Scale matters: Risk perception, return expectations, and investment propensity under different scalings*

Christoph Huber† Jürgen Huber‡

August 24, 2018

Abstract

With a novel experimental design we investigate whether risk perception, return expectations, and investment propensity are influenced by the scale of the vertical axis in charts. We explore this for two presentation formats, namely return charts and price charts, where we depict low- and high-volatility assets with distinct trends. We find that varying the scale strongly affects people's risk perception, as a narrower scale of the vertical axis leads to significantly higher perceived riskiness of an asset even if the underlying volatility is the same. Furthermore, past returns predict future return expectations almost perfectly. In our setting perceived profitability was considered more important than perceived riskiness when making investment choices. Overall we show that adapting the scale of a chart makes it easier to recognize yearly return variations *within* a single security, but at the same time makes it harder to identify differences *between* dissimilar securities. This is something regulators should be aware of and take into account in the rules they set.

JEL: D14, D18, G11, G41

Keywords: behavioral finance, judgment, risk perception, scaling, presentation format

*We thank Felix Holzmeister, Stefan Palan, Julia Rose, and Janette Walde, two anonymous referees and the editor, seminar participants at the University of Linz, conference participants at ESA 2017 in Vienna and Experimental Finance 2018 in Heidelberg, as well as workshop participants at the University of Innsbruck for helpful comments and suggestions. Financial support from the Austrian Science Fund FWF (P29362-G27 and SFB-grant P4340-015-019, both J. Huber) and the University of Innsbruck (Doktoratsstipendium Netzwerk Banking, Accounting, Auditing, Finance & IT (BAFIT) C. Huber, Förderstipendium C. Huber) is gratefully acknowledged. A prior version of this paper was presented with the title "Scaling, Volatility, and Risk Perception."

†University of Innsbruck, Department of Banking and Finance, Universitätsstrasse 15, 6020 Innsbruck, E-mail: christoph.huber@uibk.ac.at

‡University of Innsbruck, Department of Banking and Finance, Universitätsstrasse 15, 6020 Innsbruck, E-mail: juergen.huber@uibk.ac.at

1 Introduction

When Gulliver traveled to Lilliput he was a giant. On his next journey to Brobdingnag he was a dwarf. While he had not changed, the scale of everything around him had. It seems that the scale we see something in plays a major role in how we perceive it. In financial practice, the scaling of price and return charts, e.g. in documents given to customers, is an important issue – recognized by practitioners, but mostly ignored by regulators and research so far. We mention regulators as, for example, the European Union sets rules for the presentation of a security's past performance in a Key Investor Information Document (KIID; see Commission Regulation (EU) No 583/2010). According to that regulation, returns have to be shown in the form of bar graphs with a linear vertical axis. Additionally, the scale has to be adapted appropriately and *shall not compress the bars so as to make fluctuations in returns harder to distinguish* (p. 15). While the European Commission sees the potential problems of highly compressed bars, it remains unclear what consequences arise regarding the risk and profit expectations to-be-identified by investors, and hence, regarding investment decisions. Maximizing the return bars on the available space makes yearly fluctuations more distinguishable, but also involves the danger of misinterpretation of the returns as highly volatile and therefore highly risky, even when they are not. Compressing the bars, however, could lead to risk being perceived as too low, possibly exposing consumers to unexpectedly high losses.

As individuals focus on graphical and salient pieces of information in their information processing strategies (Jarvenpaa, 1989), there is a wide range of literature on graphical representations of financial time series. One strand of research tackles the question of which presentation formats (e.g. returns, prices, or distributions) increase potential investors' forecasting abilities and accuracy. Return charts are associated with lower expected returns (Glaser et al., 2018) but also with higher perceived uncertainty (Diacon and Hasseldine, 2007), compared to price charts. Stössel and Meier (2015) also discuss framing effects of different presentation formats on risk perception, but restrict themselves to different forms of graphical representations in the narrow domain of the KIID.[1] While Weber et al. (2005) find no significant improvement in perceived risk with continuous density distributions, Kaufmann et al. (2013) and Ehm et al. (2014) develop possibilities to better calibrate people's risk perception by experience sampling from return distributions. However, these efforts require and imply a known stochastic process underlying the financial instrument to be assessed. In real-world applications, however, we have to rely on historical data, which may or may not give a good estimate for future returns and volatilities.

A number of studies has investigated graphical distortions in information processing, most notably regarding corporate reports (see e.g. Beattie and Jones, 1992). They show that

[1] In developing the standardized KIID the European Commission employed a survey-based study by IFF Research and YouGov (2009). They mention different axis scales as a possible explanation for people's difficulties in assessing past performance but do not attempt to offer any evidence confirming this intuition.

a disproportionate representation of the underlying data can be misleading (Tufte, 1983) and is often purposely used to create a more favorable view.[2] However, only few studies have investigated potential effects of varying a graph's vertical axis scale without violating proportionality principles:[3] Cleveland et al. (1988) examine the 'shape parameter' of graphs – that is, the ratio of the horizontal and vertical distances spanned by the data, while holding the scale's range constant. Lawrence and O'Connor (1992, 1993) examine scale effects with regard to people's forecasting ability in financial time series. They find that large scales or high variability in the presented time series leads to overly narrow confidence intervals. To our knowledge, however, the vertical axis scale's relevance towards risk communication and investment decisions has not yet been investigated.

Our aim with this paper is to fill this gap by providing a systematic and rich analysis of the scale effect in graphical representations of financial time series. The research question we address is whether the presentation scale – narrow or wide – affects people's risk perception, return expectations, and propensity to invest. We define a chart as having a *narrow* scale when the time series depicted extends close to the upper or lower borders of the chart, while a *wide* scale leaves ample space above and below.

To explore our research question we conduct a laboratory experiment with a 2×2 design where we vary the presentation scale (narrow or wide) and the presentation format: assets are presented either as return bar charts or as price line charts. In a within-subjects design we ask participants to assess the riskiness, expected return, and attractiveness as investment of the assets. In a second task subjects make pairwise comparisons between these assets along the same three dimensions.

We find that varying the scale strongly affects people's risk perception, namely, that a narrower scale of the vertical axis leads to significantly higher perceived riskiness of an asset across price and return charts, even if the underlying volatility is the same. We demonstrate that adapting the scale to the span of the bars is reasonable with regard to recognizing yearly return variations *within* a single security, but at the same time makes it harder to identify differences *between* dissimilar securities. This result is robust for different historical return trends. We further find that past returns predict future return expectations almost perfectly irrespective of the scale. Risk perception is highly correlated with losses which in turn drive investment behavior. Concerning investment choices, subjects tend to invest in the asset they regard as more profitable, even if they think it bears higher risk.

This study extends the existing literature in several important ways: We analyze previously unexplored scale effects in a systematic and clean experimental setup; we embed these issues directly into the context of information presentation in financial markets; and we explore

[2]See Beattie and Jones (2008) for a survey of corporate reporting using graphs.

[3]Violations of proportionality in representing the underlying data include, for example, omitting the zero line and using non-linear scales. See Tufte (1983) for a comprehensive exposition of visual data representations.

different aspects of financial decision-making relating to the scale, presentation format, and underlying asset fundamentals in individual assessments as well as in pairwise comparisons.

We think our findings are also informative for regulators: As we show, adapting the scale of a chart makes it easier to recognize yearly return variations *within* a single security, but at the same time makes it harder to identify differences *between* dissimilar securities. Regulators should be aware of – and attentive to – the potentially distorting effects of different axis scales in performance charts. While return bar charts are appropriate, allowing issuers to adapt the axis scale arbitrarily leaves room for deliberate action aimed at distorting investors' perceptions about risk. Keeping the presentation scale constant across different securities enables better identification of risk and therefore easier comparisons.

2 The Experiment

2.1 Returns and Prices

To systematically vary expected return, time trend, and volatility of percentage return time series we create eight distinct return paths consisting of ten (hypothetical) annual returns each. Each return path consists of a deterministic trend (POSITIVE STABLE, NEGATIVE STABLE, INCREASING, or DECREASING) and a normally distributed noise term $\varepsilon_t \sim N(\mu, \sigma^2)$ with $\mu = 0.0\%$, $\sigma^2 = 1.4\%$ and $t = 1, ...10$. Low-volatility assets consist of a linear return path plus the noise term for each year t. High-volatility assets have the same linear return paths but with the noise term multiplied by 6 before it is added.

Fig. 1 shows each distinct return trend as a function of time, depicted as RETURN charts (left) and PRICE charts (right). Assets with a POSITIVE STABLE trend are set up to yield positive returns fluctuating around a mean of 3% per year. The return path INCREASING starts at -3% in the first year and linearly increases to $+3\%$ in the tenth year plus a noise term. Assets with trend NEGATIVE STABLE contain the returns of the asset with trend POSITIVE STABLE multiplied by -1. Analogously returns in trends DECREASING are the returns of trend INCREASING multiplied by -1. Price paths are generated by successively applying the corresponding returns to an initial price of 100.

