4): 313-323.

[17] 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.

[18] Levy, M., and Tasoff, J. (2015). *Exponential-growth bias and lifecycle consumption.* Journal of the European Economic Association, doi: 10.1111/jeea.12149.

[19] Lusardi, A., and Mitchell, O.S. (2011). *Financial literacy and planning: Implications for retirement wellbeing.* NBER Working Paper 17078.

[20] McKenzie, C.R.M., and Liersch, M.J. (2011). *Misunderstanding savings growth: Implications for retirement savings behavior.* Journal of Marketing Research, 48: 1-13.

[21] Redelmeier, D.A., and Tversky, A. (1992). *On the framing of multiple prospects.* Psychological Science, 3(3): 191-93.

[22] Samuelson, P. A. (1963). *Risk and uncertainty: A fallacy of large numbers.* Scientia, 98: 108-113.

[23] Stango, V., and Zinman, J. (2009). *Exponential growth bias and household finance.* Journal of Finance, 64(6): 2807-2849.

[24] Stutzer, M., and Grant, S. (2010). *Misperceptions of long-term investment performance: Insights from an experiment.* Journal of Behavioral Finance and Economics, 3(1): 1-20.

[25] Wagenaar, W.A., and Sagaria, S.D. (1975). *Misperception of exponential growth.* Perception & Psychophysics, 18(6): 416-422.

[26] 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.

[27] 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.

[28] Wooldridge, J.M. (2002). *Econometric analysis of cross section and panel data.* Cambridge, MA: MIT Press.

A Details on additional experimental evidence

A.1 Stylized experiment in the Introduction

A.1.1 Experimental design and exponential growth bias prediction

The experimental asset of interest, Security A, follows a binomial-tree -60%/+70% process over 12 periods that is identical to Security A in Study 1(b), with an initial price of 10,000. In three experimental sessions, 69 students at Technical University Berlin 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. Just as in Study 1(a), we ensure incentive compatibility under a wide set of preferences by using only two possible payments per choice problem—receive a bonus of €20 versus not.
Participants are told that Investment 1 “makes a gain” (in effect pays the bonus, see below) iff the selling price of Security A is between 0 and 6,400. Investment 2 makes a gain iff the selling price of Security A is between 6,400 and 12,800, Investment 3 makes a gain iff the selling price is between 12,800 and 19,200, Investment 4 makes a gain iff the selling price is between 19,200 and 25,600, and Investment 5 makes a gain iff the selling price is above 25,600.
Once a participant has chosen her investment the computer simulates the period-12 price of Security A, with a separate simulation for each participant. If the chosen investment makes a gain, the participant receives the bonus. Under any belief the decision maker should, evidently, choose the interval that has the largest probability of containing the period-12 price. Due to the price distribution’s large skew, the rational prediction is to choose Investment 1, whose corresponding interval [0 ; 6,400) contains the period-12 price with 80% chance. The other intervals thus have a far lower success chance and the monetary incentive for a participant to choose one of them is far lower. An EGB decision maker, however, 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 EGB decision maker therefore chooses Investment 3.

A.1.2 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-value<0.001, chi-square test) and indicates no tendency to choose a mode near zero. Overall, with 94% of the participants significantly more than half of the sample overestimate the true mode (p-value<0.001, one-sided binomial test). While only 6% make the optimal choice of Investment 1, 43% conform with the EGB 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 in expectation, the observed choice distribution on average earns only €2.07 in expectation per participant.

A.2 Treatment variation HELP of Study 1(b)

A.2.1 Treatment description and exponential growth bias prediction

Participants in Study 1(b) are randomly assigned to one of two treatments. Treatment NO_HELP (N=68), which is described in the main text above, presents only the basic explanation. In treatment HELP (N=60) 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.

Thus, in treatment HELP the EGB cannot influence the subjective beliefs without contradicting the available explanations. We therefore expect the misperception to disappear, i.e., $q_{0.5,i}^{HELP} = 500$.

A.2.2 Results

	Share of participants switching from A to B				
Range of subjective median for Security A	Round 1	Round 2	Round 3	Round 4	Round 5
$[0 \; ; \; 100)$	0.000	0.032	0.047	0.046	0.092
$[100 \; ; \; 500)$	0.000	0.016	0.000	0.046	0.046
$[500 \; ; \; 2,000)$	0.703	0.612	0.666	0.676	0.661
$[2,000 \; ; \; 6,000)$	0.109	0.145	0.095	0.138	0.046
$[6,000 \; ; \; 9,000)$	0.063	0.048	0.063	0.046	0.061
$[9,000 \; ; \; 12,000)$	0.063	0.064	0.063	0.000	0.030
$[12,000 \; ; \; 20,000)$	0.031	0.064	0.031	0.462	0.046
$[20,000 \; ; \; 35,000)$	0.016	0.000	0.015	0.000	0.015
$[35,000 \; ; \; 90,000)$	0.000	0.000	0.015	0.000	0.000
$[90,000 \; ; \; 250,000)$	0.000	0.000	0.000	0.000	0.000
$[250,000 \; ; \; \infty)$	0.016	0.016	0.000	0.000	0.000

Table 5: Subjective medians in HELP for rounds 1-5.

Table 5 shows that in HELP 70% of responses are at the optimal switching point of Task 3 already in Round 1. Parametric t-tests as well as non-parametric Wilcoxon rank-sum tests confirm that all round-by-round treatment comparisons between HELP and NO_HELP are statistically significant at p-values<0.001. Performance is poor under the NO_HELP condition and much better in HELP.