2.2 Experimental Tasks

In the experiment subjects have to complete two main tasks, Task I and Task II. In both tasks participants were instructed to suppose that they want to invest 5,000 euros. Subjects are then presented with charts of hypothetical assets and are asked to assess the respective riskiness and profitability of one asset at a time in Task I, and to compare two assets at a

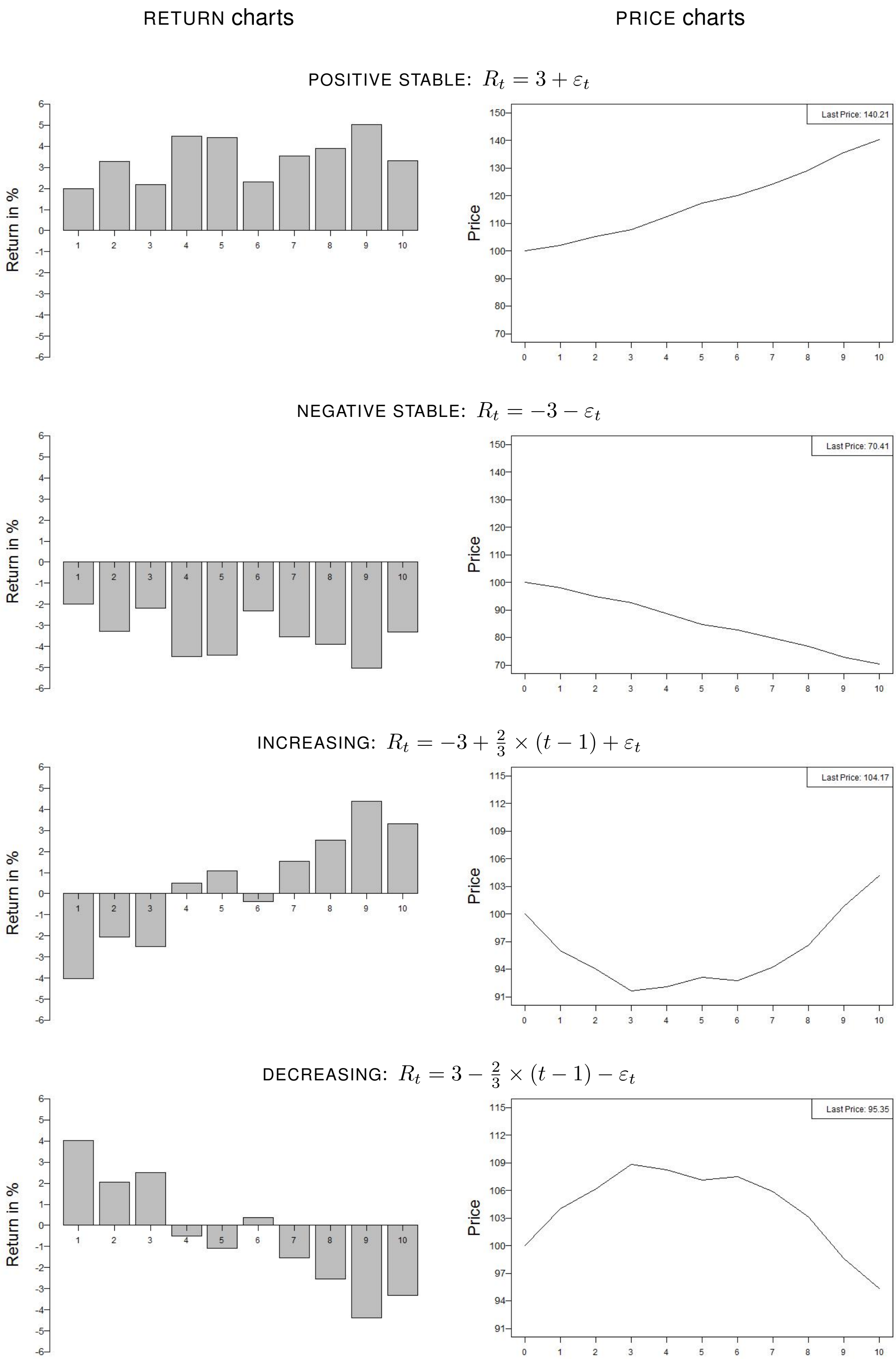

Figure 1: **Return and price paths.** This figure shows the four distinct return trends with the low volatility level as a function of time, depicted as return bar charts (left) and price line charts (right). For high-volatility assets, the added error term ε_t is multiplied by six.

time along these dimensions in Task II. There are two variants for each task: either a return bar representation (RETURN) or a line chart depicting the price development (PRICE).

Task I consists of a 2×2 treatment design in which we vary the presentation format (RETURN or PRICE) and the presentation scale of the vertical axis (NARROW or WIDE) to identify these variables' effect on risk perception as well as on return expectations and investment propensity. Subjects sequentially see eight different paths (RETURN or PRICE) and have to assess the assets' riskiness and estimate its returns over the following year and over the next five years. Whenever subjects are presented with return charts they are explicitly asked about future returns; when they see price charts they are asked to estimate future prices.[4] Each participant was presented with eight out of 16 possible return (price) charts (eight different assets in two different presentation formats, NARROW and WIDE) in which each chart has the same probability of appearing. The order in which participants saw the assets was randomized and participants were not aware of being presented only with a selection of the possible assets.

In Task II subjects make pairwise comparisons between assets regarding their riskiness and expected return. In a 2×2 design the combinations of volatility and scale of the vertical axis are varied in four distinct conditions: same scale/same volatility, same scale/different volatility, different scale/same volatility, and different scale/different volatility. Except for Condition SAME (same scale and same volatility), we name each condition after the variable in which the two charts of a pair *differ*. With this set-up we are able to generate a distinct number of 16 pairs for Condition SAME,[5] four pairs for Condition VOLATILITY, and eight pairs each for conditions SCALE and BOTH. Subjects are presented with a total of eight randomly chosen pairs – two for each condition. In this task subjects have to compare two assets at a time. They are asked to decide which of the two assets they perceive as riskier; which asset they think is more profitable; and which asset they would rather invest in. For each question there is also the possibility to choose the neutral option 'the same for both' (later also referred to as 'indifferent').

In total, 32 charts have been considered: 4 trends x 2 volatilities (high/low) x 2 formats (return/price) x 2 scales (wide/narrow) = 32; 16 price charts and 16 return charts. For both tasks there are two variants: In Tasks Ia and IIa subjects are presented with return charts, in Tasks Ib and IIb they see price charts. In Task I, each subject considers eight random return charts (Ia) and eight random price charts (Ib), and in Task II each subject considers eight random return chart comparisons (IIa) and eight random price chart comparisons (IIb).

[4] Glaser et al. (2007) and Glaser et al. (2018) discuss the impact of presenting returns vs. presenting prices and asking for returns vs. asking for prices. We consciously refrain from presenting one format and asking for the other, as the tasks we have seem demanding enough for subjects and we want to rule out potential confusion.

[5] To avoid subjects being asked to compare two identical asset representations in setting SAME (same scale and same volatility), we permute the returns of an asset in a way that preserves the asset's characteristics but generates a marginally different path for comparison. One example can be seen in the first pair shown in Fig. E2 in Online Appendix E.

Hence, each subject sees a potentially different selection of charts. The order is randomly determined and subjects are randomly assigned to one of two groups to eliminate any order effects (see Fig. 3 for the timeline and the two possible sequences).

In both presentation formats we vary the scale of the vertical axis to create a NARROW and a WIDE representation of each asset's past performance. In return charts, the maximum value on the vertical axis and the tick size of WIDE-scaled representations are three times the corresponding values of representations with scale NARROW. For price charts the scales are adapted analogously. Fig. 2 shows an example of RETURN charts (top) and PRICE charts (bottom), each with presentation scales NARROW (left) and WIDE (right). In each return (price) chart the value zero (100) as well as each tick (with precise values depending on asset and scale) is at exactly the same position in the graph to maintain consistency. Additionally, in order to reduce noise in estimating prices we provide the last price in the upper right corner of price charts (Glaser et al., 2018). In the instructions, we explicitly point out that the scale of the vertical axis might change over the course of the experiment. This note also appears when subjects review the on-screen instructions at any point in time during an experimental task. This prominently-placed reminder should ensure that our results are not driven by subjects' inattention to the scale. To guarantee subjects' understanding of the term 'return' we also include a definition stating that the return is defined as the percentage change of the price over one year. Table 1 summarizes the asset- and chart-specific variables and respective options: each chart is a distinct combination of volatility, trend, presentation format, and scale.

Table 1: **Summary of variables in each performance chart.** This table summarizes the relevant variables in specific to assets and charts: the volatility and trend of an asset, and the presentation format and scale of a chart.

	Variable	Possible Options
Asset specific	Volatility	LOW or HIGH
	Trend	POSITIVE STABLE, NEGATIVE STABLE, INCREASING, or DECREASING
Chart specific	Presentation format	RETURN or PRICE
	Scale (vertical axis)	WIDE or NARROW

2.3 Implementation of the Experiment

We conducted nine experimental sessions with a total of 193 students of business administration or economics in May and June 2017 at the Innsbruck EconLab at the University of Innsbruck. The experiment was programmed and conducted using oTree by Chen et al. (2016). Subjects were recruited with hroot by Bock et al. (2014). 45% of subjects were female;

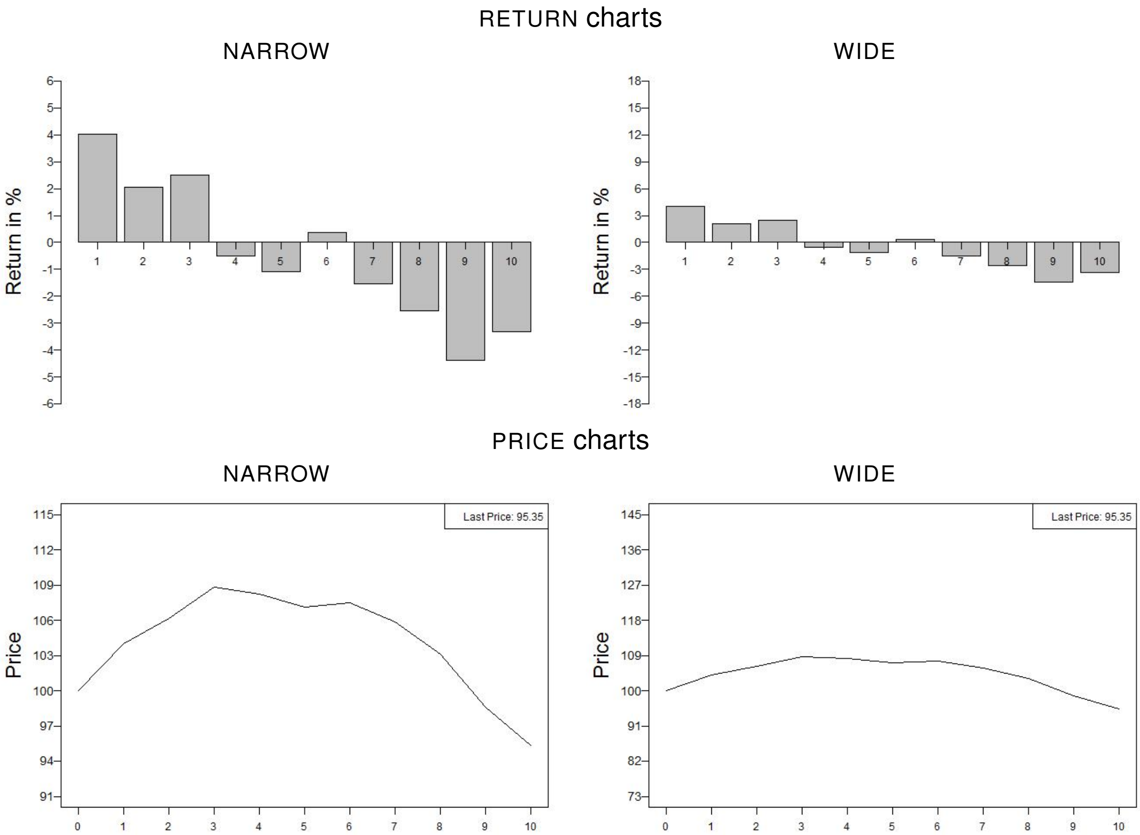

Figure 2: **Exemplary representations of the low-volatility asset with trend decreasing in a return chart (top) and a price chart (bottom) for presentation scales narrow (left) and wide (right).** For return charts the value zero and for price charts the initial price of 100 as well as each tick are at the same positions for both scales. In return bar representations the tick size on a WIDE scale is three times the one on a NARROW scale; tick sizes in price representations are adjusted accordingly.

the mean age was 23; and about 51% of subjects had completed an undergraduate course in financial management.

In total, each session lasted approximately 40 minutes. This included studying on-screen instructions for each part of the experiment as well as a multiple price list task measuring subjects' risk attitudes (Holt and Laury, 2002) and a certainty equivalence task to assess loss aversion (Gächter et al., 2007) using oTree applications by Holzmeister (2017). After the main experiment subjects completed a questionnaire assessing their risk attitudes and demographics. A graphical overview of the experimental procedure, as well as the experimental instructions, screenshots of the decision tasks, and exemplary charts for each condition are provided in Online Appendices C, D, and E.

Subjects are incentivized by being paid one randomly chosen return of the asset they chose to rather invest in in one randomly chosen pair they were presented with for both parts (prices and returns) of Task II.[6] The chosen return times two is added to an initial amount of 5 euros for each task. For example, if the chosen asset of the randomly drawn pair pays 10% in the randomly drawn year, the participant receives 5 euros $\times(1+2\times 0.10) = 6$ euros for this task. Total payouts varied between 6.30 euros and 16.30 euros with a mean of 11.65 euros; these include payouts from the risk and loss aversion tasks.

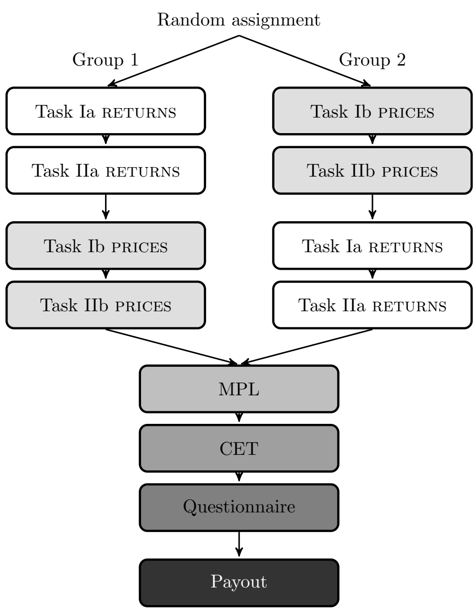

Figure 3: **Graphical overview of the experimental procedure.** Subjects are randomly assigned into two groups with Group 1 being presented with RETURN charts first (Tasks Ia and IIa) and PRICE charts second (Tasks Ib and IIb) and Group 2 vice versa. Both groups complete a multiple price list task (MPL) and a certainty equivalence task (CET) to elicit risk and loss aversion parameters, as well as a questionnaire after the experiment.

[6] The return R of year $t \in [1, 10]$ of the chosen asset of pair $p \in [1, 8]$ in presentation format $F \in \{\text{RETURN}, \text{PRICE}\}$; if subjects are indifferent between two assets, either A or B is chosen randomly. For each of the two tasks they receive 5 euros$\times(1 + 2 \times R^F_{t,p})$.

3 Results from Task I: Individual assessments

We organize the presentation of results as follows: first we analyze subjects' individual assessments (Task I), starting with perceived risk, followed by expected returns and investment propensity. Subsequently, the same structure is repeated for the analysis of pairwise comparisons (Task II).

3.1 Risk perception in individual assessments

We start our discussion by examining the influence of the scaling of the vertical axis on risk perception. We present analyses along the following dimensions: for both presentation formats (RETURN and PRICE charts) we compare the influence of scaling of the vertical axis (WIDE vs. NARROW). To get a comprehensive picture we do this separately for the four different return trends (POSITIVE STABLE, NEGATIVE STABLE, INCREASING, and DECREASING), where we have each return trend once with a low and once with high level of return volatility (LOW or HIGH).

Fig. 4 shows the differences in average perceived risk (elicited on a scale from 1 to 7) for each asset from RETURN charts (left panel) and PRICE charts (right panel). The differences are from the same asset being displayed once with a WIDE and once with a NARROW scale.[7] The four bars in each group of bars represent the four different trends; LOW volatility is shown in the left group of each panel while HIGH volatility is shown in the right group of each panel.

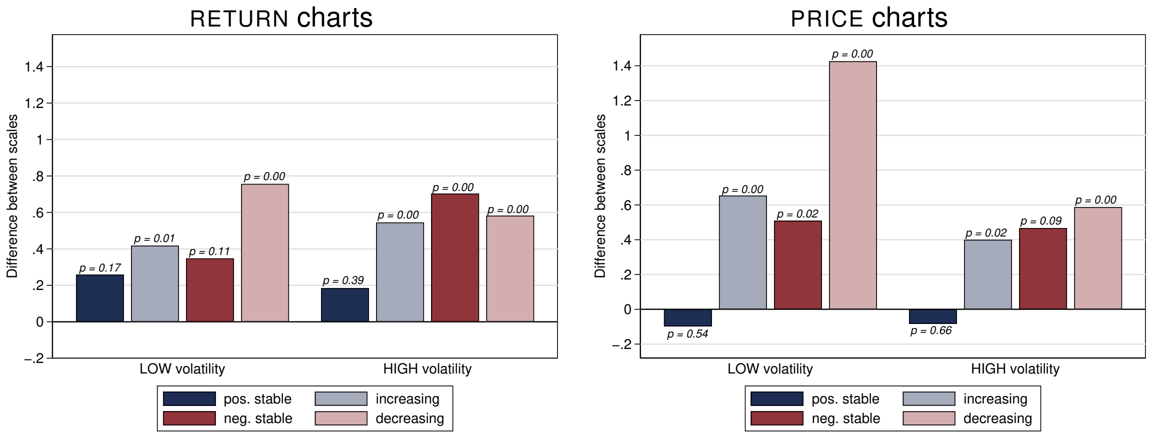

Figure 4: **Differences in average perceived risk (in NARROW minus WIDE) by trend and scale presented as return charts (left) and price charts (right).** This figure depicts differences in average perceived risk (on a scale from 1 = "not risky at all" to 7 = "very risky") for RETURN chart and PRICE chart representations of LOW (left bars in each panel) and HIGH (right bars in each panel) volatility assets. p-values above the bars are from Fisher-Pitman permutation tests on the subject-demeaned data. Each of the sixteen bars summarizes between 179 and 206 observations.

[7] As the probability of appearance of one particular chart is determined randomly for each subject, the number of observed decisions for each distinct combination of presentation format, volatility, and trend varies between 179 and 206.

Result 1. *In individual assessments assets are perceived as riskier when presented on a* NARROW *scale than when presented on a* WIDE *scale.*

Support: For all assets, except those with a POSITIVE STABLE trend, a NARROW scale leads to higher perceived risk compared to a WIDE scale. This holds for both presentation formats and both volatility levels. To test the statistical significance of the differences between charts with NARROW and WIDE axis scales, we run Fisher-Pitman permutation tests on the subject-demeaned data.[8] 10 out of 12 tests for assets other than with a POSITIVE STABLE trend deliver p-values of 0.02 or smaller (the remaining two having p-values of 0.09 and 0.11, respectively), corroborating that assets are perceived riskier when presented on a NARROW scale – we conjecture that this is the case because fluctuations are easier visible on a NARROW scale. As the POSITIVE STABLE trends show the lowest overall risk perceptions we conjecture that the non-difference in their perceived riskiness results from the fact that these always-positive returns (monotonically increasing prices, respectively) are never perceived as risky, no matter how they are displayed or whether they fluctuate more.

As we specifically make subjects aware of varying axis scales in the instructions, and as we find no relationship between the differences in risk perception and the time it took subjects to complete all related tasks,[9] we attribute the reported differences in perceived risk to the differences in axis scales, as the differences seem not to stem from subjects being unaware of the varying axes or inattentively clicking through the tasks.

Regarding different return paths we observe that the POSITIVE STABLE trend is seen as the least risky one with average assessments between 2.08 and 3.51 on a 7-point scale, whereas NEGATIVE STABLE and DECREASING trends are viewed as carrying the highest risk (average riskiness assessments between 4.52 and 5.94) with trend INCREASING being in-between across all presentation formats (see Online Appendix A).[10] This implies that an asset's volatility (standard deviation of returns) does not necessarily determine people's perceptions about its risk: e.g., for trend NEGATIVE STABLE we find no difference in perceived risk between the LOW and HIGH-volatility assets in return and price charts ($p = 0.17$ and $p = 0.40$) – even though their volatility differs by a factor of six. Furthermore, the standard deviation is the same for NEGATIVE STABLE and POSITIVE STABLE, but they are at opposite ends regarding perceived riskiness. This shows that subjects perceive losses (negative returns) as *risk*, while profits are perceived as not risky, even if they vary as much as losses do.[11]

[8]To establish independence between observations we subtract the mean across all of a subject's risk assessments from each data point. We then run non-parametric Fisher-Pitman permutation tests on these subject-demeaned data with 300,000 simulations as a more powerful alternative to Wilcoxon-Mann-Whitney tests (Kaiser, 2007). If not stated otherwise, we use the same procedure throughout the paper. Wilcoxon sign-rank tests on the matched pairs of the original observations yield very similar results, see Table A2 in Online Appendix A.

[9]Pearson's correlation coefficients ρ are between –0.10 and 0.14 across presentation formats and volatilities.

[10]$p < 0.01$ for 22 out of 24 pairwise comparisons between trends; for details see Online Appendix A.

[11]This connects nicely to recent literature, e.g. by Anzoni and Zeisberger (2016) and Huber et al. (2018) showing that risk perception is mostly driven by losses. For further details see Online Appendix B.

Additionally, one remarkable side result with potentially important implications for practitioners and regulators is that people perceive risk as significantly higher when presented with RETURN charts as compared to PRICE charts (five out of eight p-values are significant at $p < 0.01$; all differences have the same sign; details are provided in Online Appendix B).

3.2 Expected returns in individual assessments

Besides eliciting subjects' perceptions about risk we also asked participants to enter point estimates of future returns (when RETURN charts were shown) or future prices (when PRICE charts were shown) for a shorter (one year) and a longer (five year) horizon.[12] We discuss short-term forecasts first. The upper panels of Fig. 5 depict the median one-year-ahead return expectations for each asset (vertical axis) in relation to the last return (horizontal axis) in both presentation formats for scales WIDE and NARROW.

Result 2. *Expected returns are driven by the latest return. We find no systematic influence of the scale on return expectations in individual assessments.*

Support: Short-term return forecasts are not the same across assets, but strongly depend on the last return. Subjects thus seem to behave as short-term trend-followers. With an R^2 of 0.97 and a slope of 0.76 the past return almost perfectly explains return predictions when RETURNS are shown (upper-left panel of Fig. 5). When PRICES are shown (right panel) there is more dispersion, especially when the last return is negative. Still, with a slope of 0.98 and a R^2 of 0.85 the last return is a very good predictor of expected returns. This is consistent with Grosshans and Zeisberger (2018), who analyze forecasts for price paths similar to the INCREASING and DECREASING trends in the present study, as they report strong beliefs in short-term trend continuations. In both presentation formats we do not find a systematic influence of the scale (NARROW or WIDE).

We also asked subjects for their five-year return prediction (return per year); respectively price-prediction (price in five years). For RETURNS we find a very similar and consistent pattern to the one-year-predictions where the last return is again a very good predictor with a slope of 0.82 and an R^2 of 0.90 (see lower panels of Fig. 5). Returns calculated from PRICE predictions also show a strong positive relation between last return and expected return. However, with a slope of 0.42 and an R^2 of 0.50 the relation is markedly flatter and weaker than for the one-year price data or the RETURN data. In contrast to Glaser et al. (2007) and Glaser et al. (2018) we find that even for five-year-ahead forecasts, on average participants do not expect trend reversals to the extent of a change in signs – we find that the slope calculated from PRICES is only half as steep as for the one-year forecasts.

[12] As mentioned above, we consciously refrain from presenting one format and asking for the other – i.e., we ask for returns when presenting returns and ask for prices when presenting prices. For the analysis we only consider returns, either directly from subjects' return estimates or calculated as the average annual difference between price estimates and the latest price.

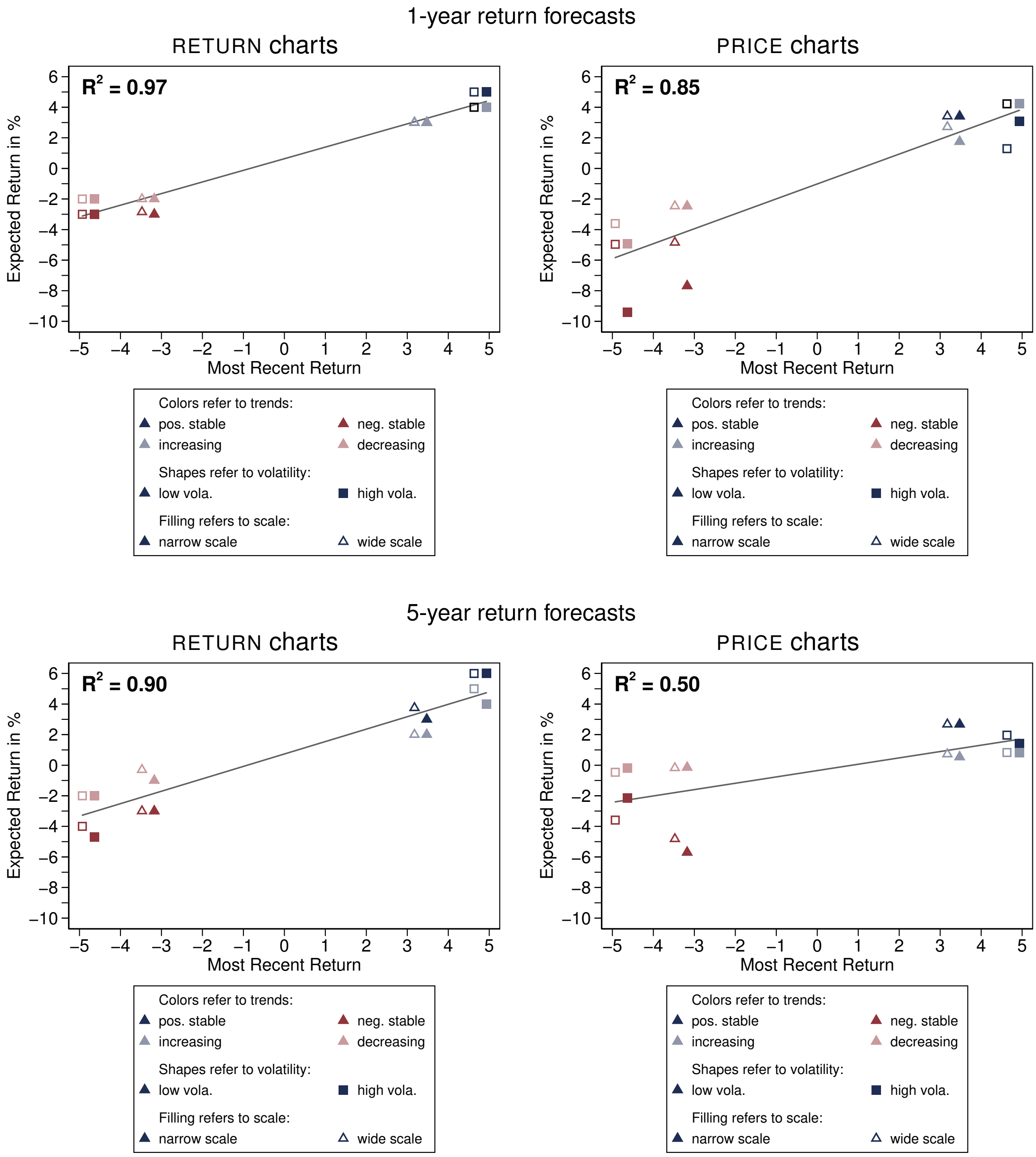

Figure 5: **Median one-year and five-year return forecasts.** This figure shows the median one-year (upper panel) and five-year (lower panel) return forecasts as a function of the most recent return, i.e. the return in Year 10. For better visibility, i.e. to avoid overlapping medians, we add 0.15% to the most recent return of scaling NARROW and deduct 0.15% for those of scaling WIDE on the horizontal axis. Each point represents the median of between 83 and 109 observations.

3.3 Investment preferences in individual assessments

We elicit subjects' propensities to invest (on a scale from 1 to 7) for each of the displayed return and price charts. We find these to be negatively related to perceived riskiness,[13] i.e. assets with a POSITIVE STABLE trend are the ones subjects would most like to invest in, while those with NEGATIVE STABLE trends are least preferred. What we are interested in, however, is, whether there are differences in investment preferences between scales (NARROW vs. WIDE), i.e. whether the differences in risk perception we report in Section 3.1 translate into differences in investment propensities.

Result 3. *Investment propensity is driven by an asset's historical return and volatility as well as by subjective risk perception and expected returns. Varying the scale, however, has almost no influence on investment propensity.*

Support: Fig. 6 summarizes subjects' answers by displaying the differences in average investment propensity (value in NARROW minus value in WIDE) by trend and scale presented as RETURN charts (left) and PRICE charts (right). We only find a significant difference for DECREASING trends in PRICE charts, i.e. there is a higher likelihood to invest when these are displayed with a WIDE scaling, while the other 14 tests do not yield significant differences. Hence, we do not find large, systematic differences between scales regarding their investment propensity.[14]

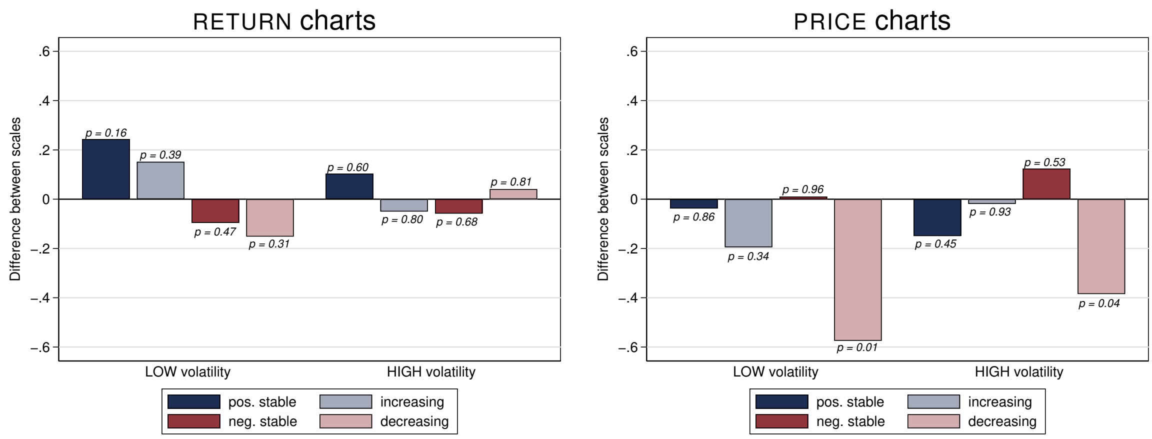

Figure 6: **Differences in average investment propensity (in NARROW minus WIDE) by trend and scale presented as RETURN charts (left) and PRICE charts (right).** This figure depicts differences in average investment propensity (on a scale from 1 = "very unlikely to invest" to 7 = "very likely to invest") for RETURN chart and PRICE chart representations of LOW (left bars in each panel) and HIGH (right bars in each panel) volatility assets. p-values above the bars are from Fisher-Pitman permutation tests on the subject-demeaned data. Each of the sixteen bars summarizes between 179 and 206 observations.

[13] With a Spearman's rank correlation coefficient of –0.68 the relationships are far from perfect, though.

[14] The average investment propensity for each asset and corresponding significance tests for differences between scales are provided in Table A5 in Online Appendix A.

In an attempt to explain investment behavior more generally, Nosić and Weber (2010) and Kaufmann et al. (2013) point out that in a behavioral risk-return framework risk taking – and therefore being willing to invest in risky assets – is driven not just by the historical return and volatility of an asset, but by the investor's risk attitude, her risk perception, and her subjective return expectation regarding the asset: thus, *Risk Taking* = f(*Perceived Return; Risk Attitude; Perceived Risk*) (also see Sarin and Weber, 1993; Jia et al., 1999). We run least squares regressions similar to Nosić and Weber (2010) to examine subjects' investment behavior. Participants' risk attitudes are captured by subject-fixed effects. Detailed results are provided in Table 2.

The first regression shows that higher historical return and lower historical volatility of the asset increase subjects' propensity to invest (measured on a scale from 1 to 7). Model (2) regresses investment propensity on subjective measures of risk and return – that is, subjects' perceptions. The estimates suggest that lower (subjective) perceived risk of an asset (given a specific presentation format) and higher subjective long-term expected returns increase the likelihood of investing. Combining (1) and (2) in Model (3) only marginally increases the model's explanatory power compared to the one with only subjective regressors, confirming the intuition put forward above: investment propensity predominantly relies on people's subjective perceptions.

In Section 3.1, we have demonstrated that the scale in which an asset's performance is presented drives people's perceptions about its risk. Therefore, we estimate an additional model, substituting risk perception by the interaction of the asset and the chart's scale; for the full data set in Model (4) as well as for return charts in Model (4a) and for price charts in Model (4b).[15] While the coefficients remain comparable in magnitude and significance, we can now run Wald tests for differences between scales for each of the eight assets. In the full-data Model (4), we find significantly lower investment propensity for the low-volatility asset with trend DECREASING when presented with a NARROW scale ($p < 0.01$). For five out of seven other assets, people's willingness to invest is also lower with a NARROW scale but the differences are insignificant.

4 Results from Task II: Pairwise comparisons

In Task II subjects are asked to compare two assets displayed on the screen at the same time. We ask for perceived riskiness ("Which of the two assets do you consider to be more risky?"), perceived profitability ("Which of the two assets do you consider to be more profitable?"), and investment propensity ("In which of the two assets would you rather invest?"). In four

[15] For these models perceived risk is substituted by 16 dummy variables for each possible Asset×Scale interaction term (eight assets presented in two different scales, NARROW and WIDE). Corresponding Wald tests are provided in Table A8 in Online Appendix A.

Table 2: **Determinants of investment behavior**. The table presents the estimated coefficients of subject-fixed effects regressions with *investment propensity* (1 = very likely to invest, 7 = not likely to invest) as the dependent variable. The assets' average return and volatility, as well as subjects' perceived risk and return expectations act as independent variables. In Model (4), (4a), and (4b), perceived risk is substituted by 16 dummy variables for each possible Asset×Scale interaction term (eight assets presented in two different scales, NARROW and WIDE).

	Dependent variable: *Investment propensity*					
	Pooled Data				RETURNS	PRICES
	(1)	(2)	(3)	(4)	(4a)	(4b)
Hist. Return	0.438***		0.312***	0.395***	0.368***	0.376***
	(0.019)		(0.017)	(0.027)	(0.036)	(0.040)
Hist. Volatility	−0.156***		−0.026*	−0.159***	−0.235***	−0.106***
	(0.016)		(0.015)	(0.023)	(0.028)	(0.035)
Risk Perception		−0.413***	−0.413***			
		(0.015)	(0.015)			
Subj. Exp. Return (1y)		0.000	0.000	0.001	0.024***	0.001
		(0.001)	(0.001)	(0.001)	(0.009)	(0.001)
Subj. Exp. Return (5y)		0.030***	0.030***	0.036***	0.050***	0.043***
		(0.003)	(0.003)	(0.004)	(0.008)	(0.005)
Constant	0.561	3.934*	2.886	0.321	1.320	−0.627
	(2.560)	(2.238)	(2.240)	(2.518)	(3.164)	(3.739)
Subject-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Asset dummies	Yes	Yes	Yes	–	–	–
Asset×Scale interaction	No	No	No	Yes	Yes	Yes
Observations	3,088	3,080	3,080	3,080	1,544	1,536
Adj. R^2	0.597	0.692	0.692	0.611	0.702	0.556

different conditions the two displayed assets vary by neither scale nor volatility (Condition SAME), only in volatility (VOLATILITY), only in the scale of the vertical axis (SCALE), or by both (BOTH), respectively, but the two assets shown always share the same presentation format and trend. Subjects compare assets eight times with RETURN charts and eight times with PRICE charts. For an investment decision this pairwise comparison could be a more natural setting than Task I, as people often consider more than one investment possibility before deciding to invest in one particular financial instrument.

Fig. 7 summarizes the results of Task II for RETURN charts (left panels) and PRICE charts (right panels). Each panel shows from left to right the four distinct trends (POSITIVE STABLE, INCREASING, NEGATIVE STABLE, and DECREASING) and from top to bottom we show data for the four conditions SAME, SCALE, VOLATILITY, and BOTH. The first bar of each group of bars within a panel always corresponds to perceived riskiness ('risk'), the second bar to perceived profitability ('profit.'), and the third bar to investment propensity ('inv.'). Each bar shows the percentage of decisions in which subjects perceive the assets as the same (light grey) or differently (black in the top row of panels; dark and light red in the second row for NARROW vs. WIDE scaling, and dark vs. light blue in the bottom two rows for the high- vs. low-volatility assets).

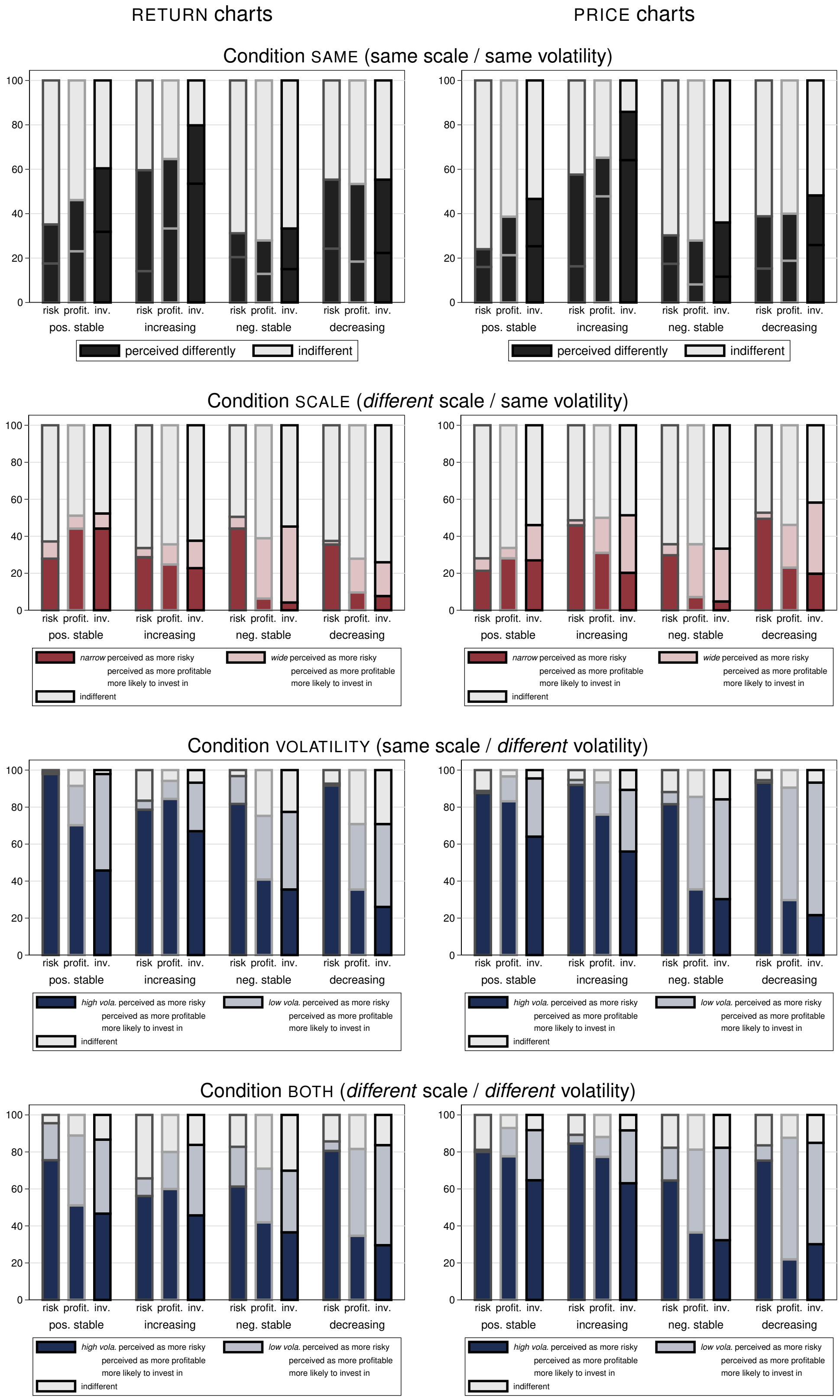

Figure 7: **Perceived riskiness, perceived profitability, and investment propensity in Task II.** This set of panels shows the percentage of decisions in which subjects perceive the riskiness (first bar in each set labelled 'risk') and the profitability (second bar; 'profit.') the same or differently, and in which subjects are more likely to invest in (third bar; 'inv.'), between different scalings and volatilities. The left panels show data for RETURN charts, while the right panels show the respective data for PRICE charts. From top to bottom we show the four different conditions, where the condition name corresponds to the variable in which the two assets of a pair *differ*: SAME (same scale / same volatility), SCALE (different scale / same volatility), VOLATILITY (same scale / different volatility), and BOTH (different scale / different volatility). In each panel data for the four distinct price trends are shown separately. Each of the eight panels summarizes between 314 and 386 observations for each variable.

4.1 Risk perception in pairwise comparisons

Result 4. *Different scaling can distort people's perception about an asset's risk in pairwise comparisons of assets with the same volatility, as assets shown on a* NARROW *scale are perceived as more risky.*

Support: From the first row of panels in Fig. 7 we see that even when the two displayed assets have the same volatility and there is no difference in the scale, a surprisingly high share of between 23% and 60% of subjects perceive risk differently between the two assets. The share who perceives the risk differently is significantly higher for INCREASING and DECREASING trends than for POSITIVE STABLE and NEGATIVE STABLE trends ($p < 0.01$).[16] This holds for RETURN as well as PRICE charts. Probably subjects saw the similarity between the two charts/price paths shown, but thought there must be some difference to find and hence considered the two paths as differently risky. Data supporting this line of argumentation is the time subjects needed until they reached a decision: even though the scales are all equal (and thus easiest to compare) in Condition SAME, subjects took significantly longer here (25.02 sec. on average) than in any other condition (between 21.74 and 23.35 sec.)

The second line of panels of Fig. 7 shows the distribution of choices for Condition SCALE, i.e. when the asset volatility is the same but the scale is different (NARROW on one side, WIDE on the other). One could argue that this is the 'trap' case where two identical assets are shown with different scaling to see whether subjects can be misled and thus perceive the asset shown with the narrower scale as more risky. This is indeed the case in 22% to 42% of all cases, while the asset with the wider scale is perceived as riskier in only 2% to 9% of all cases (in the remaining cases subjects are indifferent – most likely correctly seeing that only the scaling is different between both sides). Results are similar for RETURN and PRICE charts.

Result 5. *Different scaling can distort people's perception about an asset's risk in pairwise comparisons of assets with different volatilities. Depicted with the same scale, subjects regard the more volatile asset as riskier; with differing scales, a considerable fraction erroneously perceives the less volatile asset as riskier or views them as equally risky.*

Support: In the third row of panels of Fig. 7 we present the case where the volatility of the two assets varies by a factor of six while the scale is the same (Condition VOLATILITY). Here, both assets are displayed on exactly the same axes and it should thus be comparatively easy to identify the more volatile asset and, if volatility is perceived as 'risk', also to identify this asset as the riskier one. We find that with a POSITIVE STABLE trend in RETURN charts, almost 100% of subjects do exactly that. We also report very high shares of 90% and above identifying the more volatile asset as the riskier one for DECREASING trends, both in RETURN and PRICE charts, and for POSITIVE STABLE and INCREASING trends in PRICE charts. For INCREASING

[16] For comparisons between different choices and conditions in the pairwise comparisons of Task II we conduct two-sided Wald tests on the respective proportions and report the corresponding p-values in parentheses.

trends in RETURN charts we find 18% of subjects to be indifferent between the two assets – most likely as both trends start negative and then mostly increase, which is perceived as equally good, irrespective of volatility. An interesting case are the NEGATIVE STABLE trends, especially for RETURN (but also, to a lesser degree, for PRICE) charts: here around 20% do not see the less volatile asset as the less risky one. For RETURN charts, 16% even consider the more volatile asset as less risky. We conjecture that this is the case as all returns in the low-volatility case are clearly negative – hence, an investor always loses with this asset. With high volatility, the dispersion of returns is much wider and thus the chance of earning a positive return is also higher. The corresponding asset is therefore perceived as less risky in about every sixth decision.

Finally, the fourth row of panels in Fig. 7 displays the choices in Condition BOTH, where the shown assets differ in their volatility and additionally in their scaling. Note that in this condition, one side depicts a low-volatility asset and the other a high-volatility asset, hence different scales lead to both assets being displayed either on a WIDE or on a NARROW scale with the respective bars having comparable magnitudes. Comparing the results in this condition to the ones in the third row of panels we see that the choices are now more dispersed. While the high-volatility asset is still perceived as the more risky one in the majority of cases (between 56% and 83% of cases), 'indifferent' (up to 35%) and a preference for the low-volatility asset (up to 21% of cases) are chosen markedly more frequently than when the scaling is the same. In this cognitively demanding condition, results between and within RETURN and PRICE charts vary more than in other conditions. In particular, with trend POSITIVE STABLE almost 20% see the low-volatility asset as the riskier one in return charts but choose 'indifferent' in price charts. We conjecture that in both cases subjects are misled by the different scalings. For trend INCREASING, however, in a remarkably high share of 33% of decisions subjects are indifferent between the high- and low-volatility assets with RETURN charts (for PRICE charts the respective number is only 11%). We argue that for these subjects the main decision criterion is the clear upward trend in returns, while the details of the vertical axis scale and the exact values plays a smaller role. Assets with trend NEGATIVE STABLE (third group of bars) are again a different story: a significantly higher share of subjects picks the less volatile asset as the riskier one than in any other trend bar POSITIVE STABLE with RETURN charts (all other $p < 0.05$). This hints at losses being a driving force behind risk perception as all returns are negative in the low-volatility asset but not in the high-volatility one. Hence, in the NEGATIVE STABLE trend, having more volatile returns increases the chance that an investor could end up with a positive return. The substantial share of 'indifferent' answers (around 20%) might result from the fact that the wider scaling counterbalances the higher volatility and probably leads some subjects to judge the two assets as fairly similar in risk when their volatility actually varies by a factor of six.

4.2 Perceived profitability in pairwise comparisons

We now turn to the analysis of perceived profitability in pairwise comparisons. The respective proportions for this variable are depicted in the second bar of each group of bars in Fig. 7 ('profit.').

Result 6. *Scaling can distort people's perception about an asset's profitability in pairwise comparisons of assets with the same volatility. For trends* POSITIVE STABLE *and* INCREASING, *assets shown on a* NARROW *scale are regarded as more profitable; for trends* NEGATIVE STABLE *and* DECREASING*, the opposite holds.*

Support: In the top row of panels in Fig. 7, showing results for Condition SAME, between 35% and 72% of subjects state a difference in perceived profitability, even though the two assets are essentially identical. The patterns observed, both with RETURN charts (left panels) and PRICE charts (right panels) are almost identical to the ones from perceived riskiness (first bar, 'risk', in each group of bars).

When volatility and expected returns are the same but the charts are shown with different scaling (Condition SCALE; see second row of panels in Fig. 7), we find that between 49% and 73% of subjects (correctly) see no difference in profitability. However, up to 51% do see a difference. For the POSITIVE STABLE and INCREASING trends (first two groups of bars of each panel) the results are again similar as for perceived riskiness – the asset shown with narrow scaling is perceived as the more profitable one as the (mostly positive) bars are displayed larger here. For the NEGATIVE STABLE and DECREASING trends, however, we find a marked difference: those 35% to 50% of subjects who do perceive a difference in profitability identify the asset displayed with wide scaling as more profitable – the mostly negative returns are shown with smaller bars and subjects are misled to think these are thus more profitable. Subjects fall into this 'trap' in between 20% and 38% of all decisions.

Result 7. *Scaling can distort people's perception about an asset's profitability in pairwise comparisons of assets with different volatilities. With the same scale, more volatile assets of trends* POS. STABLE *and* INCREASING *are regarded as more profitable; for trends* NEG. STABLE *and* DECREASING*, the opposite holds. With different scales, a large share perceives the low-volatility asset as more profitable.*

Support: The third row of panels in Fig. 7 presents Condition VOLATILITY, in which for PRICE charts there are more extreme prices for HIGH-volatility assets – that is, the HIGH-volatility asset yields higher prices with trends POSITIVE STABLE and INCREASING and lower prices with trends NEGATIVE STABLE and DECREASING (compared to the respective LOW-volatility assets). While almost 100% of subjects perceive the asset with the higher volatility as the riskier one with a POSITIVE STABLE trend, we find 21% of subjects to assess the less risky asset as the more profitable one in this trend with RETURN charts. It seems that for profitability assessments subjects also take a lower volatility into account. Most notably, however, we observe the same pattern as above: for POSITIVE STABLE and INCREASING trends (first two groups of bars of each panel), most subjects perceive the more volatile asset as the more profitable one (as the returns bars are mostly positive, respectively the price mostly goes up), while for the NEGATIVE STABLE and DECREASING trends (last two groups of bars of each panel), the opposite holds and the mostly negative returns/falling prices lead subjects to select the low-volatility asset as the more profitable one. For the latter two price trends the 'indifferent' choices are also markedly higher than for the positive price trends ($p < 0.01$ for RETURN charts, $p < 0.05$ for PRICE charts) – probably because subjects see both assets markedly going down and consider this a decision 'between a rock and a hard place', i.e. a choice between two equally bad alternatives.

Finally, for Condition BOTH we find shares of 28% to 66% of decisions in which subjects consider the asset with the lower volatility to be the more profitable one in the NEGATIVE STABLE and DECREASING trends. In addition, also in the two other trends (first two groups of bars) the share of subjects considering the low-volatility asset as the more profitable one is substantial, especially when RETURN charts are displayed. These shares of up to 40% are markedly higher than the respective shares for perceived riskiness ($p < 0.01$).

4.3 Investment preferences in pairwise comparisons

In Task II we ask as third question which of the two displayed assets subjects would rather invest in and incentivize this question by paying subjects one randomly determined return from the chosen asset. This allows us to examine the behavioral consequences of the scale effects reported above. The third bar in each set of bars in Fig. 7 ('inv.') shows the respective shares of investments in one of the two displayed assets.

In Condition SAME a considerable fraction of subjects sees the two displayed assets as bearing different risks and profitabilities. A still higher share of between 33% and 86% of subjects decide to invest in either of the two. The share of 'indifferent' choices is smaller for investment preferences than for riskiness and profitability. This holds for each trend and both presentation formats. This increase in decisiveness can probably be attributed to the monetary incentives associated with this particular question.

For Condition SCALE, we find a considerable effect of the axis scale regarding investment decisions. Here, choices tend to be very similar to profitability assessments: as two identical assets are compared, assets presented on a NARROW scale are more frequently preferred for trends POSITIVE STABLE and INCREASING, whereas for trends NEGATIVE STABLE and DECREASING, the opposite holds (with the exception of INCREASING in PRICE charts). Comparing conditions SAME and SCALE vertically reveals how the scale affects subjects' behavior: we find on average a much more pronounced preference for either of the two displayed assets.

In the pairwise comparisons of conditions VOLATILITY and BOTH, we observe a number of diverging preferences – i.e., especially when return charts are displayed a large fraction chooses to invest in the HIGH-volatility asset and a similarly large fraction chooses to invest in the LOW-volatility asset. The variation of the scale between these two conditions (same or different) does not result in a systematic difference in investment preferences.

Naturally, we are interested in how perceived riskiness and perceived profitability relate to people's investment decisions. Comparing the shares of answers regarding investment preferences with the corresponding values concerning perceived riskiness (first bar, 'risk') and profitability (second bar, 'profit.') already hints at a meaningful relationship between these variables with the tendency of more investments in assets which are perceived as less risky and more profitable.

For a more thorough analysis we estimate the probability with which a subject invests in either the NARROW-scaled or in the HIGH-volatility asset, respectively, depending on which asset she perceives as more risky and as more profitable, by running probit regressions.[17] The resulting probabilities are plotted in Fig. 8.

Result 8. *Regarding one of two assets as more profitable and less risky leads to a higher probability of investing in this asset. Of the two factors profitability tends to be more important.*

Support: Regarding the effect of perceived riskiness (top panel), we observe that in cases when a subject perceives the LOW volatility or the WIDE-scaled asset as more risky, the probability that she invests in the HIGH volatility or NARROW-scaled asset is between 68% and almost 100% across trends with return charts and even higher for price charts. Conversely, only around 25% tend to invest in this asset if it is perceived as riskier in trends NEGATIVE STABLE and DECREASING, with a significantly higher number for trends POSITIVE STABLE and INCREASING.

For price charts the probability of investing in the asset with higher perceived risk is 61 and 54%, respectively, with these trends – indicating that perceived risk is not necessarily the main determinant of investment behavior in the domain with mostly positive returns. Investing in the higher-volatility asset need not be a 'wrong' choice – especially in the case of a NEGATIVE STABLE trend having more volatile returns increases the chance that an investor

[17] In particular, we run probit regressions with the pooled decisions across all conditions in which the two assets differ by at least one variable of interest – i.e., conditions SCALE, VOLATILITY, and BOTH – and consider only those decisions in which a subject chose either one of the assets to invest in. Hence, the dependent variable is a binary variable taking the value 1 when a subject chose to invest in the high-volatility or narrow-scaled asset, respectively, and 0 otherwise. Her choices regarding riskiness and profitability act as explanatory variables. For estimating probabilities depending on the choice in riskiness, choices in profitability were assumed to be at their means and vice versa. We are aware that c.p. a higher volatility might have different effects than displaying the asset on a narrow scale; however, as we find very similar patterns within each condition, we present results from the pooled data only. The estimates for each individual condition are available upon request.

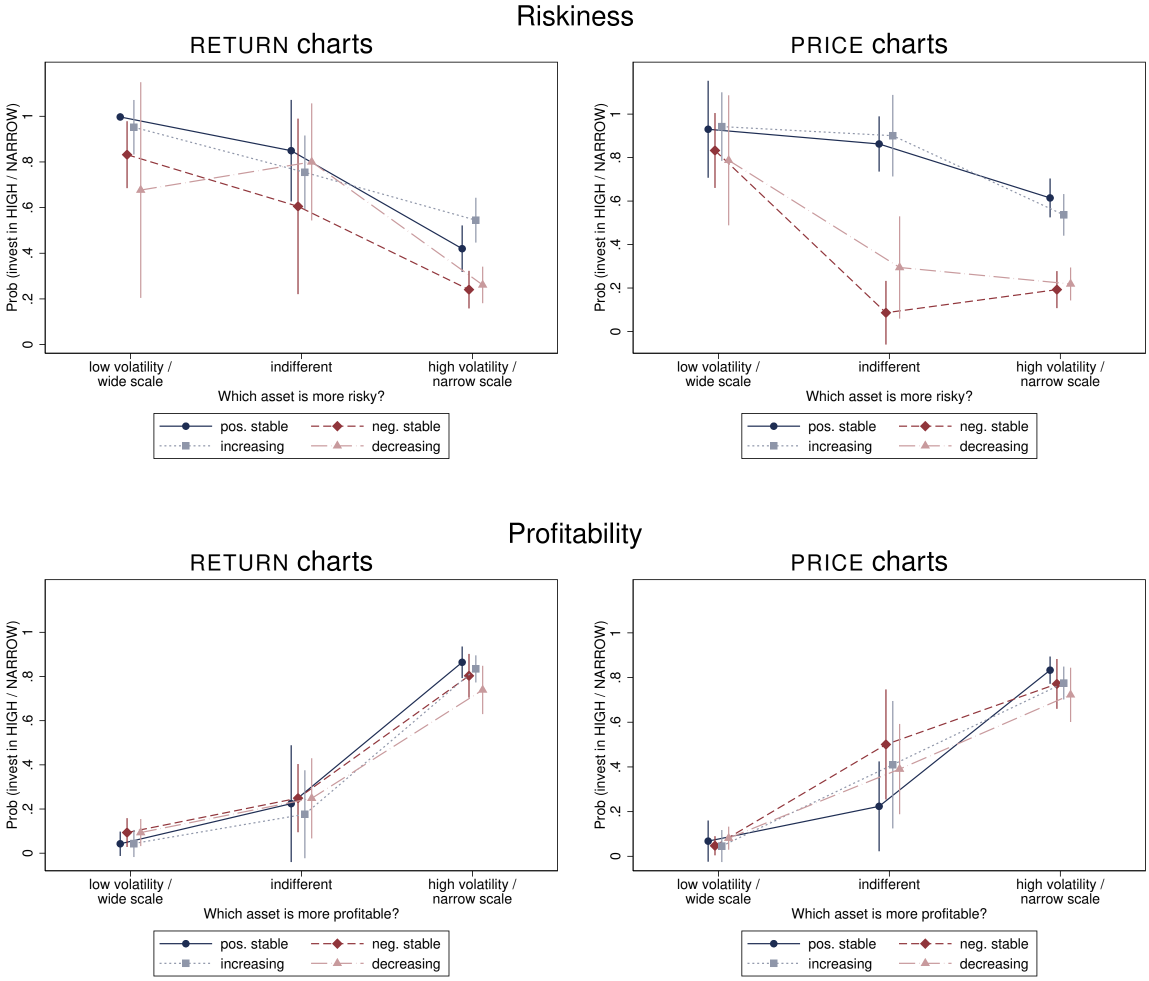

Figure 8: **Predicted probabilities of investing in the high-volatility or narrow-scaled asset.** This figure shows the predicted probabilities and 95%-confidence intervals of investing in the HIGH-volatility (NARROW-scaled) asset depending on which asset is perceived as *more risky* (top) or as *more profitable* (bottom). Probabilities are estimated from a probit model with a dummy variable indicating whether the subject would invest in the HIGH-volatility (NARROW-scaled) asset as the dependent variable and her choice regarding riskiness and profitability as independent variables. The numbers of observations for each estimation lie between 171 and 222 for different presentation formats and trends.

could end up with a positive return. Such choices are thus in line with Prospect Theory (Kahneman and Tversky, 1979) which postulates risk-seeking behavior in the loss domain, where returns of assets with a NEGATIVE STABLE trend mostly are (assuming zero return as subjects' reference point). Hence, subjects who prefer the high-volatility asset over the low-volatility one in NEGATIVE STABLE (and with lower shares also in INCREASING and DECREASING trends) should not be judged 'irrational' or 'incorrect', but can merely be risk-seeking in the loss domain.

Analyzing probabilities to invest depending on perceived profitability (bottom panel) we observe very similar estimates across all trends. If the HIGH volatility (NARROW-scaled) asset is perceived as more profitable, the probability of a subject investing in this asset is also very high (between 74 and 86%), and vice versa. As we observe comparable dynamics for all trends – also for POSITIVE STABLE and INCREASING, in particular – we conclude that perceived profitability is more important than perceived riskiness in these decisions. Subjects tend to invest in the asset which they regard as more profitable, even if they think it bears higher risk.

5 Discussion and Conclusion

In a novel experimental design we examined the impact of different vertical axis scales and presentation formats on risk perception, short- and long-term return expectations, and investment propensity. We explored return bar charts and price line charts for eight distinct assets, distinguished by either a low or a high volatility and one of four distinct return trends.

We found that varying the scale strongly affected people's risk perception. Namely, a narrower scale of the vertical axis – that is, letting return bars and the line depicting the price, respectively, fill most of the available, vertical space in a chart – lead to significantly higher perceived riskiness of an asset. This result is robust to varying the chart's presentation format (prices vs. returns) and the asset's volatility and trend. Only when returns were consistently positive, risk perception was the same across different scalings.

Assets were usually perceived as riskier when returns were shown than when prices were shown. Regulations like the European standard for investor documents (Commission Regulation (EU) No 583/2010, 2010, p. 15) demand return bar charts and a vertical axis that *shall not compress the bars so as to make fluctuations in returns harder to distinguish*. We demonstrate that adapting the scale accordingly is reasonable with regard to recognizing yearly return variations *within* a single security, but at the same time makes it harder to identify differences *between* dissimilar securities.

We further reported that past returns predicted future return expectations almost perfectly, irrespective of the presentation format. Most subjects in our setting thus act as short-term trend-followers when predicting future prices and returns.

Risk perception is highly correlated with losses, which in turn drive investment behavior. This connects nicely to recent literature which also finds that risk perception is most strongly driven by 'probability of loss', and that this drives investment intentions (Anzoni and Zeisberger, 2016) and prices (Huber et al., 2018). Concerning investment choices, subjects tend to invest in the asset which they regard as more profitable even if they assess it to be riskier. Hence, in our setting perceived profitability was considered more important than perceived riskiness when making investment choices.

With regard to policy, our results have important implications: to our knowledge, financial market regulators in the US require consumer information documents to contain return bar charts representing past performance, but do not require a standardized appearance.[18] EU regulations also demand the presentation of return bar charts and, in addition, specific criteria regarding the presentation format. Yet, neither acknowledges the potentially distorting effects of the axis scale. In particular, the EU suggests *adapting the scale to the span of the bars* (Commission Regulation (EU) No 583/2010, 2010, p. 15). As we have shown, this makes it harder to distinguish assets with different levels of volatility. hence, a well-meant regulatory rule might even have unintended negative consequences on investors' decisions. An example is the case of two passive funds with the same tracking error but with different fee structures. As funds are required to report returns net of fees, this case essentially corresponds to a change of the trend of the data generating process. Many investors will not be able to detect the fund with the better fee structure if fund companies follow the current regulatory rules and adjust their scaling to the data.

To summarize, regulators should be aware of – and attentive to – the potentially distorting effects of different axis scales in performance charts. While return bar charts are appropriate, allowing issuers to adapt the axis scale arbitrarily leaves room for deliberate action aimed at distorting investors' perceptions about risk. Keeping the presentation scale constant across different securities enables better identification of risk and therefore better comparisons and decisions.

[18] See Zimmer (2009) for US regulations regarding past performance information in prospectuses and Mercer et al. (2010) for mutual fund advertisements.

References

Anzoni, Laura, Stefan Zeisberger. 2016. What is risk? How investors perceive risk in return distributions. Working Paper.

Beattie, Vivien, M Jones. 2008. Corporate reporting using graphs: A review and synthesis. *Journal of Accounting Literature* **27** 71–110.

Beattie, Vivien, Michael John Jones. 1992. The Use and Abuse of Graphs in Annual Reports: Theoretical Framework and Empirical Study. *Accounting and Business Research* **22**(88) 291–303.

Bock, Olaf, Ingmar Baetge, Andreas Nicklisch. 2014. hroot: Hamburg registration and organization online tool. *European Economic Review* **71** 117–120.

Chen, Daniel L., Martin Schonger, Chris Wickens. 2016. oTree — An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* **9** 88–97.

Cleveland, William S., Marylyn E. McGill, Robert McGill. 1988. The shape parameter of a two-variable graph. *Journal of the American Statistical Association* **83**(402) 289–300.

Commission Regulation (EU) No 583/2010. 2010. Commission Regulation (EU) No 583/2010 of 1 July 2010 implementing Directive 2009/65/EC of the European Parliament and of the Council as regards key investor information and conditions to be met when providing key investor information or the prospectus in a durable medium other than paper or by means of a website. *Official Journal* **L 181**(10.7.2010) 1–15. http://data.europa.eu/eli/reg/2010/583/oj.

Diacon, Stephen, John Hasseldine. 2007. Framing effects and risk perception: The effect of prior performance presentation format on investment fund choice. *Journal of Economic Psychology* **28**(1) 31–52.

Ehm, Christian, Christine Kaufmann, Martin Weber. 2014. Volatility inadaptability: Investors care about risk, but can't cope with volatility. *Review of Finance* **18** 1387–1423. doi:10.1093/rof/rft032.

Gächter, Simon, E. J. Johnson, A. Hermann. 2007. Individual-level loss aversion in riskless and risky choices. CeDEx Discussion Paper No. 2007-02.

Glaser, Markus, Zwetelina Iliewa, Martin Weber. 2018. Thinking about prices versus thinking about returns in financial markets. *The Journal of Finance* **forthcoming**.

Glaser, Markus, Thomas Langer, Jens Reynders, Martin Weber. 2007. Framing effects in stock market forecasts: The difference between asking for prices and asking for returns. *Review of Finance* **11**(2) 325–357.

Grosshans, Daniel, Stefan Zeisberger. 2018. All's well that ends well? On the importance of how returns are achieved. *Journal of Banking & Finance* **87** 397–410.

Holt, Charles A., Susan K. Laury. 2002. Risk aversion and incentive effects. *American Economic Review* **92**(5) 1644–1655.

Holzmeister, Felix. 2017. oTree: Ready-made apps for risk preference elicitation methods. *Journal of Behavioral and Experimental Finance* **16** 33–38.

Huber, Jürgen, Stefan Palan, Stefan Zeisberger. 2018. Does investor risk perception drive asset prices in markets? Experimental evidence. Working Paper.

IFF Research, YouGov. 2009. UCITS Disclosure Testing Research Report. Research report, European Commission. URL `http://ec.europa.eu/internal_market/investment/docs/other_docs/research_report_en.pdf`.

Jarvenpaa, Sirkka L. 1989. The effect of task demands and graphical format on information processing strategies. *Management Science* **35**(3) 285–303.

Jia, Jianmin, James S Dyer, John C Butler. 1999. Measures of perceived risk. *Management Science* **45**(4) 519–532.

Kahneman, Daniel, Amos Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47** 263–291.

Kaiser, Johannes. 2007. An exact and a Monte Carlo proposal to the Fisher-Pitman permutation tests for paired replicates and for independent samples. *Stata Journal* **7**(3) 402–412.

Kaufmann, Christine, Martin Weber, Emily Celia Haisley. 2013. The role of experience sampling and graphical displays on one's investment risk appetite. *Management Science* **59** 323–340.

Lawrence, Michael, Marcus O'Connor. 1992. Exploring judgemental forecasting. *International Journal of Forecasting* **8**(1) 15–26.

Lawrence, Michael, Marcus O'Connor. 1993. Scale, variability, and the calibration of judgmental prediction intervals. *Organizational Behavior and Human Decision Processes* **56**(3) 441–458.

Mercer, Molly, Alan R Palmiter, Ahmed E Taha. 2010. Worthless warnings? Testing the effectiveness of disclaimers in mutual fund advertisements. *Journal of Empirical Legal Studies* **7**(3) 429–459.

Nosić, Alen, Martin Weber. 2010. How riskily do I invest? The role of risk attitudes, risk perceptions, and overconfidence. *Decision Analysis* **7**(3) 282–301.

Sarin, Rakesh K, Martin Weber. 1993. Risk-value models. *European Journal of Operational Research* **70**(2) 135–149.

Stössel, Remo, Anna Meier. 2015. Framing effects and risk perception: Testing graphical representations of risk for the KIID. Working Paper.

Tufte, Edward R. 1983. *The visual display of quantitative information*. Graphics Press.

Weber, Elke U, Niklas Siebenmorgen, Martin Weber. 2005. Communicating asset risk: how name recognition and the format of historic volatility information affect risk perception and investment decisions. *Risk Analysis* **25**(3) 597–609.

Zimmer, Sarah B. 2009. Securities and exchange commission's enhanced disclosure and new prospectus delivery option for registered mutual funds. *St. John's Law Review* **83** 1431–1468.